STUDIA POLITOLOGICZNE

Wydział Nauk Politycznych
i Studiów Międzynarodowych
Uniwersytet Warszawski

UNIWERSYTET
WARSZAWSKI

VOL. 54

POLITICAL SCIENCE STUDIES

INFOBROKERING – THE ART OF ACQUIRING, ANALYZING AND EVALUATING INFORMATION

edited by Daniel Mider

WARSAW 2019

VOL. 54

STUDIA POLITOLOGICZNE

INFOBROKERING – SZTUKA POZYSKIWANIA, ANALIZY I EWALUACJI INFORMACJI

pod redakcją Daniela Midera

WARSZAWA 2019

VOL. 54

„Studia Politologiczne" znajdują się w wykazie czasopism naukowych prowadzonym przez Ministra Nauki i Szkolnictwa Wyższego na potrzeby oceny jednostek naukowych z przyznaną liczbą 13 punktów.

„Studia Politologiczne" są dostępne w bazach danych: CEJSH, Index Copernicus, Erih Plus.

Czasopismo recenzowane przez recenzentów zewnętrznych.

Wersja pierwotna czasopisma: papierowa.
www.studiapolitologiczne.pl

ISSN 1640-8888
Nakład 600 egz.

DOM WYDAWNICZY ELIPSA

Opracowanie komputerowe, druk i oprawa:
Dom Wydawniczy ELIPSA
ul. Inflancka 15/198, 00-189 Warszawa
tel. 22 635 03 01
e-mail: elipsa@elipsa.pl, www.elipsa.pl

Spis treści

Wprowadzenie 7

STUDIA I ANALIZY

Przemysław Potocki
Informacyjne determinanty rozwoju społeczno-ekonomicznego w XXI wieku: perspektywa infobrokeringu 14

Piotr Sosnowski
Systematyzacja pojęć związanych z metodami i źródłami pozyskiwania informacji w kontekście infobrokeringu 45

Piotr Dela
Elementy propagandy w życiu publicznym 68

Urszula Kurcewicz
Znaczenie tradycyjnych źródeł informacji w działalności infobrokerskiej 96

Magdalena Tomaszewska-Michalak
Prawne aspekty pozyskiwania informacji w Internecie 116

Konrad Gałuszko, Joanna Lewczuk, Konrad Krystian Kuźma
Walidować? Weryfikować? Nie ruszać? O niestatystycznych, statystycznych i stochastycznych metodach oceny jakości danych ilościowych opowieść 135

Patrycja Hrabiec-Hojda, Justyna Trzeciakowska
Techniki wyszukiwania informacji w mediach społecznościowych dla celów białego wywiadu 175

Daniel Mider
Sztuka wyszukiwania w Internecie – autorski przegląd wybranych technik i narzędzi 191

Wojciech Mincewicz
Metadane – cichy zabójca prywatności 230

Paweł Tomczyk, Daniel Mider, Józef Grzegorczyk
Inwigilacja elektroniczna jako metoda pozyskiwania informacji – ewaluacja i prognozy 258

RECENZJE

Michael Bazzell, *Open Source Intelligence Techniques. Resources for Searching and Analyzing Online Information* (**Bartosz Biderman**) 295
Christopher Hadnagy, *Socjotechnika. Sztuka zdobywania władzy nad umysłami* (**Anna Maria Złocka**) 300
Питер Фрейз, *Четыре сценария будущего: Жизнь после капитализма* (Иванна Килюши) 305
Andrzej Wierzbicki, *Polish-Belarusian Relations, Between a Common Past and the Future* (**Яна Волчецкая**) 310

Autorzy 314

Wprowadzenie

Wytwarzanie i proliferacja informacji w Internecie zyskała kluczowe znaczenie we współczesnej polityce lokalnej, narodowej, regionalnej i globalnej. Jednocześnie obserwuje się intensyfikację negatywnych zjawisk w tym obszarze – pojęcia takie jak *fake news*, *astroturfing*, *sockpuppet*, *troll* stały się trwale obecne we współczesnym globalnym krajobrazie politycznym, dyskurs zaś w tej sferze coraz częściej można interpretować w kategoriach walk, a nawet wojen informacyjnych. Koszty fabrykowania fałszywych przekazów są znikome, a rozpowszechnianie informacji łatwe i tanie, co generuje pokusę dla rozmaitych podmiotów pragnących oddziaływać na procesy polityczne. Szczególne znaczenie ma to w reżimach, gdzie opinia publiczna stanowi podmiot gry politycznej władny oddziaływać na decyzje polityczne. Ziszcza się Tofflerowska wizja zmiany władzy – do klasycznej diady czynników konstytuujących władzę polityczną, to jest siły militarnej i ekonomicznej, dołącza trzeci – informacja[1].

Drugim z kluczowych zjawisk jest fenomen nadprodukcji informacji. W kategoriach praktycznych, operacyjnych, ale także akademickich opatruje się je pojęciem *big data*[2]. Jest to tyleż termin odnoszący się do ekstremalnie wielkich zbiorów danych, co i wyraźne oznaczenie zjawiska cywilizacyjnego wytwarzającego wiązkę efektów – zarówno pozytywnych, jak i negatywnych. Od zarania dziejów do 2003 roku ludzkość wytworzyła pięć eksabajtów danych, a obecnie – według ekspertów z IBM – dwa i pół eksabajta wytwarzamy globalnie każdego dnia[3]. Istotą *big data* jest objętość, zmienność, różnorodność, co sprawia, że pozyskiwanie z tych obszernych zbiorów danych jest trudne, wymaga kompetencji informacyjnych i narzędzi, jednak zawierają one informacje nad wyraz wartościowe. *Big data* stanowi jednocześnie źródło negatywnych zjawisk – w literatu-

---

1 A. Toffler, *Zmiana władzy. Wiedza, bogactwo i przemoc u progu XXI wieku*, przekł. P. Kwiatkowski, Poznań 2003.

2 J.S. Ward, A. Barker, *Undefined by Data: A Survey of Big Data Definitions*, http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.705.9909&rep=rep1&type=pdf (dostęp: 13.03.2019).

3 IBM Marketing Cloud, *10 Key Marketing Trends for 2017 and Ideas for Exceeding Customer Expectations*, https://www.ibm.com/downloads/cas/XKBEABLN (dostęp: 13.03.2019).

rze przedmiotu obecne są liczne sugestywne, ekspresywne i pobudzające wyobraźnię pojęcia na oznaczenie informacyjnych patologii, jak: smog informacyjny (*data smog*)[4], bomba megabitowa[5] (obecnie raczej należałoby ją nazwać bombą eksabajtową), przeładowanie informacyjne (*infobesity*, *information overload*)[6] czy zatrucie informacyjne (*infoxication*)[7], zanieczyszczenie informacyjne (*information pollution*)[8], paraliż analityczny (*analysis paralysis* bądź *overchoice* nawiązujący do paradoksu Buridana)[9]. Według Nicolasa Carra zjawisko informacyjnego przeładowania może stać się nawet czynnikiem degeneracji i zapaści cywilizacyjnej. Nadmiar informacji, przy jednoczesnym deficycie dyscypliny ich preselekcji i umiejętności obróbki, doprowadza do rozkojarzenia, a to z kolei do powierzchowności sądów i wniosków. Cywilizacyjny fundament stanowiły dotychczas zaangażowanie w informacje i narracje, sztuka pogłębionej refleksji, koncentracji i zdyscyplinowanego myślenia[10].

W takim kontekście krytyczna staje się umiejętność pozyskiwania, selekcji i ewaluacji informacji. Kompetencje te lokować należy pośród cywilizacyjnych rudymentów, obok umiejętności czytania i pisania. Jednakże sztuka i nauka wyszukiwania, ewaluacji, analizy i przetwarzania informacji pozyskiwanej z Internetu stanowi domenę względnie hermetycznych, profesjonalnych grup zawodowych. Nie zauważa się systemowych, systematycznych i skoordynowanych wysiłków na rzecz opracowania paradygmatów poruszania się w sieci, wyszukiwania, selekcji, ewaluacji i analizy informacji przeznaczonych dla maksymalnie szerokiego spektrum użytkowników. Brakuje skodyfikowanych reguł i zasad oraz programów nauczania i szkoleń w zakresie szeroko ujmowanego pozyskiwania i analizy informacji. Odgórne inicjatywy polegające na insty-

---

[4] D. Shenk, *Data Smog. Surviving the Information Glut*, Nowy Jork 1997.

[5] S. Lem, *Bomba megabitowa*, Kraków 1999.

[6] Oba pojęcia pojawiają się po raz pierwszy w nauce o zarządzaniu: P. Rogers, R. Puryear, J. Root, *Infobesity: The enemy of good decisions. How to attack information overload*, https://www.bain.com/insights/infobesity-the-enemy-of-good-decisions (dostęp: 13.02.2019).

[7] P. Dias, *From 'infoxication' to 'infosaturation': a theoretical overview of the cognitive and social effects of digital immersion*, http://ambitoscomunicacion.com/2014/from-infoxication-to-infosaturation-a-theoretical-overview-of-the-cognitive-and-social-inmersion/ (dostęp: 13.03.2019).

[8] R. Pandita, *Information Pollution, a Mounting Threat: Internet a Major Causality*, https://www.researchgate.net/publication/270564861_Information_Pollution_a_Mounting_Threat_Internet_a_Major_Causality (dostęp: 13.03.2019).

[9] B. Schwarz, *The Paradox of Choice: Why More Is Less*, Nowy Jork 2004.

[10] N.G. Carr, *The Shallows. What the Internet is Doing to Our Brains*, Nowy Jork, Londyn 2010.

tucjonalnej preselekcji, następczej eliminacji lub ograniczaniu proliferacji określonych typów informacji (między innymi w serwisach społecznościowych Facebook i Twitter) są nieefektywne (czasami, jak w przypadku Facebooka, przynoszą skutek odwrotny do zamierzonego, a wdrażające je podmioty rychło się z nich wycofują). Są one również szkodliwe (demotywują użytkowników Internetu w zakresie samodzielnych rozwiązań i poszukiwań informacji, jej analizy, samodzielnego kształtowania opinii, indywidualnego intelektualnego uodparniania się na treści fałszywe), a nawet rodzące niebezpieczeństwa (władza nad możliwością selekcji informacji docierającej do końcowych użytkowników stanowi pokusę, której wiele podmiotów może się nie oprzeć).

Przeciętny użytkownik Internetu porusza się w nim intuicyjnie, przygodnie, często nie posiada elementarnej wiedzy o jego możliwościach i zasobach. Rzadko bywa, że ma umiejętność wprawnego wyszukiwania, a następnie oceny ujawnionych zasobów. Badania internautów ujawniają niskie kompetencje użytkowników, a nawet brak elementarnej świadomości w zakresie spektrum możliwości wyszukiwania w Internecie[11]. Nieustannie doskonalone algorytmy i interfejsy człowiek – maszyna zaledwie częściowo redukują bezradność poznawczą użytkowników (choć niekiedy paradoksalnie pogłębiają ją), a przyrastająca w postępie geometrycznym liczba informacji pogłębia i utrwala lukę kompetencyjną. Brakuje jednoznacznej odpowiedzi na owe niedobory w środowisku nie tyle bogatym w informację, ile przesyconym informacją[12]. Dualizm Internetu wymusza przyswojenie wiedzy i nabycie umiejętności pochodzących z dwóch odległych od siebie obszarów – społecznego i technicznego, co niejako generuje konieczność profesjonalizacji i długi czas trwania takiego nauczania. W ramach społecznego obszaru użytkownik winien nabyć umiejętności komunikacji z rozmaitymi grupami użytkowników Internetu. Są to kom-

[11] M. Baran, E. Cichocka, P. Maranowski, W. Pander, *Cybernauci – diagnoza wiedzy, umiejętności i kompetencji dzieci i młodzieży, rodziców i opiekunów oraz nauczycieli w zakresie bezpiecznego korzystania z internetu. Raport podsumowujący badanie ex-ante*, Warszawa 2016, http://cybernauci.edu.pl/wp-content/uploads/2016/06/Cybernauci-diagnoza-wiedzy-umiejętności-i-kompetencji.-Raport.pdf [dostęp: 24.01.2019]. Warto zwrócić uwagę, że sami autorzy raportu skupiają się w nim co prawda na kompetencjach medialnych i równolegle „funkcjonalnych kompetencjach cyfrowych", jednak rozpraszając się (skądinąd również ważne, jednak nie w omawianym kontekście) kwestie bezpieczeństwa użytkowników w globalnej sieci, jak gdyby sam temat sprawnego pozyskiwania informacji nie był wystarczający. Zasygnalizujmy, że sami twórcy przytaczanego raportu nie odwołują się do prawdziwie profesjonalnych i pogłębionych procedur i algorytmów korzystania z sieci.

[12] D. Shenk, *Data Smog. Surviving the Information Glut*, Nowy Jork 1998.

petencje psychologiczne, komunikacyjne (w tym na przykład netykieta), a także językowe (przyswojenie socjolektów rozmaitych grup). Istotne są również kompetencje w zakresie ewaluacji informacji: rozpoznawania fałszu, wiedzy źródłoznawczej, posługiwania się rozmaitymi sposobami analizy informacji (ilościową i jakościową, z użyciem narzędzi informatycznych i bez nich), w tym logiką praktyczną. Spośród kompetencji technicznych nieodzowne wydaje się przyswojenie liczbowych, stochastycznych umiejętności oceny zjawisk, w szczególności fenomenów o charakterze masowym, długotrwałym lub globalnym. Niezbędne wydają się kompetencje *stricte* informatyczne, jak opanowanie języków zapytań (*query languages*), topografii Internetu, biegłość obsługi rozmaitych narzędzi służących pozyskiwaniu, przetwarzaniu i wizualizacji informacji.

Środowisko informacyjne człowieka zmienia się nieprzewidywalnie i w szybkim tempie, co winno wymusić równie głębokie zmiany w paradygmatach percepcji rzeczywistości. Zysku jednostki wyposażonej w taką wiedzę i jej umiejętności nie wolno mierzyć jedynie miarą jej erudycji czy swoistego „oświecenia". Jednostka wykazująca umiejętności sprawnego pozyskiwania, analizy i oceny informacji oraz zamiany jej na wiedzę zyskuje wolność informacyjną – niezależność od zjawisk takich jak dezinformacja i propaganda oraz zdolność do korzystania z jak najszerszego spektrum źródeł informacji. Zaniechanie wdrażania propozycji programów edukacyjnych w tym zakresie zaowocować może – w dalszej przyszłości – wieszczonym przez Stanisława Lema informacyjnym impasem rozwojowym, choć szkicowanym przezeń lekko i żartobliwie, to w swoich efektach ponurym[13].

Biorąc pod uwagę powyższe spostrzeżenia, obawy i troski przedstawiamy czytelnikowi tom zawierający efekty refleksji, ale także wyniki badań i wskazówki *stricte* praktyczne, w jaki sposób odróżniać prawdę od fałszu, w jaki sposób informacja jest manipulowana, jak ją pozyskiwać, analizować i oceniać.

Tom otwiera tekst Przemysława Potockiego zawierający refleksję teoriopoznawczą nad treścią i zakresem funkcjonowania zawodu infobrokera w cywilizacji informacyjnej i społeczeństwie informacyjnym. Autor jasno zdaje sobie sprawę ze złożoności oddziaływań na linii technologia – społeczeństwo, wyważając postawę pomiędzy dwoma fundamentalnymi nurtami myślenia o społecznych efektach technologii – determinizmie technologicznymi i jego dopełnieniu logicznym – podejściu *use tool*. Dokonuje przeglądu i systematyzacji pojęć, jednocześnie nakreślając uniwersalną

[13] S. Lem, *Wizja lokalna*, Kraków 1982, s. 93.

perspektywę zawodu infobrokera. Za ważki wkład własny w namysł nad infobrokeringiem należy uznać zaproponowany przez autora i stanowiący podsumowanie rozważań model infobrokeringu w ujęciu sekwencyjnym. Ma on znaczenie nie tylko poznawcze, ale również analityczne.

Piotr Sosnowski wykonał niewątpliwie potrzebną i żmudną pracę, porządkując i strukturyzując obszerny zbiór pojęć z kategorii danych rozpoznawczo-wywiadowczych związanych z metodyką pozyskiwania informacji. Autor wprowadza, definiuje i analizuje nazewnictwo i istotę głównych dyscyplin wywiadowczych, stosując przy tym podział na otwarte i zamknięte źródła informacji, wskazuje także dyscypliny wspierające oraz nowo powstałe (między innymi wywiad oparty na technologiach lokowania reklam w aplikacjach mobilnych, wywiad zbiorowy, wywiad rynkowy i naukowy oraz wywiad oparty na mediach społecznościowych).

Tekst Piotra Deli ogniskuje się na zjawisku propagandy, które ze względu na obfitość literatury przedmiotu może wydawać się doskonale rozpoznane. Jak uwidacznia artykuł, jest to tylko pozór – nowe zjawiska, w szczególności wykorzystanie cyberprzestrzeni jako medium komunikacyjnego, konstytuują nowe techniki zniekształcania informacji. W tekście odnajdziemy refleksję systematyzującą współczesne oblicze zjawiska propagandy – jej model oraz typologię. Autor lokuje także – co istotne – swoje rozważania nad propagandą w kontekście bezpieczeństwa informacyjnego.

Artykuł Magdy Tomaszewskiej zawiera zestawienie i systematyczną analizę obowiązujących uregulowań prawnych w zakresie pozyskiwania informacji. Autorka często sięga do perspektywy porównawczej – krajów zachodnich, co w tym wymiarze nadaje pracy walor prognostyczny określający, jakimi drogami mogą podążać rodzime pomysły i rozwiązania ustawodawcze. Niewątpliwym walorem tekstu jest fakt, iż dostarcza praktycznych wskazówek poruszania się pośród uregulowań prawnych wszystkim tym, którzy informacji w Internecie poszukują. Autorka w swoich rozważaniach – co należy uznać za zaletę – odnosi się do nowych zjawisk znajdujących się na styku prawa, pozyskiwania informacji i nowych technologii: wyłudzania informacji (*phishing*), prowokacji w celu pozyskania informacji, a także masowej dezinformacji z wykorzystaniem sfałszowanych wiadomości (*fake news*, *sockpuppets*).

Bezcennym dopełnieniem tekstów z niniejszego tomu ogniskujących się na pozyskiwaniu informacji z użyciem nowych technologii jest publikacja Urszuli Kurcewicz *Znaczenie tradycyjnych źródeł informacji w działalności infobrokerskiej*, ukazująca bogactwo, złożoność i wartość informacyjną źródeł tradycyjnych. Źródła te – jak słusznie podkreśla

autorka – mimo rozwoju technologii i masowej digitalizacji informacji nie utraciły swojej aktualności i znaczenia. W pozyskiwaniu informacji ze źródeł klasycznych, chciałoby się rzec – analogowych, jednym z kluczowych elementów są kompetencje społeczne, techniczne i merytoryczne, zwłaszcza w odniesieniu do pozyskiwaniu informacji z takich instytucji, jak państwowe i prywatne archiwa, biblioteki czy muzea.

Artykuł trojga autorów – Konrada K. Kuźmy, Joanny Lewczuk i Konrada Gałuszko – reprezentuje perspektywę ilościowego, stochastycznego podejścia do analizy zjawisk społecznych i politycznych. Badacze udanie przełamują negatywny stereotyp ilościowej analizy danych jako działania niepraktycznego i nieprzystępnego. Udowadniają, że stanowi ona narzędzie umożliwiające wygenerowanie wartościowej, zrozumiałej i przekładalnej na praktyczne wnioski wiedzy o zjawiskach i procesach politycznych. Podkreślić należy, że statystyka jest w swojej genezie i rozwoju ściśle związana z centralnym przedmiotem zainteresowań politologa – państwem. Wyłoniła się z potrzeby organizacji życia we wspólnotach, a rozwijała poza matematyką. Konstruowano i postrzegano ją jako ważny, a nawet nieodzowny filar sprawowania władzy; przez wieki była ona nierozdzielna z instytucjami państwa. Autorzy nie popadają jednocześnie w „numerolatrię" – liczbowe wyniki zaopatrzone są w tak potrzebny interpretacyjny „współczynnik humanistyczny", a surowy świat liczb wyjaśniany jest przystępnie i co najistotniejsze – praktycznie.

Autorki artykułu *Techniki wyszukiwania informacji w mediach społecznościowych dla celów białego wywiadu* – Patrycja Hrabiec-Hojda i Justyna Trzeciakowska, od lat związane z branżą infobrokerską – podzieliły się swoim bogatym doświadczeniem w zakresie eksploracji Facebooka i Twittera. Zogniskowały swoje rozważania na trzech następujących komplementarnych elementach: heurystykach wyszukiwania, egzemplifikacjach zapytań (kwerendy) oraz narzędziach wyszukiwawczych (StalkScan, Search Is Back, Recruitin.net, Onemilliontweetmap.com, Twitonomy). Tekst zawiera również wprowadzenie do technik *Google hacking*. Publikacja ma walor praktyczny, lokując się w słabo jeszcze poznanym w Polsce nurcie *Social Media Intelligence* (SMI, SOCMINT).

Artykuł *Sztuka wyszukiwania w Internecie – autorski przegląd wybranych technik i narzędzi* niżej podpisanego składa się z dwóch części tematycznych. W pierwszej przeanalizowano techniki wyszukiwania: w sensie ogólnym, to jest heurystyk oraz szczegółowym, czyli konkretnych technik należących do rodziny języków zapytań (*query languages*) – operatorów. Ich zasadniczą funkcję stanowi doprecyzowanie zapytań dla wyszukiwa-

rek. Druga część tekstu zawiera przegląd i analizę wybranych narzędzi eksploracji Internetu – wyszukiwarek internetowych.

W infobrokeringu szczególne znaczenie ma poszukiwanie wciąż nowych obszarów i sposobów pozyskiwania informacji. Taki wysiłek podjął Wojciech Mincewicz w artykule *Metadane – cichy zabójca prywatności*. Metadane to nieodłączny składnik komunikacji z użyciem sieci Internet i telefonii GSM. Jednocześnie stanowią tyleż powszechne, co nieuświadamiane zjawisko. Autor reprezentuje podejście praktyczne, jednak dokonuje również – udanej – próby usystematyzowania badanego zjawiska, czyli typologizacji metadanych. Istotne uzupełnienie tekstu stanowi wprowadzenie do zautomatyzowanej eksploracji metadanych z użyciem programu Fingerprinting Organizations with Collected Archives (FOCA).

Artykuł *Inwigilacja elektroniczna jako metoda pozyskiwania informacji – ewaluacja i prognozy* to multidyscyplinarna refleksja nad współczesnymi trendami w dziedzinie inwigilacji – diagnoza *status quo* i próba sformułowanie prognoz w zakresie trendów zjawisk związanych z naruszaniem prywatności. Autorzy podjęli wysiłek odpowiedzi na szereg istotnych w tym zakresie pytań. Po pierwsze, jakie typy negatywnych zjawisk są wytwarzane i intensyfikowane przez technologie inwigilacji elektronicznej? Po wtóre, jak głęboki jest stan „bezbronności inwigilacyjnej" współczesnych społeczeństw, to jest jakie są możliwości urządzeń służących inwigilacji? Po trzecie, czy istnieje możliwość praktycznego przeciwstawienia się im i – jeśli tak – w jaki sposób? Po czwarte, jaka jest geneza tych zjawisk i jakie spodziewane scenariusze przyszłości można szkicować na podstawie antycypacji zaobserwowanych trendów? Tak zdefiniowany zbiór pytań badawczych wymagał oglądu z dwóch perspektyw: socjologicznej i technicznej. Autorzy dostrzegają i analizują szereg negatywnych zjawisk związanych z inwigilacją elektroniczną – jej eskalację, profesjonalizację, instytucjonalizację i normalizację.

Daniel Mider

VOL. 54 STUDIA POLITOLOGICZNE
DOI: 10.33896/SPolit.2019.54.1

Przemysław Potocki
ORCID:0000-0001-8167-6747

Informacyjne determinanty rozwoju społeczno-ekonomicznego w XXI wieku: perspektywa infobrokeringu[1]

Słowa kluczowe:
informacja, społeczeństwo informacyjne, gospodarka cyfrowa, rynek pracy, infobroker

Wstęp

Informacja zawsze stanowiła w życiu społecznym cenny zasób – w polityce dostęp do niej nierzadko przesądzał o sukcesie lub porażce danego polityka, a w gospodarce stawała się przyczyną pojawienia się ogromnych fortun lub bankructw na międzynarodową skalę[2]. Dysponenci unikalnych informacji byli pożądanymi sojusznikami lub cennymi partnerami biznesowymi. Wiek XX przyniósł upowszechnienie dostępu do wiedzy, a środki masowego przekazu – jako ogólnodostępne źródło informacji – stały się jednymi z najważniejszych aktorów życia publicznego. Ich wpływ na społeczeństwo i gospodarkę stał się immanentnym elementem rzeczywistości[3]. Informacja może być czynnikiem postępu naukowego oraz przesłanką zapewnienia bezpieczeństwa narodowego

[1] Rozszerzona i zmodyfikowana wersja referatu *Wybrane aspekty zmian technologicznych na rynkach pracy państw rozwiniętych w kontekście cyberbezpieczeństwa* przedstawionego podczas IV Kongresu Politologii (18–20.09.2018, Lublin). Podziękowania dla pani Sylwii Szołuchy za przeprowadzenie kwerendy polskich źródeł dotyczących infobrokeringu.

[2] Szerzej na temat znaczenia wiedzy i informacji w rozwoju społecznym Europy: P. Burke, *Społeczna historia wiedzy*, przekł. A. Kunicka, Warszawa 2016, s. 141–211.

[3] Zob. L. Gorman, D. McLean, *Media i społeczeństwo. Wprowadzenie historyczne*, przekł. A. Sadza, Kraków 2010.

– od charakteru aktywności dysponentów danej informacji zależy, jakie będą skutki jej wykorzystania[4]. Głównym celem tego artykułu jest wskazanie teoretycznych, społecznych i ekonomicznych przesłanek, które przyczyniły się do powstania takich stosunków społeczno-ekonomicznych na początku XXI wieku, w których dostęp decydenta do informacji jest uznawany za jeden z kluczowych czynników eksplanacyjnych. Oczywiście jest to zagadnienie interdyscyplinarne, które zostało potraktowane w sposób aspektowy ze względu na bardzo liczne dotyczące go źródła. Infobrokering został omówiony jako jedna z egzemplifikacji współzależności zmian w społeczeństwie informacyjnym i gospodarce cyfrowej.

W pierwszej części przedstawiono sposoby rozumienia pojęcia „informacja" w naukach społecznych. Zostały nakreślone różne jego ujęcia teoretyczne. Przedmiotem analizy stał się również ilościowy i jakościowy aspekt informacji. Uwzględniono także jej wymiar procesualny i treściowy. Część druga, dotycząca społeczeństwa informacyjnego, zawiera omówienie etapowego charakteru tej fazy rozwoju społecznego. Zaprezentowano wybrane pozytywne i negatywne strony tego procesu, a jednocześnie ukazano jego empiryczne konsekwencje – dane statystyczne. W części trzeciej, która skupiła się na omówieniu wybranych danych ilościowych dotyczących obecnych form gospodarki cyfrowej, ukazano także dane statystyczne obejmujące wybrane aspekty gospodarki cyfrowej Unii Europejskiej. W części czwartej zaprezentowano koncepcję „ekosystemu informacyjnego" oraz rolę, jaką pełnią w nim profesjonaliści informacji. Omówiono definicyjne aspekty zjawiska infobrokeringu. Zostały także przedstawione praktyczne aspekty prowadzenia aktywności infobrokerskiej: źródła informacji, wymagania zawodowe. Rozważania w tej części zamyka autorski model infobrokeringu w ujęciu sekwencyjnym. Cześć wnioskowa stanowi projekcję zmian uwarunkowań działalności infobrokerskiej w kontekście ewolucji społeczeństwa informacyjnego i gospodarki cyfrowej w XXI wieku.

Informacja jako element procesów społecznych i ekonomicznych

Informacja jest pojęciem, które należy do terminologicznego kanonu nauk społecznych. Pojawia się w analizach i badaniach z zakresu nauki o polityce i administracji, o bezpieczeństwie, nauki o komunikacji społecznej i mediach, o zarządzaniu, socjologii, pedagogiki, psychologii.

[4] Zob. M. Minkina, *Sztuka wywiadu w państwie współczesnym*, Warszawa 2014, s. 167.

Zalicza się do terminów podstawowych, które są rozmaicie definiowane – w zależności od orientacji teoretycznej badacza i jego celów poznawczych. Jak pisze Tomasz Goban-Klas: „W jej najszerszym rozumieniu informacja, to tyle, co treści przekazywane w procesie komunikacji międzyludzkiej”[5]. Informacja ma zatem relacyjny charakter i pojawia się w momencie interakcji między jednostkami w formie danych zmysłowych (dwa najczęściej wykorzystywane źródła danych to zmysł wzroku i zmysł słuchu). Należy także zaznaczyć, że informacja – w tym szerokim ujęciu – pojawia się zawsze w polu świadomości jednostki i pełni funkcję regulacyjną, czyli określa parametry jej zachowania w danej sytuacji społecznej. Z etymologicznego punktu widzenia znaczenia terminu „informacja” należy szukać w języku łacińskim w dwóch słowach: 1. *informatio* – przedstawienie, wizerunek, 2. *informare* – kształtować, przedstawiać[6]. Zakres znaczeniowy tego terminu ma charakter czynnościowy i wskazuje na aktywność poznawczą podmiotów, które uczestniczą w procesie informacyjnym. Termin „informacja” może być rozumiany dwojako: w ujęciu obiektywistycznym jest to określony stan otoczenia, a w ujęciu subiektywistycznym to akt o charakterze percepcyjno-twórczym[7]. Badacze, którzy stosują termin „informacja” w swoich analizach, posługują się jednym z tych dwóch sposobów jego rozumienia. Ujęcie obiektywistyczne jest powiązane z ilościowym sposobem rozumienia istoty informacji (cybernetyka, informatyka, matematyka), natomiast ujęcie subiektywistyczne opiera się na jakościowym podejściu do informacji (co jest typowe dla nauk społecznych).

Informacja w ujęciu ogólnym ma następujące cechy: 1. jest obiektywną własnością świata związaną z jego relacyjnym charakterem; 2. odwzorowuje jeden układ w innym układzie przy zachowaniu określonych cech układu pierwotnego; 3. własność procesów poznawczych związana jest z przetwarzaniem sygnałów z otoczenia zewnętrznego; 4. ma pragmatyczną własność czynności sterowania różnymi układami[8]. Uzyskanie dostępu do informacji zapewnia aktorom życia społecznego możliwość zredukowania poziomu niepewności (pojawienia się zdarzeń przypadko-

---

5 T. Goban-Klas, Hasło: *Informacja. Manipulacja informacją*, [w:] W. Szewczuk (red.), *Encyklopedia psychologii*, Warszawa 1998, s. 126.

6 Hasło: *Informacja*, [w:] *Powszechna encyklopedia filozofii (Internetowa wersja)*, s. 1, Polskie Towarzystwo Tomasza z Akwinu, http://www.ptta.pl/pef/pdf/i/Informacja.pdf (dostęp: 12.01.2019).

7 Zob. J. Mikułowski Pomorski, *Informacja i komunikacja: pojęcia, wzajemne relacje*, Wrocław 1988, s. 18.

8 M. Hetmański, *Epistemologia informacji*, Kraków 2013, s. 103.

wych, czyli takich, którym przypisujemy prawdopodobieństwo pojawienia się mniejsze od jeden)[9] i wyboru takiej alternatywy decyzyjnej, której realizacja zapewni osiągnięcie celu działania.

Witold Nawrocki na podstawie analizy definicji pojęcia „informacja" wyodrębnił trzy podejścia analityczne występujące w literaturze przedmiotu: 1. podejście strukturalistyczne – informacja jest zawarta w każdym bycie i stanowi (wraz z materią i energią) trzeci element każdego bytu, a poznanie pełnej informacji o danym obiekcie jest niemożliwe; 2. podejście interakcyjne – informacja jest generowana w ramach procesów interakcji między bytami, a jej pojawienie się wywołuje reakcje odbiorcy (informacja pojawia się w układzie relacyjnym), 3. podejście komunikatywistyczne – nośnikiem informacji jest sygnał, ma zarówno treść jak i znaczenie, informacja pojawia się w procesie interpretacji sygnału przez odbiorcę[10]. W podejściu strukturalistycznym kluczową rolę odgrywa więc nośnik informacji, w podejściu interakcyjnym podkreśla się sytuacyjność i zmienność informacji, a podejście komunikatywistyczne podkreśla znaczenie umiejętności poznawczych odbiorcy.

Informacja, zarówno w aspekcie ilościowym jak i jakościowym, jest tworzona w ramach określonego systemu komunikacyjnego. Za jeden z najczęściej omawianych przez badaczy model systemu komunikacyjnego należy uznać zaproponowany w 1948 roku przez Claude'a E. Shannona model ogólnego systemu komunikacyjnego (ang. *general communication system*), który składa się z pięciu elementów: 1. źródło informacji – wytwarza określoną wiadomość lub zespół wiadomości (składa się z sygnałów); 2. nadajnik – przetwarza wiadomość w celu dokonania jej transmisji w postaci sygnału; 3. kanał komunikacyjny – pełni funkcję pośredniczącą w procesie transmisji sygnału od nadajnika do odbiornika; 4. odbiornik – odbiera sygnał i przekształca go w wiadomość; 5. adresat wiadomości[11]. W kanale komunikacyjnym mogą pojawiać się szumy komunikacyjne, które zmieniają zakres treściowy pierwotnego komunikatu, mogąc prowadzić do jego zmiany w odbiorniku. Im większy zakres szumów komunikacyjnych, tym wyższy poziom zniekształcenia treści wiadomości odebranej w porównaniu z treścią wiadomości nadanej.

---

[9] M. Heller, *Filozofia przypadku. Kosmiczna fuga z preludium i codą*, Kraków 2016, s. 152.

[10] W. Nawrocki, *W poszukiwaniu istoty informacji*, [w:] J.J. Jadacki (red.), *Analiza pojęcia informacja*, Warszawa 2003, s. 53.

[11] C.E. Shannon, *A Mathematical Theory of Communication*, „The Bell System Technical Journal" (Reprinted with corrections), t. 27, s. 379–423, 623–656, lipiec, sierpień 1948, s. 2, http://math.harvard.edu/~ctm/home/text/others/shannon/entropy/entropy.pdf (dostęp: 12.01.2019).

Wprawdzie pierwotnie za źródło szumów uznawano jedynie parametry techniczne kanału komunikacyjnego, ale adaptacja tego modelu przez przedstawicieli nauk społecznych rozszerzyła listę czynników w ramach zbioru czynników zniekształcających sygnał (na przykład szumy semantyczne).

Informacje przekazywane w formie wiadomości zawierają określoną treść, która jest uwarunkowana przez cele komunikacyjne nadawcy, charakterystykę procesu informacyjnego, kontekst komunikacyjny, aspekt strukturalny i formalny. Do najczęściej spotykanych typów wiadomości wyróżnionych ze względu na charakterystykę treści należą informacje:

- faktograficzne – opis obiektu w ujęciu aspektowym w określonym przedziale czasu;
- komparatywne – porównania danego obiektu w różnych przedziałach czasu lub porównania więcej niż jednego obiektu;
- semantyczne – znaczenie posiadane przez dany obiekt (na przykład definicje);
- normatywne – określają warunki funkcjonowania obiektu;
- klasyfikacyjne – kryteria przypisania obiektu do danej klasy obiektów;
- strukturalne – opis wewnętrznych powiązań elementów obiektu;
- przestrzenne – opis obiektu ze względu na jego położenie w przestrzeni (dwa lub trzy wymiary);
- imperatywne – nakazują obiektowi określone działanie;
- kwerencyjne – pytania na temat obiektów i stanów rzeczywistości;
- fatyczne – pełnią funkcję regulacyjną w życiu społecznym[12].

Tworzenie treści informacji wynika z funkcji, jaką ma ona pełnić w życiu społecznym, Uwzględniając wcześniejsze rozważania, należy zauważyć, że w ujęciu obiektywistycznym informacja stanowi niezależny od obserwatora (odbiorcy) element rzeczywistości w wymiarze materialnym, ma wymiar ilościowy. Może zatem podlegać procesom kwantyfikacji i pomiaru, określania wartości w wymiarze pieniężnym, przekształceniom według reguł matematycznych. Ma procesualny charakter, jest to bowiem „[…] zdarzenie, które dokonuje wyboru stanów systemu"[13]. Informacja pochodząca z obiektywnego źródła i mająca ilościowy charakter pełni funkcję regulacyjną wobec struktury społecznej oraz zbiorowych aktorów życia społecznego. W ujęciu drugim, nazwanym subiektywistycznym,

[12] B. Stefanowicz, *Informacja. Wiedza. Mądrość*, Biblioteka Wiadomości Statystycznych, t. 66, Warszawa 2013, s. 28–32, http://stat.gov.pl/cps/rde/xbcr/gus/OZ_Informacja_Wiedza_Madrosc_180413.pdf (dostęp: 21.06.2018).

[13] N. Luhmann, *Systemy społeczne. Zarys ogólnej teorii*, przekł. M. Kaczmarczyk, Kraków 2007, s. 69.

informacja powstaje jako efekt relacji pomiędzy nadawcą a obserwatorem, każdorazowo jej charakterystyka jest nieco inna, a więc identyfikacja jej treści dokonuje się w wymiarze jakościowym. Taka informacja jest zjawiskiem społecznym, czyli każdorazowo kształtuje się w innych warunkach komunikacyjnych. Spełnia funkcję kreacyjną, czyli warunkuje okoliczności, w których jednostka może odbierać informacje, aby utrzymać homeostazę na poziomie osobniczym[14]. Określa pole percepcji jednostki poprzez jej kompetencje językowe i kulturowe, które zapewniają interpretację danych zmysłowych na poziomie symbolicznym.

Marek Hetmański podkreśla, że w epistemologii informacji fundamentalne znaczenie ma kwestia procesualnego charakteru informacji, która pojawia się w wyniku zajścia określonej sekwencji działań ze strony podmiotu dążącego do uzyskania informacji: „Bez obliczania nie ma informacji, są tylko sygnały, znaki, symbole, dane i dowolne ich konfiguracje, których największe nawet ilościowe parametry (rozległość, szybkość, złożoność) nie decydują o informacji"[15]. Transformacja danych w informacje dokonuje się według określonych reguł – matematycznych, językowych, czynnościowych. Od celów poznawczych i możliwości technicznych podmiotu dążącego do uzyskania konkretnej informacji zależy, jakiego rodzaju reguły zostaną w procesie obliczania wykorzystane.

W matematycznej teorii informacji za podstawowe kryterium informacyjnej wartości wiadomości należy uznać poziom jej nieokreśloności (entropii[16]): im większa niepewność co do powstania danej wiadomości ze zbioru wiadomości potencjalnie dostępnych dla odbiorcy, tym większa wartość takiej wiadomości – cenniejszą wiadomością jest taka, którą uzyskamy ze zbioru dziesięciu niż taka, którą pozyskamy ze zbioru składającego się z miliona elementów[17]. Zatem w społeczeństwie informacyjnym i gospodarce cyfrowej, które bazują na informacji jako kluczowym zasobie, wartość informacji jest o wiele wyższa niż w przypadku wcześniejszych faz rozwoju społeczeństwa i gospodarki ze względu na rosnący w sposób wykładniczy zbiór danych, które wymagają przetworzenia.

Ze względu na kryterium pomiaru informacji można wyróżnić dwa podstawowe podejścia teoretyczne umożliwiające identyfikację założeń

---

[14] Zob. I. Ihnatowicz, *Człowiek. Informacja. Społeczeństwo*, Warszawa 1989, s. 95.

[15] M. Hetmański, *Epistemologia...*, s. 320.

[16] Pojęcie to rozumiane jest tu jako „entropia informacyjna", czyli uwzględnia stan wiedzy odbiorcy informacji. Zob. L. Ciborowski, *Walka informacyjna*, Toruń 1999, s. 58.

[17] J.R. Pierce, *Symbole, sygnały i szumy, Wprowadzenie do teorii informacji*, przekł. J. Mieścicki, R. Gomulicki, Warszawa 1967, s. 41.

ontologicznych i epistemologicznych, którymi kierują się badacze zajmujący się zjawiskiem informacji w rzeczywistości społecznej:

- ilościowe (matematyczne, cybernetyczne) – informacja może być przedmiotem kwantyfikacji, podlega procesom redukcji lub akumulacji, a jej nadmiar stanowi źródło szumu informacyjnego;
- jakościowe (prawo, psychologia, ekonomia) – informacja odzwierciedla niematerialne elementy rzeczywistości społecznej i może być odbierana w sposób wieloznaczny.

W podejściu ilościowym „Jeden bit informacji określa wybór jednej z dwu jednakowo prawdopodobnych wielkości, którymi mogą być liczby lub wiadomości przeznaczone do przesłania"[18]. Natomiast w podejściu jakościowym informacja ma charakter subiektywny i jest uzależniona od takich czynników, jak: czas do namysłu, treść wiedzy użytkownika, kontekst poszukiwania informacji, stan emocjonalny użytkownika, okoliczności odbioru[19]. Podejścia ilościowe i jakościowe należy uznać za komplementarne. Jeśli odbiorcą informacji jest jednostka, zakres informacji uzyskiwanej przez nią z otoczenia zewnętrznego podlega regulacjom wynikającym z norm prawnych, kulturowych, czynników psychologicznych. Wydaje się, że aspektowość informacji trafnie można ująć w następującej definicji: informacja to sens znaczeniowy nadany przez jej odbiorcę treściom odebranego przezeń sygnału[20]. Docierające do odbiorcy dane zmysłowe są przedmiotem przekształceń w systemie poznawczym jednostki, które są uwarunkowane zarówno charakterystyką fizyczną, jak i semantyczną komunikatu. Im wyższe kompetencje poznawcze jednostki i potencjał jej psychofizyczny, tym wyższą wartość informacyjną może uzyskać komunikat, którego adresatem jest ta jednostka.

Wartość informacji zależy od dwóch jej parametrów: ilości i jakości[21]. W modelowym ujęciu możliwe są cztery stany informacji wynikające z dwóch wartości obu parametrów. Stan pierwszy to informacja o wysokim poziomie zawartości treściowej i wysokim poziomie jakości (to sytuacja najbardziej korzystna z punktu widzenia odbiorcy informacji – umożliwi mu bowiem sformułowanie wielu alternatyw decyzyjnych). Stan drugi

---

18 Tamże, s. 23.

19 B. Stefanowicz, *Informacja…*, s. 13–14.

20 W. Nawrocki, *W poszukiwaniu istoty…*, s. 56.

21 W ujęciu teoretycznym dotyczącym kooperacji negatywnej zasoby informacyjne mają dwa rodzaje wartości: 1. wymienną – ma określoną wartość rynkową możliwą do wyznaczenia w formie ceny, 2. operacyjną – zakres korzyści możliwych do uzyskania dzięki uzyskaniu dostępu do informacji. Zob. D.E. Denning, *Wojna informacyjna i bezpieczeństwo informacji*, przekł. J. Bloch, Warszawa 2002, s. 25–28.

to informacja o niskim poziomie zawartości treściowej i niskim poziomie jakości (decydent uzna poziom jej użyteczności za uniemożliwiający zaspokojenie jego potrzeb informacyjnych). Stany pośrednie: wysoki poziom zawartości treściowej połączony z niskim poziomem jakości lub niski poziom zawartości treściowej występujący równocześnie z wysokim poziomem jakości. Określenie poziomu jakości informacji jest możliwe poprzez dokonanie analizy atrybutów informacji: relewancji, dokładności, kompletności, aktualności, spójności, odpowiedniości formatu, dostępności, kompatybilności, bezpieczeństwa, wiarygodności[22]. Wydaje się, że za najważniejsze należy uznać dwa atrybuty: relewancję i wiarygodność – informacja ma znaczenie praktyczne dla odbiorcy i w prawdziwy sposób przedstawia element rzeczywistości społecznej, którego dotyczą jego potrzeby informacyjne.

Informacja pojawia się jako rezultat wystąpienia określonej sekwencji działań, w której żadna faza (etap) nie może zostać pominięta, a zaspokojenie potrzeb informacyjnych decydenta wymaga przejścia danych wyjściowych przez wszystkie etapy tego procesu. Jest on określany jako proces informacyjny, na który składają się następujące elementy: generowanie, gromadzenie, przechowywanie, transmisja, transformacja, udostępnianie, interpretacja, wykorzystywanie[23]. Przykładowo: kandydat w wyborach parlamentarnych zamawia raport poparcia wyborczego konkurentów w ujęciu przestrzennym. Faza generowania to uzyskanie danych wyborczych od Państwowej Komisji Wyborczej. Faza gromadzenia to stworzenie pliku Excela z danymi wyborczymi. Faza przechowywania to zapisanie takiego pliku na komputerze analityka. Następnie następuje przesłanie danych zawartych w pliku Excela do programu komputerowego umożliwiającego ilościową analizę danych – jest to faza transmisji. W kolejnej fazie – transformacji – dane poparcia wyborczego zostają poddane analizie statystycznej. Efekty tej analizy w formie graficznej i tekstowej zostają zapisane w postaci raportu, który trafia do klienta – faza udostępniania. Raport zostaje przedstawiony przez analityka klientowi w ramach fazy interpretacji. Wnioski z takiej prezentacji są implementowane podczas fazy wykorzystywania w trakcie kampanii wyborczej (na przykład służą do geotargetowania).

---

[22] Zob. I. Swoboda, *Jakość informacji*, [w:] S. Cisek, A. Januszko-Szakiel (red.), *Zawód infobroker. Polski rynek informacji*, Warszawa 2015, s. 244–246.

[23] J. Oleński, *Ekonomika informacji. Metody*, Warszawa 2003, s. 41.

Społeczeństwo informacyjne: geneza, charakterystyczne cechy, rozwój

Informacja jest zarówno rezultatem jak i determinantą procesów społecznych. Przez Józefa Oleńskiego jest ona określana jako funkcjonalne minimum informacyjne. Należy podkreślić, że w społeczeństwie informacyjnym, w przeciwieństwie do poprzednich faz rozwoju życia społecznego (faza społeczeństwa rolniczego, faza społeczeństwa przemysłowego), dostępne zasoby informacji stają się dysfunkcjonalne, gdyż: „[…] minimum funkcjonalne znacznie przekracza zdolności percepcyjne człowieka. Przekracza ono często zasoby informacyjne oraz możliwości organizacyjne, techniczne i gospodarcze jednostki organizacyjnej”[24]. Jeśli przyjmiemy perspektywę badawczą, którą proponuje nurt determinizmu technologicznego[25], iż rozwój techniczny społeczeństw jest główną determinantą rozwoju nowych form życia społecznego, to za jeden z głównych czynników aktywizujących przejście od fazy społeczeństwa industrialnego do społeczeństwa informacyjnego należy uznać zmianę relacji pomiędzy trzema sferami życia społecznego: infosferą, technosferą i socjosferą (w fazie społeczeństwa przemysłowego wszystkie trzy sfery tworzyły komplementarny system[26]) – technosfera zaczęła dominować nad dwiema pozostałymi sferami po zakończeniu II wojny światowej. Polityczny aspekt tego procesu reorganizacji relacji pomiędzy tymi trzema sferami życia społecznego najlepiej obrazuje rezultat technologicznego wyścigu w ramach zimnej wojny, którą przegrał Związek Radziecki – nie był w stanie wytworzyć sprzętu informatycznego dysponującego mocą obliczeniową równą tej, którą posiadały Stany Zjednoczone[27]. W społeczeństwie informacyjnym głównym sektorem gospodarki, jeśli chodzi o skalę zatrudnienia, jest sektor usług, na drugim miejscu znajduje się przemysł, a na ostatnim miejscu jest rolnictwo[28].

Za kluczowy element strukturalny społeczeństwa informacyjnego (ang. *information society*) na obecnym etapie jego rozwoju należy uznać

---

[24] Tamże, s. 277.

[25] Zob. M. McLuhan, *Zrozumieć media: przedłużenia człowieka*, przekł. N. Szczucka, Warszawa 2004.

[26] A. Toffler, *Trzecia fala*, przekł. E. Woydyłło, Warszawa 1997, s. 78–79.

[27] Szerzej na ten temat: M. Castells, *Koniec tysiąclecia*, przekł. J. Stawiński, S. Szymański, Warszawa 2009, s. 20–30.

[28] Por. International Labour Organisation, *World Employment and Social Outlook: Trends 2018*, 22.01.2018, s. 30, https://www.ilo.org/wcmsp5/groups/public/---dgreports/---dcomm/---publ/documents/publication/wcms_615594.pdf (dostęp: 9.10.2018).

Internet, którego geneza jest nierozerwalnie związana z wyścigiem zbrojeń w ramach zimnej wojny. Stanowi on główne źródło generowania informacji w tej fazie rozwoju życia społecznego. Przeciętna dzienna aktywność informacyjna wszystkich użytkowników Internetu w 2016 roku obejmowała następujące rodzaje aktywności komunikacyjnej: 8,8 mld odsłon plików w serwisie YouTube, 4,2 mld wyszukań w serwisie internetowym Google, 2,3 mld GB przesłanych plików, 207 mld wysłanych maili, 803 mln tweetów, 186 mln fotografii umieszczonych w fotograficznym serwisie społecznościowym Instagram, 152 mln rozmów w komunikatorze internetowym Skype, 36 mln transakcji w sklepie internetowym Amazon[29]. To statystyki dotyczące powszechnie dostępnej części Internetu, ale aktywność użytkowników tego medium równie ożywiona, choć o mniejszej skali występowania, dotyczy „ukrytej sieci" (ang. *deep web*) oraz tzw. *dark web*. W przypadku tej ostatniej szacuje się, że w Polsce korzysta z niej nie mniej niż 25 tys. osób[30], które mogą stanowić grono potencjalnych inicjatorów działań naruszających obszar cyberbezpieczeństwa.

Koncepcja teoretyczna społeczeństwa informacyjnego pojawiła się w latach 60. XX wieku w Japonii, aczkolwiek trudno jednoznacznie wskazać, kto jest jej głównym twórcą – wydaje się, iż prawdziwa będzie konstatacja o zbiorowym jej autorstwie[31]. Teoretyczne podstawy uległy od czasu powstania rozszerzeniu oraz wzbogaceniu o dane empiryczne. Z teoretycznego punktu widzenia można przyjąć, za Manuelem Castellsem i Pekką Himanenem, że: „Podobnie jak społeczeństwa przemysłowe, społeczeństwa informacyjne na całym świecie mają pewne wspólne cechy strukturalne: zasadzają się na generowaniu wiedzy oraz przetwarzaniu informacji za pomocą technologii informacyjnych opartych na mikroelektronice. Są one zorganizowane w sieci, a ich centralne czynności powiązane są w sieć na skalę globalną, operują zaś, dzięki infrastrukturze telekomunikacyjnej oraz transportowej, jako jednostka w czasie rzeczy-

---

[29] World Bank, *World Development Report 2016: Digital Dividends*, Waszyngton 2016, s. 6, http://documents.worldbank.org/curated/en/896971468194972881/pdf/102725-PUB-Replacement-PUBLIC.pdf (dostęp: 16.06.2018).

[30] K. Kucharczyk, *Ciemna strona internetu pod lupą*, „Rzeczpospolita", 28.02.2018, http://www.rp.pl/Media-i-internet/180229305-Ciemna-strona-internetu-pod-lupa.html (dostęp: 14.06.2018).

[31] L.Z. Karvalics, *Information Society – What is it Exactly? (The Meaning, History and Conceptual Framework of an Expression)*, Budapeszt 2007, s. 5, http://www.msu.ac.zw/elearning/material/1349116439Information-Society-whatis.pdf (dostęp: 20.11.2017).

wistym"[32]. Tak więc społeczeństwo informacyjne tworzy się, reprodukuje i rozwija w sytuacji, gdy traktuje swoje otoczenie zewnętrzne jako swoisty „ekosystem informacyjny", pobierając z niego te informacje, które umożliwiają mu zachowanie stanu funkcjonalnej równowagi wewnątrzsystemowej. Podstawowymi elementami tego społeczeństwa są sieci, które można postrzegać jako heterogeniczne zbiory elementów organizacyjnych. Są one zdolne – dzięki dostępowi do informacji – do zorganizowania się w homogeniczne (z funkcjonalnego punktu widzenia) niehierarchiczne struktury zadaniowe działające w trybie „online".

Według definicji Głównego Urzędu Statystycznego, którą można traktować jako operacjonalizacyjne ujęcie przedstawionych powyżej zależności między informacją a siecią, społeczeństwo informacyjne jest to „Społeczeństwo znajdujące się na takim etapie rozwoju, na którym osiągnięty poziom techniki informatyczno-telekomunikacyjnej stwarza warunki techniczne, ekonomiczne, edukacyjne i inne do wykorzystywania informacji w produkcji wyrobów i świadczeniu usług. Społeczeństwo takie zapewnia obywatelom powszechny dostęp i umiejętność korzystania z technologii informacyjnych w ich działalności zawodowej, społecznej, w celu podnoszenia i aktualizacji wiedzy, korzystania ze zdobyczy kultury, ochrony zdrowia oraz spędzania wolnego czasu i innych usług mających wpływ na wyższą jakość życia"[33]. W treści tej definicji zwraca uwagę element powszechnego charakteru dostępu do technologii informacyjnych – to makroaspekt funkcjonowania tego typu społeczeństwa, a także umiejętność korzystania przez jednostki do realizacji swoich celów – jest to zatem mikroaspekt. W sytuacji, gdy zarówno makroaspekt, jak i mikroaspekt mają charakter sektorowy, czyli pozbawiony charakteru powszechności, społeczeństwo informacyjne można traktować jako typ idealny, który jest punktem odniesienia dla analiz porównawczych o charakterze ilościowym.

Poziom rozwoju społeczeństwa informacyjnego można określić poprzez kwantyfikację trzech wymiarów odnoszących się do technologii informacyjno-komunikacyjnych (ang. *information and communication technologies* – skrót ICTs), które tworzą Indeks rozwoju ICTs. Te wymiary to: 1. dostęp do ICTs (mierzony dostępem ludności do telefonii stacjonarnej oraz do telefonii komórkowej, szybkością transferu danych w Interne-

[32] M. Castells, P. Himanen, *Społeczeństwo informacyjne i państwo dobrobytu. Model fiński*, przekł. M. Penkala, M. Sutowski, Warszawa 2009, s. 21.

[33] Hasło: *Społeczeństwo informacyjne*, [w:] Główny Urząd Statystyczny, *Pojęcia stosowane w statystyce publicznej*, http://stat.gov.pl/metainformacje/slownik-pojec/pojecia-stosowane-w-statystyce-publicznej/1869,pojecie.html (dostęp: 14.06.2018).

cie, odsetkiem gospodarstw domowych posiadających komputer, odsetkiem gospodarstw domowych dysponujących dostępem do Internetu); 2. korzystanie z ICTs (odsetek osób korzystających z Internetu, odsetek osób korzystających z szybkiego Internetu, odsetek osób korzystających z dostępu do mobilnego Internetu); 3. umiejętności związane z wykorzystaniem ICTs (liczba lat w szkole, odsetek osób rozpoczynających szkołę średnią, odsetek osób rozpoczynających szkołę wyższą)[34]. W Indeksie rozwoju ICTs pierwszy i drugi wymiar są uznane za ważniejsze składowe (waga 0,4), a wymiar trzeci ma charakter drugoplanowy (waga 0,2). W Indeksie rozwoju ICTs dla roku 2017[35], który przygotowała Międzynarodowa Unia Telekomunikacyjna (ang. International Telecommunication Union), na pierwszych 10 miejscach znalazły się kolejno: Islandia, Korea Południowa, Szwajcaria, Dania, Wielka Brytania, Hong Kong, Holandia, Norwegia, Luksemburg, Japonia. Polska była na 49. miejscu, a przed nami znalazły się pozostałe państwa Europy Środkowej – Czechy na 43. miejscu, Słowacja na 46., Węgry na 48.). Polska nie była zatem regionalnym liderem w 2017 roku, a różnica wielkości wskaźnika między nią i Czechami jest większa niż pomiędzy Polską i Węgrami[36].

Rozwój społeczeństwa informacyjnego przebiega w określonym kontekście demograficznym. O ile rozwój społeczeństwa rolniczego był możliwy dzięki zachowaniu stabilności istniejącej struktury demograficznej (utrzymująca się wyraźna przewaga osób w wieku przedprodukcyjnym nad osobami w wieku poprodukcyjnym), a społeczeństwa przemysłowego dzięki dynamicznemu przyrostowi osób w wieku produkcyjnym, o tyle rozwojowi społeczeństwa informacyjnego (bo trudno uznać to za czynnik determinujący ten rozwój) towarzyszy wzrost liczby osób w wieku poprodukcyjnym, który w drugiej połowie tego wieku doprowadzi do wyraźnej przewagi ilościowej osób w wieku poprodukcyjnym nad osobami w wieku przedprodukcyjnym – po raz pierwszy w historii ludzkości. Prognozy Organizacji Narodów Zjednoczonych w tym zakresie są dość

---

[34] International Telecommunication Union, *Measuring the Information Society Report*, t. 1, Genewa 2017, s. 27, https://www.itu.int/en/ITU-D/Statistics/Documents/publications/misr2017/MISR2017_Volume1.pdf (dostęp: 12.06.2018).

[35] Tamże, s. 31.

[36] W Polsce w 2017 r. z Internetu korzystało 76,0% osób w wieku 16–74 lat (w porównaniu z 2010 r. wzrost o 17,2 punktów proc.) – w miastach 80,3% (wzrost o 15,2 punktów proc.), na wsi 69,3% (wzrost o 21,2 punktów proc.), zob. Główny Urząd Statystyczny, *Rocznik Statystyczny Rzeczypospolitej Polskiej 2018*, Warszawa 2018, s. 440; https://stat.gov.pl/download/gfx/portalinformacyjny/pl/defaultaktualnosci/5515/2/18/1/rocznik_statystyczny_rzeczypospolitej_polskiej_2018.pdf (dostęp: 13.01.2019).

jednoznaczne – w porównaniu z rokiem 2017 liczba osób na świecie w wieku 60 lat lub więcej podwoi się do roku 2050, a do roku 2100 potroi, osiągając w roku 2050 liczbę 2,1 mld osób, a w 2100 – 3,1 mld. Liczba osób sędziwych (w wieku 80 lat lub więcej) potroi się do roku 2050 i zwiększy się siedmiokrotnie do roku 2100[37]. Takie zmiany w strukturze demograficznej poszczególnych państw (najpierw w Europie, Japonii, Ameryce Północnej, a do roku 2100 na całym świecie) oznaczają konieczność zredefiniowania i przebudowania rynków pracy (na przykład poprzez robotyzację) oraz systemu dostępu do usług publicznych.

Ze zjawiskiem starzenia się populacji społeczeństw rozwiniętych współwystępuje zjawisko cyfrowego podziału (ang. *digital divide*), które jest operacjonalizowane przez badaczy jako „cyfrowe wykluczenie: […] dotyczy różnic pomiędzy osobami, które mają regularny dostęp do technologii informacyjno-komunikacyjnych i potrafią efektywnie z niego skorzystać a tymi, które tego dostępu nie mają"[38]. Zjawisko cyfrowego podziału ma dwa wymiary: fizyczny (bezpośredni kontakt jednostki z technologiami informacyjno-komunikacyjnymi, dostęp gospodarstw domowych do infrastruktury sieci teleinformatycznych) oraz psychologiczny (umiejętności posługiwania się przez jednostkę ICT's w życiu prywatnym i zawodowym). Do grup demograficznych i zawodowych, które w największym stopniu są narażone na negatywne konsekwencje występowania cyfrowego podziału, należy zaliczyć: seniorów, osoby o niskim poziomie wykształcenia, osoby o niskim poziomie dochodów, mieszkańców terenów wiejskich, rolników, osoby pozostające bez pracy, osoby niepełnosprawne. Na podstawie kryterium poziomu umiejętności cyfrowych (ang. *digital skills*) wyróżnia się trzy główne klasy społeczne w społeczeństwie informacyjnym:

- proletariat – osoby, które nie korzystają z technologii informacyjno-komunikacyjnych;
- digitariat – osoby, które korzystają z tych technologii w sposób powierzchowny;

[37] United Nations, *World Population Prospects: The 2017 Revision, Key Findings and Advance Tables*, ESA/P/WP/248, Nowy Jork, 2017, s. 13, https://esa.un.org/unpd/wpp/publications/Files/WPP2017_KeyFindings.pdf (dostęp: 7.06.2018).

[38] D. Batorski, A. Płoszaj, *Diagnoza i rekomendacje w obszarze kompetencji cyfrowych społeczeństwa i przeciwdziałania wykluczeniu cyfrowemu w kontekście zaprogramowania wsparcia w latach 2014–2020*, Warszawa 2012, s. 7, https://kometa.edu.pl/uploads/publication/379/10f1_A_KompetencjeCyfrowe_ost.pdf?v2.6 (dostęp: 12.01.2019).

- kognitariat – osoby potrafiące używać technologii informacyjno-komunikacyjnych do złożonych operacji poznawczych w stosunku do danych[39].

Opierając się na tej typologii klas w społeczeństwie informacyjnym, możliwe jest dokonanie analizy danych statystycznych Eurostatu (państwa członkowskie Unii Europejskiej) i Głównego Urzędu Statystycznego (polskie społeczeństwo) dotyczących poziomu umiejętności cyfrowych wybranych państw europejskich w ostatnich latach. W przypadku państw członkowskich UE charakterystyka czasowa tych zmian jest następująca (lata 2015–2017):

- w całej Unii Europejskiej osoby należące do digitariatu lub kognitariatu (obie kategorie są w przypadku tych danych zagregowane) stanowiły 57% mieszkańców UE w wieku 16–74 lata w 2017 roku (55% w 2015);
- wskaźnik przynależności do digitariatu lub kognitariatu na poziomie poniżej 50% dla danych z 2017 roku odnotowano w odniesieniu do Bułgarii, Rumunii, Chorwacji, Grecji, Polski, Irlandii, Łotwy (w tej grupie w latach 2015–2017 największy wzrost wystąpił w przypadku Polski, wyniósł 6 punktów proc.);
- do grupy państw członkowskich UE, w których 3/4 lub więcej osób w wieku 16–74 można zaliczyć do digitariatu lub kognitariatu, należą: Luksemburg, Holandia, Szwecja, Finlandia[40]. W przypadku Polski można zaklasyfikować (osoby w wieku 16–74 będące użytkownikami Internetu) do klasy proletariatu 28,5% osób, do klasy digitariatu – 25,2%, a do klasy kognitariatu – 21,1% badanych[41].

Należy zauważyć, że istotną przesłanką dla rozwoju społeczeństwa informacyjnego w Polsce jest modernizacja technologiczna administracji publicznej, która – od przystąpienia do Unii w 2004 roku – nadal nie jest zsynchronizowana z rozwojem gospodarki cyfrowej. Świadczą o tym takie negatywne zjawiska, jak: dostęp obywateli do zasobów informacyjnych państwa jest ograniczony; zgromadzone dane nie są ponownie wykorzystywane; brak jest interoperacyjności i kompatybilności publicz-

---

[39] A. Waligórska-Kotfas, *Etyczny wymiar usług infobrokerskich w gospodarce opartej na wiedzy*, „Konińskie Studia Społeczno-Ekonomiczne" 2016, nr 3, s. 231.

[40] Eurostat, *Individuals' Level of Digital Skills*, Last update: 15.03.2018, https://ec.europa.eu/eurostat/data/database (dostęp: 14.06.2018).

[41] Główny Urząd Statystyczny, *Społeczeństwo informacyjne w Polsce. Wyniki badań statystycznych z lat 2013–2017*, Warszawa – Szczecin 2017, s. 153, https://stat.gov.pl/files/gfx/portalinformacyjny/pl/defaultaktualnosci/5497/1/11/1/spoleczenstwo_informacyjne_w_polsce._wyniki_badan_statystycznych_z_lat_2013-2017.pdf (dostęp: 11.10.2018).

nych systemów informatycznych i rejestrów; zarządzanie zasobami informatycznymi jest zdecentralizowane; występuje redundancja informacji publicznych; brak jest racjonalności w wydatkowaniu środków na informatyzację[42].

Gospodarka cyfrowa: geneza, charakterystyczne cechy, rozwój

Rozwój gospodarki cyfrowej był procesem, który miał charakter równoległy w stosunku do rozwoju społeczeństwa informacyjnego i należy go wiązać z rozwojem sektora usług w państwach rozwiniętych (usługi edukacyjne, zdrowotne, działalność rozwojowo-badawcza[43]), który absorbował nadwyżki zasobów pracy pojawiające się w efekcie procesów demograficznych po II wojnie światowej. Jednym z głównym czynników technologicznych stymulujących zmiany ilościowe i jakościowe w sektorze usług było stopniowe wprowadzenie komputerów, które umożliwiały optymalizację procesów zarządzania i czynności wykonywanych przez pracowników (zgodnie z prawem Moore'a – liczba tranzystorów w układzie scalonym podwaja się w określonym przedziale czasu): w XX wieku w latach 50. – komputery lampowe, w pierwszej połowie lat 60. – komputery tranzystorowe, w drugiej połowie 60. – komputery oparte na układach scalonych, w latach 80. i 90. – komputery zawierające układy scalone o coraz wyższej skali integracji, obecnie – komputery wieloprocesorowe[44].

Castells dokonał syntezy wniosków różnych badaczy, które dotyczą podstawowych wymiarów przeobrażania gospodarki kapitalistycznej od fazy industrialnej do fazy informacyjnej, i wyodrębnił pięć wspólnych dla różnych analiz elementów: 1. przełom w organizacji produkcji i rynków nastąpił w latach 70. XX wieku; 2. zmiany organizacyjne w firmach wystąpiły wcześniej niż upowszechnienie się w nich technologii infor-

[42] Ministerstwo Cyfryzacji, *Program zintegrowanej informatyzacji państwa*, Warszawa 2018, s. 10, https://www.gov.pl/documents/31305/0/PZIP+wrzesie%C5%84+2016+r..pdf/9be4a5fe-2395-8904-5dbc-73e46deb284a (dostęp: 12.01.2019).

[43] D. Bell, *The Coming of the Post-Industrial Society*, „The Educational Forum" 1976, t. 40, nr 4, s. 577.

[44] Do głównych innowacji z zakresu informatyki i telekomunikacji, które wpływają na obecną strukturę gospodarki cyfrowej, należy zaliczyć: superkomputery, systemy operacyjne dla komputerów osobistych, wydajne serwery sieciowe, sieci neuronowe, inżynierię wiedzy, sieci szerokopasmowe, komunikację satelitarną, sieci światłowodowe, telefonię komórkową, zob. T. Goban-Klas, P. Sienkiewicz, *Społeczeństwo informacyjne: Szanse, zagrożenia, wyzwania*, Kraków 1999, s. 172–173, http://informacjacyfrowa.wsb.edu.pl/pdfs/SpoleczenstwoInformacyjne.pdf (dostęp: 5.06.2018).

macyjnych; 3. zmiany organizacyjne w firmach były reakcją na zmiany w ich całościowym otoczeniu zewnętrznym; 4. doszło do zmiany logiki funkcjonowania procesów pracy między innymi poprzez zastosowanie automatyzacji pracy; 5. zarządzanie wiedzą i przetwarzanie informacji stanowią kluczowe przesłanki skuteczności firm w globalnej gospodarce informacyjnej[45]. Wnioski M. Castellsa sugerują odwrócenie porządku relacji następstwa czasowego: najpierw doszło do zmian organizacyjnych w firmach, a dopiero potem zaadaptowały się one do wymogów związanych z korzystaniem z ICTs.

Gospodarka cyfrowa (ang. *digital economy*) jest terminem, który nie ma ogólnie akceptowalnej definicji, jej treść konstytutywna obejmuje czynności i podmioty związane z tworzeniem, akumulacją i przekształceniem informacji w trakcie procesów ekonomicznych[46]. Definicję realną tego terminu, która mieści się w sposobie ujęcia procesów informacyjnych przyjętych w tym artykule, zawiera *World Investment Report 2017. Investment and Digital Economy*, brzmi ona następująco: „zastosowanie internetowych technologii cyfrowych w produkcji i obrocie handlowym dobrami i usługami"[47]. Atrybutem takiego sposobu realizacji procesów ekonomicznych jest generowanie zysków w ramach pracy kognitywnej i produkcji wiedzy poprzez zastosowanie informacji jako kluczowego czynnika produkcji, zwłaszcza przez podmioty gospodarcze działające w sektorze informatycznym i przemyśle wysokiej technologii[48]. Oznacza to następujące postrzeganie udziału gospodarki cyfrowej (którą należy traktować jako jeden z segmentów całego systemu gospodarczego) – im wyższy udział sektora technologii informacyjnych i komunikacyjnych[49] w wytwarzaniu produktu krajowego brutto (PKB) i zatrudnieniu, tym wyższy poziom rozwoju tego modelu gospodarczego. Obecna faza rozwoju

---

[45] M. Castells, *Społeczeństwo sieci*, przekł. M. Marody i in., Warszawa 2008, s. 161.

[46] International Monetary Fund, *Measuring the Digital Economy*, 3.04.2018, s. 7, https://www.imf.org/~/media/Files/Publications/PP/2018/022818MeasuringDigitalEconomy.ashx (dostęp: 9.10.2018).

[47] United Nations Confrence on Trade and Development, *World Investment Report 2017. Investment and the Digital Economy*, Genewa 2017, s. 156, https://unctad.org/en/PublicationsLibrary/wir2017_en.pdf (dostęp: 9.10.2018).

[48] Zob. M. Ratajczak, *Wprowadzenie do teorii kapitalizmu kognitywnego: kapitalizm kognitywny jako reżim akumulacji*, „Praktyka Teoretyczna" 2015, nr 1(15), http://numery.praktykateoretyczna.pl/PT_nr15_2015_Praca_i_wartosc/02.Ratajczak.pdf (dostęp: 12.12.2018).

[49] Do sektora ICT zaliczane są przedsiębiorstwa: 1. produkujące dobra, które umożliwiają elektroniczne przetwarzanie informacji i komunikację; 2. świadczące usługi pozwalające na elektroniczne przetwarzanie informacji i komunikację; zob. Główny Urząd Statystyczny, *Społeczeństwo informacyjne…*, s. 19.

gospodarki cyfrowej przebiega w coraz bardziej złożonym ekosystemie informacyjnym, którego rozwój wynika z procesów wdrażania cyfrowych modeli biznesowych w tradycyjnych przedsiębiorstwach. Czynnikami stymulującymi ten wzrost są: Internet rzeczy (ang. *Internet of Things*, skrót IoT) – wymiana informacji pomiędzy urządzeniami bez pośrednictwa człowieka (na przykład zarządzanie zasobami), big data – analiza dużych zbiorów danych w celu wykrycia prawidłowości statystycznych o dużej zmienności (jak detekcja nowych wzorów zachowań konsumentów), sztuczna inteligencja – automatyzacja prostych czynności poznawczych (na przykład w logistyce)[50].

Gospodarka cyfrowa w przeciwieństwie do gospodarek poprzednich faz rozwoju cywilizacyjnego, które zapewniały miejsca pracy dla nowych pokoleń będących beneficjentami postępu medycyny i które coraz częściej docierają do okresu biologicznej starości, już nie tworzy automatycznie organizacyjnych rozwiązań zapewniających osiąganie stanu równowagi pomiędzy podażą a popytem na rynku pracy. W syntetyczny sposób ujmuje to futurolog Jeremy Rifkin: „W przeszłości, gdy w danym sektorze gospodarki nowa technologia wypierała zatrudnionych, zawsze pojawiały się nowe sektory, które wchłaniały zwolnionych robotników. Dziś wszystkie trzy tradycyjne sektory gospodarki – rolnictwo, przemysł i usługi – doświadczają zmian technologicznych i skazują miliony ludzi na bezrobocie. Jedynym nowo powstałym sektorem jest sektor naukowo-techniczny, złożony z nielicznej elity przedsiębiorców, naukowców, techników, programistów, specjalistów, instruktorów i konsultantów”[51]. Zmiana technologiczna w przedstawionym powyżej ujęciu nie jest postrzegana jako szansa rozwojowa, ale zagrożenie egzystencjalne. Ten dylemat jest przedmiotem analiz od momentu, gdy koncepcja społeczeństwa informacyjnego wkroczyła na trwałe do dyskursu naukowego.

Organizacyjna adaptacja podmiotów gospodarki do fazy cyfrowego rozwoju nie byłaby możliwa bez zmian na poziomie mikro, czyli tych, które dotyczą jednostek. Chodzi tutaj zarówno o kompetencje komunikacyjne wynikające z funkcjonowania jednostki jako odbiorcy/twórcy przekazów medialnych, jak i o umiejętności zawodowe wynikające z wykorzystania ICTs w miejscu pracy. Zdaniem ekspertów Banku Światowego we współczesnej gospodarce jednostka musi posiadać trzy typy umiejętności: 1. kognitywne (piśmienność, umiejętność liczenia, kreatywne

[50] International Telecommunication Union, *Measuring the Information…*, s. 95.

[51] J. Rifkin, *Koniec pracy. Schyłek siły roboczej na świecie i początek ery postrynkowej*, przekł. E. Kania, Wrocław 2001, s. 13.

myślenie, rozwiązywanie problemów oparte na wiedzy, werbalna biegłość, dobra pamięć, szybkość działania); 2. społeczne i behawioralne (dotyczące emocji, otwartość na doświadczenia, sumienność, ekstrawersja, umiejętność samoregulacji zachowania, podejmowanie decyzji); 3. techniczne (fizyczna zręczność, wykorzystywanie dostępnych metod i materiałów, dotyczące wymogów konkretnych zawodów)[52]. Warto zauważyć, iż wszystkie trzy typy umiejętności jednostka nabywa podczas długiego procesu socjalizacji i nauki w ramach wielostopniowego systemu edukacyjnego, co powoduje, że jej wejście na rynek pracy odbywa się często w późniejszym okresie jej życia niż miało to miejsce jeszcze około 100 lat temu.

Gospodarka cyfrowa bazuje na tych trzech typach, ale stanowią one jedynie podstawowe wersje umiejętności, które są niezbędne dla pracownika w gospodarce cyfrowej. Komisja Europejska wskazuje pięć kluczowych obszarów kompetencji cyfrowych: umiejętność odczytania i przetwarzania informacji; komunikacja i współpraca przy wykorzystaniu technologii cyfrowych; tworzenie cyfrowych treści; wielowymiarowa ochrona w zakresie bezpieczeństwa jednostki i jej otoczenia; rozwiązywanie problemów i wyzwań pojawiających się w związku z używaniem cyfrowych technologii[53]. Przegląd treściowego zakresu tych obszarów wskazuje, że jednostka obecna na rynku pracy w gospodarce cyfrowej powinna być wewnątrzsterowna i autonomiczna jako decydent w miejscu pracy. Produktywna aktywność jednostki na rynku pracy jest w pełni uzależniona od technologicznej sprawności jej środowiska pracy – w sytuacji, gdy ta sprawność zostaje przerwana, może dojść do czasowego przerwania procesu wytwarzania produktów lub usług w danym zakładzie pracy lub, w skrajnym przypadku, do zatrzymania tego procesu nawet w całej branży (w sytuacji cyberataku w makroskali).

Według obliczeń różnych analityków i ekspertów potencjał finansowo-rozwojowy wybranych zaawansowanych technologii informacyjno-komunikacyjnych w latach 2015–2025 zwiększy się znacząco: 1. Internet rzeczy – od 193,5 mld $ w 2015 roku do 640 mld $ w 2025; 2. big data – od 27,3 mld $ w 2015 roku do 88,5 mld $ w 2025; 3. chmura obliczeniowa – od 75,3 mld $ w 2015 roku do 489 mld $ w 2025; 4. sztuczna inteligencja – od 644 mln $ w 2015 roku do 36,8 mld $ w 2025[54]. Najwięk-

[52] World Bank, *World Development…*, s. 33.

[53] European Commission, *The European Digital Competence Framework for Citizens*, Luxembourg 2016, s. 4, http://ec.europa.eu/social/BlobServlet?docId=15688&langId=en (dostęp: 14.06.2018).

[54] International Telecommunication Union, *Measuring the Information…*, s. 107.

szy potencjał wzrostu w ujęciu wartościowym ma zatem IoT, a w ujęciu procentowym sztuczna inteligencja. W przypadku potencjału finansowo-rozwojowego polskiej gospodarki w aspekcie jej cyfryzacji należy zgodzić się z pesymistyczną tezą wyrażaną przez niektórych analityków, że brak jest przesłanek do stawiania wniosku o tym, że mamy już do czynienia z jej cyfrową transformacją[55]. Aczkolwiek należy zauważyć, iż decydenci publiczni wreszcie dostrzegli ten problem, gdyż 30 czerwca 2016 roku ówczesny wicepremier Mateusz Morawiecki podpisał zarządzenie w sprawie powołania Zespołu do spraw Transformacji Przemysłowej.

Obecny etap rozwoju gospodarki cyfrowej ze względu na wzrost tempa generowania informacji w obrębie ekosystemu informacyjnego jest nazywany rewolucją cyfrową, a do jego kluczowych trendów zalicza się: konwergencję sieci, autonomizację urządzeń elektronicznych, cyborgizację, spadające znaczenie pośredników, tworzenie różnych platform wymiany informacji, wzrost znaczenia globalnej konkurencji w obrocie gospodarczym, rosnące znaczenie danych i wiedzy w procesie zarządzania, deficyt uwagi odbiorców informacji, wzrost liczebności grupy prosumentów, *crowdsourcing*, automatyzację procesów produkcyjnych[56]. Warto zaznaczyć, że rewolucja cyfrowa przebiega odmiennie na poziomie kontynentów, ale również w odmienny sposób wpływa na życie różnych grup demograficznych, zawodowych i społecznych na poziomie państw. W celu wzmocnienia pozytywnych efektów i minimalizacji negatywnych efektów tego procesu pojawiają się inicjatywy zmierzające do ukierunkowania tych trendów poprzez ich programowanie – na przykład w UE dąży się do stworzenia jednolitego rynku cyfrowego w ramach projektu przewodniego „Europejska agenda cyfrowa" będącego częścią strategii „Europa 2020" przyjętej przez Radę Europejską w 2010 roku (ang. Single Digital Market)[57].

---

[55] B. Michałowski i in., *Internet of Things (IoT) i Artificial Intelligence (AI) w Polsce. Jak wykorzystać rewolucję technologiczną Internetu Rzeczy i Sztucznej Inteligencji w rozwoju Polski*, Warszawa 2018, s. 20, http://www.sobieski.org.pl/wp-content/uploads/Raport-IoT-i-AI-Instytut-Sobieskiego.pdf (dostęp: 12.06.2018).

[56] D. Batorski (red.), *Cyfrowa gospodarka. Kluczowe trendy rewolucji cyfrowej. Diagnoza, prognozy, strategie reakcji*, Warszawa 2012, s. 13–61, http://www.euroreg.uw.edu.pl/dane/web_euroreg_publications_files/1335/cyfrowa_gospodarka_kluczowe_trendy_rewolucji_cyfrowej.pdf (dostęp: 14.08.2018).

[57] Komisja Europejska, *Europejska agenda cyfrowa*. KOM(2010)245 wersja ostateczna/2, Bruksela, 26.8.2010, https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52010DC0245R(01) (dostęp: 15.01.2019); European Commission, *Digitising European Industry: Reaping the full benefits of a Digital Single Market*, Bruksela,

Dane statystyczne dotyczące tempa wzrostu, udziału w PKB oraz rynku pracy sektora technologii informacyjnych i komunikacyjnych na poziomie całej UE, jak również jej państw członkowskich wskazują, że wzrost gospodarki cyfrowej nie ma homogenicznego charakteru w wymiarze przestrzennym. Z danych Eurostatu wynika, iż: 1. w latach 2012–2017 wystąpił trend wzrostowy zatrudnienia w sektorze ICT w porównaniu w zatrudnieniem w ramach całego rynku pracy UE; 2. w latach 2007–2017 nie zmieniły się proporcje w zakresie zatrudnienia kobiet i mężczyzn w tym sektorze (ponad 3/4 pracowników to mężczyźni); 3. w 2017 roku w tym sektorze w całej UE pracowało 8 mln 385 tys. osób (wzrost o ponad 2 mln 103 tys. w porównaniu z 2008); 4. w 2017 roku w porównaniu z 2008 największy przyrost liczby pracowników (50% lub więcej) wystąpił w Irlandii, Belgii, Bułgarii, Francji, Niemczech, Estonii, Łotwie, Portugalii; 5. udział pracowników sektora ICT stanowił 3,7% zatrudnienia w całej UE (2,8% w 2008); 6. wśród osób zatrudnionych w tym sektorze w 2017 roku 57% było zlokalizowanych w czterech państwach członkowskich UE – Wielkiej Brytanii, Niemczech, Francji, Włoszech; 7. Polska w 2017 roku była pod względem liczby zatrudnionych w sektorze ICT na 6. miejscu w Unii, a odsetek zatrudnionych – w porównaniu z rokiem 2008 – nie zmienił się; 8. w 2017 roku największy odsetek zatrudnionych (między 6,8 a 5,0) w sektorze ICT odnotowano w Finlandii, Szwecji, Estonii, Wielkiej Brytanii, Luksemburgu; 9. najmniejszy odsetek zatrudnionych w tym sektorze w 2017 roku (poniżej 3) wystąpił w Hiszpanii, Polsce, Słowacji, na Litwie, we Włoszech, w Bułgarii, na Cyprze, w Łotwie, Portugalii, Rumunii, Grecji[58].

Infobrokering jako rezultat rozwoju społeczeństwa informacyjnego i gospodarki cyfrowej

Społeczeństwo informacyjne i gospodarka cyfrowa są dwoma podsystemami ekosystemu informacyjnego[59]. Ze strukturalnego punktu widzenia

19.4.2016, http://ec.europa.eu/transparency/regdoc/rep/1/2016/EN/1-2016-180-EN-F1-1.PDF (dostęp: 14.06.2018).

[58] Eurostat, *ICT Specialists in Employment*, czerwiec 2018, https://ec.europa.eu/eurostat/statistics-explained/index.php/ICT_specialists_in_employment#Number_of_ICT_specialists (dostęp: 15.01.2019).

[59] Parametry jego funkcjonowania są przedmiotem oddziaływania państwa za pośrednictwem sterowania informacyjnego obejmującego: 1. emocje, 2. zachowania konformistyczne, 3. zachowania rutynowe, 4. zachowania racjonalne. Zob. K. Mikołajewski,

można uznać, że ważniejszy jest podsystem „społeczny", jego potencjał demograficzny bowiem determinuje poziom popytu na informacje, które wytwarza podsystem „gospodarczy": im więcej osób znajduje się w klasie kognitariatu, tym większe możliwości tworzenia modeli biznesowych opartych na generowaniu i wykorzystywaniu danych. Podsystem „społeczny" spełnia także funkcję kontrolną w zakresie sposobów zdobywania i wykorzystania informacji przez podsystem „gospodarczy" – poprzez normy prawne[60] oraz wzorce zachowań obecne w kulturze politycznej. Natomiast podsystem „gospodarczy" korzysta z zasobów pracy społeczeństwa informacyjnego w celu zatrudnienia takich pracowników, których informacyjno-komunikacyjne kompetencje zapewniają podmiotom gospodarczym uzyskanie przewagi rynkowej (na przykład wykorzystując zjawisko asymetrii informacyjnej lub uzależnienia od mediów społecznościowych) poprzez generowanie informacji – to funkcja kreacyjna tego podsystemu. Demokratyczny system polityczny wyznacza sposoby przepływu informacji w ramach całego ekosystemu informacyjnego poprzez prawną delimitację sfery publicznej i sfery prywatnej[61].

Pośrednikami w procesie obiegu informacji w ekosystemie informacyjnym są osoby określane jako profesjonaliści informacji. W tej grupie znajdują się przedstawiciele takich zawodów, jak: analityk danych, badacz rynku, infobroker (broker informacji), specjalista białego wywiadu (ang. *open source intelligence*, OSINT), osoby zajmujące się wywiadem gospodarczym[62]. Specjalizują się oni w dostarczaniu usług i produktów

---

Pragmatyczne i moralne granice sterowania politycznego. Ujęcie systemowo-cybernetyczne, Warszawa 2010, s. 148–162.

60 Proces tworzenia prawa jako element realizacji tej funkcji kontrolnej może stać się istotnym źródłem szumu informacyjnego – np. w polskim porządku prawnym w 2017 r. pojawiło się 27 118 stron maszynopisu nowych aktów prawnych, a ich całościowa lektura zajęłaby 3 godziny i 37 minut codziennie! Zob. Grant Thornton, *Produkcja prawa w Polsce spowolniła, ale nadal przytłacza. Barometr stabilności otoczenia prawnego w polskiej gospodarce*, Barometr prawa. Edycja 2018, s. 5, http://barometrprawa.pl/wp-content/uploads/2018/02/barometr_prawa_2018.pdf (dostęp: 12.01.2019).

61 Przykład takich działań władczych to Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2016/679 z 27 kwietnia 2016 r. w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych i w sprawie swobodnego przepływu takich danych oraz uchylenia dyrektywy 95/46/WE (ogólne rozporządzenie o ochronie danych), https://eur-lex.europa.eu/legal-content/PL/TXT/?uri=celex%3A32016R0679 (dostęp: 10.01.2019).

62 S. Cisek, *Infobrokering i wywiad rynkowy. Podstawy*, 16.02.2017, s. 53, https://www.researchgate.net/publication/313820834_Infobrokering_i_wywiad_rynkowy_podstawy_16_17 (dostęp: 23.06.2018).

opartych na informacji rozumianej jako zasób mający określoną użyteczność dla organizacji zamawiającej ich usługi lub produkty. Jaka jest zatem pozycja infobrokera w tej grupie? Aby odpowiedzieć na to pytanie, należy zidentyfikować i wskazać dwa kryteria, które umożliwią określenie pozycji danego zawodu w stosunku do użytkownika finalnego informacji (organizacja) i charakteru wykorzystywanych przez niego źródeł informacji. Kryterium pierwsze – umiejscowienie w stosunku do użytkownika finalnego (organizacji) – zawiera dwa parametry: wewnątrz organizacji (jest jej pracownikiem) oraz w otoczeniu organizacji (działa na rzecz różnych podmiotów). Kryterium drugie – zakres dostępu do informacji – również odnosi się do dwóch parametrów, są to informacja jawna oraz informacja ukryta. Typologia, która powstaje poprzez skrzyżowanie tych dwóch kryteriów, zawiera cztery typy aktywności zawodowej:

- pozyskiwanie jawnych informacji wewnątrz organizacji (infobroker systemowy),
- pozyskiwanie informacji ukrytych wewnątrz organizacji (badacz rynku),
- pozyskiwanie informacji jawnych występujących w otoczeniu organizacji (infobroker rozumiany w tradycyjny sposób[63]),
- pozyskiwanie informacji ukrytych występujących w otoczeniu organizacji (wywiad gospodarczy).

Przyjęto, że specjalista od białego wywiadu (OSINT) działa na rzecz użytkownika finalnego znajdującego się poza „ekosystemem informacyjnym", czyli państwa[64]. Wszystkie zidentyfikowane typy aktywności zawodowej realizowane przez profesjonalistów informacji stanowią początkowy etap procesu decyzyjnego, którego inicjatorem jest użytkownik finalny informacji, gdyż proces informacyjny poprzedza proces decyzyjny[65].

Aby zidentyfikować charakterystykę działań składających się na infobrokering, należy najpierw wskazać, jakie cechy zasadnicze składają się na treść pojęcia „infobroker". M można przyjąć, że jest to „[…] podmiot, który na zlecenie i odpłatnie wyszukuje, ocenia, opracowuje, przetwarza

---

[63] Przegląd polskiej literatury na temat infobrokeringu w: A. Walczak-Niewiadomska, G. Czapnik, Z. Gruszka, *Brokerstwo informacyjne w Polsce – przegląd publikacji*, „Acta Universitatis Lodziensis. Folia Librorum" 2013, nr 17, http://cejsh.icm.edu.pl/cejsh/element/bwmeta1.element.hdl_11089_2334/c/13-walczak.pdf (dostęp: 7.07.2018).

[64] Na temat pojęcia „białego wywiadu" i jego charakterystyki zob. B. Saramak, *Wykorzystanie otwartych źródeł informacji w działalności wywiadowczej: historia, praktyka, perspektywy*, Warszawa 2015, s. 15–43, https://wnpism.uw.edu.pl/media/SOBN/Archiwum%20WDiNP/Wykorzystanie%20otwartych%20zrodel.pdf (dostęp: 2.01.2019).

[65] E. Knosala, *Zarys teorii decyzji w nauce administracji*, Warszawa 2011, s. 98.

i udostępnia informację"[66]. Jest to zatem zawód, który wraz z rozwojem społeczeństwa informacyjnego i gospodarki cyfrowej powinien zyskiwać na popularności w obrębie rynku pracy i wśród potencjalnych klientów. Działania realizowane przez infobrokera to: „[…] zawodowe, komercyjne pośrednictwo (mediacja) w świecie informacji"[67]. Może on oferować swoje usługi każdemu z podmiotów obecnych w ekosystemie informacyjnym. Barierą dla popytu na jego usługi jest zakres rozwoju społeczeństwa informacyjnego, świadomość decydentów biznesowych, uwarunkowania prawne, potencjał finansowy firm i osób fizycznych.

Genezy uwarunkowań powstania zawodu infobrokera należy upatrywać w rozwoju systemów informacyjnych, które rozwijały się po II wojnie światowej wraz odtajnieniem zasobów informacyjnych zgromadzonych do celów militarnych[68]. Infobrokering zaczął rozwijać się jako odrębny zawód w Stanach Zjednoczonych od lat 60. ubiegłego wieku, a w 1987 roku powstała tam organizacja skupiająca osoby wykonujące ten zawód – The Association of Independent Information Professionals (AIIP), w Polsce zaś infobrokerzy zaczęli oferować swoje usługi pod koniec lat 90.[69] Na koniec 2015 roku usługi infobrokerskie lub usługi zawierające działania informacyjne oferowało w Polsce ponad 700 firm lub osób[70]. Rozwój infobrokeringu jest związany z rozwojem Internetu, który zwiększył dostępne zasoby informacyjne do poziomu wymagającego kompetencji komunikacyjnych, którymi nie dysponowali reprezentanci klasy informacyjnego proletariatu lub digitariatu – bibliotekarz przestał pełnić rolę pośrednika w procesie cyrkulacji informacji w życiu społecznym i gospodarczym[71]. Główna różnica między infobrokerem i bibliotekarzem polega na tym, iż ten pierwszy dokonuje akredytacji informacji, czyli ocenia jakość pozyskanej informacji w sposób umożliwiający przypisanie mu autorstwa takiej

---

66 I. Bałos, S. Cisek, A. Januszko-Szakiel, *Wprowadzenie do infobrokeringu. Wybrane aspekty*, [w:] S. Cisek, A. Januszko-Szakiel (red.), *Zawód infobroker…*, s. 14.

67 Tamże, s. 13.

68 B. Sosińska-Kalata, *Obszary badań współczesnej informatologii (nauki o informacji)*, „Zagadnienia Informacji Naukowej – Studia Informacyjne" 2013, nr 51, s. 14, http://eprints.rclis.org/28291/1/ZIN_02-2013_s.9-41_BSosinska-Kalata.pdf (dostęp: 1.10.2018).

69 S. Cisek, *Infobrokerstwo w Polsce. Stan obecny i perspektywy*, IX Forum INT, Zakopane, 25–28.09.2007, s. 5, http://www.ptin.us.edu.pl/konferencje/9forum/repoz/cisek.pdf (dostęp: 2.01.2019).

70 M. Blim, *Bezpieczeństwo informacji a infobrokerstwo (część 1). Broker informacji – jego rodowód i użyteczność zawodowa*, „Zabezpieczenia", 13.12.2017, https://www.zabezpieczenia.com.pl/ochrona-informacji/bezpieczenstwo-informacji-infobrokerstwo-czesc-1-broker-informacji-jego-rodowod-i (dostęp: 2.01.2019).

71 W Polsce temu zjawisku towarzyszy także malejący poziom czytelnictwa.

oceny[72]. Zapewnienie prawidłowego procesu akredytacji informacji jest ściśle związane z przestrzeganiem przez infobrokera określonych zasad etycznych dotyczących pozyskiwania, przetwarzania i prezentowania użytkownikowi finalnemu produktu lub świadczenia usługi[73].

Tadeusz Wojewódzki, biorąc pod uwagę kryterium zakresu realizowanych zadań, wyodrębnia dwie profesje związane z realizacją zadań infobrokerskich: infobroker klasyczny i infobroker systemowy[74]. Infobroker klasyczny to zawód, którego przedstawiciele zajmują się technologią procesów informacyjnych, pełnią rolę pośrednika między zasobami informacji a jej potencjalnymi użytkownikami, skupiają się na potrzebach informacyjnych klienta, pracują samodzielnie lub w firmie infobrokerskiej. Infobroker systemowy realizuje zadania informacyjne będące reakcją na określone problemy informacyjne danej organizacji, posiada kompetencje z zakresu zarządzania wiedzą i/lub kompetencje odnoszące się do kwestii informatycznych, działa w sposób interdyscyplinarny, jest zdolny do aktywności opartej na systemowym podejściu do organizacji, skupia się na potrzebach i problemach organizacji, efekty jego pracy są dopasowane do sposobów analizy wiedzy typowych dla danej organizacji. Aktywność infobrokera systemowego jest w większym stopniu oparta na powtarzalnych i typowych wzorach interakcji w organizacji oraz nakierowana na osiągnięcie jej celów strategicznych[75]. Efektami uzyskania przez organizację nowych informacji jest wytworzenie przez nią nowej wiedzy – jawnej i ukrytej[76].

---

[72] K.R. Fiałkowski, *Broker informacji – definicja misji*, [w:] B. Sosińska-Kalata, E. Chuchro, W. Daszewski (red.), *Informacja w sieci. Problemy, metody, technologie*, Warszawa 2006, s. 34, http://bbc.uw.edu.pl/Content/3/INFORMACJA+W+SIECI.pdf (dostęp: 2.01.2019).

[73] Por. The Association of Independent Information Professionals, *Professsional Standards*, https://www.aiip.org/About/Professional-Standards (dostęp: 2.01.2019).

[74] T. Wojewódzki, *Infobrokerstwo systemowe – kontekst niezbędności infobrokerskiej roboty*, [w:] S. Cisek, A. Januszko-Szakiel (red.), *Zawód infobroker…*, s. 158–159.

[75] Jego funkcjonowanie w ramach powtarzalnych wzorców zachowań danej kultury organizacyjnej może powodować wewnątrz organizacji aktywizację negatywnych schematów postępowania z informacją. Zob. J. Boruszewski, *Identyfikacja problemów w infobrokerstwie systemowym*, „Studia Metodologiczne" 2014, t. 32, nr 32, https://repozytorium.amu.edu.pl/bitstream/10593/13781/1/Studia_metod_32_2014_Jaroslaw_Boruszewski.pdf (dostęp: 10.01.2019).

[76] Skuteczność implementacji nowych informacji warunkuje sposób realizacji procesów zarządzania wiedzą. Zob. G. Probst, S. Raub, K. Romhardt, *Zarządzanie wiedzą w organizacji*, przekł. K. Wacowska, Kraków 2002, s. 40–46.

Podstawową przesłanką skuteczności aktywności zawodowej infobrokera jest umiejętność korzystania ze źródeł informacji, które umożliwią optymalizację poziomu satysfakcji użytkownika finalnego z dostarczonego produktu lub usługi. Te źródła można podzielić na dwa podstawowe (według kryterium sposobu powiązań z innymi źródłami informacji): indywidualne oraz sieciowe. Pierwsza grupa źródeł obejmuje takie, które – ze względu na swoją charakterystykę fizyczną lub sposób zapisu danych – w procesie analizy nie mogą być powiązane w bezpośredni sposób z innymi źródłami informacji w celu uzyskania z nich informacji finalnej. Pojawia się ona dopiero dzięki scaleniu informacji cząstkowych. Natomiast druga grupa zawiera takie źródła, które w trakcie procesu analizy mogą być powiązane w bezpośredni sposób (zależności przyczynowe, strukturalne, funkcjonalne) z innymi źródłami informacji w celu przygotowania informacji finalnej. Warto zaznaczyć, iż sieciowy charakter źródeł informacji jest czynnikiem, który może obniżyć koszt wytworzenia informacji finalnej, ale jednocześnie może wydłużyć proces jej akredytacji (na przykład ze względu na konieczność weryfikacji informacji finalnej w źródłach o indywidualnym charakterze)[77]. Infobrokerzy mogą specjalizować się w pozyskiwaniu określonych rodzajów informacji – na przykład gospodarczych, medycznych, prawniczych, technicznych.

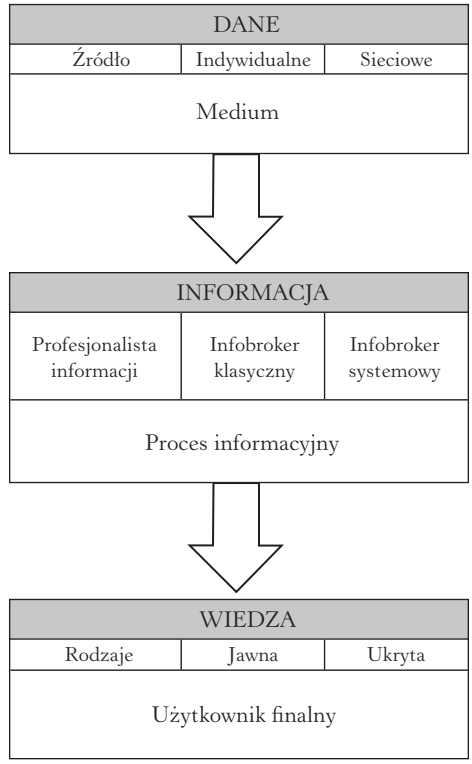
Zgodnie z *Krajowym standardem kompetencji zawodowych: broker informacji (researcher)* z 2013 roku[78] infobroker powinien posiadać następujące kompetencje zawodowe (wiedza, umiejętności, kompetencje społeczne): 1. pozyskiwanie, weryfikowanie oraz przechowywanie informacji; 2. analiza informacji przy użyciu odpowiednich metod analizowania i syntezy danych; 3. udostępnianie informacji; 4. kompetencje społeczne służące do komunikowania się z otoczeniem wewnętrznym i zewnętrznym. W profilu kluczowych kompetencji za najważniejsze zostały uznane następujące: rozwiązywanie problemów, ustna komunikacja, planowanie i organizowanie pracy, umiejętność czytania i pisania ze zrozumieniem, umiejętność obsługi komputera i wykorzystania Internetu. Należy zazna-

[77] Proces akredytacji może skrócić korzystanie z tzw. szarych źródeł informacji, czyli tych, które są skierowane do osób należących do klasy kognitariatu (np. naukowcy). Zob. A. Strojek, *Znaczenie terminu szara literatura*, „Zagadnienia Informacji Naukowej" 2000, nr 1, s. 64–76, http://bbc.uw.edu.pl/Content/1691/z2000_1_05.pdf (dostęp: 10.01.2019).

[78] Centrum Rozwoju Zasobów Ludzkich, *Krajowy standard kompetencji zawodowych: Broker informacji (researcher)*, Warszawa 2013, ftp://kwalifikacje.praca.gov.pl/standardy%20kompetencji%20zawodowych/77_262204_broker_informacji_researcher.pdf (dostęp: 2.01.2019).

czyć, że dostęp do Internetu jest warunkiem koniecznym prawidłowej realizacji zadań zawodowych przez infobrokera ze względu na zalety tego źródła: niski koszt dostępu do informacji, krótki czas jej pozyskania, wielkość zasobów informacyjnych[79].

Rysunek. Model infobrokeringu w sekwencyjnym ujęciu



Źródło: opracowanie własne.

Podsumowanie dotychczasowych rozważań zawiera przedstawiony graficznie model infobrokeringu w ujęciu sekwencyjnym. Składa się on z trzech faz, które są powiązane w aspekcie czasowym i podmiotowo-informacyjnym[80]. Każda z nich składa się z dwóch poziomów: informa-

[79] Potwierdzają to badania ankietowe zrealizowane w marcu 2010 r. wśród polskich firm infobrokerskich. Zob. K. Nizioł, *Infobrokering w Polsce – wyniki badań w środowisku praktyków zawodu*, „PTINT Praktyka i Teoria Informacji Naukowej i Technicznej" 2010, t. 18, nr 4, s. 12–13, http://www.ptin.us.edu.pl/pelne_teksty/2010_4.pdf (dostęp: 2.01.2018).

[80] Por. model procesu pozyskiwania informacji, w którym za najważniejszy element uznano centralną pozycję samego infobrokera; D. Christozov, S. Toleva-Stoimenova, *The Role of Information Brokers in Knowledge Management*, „Online Journal of Applied Knowledge

cyjnego oraz podmiotowego. W pierwszej fazie dane (nieustrukturyzowane informacje) są przechowywane/emitowane przez medium (Internet, prasa, radio, telewizja, osobowe źródła informacji, książki). Stanowią one określony zasób danych, który w drugiej fazie jest obiektem procesu informacyjnego realizowanego przez infobrokera (na zlecenie podmiotu zewnętrznego lub w organizacji): identyfikacja, zebranie, analiza, akredytacja, dostarczenie użytkownikowi finalnemu (jednostce lub organizacji). W ostatniej fazie ten użytkownik wykorzystuje w procesie decyzyjnym dostarczone informacje, które przyczyniają się do akumulacji posiadanej przez niego wiedzy jawnej lub ukrytej. Pomiędzy poszczególnymi fazami mogą pojawić się szumy informacyjne, wpływając niekorzystnie na procesy konwersji danych w informacje i konwersji informacji w wiedzę (zmniejszenie poziomu jakości).

Wnioski

Informacja jako jeden z warunków prawidłowego (podmiotowego) funkcjonowania jednostki w ramach społeczeństwa informacyjnego (obywatel[81]) i gospodarki cyfrowej (konsument) w odmienny sposób wpływa na zachowania jednostek i działania aktorów życia społecznego w każdym z tych dwóch podsystemów ekosystemu informacyjnego. W społeczeństwie informacyjnym dostęp do informacji publicznej gwarantuje jednostce uzyskanie instrumentów kontroli nad działaniami władzy publicznej – zakres i skuteczność tej kontroli bywają jednak ograniczone. Umiejętności cyfrowe są przedmiotem zainteresowania państwa w trakcie procesu edukacji, co stawia w lepszej sytuacji dzieci oraz młodzież i prowadzi do zjawiska cyfrowego wykluczenia seniorów. Wyższy poziom umiejętności cyfrowych uzyskanych przez dzieci („cyfrowych tubylców") wpływa pozytywnie na ich rozwój poznawczy[82]. Jednostki są narażone na

Management" 2014, nr 2, s. 113–115, http://www.iiakm.org/ojakm/articles/2014/volume2_2/OJAKM_Volume2_2pp109-119.pdf (dostęp: 2.01.2019).

81 Internet odegra kluczową rolę tym zakresie – w wymiarze tak pozytywnym (źródło rozwoju demokracji), jak i negatywnym (źródło atrofii demokracji lub totalitarnej kontroli nad społeczeństwem). Szerzej na ten temat: D. Mider, *Partycypacja polityczna w Internecie. Studium politologiczne*, Warszawa 2008, s. 320–371.

82 D. Di Giacomo, J. Ranieri, P. Lacasa, *Digital Learning As Enhanced Learning Processing? Cognitive Evidence for New insight of Smart Learning*, „Frontiers in Psychology" 2017, nr 8, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5540899/pdf/fpsyg-08-01329.pdf (dostęp: 21.01.2019).

malejący zakres prywatności wynikający z cyfryzacji wiedzy na ich temat, którą uzyskują organy władzy publicznej[83]. Pozycja jednostki w strukturze społecznej determinuje zasięg treściowy i podmiotowy jej dostępu do informacji w przestrzeni wirtualnej (Facebook i Twitter) – zjawisko „baniek informacyjnych". W gospodarce cyfrowej o przewadze rynkowej danej firmy decydują jej możliwości pozyskiwania i przetwarzania danych z otoczenia zewnętrznego (big data), a następnie – jej zdolność zarówno konwersji tych danych w informacje biznesowe[84], jak i akumulacji w postaci wiedzy organizacyjnej. Identyfikacja potencjalnych konsumentów/wyborców następuje dzięki agregacji informacji pochodzących z różnych źródeł (na przykład mediów społecznościowych), która bywa realizowana poza kontrolą społeczeństwa obywatelskiego i państwa[85]. Osobiste uczestnictwo w procesach społecznych i gospodarczych jest w coraz większym stopniu zastępowane przez uczestnictwo wirtualne – przykłady takiej fantomowej partycypacji to: podpisywanie petycji, zainteresowanie wydarzeniami publikowanym na Facebooku, retweety na Twitterze.

Uwarunkowania aktywności zawodowej infobrokera również ulegają ewolucji. Po pierwsze, należy wskazać na rosnącą dostępność różnego rodzaju informacji naukowych w formule otwartego dostępu (ang. *open access*, OA) – analiza dostępności danych zawartych w formule OA w bazie danych naukowych Scopus wskazuje na rosnący udział publikacji typu OA w całym zbiorze: w 2009 roku – 20,6%, a w 2016 – 28,3%[86]. Infobrokerzy uzyskują zatem coraz szerszy dostęp do przetworzonej, wysoko specjalistycznej informacji, która wymaga umiejętności prawidłowej syntezy przed etapem jej dostarczenia użytkownikowi finalnemu.

---

[83] Przykład Polski: Ministerstwo Cyfryzacji, Serwis OBYWATEL.GOV.PL, https://obywatel.gov.pl/o-serwisie (dostęp: 21.01.2019).

[84] Istnieją różne typy rynków informacji wykorzystywanej w działalności gospodarczej: 1. rynki informacji długotrwałej (licencje, patenty, oprogramowanie, usługi oświatowe), 2. rynki informacji krótkotrwałej (m.in. usługi doradcze, usługi wywiadowcze, 3. rynki informacji nabywanej w celach osobistych. Zob. S. Forlicz, *Informacja w biznesie*, Warszawa 2008, s. 57–62.

[85] Przykładem jest afera związana z firmą Cambridge Analytica ujawniona w 2018 r. przez brytyjski dziennik „The Guardian". Zob. C. Cadwalladr, E. Graham-Harrison, *Revealed: 50 million Facebook Profiles Harvested for Cambridge Analytica in Major Data Breach*, „The Guardian", 17.03.2018, https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election (dostęp: 21.01.2019).

[86] European Commission, *Trends for Open Access to Publications*, https://ec.europa.eu/info/research-and-innovation/strategy/goals-research-and-innovation-policy/open-science/open-science-monitor/trends-open-access-publications_en (dostęp: 21.01.2019).

Drugie interesujące zjawisko to utrzymująca się na wysokim poziomie w latach 2015–2018 gotowość użytkowników mediów społecznościowych (Facebooka i Twittera) do zapoznawania się z tzw. fałszywymi informacjami (ang. *fake news*), które opierają się na mechanizmie dezinformacji[87]. Gotowość do czytania i udostępnienia tych fałszywych informacji może utrudniać lub wydłużać proces akredytacji prawdziwych informacji (to jest zgodnych z faktami). Trzeci czynnik, na który warto zwrócić uwagę w kontekście czynników warunkujących skuteczność pracy infobrokera, to rosnąca dostępność i znaczenie narzędzi informatycznych służących do wizualizacji informacji – według oszacowań firmy badawczej Mordor Intelligence wartość rynku dostarczającego takie oprogramowanie wynosiła 4,51 mld dolarów w 2017 roku (a do roku 2023 osiągnie 7,76 mld dolarów)[88]. Podsumowując, aktywność informacyjna podmiotów społeczeństwa informacyjnego i gospodarki cyfrowej będzie tworzyła w najbliższych latach nowe możliwości działania dla infobrokerów oraz dostarczała im coraz bardziej złożonych metod i narzędzi do tworzenia produktów i usług informacyjnych.

Streszczenie

Artykuł przedstawia wybrane teoretyczne i empiryczne aspekty istnienia informacji w społecznym i gospodarczym wymiarze współczesnych demokratycznych systemów politycznych. Wskazano na poziom ilościowy i jakościowy analizy informacji. Ukazano nowe formy rozwojowe społeczeństwa informacyjnego. Przedstawiono kompetencje odnoszące się do korzystania z technologii informacyjno-komunikacyjnych. Omówiono sytuację na rynku pracy w gospodarce cyfrowej. Jako egzemplifikację zaprezentowano zawód infobrokera. Zaproponowano model infobrokeringu w sekwencyjnym ujęciu.

[87] H. Allcott, M. Gentzkow, Ch. Yu, *Trends in the Diffusion of Misinformation on Social Media*, 2018, https://web.stanford.edu/~gentzkow/research/fake-news-trends.pdf (dostęp: 21.01.2019).

[88] Mordor Intelligence, *Data Visualization Market*, 2018, https://www.mordorintelligence.com/industry-reports/data-visualization-applications-market-future-of-decision-making-industry (dostęp: 21.01.2019).

Przemysław Potocki

Informative Determinants of Social and Economical Development in the 21st Century: the Perspective of Information Broker

An article presents selected theoretical and empirical aspects of functioning information in social and economical dimensions of contemporary democratic political systems. Quantitative and qualitative levels of analysis were highlighted. New developmental forms of the information society were the object of consideration. Some skills related to the use of information and communication technologies were presented. The labour market's situation in the digital economy was include in the analysis. Profession of the information broker was used as an empirical example of that situation. Lastly, a sequential model of information broker's activity was proposed.

Key words: *information, information society, digital economy, labour market, information broker*

Bibliografia

Allcott H., Gentzkow M., Yu Ch., *Trends in the Diffusion of Misinformation on Social Media*, 2018, https://web.stanford.edu/~gentzkow/research/fake-news-trends.pdf (dostęp: 21.01.2019).

Batorski D., Płoszaj A., *Diagnoza i rekomendacje w obszarze kompetencji cyfrowych społeczeństwa i przeciwdziałania wykluczeniu cyfrowemu w kontekście zaprogramowania wsparcia w latach 2014–2020*, Warszawa 2012, https://kometa.edu.pl/uploads/publication/379/10f1_A_KompetencjeCyfrowe_ost.pdf?v2.6 (dostęp: 12.01.2019).

Blim M., *Bezpieczeństwo informacji a infobrokerstwo (część 1). Broker informacji – jego rodowód i użyteczność zawodowa*, „Zabezpieczenia", 13.12.2017, https://www.zabezpieczenia.com.pl/ochrona-informacji/bezpieczenstwo-informacji-infobrokerstwo-czesc-1-broker-informacji-jego-rodowod-i (dostęp: 2.01.2019).

Boruszewski J., *Identyfikacja problemów w infobrokerstwie systemowym*, „Studia Metodologiczne" 2014, t. 32, https://repozytorium.amu.edu.pl/bitstream/10593/13781/1/Studia_metod_32_2014_Jaroslaw_Boruszewski.pdf (dostęp: 10.01.2019).

Cadwalladr C., Graham-Harrison E., *Revealed: 50 million Facebook Profiles Harvested for Cambridge Analytica in Major Data Breach*, „The Guardian" 17.03.2018, https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election (dostęp: 21.01.2019).

Castells M., *Społeczeństwo sieci*, przekł. M. Marody i in., Warszawa 2008.

Castells M., *Koniec tysiąclecia*, przekł. J. Stawiński, S. Szymański, Warszawa 2009.

Christozov D., Toleva-Stoimenova S., *The Role of Information Brokers in Knowledge Management*, „Online Journal of Applied Knowledge Management" 2014, nr 2, http://www.iiakm.org/ojakm/articles/2014/volume2_2/OJAKM_Volume2_2pp109-119.pdf (dostęp: 2.01.2019).

Ciborowski L., *Walka informacyjna*, Toruń 1999.

Cisek S., *Infobrokerstwo w Polsce. Stan obecny i perspektywy*, IX Forum INT, Zakopane, 25–28.09.2007, http://www.ptin.us.edu.pl/konferencje/9forum/repoz/cisek.pdf (dostęp: 2.01.2019).

Cisek S., Januszko-Szakiel A. (red.), *Zawód infobroker. Polski rynek informacji*, Warszawa 2015.

Denning D.E., *Wojna informacyjna i bezpieczeństwo informacji*, przekł. J. Bloch, Warszawa 2002.

Di Giacomo D., Ranieri J., Lacasa P., *Digital Learning As Enhanced Learning Processing? Cognitive Evidence for New insight of Smart Learning*, „Frontiers in Psychology" 2017, nr 8, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5540899/pdf/fpsyg-08-01329.pdf (dostęp: 21.01.2019).

European Commission, *Digitising European Industry: Reaping the Full Benefits of a Digital Single Market*, Brussels, 19.4.2016, http://ec.europa.eu/transparency/regdoc/rep/1/2016/EN/1-2016-180-EN-F1-1.PDF (dostęp: 14.06.2018).

Fiałkowski K.R., *Broker informacji – definicja misji*, [w:] B. Sosińska-Kalata, E. Chuchro, W. Daszewski (red.), *Informacja w sieci. Problemy, metody, technologie*, Warszawa 2006, http://bbc.uw.edu.pl/Content/3/INFORMACJA+W+SIECI.pdf (dostęp: 2.01.2019).

Forlicz S., *Informacja w biznesie*, Warszawa 2008.

Karvalics L.Z., *Information Society – What is it Exactly? (The Meaning, History and Conceptual Framework of an Expression)*, Budapeszt 2007, http://www.msu.ac.zw/elearning/material/1349116439Information-Society-whatis.pdf (dostęp: 20.11.2017).

Michałowski B. i in., *Internet of Things (IoT) i Artificial Intelligence (AI) w Polsce. Jak wykorzystać rewolucję technologiczną Internetu Rzeczy i Sztucznej Inteligencji w rozwoju Polski*, Warszawa 2018, http://www.sobieski.org.pl/wp-content/uploads/Raport-IoT-i-AI-Instytut-Sobieskiego.pdf (dostęp: 12.06.2018).

Mider D., *Partycypacja polityczna w Internecie. Studium politologiczne*, Warszawa 2008.

Nizioł K., *Infobrokering w Polsce – wyniki badań w środowisku praktyków zawodu*, „PTINT Praktyka i Teoria Informacji Naukowej i Technicznej" 2010, t. 18, nr 4, http://www.ptin.us.edu.pl/pelne_teksty/2010_4.pdf (dostęp: 2.01.2018).

Saramak B., *Wykorzystanie otwartych źródeł informacji w działalności wywiadowczej: historia, praktyka, perspektywy*, Warszawa 2015, https://wnpism.uw.edu.pl/media/SOBN/Archiwum%20WDiNP/Wykorzystanie%20otwartych%20zrodel.pdf (dostęp: 2.01.2019).

Shannon C.E., *A Mathematical Theory of Communication*, „The Bell System Technical Journal" (Reprinted with corrections), t. 27, lipiec, sierpień, 1948, http://math.harvard.edu/~ctm/home/text/others/shannon/entropy/entropy.pdf (dostęp: 12.01.2019).

Sosińska-Kalata B., *Obszary badań współczesnej informatologii (nauki o informacji)*, „Zagadnienia Informacji Naukowej – Studia Informacyjne" 2013, nr 51, http://eprints.rclis.org/28291/1/ZIN_02-2013_s.9-41_BSosinska-Kalata.pdf (dostęp: 1.10.2018).

Walczak-Niewiadomska A., Czapnik G., Gruszka Z., *Brokerstwo informacyjne w Polsce – przegląd publikacji*, „Acta Universitatis Lodziensis. Folia Librorum" 2013, nr 17, http://cejsh.icm.edu.pl/cejsh/element/bwmeta1.element.hdl_11089_2334/c/13-walczak.pdf (dostęp: 7.07.2018).

Waligórska-Kotfas A., *Etyczny wymiar usług infobrokerskich w gospodarce opartej na wiedzy*, „Konińskie Studia Społeczno-Ekonomiczne" 2016, nr 3.

STUDIA POLITOLOGICZNE VOL. 54
DOI: 10.33896/SPolit.2019.54.2

Piotr Sosnowski
ORCID: 0000-0001-6985-4555

Systematyzacja pojęć związanych z metodami i źródłami pozyskiwania informacji w kontekście infobrokeringu

Słowa kluczowe:
infobrokering, pozyskiwanie informacji, źródła informacji, dyscypliny wywiadowcze, dyscypliny rozpoznania

STUDIA I ANALIZY

Wprowadzenie

Zdolność do sprawnego pozyskiwania, przetwarzania i zarządzania informacjami odgrywa dziś kluczową rolę w funkcjonowaniu organizacji. W procesie zarządzania dostęp do informacji pozwala na podejmowanie lepszych decyzji i tym samym na ograniczenie ryzyka związanego z działalnością oraz przeciwdziałanie zagrożeniom pochodzącym z otoczenia podmiotu. Globalna komunikacja i przetwarzanie danych w chmurze wraz z konwergencją sieci poszerzają zakres dostępnych informacji, czego konsekwencją jest dynamiczny rozwój metod ich pozyskiwania, analizy i oceny. Wzrost liczby dostępnych danych oraz szeroka oferta usług stwarzają konieczność uporządkowania terminologicznego i typologicznego chaosu w tym obszarze. Z tej perspektywy celem artykułu uczyniono przegląd i systematyzację pojęć związanych z metodami i źródłami pozyskiwania informacji w kontekście infobrokeringu.

Pierwsza część artykułu stanowi refleksję na temat dynamiki rozwoju technologicznego, która determinuje nowe metody i źródła pozyskiwania informacji. Druga dotyczy podziału ze względu na legalność źródeł pozyskiwanych informacji. Trzecia część jest poświęcona klasycznemu

podziałowi na dyscypliny według Paktu Północnoatlantyckiego[1]. Czwarta – dyscyplinom wspierającym zdefiniowanym w doktrynie amerykańskiej armii (ADP 2-0)[2]. Piąta część dotyczy nowych dyscyplin i subdyscyplin, które zdefiniowano w ostatniej dekadzie. Artykuł zamyka refleksja na temat wykorzystania „klasycznych" i nowych źródeł informacji w pracy brokera informacji.

Rozwój technologiczny jako źródło nowych metod i obszarów pozyskiwania informacji

Informacja jest uważana za najważniejszy czynnik kształtujący rozwój ludzkiej społeczności, a dynamiczne zmiany w dziedzinie jej przekazywania, jak na przykład wynalezienie pisma, druku, telegrafu czy Internetu, prowadziły do cywilizacyjnych przemian. W 1980 roku pierwszy dysk twardy, którego pamięć przekroczyła pojemność 1 GB, ważył 250 kg. Obecnie na rynku są dostępne smartfony dysponujące pamięcią 1 TB. Według danych opublikowanych przez Międzynarodowy Związek Telekomunikacyjny w 1998 roku z sieci Internet korzystało 3% światowej populacji, na koniec 2018 roku ponad połowa ludzkości była online (51,2%), a dziś już niemal cała żyje w zasięgu telefonii komórkowej[3]. W pierwszej połowie 2018 roku analitycy amerykańskiej spółki DOMO[4] oszacowali, że w każdej minucie użytkownicy Twittera przesyłali 473,3 tys. tweetów, a na Instagramie pojawiało się niemal 50 tys. nowych zdjęć. Ludzkość wytwarza ponad 2,5 kwintylion danych dziennie, a do 2020 roku na każdego mieszkańca planety będzie przypadać 1,7 MB wytwarzanych co sekundę[5]. Konsekwencje takiego stanu rzeczy są zbliżone do opisanej przez Alana Turinga koncepcji *infinite computing* i można je nazwać „rewolucją nieogra-

---

1 *AAP-06. NATO Glossary of Terms and Definitions*, NATO 2018, https://nso.nato.int/nso/zPublic/ap/PROM/AAP-06%202018%20EF.pdf (dostęp: 19.02.2019).

2 *Complementary Intelligence Capabilities*, Army Doctrine Publication ADP 2-0 for Army Intelligence Activities, Department of the Army, Waszyngton 2018, https://fas.org/irp/doddir/army/adp2_0.pdf (dostęp: 6.01.2019).

3 International Telecommunication Union, *Measuring the Information Society Report*, 2018, t. 1, s. 2–3, https://www.itu.int/en/ITU-D/Statistics/Pages/publications/misr2018.aspx (dostęp: 19.02.2019).

4 DOMO, Inc. – amerykańska spółka specjalizująca się w narzędziach do analizy biznesowej i wizualizacji danych.

5 *Data Never Sleeps* 6.0, infografika, https://www.domo.com/blog/wp-content/uploads/2018/06/18_domo_data-never-sleeps-6verticals.pdf (dostęp: 19.02.2019).

niczonego przetwarzania danych"[6]. Owa rewolucja jest wynikiem wzrostu mocy obliczeniowych oraz coraz szerszego dostępu do danych[7]. Yochai Benkler zauważył, że przetwarzanie i pozyskiwanie informacji przy użyciu zasobów sieci stało się tak relatywnie proste i tanie, że wyzwaniem nie jest już samo odnalezienie informacji, lecz ocena jej wartości i wiarygodności[8]. Podobną opinię wyraził dowódca Dowództwa Strategicznego USA (U.S. Strategic Command) gen. Robert Kehler, który w 2011 roku oświadczył, że w latach 2007–2011 liczba danych zebranych w obszarze rozpoznania geoprzestrzennego (*Geospatial Intelligence*, patrz: tabela 1) wzrosła o 1500%, a zdolność do ich analizy tylko o 30%[9].

W 2011 roku Henning Kagermann[10] wprowadził koncepcję Przemysłu 4.0[11] (*Industrie 4.0*, skrót – I4.0). Przemysł 4.0 opiera się na cyfryzacji *end-to-end*[12] ogółu aktywów fizycznych oraz integracji ekosystemów cyfrowych ze wszystkimi działaniami dotyczącymi produktu[13]. Jego istotą jest powiązanie technologii teleinformatycznych, przemysłu i Internetu rzeczy w celu obniżenia kosztów, poprawy wydajności oraz dostosowania produktu i usług do preferencji i zachowań konsumentów. Opiera się przede wszystkim na przetwarzaniu danych w czasie rzeczywistym z wykorzystaniem nowych technologii, między innymi sztucznej inteligencji, systemów cyberfizycznych, chmur obliczeniowych, robotyki, druku 3D, technologii addytywnych, rozszerzonej rzeczywistości i analityki biznesowej.

Procesy związane z pozyskiwaniem, przetwarzaniem i interpretacją danych od zarania dziejów towarzyszyły działalności handlowej, ale

---

[6] A.M. Turing, *On Computable Numbers, with an Application to the Entscheidungsproblem*, „Proceedings of the London Mathematical Society" 1937, nr 1, s. 230–265.

[7] P. Płoszajski, *Big Data: nowe źródło przewag i wzrostu firm*, „E-mentor" 2013, nr 3(50), s. 6.

[8] Y. Benkler, *Bogactwo sieci. Jak produkcja społeczna zmienia rynki i wolność*, Warszawa 2008.

[9] P.B. de Selding, *Pentagon Struggles with Avalanche of Data*, SpaceNews, 29.11.2011, https://spacenews.com/pentagon-struggles-avalanche-data (dostęp: 19.01.2019).

[10] Profesor fizyki i były prezes zarządu niemieckiego koncernu SAP (Systeme, Anwendungen und Produkte in der Datenverarbeitung).

[11] Inne kraje europejskie podjęły podobne inicjatywy, na przykład *Smart Industry* (Holandia), *Catapults* (Wielka Brytania), *Industrie du Futur* (Francja).

[12] W polskojęzycznej literaturze przedmiotu występuje tłumaczenie „od punktu wyjścia do punktu docelowego".

[13] Łańcuch wartości (ang. *value chain*) to w uproszczeniu sekwencja działań podejmowanych przez firmę, aby opracować, wytworzyć, sprzedać i dostarczyć produkt, a następnie świadczyć usługi posprzedażowe. M.E. Porter, *Competitive Advantage: Creating and Sustaining Superior Performance*, Nowy Jork 2008, s. 33.

dopiero współcześnie stały się odrębną i znaczącą gałęzią gospodarki. Przykładem może być ogromny udział w rynku sektora usług bezpłatnych dla konsumentów (jak na przykład media społecznościowe), którego zyski pochodzą w znacznej mierze z monetyzacji danych użytkowników[14]. Główny mechanizm ich funkcjonowania w uproszczeniu można porównać do cyklu wywiadowczego w ujęciu Roberta M. Clarka. Składa się z sześciu elementów: określenia potrzeb, planowania i kierowania, gromadzenia, przetwarzania, analizy i opracowania, rozpowszechniania, którego skutkiem w tym przypadku jest utowarowienie (monetyzacja) danych. Zatem na przykład podmiot z sektora usług marketingowych informuje podmiot administrujący platformą społecznościową o konkretnym zapotrzebowaniu na produkt informacyjny; platforma planuje i kieruje odpowiednie instrumenty, które gromadzą, a następnie przetwarzają dane użytkowników; odpowiednie systemy realizują zadania z zakresu analizy i opracowania zebranych danych. Wyniki trafiają jednocześnie na zewnątrz do klienta w postaci produktu i do wnętrza organizacji jako wnioski dotyczące funkcjonowania całego cyklu. Platforma nieustannie uczy się i rozwija możliwości, na przykład w dziedzinie utrzymania uwagi użytkowników. Im więcej interakcji użytkowników z platformą, tym większa jej wartość.

W sferze marketingu wykorzystanie wielkich zbiorów danych, w tym personalnych, do przygotowania treści reklamowej nazywa się profilowaniem. Można dopatrywać się analogii w dziedzinie wojskowości, w której dokładne uderzenie środków ogniowych na zamierzone cele potocznie nazywa się „precyzyjnym cięciem chirurgicznym". Oba zamierzenia polegają na wykryciu i wejściu w interakcję tylko z obiektem będącym w kręgu zainteresowań organizacji i wymagają zaawansowanych, precyzyjnie zorganizowanych metod pozyskiwania i analizy danych. Jest to tylko jeden z wielu elementów wskazujących na wspólną tożsamość i wzajemne przenikanie się metod i narzędzi procesu rozpoznania realizowanego przez wojsko i wywiadu biznesowego (*Business Intelligence*).

Podział na otwarte i zamknięte źródła informacji

W uproszczeniu źródła informacji dzieli się na jawne i niejawne. Te pierwsze odnoszą się do źródeł otwartych. Dostęp do nich jest

[14] Zob. szerzej: S. Elvy, *Paying for Privacy and the Personal Data Economy*, „Columbia Law Review" 2017, t. 117, nr 6, s. 1369–1459.

powszechny i nie łamie prawa. Pozyskiwanie informacji z tych źródeł w ujęciu procesowym nazywa się białym wywiadem. Drugie nazywane są czarnym wywiadem i dotyczą informacji zbieranych metodami wywiadu operacyjnego przez agenturę lub funkcjonariuszy mających specjalne uprawnienia. Wykorzystuje urządzenia techniczne (podsłuch, podgląd, kontrola korespondencji), dane z satelitów szpiegowskich, skryte pozyskiwanie utajonych informacji itp.[15] Pochodzą ze źródeł zamkniętych. Zdaniem Bartosza Saramaka tym, co odróżnia biały wywiad od czarnego wydaje się być brak konieczności naruszania czyjejś prywatności czy łamania prawa chroniącego poufne bądź tajne informacje. Immamentną cechą czarnego wywiadu jest konieczność utrzymania go w tajemnicy. W przypadku ujawnienia takiej działalności organizacja (rząd, podmiot prywatny) zaprzecza, jakoby ją realizował.

Powyższe rozróżnienie zarysowuje dwa przeciwległe bieguny osi, na której praktyka pozyskiwania informacji oprócz czerni i bieli może też przyjąć najróżniejsze odcienie szarości. Zarysowanie wyraźnej granicy między środkami kwalifikowanymi jako operacyjne a wszelkimi innymi, które można wykorzystać do zbierania informacji, wydaje się niemożliwe. Różne podmioty prywatne (na przykład wywiadownie gospodarcze, firmy detektywistyczne) nie korzystają jedynie z materiałów jawnych i ogólnodostępnych. Pracownicy tych podmiotów mimo braku odpowiednich uprawnień śledczych lub posiadania ich w węższym zakresie niż funkcjonariusze państwowi również stosują techniki operacyjne. Taką działalność podmiotów niepaństwowych nazywa się potocznie szarym wywiadem.

Istotą tej typologii jest rozróżnienie między trzema rodzajami działalności wywiadowczej: legalną i zgodną z normami etycznymi (biały wywiad), legalną, ale wątpliwą etycznie (szary wywiad) oraz nielegalną i nieetyczną (czarny wywiad). Niestety przypisywanie tych pojęć współczesnym niezwykle złożonym problemom staje się coraz trudniejsze. Być może w momencie, kiedy postrzeganie granicy między rzeczywistością wirtualną a fizyczną stało się błędem poznawczym, zaistniała potrzeba konceptualizacji dodatkowych kolorów w tej klasyfikacji. Coraz częściej opinia publiczna dowiaduje się o nieprawidłowościach[16] w funkcjonowa-

[15] T.A. Aleksandrowicz, *Biały wywiad w walce z terroryzmem*, [w:] K. Liedel, P. Piasecka (red.), *Rola mediów w przeciwdziałaniu terroryzmowi*, Warszawa 2009, s. 81; T.R. Aleksandrowicz, *Podstawy walki informacyjnej*, Warszawa 2016, s. 72.

[16] Nieprawidłowości w percepcji opinii publicznej. W percepcji podmiotów świadczących usługę mogą to być funkcjonalności, o których celowo nie informuje się użytkowników w sposób precyzyjny, gdyż przynoszą dochody, np. sprzedaż metadanych firmom marketingowym.

niu technologii, które pozwalają nieprzewidzianym podmiotom na dostęp do prywatnych danych użytkowników. Przykładem mogą być kontrowersje związane z ujawnieniem procederu wykorzystywania precyzyjnych danych lokalizacji smartfonów przez łowców nagród. Firmy telekomunikacyjne sprzedawały dane pochodzące z usługi Assisted-GPS podmiotom zewnętrznym, które z kolei sprzedawały je osobom prywatnym[17]. W tym przypadku jest jasne, że taki proceder nie powinien mieć miejsca i jest społecznie nieakceptowalny. Jednak nie jest do końca jasne, czy którykolwiek z podmiotów złamał prawo. Zakwalifikowanie działań łowców nagród jako szary wywiad lub czarny wywiad zależeć będzie od ewentualnej decyzji amerykańskiego sądu. Ten przykład obrazuje trudność w stosowaniu tej typologii podczas analizy powszechnych dziś zjawisk związanych z pozyskiwaniem informacji pochodzących z urządzeń elektronicznych przez podmioty trzecie, które w założeniu nie powinny mieć do nich dostępu, ale go uzyskują, nie łamiąc przy tym prawa.

Rozróżnienie na biały, szary i czarny wywiad rzadko pojawia się w zagranicznych opracowaniach teoretycznych w znaczeniu wyżej opisywanym[18]. Nie występuje ani w dokumentach doktrynalnych i standaryzacyjnych Sojuszu Północnoatlantyckiego, ani w oficjalnych i publikowanych przez państwa dokumentach i aktach prawnych dotyczących tej dziedziny. Najczęściej stosowana jest typologia ze względu na źródła pozyskanej informacji. Nazywana jest podziałem na dyscypliny wywiadowcze.

Klasyczne dyscypliny wywiadowcze

W amerykańskiej terminologii wojskowo-wywiadowczej ukształtowało się ujęcie procesowe oparte jednocześnie na metodzie i źródle pozyskiwanej informacji oraz na problemie, którego dotyczy. Dyscypliny są nazywane według schematu polegającego na połączeniu dwóch akronimów. Pierwszy pochodzi od określenia problemu, którego dotyczy zakres

[17] M. Giles, *Bounty Hunters Tracked People Secretly Using US Phone Giants Location Data*, „MIT Technology Review", 7.02.2019, https://www.technologyreview.com/the-download/612907/bounty-hunters-tracked-people-secretly-using-us-phone-giants-location-data (dostęp: 8.02.2019).

[18] *Grey Intelligence* pojawia się np. jako koncepcja opisująca zjawisko zatarcia granic między działalnością w sferze bezpieczeństwa realizowaną przez państwo a tą realizowaną przez podmioty prywatne. Odnosi się do procesów politycznych, a nie do działań wywiadowczych.

przedsięwzięć, na przykład otwarte źródła informacji – *open source*, a drugi od terminu *intelligence*, co w połączeniu daje OSINT.

Angielskojęzyczny termin *intelligence* jest wieloznaczny i może określać wiedzę, mądrość lub informację. W polskiej literaturze przedmiotu jest tłumaczony najczęściej jako wywiad, gdyż w tym kontekście odnosi się nie do zasobu wiedzy, lecz do procesu, który obejmuje szereg podmiotów włączonych we wspólny łańcuch wymiany informacji[19]. W NATO definiowaniem dyscyplin wywiadowczych zajmuje się Połączona Rada Standaryzacyjna Komitetu Wojskowego (PRSKW)[20]. Według autorów polskiego tłumaczenia słownika terminów NATO z 2014 roku *intelligence* oznacza „[…] produkt wynikający z przetworzenia informacji dotyczących innych państw, wrogich lub potencjalnie wrogich sił lub elementów albo obszarów rzeczywistych lub potencjalnych działań. Termin ten jest stosowany także do określania działania, którego wynikiem jest ten produkt oraz do struktur zaangażowanych w taką działalność"[21]. Zatem może być rozumiany zarówno jako zasób (dane wywiadowcze lub dane rozpoznawcze), jak i czynność (wywiad lub rozpoznanie). Wybór terminu zależy od poziomu prowadzonych działań lub stosowanych procedur. Definicja opublikowana w tym słowniku pochodziła z 1981 roku.

W słowniku z 2018 roku znajduje się definicja PRSKW z 2013 roku, wedle której *intelligence* jest „produktem powstałym wskutek prowadzonego gromadzenia i przetwarzania informacji dotyczących środowiska, możliwości i intencji aktorów w celu identyfikacji zagrożeń i przedstawienia decydentom możliwości do wykorzystania"[22]. W tym samym słowniku środowisko (*environment*) jest zdefiniowane jako otoczenie działania organizacji; obejmuje powietrze, wodę, ląd, zasoby naturalne, florę, faunę oraz ludzi wraz z ich interakcjami[23]. W tej definicji zastanawia brak wyszczególnienia na przykład cyberprzestrzeni jako obszaru, gdzie zachodzą interakcje między urządzeniami.

---

[19] M. Ciecierski, *Wywiad biznesowy w korporacjach transnarodowych. Teoria i praktyka*, Toruń 2009, s. 113.

[20] Military Committee Joint Standardization Board (MCJSB).

[21] *AAP-6. Słownik terminów i definicji NATO. Zawierający wojskowe terminy i ich definicje stosowane w NATO*, 2014, http://wcnjk.wp.mil.pl/plik/file/N_20130808_AAP6PL.pdf (dostęp: 7.02.2019).

[22] „The product resulting from the directed collection and processing of information regarding the environment and the capabilities and intentions of actors, in order to identify threats and offer opportunities for exploitation by decision-makers". Źródło: *AAP-06. NATO Glossary of Terms…*, s. 66.

[23] „The surroundings in which an organization operates, including air, water, land, natural resources, flora, fauna, humans, and their interrelations", tamże, s. 48.

Tabela 1. Dyscypliny wywiadowcze według Paktu Północnoatlantyckiego

Akronim	Nazwa dyscypliny w oryginale	Nazwa dyscypliny w języku polskim	Problem
ACOUSTINT /ACINT	*Acoustical Intelligence*	rozpoznanie akustyczne	Zjawiska akustyczne.
CI	*Counter Intelligence*	kontrwywiad	Zagrożenia związane ze szpiegostwem, sabotażem, działalnością wywrotową i terroryzmem.
COMINT	*Communication Intelligence*	rozpoznanie komunikacyjne	Komunikacja międzyludzka realizowana za pomocą sygnałów elektromagnetycznych
ELINT	*Electronic Intelligence*	rozpoznanie elektroniczne	Sygnały elektromagnetyczne niezwiązane z komunikacją międzyludzką.
GEOINT	*Geospatial Intelligence*	rozpoznanie geoprzestrzenne	Wizualizacja problemu na podstawie powiązania danych geoprzestrzennych z danymi obrazowymi i/lub z danymi pochodzącymi z innych dyscyplin.
HUMINT	*Human Intelligence*	rozpoznanie osobowe	Informacje zbierane i dostarczane przez źródła osobowe.
MASINT	*Measurement and Signature Intelligence*	rozpoznanie pomiarowo-badawcze	Identyfikacja źródła, emitera lub nadawcy na podstawie danych pozyskanych z przyrządów pomiarowych.
MEDINT	*Medical Intelligence*	rozpoznanie medyczne	Dane: z dziedziny nauk biologicznych, medyczne, epidemiologiczne, środowiskowe i inne związane ze zdrowiem ludzi lub zwierząt.
OSINT	*Open Source Intelligence*	rozpoznanie ze źródeł jawnych	Informacje dostępne publicznie lub o ograniczonym dostępie publicznym, ale jawne.
SIGINT	*Signal Intelligence*	rozpoznanie elektromagnetyczne	Sygnały radiowe i elektroniczne. Określenia używa się tylko wtedy, kiedy występują razem.
TECHINT	*Technical Intelligence*	rozpoznanie techniczne	Rozwój zagranicznych technologii oraz wydajność i możliwości zagranicznych materiałów, które mają lub mogą mieć zastosowanie wojskowe.

Źródło: opracowanie własne na podstawie: *AAP-06. NATO Glossary of Terms and Definitions*, NATO 2018, https://nso.nato.int/nso/zPublic/ap/PROM/AAP-06%202018%20EF.pdf [dostęp: 1.02.2019].

Rada standaryzacyjna NATO zdefiniowała 12 dyscyplin wywiadowczych (tabela 1). Według definicji przyjętej i przez NATO, i przez amerykańską armię OSINT „dotyczy danych pochodzących zarówno z publicznie dostępnych informacji, jak i innych jawnych informacji o ograniczonym rozpowszechnianiu lub dostępie"[24]. Jednak amerykańska doktryna wywiadowcza precyzuje, że samo pozyskiwanie i przetwarzanie informacji z otwartych źródeł jest na tyle powszechną czynnością, że nie należy kwalifikować jej jako OSINT, który jest realizowany wyłącznie przez odpowiednio przygotowany personel i ma na celu wsparcie innych dyscyplin rozpoznania, operacji wywiadu oraz decyzji podejmowanych przez dowódców[25]. Szczegółowej charakterystyce tej dyscypliny została poświęcona amerykańska „Army Techniques Publication" – ATP 2-22.9[26].

Definicje występujące w dokumentach doktrynalnych kładą nacisk na ścisły związek produktu informacyjnego ze zgłoszonym wcześniej zapotrzebowaniem oraz na jego weryfikację[27]. Generalnie nie przedstawiają sztywnej metodyki, lecz jedynie zarysowują ogólne ramy, wedle których poszczególne organizacje wypracowują własne standardy[28]. Departament Armii Stanów Zjednoczonych definiuje OSINT jako „pozyskiwanie informacji (publicznie dostępnych, o które każdy może legalnie poprosić, kupić je lub zdobyć w wyniku obserwacji) o dużym znaczeniu, systematycznie gromadzonych w bazach danych, przetwarzanych i analizowanych. Opracowany materiał stanowi zaś odpowiedź na zapytanie wywiadowcze o określonych wymaganiach. Informację jawnoźródłową stanowi każda informacja, którą każdy członek społeczeństwa może legalnie uzyskać na żądanie, za pomocą obserwacji, a także inne nieklasyfikowane jawne informacje o ograniczonym dostępie publicznym. Za informacje ogólnodostępne (publiczne) uznaje się powszechnie dostępne informacje, które spełniają następujące warunki: zostały opublikowane lub zezwolono na ich publiczne wykorzystanie; są dostępne na życzenie, dostępne online, poprzez subskrypcję lub zakup; każdy potencjalny obserwator może je zobaczyć lub usłyszeć; są udostępniane na publicznym spotkaniu lub

---

[24] *Complementary Intelligence Capabilities…*, roz. 4, s. 7–8, https://fas.org/irp/doddir/army/adp2_0.pdf (dostęp: 19.02.2019).

[25] Tamże.

[26] *Open Source Intelligence*, „Army Techniques Publication" No. 2-22.9 (FMI 2-22.9), Department of the Army, Waszyngton 2012, https://fas.org/irp/doddir/army/atp2-22-9.pdf (dostęp: 20.01.2019).

[27] B. Saramak, *Wykorzystanie otwartych źródeł informacji w działalności wywiadowczej: historia, praktyka, perspektywy*, Warszawa 2015, s. 21.

[28] Tamże, s. 20–21.

uzyskiwane przez odwiedzanie dowolnego miejsca albo uczestnictwo w każdym wydarzeniu otwartym dla publiczności. Publicznie dostępne informacje obejmują także informacje ogólnie dostępne dla społeczności wojskowej, chociaż społeczność wojskowa nie jest otwarta dla ludności cywilnej. Otwarte informacje mogą występować na przykład w formie interakcji społecznych, materiałów drukowanych, przekazu medialnego, Internetu, otwartego forum. Pozyskiwanie takich informacji powinno być nieinwazyjne (ang. *nonintrusive*)[29].

Niezależnie od terminologii wojskowej sfera cywilna interpretuje OSINT w sposób tożsamy z polskim białym wywiadem, czyli po prostu jako wykorzystanie otwartych (jawnych) źródeł informacji. W tym ujęciu informacje mogą mieć różne źródło i charakter, przez co cywilny OSINT przenika się z innymi dyscyplinami, które w nomenklaturze wojskowej mogłyby zostać uznane za odrębne.

Armia amerykańska definiuje HUMINT (ang. *Human Intelligence*, pol. rozpoznanie osobowe) jako pozyskiwanie informacji ze źródeł osobowych i multimediów przez wyszkolonego w tym zakresie specjalistę w celu identyfikacji: obcych podmiotów, ich intencji, części składowych, sił, zadań, taktyki, wyposażenia i możliwości. Źródłem osobowym (osobą dostarczającą informacji) mogą być na przykład więźniowie, jeńcy wojenni, uchodźcy, przesiedleńcy, lokalni mieszkańcy, siły przeciwnika, członkowie zagranicznych organizacji rządowych i pozarządowych. Proces pozyskiwania informacji może być zarówno jawny jak i niejawny, opiera się na metodach związanych z monitorowaniem, prowadzeniem przesłuchań, rozpytywaniem, przetwarzaniem informacji nieosobowych (na przykład multimediów) dotyczących źródeł osobowych oraz na współpracy łącznikowej z organizacjami państwowymi i pozarządowymi[30].

Dyscyplina MASINT (ang. *Measurement and Signature Intelligence*, pol. rozpoznanie pomiarowo-badawcze) dotyczy produktu informacyjnego powstałego w wyniku analizy naukowej i technicznej danych pochodzących z przyrządów pomiarowych. Dzięki rozpoznaniu charakterystycznych cech związanych ze źródłem, emiterem lub nadawcą możliwa jest jego identyfikacja[31]. Często dotyczy tych samych problemów co inne dyscypliny, ale różni się od nich stosowaniem metod naukowych. Na przykład SIGINT (ang. *Signal Intelligence*, pol. rozpoznanie radioelek-

[29] *Complementary Intelligence Capabilities*…, roz. 4, s. 4–8.

[30] *FM 2-22.3 (FM 34-52) Human Intelligence Collector Operations Headquarters*, Department of The Army, Waszyngton 2006, https://fas.org/irp/doddir/army/fm2-22-3.pdf (dostęp: 20.01.2019).

[31] *AAP-06. NATO Glossary of Terms*…

troniczne) koncentruje się na samym przechwytywaniu i przetwarzaniu danych, a MASINT na szczegółowej analizie opartej na metodach jakościowych i ilościowych. W dużym uproszczeniu można to zobrazować przykładem: analityk SIGINT na podstawie podobieństwa przechwyconych sygnałów do posiadanych wzorów określa typ urządzenia, które je emituje. Odpowiednio dobrany analityk MASINT może na przykład dojść do wniosku, że ma do czynienia z dezinformacją (na podstawie przyczyn nieznanych analitykowi SIGINT), a ponadto może być zdolny do umieszczenia danego zdarzenia w szerszym kontekście. W przypadku wywiadu naukowo-technicznego ocena danych jest realizowana przez ekspertów z danej dziedziny. Amerykańska doktryna wywiadu z 2018 roku podaje praktyczne przykłady zastosowania tej dyscypliny, na przykład wykrywanie i penetracja kamuflażu, wykrywanie zmian na powierzchni ziemi, identyfikacja swój–obcy, wykrywanie broni masowego rażenia. Matthew M. Aid wyróżnił sześć subdyscyplin składających się na MASINT[32]:

1. rozpoznanie radiooptyczne (*Radiooptical Intelligence*, EOINT), w ramach którego wyróżnia się dyscypliny podrzędne:
   a. rozpoznanie promieniowania podczerwieni (*Infrared Intelligence*, IRINT);
   b. rozpoznanie optyczne (*Optical Intelligence*, OPTINT) dotyczy fal elektromagnetycznych widzialnych – na przykład ultrafiolet, bliska podczerwień;
   c. rozpoznanie promieniowania laserowego (*Laser Intelligence*, LASINT) dotyczy systemów komunikacji laserowej, naprowadzania laserowego itp.;
2. rozpoznanie radiolokacyjne (*Radar Intelligence*, RADINT);
3. rozpoznanie fal radiowych (*Radio-frequency Intelligence*, RF), w ramach którego wyróżnia się dyscypliny podrzędne:
   a. rozpoznanie fal radiowych (RF) i/lub impulsów elektromagnetycznych (*Electro-magnetic Pulse*, EMP);
   b. rozpoznanie niezamierzonego promieniowania (*Unintentional Radiation Intelligence*, RINT);
4. rozpoznanie geofizyczne (*Geophysical Intelligence*, GEOINT)[33];
5. rozpoznanie materiałowe (*Materials Intelligence*);
6. rozpoznanie jądrowe (*Nuclear Intelligence*, NUCINT).

---

[32] M.M. Aid, *Measurement and Signature Intelligence*, [w:] R. Dover, M.S. Goodman, C. Hillebrand, *Routledge Companion to Intelligence Studies*, Londyn 2014, s. 121.

[33] Dokumenty NATO AAP-06 (2018) i ADP 2-0 (2019) używają tego samego akronimu do opisu rozpoznania geoprzestrzennego (*Geospatial Intelligence*).

Amerykańska doktryna wywiadu oprócz powyższych obszarów MASINT wyróżnia też metody pozyskiwania danych dotyczących broni chemicznej, biologicznej i radiologicznej. Specyfika tych subdyscyplin wynika z umieszczenia ich w kontekście wywiadu pomiarowo-badawczego. W zależności od uwarunkowań doktrynalnych, organizacyjnych czy specyfiki problemu powyższe subdyscypliny MASINT mogą występować jako samodzielne dyscypliny.

Słownik terminów i definicji NATO definiuje również rozpoznanie akustyczne (ang. *Acoustic Intelligence*, ACINT, dawniej ACOUSINT[34]). W literaturze i w brytyjskiej doktrynie jest ono kwalifikowane jako subdyscyplina MASINT[35]. Dotyczy pozyskiwania i przetwarzania informacji generowanych przez zjawiska akustyczne[36]. Brytyjska doktryna definiuje również rozpoznanie obrazowe (*Imagery Intelligence*, IMINT) i kwalifikuje je jako subdyscyplinę rozpoznania geoprzestrzennego. Niektórzy autorzy przyporządkowują IMINT jako subdyscyplinę TECHINT (*Technical Intelligence*)[37] lub jako samodzielną dyscyplinę. Dotyczy ona obrazów (zdjęć) dostarczonych na przykład przez samoloty, satelity, pojazdy itp.

W nomenklaturze wojskowej rozróżnia się trzy subdyscypliny wchodzące w skład SIGINT: rozpoznanie elektroniczne (*Electronic Intelligence*, ELINT), wywiad radiowy[38] (*Communications Intelligence*, COMINT) oraz *Foreign Instrumentation Signals Intelligence* (FISINT). Mimo takiego przyporządkowania ELINT i COMINT występują w literaturze przedmiotu jako samodzielne dyscypliny[39]. FISINT zaś zazwyczaj jest kwalifikowany jako subdyscyplina SIGINT lub MASINT, w zależności od podejścia podmiotu, który realizuje działanie i zastosowanych technik. FISINT dotyczy informacji technicznych i danych wywiadowczych pochodzących z przechwycenia emisji elektromagnetycznych związanych z testami i rozmieszczaniem operacyjnym obcych systemów w powietrzu, na lądzie i pod wodą.

---

34 OPSEC, *Intelligence Threat Handbook*, 1996, s. 1–2 (dostęp: 12.06.2019).

35 Ministry of Defence, *Joint Doctrine Publication 2-00 (JDP 2-00) Understanding and Intelligence Support to Joint Operations*, 3rd ed., 2011, s. 11, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/311572/20110830_jdp2_00_ed3_with_change1.pdf (dostęp: 19.01.2019).

36 *AAP-06. NATO Glossary of Terms…*

37 L.K. Johnson, *Introduction*, [w:] L.K. Johnson (red.), *Handbook of Intelligence Studies*, Nowy Jork 2016, s. 6.

38 Polskie tłumaczenie na podstawie: *AAP-6. Słownik terminów…*, s. 102.

39 A.D.M. Svendsen, *Collective Intelligence (COLINT)*, [w:] *Encyclopedia of U.S. Intelligence*, Nowy Jork 2015, s. 114.

COMINT dotyczy „pozyskiwania danych wywiadowczych (rozpoznawczych) za pomocą środków i systemów łączności radiowej przez osoby nie będące właściwymi odbiorcami lub użytkownikami”[40]. Amerykańska doktryna wywiadu rozróżnia informacje techniczne od wywiadowczych i uwzględnia je jako obszar zainteresowania tej dyscypliny. Ponadto zwraca uwagę, że COMINT obejmuje też zbieranie danych pochodzących z systemów zautomatyzowanych oraz sieci informatycznych przeciwnika. Może również dotyczyć zdjęć, jeśli są przetwarzane za pomocą sieci komputerowych lub urządzeń radiowych[41].

ELINT jest definiowany jako potencjał rozpoznawczy oraz proces pozyskiwania danych i informacji rozpoznawczych z systemów niekomunikacyjnych przechwyconych przez innych odbiorców niż tych, do których są one adresowane. Amerykańska doktryna wywiadu wskazuje, że dotyczy informacji technicznych i geolokalizacyjnych, których źródłem jest promieniowanie elektromagnetyczne, które jednocześnie nie stanowi komunikacji międzyludzkiej ani nie jest skutkiem promieniowania lub eksplozji nuklearnej. W jego skład wchodzą dwie subkategorie: operacyjna i techniczna. Pierwsza z nich dotyczy informacji istotnych dla operacji, są to na przykład lokalizacja, ruch, wykorzystanie, taktyka i działalność obcych emiterów i systemów uzbrojenia niesłużących komunikacji. Druga dotyczy technicznych cech emitowanych sygnałów, jak na przykład charakterystyka sygnału, tryb, funkcje, powiązania, możliwości, ograniczenia, luki w zabezpieczeniach i poziom technologii.

Zwraca uwagę multidyscyplinarne podejście do analizy wywiadowczej, które jest efektem integracji dyscyplin wywiadowczych. Wykorzystanie wszystkich dostępnych źródeł i metod jest nazywane „wywiadem ze wszystkich źródeł” (*all source intelligence*). Mimo iż proces rozpoznania, fuzji i produkcji informacji w tym podejściu jest bardziej skomplikowany i czasochłonny, to produkt informacyjny jest lepszy jakościowo i bardziej wiarygodny niż uzyskany w ramach pojedynczej dyscypliny.

---

[40] *AAP-6. Słownik terminów…*, s. 102.

[41] „Communications intelligence is technical information and intelligence derived from foreign communications by other than the intended recipients (JP 2-0). Communications intelligence includes collecting data from target or adversary automated information systems or networks. It also may include imagery when pictures or diagrams are encoded by a computer network or radio frequency method for storage or transmission. The imagery can be static or streaming”, *AAP-06. NATO Glossary of Terms…*

Dyscypliny wspierające

Doktryna wywiadu armii amerykańskiej z 2018 roku (*Army Doctrine Publication 2.0 Intelligence*) wymienia siedem dyscyplin wywiadowczych: CI (*Competitive Intelligence*), GEOINT, HUMINT, MASINT, OSINT[42], SIGINT i TECHINT. Oprócz nich definiuje również cztery dyscypliny wspierające (tabela 2).

Tabela 2. Dyscypliny wspierające

Akronim	Nazwa dyscypliny w oryginale	Nazwa dyscypliny w języku polskim	Problem
BEI	*Biometrics-enabled Intelligence*	Wywiad umożliwiony biometrią	Ustalenie tożsamości na podstawie danych biometrycznych.
_	*Cyber-enabled Intelligence*	Wywiad umożliwiony cyberprzestrzenią i spektrum elektromagnetycznym	Cyberprzestrzeń oraz spektrum elektromagnetyczne.
DOMEX	*Document and Media Exploitation*	Wywiad umożliwiony wykorzystaniem dokumentów i mediów	Zbiory dokumentów i mediów elektronicznych będących w posiadaniu rządu USA, ale niedostępnych publicznie.
FEI	*Forensic-enabled Intelligence*	Wywiad umożliwiony analizą kryminalistyczną	Analiza kryminalistyczna w celu ustalenia relacji między osobami, zdarzeniami, miejscami i przedmiotami.

Źródło: opracowanie własne na podstawie *Complementary Intelligence Capabilities*, Army Doctrine Publication (ADP 2-0) for Army Intelligence Activities, Department of the Army, Waszyngton 2018, https://fas.org/irp/doddir/army/adp2_0.pdf (dostęp: 19.02.2019).

Wywiad umożliwiony cyberprzestrzenią dotyczy zarówno pozyskiwania i przetwarzania danych z cyberprzestrzeni, jak i spektrum elektromagnetycznego. Mogą one dotyczyć środowiska informacyjnego i przestrzeni fizycznej (na przykład infrastruktury).

[42] Niekiedy opisywany jako *Digital Intelligence* (DIGINT), zob. szerzej: S.C. Mercado, *Sailing the Sea of OSINT in the Information Age*, „Studies in Intelligence" (CIA Journal) 2004, t. 48, nr 3, s. 45–55. W niektórych dokumentach i publikacjach DIGINT jest utożsamiany z CYBINT (*Cybespace Intelligence*).

Istotą wywiadu umożliwionego wykorzystaniem dokumentów i mediów (DOMEX) jest wsparcie zespołu lingwistów, których zadaniem jest archiwizacja, tłumaczenie i wstępna ocena dokumentów i materiałów multimedialnych. Produkt informacyjny może zostać wykorzystany zarówno na szczeblu taktycznym, jak i strategicznym.

Wywiad umożliwiony biometrią (BEI) polega na połączeniu informacji biometrycznych z innymi danymi wywiadowczymi, informacjami o zagrożeniach lub dotyczącymi innych aspektów środowiska operacyjnego w celu uzyskania odpowiedzi na pytania wywiadowcze. Sama biometria jest definiowana jako proces potwierdzania tożsamości na podstawie mierzalnych cech anatomicznych, fizjologicznych i behawioralnych. Warto zauważyć, że w otoczeniu współczesnego człowieka znajduje się coraz więcej urządzeń, które takie dane zbierają (na przykład monitoring miejski, urządzenia mobilne).

Wywiad umożliwiony analizą kryminalistyczną (FEI) ma na celu ustalenie relacji między osobami, zdarzeniami, miejscami i przedmiotami. Jego istotą jest integracja materiałów poddanych naukowej analizie z innymi danymi (na przykład biometrycznymi: odciski palców, DNA). Może obejmować między innymi badanie pochodzenia na przykład dokumentów, dokumentacji miejsca incydentów czy analizę przebiegu wydarzeń itd. W doktrynie brytyjskiej występuje dyscyplina będąca połączeniem analizy kryminalistycznej i biometrii (*Forensic and Biometric Intelligence*, FABINT)[43].

Podobnie jak w przypadku innych zdefiniowanych przez wojsko klasyfikacji, o przynależności do konkretnej dyscypliny decydują wewnętrzne przepisy i wytyczne. Narzędzia (na przykład technologie informatyczne) i źródła informacji mogą być wspólne dla wielu dyscyplin i dyscyplin wpierających. Powyższe zestawienie jest jedynie pewnym skrótowym omówieniem nomenklatury obowiązującej w NATO i Armii USA. Stanowi ważne źródło inspiracji i wyznacza istotne trendy w literaturze przedmiotu. Z racji swojej specyfiki sfera cywilna wypracowuje pojęcia, które nie muszą być kompatybilne z wojskowymi. Jest to zależne od konkretnego podmiotu oraz jego specyficznych potrzeb i możliwości.

[43] Ministry of Defence, *Joint Doctrine Publication 2-00 (JDP 2-00)…*

Nowe dyscypliny

Dynamicznie zmieniająca się rzeczywistość, a przede wszystkim tzw. Gospodarka 4.0, przyczyniła się do zdefiniowania nowych dyscyplin dotyczących pozyskiwania informacji. Są one oparte na źródłach i problemach, które albo wcześniej nie występowały (jak Internet rzeczy), albo brakowało odpowiednich technologii pozwalających na ich wyodrębnienie.

Tabela 3 wymienia dyscypliny, które zostały zdefiniowane w literaturze przedmiotu, ale nie zostały uwzględnione jako samodzielne dyscypliny we wspomnianych wyżej dokumentach doktrynalnych.

Tabela 3. Nowe dyscypliny pozyskiwania informacji

Akronim	Nazwa dyscypliny w oryginale	Nazwa dyscypliny w języku polskim	Problem
ADINT	*Advertising-based Intelligence*	Wywiad oparty na technologiach kierowania (targetowania) reklam w aplikacjach mobilnych	Realizowany z wykorzystaniem sieci reklamowych (infrastruktury, której zadaniem jest łączyć reklamodawców z miejscem wyświetlania reklamy, jak na przykład aplikacja mobilna).
COLINT	*Collective Intelligence*	Wywiad zbiorowy	Współpraca i/lub rywalizacja podmiotów z zakresu wszystkich dyscyplin wywiadowczych[a].
IoTINT	*Internet of Things Intelligence*	Wywiad oparty na Internecie rzeczy	Dane zebrane przez sensory urządzeń Internetu rzeczy.
MARKINT/ MI	*Market Intelligence*	Wywiad rynkowy	Dane dotyczące działalności rynkowej i konkurencji.
RESINT[b]	*Research Intelligence*	Wywiad naukowy	Informacja i wiedza wywodząca się z działalności badawczej.
SOCMINT	*Social Media Intelligence*	Wywiad oparty na mediach społecznościowych	Interakcje międzyludzkie w mediach społecznościowych.

[a] COLINT różni się od wywiadu ze wszystkich źródeł (*all source intelligence*) tym, że jego immamentną cechą jest głęboka współpraca między podmiotami specjalizującymi się w różnych dyscyplinach.

[b] A.D.M. Svendsen, *Introducing RESINT: A Missing and Undervalued „INT" in All-Source Intelligence Efforts*, „International Journal of Intelligence and CounterIntelligence" 2013, nr 26, s. 777–794.

Źródło: opracowanie własne.

Wywiad rynkowy (*Market Intelligence*, MARKINT[44]) jest definiowany jako proces, którego zadaniem jest ciągłe generowanie wiedzy ze źródeł rozproszonych na użytek zarządzania strategicznego organizacją. Celem jest wsparcie procesu podejmowania decyzji odnoszących się do konkurowania w środowisku biznesowym[45]. W gestii zainteresowania tej dyscypliny znajdują się dane dotyczące na przykład pozycji rynkowej, działalności konkurencji, strategii sprzedaży itp.

Adam D.M. Svendsen definiuje dwie nowe dyscypliny: COLINT i RESINT. Pierwsza polega na ustanowieniu w instytucjonalnej formie głębokiej współpracy między podmiotami specjalizującymi się we wszystkich możliwych dyscyplinach. Określa produkt informacyjny, który powstał w wyniku pracy z zakresu różnych dyscyplin oraz we współpracy i/lub konkurencji wszystkich możliwych podmiotów[46]. W tym zakresie jest tożsamy z wywiadem opartym na wszystkich źródłach (*all source intelligence*). Ponadto odnosi się do instytucji zajmującej się pozyskiwaniem i analizą informacji, która powstała z inicjatywy kilku różnych podmiotów i jest przez nie utrzymywana. Druga dyscyplina (*Research Intelligence*, RESINT) dotyczy informacji, które powstają w wyniku analizy badań naukowych[47]. Wywiad naukowy może być realizowany na przykład w celu dostarczenia odpowiedniego opracowania wyników badań do odbiorcy, który jest zainteresowany ich wykorzystaniem. A.D.M. Svendsen wskazuje, że RESINT dotyka szerszego obszaru niż OSINT ze względu na możliwość wykorzystania go do optymalizacji całego procesu pozyskiwania, przetwarzania, dystrybucji i wykorzystania informacji[48].

W 2016 roku amerykańska Rada Naukowa Armii (Army Science Board) dyskutowała na temat nowej dyscypliny specjalizującej się w pozyskiwaniu informacji za pomocą Internetu rzeczy[49]. Zidentyfikowanie tego obszaru ma prowadzić do zwiększenia świadomości sytuacyjnej żołnierzy operujących na terenie zurbanizowanym. Może on dostarczyć informacji

---

[44] Niektórzy autorzy stosują określenie *Competitive Intelligence*, CI.

[45] G.L. Jamil, L.H.R. Santos, M.L. Alves, L. Furbino, *A Design Framework for a Market Intelligence System for Healthcare Sector: a Support Decision Tool in an Emergent Economy*, [w:] *Handbook on Research of ICTs for Social Services and Healthcare: Developments and Applications*, Hershey 2012.

[46] A.D.M. Svendsen, *Collective Intelligence…*, s. 114–119.

[47] A.D.M. Svendsen, *Introducing RESINT: A Missing and Undervalued „INT" in All-Source Intelligence Efforts*, „International Journal of Intelligence and Counter Intelligence" 2013, nr 26, s. 778.

[48] Tamże, s. 780–781.

[49] M.L. Loper, *Situational Awareness in Megacities*, [w:] M.E. Kosal (red.), *Technology and the Intelligence Community Challenges and Advances for the 21st Century*, Cham 2018, s. 226.

o: lokalizacji, ładunku i pasażerach pojazdów; budynkach wraz z identyfikacją osób, które w nich przebywają; zasobach (na przykład żywności); informacjach medycznych dotyczących populacji miasta itp. Może również pomóc rozróżnić cywili od żołnierzy przeciwnika.

Dostęp do danych publikowanych za pośrednictwem portali społecznościowych spowodował zdefiniowanie odpowiedniej dyscypliny – *Social Media Intelligence* (SOCMINT)[50]. W odróżnieniu do choćby HUMINT (rozpoznania osobowego) pozwala na opracowanie informacji o konkretnych grupach osób odnoszących się do konkretnych wydarzeń w czasie rzeczywistym (jak trendy zachowań użytkowników Twittera)[51]. Zdaniem publicystów rosyjski wywiad wykorzystał głównie dane zebrane za pośrednictwem mediów społecznościowych w celu wywarcia wpływu na wybory prezydenckie w Stanach Zjednoczonych w 2016 roku[52]. Teoretycy kwalifikują SOCMINT jako subdyscyplinę OSINT-u[53].

Na szczególną uwagę zasługują wyniki badań naukowców z Uniwersytetu Waszyngtonu. W celu zwrócenia uwagi na zagadnienie prywatności przeprowadzili oni eksperyment, w którym wykazali, że istnieje możliwość pozyskania danych wrażliwych użytkowników urządzeń mobilnych przy wykorzystaniu środowiska reklam internetowych. Zakupili usługę wyświetlania reklam w aplikacji mobilnej i za jej pomocą uzyskali dane lokalizacji urządzenia należącego do określonej osoby[54]. Zdefiniowali *Adware-based Intelligence* (ADINT) jako metodę wykorzystywania ekosystemu reklamowego przez nabywcę reklam do zbierania informacji o osobach docelowych. Przy użyciu podobnego mechanizmu nabywca reklam może uzyskać też inne dane zbierane i udostępniane sieciom reklamowym przez aplikacje – na przykład lista zainstalowanych aplikacji może wskazać zainteresowania obiektu, płeć, styl życia, światopogląd czy orientację seksualną. Aplikacje często mają też dostęp do kontaktów, czujników, mikrofonu, aparatu itp. Te dane są zanonimizowane, ale jak

[50] D. Omand, J. Bartlett, C. Miller, *Introducing Social Media Intelligence (SOCMINT)*, „Intelligence and National Security" 2012, nr 27(6), s. 801–823.

[51] K.P. Jani, A. Soni, *Promise and Perils of Big Data Science for Intelligence Community*, [w:] M.E. Kosal (red.), *Technology and the Intelligence*…, s. 192.

[52] Źródło: S. Shane, *These are the Ads Russia Bought on Facebook in 2016*, „The New York Times", 1.11.2017, https://www.nytimes.com/2017/11/01/us/politics/russia-2016-election-facebook.html (dostęp: 19.02.2019).

[53] A.N. Liaropoulos, *The Challenge of Social Media for the Intelligence Community*, „Journal of Mediterranean and Balkan Intelligence" 2013, nr 1(1), s. 6.

[54] P. Vines, F. Roesner, T. Kohno, *Exploring ADINT: Using Ad Targeting for Surveillance on a Budget – or – How Alice Can Buy Ads to Track Bob*, Waszyngton 2017, https://adint.cs.washington.edu/ADINT.pdf (dostęp: 15.07.2018).

pokazuje wynik wyżej wspomnianego eksperymentu, dysponując odpowiednimi informacjami (na przykład dotyczącymi miejsca pracy i miejsca zamieszkania) możliwe jest powiązanie ich z konkretną osobą. Opisana metoda wykorzystuje ekosystem reklamowy urządzeń mobilnych. Zatem można ją uznać za charakterystyczną dla IoTINT. Zakwalifikowanie jej jako oddzielnej dyscypliny powinno być poprzedzone badaniami dotyczącymi wykorzystywania tej metody oraz dyskusją.

Użytkownicy urządzeń mobilnych, korzystając z ich szerokich możliwości, zostawiają na nich pewnego rodzaju cyfrową kopię swojej świadomości. Jacek Dukaj nazwał to „protezą umysłu"[55], co jest kontynuacją idei przedstawionej w pracy *The Extended Mind* z 1998 roku, której autorzy próbowali zdefiniować granicę ludzkiej świadomości[56]. Trudno wykluczyć, że w niedalekiej przyszłości broker informacji będzie w stanie nie tylko precyzyjnie udzielić odpowiedzi na pytanie, kim jest dana osoba, ale również kim będzie i jak postąpi w przyszłości.

Podsumowanie

Twórca koncepcji społeczeństwa informacyjnego Manuel Castells wysnuł tezę, że wytwarzanie, przetwarzanie i transmisja informacji stanowią podstawowe źródło produktywności i bogactwa[57]. Wraz ze wzrastającą ilością różnych urządzeń w naszym otoczeniu wzrasta też liczba przetwarzanych danych, by – teoretycznie – lepiej nam służyć. Zebrane dane są przetwarzane i wykorzystywane nie tylko po to, by usprawnić urządzenia w naszym otoczeniu, ale by zdobyć naszą uwagę i zasugerować pewne wybory. W obecnie rozwijającym się modelu gospodarczym coraz bardziej widoczny jest proces masowego utowarowienia intymnych aspektów ludzkiego życia. Na masową skalę rynek eksploruje obszary ludzkiej osobowości, emocji i życia intymnego[58]. Zjawisko monetyzacji

---

55 *Smartfon jest protezą naszego umysłu – Jacek Dukaj*, „Rozmowy o Przyszłości", Onet News (wideo), https://www.youtube.com/watch?v=UuEPplXAtJQ (dostęp: 19.01.2019). Prawdopodobnie J. Dukaj użył tego terminu w nawiązaniu do badania: N. Barr, G. Pennycook, J.A. Stolz, J.A. Fugelsang, *The Brain in Your Pocket: Evidence that Smartphones are Used to Supplant Thinking*, „Computers in Human Behavior" 2015, nr 48, s. 473–480.

56 Zob. szerzej: A. Clark, D. Chalmers, *The Extended Mind*, „Analysis" 1998, t. 58, nr 1, s. 7–19.

57 M. Castells, *The Rise of the Network Society. The Information Age. Economy, Society and Culture. Volume 1*, Oxford 1996, s. 17.

58 Zob. E. Illouz, *Uczucia w dobie kapitalizmu*, przekł. Z. Simbierowicz, Warszawa 2010.

uwagi starają się wyjaśnić koncepcje kapitalizmu kognitywnego[59] i ekonomii uwagi[60].

Zwiększenie możliwości pozyskiwania informacji poszerza obszary zainteresowania informacyjnego i powoduje wzrost wymagań wobec systemów przetwarzania wielkich zbiorów danych. Wielość codziennych czynności, które człowiek wykonuje w asyście lub przy pomocy urządzeń mających możliwość zbierania danych i komunikacji z użytkownikiem (jak smartfony), umożliwiła oddziaływanie w czasie rzeczywistym na jego wybory poprzez sprzężenie zwrotne systemów profilowania z algorytmami dokonującymi wyboru wyświetlanych treści w mediach społecznościowych.

Charakter rynku opartego na informacji powoduje, że te powszechnie dostępne są zazwyczaj powierzchowne, a dostawcy usług informacyjnych coraz częściej dostarczają odbiorcy gotową interpretację zamiast produktu informacyjnego, analizę dokonywaną z własnego punktu widzenia i wynikające z niej wnioski. Wykorzystanie danych surowych jest coraz bardziej skomplikowane i wymaga specjalistycznej wiedzy[61]. Analiza takich danych jest ogromnym wyzwaniem dla brokerów informacji, a podmioty projektujące systemy zbierania danych na przykład z określonej rodziny urządzeń będą coraz częściej utrudniać pracę niezależnych infobrokerów w celu choćby skłonienia ich do zakupu dedykowanego oprogramowania do analizy lub udziału w kosztownych szkoleniach. Jednak z drugiej strony analiza danych w pewnych obszarach stała się obecnie dostępna dla każdego. To generuje też wyzwania terminologiczne, gdyż coraz trudniej zakwalifikować działalność na przykład aktywistów dokonujących triangulacji danych zebranych z otwartych źródeł i zamkniętych grup dyskusyjnych, którzy następnie poddają je wielowarstwowej analizie (na przykład obraz, dźwięk, metadane, artefakty[62]) do określonego obszaru[63].

---

[59] Termin zaproponował Lorenzo Cillario: *Il capitalismo cognitivo: Saper, sfruttamento e accumulazione dopo la rivoluzione informatica*, [w:] L. Cillario i in. (red.), *Trasformazione e persistenza. Saggi sulla storicità*, Mediolan 1990. Koncepcję rozwinął m.in.: Y. Moulier Boutang, *Cognitive Capitalism*, Amsterdam 2011.

[60] M.H. Goldhaber, *The Attention Economy and the Net*, „First Monday" 1997, t. 2, nr 4.

[61] T.R. Aleksandrowicz, *Podstawy walki…*, s. 53.

[62] Wady powstające najczęściej podczas kompresji lub zmiany formatu danych. Ich analiza pozwala wykryć, czy np. zdjęcie satelitarne lub nagranie zostało poddane manipulacji.

[63] Przykładem takiej działalności jest założony przez brytyjskiego dziennikarza Eliota Higginsa portal bellingcat.com. Jedną z najbardziej znanych opublikowanych analiz jest dotycząca zestrzelenia holenderskiego samolotu pasażerskiego MH17 przez rosyjskie wojska nad Ukrainą w 2014 r. Źródło: https://www.bellingcat.com/news/uk-and-europe/2017/07/17/mh17-open-source-investigation-three-years-later/ (dostęp: 19.02.2019).

Dzięki nowym obszarom i narzędziom pozyskiwania informacji rola brokera może znacząco wzrosnąć w najbliższej przyszłości. Z jednej strony zaawansowane algorytmy będą mogły odpowiadać na coraz bardziej abstrakcyjne zapytania, ale nie oznacza to, że zastąpią człowieka, który dzięki swojej wiedzy i umiejętnościom dobierze odpowiednie narzędzia, zweryfikuje i dokona ostatecznej obróbki produktu informacyjnego. Umiejętność doboru narzędzi i zdolność do weryfikacji są już dziś kluczowymi kompetencjami. Nie sposób jednak określić precyzyjnie, jak infobrokering będzie rozwijać się w przyszłości – można założyć, że jednym z istotnych kierunków będzie ten wyznaczony przez nanotechnologie, biotechnologie, informatykę i kognitywistykę, które wspólnie podążają kierunkiem modyfikacji ludzkiego mózgu.

STRESZCZENIE

Dynamiczny rozwój technologii stwarza potrzebę uporządkowania istniejących i konceptualizacji nowych metod pozyskiwania informacji. Aktualna typologia zawarta w dokumentach doktrynalnych NATO i Armii Stanów Zjednoczonych częściowo koresponduje z możliwościami, jakie dają nowe technologie. Na uwagę zasługują również propozycje nowych dyscyplin, które w ostatnich latach pojawiły się w literaturze przedmiotu. Analizie towarzyszy refleksja na temat nowych trendów w pracy brokera informacji.

Piotr Sosnowski

SYSTEMATISATION OF CONCEPTS RELATED TO METHODS AND SOURCES OF GATHERING INFORMATION FOR THE INFOBROKERING

The dynamic development of technology causes the need to organize existing and conceptualize new methods of gathering information. Current typology contained in the doctrinal documents of NATO and the United States Army partially corresponds to the possibilities offered by new technologies. Noteworthy are also the proposals for new disciplines, which in recent years appeared in the literature. The analysis includes reflection about the new trends in the work of the information broker.

KEY WORDS: *information brokering, information gathering, information sources, intelligence collection*

Bibliografia

Aid M.M., *Measurement and Signature Intelligence*, [w:] R. Dover, M.S. Goodman, C. Hillebrand (red.), *Routledge Companion to Intelligence Studies*, Londyn 2014.

Aleksandrowicz T.R., *Podstawy walki informacyjnej*, Warszawa 2016.

Aleksandrowicz T.A., *Biały wywiad w walce z terroryzmem*, [w:] K. Liedel, P. Piasecka (red.), *Rola mediów w przeciwdziałaniu terroryzmowi*, Warszawa 2009.

Barr N., Pennycook G., Stolz J.A., Fugelsang J.A., *The Brain in Your Pocket: Evidence that Smartphones are Used to Supplant Thinking*, „Computers in Human Behavior" 2015, nr 48.

Castells M., *The Rise of the Network Society. The Information Age. Economy, Society and Culture. Volume 1*, Oxford 1996.

Ciecierski M., *Wywiad biznesowy w korporacjach transnarodowych. Teoria i praktyka*, Toruń 2009.

Cillario L., *Il capitalismo cognitivo: Saper, sfruttamento e accumulazione dopo la rivoluzione informatica*, [w:] L. Cillario i in. (red.), *Trasformazione e persistenza. Saggi sulla storicità*, Mediolan 1990.

Clark A., Chalmers D., *The Extended Mind*, „Analysis" 1998, t. 58, nr 1.

Dukaj J., *Smartfon jest protezą naszego umysłu*, „Rozmowy o Przyszłości", Onet News, (wideo), https://www.youtube.com/watch?v=UuEPplXAtJQ (dostęp: 19.01.2019).

Elvy S., *Paying for Privacy and the Personal Data Economy*, „Columbia Law Review" 2017, t. 117, nr 6.

Favarel-Garrigues G., Godefroy T., Lascoumes P., *Reluctant Partners? Banks in the Fight against Money Laundering and Terrorism Financing in France*, „Security Dialogue" 2011, nr 42(2).

Giles M., *Bounty Hunters Tracked People Secretly Using US Phone Giants Location Data*, „MIT Technology Review", 7.02.2019, https://www.technologyreview.com/the-download/612907/bounty-hunters-tracked-people-secretly-using-us-phone-giants-location-data (dostęp: 8.02.2019).

Hurley M.M., *For and from Cyberspace Conceptualizing Cyber Intelligence, Surveillance, and Reconnaissance*, „Air and Space Power Journal" 2012, nr 26(6).

Jamil G.L., Santos L.H.R., Alves M.L., Furbino L., *A Design Framework for a Market Intelligence System for Healthcare Sector: a Support Decision Tool in an Emergent Economy*, [w:] *Handbook on Research of ICTs for Social Services and Healthcare: Developments and Applications*, Hershey 2012.

Jani K.P., Soni A., *Promise and Perils of Big Data Science for Intelligence Community*, [w:] M.E. Kosal (red.), *Technology and the Intelligence Community Challenges and Advances for the 21st Century*, Cham 2018.

Liaropoulos A.N., *The Challenge of Social Media for the Intelligence Community*, „Journal of Mediterranean and Balkan Intelligence" 2013, nr 1(1).

Loper M.L., *Situational Awareness in Megacities*, [w:] M.E. Kosal (red.), *Technology and the Intelligence Community Challenges and Advances for the 21st Century*, Cham 2018.

Mercado S.C., *Sailing the Sea of OSINT in the Information Age*, „Studies in Intelligence" (CIA Journal) 2004, t. 48, nr 3.

Murch R., *A Perspective on the Strategy of Intelligence*, [w:] V. Radosavljevic, I. Banjari, G. Belojevic (red.), *Defence Against Bioterrorism. Methods for Prevention and Control*, Dordrecht 2018.

Omand D., Bartlett J., Miller C., *Introducing Social Media Intelligence (SOCMINT)*, „Intelligence and National Security" 2012, nr 27(6).

Płoszajski P., *Big Data: nowe źródło przewag i wzrostu firm*, „E-mentor" 2013, nr 3(50).

Porter M.E., *Competitive Advantage: Creating and Sustaining Superior Performance*, Nowy Jork 2008.

Saramak B., *Wykorzystanie otwartych źródeł informacji w działalności wywiadowczej: historia, praktyka, perspektywy*, Warszawa 2015.

Selding P.B. de, *Pentagon Struggles with Avalanche of Data*, SpaceNews, 29.11.2011, https://spacenews.com/pentagon-struggles-avalanche-data (dostęp: 19.10.2011).

Shane S., *These are the Ads Russia Bought on Facebook in 2016*, „The New York Times", 1.11.2017, https://www.nytimes.com/2017/11/01/us/politics/russia-2016-election-facebook.html (dostęp: 19.02.2019).

Svendsen A.D.M., *Collective Intelligence* (*COLINT*), [w:] *Encyclopedia of U.S. Intelligence*, Nowy Jork 2015.

Svendsen A.D.M., *Introducing RESINT: A Missing and Undervalued „INT" in All-Source Intelligence Efforts*, „International Journal of Intelligence and CounterIntelligence" 2013, nr 26.

Vines P., Roesner F., Kohno T., *Exploring ADINT: Using Ad Targeting for Surveillance on a Budget – or – How Alice Can Buy Ads to Track Bob*, Waszyngton 2017, https://adint.cs.washington.edu/ADINT.pdf (dostęp: 15.07.2018).

VOL. 54 STUDIA POLITOLOGICZNE
DOI: 10.33896/SPolit.2019.54.3

Piotr Dela
ORCID:0000-0003-3643-3759

Elementy propagandy w życiu publicznym

Słowa kluczowe:
propaganda, dezinformacja, bezpieczeństwo informacyjne

STUDIA I ANALIZY

Wprowadzenie

Obserwując współczesne wydarzenia, można wyciągnąć wniosek, że cokolwiek stanie się na świecie, natychmiast jesteśmy o tym poinformowani. Niemniej jednak informacja, która dociera do społeczeństwa, nie jest dokładnym odzwierciedleniem rzeczywistości, lecz jedynie jej interpretacją, która wykorzystywana jest do stworzenia odpowiedniej reakcji jej odbiorców. Informacja, bo ona ma tutaj największe znaczenie, to narzędzie polityki wykorzystywane zarówno na poziomie lokalnym (państwa), jak i globalnym (międzynarodowym). Przekazy informacyjne, zarówno współczesne jak i na przestrzeni wieków, są elementem szeroko pojmowanej walki politycznej, walki informacyjnej (w ujęciu militarnym i niemilitarnym), a także propagandy stanowiącej ich nieodzowny element. Niniejszy tekst skupia się na najważniejszych aspektach współczesnej propagandy. Jej oblicze jest niezwykle zmienne i ma odzwierciedlenie w środowisku, w którym jest ona realizowana. Integralnymi częściami tego środowiska są przestrzeń bezpieczeństwa państwa, cyberprzestrzeń i przestrzeń informacyjna. Oczywiście przestrzenie te są ze sobą ściśle powiązane i wywierają na siebie wzajemny wpływ. Głównym celem opracowania jest zidentyfikowanie oblicza współczesnej propagandy i jej znaczenia w życiu publicznym. Realizuje ona bowiem cele polityczne,

prowadzona jest zgodnie z pewnymi utartymi technikami i etapami w środowisku, w którym znaczącą rolę odgrywają media publiczne i komercyjne, ale także z coraz większą rolą cyberprzestrzeni, ze szczególnym uwzględnieniem mediów społecznościowych. Realizacja powyższego celu wymaga w pierwszej kolejności identyfikacji podstawowych terminów, takich jak informacja, przestrzeń bezpieczeństwa państwa, przestrzeń informacyjna i cyberprzestrzeń, bezpieczeństwo informacyjne, walka informacyjna, dezinformacja i propaganda. Następnie przedstawione zostaną podstawowe rodzaje propagandy, techniki, przy pomocy których jest ona realizowana, oraz przebieg klasycznej kampanii informacyjnej (propagandowej).

Informacja – uwarunkowania środowiskowe

Współcześnie istnieje szereg różnorodnych definicji oraz ujęć terminu informacja. Przez wielu teoretyków jest ona uważana za pojęcie pierwotne, niedające się zdefiniować. Ponadto funkcjonowanie informacji w różnorodnym środowisku (politycznym, fizycznym, matematycznym, ekonomicznym itp.) spowodowało, iż do tej pory nie doczekała się ona jednoznacznej definicji wyrażającej jej istotę. Część autorów rezygnuje z jej definiowania, poprzestając na intuicyjnym i potocznym jej rozumieniu. Niemniej jednak należy zauważyć, że samo słowo informacja wywodzi się z łacińskiego słowa *informatio*, co oznacza wyobrażenie, wyjaśnienie, zawiadomienie. Interesujący zbiór definicji terminu informacja przedstawia w jednej ze swoich publikacji Wiesław Flakiewicz[1]. Przytacza on najpopularniejsze definicje tego pojęcia:

- Informacja to komunikacja, łączność, w wyniku której likwiduje się nieokreśloność (Claude E. Shannon).
- Informacja jest nazwą treści zaczerpniętej ze świata zewnętrznego, nie jest więc ani materią, ani energią (Norbert Wiener).
- Informacja jest to czynnik sterujący strumieniami zasileń, wykorzystywany w organizmach żywych lub maszynach do bardziej sprawnego, efektywnego i celowego działania (Edward Kowalczyk).
- Informacja jest to treść przekazywanych od nadawcy do odbiorcy wiadomości, będąca opisem, poleceniem, zakazem, nakazem lub zleceniem (Janusz Gościński).

[1] W. Flakiewicz, *Podejmowanie decyzji kierowniczych*, Warszawa 1973, s. 38.

- Jest to przekazywanie wiedzy do odbiorcy informacji, umożliwiające ze względu na jej wartość zmniejszenie niepewności działania odbiorcy informacji (Russell L. Ackoff).
- Informacja jest to wiedza przekazywana przez innych ludzi bądź uzyskiwana przez studia, obserwacje, badania (A. Webster).

Także Leopold Ciborowski w swojej książce *Walka informacyjna*, podejmując się analizy pojęcia informacja, powołał się na kilkanaście definicji stworzonych przez naukowców z dziedziny cybernetyki i fizyki. Podsumowując cytowane definicje zauważył on, że informacja jest niejako bodźcem oddziałującym na układ recepcyjny człowieka, powodującym w jego wyobraźni wytworzenie przedmiotu myślowego, który odzwierciedla obraz rzeczy materialnej i abstrakcyjnej oraz w jego przekazie kojarzy się z tym bodźcem. Ponadto Ciborowski uważa, że związek między człowiekiem i informacją jest nierozerwalny: „Tak jak foton nie może istnieć bez pędu, tak informacja nie może istnieć bez umysłu ludzkiego. Tylko ten organ natury ludzkiej dostosowany jest do nieskończonego przetwarzania transformowanych doznań recepcyjnych w wyobrażenia informacyjne. […] każda informacja jest szczególną formą sygnału, która oprócz wspólnych cech wyróżnialności, właściwych dla sygnału i informacji, posiada jeszcze tę właściwość, że inspiruje umysł ludzki do tworzenia pewnej wyobraźni”[2].

Interesującą definicję informacji przedstawił także Piotr Sienkiewicz, dla którego informacja to „zbiór faktów, zdarzeń, cech, obiektów ujęty w takiej formie, że pozwala odbiorcy ustosunkować się do zaistniałej sytuacji i podjąć odpowiednie działania umysłowe lub fizyczne”[3].

W literaturze można także zauważyć, że sama informacja ma wymiar strukturalny. Podejście strukturalne mówi o tym, że w każdej strukturze zawiera się informacja warunkująca zachowanie odpowiedniej (właściwej) formy tej struktury, odnosząca się do celów, wartości oraz sposobów funkcjonowania danej struktury (organizacji, społeczności).

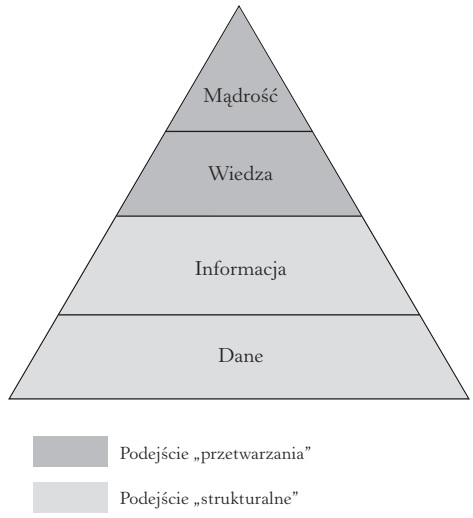
Zapoznając się z przedstawionymi na rysunku 1 elementami – zaczynając od dołu do góry, czyli od surowych danych (ich fizycznego aspektu – fizycznej postaci), przez zorganizowane informacje (dane mogące być zinterpretowane przy pomocy aparatu poznawczego) i informacje przetworzone, aż do poziomu wiedzy (zinterpretowane i przyswojone) oraz najwyższego poziomu informacji, czyli mądrości (bazującej na wiedzy, wykształceniu i doświadczeniu) – należy zauważyć, że dane, aby mogły

[2] L. Ciborowski, *Walka informacyjna*, Toruń 1999, s. 59.

[3] P. Sienkiewicz, *Systemy kierowania*, Warszawa 1989, s. 128.

stać się informacją, muszą zostać przetworzone, a do tego niezbędna jest wiedza. Dwa najniższe szczeble odnoszą się do podejścia przetwarzania, a dwa górne do podejścia strukturalnego. Szczeble te nie są od siebie zależne, ponieważ większa liczba danych nie oznacza większej liczby informacji. Podobnie większa liczba informacji nie jest tożsama z większą wiedzą, a czasami wręcz przeciwnie – zbyt wielka liczba informacji ogranicza wiedzę, a nawet paraliżuje jej działania.

Rysunek 1. Piramida informacyjna



Źródło: P. Sienkiewicz, *Wiek informacji*, Warszawa 2000.

Informacja jest podstawą do budowania każdego systemu wiedzy, warunkuje sukces. Możemy zauważyć, że informacja nabiera ogromnego znaczenia w każdej dziedzinie działalności człowieka, a w szczególności w takich sferach, jak: społeczna, gospodarcza, zarządzania, polityczna, kulturowo-religijna, bezpieczeństwa.

Przedstawione powyżej rozważania uprawniają do stwierdzenia, że informacja jest czymś więcej niż tylko wiadomością, znakiem lub inną formą komunikowania. W swojej istocie jest zarówno opisem rzeczywistości, jak i odzwierciedleniem stanu systemu oraz jego elementów, który może podlegać procesom pozyskiwania, przetwarzania, gromadzenia lub dystrybucji. Uwarunkowania te są niezwykle istotne z punktu widzenia działań, jakie podejmowane są w trakcie kampanii propagandowych. Informacja bowiem jest odbiciem obrazu obserwowanej rzeczywistości, która nie musi być odzwierciedleniem prawdy lub stanu faktycznego. Na bazie

tego odbicia budowana jest wiedza, która jest równocześnie pochodną wielu różnorodnych czynników, takich jak: wykształcenie, doświadczenie, przekonanie obserwatora itp. Wiedza, najogólniej rzecz ujmując, to ogół wiarygodnych informacji o rzeczywistości wraz z umiejętnością ich wykorzystywania[4]. Przymiotnik „wiarygodnych" jest gwarantem poprawności interpretacji informacji i odpowiedniego ustosunkowania się do niej. Jeżeli uzyskany obraz rzeczywistości odbiegać będzie w jakimś stopniu od prawdy, jeżeli uzyskana informacja nie będzie cechowała się wiarygodnością, to stworzona na jej podstawie wiedza nie będzie gwarantowała prawidłowej interpretacji rzeczywistości. W takie podejście do przekazu informacji wpisuje się propaganda.

Istotnym terminem związanym bezpośrednio z informacją jest przestrzeń informacyjna, której rozwój nierozerwalnie jest związany z rozwojem ludzkości. Pierwotnie przestrzeń ta obejmowała swym zasięgiem lokalne społeczności, takie jak plemiona. Z upływem czasu obszar jej oddziaływania zwiększał się wraz z rozwojem struktur społecznych i wykorzystywanych narzędzi komunikacyjnych. Narzędzia te jednak miały najczęściej ograniczenia związane z zasięgiem i czasem przekazywania informacji. Dopiero rozwój telekomunikacji, informatyki i teleinformatyki spowodował diametralne zmiany w przestrzeni informacyjnej i nie może być ona dziś kojarzona z jakąkolwiek społecznością lokalną. Obejmuje swym zasięgiem praktycznie całą powierzchnię kuli ziemskiej przy jednoczesnym skróceniu czasu przesyłu informacji do minimum. Na rysunku 2 przedstawiono drzewo rozwoju domen komunikacyjnych, które były i są istotnym elementem każdej kampanii propagandowej. Należy zwrócić uwagę, że najwyżej w hierarchii znajduje się cyberprzestrzeń, która stała się najważniejszym elementem informacyjnego oddziaływania we współczesnym świecie.

Najogólniej rzecz ujmując, współczesna przestrzeń informacyjna jest złożoną całością interakcji, wartości i funkcji łączących światy realne, wirtualne, indywidualne, społeczne, przeszłe, obecne i przyszłe[5]. To nie tylko swoiste narzędzie wymiany informacji pomiędzy elementami składowymi systemu bezpieczeństwa podmiotu, ale również odbicie ich wzajemnych powiązań i relacji[6].

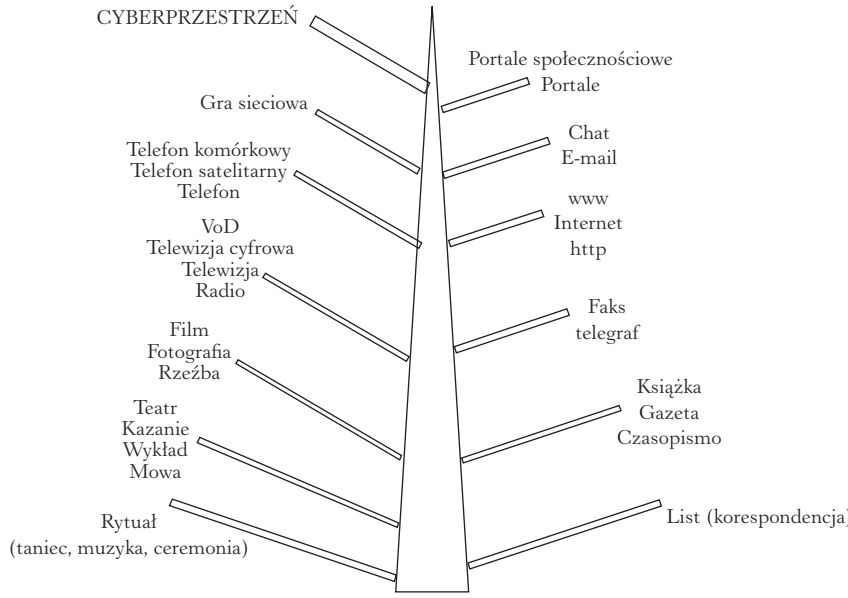
---

[4] Hasło: *Wiedza*, [w:] *Encyklopedia PWN*, http://encyklopedia.pwn.pl/haslo/3995573/wiedza.html (dostęp: 22.12.2018).

[5] Por. S. Jaskuła, *Informacyjna przestrzeń tożsamości*, Warszawa 2010, s. 1.

[6] Por. R. Kwećka, *Strategia bezpieczeństwa informacyjnego państwa*, Warszawa 2014, s. 25.

Rysunek 2. Studium rozwoju domen komunikacyjnych



Źródło: na podstawie: B. Siemieniecki (red.), *Pedagogika medialna*, t. I, Warszawa 2007, s. 154.

Przechodząc do identyfikacji następnego istotnego terminu – cyberprzestrzeni, należy zauważyć, że jest to środowisko, które swój początek miało wraz z powstaniem sieci Internet, najczęściej definiowanej jako ogólnoświatowa sieć komputerowa, tak zwana sieć sieci. Początki Internetu sięgają końca lat 60. XX wieku i związane są bezpośrednio z powstaniem sieci rozległej ARPANET. W wyniku rozwoju technologii i architektury, poprzez łączenie coraz to nowych sieci lokalnych, sieć ta pokryła swoim zasięgiem praktycznie cały obszar kuli ziemskiej i stała się ogólnie dostępna dla większości populacji. Przełomowymi momentami jej rozwoju było opracowanie technologii WWW i pojawienie się portali społecznościowych, które zapoczątkowały rewolucję w dostępie społeczeństwa do informacji. Tym samym system techniczny, jakim jest sieć Internet, przerodził się w system społeczny, jaką jest cyberprzestrzeń.

Pierwsze użycie terminu cyberprzestrzeń przypisuje się Williamowi Gibsonowi, który w 1984 roku użył go w powieści *Neuromancer*[7] dla opisania nowego świata. Wprawdzie użyte przez Gibsona określenie cyberprzestrzeni trudno odnieść do świata współczesnej nauki, jednak należy zauważyć, że przed nim nikt inny nie użył takiego terminu w celu

[7] W. Gibson, *Neuromancer*, Londyn 1984.

wyartykułowania bytu służącego do przetwarzania, przesyłania i przechowywania informacji przez miliardy użytkowników.

W polskim systemie prawnym występuje wiele definicji cyberprzestrzeni. Jedna z nich jest zawarta w nowelizacjach ustaw o stanach nadzwyczajnych[8], która postrzega cyberprzestrzeń jako przestrzeń przetwarzania i wymiany informacji tworzoną przez systemy teleinformatyczne (w rozumieniu art. 3 pkt 3 ustawy o informatyzacji działalności podmiotów realizujących zadania publiczne[9]) wraz z powiązaniami pomiędzy nimi oraz relacjami z użytkownikami. System teleinformatyczny według wspomnianej ustawy o informatyzacji to zespół współpracujących ze sobą urządzeń informatycznych i oprogramowania zapewniający przetwarzanie, przechowywanie, a także wysyłanie i odbieranie danych przez sieci telekomunikacyjne za pomocą właściwego dla danego rodzaju sieci telekomunikacyjnego urządzenia końcowego (w rozumieniu przepisów ustawy Prawo telekomunikacyjne[10]). Urządzeniem końcowym w tym ujęciu są urządzenia telekomunikacyjne przeznaczone do podłączenia bezpośrednio lub pośrednio do zakończeń sieci.

Cyberprzestrzeń definiowana jest także w środowisku międzynarodowym. Spośród wielu różnorodnych definicji na uwagę zasługują trzy najważniejsze. Jedną z powszechnie cytowanych definicji jest ta opracowana przez Departament Obrony USA, według której cyberprzestrzeń jest „globalną domeną środowiska informacyjnego składającą się z współzależności (IT) oraz zawartych w nich danych, włączając Internet, sieci telekomunikacyjne a także osadzone w nich procesory i kontrolery”[11]. Należy zauważyć, że definicja ta odnosi się tylko i wyłącznie do sfery technicznej cyberprzestrzeni i nie uwzględnia aspektu społecznego – jej roli i znaczenia dla współczesnych społeczeństw. W Unii Europejskiej cyberprzestrzeń definiowana jest jako „wirtualna przestrzeń, w której krążą elektroniczne dane przetwarzane przez komputery PC z całego świata”[12].

[8] Ustawa z 30 sierpnia 2011 r. o zmianie ustawy o stanie wojennym oraz o kompetencjach Naczelnego Dowódcy Sił Zbrojnych i zasadach jego podległości konstytucyjnym organom Rzeczypospolitej Polskiej oraz niektórych innych ustaw (Dz.U. z 2011 r. Nr 222, poz. 1323).

[9] Ustawa z 17 lutego 2005 r. o informatyzacji działalności podmiotów realizujących zadania publiczne (tekst jedn. Dz.U. z 2019 r., poz. 700).

[10] Ustawa z 16 lipca 2004 r. Prawo telekomunikacyjne (t.j. Dz.U. z 2018 r., poz. 1954, ze zm.).

[11] *Dictionary of Military and Associated Terms*, Joint Publication 1-02, DoD, listopad 2010, s. 63.

[12] J. Wasilewski, *Zarys definicji cyberprzestrzeni*, „Przegląd Bezpieczeństwa Wewnętrznego” 2013, nr 9, s. 229.

Podobnie jak w definicji amerykańskiej sfera użytkownika została także całkowicie pominięta. Według definicji opracowanej na potrzeby NATO w Centrum Doskonalenia Cyberobrony w Tallinie cyberprzestrzeń jest „zależnym od czasu zbiorem połączonych systemów informacyjnych oraz ludzi/użytkowników wchodzących w interakcje z tymi systemami”[13]. W tym ujęciu użytkownik jest ważnym elementem cyberprzestrzeni.

Próbując rozstrzygnąć, czym w takim razie jest cyberprzestrzeń i jakie są jej relacje w stosunku do sieci Internet, należałoby uwzględnić aspekty z obszaru ontologii i teorii systemów. Ale nie o tym tutaj mowa. Najogólniej rzecz ujmując, cyberprzestrzeń stała się środowiskiem umożliwiającym zarówno jednostkom, jak i całym społeczeństwom tworzenie nowych form relacji, kooperacji i funkcjonowania. Należy podkreślić, że główną domeną cyberprzestrzeni, jej istotą, jest informacja. Bez informacji cyberprzestrzeń nie istnieje. A zatem możemy zdefiniować cyberprzestrzeń jako przestrzeń kooperacji międzyludzkich z wykorzystaniem urządzeń elektronicznych do wytwarzania, przechowywania, przekazywania i przetwarzania informacji. Dodatkowo należy podkreślić, że kooperacje te mogę odbywać się w trzech głównych relacjach: człowiek – człowiek, człowiek – cyberprzestrzeń i cyberprzestrzeń – człowiek. W pierwszym przypadku cyberprzestrzeń jest środowiskiem informacyjnym, umożliwiającym wymianę informacji pomiędzy jej użytkownikami. W drugim przypadku cyberprzestrzeń jest środowiskiem, w którym człowiek tworzy swój własny „wirtualny” świat – „wirtualną rzeczywistość”. Trzeci przypadek związany jest z wykorzystaniem sztucznej inteligencji jako elementu cyberprzestrzeni zdolnego do samodzielnego rozwiązywania problemów postawionych przez użytkowników. Powstanie czwartej relacji, cyberprzestrzeń – cyberprzestrzeń, oznaczać będzie koniec świata jaki znamy, ponieważ wiąże się to z usamodzielnieniem i uniezależnieniem sztucznej inteligencji w cyberprzestrzeni. Człowiek w tym przypadku będzie elementem zbędnym. Zidentyfikowanie powyższych realizacji jest niezwykle istotne z punktu widzenia informacyjnego oddziaływania na użytkowników cyberprzestrzeni. Takie narzędzia (sposoby, mechanizmy) jak *trolling*, *boty* czy też *astroturfing* są dzisiaj powszechnie i skutecznie wykorzystywane w kampaniach informacyjnych. Ponadto jeżeli w przedstawionej powyżej definicji cyberprzestrzeni zamienimy słowo „międzyludzkich” na „podmiotów bezpieczeństwa” i wpiszemy jako cel takiej działalności „realizację własnych interesów”, to możemy zauważyć, że cyberprzestrzeń staje się polem rywalizacji. Kooperacje mogą być bowiem i pozytywne,

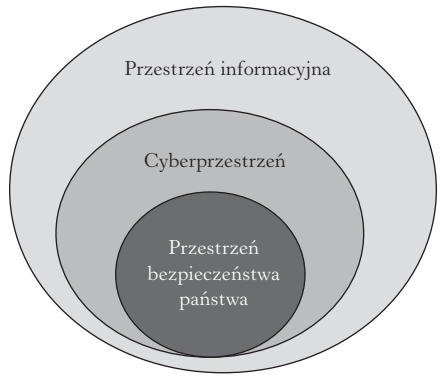
[13] R. Ottis, P. Lorents, *Cyberspace: Definition and Implications,* Tallin 2010.

i negatywne. Pierwsze ukierunkowane są na rozwój, te drugie na walkę. Mając na uwadze powyższe, możemy zdefiniować cyberprzestrzeń jako przestrzeń kooperacji podmiotów bezpieczeństwa z wykorzystaniem urządzeń elektronicznych w celu realizacji własnych interesów.

W zaproponowanej definicji użyto określenia „urządzenia elektroniczne" rozumiane jako elementy infrastruktury teleinformatycznej tworzącej środowisko wymiany i przetwarzania informacji. Środowiskiem tym jest sieć Internet i inne sieci teleinformatyczne wykorzystywane do przesyłania, przetwarzania i przechowywania informacji.

Biorąc pod rozwagę znaczenie przestrzeni informacyjnej i cyberprzestrzeni dla funkcjonowania współczesnych społeczeństw, należy skupić się na ich wzajemnych relacjach z następnym ważnym elementem, a mianowicie przestrzenią bezpieczeństwa. Definiowana jest ona jako „niejednorodny zbiór otwarty złożony z podprzestrzeni umożliwiających osiągnięcie zakładanych celów działań, dla których owa przestrzeń została stworzona"[14]. W dobie rozwoju technologii teleinformatycznych i coraz większego zasięgu sieci teleinformatycznych niezwykle trudno jest ustalić, jakie są wzajemne relacje przestrzeni informacyjnej, cyberprzestrzeni i przestrzeni bezpieczeństwa dla podmiotu, jakim jest państwo. Czy przestrzeń informacyjna jest elementem nadrzędnym w stosunku do przestrzeni bezpieczeństwa, czy też relacje te są odwrotne? Zasadna wydaje się odpowiedź, że przestrzeń bezpieczeństwa państwa w obszarze informacyjnym jest elementem składowym cyberprzestrzeni, która z kolei wchodzi w skład przestrzeni informacyjnej, co przedstawiono na rysunku 3.

Rysunek 3. Wzajemne relacje przestrzeni bezpieczeństwa, cyberprzestrzeni i przestrzeni informacyjnej



Źródło: opracowanie własne.

[14] R. Kwećka, *Strategia bezpieczeństwa…*, s. 20.

Bezpieczeństwo informacyjne – obszary aktywności

Nie można mówić o propagandzie w oderwaniu od bezpieczeństwa informacyjnego, którego integralnymi elementami są walka informacyjna i jej składowa – dezinformacja. Bezpieczeństwo informacyjne, w odróżnieniu od bezpieczeństwa informacji, jest pojęciem złożonym i dużo trudniejszym do uchwycenia. To drugie bowiem najczęściej ogranicza się do spełnienia trzech atrybutów informacji, takich jak poufność, dostępność i integralność. Osiągane jest w trzech obszarach: organizacyjnym, technicznym i fizycznym poprzez implementację między innymi systemu zarządzania bezpieczeństwem informacji bazującym na przyjętych normach, na przykład ISO 27001. Natomiast bezpieczeństwo informacyjne wykracza w istotny sposób poza ramy obowiązujących norm. Informacja funkcjonująca w przestrzeni informacyjnej stała się narzędziem i środkiem realizacji przyjętych celów działania. W bezpieczeństwie informacyjnym niezwykle istotne jest, oprócz zapewnienia bezpieczeństwa informacji, stosowanie elementów walki informacyjnej z elementami dezinformacji i propagandy. Ich umiejętne wykorzystanie pozwala nie tylko na działanie z pozycji dodatniej w stosunku do potencjalnych adwersarzy, ale także pozwala kreować rzeczywistość stawiającą dowolny podmiot bezpieczeństwa zarówno w korzystnym, jak i niekorzystnym świetle. Co więcej, każdy podmiot bezpieczeństwa musi być równocześnie świadomy, że także on może być (a nawet jest) celem oddziaływania informacyjnego, celem propagandy ukierunkowanej na dyskredytację, zmniejszenie wpływów, poniesienie wymiernych strat. Zdolność skutecznego przeciwstawienia się wrogiej kampanii informacyjnej i identyfikacji tego, kto za nią stoi, staje się newralgicznym elementem funkcjonowania każdego państwa i najważniejszym czynnikiem skutecznego działania. Spłaszczanie problemu tylko i wyłącznie do ochrony zasobów informacyjnych nie gwarantuje w XXI wieku działania z pozycji dodatniej.

Bezpieczeństwo informacyjne to transsektorowy obszar bezpieczeństwa, który odnosi się do środowiska informacyjnego (przestrzeni informacyjnej). Jest to proces, którego celem jest zapewnienie prawidłowego i zarazem bezpiecznego funkcjonowania podmiotu bezpieczeństwa w przestrzeni informacyjnej poprzez panowanie we własnej infosferze w celu zabezpieczenia własnych interesów, w przypadku państwa – interesów narodowych. Realizacja powyższego wymaga: zapewnienia odpowiedniej ochrony posiadanych zasobów informacyjnych, przeciwstawienia się wrogim działaniom dezinformacyjnym i propagandzie oraz prowadzenia aktywnych informacyjnych działań ofensywnych wobec adwersarzy.

Zadania te powinny znaleźć odzwierciedlenie w strategii każdego państwa, a ich implementacja powinna umożliwić stworzenie systemu bezpieczeństwa informacyjnego[15]. Bezpieczeństwo informacyjne występuje w środowisku wewnętrznym, zewnętrznym, militarnym, niemilitarnym, osobowym, społecznym, technologicznym i wielu innych. Dla zapewnienia całościowego bezpieczeństwa informacyjnego ważne jest zidentyfikowanie celów realnych i głównych niewiadomych, które powinny być rozpatrywane z uwzględnieniem obszarów: zagrożeń, wyzwań, szans i ryzyka.

Bezpieczeństwo informacyjne, w ocenie autora, jest najbardziej wrażliwym obszarem bezpieczeństwa zarówno w państwie, jak również na arenie międzynarodowej. Ma wpływ na skuteczność i efektywność funkcjonowania całego systemu bezpieczeństwa tak narodowego, jak i międzynarodowego. Dodatkowo należy zauważyć, że działania podejmowane w obszarze bezpieczeństwa informacyjnego muszą być podejmowane z uwzględnieniem praw człowieka i obywatela, a w szczególności z poszanowaniem prawa do prywatności i wolności słowa, co nie jest takie proste i oczywiste.

Nieodzownym elementem bezpieczeństwa informacyjnego jest walka informacyjna. Pierwotnie walka informacyjna kojarzona była z prowadzeniem wojny. Współcześnie jest ona elementem polityki podmiotów bezpieczeństwa takich jak państwa czy też korporacje, a ich głównym celem jest działanie z pozycji dodatniej w stosunku do konkurenta (przeciwnika). Zdobywanie informacji, ochrona własnych zasobów informacyjnych oraz prowadzenie kampanii informacyjnych (z elementami propagandy i dezinformacji, ang. *denial and deception*) będą odgrywały coraz większą rolę w otaczającym nas świecie.

Termin walka informacyjna po raz pierwszy został użyty dopiero w latach 90. XX wieku. Amerykański ekspert Winn Schwartau w książce *Information Warfare*[16] zdefiniował walkę informacyjną jako „działania ukierunkowane na ochronę, uszkodzenie, wykorzystanie, zniszczenie informacji, jak i zasobów informacji, ale i również zaprzeczenie informacjom po to, aby osiągnąć swój cel, korzyści lub nawet zwycięstwo nad przeciwnikiem"[17]. Niemniej jednak najstarsze zapisy, które odnoszą się do elementów walki informacyjnej, pojawiły się już w VI wieku p.n.e.

---

[15] Z. Nowakowski, H. Szafran, R. Szafran, *Bezpieczeństwo w XXI wieku. Strategia bezpieczeństwa narodowego Polski i wybranych państw*, Rzeszów 2009, s. 67.

[16] W. Schwartau, *Information Warfare: Chaos on the Electronic Superhighway*, Nowy Jork 1994.

[17] P. Bączek, *Zagrożenia informacyjne a bezpieczeństwo państwa polskiego*, Toruń 2006, s. 132.

w książce *Sztuka wojny* Sun Tzu. Można w niej znaleźć zapis: „jeśli wiem, że moje oddziały mogą uderzyć na wroga, lecz nie wiem, czy wróg jest przygotowany do odparcia, to szansa przegranej i wygranej jest jak jeden do jednego. Tak samo, jeśli wiem, że wróg nie jest przygotowany na atak, lecz nie wiem czy moje oddziały są gotowe do uderzenia, szansa zwycięstwa i porażki jest jak jeden do jednego. Jeśli wiem, że moje oddziały mogą uderzyć i wróg nie jest przygotowany na atak, lecz nie rozpoznałem dobrze ułożenia terenu bitwy, szansa zwycięstwa i porażki jest jak jeden do jednego. Dlatego też twierdzę: Poznaj warunki terenu i pogody, wtedy Twoje zwycięstwo będzie całkowite"[18].

Z powyższego można wnioskować, że działania te ukierunkowane były na:

- poznanie wroga, pogody i terenu, czyli dzisiejsze rozpoznanie;
- poznanie siebie, czyli zapewnienie niezakłóconego funkcjonowania własnego systemu informacyjnego;
- wprowadzanie komunikatów zakłócających, czyli dzisiejsza dezinformacja.

Walkę informacyjną określa się także jako wojnę informacyjną (i odwrotnie). Jest to określenie często wykorzystywane przez polityków, którzy w ten sposób chcą wyrazić swoje nastawienie do zaistniałej sytuacji. Zjawisko to jest już tak powszechne, że obecnie trudno określić, co jest wojną informacyjną, a co wojną informacyjną jeszcze nie jest.

Elementem składowym walki informacyjnej jest dezinformacja, definiowana współcześnie jako „wprowadzenie kogoś w błąd przez podanie mylących lub fałszywych informacji"[19]. Termin ten pojawił się po raz pierwszy w literaturze anglojęzycznej w 1926 roku w londyńskim miesięczniku „The Whitehall Gazette & St. James's Review"[20]. Określenia tego użyto w celu opisania działań sowieckich służb specjalnych. Rok później w piśmie „Siegodnia" wspomniano, iż należy ona do głównych działań GPU (Państwowego Zarządu Politycznego przy Ludowym Komisariacie Spraw Wewnętrznych Rosyjskiej Federacyjnej Socjalistycznej Republiki Radzieckiej)[21].

Kolejną interpretacją zjawiska dezinformacji jest definicja podana przez Vladimira Volkoffa, który postrzega ją „w wąskim lub szerszym

---

[18] Sun Tzu, *Sztuka wojny*, przekł. K.A.M., Warszawa 1994, s. 116.

[19] Hasło: *Dezinformacja*, „Słownik języka polskiego", https://sjp.pwn.pl/sjp/dezinformacja;2554971.html (dostęp: 22.12.2018).

[20] *The Embassies and Foreign Affairs*, „The Whitehall Gazette & St. James's Review" 1926, maj, s. 8.

[21] J.J. Dziak, *Chekisty: A History of the KGB*, Lexington 1987, s. 42.

znaczeniu. W wąskim tego słowa znaczeniu mieści się ona w połowie drogi między wprowadzeniem w błąd a wpływaniem. Podczas gdy wprowadzanie w błąd […] jest czynnością jednorazową, związaną z konkretnym zadaniem, dopuszcza pewną amatorszczyznę, wykorzystuje najprzeróżniejsze środki i zmierza do wmówienia pewnych określonych rzeczy określonym osobom, dezinformacja jest prowadzona w sposób systematyczny, fachowy, zawsze za pośrednictwem mass mediów i jest adresowana do opinii publicznej, a nie sztabów krajów – obiektów działań. I analogicznie: podczas gdy wpływanie przejawia się w działaniach pozornie niezorganizowanych, oportunistycznych, głównie ilościowych, dezinformacja stawia sobie za cel realizację konsekwentnego programu zmierzającego do zastąpienia w świadomości, a przede wszystkim podświadomości mas będących przedmiotem tych działań, poglądów uznanych za niekorzystne dla dezinformatora takimi, które uważa on za korzystne dla siebie"[22]. Volkoff wymienił także kilka metod dezinformacji, takich jak: odwrócenie faktów, negacja faktów, mieszanie prawdy i kłamstwa, rozmycie, kamuflaż, interpretacja, generalizacja, ilustracja, nierówna reprezentacja, równa reprezentacja[23]. Techniki te, wykorzystywane we wszelkiego rodzaju manipulacjach informacyjnych oraz kampaniach dezinformacyjnych, nie wyczerpują całości metod i narzędzi wykorzystywanych w walce informacyjnej.

Początków dezinformacji możemy doszukiwać się patrząc daleko wstecz, do najstarszych rozwiniętych cywilizacji czy też początków zorganizowanych konfliktów zbrojnych. Przykładem może być myśl Sun Tzu, który w cytowanej już wcześniej książce *Sztuka wojny* stwierdził, że „wojna jest to wprowadzanie w błąd. Jeśli zatem jesteś zdolny, udawaj mało zdolnego. Gdy porywasz swoje wojska do działania, udawaj bierność. Jeżeli twój cel jest bliski, zachowuj się tak, jakby był odległy. A gdy jest odległy, udawaj, że jest bliski"[24]. W każdej bowiem wojnie czy też konflikcie zbrojnym niezbędnym newralgicznym elementem jest umiejętność i zdolność wprowadzania przeciwnika w błąd, poprzez podawanie nieprawdziwych informacji i tworzenie fałszywej rzeczywistości. Dezinformacja jest zjawiskiem ponadczasowym, bardzo mocno zapisanym w naszej przeszłości, teraźniejszości i zapewne w przyszłości. Można wręcz pokusić się o tezę, że jest naturalnym elementem rozwoju ludzkości.

---

[22] V. Volkoff, *Psychosocjotechnika, dezinformacja – oręż wojny*, przekł. A. Arciuch, Warszawa 1999, s. 8.

[23] Tamże, s. 157–172.

[24] Sun Tzu, *Sztuka wojny*, przekł. K.A.M.,, Warszawa 1994, s. 61.

Dezinformacja może mieć zarówno wymiar strategiczny, jak i taktyczny. Dezinformacja taktyczna trwa zazwyczaj stosunkowo krótko, kilka miesięcy, a jej głównym celem jest wprowadzenie w błąd w jednej lub kilku powiązanych ze sobą kwestiach. Przykładami tego rodzaju dezinformacji mogą być:

- podsunięcie nieprawdziwych danych technicznych nowej broni;
- zmodyfikowanie danych statystycznych celem wywołania wrażenia, iż stan ekonomii danego państwa jest lepszy lub gorszy od rzeczywistego;
- opublikowanie sfabrykowanych materiałów w celu skompromitowania polityka, partii czy też nawet rządu.

Z kolei dezinformacja strategiczna polega na systematycznym przekazywaniu sfabrykowanej informacji oraz fałszywych sygnałów politycznych. Jej głównym celem jest wytworzenie wypaczonego obrazu rzeczywistości powodującego błędną analizę sytuacji. To działanie ukierunkowane na wprowadzenie w błąd przeciwnika co do podstawowych kwestii polityki państwa[25]. W przeciwieństwie do dezinformacji taktycznej trwa ona nawet kilkadziesiąt lat i należy, w głównej mierze, do sfery walki służb specjalnych. Więcej informacji na ten temat można znaleźć w doskonałej książce Roya Godsona i Jamesa J. Wirtza *Strategic Denial and Deception*[26].

Istota i elementy propagandy

Przechodząc do tematu propagandy, można zauważyć, że także ona jest bezpośrednio związana z historią ludzkości. Z pierwszymi przykładami takich działań mieliśmy do czynienia już w kulturze mezoamerykańskiej i egipskiej, gdzie zapisy hieroglificzne składające się głównie z symboli i obrazów przedstawiały historię w sposób korzystny dla władzy. Było to możliwe dzięki temu, że jedynie władcy i ich kapłani posiadali umiejętność tworzenia hieroglifów, przez co perswazja była jednokierunkowa, skierowana od władcy do mas. Pierwsze zapisy, w których użyto określenia propaganda, odnotowano w 1622 roku, wraz z powołaniem przez Papieża Grzegorza XV Kongregacji Propagandy Wiary (Congregatio de Propaganda Fide). Do powszechnego obiegu pojęcie to weszło

[25] S. Lewczenko, *Private Channel*, [w:] *Influence: A KGB Disinformation Tool*, „Counterpoint: A Monthly Report on Soviet Active Measures" 1988, t. 3, nr 11, s. 1.

[26] R. Godston, J.J. Wirtz, *Strategic Denial and Deception. The Twenty-First Century Challenge*, Waszyngton 2012.

podczas I wojny światowej jako nowa technika perswazji, oddziaływania na społeczeństwo[27]. Propagandę postrzegano wtedy jako rozprzestrzenianie stronniczych idei i poglądów, w praktyce często używając przy tym kłamstwa, manipulacji i podstępu. Największy rozwój propagandy związany był z systemami totalitarnymi: faszyzmem i komunizmem. Od tego czasu obejmuje ona również sugestię i wywieranie wpływu przy użyciu manipulacji symbolami i mechanizmami psychologicznymi. Propaganda to także zręczne posługiwanie się obrazami i sloganami, z odwoływaniem się przy tym do uprzedzeń czy emocji. Jest to pewnego rodzaju komunikat stworzony na potrzeby zdefiniowanego (określonego) punktu widzenia – celem jest skłonienie odbiorcy tego komunikatu do dobrowolnego przyjęcia zawartego w nim punktu widzenia i uznania go za swój[28]. Współcześnie propaganda definiowana jest jako „proces składający się z planowanego użycia każdej formy publicznych lub masowo wytwarzanych komunikatów zaprojektowanych tak, aby wpływały na umysły i emocje wybranej grupy odbiorców, w z góry określonym celu (społecznym, militarnym, ekonomicznym, politycznym)"[29]. Kluczowymi elementami propagandy są przyjęty plan działania i cele, jakie ten plan realizuje. Bez planu i celowości działania nie można mówić o propagandzie, lecz jedynie o kłamstwie lub manipulacji. Nie każde kłamstwo i manipulacja stosowane w życiu publicznym są propagandą, ale każda kampania propagandowa jest oparta na kłamstwach i manipulacji.

Mówiąc o schemacie, według którego realizowana jest kampania propagandowa, należy odnieść się do zapisów przedstawionych przez Paula M.A. Linebargera w książce *Psychological Warfare*. Wymienia on pięć najważniejszych etapów propagandy, określanych łącznie jako STASM od pierwszych liter ich angielskich nazw, a mianowicie[30]:

- S – *source* (źródła – kanały informacyjne propagandy);
- T – *time* (czas rozpoczęcia i prowadzenia kampanii propagandowej);
- A – *audience* (publiczność – odbiorcy propagandy);
- S – *subject* (temat – sprawa, której dotyka kampania propagandowa);
- M – *mission* (misja – cel kampanii propagandowej powiązany z celem politycznym).

---

[27] E. Aronson, A. Pratkanis, *Wiek propagandy*, przekł. J. Radzicki, M. Szuster, Warszawa 2004, s. 17.

[28] A. Chorobiński, *Walka informacyjna jako fundamentalny składnik działalności terrorystycznej w przyszłości*, s. 4, http://konkursy.byd.pl/userfiles/files/chorobinski.pdf (dostęp: 22.12.2018).

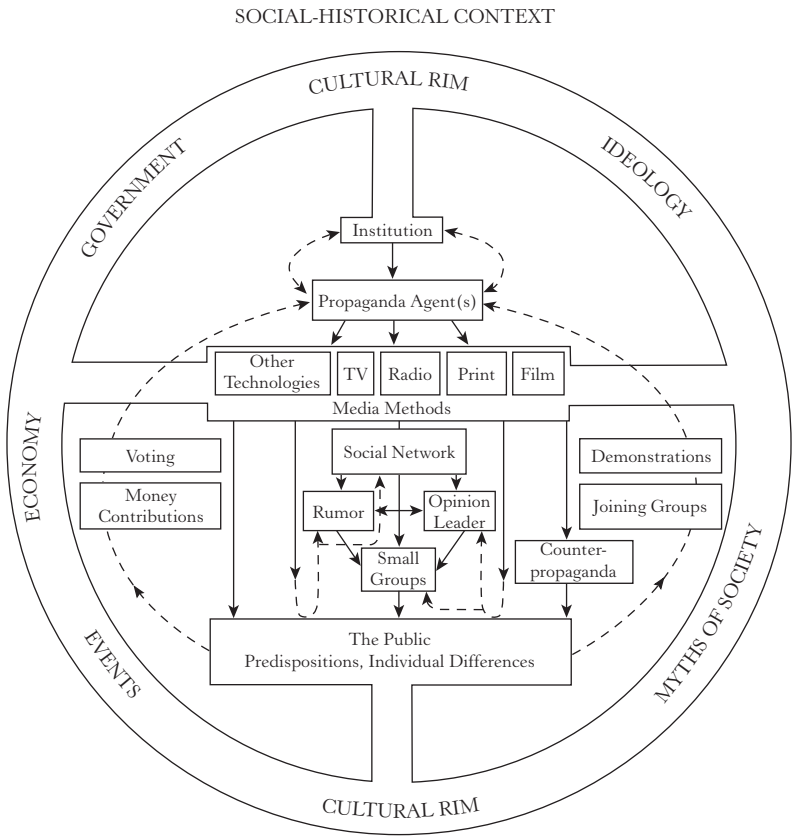
[29] P.M.A. Linebarger, *Psychological Warfare*, Waszyngton 1954 (1972), s. 61.

[30] Tamże, s. 67.

Pierwszy element dotyczy źródeł informacji (kanałów informacyjnych), przez które będzie prowadzona kampania propagandowa. Współcześnie są to: prasa, telewizja, ulotki, Internet, wiece, bezpośrednie spotkania, fałszywe autorytety, agenci wpływu itp. „Czas" to nic innego jak określenie, kiedy i jak długo kampania będzie prowadzona. Określenie czasu jest bardzo istotne, ponieważ łatwiej jest wpływać na społeczeństwo będące w kryzysie, stojące przed wyzwaniami lub wyborami. Dobór czasu jest uwarunkowany w głównej mierze percepcją grupy docelowej i budowaniem w niej odpowiednich przekonań i poglądów (psychologia tłumów). Publiczność – grupa docelowa kampanii – charakteryzuje się swoimi systemami wartości, przekonaniami, nawykami, środowiskiem informacyjnym, z którego pozyskuje wiedzę na temat otaczającego świata. Inny przekaz będzie kierowany do ludzi wykształconych, świadomie korzystających ze sprawdzonych źródeł informacji, inny z kolei do grupy osób budujących swoje spojrzenie na świat na podstawie na przykład mediów społecznościowych czy też konkretnych (ulubionych) kanałów informacyjnych. Temat propagandy powinien być bezpośrednio związany z oczekiwaniami wybranej grupy docelowej, niejako rozwiązujący stojące przed nią problemy. Ostatni aspekt to cel (misja), dla którego prowadzone są działania propagandowe. Jest to podstawowy element, który musi zostać określony zaraz na początku planowania kampanii propagandowej. Oczywiście nie jest to cel sam w sobie – jest ściśle powiązany z celem politycznym, militarnym, ekonomicznym, marketingowym, społecznym.

Patrząc przez pryzmat współczesnego społeczeństwa i środowiska informacyjnego, w którym ono funkcjonuje, można zauważyć, że proces propagandy uwarunkowany jest wieloma czynnikami, takimi jak: działalność propagandystów, media społecznego przekazu, sieci i media społecznościowe, organizacje rządowe i pozarządowe, grupy społeczne, systemy wartości, kultura i tradycje, systemy ekonomiczne i polityczne oraz wiele, wiele innych. Komunikaty wysyłane w propagandzie funkcjonują w szeroko pojmowanej obręczy kulturowej (ang. *cultural rim*), która z kolei uwarunkowana jest kontekstem społeczno-historycznym (ang. *social-historical context*). Z tego powodu nie można rozpatrywać propagandy w oderwaniu od historycznych ram kulturowych. Każde społeczeństwo funkcjonuje bazując na innych systemach wartości, uwarunkowanych kulturą, doświadczeniami, przekonaniami i historią. Propaganda uwzględniająca powyższe aspekty ma wpływ na kształt kultury, ale także kultura ma wpływ na kształt propagandy. Na rysunku 4 przedstawiono model współczesnej propagandy.

Rysunek 4. Model współczesnej propagandy



Źródło: G.S. Jowert, V. O'Donnell, *Propaganda and Persuasion*, Waszyngton 2006, s. 361.

Na rysunku 4 uwidocznione zostały ścisłe powiązanie procesu propagandy ze społeczeństwem, w którym jest ona realizowana, a zwłaszcza z jego społeczno-historycznymi i kulturowymi uwarunkowaniami. Dlatego propagandyści, chcąc być skuteczni, muszą uwzględniać wszystkie aspekty dziedzictwa narodowego społeczeństwa, mające odzwierciedlenie w stylu, sposobie i narracji przekazu. Tym samym także identyfikowanie kampanii propagandowych wymaga uwzględnienia kontekstu społeczno-historycznego i kulturowego społeczeństwa, w którym była ona prowadzona. Patrząc przez pryzmat ostatnich wydarzeń w naszej części Europy, zauważamy, że w rosyjskiej propagandzie inna narracja wykorzystywana jest w kampanii wymierzonej w realizację celów politycznych na Ukrainie, a inna w Polsce. Adresatem przekazu jest nie tylko społeczeństwo danego kraju, ale także

społeczeństwo rosyjskie, rosyjskojęzyczne i międzynarodowe. W przypadku Ukrainy głównym tematem narracji jest nacjonalizm ukraiński i porównywanie Ukraińców do faszystów i banderowców, wsparte dodatkowo przekazem o upadku korupcyjnego państwa, w którym oligarchowie i władza kradną, a reszta społeczeństwa klepie biedę. W przypadku Polski narracja ukierunkowana jest na rusofobię, antysemityzm, niewdzięczność za wyzwolenie, nacjonalizm Polaków oraz zgubny wpływ Unii Europejskiej na tożsamość i niezależność narodową. Ponadto w przypadku Polski, w myśl zasady dziel i rządź, w propagandzie wykorzystywany jest temat katastrofy smoleńskiej, którego najistotniejszym elementem jest wrak polskiego samolotu, polaryzujący polskie społeczeństwo. Faktem jest to, że dopóki wrak samolotu będzie dzielił polskie społeczeństwo, dopóty Rosja nie przekaże go Polakom. W tym przypadku przekazy propagandowe wspierane są dodatkowo agenturą wpływu mającą swoje korzenie we wspólnej rosyjsko-polskiej historii. W czasach zaborów, okupacji i PRL Rosjanie budowali w polskim społeczeństwie siatkę agentów i donosicieli, zarówno wśród opozycji jak i w obozie rządzącym. Dzisiaj ich teczki znajdują się w Moskwie i są podstawą budowania silniej agentury wpływu. Oczywiście nie ma na to jednoznacznych dowodów, ale które z państw zrezygnowałoby z agentury wpływu, mając w swoich archiwach narzędzia oddziaływania i szantażu? (Tak na marginesie, w przypadku Polski osobnym zagadnieniem wymagającym zbadania jest niemiecka agentura wpływu, budowana na bazie agentury III Rzeszy i Stasi).

Kończąc powyższe rozważania, należy jeszcze raz podkreślić, że propaganda będzie wtedy skuteczna, jeżeli jest zaplanowana długofalowo, a u podstaw jej tworzenia zostaną uwzględnione uwarunkowania społeczno-historyczne i kulturowe danego społeczeństwa. Najlepszym przykładem takiego oddziaływania jest zakorzenienie w społeczeństwach zachodnich określenia „polski obóz śmierci". I chociaż jest to oczywiste kłamstwo historyczne, to jest w dalszym ciągu wykorzystywane przeciwko Polsce i Polakom. Szczytem zarówno historycznej ignorancji, jak i skuteczności antypolskiej propagandy była wypowiedź prezydenta Baracka Obamy z 2012 roku, kiedy użył określenia *polish death camps*[31].

[31] *Obama Angers Poles with „Death Camp" Remark*, BBC News, 30.05.2012, https://www.bbc.com/news/world-europe-18264036 (dostęp: 22.12.2018); *Obama's Words at Medal Ceremony Cause Trouble with Poland*, CNN Wire Staff, 30.05.2012, https://edition.cnn.com/2012/05/30/politics/obama-death-camps/index.html (dostęp: 22.12.2013); M. Landler, *Polish Premier Denounces Obama for Referring to a „Polish Death Camp"*, „The New York Times", 30.05.2012, https://www.nytimes.com/2012/05/31/world/europe/poland-bristles-as-obama-says-polish-death-camps.html (dostęp: 22.12.2018).

Kolejnym obszarem wymagającym zidentyfikowania w propagandzie jest jej źródło i stopień prawdziwości czy też zakłamania, co pozwala wyróżnić trzy podstawowe rodzaje propagandy[32]: białą (ang. *overt or white propaganda*), szarą (ang. *gray propaganda*) i czarną (ang. *covert or black propaganda*).

Pierwsza z nich, propaganda biała, to działania, w których źródło i pochodzenie informacji są znane odbiorcy. Jest to nic innego jak przekazanie oficjalnego stanowiska rządu, agencji rządowej czy też organizacji w sposób jawny, z tym że przekazywane w nich informacje są podawane w sposób wybiórczy, z uwypukleniem tych sprzyjających stronie przekazującej i jednocześnie z pomijaniem faktów niewygodnych dla niej.

Szara propaganda cechuje się tym, że źródło informacji nie jest do końca znane, jedynie można się domyślać jego pochodzenia, przy tym przekazywana w niej informacja może być bądź korzystna, bądź niekorzystna dla określonej grupy odbiorców. Informacje podawane są w taki sposób, aby nosiły one zmamiona przecieku informacyjnego, nieoficjalnego stanowiska rządu czy też organizacji.

Ostatni rodzaj – czarna propaganda – jest z założenia budowana na kłamstwie i fałszu. Zarówno źródło informacji, jak i sama informacja są fałszywe. Celem tego typu działań jest uzyskanie w grupie odbiorców odpowiedniego efektu psychologicznego polegającego na całkowitej zmianie postrzegania bieżącej sytuacji politycznej, społecznej, ekonomicznej, militarnej itp. W dobie cyberprzestrzeni czarna propaganda jest narzędziem powszechnie wykorzystywanym, chociażby w czasie wyborów, referendów czy w innych ważnych dla danego społeczeństwa sytuacjach.

Podobnie jak w przypadku dezinformacji, także w przypadku propagandy można wyróżnić jej dwie zasadnicze kategorie, a mianowicie propagandę taktyczną i propagandę strategiczną[33]. Są one nierozerwalnie związane z czasem ich prowadzenia i przyjętym celem oddziaływania. Najczęściej formy te wykorzystywane są w trakcie operacji militarnych, ale ostatnio także w elementach tak zwanej wojny hybrydowej. Propaganda strategiczna ukierunkowana jest na realizację celów strategicznych, rozłożonych w długim przedziale czasu, często sięgającym nawet lat, a jej głównym celem jest zbudowanie nowej świadomości (nowego postrzegania bieżącej sytuacji) wśród grupy docelowej. Propaganda taktyczna skierowana jest bezpośrednio do wybranej grupy odbiorców, na przykład żołnierzy biorących udział w danej operacji, w celu wywarcia na nich

[32] G.S. Jowert, V. O'Donnell, *Propaganda and Persuasion*, Waszyngton 2006, s. 17.

[33] P.M.A. Linebarger, *Psychological…*, s. 69.

odpowiedniego wrażenia, przekonania do własnej racji, zdemotywowania do działania i poświęceń. Jej czas oddziaływania związany jest bezpośrednio z czasem realizacji danej operacji militarnej.

W literaturze przedmiotu można zidentyfikować także inne jej rodzaje i formy. I tak, w zależności od rodzaju aktywności stron prowadzących kampanie propagandowe, wyróżniamy[34] propagandę defensywną (ang. *defensive propaganda*) i propagandę ofensywną (ang. *offensive propaganda*). Pierwsza z nich wykorzystywana jest do wzmocnienia pozytywnego nastawienia społeczeństwa i akceptacji w stosunku do planów i przedsięwzięć podejmowanych, realizowanych przez państwo (rząd, organizację). Drugi rodzaj propagandy wykorzystywany jest w celu przerwania i zmiany negatywnego nastawienia społeczeństwa, niezgodnego z celami propagandystów.

Z punktu widzenia celu oddziaływania propagandy na społeczeństwo wyróżniane są następujące jej rodzaje[35]:

- propaganda konwersyjna (ang. *conversionary propaganda*) – jej głównym celem jest zmiana świadomości, poglądów i lojalności jednostki w stosunku do danej grupy społecznej i ich przekierowania na inną, pożądaną grupę społeczną;
- propaganda dzieląca (ang. *divisive propaganda*) – ukierunkowana jest na rozbicie wybranej społeczności (grupy), tak aby powstały w niej podziały (mniejsze grupy), a tym samym zmniejszyła się jej siła oddziaływania;
- propaganda scalająca (ang. *consolidation propaganda*) – charakteryzują ją odmienne cele, jej głównym zadaniem jest likwidacja podziałów i zjednoczenie społeczeństwa wokół wspólnej sprawy, wspólnego celu;
- przeciwpropaganda (ang. *counter propaganda*) – jest w swej istocie nastawiona na osłabienie lub też całkowite uniemożliwienie realizacji celów propagandy stosowanej przez przeciwnika zarówno militarnego, politycznego czy gospodarczego.

Następnym istotnym elementem wymagającym zidentyfikowania są techniki, przy pomocy których propaganda jest realizowana. Do najważniejszych z nich można zaliczyć[36]: zapewnienie (ang. *assertion*), owczy pęd (ang. *bandwagon*), naciąganie faktów (ang. *card stacking*), błyskotki (ang. *glittering generalities*), mniejsze zło (ang. *lesser of two evils*), docze-

[34] Tamże, s. 70.

[35] Tamże.

[36] Na podstawie materiałów uzyskanych przez autora w ramach przedmiotu *Foreign Propaganda, Perceptions, and Policy*, realizowanego w semestrze jesiennym 2017 r. w The Institute of World Politics, Washington D.C.

pianie ogólników (ang. *name calling*), wskazywanie wroga (ang. *pinpointing the enemy*), ludowość (ang. *plain folks*), uproszczenie (ang. *simplification*), referencja (ang. *testimonials*), przenoszenie (ang. *transfer*).

Zapewnienie (ang. *assertion*) jest współczesną techniką propagandy powszechnie stosowaną w reklamie. Przekazywane informacje mają entuzjastyczny, radosny, energiczny charakter, a treści w nich zawarte, niekoniecznie prawdziwe, odwołują się do przekonań i nastawienia odbiorców. Coś jest z natury dobre, bo jest, bez żadnego udawadniania i odwoływania się do faktów, bez zadawania zbędnych pytań. Za każdym razem, gdy reklamodawca stwierdzi, że jego produkt jest najlepszy, bez dostarczania naukowych dowodów, korzysta z techniki zapewnienia. Odbiorca informacji zgadza się z otrzymywanym przekazem bezrefleksyjnie, nie dążąc do zidentyfikowania, jak jest naprawdę. Zapewnienie, choć dosyć łatwe do wykrycia, jest coraz częściej wykorzystywane w polityce do sterowania emocjami społecznymi, a narracja takiego przekazu jest, z natury rzeczy, budowana na fałszu i kłamstwie[37].

Owczy pęd (ang. *bandwagon*) jest jedną z najpopularniejszych technik propagandy stosowaną zarówno w czasie wojny, jak i pokoju i odgrywa bardzo ważną rolę we współczesnej inżynierii społecznej. Polega na apelu do odbiorcy o podążanie za tłumem, za większością, ponieważ inni też to robią i będzie to tylko i wyłącznie z korzyścią dla niego. W tym przypadku propagandysta zasadniczo stara się przekonać odbiorcę, że tylko jedna ze stron, ta właściwa, jest zwycięską, ponieważ dołączyło do niej więcej osób i z tego też względu tylko oni mają rację. Odbiorcy mają wierzyć, że skoro tak wielu ludzi dołączyło, to zwycięstwo jest nieuniknione, a porażka niemożliwa. Przeciętna osoba zawsze chce być po zwycięskiej stronie, więc jest zmuszona presją tłumu do przyłączenia się. Przeciwdziałanie tej technice propagandy polega w głównej mierze na szczegółowym rozważeniu plusów i minusów przyłączania się do sugerowanej idei bez uwzględniania liczby osób, które już się do niej dołączyły. Odbiorca powinien poszukiwać więcej informacji, uzyskanych z wielu niezależnych źródeł[38].

Technika naciągania faktów (ang. *card stacking*) polega na przedstawianiu informacji jedynie pozytywnych, sprzyjających propagandystom, z pominięciem informacji dla nich negatywnych. Ze względu na swoją wysoką skuteczność jest ona powszechnie stosowana w życiu społecznym

[37] H.T. Conserva, *Propaganda Techniques*, Bloomington 2003, s. 25–27.

[38] W. Garber *Propaganda Analysis – to What Ends*, „The American Journal of Sociology" 1942, t. 48, nr 2, s. 240.

i politycznym. Chociaż większość informacji przedstawionych przez propagandystę jest prawdziwa, to jednak pomija się w przekazie to, co jest istotne, ale równocześnie niekorzystne dla propagandysty. Najlepszym sposobem radzenia sobie z naciąganiem faktów jest zdobycie większej liczby informacji[39].

Błyskotki lub też błyszczące ogólniki (ang. *glittering generalities*) to technika wykorzystywana najczęściej w polityce i politycznej propagandzie polegająca na używaniu słów mających pozytywne znaczenie dla poszczególnych odbiorców przekazu propagandowego i powiązanych bezpośrednio z wysoko cenionymi wartościami. Użyte słownictwo jest przyjmowane (akceptowane) bezrefleksyjnie, automatycznie, bez zastanowienia, tylko i wyłącznie dlatego, że dotyczy ważnego dla odbiorcy pojęcia. Słowa często używane w technice błyskotek to: honor, ojczyzna, bóg, chwała, miłość do kraju, demokracja, wolność, niepodległość, suwerenność. Technika ta nie jest zła sama w sobie, a jej głównym celem jest jednoczenie społeczeństwa wokół wspólnych wartości. Niemniej jednak może być nadużywana przez polityków na potrzeby walki politycznej i wygrania wyborów[40].

Technika mniejszego zła (ang. *lesser of two evils*) polega na przekonaniu odbiorców, że spośród wielu złych rozwiązań jedno jest najlepsze i powinno być zaakceptowane przez społeczeństwo. Stosowana jest najczęściej w czasie wojny, w celu przekonania obywateli o potrzebie poświęcenia i poniesienia niezbędnych dla zwycięstwa ofiar. W czasie pokoju jej przekaz nie jest już taki jednoznaczny, a główny nacisk ukierunkowany jest na przekonanie społeczeństwa do przeprowadzenia niezbędnych kosztownych społecznie reform, na przykład reformy emerytalnej. Ponadto często winą za wszelkie trudy i wyrzeczenia obarczany jest wrogi kraj (w czasie wojny) lub wroga grupa polityczna (w czasie pokoju). Prezentowana idea jest zwykle przedstawiana jako jedyna opcja lub jako rozwiązanie, które należy wybrać dlatego, że każde inne jest gorsze. W konfrontacji z tą techniką podmiot, na który jest ona ukierunkowana, powinien brać pod rozwagę wszystkie możliwe rozwiązania stojących przed nim problemów, a nie tylko zaproponowane w przekazie propagandowym.

Doczepianie ogólników (ang. *name calling*) występuje często w scenariuszach politycznych i wojennych, ale bardzo rzadko w reklamie. Jest to użycie obraźliwego języka lub słów, które niosą negatywną konotację

[39] Tamże, s. 244.

[40] Tamże, s. 242.

przy opisie przeciwnika politycznego czy też militarnego. Propagandyści usiłują wzbudzić wśród opinii publicznej uprzedzenie do wybranej grupy, nadając przeciwnikowi taką nazwę, która źle się kojarzy społeczeństwu lub której nie lubi. Często nazwę taką nadaje się za pomocą sarkazmu i ośmieszania, stosując określenia typu złodzieje, komuniści, faszyści, zdrajcy. Identyfikacja tej techniki polega na oddzieleniu emocji od odbieranego przekazu informacyjnego i poszukiwaniu merytorycznych przesłanek.

Wskazywanie wroga (ang. *pinpointing the enemy*) jest techniką wykorzystywaną zarówno w czasie wojny, jak również w kampaniach politycznych i debatach społecznych. Jest to działanie ukierunkowane na przedstawienie jednej konkretnej osoby lub grupy jako wroga społeczeństwa, narodu. Identyfikacja tej techniki propagandy polega na postrzeganiu przekazu informacyjnego i kierowanych w nim oskarżeń w kategoriach jednoznacznego dobra i zła. Do tego niezbędna jest analiza wszystkich dostępnych informacji na ten temat i ich odniesienia do uniwersalnego systemu wartości. Społeczeństwo dobrze wyedukowane i świadome otaczającej rzeczywistości, mechanizmów zachodzących w jego otoczeniu, jest o wiele mniej podatne na tego rodzaju propagandę.

Ludowość (ang. *plain folks*) to technika propagandowa, która ma na celu przekonać opinię publiczną, że poglądy propagandysty odzwierciedlają poglądy zwykłego, prostego ludu i z tego też względu działają również na korzyść prostych ludzi. Propagandysta często stosuje akcent i charakterystyczne zwroty określonej grupy odbiorców, wspierając przekaz żartami, gestami i zachowaniami. Dodatkowo podczas wystąpień język użyty przez propagandystę tworzy iluzję niedoskonałości mówcy poprzez błędy wymowy, jąkanie się i ograniczone słownictwo, co sugeruje jego szczerość, spontaniczność i swojskość. Technika ta występuje często w połączeniu z techniką błyskotek, przez co wzmacniany jest przekaz, a główny nacisk kładziony jest na wspólne wartości. W zderzeniu z tą techniką propagandy najlepszym sposobem jej przeciwdziałania jest oddzielenie idei prezentowanych przez propagandystę od jego osobowości i wyuczonych technik wystąpień publicznych.

Uproszczenie (ang. *simplification*) jest w swej istocie zbieżne z techniką wskazywania wroga, ponieważ często redukuje złożoną sytuację do wyraźnego wyboru pomiędzy dobrem a złem. Technika ta jest wykorzystywana w manipulowaniu niewykształconymi odbiorcami, postrzegającymi świat przez pryzmat ulubionych kanałów informacyjnych i powielającymi otrzymane przekazy informacyjne w bezrefleksyjny sposób. Jest taka rzeczywistość a nie inna, bo tak mówili w telewizji lub tak wyczytałem

w Internecie. W walce z uproszczeniem niezbędne jest zbadanie innych występujących czynników i przekazów, czyli, podobnie jak w przypadku wszystkich innych technik propagandy, uzyskanie dodatkowych informacji z różnych kanałów komunikacyjnych[41].

Referencja (ang. *testionials*) to technika wykorzystywana często w trakcie kampanii reklamowych i politycznych polegająca na odniesieniu się, w kontekście propagandy lub poza jej kontekstem, do cytatów lub rekomendacji, które próbują połączyć znaną lub szanowaną osobę z produktem, przedmiotem lub ideą. To nic innego jak odniesienie się do autorytetów (lub pseudoautorytetów, choćby celebrytów, a także fałszywych autorytetów będących na przykład agentami wpływu lub przywódcami partii politycznych) i przekonanie odbiorców, że jeżeli „autorytet" ma takie poglądy, to oni powinni mieć takie same, bez względu na fakty, koszty i okoliczności. Nie trzeba nic uzasadniać, autorytet przecież wie lepiej, co jest dobre, a co złe. Przeciwdziałanie tej technice propagandy polega głównie na obiektywnej ocenie wszystkich wad i zalet przedmiotu, oferty czy idei i wyciągnięciu bezstronnych wniosków, niepodyktowanych żadnymi subiektywnymi względami lub zewnętrznymi naciskami.

Ostatnią z wymienionych technik propagandy, stosowaną zarówno w polityce jak i podczas wojny, jest przenoszenie (ang. *transfer*). Polega na próbie stworzenia i przeniesienia autorytetu z jednej rzeczy czy też osoby (grupy społecznej) na inną rzecz czy osobę (grupę społeczną) w celu zdobycia przychylności. Istotnego znaczenia nabiera wykorzystanie symboli, które niosą za sobą konkretne znaczenie. Stąd też politycy chętnie robią sobie zdjęcia na tle flagi narodowej, zyskując poparcie w propatriotycznej części społeczeństwa. W innym przypadku popierają palenie flag instytucji międzynarodowych, na przykład Unii Europejskiej, zyskując poparcie wśród eurosceptyków i narodowców. To także robienie zdjęć z osobami będącymi autorytetami, na przykład z Lechem Wałęsą w trakcie wyborów w 1989 roku czy też z papieżem. Technika transferu jest także wykorzystywana w budowaniu negatywnego postrzegania organizacji poprzez doklejanie im łatek związanych z ideologią, która w danym społeczeństwie kojarzy się w sposób jednoznacznie negatywny, jak faszyzm czy komunizm.

Następnym elementem związanym z propagandą jest sposób jej identyfikowania. Jest to proces złożony, wymagający wnikliwych i długotrwałych badań historycznych, ponieważ tylko w ten sposób można zidentyfikować jej kontekst i prawdziwy cel zastosowania. Podejście to

[41] H.T. Conserva, *Propaganda*…, s. 74–76.

wymaga badania wszystkich przekazów propagandowych, we wszystkich kanałach komunikacyjnych, reakcji jej odbiorców, zastosowanych technik i narzędzi, w długim przedziale czasu. Badanie efektów propagandy w krótkim przedziale czasu może być zasadne tylko w analizie propagandy taktycznej, wykorzystywanej na potrzeby doraźnych, taktycznych celów, na przykład kampanii wyborczej. Propaganda bazuje najczęściej na zakorzenionych w danej społeczności i kulturze tradycjach, mitach i stereotypach, co powoduje, że odróżnienie tego, co jest propagandą od tego, co nią nie jest, stanowi bardzo trudne zadanie, tym bardziej w krótkim przedziale czasu. Jak wspomniano wcześniej, propaganda jest działalnością celową, często prowadzoną przez lata i poszukując tego celu należy ukierunkować badania na długi przedział czasowy.

W cytowanej książce *Propaganda and Persuasion* przedstawiono 10-etapową procedurę analizy najważniejszych elementów propagandy. Jak podkreślają autorzy, jej wykorzystanie w analizie aktualnie prowadzonej propagandy jest złożone ze względu na nie do końca zidentyfikowany cel jej prowadzenia przy jednoczesnym braku danych do oceny jej skuteczności. Z drugiej jednak strony analiza bieżącej propagandy pozwala na zaobserwowanie sposobów wykorzystania mediów i bieżącej reakcji jej odbiorców. Poszczególne etapy analizy propagandy dotyczą następujących kategorii[42]:

1. ideologia i cel kampanii propagandowej (ang. *the ideology and purpose of the propaganda campaign*);
2. kontekst, w którym propaganda jest wykorzystywana (ang. *the context in which the propaganda occurs*);
3. identyfikacja propagandystów (ang. *identification of the propagandist*);
4. struktura organizacji propagandowej (ang. *the structure of the propaganda organization*);
5. odbiorcy (ang. *the target audience*);
6. techniki wykorzystania mediów (ang. *media utilization techniques*);
7. techniki specjalne do zwiększenia efektu (ang. *special techniques to maximize effect*);
8. reakcja odbiorców na poszczególne techniki propagandy (ang. *audience reaction to various techniques*);
9. przeciwpropaganda, jeżeli występuje (ang. *counterpropaganda, if present*);
10. efekty propagandy i jej ocena (ang. *effects and evaluation*).

[42] G.S. Jowert, V.O'Donnell, *Propaganda and…*, s. 290.

Powyższy schemat poddano szczegółowemu opisowi w rozdziale szóstym *How to analyze propaganda* przywołanej książki. Szczegółowa analiza zawartych tam zapisów jednoznacznie uwidacznia, że kampania propagandowa jest skomplikowanym procesem, bazującym na istniejącej w danym społeczeństwie ideologii w odniesieniu do specyfiki czasów, w których jest realizowana. Dysponuje ona odpowiednimi strukturami organizacyjnymi wspierającymi propagandystów w dotarciu do określonych grup odbiorców. Wykorzystywane są w niej wszystkie możliwe media społecznego przekazu, najlepiej znajdujące się w rękach propagandystów, wsparte narzędziami zwiększającymi efekty oddziaływania, takimi jak na przykład wymienione wcześniej techniki propagandy. Ich dobór podyktowany jest specyfiką grupy docelowej oddziaływania informacyjnego i jej reakcją na stosowane techniki. Należy pamiętać także o tym, że zgodnie z prawem akcji i reakcji będziemy mieli do czynienia z przeciwpropagandą, czego najczęstszym przejawem w naszym codziennym życiu jest walka polityczna obozu rządzącego z opozycją i odwrotnie. Skuteczna propaganda wymaga także wyciągania wniosków z przeszłości. Analiza przeszłych kampanii propagandowych pozwala zidentyfikować, które z metod i narzędzi były skuteczne, w jakim środowisku społecznym propaganda była realizowana, w jakim czasie i jakimi kanałami komunikacyjnymi odbywał się przekaz i co należy zmienić w przyszłości, aby zmaksymalizować przekaz i tym samym skuteczność propagandy.

Podsumowanie

Propaganda we współczesnym społeczeństwie jest coraz powszechniej wykorzystywana do realizacji celów politycznych czy też nawet interesów narodowych. Jej rozwój związany jest bezpośrednio z rozwojem ludzkości, a zwłaszcza z wykorzystywanymi kanałami komunikacyjnymi. W czasach starożytnych i we wczesnym średniowieczu jej zasięg był mocno ograniczony, tak jak ograniczone były środki przekazu. Wraz z rozwojem społecznym propaganda zyskiwała na znaczeniu i stała się skutecznym narzędziem wpływu. Można zidentyfikować kilka głównych punktów przełomowych w rozwoju ludzkości, które miały wpływ na skuteczność i zasięg propagandy. Zawsze były związane z rozwojem technicznym i nowymi wynalazkami, takimi jak: druk, telegraf, telefon, radio, telewizja, sieć Internet. Co przyniesie przyszłość – zobaczymy. Najprawdopodobniej będzie związana z elementami sztucznej inteligencji zdolnymi do samodzielnego kreowania przekazów informacyjnych.

Autor zdaje sobie sprawę z tego, że problematyka propagandy we współczesnym społeczeństwie jest bardzo skomplikowana i niejednorodna. Często trudno odróżnić, co jest propagandą, a co nią nie jest. Tym bardziej, że Internet, a zwłaszcza portale społecznościowe, są doskonałym środowiskiem propagandy ukierunkowanej na indywidualne preferencje jej użytkowników, często noszącej znamiona czarnej propagandy, współcześnie określane jako *fake news*. Jesteśmy zalewani informacjami z wielu źródeł i to, czy uznamy je za prawdziwe, zależy tylko i wyłącznie od nas. A do tego potrzebne są kompetencje, świadomość i wiedza. Paradoksem współczesnych czasów jest to, że mówimy o społeczeństwie informacyjnym, w którym *de facto* wiedza, jaką posiada rzeczone społeczeństwo, ma najmniejsze znaczenie. Nie mówimy tutaj o podmiotach takich jak państwa, instytucje międzynarodowe czy też korporacje. Dla nich wiedza jest niezbędna dla skutecznego działania. Ale te działania są tym skuteczniejsze, im uboższy jest intelektualny poziom środowiska oddziaływania.

Osobnym zagadnieniem, a raczej problemem i wyzwaniem, jest zjawisko konfliktu hybrydowego (wojny hybrydowej), w którym podmioty prawa międzynarodowego, ale nie tylko, wykorzystują informację jako narzędzie oddziaływania na obiekt, którym jest społeczeństwo, w celu realizacji przyjętych celów politycznych. Podyktowane to jest oczywiście pragmatyką związaną z redukcją kosztów prowadzenia konfliktu (wojny), zarówno w wymiarze ekonomicznym, jak również społecznym. Konflikt bowiem musi się opłacać i być społecznie akceptowalny. W tym obszarze propaganda będzie nabierała coraz większego znaczenia.

STRESZCZENIE

W artykule przedstawiono najważniejsze elementy propagandy wykorzystywanej jako narzędzie kreowania postaw w społeczeństwie. Zidentyfikowano środowiskowe uwarunkowania informacji oraz przedstawiono zakres odpowiedzialności bezpieczeństwa informacyjnego, uwzględniając takie obszary, jak bezpieczeństwo informacji, walkę informacyjną, dezinformację. Główny nacisk położono na zidentyfikowanie elementów składowych propagandy, schematu, według którego jest ona realizowana, oraz jej uwarunkowań społeczno-historycznych i kulturowych. Przedstawiono najważniejsze techniki manipulacji stosowane w propagandzie, a także sposób identyfikowania kampanii propagandowych.

Piotr Dela

Elements of Propaganda in Public Life

The chapter presents the most important element of propaganda used as a tool to create attitudes in society. The environmental determinants of information were identified and the scope of information security responsibilities was presented, taking into account such areas as information security, information warfare, disinformation. The main focus was on identifying the components of propaganda, the scheme by which it is implemented, and its socio- historical and cultural determinants. The most important techniques of manipulation used in propaganda as well as the way of identifying propaganda campaigns are presented.

Key words: *propaganda, disinformation, information security*

Bibliografia

Bączek P., *Zagrożenia informacyjne a bezpieczeństwo państwa polskiego*, Toruń 2006.

Chorobiński A., *Walka informacyjna jako fundamentalny składnik działalności terrorystycznej w przyszłości*, http://konkursy.byd.pl/userfiles/files/chorobinski.pdf (dostęp: 22.12.2018).

Ciborowski L., *Walka informacyjna*, Toruń 1999.

Conserva H.T., *Propaganda Techniques*, Bloomington 2003.

Gibson W., *Neuromancer*, przekł. P. Cholewa, Katowice 2009.

Godston R., Wirtz J.J., *Strategic Denial and Deception. The Twenty-First Century Challenge*, Waszyngton 2012.

Jowert G.S., O'Donnell V., *Propaganda and Persuasion*, Waszyngton 2006.

Kwećka R., *Strategia bezpieczeństwa informacyjnego państwa*, Warszawa 2014.

Ottis R., Lorents P., *Cyberspace: Definition and Implications*, Tallin 2010.

Volkoff V., *Psychosocjotechnika, dezinformacja – oręż wojny*, przekł. A. Arciuch, Warszawa 1999.

Wasilewski J., *Zarys definicji cyberprzestrzeni*, „Przegląd Bezpieczeństwa Wewnętrznego" 2013, nr 9.

VOL. 54 STUDIA POLITOLOGICZNE
DOI: 10.33896/SPolit.2019.54.4

Urszula Kurcewicz
ORCID: 0000-0001-5808-9851

Znaczenie tradycyjnych źródeł informacji w działalności infobrokerskiej

Słowa kluczowe:
infobroker, tradycyjne źródło informacji, proces informacyjny, archiwa, biblioteczne zbiory specjalne

STUDIA I ANALIZY

Wprowadzenie

Rozwój techniki digitalizacji pozwolił na udostępnianie na masową skalę informacji przechowywanych w pierwotnej formie na tradycyjnych nośnikach, głównie na papierze czy w różnego rodzaju formach utrwalania obrazu i dźwięku. Pomimo ogromnego postępu w dziedzinie digitalizacji znaczna liczba informacji dostępna jest nadal wyłącznie w pierwotnej formie. Dostęp do tej części zasobów wiedzy ludzkiej dla infobrokera możliwy jest za pośrednictwem instytucji, zarówno państwowych, jak i prywatnych, wyspecjalizowanych w zabezpieczaniu, przechowywaniu i udostępnianiu pierwotnych form informacji czy też stworzonych na ich podstawie dokumentów pochodnych i wtórnych.

Badania polskiego rynku informacji wskazują, że wśród najczęściej otrzymywanych przez firmy infobrokerskie zleceń znajdują się: wyszukiwanie informacji i publikacji naukowych, tworzenie bibliografii do prac naukowych, wyszukiwanie danych historycznych/archiwalnych czy przygotowywanie opracowań prasowych[1]. Rzetelne sprostanie tego typu

[1] B. Baczyńska, K. Grabarz, S. Machlowski, *Jak zostać brokerem informacji? Wybrane aspekty praktyczne*, [w:] S. Cisek, A. Januszko-Szakiel (red.), *Zawód infobroker. Polski rynek informacji*, Warszawa 2015, s. 100.

zamówieniom klientów wymaga od infobrokera umiejętności wyszukiwania, klasyfikowania i przetwarzania informacji pochodzących z tradycyjnych źródeł informacji, Internet zaś staje się dla nich źródłem komplementarnym. Infobroker powinien posiadać wiedzę o sieci krajowych i zagranicznych archiwów, bibliotecznych zasobach specjalistycznych czy materiałach przechowywanych w różnego rodzaju repozytoriach danych. Ważne jest, aby również orientował się w zasadach udostępniania i przetwarzania zgromadzonych przez te instytucje materiałów.

Celem artykułu jest wskazanie wagi biegłego wyszukiwania i przetwarzania informacji pochodzących z ich tradycyjnych źródeł jako niezbędnego elementu bazy umiejętności osoby wykonującej zawód infobrokera. W tym celu autorka podjęła się systematyzacji typologii tradycyjnych źródeł informacji oraz uporządkowania terminologii omawianej problematyki. W artykule ukazane zostały poszczególne etapy procesu informacyjnego z perspektywy kształcenia przyszłych adeptów infobrokeringu. Druga część tekstu odnosi się do tworzenia oraz doskonalenia indywidualnego warsztatu informacyjnego infobrokera.

Informacja i proces informacyjny

Istotą rozwoju komunikowania międzyludzkiego jest pojawianie się nowych kanałów komunikacyjnych. Szczególny postęp w tej dziedzinie przyniosła rewolucja technologiczna XX wieku, wynalazek radia i telewizji, a następnie Internetu i telefonii cyfrowej. Wprowadzenie nowych środków przekazu doprowadziło do powstania podziału na tzw. stare i nowe media. Analogicznie do niego w rozróżnieniu źródeł informacji zaczęto używać terminów „stare" i „nowe" źródła informacji. Jednakże pojawienie się tych nowych nie doprowadziło do unicestwienia starych, gdyż to one zachowują w sobie większość zgromadzonej wiedzy ludzkiej. Jako dziedzictwo kulturowe ludzkości przechowywane są pieczołowicie w bibliotekach, archiwach czy muzeach, służąc kolejnym pokoleniom badaczy. Wypracowano natomiast nowe formy dostępu do informacji w nich zawartych, których funkcjonowanie w obiegu informacyjnym umożliwiła technologia cyfrowa.

Termin „informacja" zarówno w naukowym znaczeniu, jak i w potocznym rozumieniu ma wiele interpretacji. Jest to pojęcie, którego naukowe zdefiniowanie napotyka na znaczne trudności. W poszczególnych dyscyplinach naukowych jego definicje istotnie od siebie odbiegają. Jerzy Ratajewski podaje, że do lat 80. XX wieku w nauce funkcjonowało już

ponad 400 różnych definicji tego pojęcia[2]. Wśród nich wyróżnić można dwie główne grupy: definicje cybernetyczne i definicje informatyczne.

W ujęciu cybernetycznym uznaje się, że informację stanowi: „[…] każdy czynnik zmniejszający stopień niewiedzy o badanym zjawisku, umożliwiający człowiekowi i innemu organizmowi żywemu lub urządzeniu lepsze poznanie otoczenia i przeprowadzenie w sprawniejszy sposób celowego działania. Inaczej mówiąc, informacją będzie wszelka treść zaczerpnięta z otaczającej rzeczywistości, zmieniająca stan danego obiektu na inny"[3].

Ujęcie informatyczne zakłada, że: „[…] informacją jest znaczenie (treść) przypisywane danym, przez które rozumie się liczby, fakty, pojęcia lub rozkazy przedstawiane w sposób wygodny do przesyłania, interpretacji lub przetwarzania metodami ręcznymi względnie automatycznymi"[4].

Cechami informacji jako zasobu społecznego są: niematerialność, niezużywalność, kumulowalność oraz odnawialność.

Podkreślić należy, iż termin „informacja" może być rozumiany trojako:

- jako wiadomość (wskazówka, pouczenie), czyli każda treść mająca znaczenie i stanowiąca odbicie rzeczywistości (zaczerpnięta ze świata zewnętrznego człowieka) lub ją zastępująca (ze świata wewnętrznego);
- czynność powiadomienia lub zakomunikowania czegoś;
- instytucja informująca o czymś[5].

W odniesieniu do infobrokeringu informację powinno traktować się jako wiadomość uporządkowaną, przeanalizowaną, przekazaną odbiorcy w odpowiedniej, zrozumiałej dla niego postaci, na którą ten zgłaszał wcześniej potrzebę w związku z realizacją określonych celów[6].

Jak zaznacza Aneta Januszko-Szakiel, działalność brokera informacji polega na: „[…] pozyskiwaniu, ocenie, weryfikacji, analizie i odpłatnym dostarczaniu informacji jawnoźródłowych (tj. jawnych i pozyskanych

---

[2] J. Ratajewski, *Informologia – nauka o informacji*, [w:] A. Jarosz (red.), *Informacja naukowa, bibliotekarstwo, zagadnienia wydawnicze*, Prace Naukowe Uniwersytetu Śląskiego. Studia Bibliologiczne, t. V, Katowice 1992, s. 9.

[3] H. Batorowska, B. Czubała, *Wybrane zagadnienia nauki o informacji i technologii informacyjnej*, Kraków 1996, s. 9.

[4] Tamże.

[5] J. Ratajewski, *Informologia…*, s. 10–11.

[6] G.K. Świderska (red.), *Informacja zarządcza w procesie formułowania i realizacji strategii firmy. Wyzwania dla polskich przedsiębiorstw*, Warszawa 2003, s. 52.

legalnie) ludziom i organizacjom zgłaszającym zapotrzebowanie na usługę infobrokerską"[7].

Termin „przetwarzanie informacji" oznacza wszystkie operacje przeprowadzone w procesie obiegu informacji. Takimi procesami są: wydobywanie informacji ze źródła, przekazywanie oraz utrwalanie[8].

Działalność informacyjna polega na organizowaniu i doskonaleniu przepływu informacji, przy czym występuje zastrzeżenie, iż przekazywanie informacji nie jest czynnością jednorazową, lecz wielokrotną i powtarzalną. Inaczej ujmując, działalność informacyjna określana jest jako zorganizowana działalność, której zadaniem jest gromadzenie, opracowanie i udostępnienie informacji o osiągnięciach nauki, techniki i innych dziedzinach życia społecznego[9].

Termin „służba informacyjna" odnosi się bądź do instytucji prowadzących działalność informacyjną, bądź do zespołu pracowników zatrudnionych w instytucjach prowadzących działalność informacyjną. Do najważniejszych zadań służby informacyjnej należą:

- informacja o dokumentach gromadzonych i przechowywanych w danej placówce;
- informacja o treści tych dokumentów;
- dokumentacja zbiorów własnych i innych placówek poprzez opracowania formalne, rzeczowe, analityczno-syntetyczne tych zbiorów;
- działalność wydawnicza;
- działalność dydaktyczno-metodyczna;
- propaganda ośrodka informacji i jego zbiorów[10].

---

[7] A. Januszko-Szakiel, *Informacja tworzywem przekazów infobrokerskich. Wybrane zagadnienia*, [w:] S. Cisek, A. Januszko-Szakiel (red.) *Zawód infobroker…*, s. 278.

[8] W procesie obiegu informacja ulega różnego rodzaju przekształceniom, np. nadawca może modyfikować postać informacji w celu jej zapisania lub przekazania, odbiorca zaś w celu jej odebrania i zrozumienia. E. Chmielewska-Gorczyca, B. Sosińska-Kalata, *Informacja naukowa z elementami naukoznawstwa*, Warszawa 1991, s. 27.

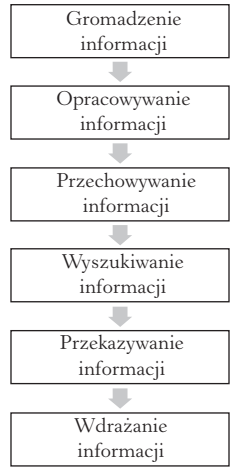
[9] W kontekście tematu niniejszego artykułu działalność informacyjną należy oddzielić od działalności dokumentacyjnej, która jest pojęciem węższym i najczęściej jest rozumiana dwojako. Po pierwsze – jako część działalności informacyjnej związanej z gromadzeniem i opracowywaniem dokumentów, po drugie – jako zbiór lub spis dokumentów dotyczących określonego zagadnienia bądź dobranych według innych kryteriów, np. dokumentacja zbiorów własnych biblioteki. Podstawę działalności dokumentacyjnej stanowią dokumenty, czyli materialne nośniki, na których została utrwalona informacja w znaczeniu wiadomości. M. Dembowska, *Dokumentacja i informacja naukowa. Zarys problematyki i kierunki rozwoju*, Warszawa 1965, s. 51–52.

[10] H. Batorowska, B. Czubała, *Wybrane zagadnienia nauki…*, s. 11.

Infobroker w realizacji zapotrzebowań informacyjnych klientów powinien biegle korzystać z informacji naukowej – stanowi to jedną z podstawowych kompetencji w ramach wyszukiwania źródeł informacji. Termin „informacja naukowa" oznacza: informację o osiągnięciach nauki; informację przeznaczoną dla pracowników nauki; informację opracowaną metodą naukową; dziedzinę wiedzy obejmującą całokształt zagadnień teoretycznych i praktycznych związanych z działalnością informacyjną[11]. W tym ostatnim znaczeniu termin „informacja naukowa" w literaturze problemu stosowany jest wymiennie z terminami: „informatologia" bądź „informologia"[12].

Proces informacyjny składa się z kolejnych etapów: gromadzenia, opracowywania, przechowywania, wyszukiwania, przekazywania i wdrażania informacji.

Rysunek 1. Etapy procesu informacyjnego



Źródło: opracowanie własne na podstawie: H. Batorowska, B. Czubała, *Wybrane zagadnienia nauki o informacji i technologii informacyjnej*, Kraków 1996, s. 32–33.

[11] Tamże, s. 11–12.

[12] Jerzy Ratajewski definiuje informologię jako naukę o informacji i informowaniu, obejmującą zarówno zagadnienia ogólnej teorii informologii, jak i poszczególnych dziedzin ze szczególnym uwzględnieniem informologii nauki. J. Ratajewski, *Wybrane problemy metodologiczne informologii nauki (informacji naukowej)*, Katowice 1994, s. 32–49. Z kolei Maria Dembowska podaje, że: „Przedmiotem informatologii jest działalność naukowo-informacyjna, której zadanie polega na udostępnianiu wyników nauki lub osiągnięć praktyki w celu wykorzystania tych zdobyczy dla dalszego rozwoju nauki, kultury, gospodarki. Informatologia zajmuje się całokształtem zagadnień teoretycznych i praktycznych związanych z działalnością naukowo-informacyjną", M. Dembowska, *Nauka o informacji naukowej (informatologia). Organizacja i problematyka badań w Polsce*, Warszawa 1991, s. 22–26.

Pierwszy etap procesu informacyjnego stanowi gromadzenie informacji obejmujące zespół czynności, takich jak: uzyskanie informacji o dokumencie, jego nazwanie i finalnie wprowadzenie do zbioru informacyjnego[13]. W drugim etapie następuje przygotowanie zawartości zbioru informacyjnego do udostępniania, polega ono na: wprowadzaniu, katalogowaniu, klasyfikowaniu, analizowaniu, przygotowaniu technicznym dokumentów oraz sporządzeniu opracowań dokumentacyjnych, na podstawie których możliwy będzie dostęp do konkretnych dokumentów. Przechowywanie informacji jest trzecim etapem procesu informacyjnego, odbywa się w sposób adekwatny dla danych nośników pamięci, na przykład w formie papierowej pierwotnej czy w formie mikrofilmów, a obecnie również pod postacią cyfrową. Na etapie czwartym – wyszukiwania informacji – przeprowadzana jest kwerenda informacyjna, czyli proces polegający na porównywaniu zapytań informacyjnych użytkowników, sformułowanych na przykład w postaci zapytań hasłowych, z zasobami zbioru informacyjnego. Przedostatni etap to przekazywanie (udostępnianie) informacji określonemu odbiorcy. Ostatni etap procesu informacyjnego to wdrażanie informacji, podczas którego użytkownik sprawdza przydatność uzyskanych informacji, wartościuje je i wykorzystuje w praktyce[14].

Miejsce infobrokera w procesie informacyjnym może być dwojakie. Z jednej strony jest on użytkownikiem zasobów wyspecjalizowanych instytucji, takich jak biblioteki czy archiwa, z drugiej strony, realizując zamówienie klienta, przechodzi przez wszystkie etapy procesu informacyjnego, gromadząc i opracowując informację do wdrożenia w formie końcowej, która optymalnie zaspokaja potrzeby informacyjne zamawiającego.

Klasyfikacja źródeł informacji

Współczesna nauka o informacji czerpie z dorobku nauk historycznych, w odniesieniu do metodologii historii formułuje definicje dokumentu oraz źródła informacji oraz proponuje ich klasyfikację. W naukach historycznych obok terminu „źródło historyczne" równoprawnie funkcjonuje określenie „nośnik pamięci historycznej"[15]. Jak pisze Henryk Dominiczak:

---

[13] Zbiór informacyjny definiowany jest jako uporządkowany zbiór dokumentów stanowiący podstawę działalności informacyjnej.

[14] H. Batorowska, B. Czubała, *Wybrane zagadnienia nauki…*, s. 32–33.

[15] Szerzej: M. Kula, *Nośniki pamięci historycznej*, Warszawa 2002.

„Źródłem jest to wszystko, co zachowało się z czasów minionych, co może służyć badaniom historyka, co odzwierciedla splot określonych stosunków, które niegdyś istniały w określonym miejscu, czasie i środowisku”[16].

Sformułowania uniwersalnej definicji pojęcia źródła historycznego na przestrzeni wieków podejmowało się wielu badaczy. W nauce polskiej krytyczny przegląd tych definicji przeprowadził Jerzy Topolski, a do jego ustaleń odwołują się kolejne generacje historyków. Jak zaznacza J. Topolski, definicje źródła historycznego generalnie można podzielić na jedno- i dwuczłonowe. W definicjach jednoczłonowych historycy wskazywali tylko na ślady (synonimicznie rezultaty, wytwory, pozostałości) działania, ewentualnie także samego istnienia człowieka, natomiast w definicjach dwuczłonowych źródło historyczne jest nie tylko pozostałością po celowej pracy człowieka, lecz stanowi także odbicie tej działalności w aspekcie poziomu świadomości społeczeństwa oraz w aspekcie przejawów rozwoju kultury w czasie, kiedy to źródło powstało[17].

Sam J. Topolski proponuje szerokie ujęcie pojęcia, w którym: „[…] źródłem historycznym są wszelkie źródła poznania historycznego (bezpośredniego i pośredniego), tzn. wszelkie informacje (w rozumieniu teorioinformacyjnym) o przeszłości społecznej, gdziekolwiek one się znajdują, wraz z tym, co owe informacje przekazuje (kanałem informacyjnym). Przeszłość społeczna, rzecz jasna, rozumiana jest tutaj szeroko, obejmując również warunki naturalne, w których żył człowiek”[18].

Klasyfikacje źródeł poznania historycznego, podobnie jak omówione podejścia definicyjne, kształtowały się od epoki średniowiecza, dzieląc środowisko historyków na zwolenników odmiennych koncepcji typologizacji. J. Topolski wskazuje na dwie wiodące klasyfikacje dychotomiczne:

---

16 H. Dominiczak, *Zarys metodologii historii*, Częstochowa 1995, s. 91.

17 J. Topolski, *Metodologia historii*, Warszawa 1984, s. 323. Przykładem definicji jednoczłonowej jest propozycja Marcelego Handelsmana zawarta w pracy z 1928 r., w której uczony źródłem historycznym nazywa: „utrwalony i zachowany ślad myśli, działania lub najogólniej życia ludzkiego”, M. Handelsman, *Historyka. Zasady metodologii i teorii poznania historycznego*, oprac. P. Węcowski, Warszawa 2010, s. 44. Uznaną w literaturze polskiej definicję dwuczłonową źródła historycznego w okresie powojennym zaproponował Gerard Labuda, który pisze, iż tym określeniem: „nazwiemy wszystkie pozostałości psychofizyczne i społeczne, które będąc wytworem pracy ludzkiej, a zarazem uczestnicząc w rozwoju życia społeczeństwa, nabierają przez to zdolności odbijania tego rozwoju. Wskutek tych swych właściwości (tj. wytworu pracy i zdolności odbijania) źródło jest środkiem poznawczym, umożliwiającym naukowe odtworzenie rozwoju społeczeństwa we wszystkich jego przejawach”, G. Labuda, *Próba nowej systematyki i nowej interpretacji źródeł historycznych*, „Studia Źródłoznawcze” 1957, t. I, s. 22.

18 J. Topolski, *Metodologia…*, s. 324.

- podział na źródła bezpośrednie i pośrednie;
- podział na źródła pisane i niepisane[19].

Do zwolenników pierwszego podziału należeli między innymi Johann Gustav Bernhard Droysen i Ernst Bernheim, w Polsce opowiadano się głównie za drugim podejściem, jego rzecznikami byli między innymi Joachim Lelewel i Stanisław Kościałkowski[20]. Marceli Handelsman uznawał za zasadne oba podziały. W kwestii klasyfikacji źródeł informacji widoczne są odniesienia do prac tego historyka we współczesnych polskich studiach informatologicznych. Stąd w niniejszym tekście zasadne wydaje się przytoczenie klasyfikacji źródeł przez niego opracowanej (i zobrazowanie jej odpowiednim schematem – rys. 2)[21]. M. Handelsman dzielił źródła historyczne na niepisane, rzeczowe, pozostałościowe (źródła materialne) oraz na materiały źródłowe w formie pisanej lub drukowanej. Do źródeł materialnych niepisanych zaliczał:

- środowisko naturalne przekształcone wskutek celowej działalności człowieka (na przykład sieć dróg komunikacyjnych);
- narzędzia pracy służące do przeobrażania środowiska naturalnego (na przykład narzędzia rolnicze);
- wytwory użytkowe, takie jak: ubiory, przedmioty codziennego użytku czy broń;
- zabytki architektury drewnianej i murowanej;
- miejsca pochówku (cmentarzyska i grobowce);
- pomniki wystawiane ku czci wybitnych postaci czy upamiętniające wydarzenia historyczne[22].

Drugą główną grupę źródeł w podziale Handelsmana stanowiły źródła pisane lub drukowane. W ich obrębie historyk wyróżniał opisy i akta. Akta rozumiał jako dokumenty bądź ustanawiające pewną czynność lub stwierdzające jej istnienie w stosunkach między ludźmi, bądź będące pisanym wyrazem czynności, w szczególności czynności prawnej. Aktami są wszelkiego rodzaju świadectwa i zaświadczenia urzędowe, umowy najmu, umowy o pracę itd. Wśród źródeł opisowych J. Handelsman wyróżniał trzy kategorie: opisy dziejopisarskie (roczniki, kroniki, historie), opisy pamiętnikarskie (dzienniki, biografie, wspomnienia, listy, pamiętniki), zbiory wiadomości (prasa – dzienniki, tygodniki i miesięczniki).

---

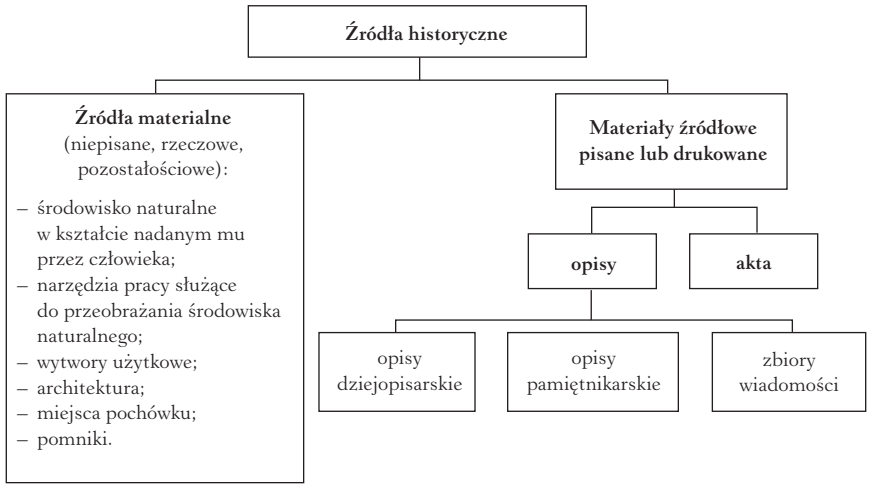
19 Tamże, s. 328–329.

20 H. Dominiczak, *Zarys metodologii…*, s. 103.

21 M. Handelsman, *Historyka. Zasady metodologii…*, s. 44–50.

22 Tamże.

Rysunek 2. Podział źródeł historycznych według Marcelego Handelsmana



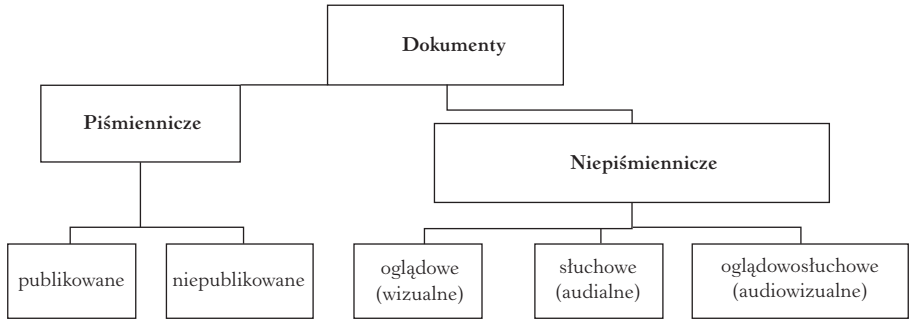
Źródło: opracowanie własne na podstawie: M. Handelsman, *Historyka. Zasady metodologii i teorii poznania historycznego*, oprac. P. Węcowski, Warszawa 2010, s. 44–50.

W naukach historycznych za bardzo istotną poznawczo grupę źródeł uznaje się ikonografię – ryciny, obrazy, przedstawienia miast, pieczęcie, herby, środki płatnicze itd. – jest to grupa źródeł historycznych ułatwiająca poznanie wielu zjawisk przeszłości, które w formie oryginalnej nie zachowały się do czasów współczesnych. Rozwój techniki w XIX i XX wieku wprowadził szereg nowych nośników pamięci historycznej: fotografię, radio, kinematografię, które stanowią nieocenione źródła informacji o minionym czasie ze względu na fakt, iż pozwalały na bieżące rejestrowanie zjawisk.

We współczesnej nauce o informacji za dokument uważa się informację wraz z materiałem, na którym została utrwalona. Zgodnie z kryterium formy zapisu wyróżnia się dokumenty piśmiennicze i dokumenty niepiśmiennicze. Dokumenty piśmiennicze zawierają treść utrwaloną na dowolnym materiale (w czasach nowożytnych najczęściej papierze) pod postacią tekstu słownego (pisma lub druku) bez względu na kod (język), w jakim zostały zapisane. W dokumentach niepiśmienniczych treść wyrażona jest w postaci obrazu lub dźwięku bądź w formie mieszanej. Pierwsza grupa, dokumenty oglądowe (wizualne), odbierane są przez adresata wyłącznie za pośrednictwem wzroku, należą do niej fotografie, przeźrocza, rysunki, obrazy, grafiki, modele, eksponaty muzealne, nieme filmy itd. Dokumenty słuchowe (audialne) przekazują treść za pomocą dźwięku i odbierane są poprzez zmysł słuchu, do tej grupy zalicza się

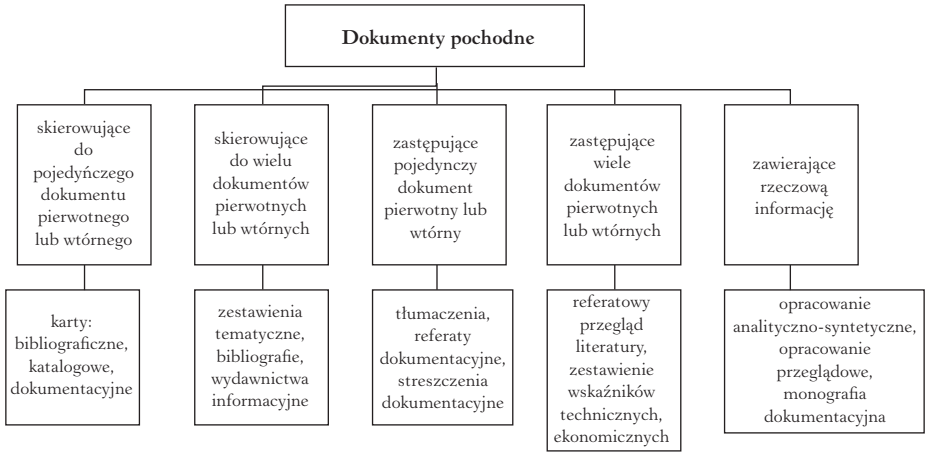
formy zapisu dźwiękowego, takie jak płyty gramofonowe, różnego rodzaju taśmy. Najbardziej typowym przykładem dokumentów oglądowosłuchowych (audiowizualnych) są filmy dźwiękowe. W związku z rozwojem technik reprodukcji podział dokumentów piśmienniczych na publikowane i niepublikowane ma coraz bardziej umowny charakter.

Rysunek 3. Klasyfikacja dokumentów według formy zapisu



Źródło: opracowanie własne na podstawie: H. Batorowska, B. Czubała, *Wybrane zagadnienia nauki o informacji i technologii informacyjnej*, Kraków 1996, s. 35.

Rysunek 4. Typy dokumentów pochodnych



Źródło: opracowanie własne na podstawie: K. Sosnowska, *Pomoce do nauczania przedmiotu „Opracowanie dokumentacyjne źródeł informacji". Ćwiczenia. Sprawdziany*, Warszawa 1980, s. 21.

Drugim podstawowym kryterium podziału dokumentów są sposób przygotowania i stopień przetworzenia treści. Na ich podstawie

Tabela 1. Systematyka typologii dokumentów

Kryterium	Typy dokumentów
Kryterium zastosowanego kodu	• piśmiennicze • niepiśmiennicze
Kryterium pochodzenia i sposobu powstawania	• pierwotne (prymarne) • pochodne • wtórne
Kryterium nośnika fizycznego	• konwencjonalne • zminiaturyzowane
Kryterium zasięgu upowszechniania	• opublikowane • nieopublikowane
Kryterium stopnia dostępności	• publikacje powszechnie dostępne • publikacje o ograniczonym udostępnianiu • publikacje do użytku wewnętrznego • publikacje dostępne tylko bibliograficznie
Kryterium cech wydawniczych	• periodyczne • nieperiodyczne
Kryterium formy wydawniczej	• wydawnictwa samoistne • utwory • fragmenty
Kryterium formy piśmienniczej	• monografie (oryginalne i kompilacyjne) • doniesienia, komunikaty • teksty źródłowe • publicystyka • podręczniki • spisy bibliograficzne • encyklopedie • słowniki
Kryterium biblioteczne	• wydawnictwa zwarte • wydawnictwa ciągłe • zbiory specjalne
Kryterium przeznaczenia czytelniczego i zaspokajania potrzeb	• naukowe • techniczno-ekonomiczne • społeczno-polityczne • artystyczne • społeczno-kulturalne • dokumenty specjalne, szczególnego rodzaju i przeznaczenia

Źródło: opracowanie własne.

wyróżnia się dokumenty: pierwotne, pochodne i wtórne[23]. Dokumenty pierwotne (prymarne) występują w formie oryginalnej, w jakiej zostały sporządzone przez autora, stanowią podstawę sporządzania dokumentów pochodnych i wtórnych. Dokumenty pochodne sporządzane są na podstawie dokumentów pierwotnych lub wtórnych, zawierają ich charakterystykę formalną lub treściową (bądź obie łącznie). Przedstawiają charakterystykę dokumentu pierwotnego i jego zawartości.

Dokumenty wtórne to dokumenty sporządzone na podstawie dokumentów prymarnych bądź pochodnych, są z nimi identyczne pod względem zawartości treści, często natomiast różnią się w formie zewnętrznej, czego przykładem są mikrofilmy.

We współczesnej literaturze problemu oprócz omówionych powyżej dwóch typologizacji dokumentów wyróżnia się ich podziały na podstawie takich kryteriów, jak: nośnik fizyczny, zasięg upowszechniania, stopień dostępności, cechy wydawnicze, forma wydawnicza, forma piśmiennicza, kryterium biblioteczne oraz kryterium ludzkiej działalności (zaspokajania potrzeb i przeznaczenia czytelniczego). Typologie te zbiorczo ukazuje tabela 1.

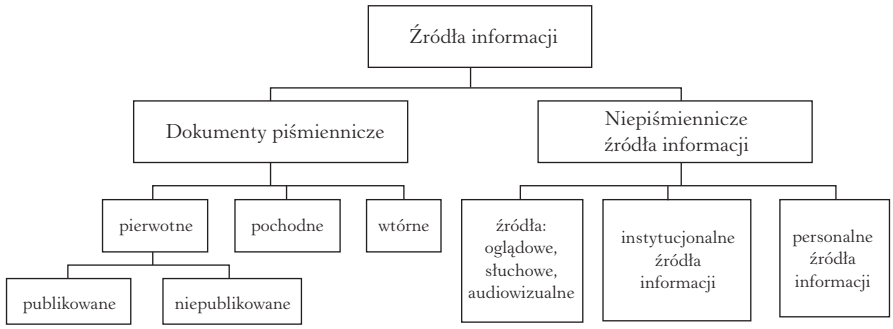
Pojęcie „źródło informacji" jest pojęciem szerszym niż dokument, gdyż oznacza dowolny system wypracowujący informację lub zawierający informację przeznaczoną do przekazania w celu zaspokojenia potrzeb informacyjnych[24]. Podstawowym współcześnie funkcjonującym podziałem źródeł informacji jest rozróżnienie na źródła dokumentalne (omówione w pierwszej części podrozdziału), źródła instytucjonalne oraz źródła personalne. Dotarcie do odpowiednich źródeł personalnych i instytucjonalnych jest niezwykle cenne w przypadku działalności infobrokerskiej. Personalnymi źródłami informacji mogą być świadkowie wydarzeń historycznych, naukowcy, eksperci, rzecznicy prasowi itd.

Infobroker uzyska kompletny produkt zaspokajający potrzeby klienta pod warunkiem wykorzystania maksymalnej ilości typów źródeł informacji. Ich podstawowe rodzaje obrazuje rysunek 5.

[23] H. Batorowska, B. Czubała, *Wybrane zagadnienia nauki*…, s. 35.

[24] E. Chmielewska-Gorczyca, B. Sosińska-Kalata, *Informacja naukowa*…, s. 69.

Rysunek 5. Klasyfikacja źródeł informacji



Źródło: opracowanie własne na podstawie: H. Batorowska, B. Czubała, *Wybrane zagadnienia nauki o informacji i technologii informacyjnej*, Kraków 1996, s. 36.

Dokumenty piśmiennicze pierwotne dzielą się na dokumenty piśmiennicze publikowane (do których należą wydawnictwa zwarte – książki, broszury oraz wydawnictwa ciągłe – czasopisma, wydawnictwa seryjne) oraz specjalne rodzaje wydawnictw (opisy patentowe, normy, literatura firmowa oraz dokumenty życia społecznego[25]). Wśród dokumentów piśmienniczych niepublikowanych znajdują się materiały archiwalne[26], rękopisy (maszynopisy), sprawozdania ze zjazdów i konferencji naukowych oraz dysertacje naukowe (nieopublikowane – głównie prace magisterskie i doktorskie)[27].

Warsztat infobrokera w zakresie wykorzystania tradycyjnych źródeł informacji

Baza źródłowa wykorzystywana w pracy infobrokera jest w dużym stopniu niejednolita, jej bogactwo wzrastało wraz z następującymi po

[25] Terminem „dokumenty życia społecznego" określa się druki wydawane w celu osiągnięcia doraźnych celów informacyjnych, propagandowych, reklamowych i normatywnych, odzwierciedlające wewnętrzną działalność różnego rodzaju organizacji, stowarzyszeń i instytucji oraz grup społecznych. Są to różnego rodzaju ulotki, prospekty, afisze, reklamy. Zwykle przeznaczone są dla określonego kręgu odbiorców i najczęściej do pozaksięgarskiego sposobu rozpowszechniania. Od dokumentów życia społecznego należy odróżnić literaturę patentową, a tzw. literatura firmowa pod postacią katalogów fabrycznych, prospektów, ulotek i folderów może stanowić dokumenty życia społecznego.

[26] Materiały archiwalne mogą występować pod postacią drukowanych dokumentów wtórnych, są to najczęściej tematyczne zbiory dokumentów źródłowych.

[27] H. Batorowska, B. Czubała, *Wybrane zagadnienia nauki*…, s. 36.

sobie epokami, rozwojem techniki i stosunków międzyludzkich. Można stwierdzić wręcz, że badacz czasów najnowszych styka się z nadmiarem i „nadróżnorodnością" źródeł, wśród których wyselekcjonowanie tych stanowiących jądro badanego problemu może stanowić nie lada wyzwanie. Zatem wykonywanie zawodu infobrokera wymaga biegłego poruszania się w dorobku instytucji naukowych powołanych i wyspecjalizowanych w gromadzeniu, przechowywaniu i udostępnianiu źródeł. Wśród nich należy wymienić instytucje obsługujące naukę, takie jak: biblioteki, archiwa naukowe, ośrodki informacji naukowej, technicznej i ekonomicznej. W świetle tytułu niniejszego artykułu szczególną uwagę należy poświecić dwóm instytucjom: bibliotekom i archiwom.

Biblioteka jest najstarszym typem instytucji, która w sposób zorganizowany i systematyczny zajmuje się gromadzeniem, przechowywaniem, opracowywaniem i udostępnianiem informacji. Coraz częściej na czoło zadań pełnionych przez biblioteki wysuwa się ich funkcja informowania o piśmiennictwie, zbiorach bibliotecznych i ich treści. Współcześnie polskie biblioteki wchodzące w skład ogólnokrajowej sieci bibliotecznej mają różny status, są to między innymi biblioteki: naukowe, fachowe, pedagogiczne, publiczne, działające przy szkołach i uczelniach wyższych, instytucji badawczych. W praktyce jedna biblioteka z reguły łączy w sobie cechy dwóch lub kilku rodzajów – na przykład biblioteki naukowej i biblioteki publicznej.

Szczególne miejsce w polskiej sieci bibliotecznej zajmuje Biblioteka Narodowa (BN) powołana po odzyskaniu niepodległości rozporządzeniem Prezydenta Rzeczpospolitej Polskiej z 24 lutego 1928 roku jako centralna biblioteka państwowa[28]. Obecnie zakres działania BN jest bardzo rozległy. Jako centralna biblioteka państwowa inicjuje, organizuje i wykonuje prace międzybiblioteczne, które służą centralnej ewidencji, wymianie i udostępnianiu zbiorów bibliotek z terenu kraju, prowadzi również szeroko rozumianą działalność informacyjną. Biblioteka Narodowa gromadzi i udostępnia czytelnikom: książki z XIX–XXI wieku, czasopisma z XIX–XXI wieku, dokumenty życia społecznego, stare druki, rękopisy, zbiory muzyczne, zbiory dźwiękowe i audiowizualne, dokumenty elektroniczne, zbiory kartograficzne, zbiory ikonograficzne, zbiory bibliologiczne oraz mikrofilmy[29].

---

[28] Rozporządzenie Prezydenta Rzeczpospolitej Polskiej z 24 lutego 1928 r. O Bibljotece Narodowej, Dz.U. 1928, poz. 183.

[29] Oficjalna strona Biblioteki Narodowej w Warszawie – „Zbiory", https://www.bn.org.pl/o-nas/zbiory-bn/zbiory (dostęp: 11.12.2018).

Dla infobrokera szczególnie cennym przewodnikiem po źródłach są wydawnictwa informacyjne, w tym bibliografie[30]. Zgodnie z wytycznymi polskiej normy PN-89/N-01225 *Rodzaje i części składowe bibliografii* uznaje się dwa znaczenia terminu „bibliografia". W pierwszym jest to uporządkowany zbiór opisów bibliograficznych dokumentów dobranych według określonych kryteriów, którego celem jest informowanie o istnieniu tych dokumentów, na ogół bez względu na miejsce ich przechowywania. W drugim znaczeniu jest to dziedzina wiedzy i działalności praktycznej obejmująca swoim zakresem problemy opisywania dokumentów w celu ich identyfikacji oraz zasady tworzenia i użytkowania bibliografii[31].

Biblioteka Narodowa prowadzi Polską Bibliografię Narodową, której podstawowym bieżącym członem jest „Przewodnik Bibliograficzny" rejestrujący wydawnictwa zwarte. Do końca 2009 roku wychodził drukiem, obecnie ogłaszany jest w formie plików PDF[32].

W obrębie informacji o pozostałych wydawnictwach książkowych BN ogłasza w formie drukowanej „Bibliografię podziemnych druków zwartych z lat 1976–1989" rejestrującą druki zwarte, które zostały wydane poza cenzurą przez nielegalne wydawnictwa, partie polityczne, grupy i ruchy religijne itp. w latach 1976–1989. Bibliografia obejmuje również wydawnictwa Niezależnego Samorządnego Związku Zawodowego „Solidarność" w okresie legalnej jego działalności[33]. Bibliografia „Polonica zagraniczne" prezentuje możliwie najbardziej kompletne i aktualne informacje na temat obecności Polski, Polaków i kultury polskiej w światowej produkcji wydawniczej[34]. W ramach informacji o zawartości czasopism prowadzone są: „Bibliografia Wydawnictw Ciągłych", „Bibliografia Wydawnictw Ciągłych Nowych, Zawieszonych i Zmieniających Tytuł",

[30] Praktyczny wymiar korzystania z bibliografii opisał m.in. Andrzej Chodubski, odsyłając jednocześnie do najważniejszych pozycji bibliograficznych w polskiej literaturze, A.J. Chodubski, *Wstęp do badań politologicznych*, Gdańsk 2004, s. 84–89.

[31] Z. Żmigrodzki, *Bibliografia. Metodyka i organizacja*, Warszawa 2000, s. 13.

[32] Oficjalna strona Biblioteki Narodowej w Warszawie – „Przewodnik Bibliograficzny", https://www.bn.org.pl/bibliografie/bibliografia-narodowa/przewodnik-bibliograficzny (dostęp: 11.12.2018).

[33] Oficjalna strona Biblioteki Narodowej w Warszawie – „Książki polskie podziemne (1976–1989)", https://www.bn.org.pl/bibliografie/bibliografia-narodowa/ksiazki-polskie-podziemne-(1976-1989) (dostęp: 11.12.2018).

[34] Oficjalna strona Biblioteki Narodowej w Warszawie – „Polonica zagraniczne", https://www.bn.org.pl/bibliografie/bibliografia-narodowa/polonica-zagraniczne (dostęp: 11.12.2018).

„Bibliografia niezależnych wydawnictw ciągłych z lat 1976–1990" oraz „Bibliografia Zawartości Czasopism"[35].

Drugi typ instytucji gromadzących, opracowujących i udostępniających dokumenty źródłowe stanowią, obok bibliotek, archiwa naukowe. W Polsce organem zarządzającym archiwami jest Naczelny Dyrektor Archiwów Państwowych podlegający ministrowi kultury i dziedzictwa narodowego[36]. Centralnym urzędem obsługującym Naczelnego Dyrektora Archiwów Państwowych jest Naczelna Dyrekcja Archiwów Państwowych sprawująca kontrolę nad siecią archiwów państwowych, w skład której wchodzą archiwa centralne: Archiwum Główne Akt Dawnych, Archiwum Akt Nowych, Narodowe Archiwum Cyfrowe[37] oraz archiwa państwowe zlokalizowane w miastach wojewódzkich wraz z oddziałami i ekspozyturami[38].

Archiwum Główne Akt Dawnych dysponuje zasobem historycznym obejmującym materiały archiwalne z okresu od XII wieku do I wojny światowej. Oprócz polskich dokumentów przechowywane są w nim archiwalia władz zaborczych oraz archiwa rodzin i osób prywatnych o szczególnym znaczeniu z terenów dawnej Rzeczypospolitej – Korony i Litwy, oraz z poszczególnych zaborów[39].

Archiwum Akt Nowych przechowuje i udostępnia dokumentację archiwalną wytworzoną po 1918 roku. W jego zasobach znajduje się spuścizna czasów Polski Ludowej. W 1990 roku archiwum przejęło zasób Archiwum Komitetu Centralnego Polskiej Zjednoczonej Partii Robotniczej zawierający akta organów wybieralnych i partyjnego aparatu

---

35 Oficjalna strona Biblioteki Narodowej w Warszawie – Bibliografia Narodowa, https://www.bn.org.pl/bibliografie/bibliografia-narodowa (dostęp: 11.12.2018).

36 Ustawa z 14 lipca 1983 r. o narodowym zasobie archiwalnym i archiwach (tekst jedn. Dz.U. 2018, poz. 217).

37 Narodowe Archiwum Cyfrowe powstało w 2008 r. w wyniku przekształcenia Archiwum Dokumentacji Mechanicznej. Obecnie odpowiedzialne jest za digitalizację i dokumentację elektroniczną oraz przechowuje dokumentację fotograficzną i audiowizualną.

38 Oficjalna strona Naczelnej Dyrekcji Archiwów Państwowych, https://www.archiwa.gov.pl/pl/o-nas (dostęp: 10.12.2018).

39 Dużym ułatwieniem w pracy infobrokera są przewodniki i informatory dotyczące zasobu zgromadzonego w Archiwum Głównym Akt Dawnych: D. Lewandowska (red.), *Archiwum Główne Akt Dawnych w Warszawie. Informator o zasobie archiwalnym*, Warszawa 2008; T. Zielińska (red.), *Archiwum Główne Akt Dawnych. Informator o zasobie*, Warszawa 1992; J. Karwasińska (red.), *Archiwum Główne Akt Dawnych w Warszawie. Przewodnik po zespołach*, t. I, *Archiwa dawnej Rzeczypospolitej*, Warszawa 1975; F. Ramotowska (red.), *Archiwum Główne Akt Dawnych w Warszawie. Przewodnik po zasobie*, t. II, *Epoka porozbiorowa*, Warszawa 1988.

wykonawczego – materiały zjazdów, konferencji, narad, plenarnych posiedzeń, w tym Centralnej Komisji Rewizyjnej i Centralnej Komisji Kontroli Partyjnej oraz protokoły Biura Organizacyjnego, Biura Politycznego i Sekretariatu KC PZPR. Do archiwum przekazywana jest dokumentacja archiwalna naczelnych organów władzy państwowej i urzędów centralnych wytworzona po 1989 roku.

Poszukując materiałów archiwalnych, infobroker może zwrócić się również do innych instytucji archiwalnych funkcjonujących w kraju, takich jak:

- Archiwum i Muzeum Pomorskie Armii Krajowej oraz Wojskowej Służby Polek (Fundacja Generał Elżbiety Zawackiej);
- Archiwum Nauki PAN i PAU w Krakowie;
- Archiwum Ośrodka KARTA;
- Archiwum Państwowego Muzeum Auschwitz-Birkenau w Oświęcimiu;
- Archiwum Państwowego Muzeum na Majdanku;
- Archiwum Polskiego Radia;
- Archiwum Polskiej Akademii Nauk w Warszawie (oraz Oddział w Poznaniu);
- Archiwum Sejmu;
- Archiwum Zamku Królewskiego w Warszawie;
- Archiwum Żydowskiego Instytutu Historycznego;
- Instytut Pamięci Narodowej;
- Polskie Centrum Informacji Muzycznej;
- Repozytorium Cyfrowe Filmoteki Narodowej;
- Stowarzyszenie „Archiwum Solidarności";
- Wojskowe Biuro Historyczne – Centralne Archiwum Wojskowe.

Infobroker prowadząc kwerendę archiwalną może korzystać również z zasobów instytucji polonijnych. Wśród nich należy wymienić te posiadające wyjątkowo cenne zbiory:

- Instytut Józefa Piłsudskiego w Ameryce – Nowy Jork;
- Instytut Literacki Kultura w Paryżu;
- Muzeum Polskie w Rapperswilu;
- Ośrodek Dokumentacji Pontyfikatu Jana Pawła II w Rzymie;
- Polski Instytut Naukowy w Ameryce (PIASA) – Nowy Jork;
- Stała Konferencja Muzeów, Bibliotek i Archiwów Polskich na Zachodzie.

Biblioteki i archiwa stanowią tylko część bardzo rozbudowanej sieci placówek informacji, w której skład wchodzą również instytucje publiczne i społeczne, te jednak coraz częściej preferują nowe formy komunikacji

z odbiorcami za pośrednictwem takich kanałów przekazu, jak Internet czy telefonia komórkowa.

Konkluzje

We współczesnych czasach ogromny zasób informacji, intensywny rozwój nauki oraz tempo postępu techniki stwarzają potrzebę istnienia wielu równolegle działających kanałów informacyjnych. Pojawienie się nowego zawodu, jakim jest broker informacji, było odpowiedzią na potrzeby współczesnych społeczeństw informacyjnych. Zaznaczyć należy, że pomimo iż cechą rynku informacji jest fakt, że wprowadza się nań nowe formy i techniki, z reguły bardziej sprawne, szybsze i tańsze, to nie eliminują one dotychczasowych tradycyjnych form przekazu. Na rynku informacyjnym nowe środki komunikacji współegzystują z istniejącymi od początku cywilizacji. Zatem, tak jak druk nie wyeliminował pisma ręcznego, a pojawienie się radia i telewizji nie zlikwidowało słowa drukowanego, tak można zakładać, że pojawienie się nowych źródeł informacji nie doprowadzi do całkowitego zaniku źródeł tradycyjnych.

Kompetentny infobroker musi wykorzystywać całość zasobu informacyjnego. W czasach społeczeństwa cyfrowego, jak pisali pod koniec ubiegłego wieku Bill Kovach i Tom Rosenstiel, wytworzyła się: „[…] kultura niepopartych zarzutów i twierdzeń, która rozwija się kosztem dawnej kultury weryfikacji”[40]. Na rynku informacyjnym pojawiają się nowe zjawiska niebezpieczne dla kultury weryfikacji, jak na przykład anonimowe źródła informacji zapełniające Internet. Już w 1920 roku Walter Lippmann upominał, że: „[…] zarówno publiczne, jak i prywatne poglądy zależą od precyzyjnych, wiarygodnych relacji o przebiegu wydarzeń. Nie to, co ktoś mówi, i nie to, co ktoś chciałby, aby było prawdą, ale to, co nią jest, daje właściwy osąd”[41]. Stąd na zawód infobrokera nałożona została wielka odpowiedzialność weryfikacji informacji i aby jej sprostać, konieczne jest dysponowanie odpowiednim warsztatem źródłowym.

[40] B. Kovach, T. Rosentiel, *Warp Speed. America in the Age of Mixed Media*, Nowy Jork 1999, cyt za: A. Briggs, P. Burke, *Społeczna historia mediów. Od Gutenberga do Internetu*, Warszawa 2010, s. 361.

[41] Cyt za: tamże.

Streszczenie

Współcześnie Internet stał się wiodącym źródłem informacji, nie można jednak umniejszać roli informacji pozyskiwanych z tradycyjnych źródeł, nie straciły one swojej aktualności i znaczenia. Za właściwą strategię infobrokera należy uznać taką, która uwzględnia techniki wyszukiwawcze typowe dla danego rodzaju źródła. W pierwszej części tekstu skupiono się na metodologii nauki o informacji oraz wskazano miejsce infobrokera w procesie informacyjnym. Jednym z kluczowych aspektów prezentowanej analizy jest przedstawienie wiodących definicji i klasyfikacji tradycyjnych źródeł informacji z zakresu wiedzy historycznej i informatologii. Druga część tekstu odnosi się do kompetencji osoby wykonującej zawód brokera informacji, do których zaliczają się: wyszukiwanie, gromadzenie i opracowywanie informacji z takich instytucji, jak państwowe i prywatne archiwa, biblioteki czy muzea.

Urszula Kurcewicz

The Importance of the Traditional Sources of Information in Information Broker Activity

Though in the present age the Internet is an important resource for gathering information, there are also plenty of significant and helpful traditional sources. The best data collection strategy in information broker activity is one that employs techniques from both. The first part of the text is focused on the methodology of obtaining information. The author points out the place of information broker in the information process. One of the key aspects of the following analysis is an attempt to indicate the main definitions and classifications of the traditional sources of information in a field of historical knowledge as well as in the information science. The second part of the article considers information broker professional qualifications, since for him is important to know how to search, gather and integrate information from institution such as the national and private archives, libraries or museums.

Key words: *information broker, traditional sources of information, information process, archives, special library material*

Bibliografia

Baczyńska B., Grabarz K., Machlowski S., *Jak zostać brokerem informacji? Wybrane aspekty praktyczne*, [w:] S. Cisek, A. Januszko-Szakiel (red.), *Zawód infobroker. Polski rynek informacji*, Warszawa 2015.

Batorowska H., Czubała B., *Wybrane zagadnienia nauki o informacji i technologii informacyjnej*, Kraków 1996.

Chmielewska-Gorczyca E., Sosińska-Kalata B., *Informacja naukowa z elementami naukoznawstwa*, Warszawa 1991.

Dembowska M., *Dokumentacja i informacja naukowa. Zarys problematyki i kierunki rozwoju*, Warszawa 1965.

Dembowska M., *Nauka o informacji naukowej (informatologia). Organizacja i problematyka badań w Polsce*, Warszawa 1991.

Dominiczak H., *Zarys metodologii historii*, Częstochowa 1995.

Handelsman M., *Historyka. Zasady metodologii i teorii poznania historycznego*, oprac. P. Węcowski, Warszawa 2010.

Januszko-Szakiel A., *Informacja tworzywem przekazów infobrokerskich. Wybrane zagadnienia*, [w:] S. Cisek, A. Januszko-Szakiel (red.), *Zawód infobroker. Polski rynek informacji*, Warszawa 2015.

Kovach B., Rosentiel T., *Warp Speed. America in the Age of Mixed Media*, Nowy Jork 1999.

Kula M., *Nośniki pamięci historycznej*, Warszawa 2002.

Labuda G., *Próba nowej systematyki i nowej interpretacji źródeł historycznych*, „Studia Źródłoznawcze" 1957, t. I.

Ratajewski J., *Wybrane problemy metodologiczne informologii nauki (informacji naukowej)*, Katowice 1994.

Ratajewski J., *Informologia – nauka o informacji*, [w:] A. Jarosz (red.), *Informacja naukowa, bibliotekarstwo, zagadnienia wydawnicze*, Prace Naukowe Uniwersytetu Śląskiego. Studia Bibliologiczne, t. V, Katowice 1992.

Sosnowska K., *Pomoce do nauczania przedmiotu „Opracowanie dokumentacyjne źródeł informacji". Ćwiczenia. Sprawdziany*, Warszawa 1980.

Świderska G.K. (red.), *Informacja zarządcza w procesie formułowania i realizacji strategii firmy. Wyzwania dla polskich przedsiębiorstw*, Warszawa 2003.

Topolski J., *Metodologia historii*, Warszawa 1984.

Żmigrodzki Z, *Bibliografia. Metodyka i organizacja*, Warszawa 2000.

VOL. 54 STUDIA POLITOLOGICZNE
DOI: 10.33896/SPolit.2019.54.5

Magdalena Tomaszewska-Michalak
ORCID: 0000-0001-5441-0396

Prawne aspekty pozyskiwania informacji w Internecie

Słowa kluczowe:
prawne aspekty infobrokeringu, pozyskiwanie informacji w Internecie, phishing, sock puppetry

Wprowadzenie

Nie jest tajemnicą, że pozyskiwanie informacji z różnych źródeł zawsze leżało w zakresie zainteresowania takich instytucji, jak służby chroniące porządek publiczny czy agencje wywiadowcze. Wejście w posiadanie odpowiednich informacji było i jest również istotne z punktu widzenia podmiotów prywatnych. Firmy sprawdzające swoich kontrahentów czy pracodawcy usiłujący pozyskać dodatkową wiedzę o kandydatach do pracy nie stanowią obecnie przypadków odosobnionych. Pozyskiwanie informacji nie występuje jednakże jedynie na poziomie instytucjonalnym czy pracowniczym. Szybki postęp technologiczny bowiem sprawił, że nierzadko każdy z nas ma możliwość znalezienia w Internecie informacji dotyczących konkretnego zagadnienia czy interesującej go osoby. Obecne czasy każą zatem spojrzeć na informację jako dobro, które w niepowołanych rękach może stanowić zagrożenie lub działać na niekorzyść jednostki. Dlatego też na zagadnienie pozyskiwania informacji powinno spojrzeć się szerzej – nie tylko z perspektywy technicznych możliwości, jakie istnieją w tym zakresie, ale również jak na zjawisko społeczne mające określone konsekwencje oraz na działanie podlegające odpowiednim ograniczeniom prawnym. Celem niniejszego artykułu jest zatem wskazanie społeczno-prawnych problemów związa-

nych z pozyskiwaniem oraz wykorzystywaniem informacji zamieszczonych w Internecie.

Analizując zagadnienie pozyskiwania informacji w Internecie, należy zwrócić uwagę na pojęcia takie jak tzw. biały wywiad / wywiad jawnoźródłowy (ang. OSINT – *open source intelligence*) oraz przeciwstawiany mu tzw. czarny wywiad. Nie istnieje jedna oficjalna definicja białego wywiadu[1], jednakże pojęcie to w literaturze tłumaczone jest jako pozyskiwanie informacji z jawnych oraz ogólnodostępnych źródeł. Zgodnie z definicją przyjętą przez NATO w 2002 roku biały wywiad polega na poszukiwaniu informacji pochodzących z jawnych źródeł za pomocą legalnych metod i środków[2]. Przywołana publikacja odwołuje się co prawda do pozyskiwania informacji w celu usprawnienia pracy wywiadu, jednakże zawarte w niej zostało również twierdzenie świadczące o tym, że jej autorzy dostrzegli możliwość korzystania z OSTINT nie tylko przez osoby pracujące w wywiadzie[3].

Przeciwieństwem białego wywiadu jest tzw. czarny wywiad, który polega na pozyskiwaniu informacji przy wykorzystaniu metod prawnie zakazanych użytkownikowi Internetu[4]. Poza białym i czarnym wywiadem istnieje pojęcie pośrednie – tzw. szary wywiad, który różni się od wcześniej wymienionych tym, że z jednej strony pozyskanie informacji w jego ramach jest nadal legalne, jednakże dostęp do szukanych danych może być trudniejszy niż w przypadku białego wywiadu (na przykład do materiałów konferencyjnych, archiwalnych informacji zamieszczonych na stronie internetowej)[5]. Ponadto niektóre działania podejmowane w ramach szarego wywiadu mogą zostać uznane za nieetyczne (na przykład wykorzystanie socjotechnik do pozyskania danych)[6]. Analizując możliwości pozyskania danych w Internecie, należy pamiętać, że niektóre metody szukania informacji pod względem legalności klasyfikowane są z uwzględnieniem kategorii podmiotu podejmującego wskazane działania.

---

1 B. Stromczyński, P. Waszkiewicz, *Biały wywiad w praktyce pracy organów ścigania na przykładzie wykorzystania serwisów społecznościowych*, „Prokuratura i Prawo" 2014, nr 5, s. 148–150.

2 *NATO Open Source Intelligence Handbook*, 2002, s. 5, http://www.au.af.mil/au/awc/awcgate/nato/osint_reader.pdf (dostęp: 21.01.2019).

3 Tamże.

4 Mogą to być jednakże metody, które są dopuszczalne podczas prowadzenia czynności wywiadowczych, takie jak podsłuchy czy kontrola korespondencji.

5 B. Stromczyński, P. Waszkiewicz, *Biały wywiad w praktyce…*, s. 150.

6 D. Mider, J. Garlicki, W. Mincewicz, *Pozyskiwanie informacji w Internecie metodą Google Hacking – biały, szary czy czarny wywiad?*, „Przegląd Bezpieczeństwa Wewnętrznego" 2019, nr 20, s. 68–91.

Artykuł niniejszy skupiać będzie się na osobach fizycznych niebędących pracownikami służb czy wywiadów.

Prawne granice pozyskiwania informacji

Jednym z problemów, jaki pojawia się w omawianym zakresie, są prawne granice pozyskiwania informacji. Analizując polski kodeks karny[7] (dalej jako kk) zauważyć można, że przestępstwa z wykorzystaniem komputera do pozyskiwania danych opisane zostały w większości w rozdziale dotyczącym przestępstw przeciwko ochronie informacji. I tak artykuł 267 kk stanowi: „§ 1. Kto bez uprawnienia uzyskuje dostęp do informacji dla niego nieprzeznaczonej, otwierając zamknięte pismo, podłączając się do sieci telekomunikacyjnej lub przełamując albo omijając elektroniczne, magnetyczne, informatyczne lub inne szczególne jej zabezpieczenie, podlega grzywnie, karze ograniczenia wolności albo pozbawienia wolności do lat 2. § 2. Tej samej karze podlega, kto bez uprawnienia uzyskuje dostęp do całości lub części systemu informatycznego. § 3. Tej samej karze podlega, kto w celu uzyskania informacji, do której nie jest uprawniony, zakłada lub posługuje się urządzeniem podsłuchowym, wizualnym albo innym urządzeniem lub oprogramowaniem".

Cytowany przepis jasno wskazuje, że próba uzyskania informacji poprzez łamanie czy też omijanie zabezpieczeń podlega karze. Obok tradycyjnych form nielegalnego uzyskiwania informacji (takich jak na przykład nieuprawnione otwieranie cudzej korespondencji) ustawodawca zauważył potrzebę włączenia do katalogu również nielegalnego uzyskania informacji poprzez przełamanie systemu informatycznego. Ponadto w polskim prawie karalne jest także wykorzystywanie podsłuchów lub oprogramowania pozwalającego na obserwowanie innej osoby. Ustawodawca polski penalizuje też samą ingerencję w zamieszczone dane. Stanowią o tym kolejne artykuły kk: „Art. 268. § 1. Kto, nie będąc do tego uprawnionym, niszczy, uszkadza, usuwa lub zmienia zapis istotnej informacji albo w inny sposób udaremnia lub znacznie utrudnia osobie uprawnionej zapoznanie się z nią, podlega grzywnie, karze ograniczenia wolności albo pozbawienia wolności do lat 2. § 2. Jeżeli czyn określony w § 1 dotyczy zapisu na informatycznym nośniku danych, sprawca podlega karze pozbawienia wolności do lat 3. […] Art. 268a. § 1. Kto, nie będąc do tego uprawnionym, niszczy, uszkadza, usuwa, zmienia lub

[7] Ustawa z 6 czerwca 1997 roku – Kodeks karny t.j. Dz.U. z 2018 r., poz. 1600, ze zm.

utrudnia dostęp do danych informatycznych albo w istotnym stopniu zakłóca lub uniemożliwia automatyczne przetwarzanie, gromadzenie lub przekazywanie takich danych, podlega karze pozbawienia wolności do lat 3. […]".

Warto zwrócić uwagę na fakt, że krajowy prawodawca przewiduje surowszą odpowiedzialność w przypadku ingerencji w informatyczny nośnik danych, co podkreśla, jak istotne znaczenie ma tego rodzaju przestępstwo dla prawidłowej ochrony informacji.

Zgodnie z art. 269a kk również czynności prowadzące do zakłócania działania systemu informatycznego uznane zostały za nielegalne i mogą podlegać karze. Warto wspomnieć, że wprowadzeniu wskazanych wyżej przepisów nie towarzyszyło wyłączenie odpowiedzialności w wyjątkowych przypadkach, takich jak na przykład nieuprawniony dostęp do systemu informatycznego w celu testowania efektywności wprowadzonych zabezpieczeń. Błąd ten jednak naprawiony został w art. 269b § 1a kk („Nie popełnia przestępstwa określonego w § 1, kto działa wyłącznie w celu zabezpieczenia systemu informatycznego, systemu teleinformatycznego lub sieci teleinformatycznej przed popełnieniem przestępstwa wymienionego w tym przepisie albo opracowania metody takiego zabezpieczenia), a także w art. 269c kk („Nie podlega karze za przestępstwo określone w art. 267 § 2 lub art. 269a, kto działa wyłącznie w celu zabezpieczenia systemu informatycznego, systemu teleinformatycznego lub sieci teleinformatycznej albo opracowania metody takiego zabezpieczenia i niezwłocznie powiadomił dysponenta tego systemu lub sieci o ujawnionych zagrożeniach, a jego działanie nie naruszyło interesu publicznego lub prywatnego i nie wyrządziło szkody"). Wprowadzenie obu przepisów było konieczne dla skutecznego działania osób zajmujących się „legalnym hackingiem", a więc działaniem zmierzającym do poprawy bezpieczeństwa informatycznego podmiotów publicznych oraz firm prywatnych.

Warto podkreślić, że wszystkie omówione przepisy stanowią przestępstwa powszechne, a więc takie, które popełnione mogą być przez każdego człowieka. Oznacza to, że ustawodawca wskazał prawną granicę dopuszczalności pozyskiwania informacji w sposób legalny. Powyższe przepisy nie dają jednakże odpowiedzi na wszystkie sytuacje związane z możliwością pozyskiwania danych. Działania takie jak włamania do systemu informatycznego czy przełamywanie zabezpieczeń kryptograficznych nie budzą wątpliwości co do ich nielegalności. Wydaje się, że ustawodawca ustosunkował się również negatywnie do sytuacji pozyskiwania informacji na przykład w drodze uzyskania dostępu do systemu informatycznego w wyniku wejścia w posiadanie hasła zdobytego na skutek zastosowania

socjotechnik (art. 267 § 2 kk). Wskazuje on bowiem, że uzyskanie dostępu do systemu informatycznego przez osobę nieuprawnioną jest czynnością podlegającą odpowiedzialności karnej. W przeciwieństwie do § 1, kolejny § 2 art. 267 kk nie wymienia jednak konkretnych działań, które muszą zostać podjęte, aby nieuprawniony dostęp stał się karalny. Oznacza to penalizowanie samej sytuacji uzyskania nieuprawionego dostępu do danych bez względu na metody wykorzystane do osiągnięcia celu. W tym kontekście warto zwrócić dodatkowo uwagę na art. 287 § 1 kk stanowiący, że: „Kto, w celu osiągnięcia korzyści majątkowej lub wyrządzenia innej osobie szkody, bez upoważnienia, wpływa na automatyczne przetwarzanie, gromadzenie lub przekazywanie danych informatycznych lub zmienia, usuwa albo wprowadza nowy zapis danych informatycznych, podlega karze pozbawienia wolności od 3 miesięcy do lat 5".

Przepis ten penalizuje między innymi przestępstwo phishingu, a więc działania polegającego na skopiowaniu strony internetowej określonej instytucji (na przykład banku) w celu przekonania podmiotu korzystającego z witryny o jej prawdziwości i uzyskaniu w ten sposób pożądanych informacji. Formą phishingu może być również przesłanie wiadomości e-mailem z podaniem fałszywego nadawcy – podszywanie się pod podmiot, do którego odbiorca ma zaufanie (na przykład administrator poczty, instytucja państwowa)[8]. Współczesną odmianą phishingu jest tzw. *spear phishing* polegający na celowym doborze osób, które mają zostać ofiarami oszustwa[9]. Typowanie potencjalnych ofiar odbywa się na podstawie przeprowadzonego przez oszusta wywiadu środowiskowego (nierzadko opartego na białym wywiadzie).

Zgodnie ze statystykami umieszczonymi na stronie policji, w Polsce liczba przestępstw zawierających znamiona tzw. oszustwa komputerowego z roku na rok rośnie.

Warto zwrócić uwagę, że statystyka dotyczy jedynie postępowań wszczętych, a więc odnosi się tylko do tych przypadków, w których użytkownik zgłosił próbę oszustwa organom ścigania. Pogląd na to, jak dużą skalę mogą osiągnąć ataki phishingowe, może dać materiał przygotowany przez firmę Kaspersky Lab. W raporcie dotyczącym pierwszego kwartału 2018 roku podkreślone zostało, że tylko we wskazanym czasie program antyphishingowy Kaspersky Lab zapobiegł 90 245 060 próbom

8 Hasło: *Phishing*, „Słownik Języka Polskiego", https://sjp.pl/phishing (dostęp: 12.11.2018).

9 K. Jasiołek, *Spear phishing, czyli ataki spersonalizowane*, Komputer Świat, 13.08.2013, https://www.komputerswiat.pl/aktualnosci/bezpieczenstwo/spear-phishing-czyli-ataki-spersonalizowane/m5th9v9 (dostęp: 12.11.2018).

przekierowania użytkownika do witryn osób przeprowadzających atak[10]. Oczywiście pamiętać należy, że jest to jedno z wielu oprogramowań komercyjnych służących do zabezpieczania komputera użytkownika, nie jest to więc pełen obraz problemu.

Tabela 1. Statystyki przestępstwa w postaci oszustwa komputerowego (art. 287 kk)

Rok	Liczba wszczętych postępowań	Liczba stwierdzonych przestępstw
2016	4103	4207
2015	4105	3282
2014	2567	2154
2013	1768	1573
2012	1285	1351
2011	1012	1364
2010	838	623
2009	673	978
2008	472	404
2007	322	492
2006	285	444
2005	326	568
2004	229	390
2003	219	168
2002	114	368
2001	59	171

Źródło: http://statystyka.policja.pl (dostęp: 12.02.2019)

Kolejną istotną kwestią związaną z korzystaniem z pozyskiwanych informacji jest możliwość wykorzystania ich w charakterze dowodu w postępowaniu karnym. Artykuł 168a kpk, wprowadzony do kodeksu postępowania karnego nowelą z 11 marca 2016 roku[11], stanowi, że „Dowodu nie można uznać za niedopuszczalny wyłącznie na tej podstawie, że został uzyskany z naruszeniem przepisów postępowania lub

[10] N. Demidova, T. Shcherbakova, M. Vergelis, *Spam and Phishing in Q1 2018*, Kaspersky.com, 23.05.2018, https://securelist.com/spam-and-phishing-in-q1-2018/85650/ (dostęp: 12.11.2018).

[11] Ustawa z 11 marca 2016 r. o zmianie ustawy – Kodeks postępowania karnego oraz niektórych innych ustaw, Dz.U. z 2016 r., poz. 437.

za pomocą czynu zabronionego, o którym mowa w art. 1 § 1 Kodeksu karnego, chyba że dowód został uzyskany w związku z pełnieniem przez funkcjonariusza publicznego obowiązków służbowych, w wyniku: zabójstwa, umyślnego spowodowania uszczerbku na zdrowiu lub pozbawienia wolności". Przepis ten oznacza, że nawet dowody pozyskane w drodze przestępstwa (na przykład włamania się do systemu informatycznego) mogą zostać legalnie wykorzystane w procesie karnym. Pamiętać jednak należy, że doktryna dopuszczenia tzw. owocu zatrutego drzewa nie wyklucza postawienia osobie zarzutów w związku ze sposobem zdobycia określonego dowodu.

Bardziej skomplikowana sytuacja w tym względzie rysuje się na gruncie postępowania cywilnego. Procedura ta warta jest w tym miejscu wspomnienia, gdyż nierzadko w toczących się przed sądem sprawach strony wykorzystują dowody zdobyte w sposób naruszający prawo (na przykład nagranie osoby bez jej zgody, wykorzystanie prywatnych e-maili czy informacji z prywatnej rozmowy toczącej się na portalu społecznościowym). W ramach procesu cywilnego nie ma zasady mówiącej o dopuszczalności lub niedopuszczalności dowodów zdobytych w sposób sprzeczny z obowiązującymi normami prawnymi. Dopuszczenie dowodu zależy zatem od rozważenia przez sąd określonych dóbr, a więc z jednej strony ustalenia prawdy materialnej pomocnej w rozstrzygnięciu sporu, a z drugiej naruszenia określonych praw jednostki (na przykład prawa do prywatności, prawa do tajemnicy korespondencji). O tym, że nie wyklucza się przez sąd wykorzystania „owocu zatrutego drzewa" świadczy chociażby wyrok Sądu Najwyższego z 23 kwietnia 2003 roku, w którym sędziowie uznali, że „nie ma zasadniczych powodów do całkowitej dyskwalifikacji kwestionowanego przez pozwaną dowodu z nagrań rozmów telefonicznych, nawet jeżeli nagrań tych dokonywano bez wiedzy jednego z rozmówców. Skoro strona pozwana nie zakwestionowała skutecznie w toku postępowania autentyczności omawianego materiału, to mógł on służyć za podstawę oceny zachowania się pozwanej w stosunku do pozwanego i możliwości sformułowania wniosku o nadużywaniu alkoholu przez pozwaną"[12]. Wskazany wyrok, dotyczący nagrań telefonicznych utrwalonych bez zgody jednej ze stron, przenieść można również na inne sytuacje, w których dowody zostały zdobyte z naruszeniem praw osób trzecich. Pamiętać jednakże należy, że niezależnie od przepisów prawnych wykorzystanie informacji pozyskanych w sposób niezgodny z prawem może zostać uznane za sprzeczne z zasadami etycznymi.

[12] Wyrok Sądu Najwyższego IV CKN 94/01.

Internetowe manipulacje – ochrona praw jednostki

Ze względu na postęp technologiczny i nowe sposoby wykorzystywania potencjału Internetu polski ustawodawca nie jest w stanie wprowadzić przepisów będących odpowiedzią na każdą sytuację rodzącą prawne wątpliwości w zakresie pozyskiwania informacji. Problem pojawia się na przykład podczas podszywania się pod inną osobę lub zakładania fałszywych kont w Internecie w celu bądź pozyskania określonych informacji, bądź ingerencji w nie. Ciekawym przykładem na polskim gruncie jest orzeczenie Sądu Rejonowego w Olsztynie z 21 lipca 2015 roku[13]. Oskarżoną w sprawie była kobieta, która chciała potwierdzić podejrzenia co do homoseksualnych skłonności swojego partnera. W tym celu wykorzystała znalezione wcześniej w Internecie zdjęcie przypadkowego mężczyzny. Wizerunek posłużył kobiecie do utworzenia fikcyjnego konta na portalu przeznaczonym do nawiązywania homoseksualnych kontaktów. Pod koniec 2014 roku mężczyzna, którego zdjęcie zostało wykorzystanie do założenia profilu, dowiedział się o tym fakcie i złożył zawiadomienie o popełnieniu przestępstwa. Kobiecie postawiono zarzut z art. 190a kk (tzw. stalking)[14], wskazując, że podszywanie się pod pokrzywdzonego i wykorzystanie jego wizerunku spowodowało wyrządzenie pokrzywdzonemu „[…] szkody osobistej poprzez przedstawienie go jako osoby o skłonnościach homoseksualnych poszukującej partnera seksualnego […]"[15]. We wskazanym przypadku sąd uniewinnił oskarżoną, uznając, że popełniony przez nią czyn nie wykazuje znamion przestępstwa z art. 190a, gdyż nie znała ona pokrzywdzonego i jej celem nie było wyrządzenie mu krzywdy. Sąd zatem nie uznał samego faktu wykorzystania czyjegoś wizerunku w celu założenia fikcyjnego konta za przestępstwo, które powinno być ścigane na podstawie przepisów prawa karnego, mimo iż mogła zostać naruszona reputacja pokrzywdzonego. Wskazany wyrok nie wyklucza jednak wystą-

[13] Wyrok Sądu Rejonowego w Olsztynie II K 497/15.

[14] Art. 190a kk: „§ 1. Kto przez uporczywe nękanie innej osoby lub osoby jej najbliższej wzbudza u niej uzasadnione okolicznościami poczucie zagrożenia lub istotnie narusza jej prywatność, podlega karze pozbawienia wolności do lat 3. § 2. Tej samej karze podlega, kto, podszywając się pod inną osobę, wykorzystuje jej wizerunek lub inne jej dane osobowe w celu wyrządzenia jej szkody majątkowej lub osobistej. § 3. Jeżeli następstwem czynu określonego w § 1 lub 2 jest targnięcie się pokrzywdzonego na własne życie, sprawca podlega karze pozbawienia wolności od roku do lat 10.§ 4. Ściganie przestępstwa określonego w § 1 lub 2 następuje na wniosek pokrzywdzonego".

[15] Wyrok Sądu Rejonowego w Olsztynie II K 497/15.

pienia z powództwem o naruszenie dóbr osobistych na gruncie prawa cywilnego[16].

Kwestia wyrządzania drugiej osobie szkody jest istotnym elementem przestępstwa podszywania się, który musiał wziąć pod uwagę sąd amerykański. Sprawa People v. Golb[17] dotyczyła nowojorskiego prawnika Ralpha Golba, który chcąc pomóc swojemu ojcu próbującemu rozpowszechnić jedną z głoszonych, a niedocenianych przez środowisko teorii naukowych, podszył się pod prof. Lawrence'a Schiffmana. Celem takiego postępowania było rozesłanie e-maili do pracowników naukowych, w których rzekomy prof. Schiffman przyznawał się do popełnienia plagiatu w stosunku do prac ojca R. Golba. Ponadto R. Golb stworzył kilka fikcyjnych tożsamości i na licznych blogach oskarżał prof. Schiffmana o kradzież pomysłów swojego ojca. Proces R. Golba toczył się kilka lat i ostatecznie zakończył skazaniem na karę dwóch miesięcy więzienia za przestępstwo podszywania się. Czyn, którego dopuścił się R. Golb, jest jedną z form tzw. *sock puppetry*, a więc działania polegającego na wykorzystywaniu fikcyjnej tożsamości w Internecie. Zasadniczo *sock puppet* ma znaczenie pejoratywne, gdyż wiąże się z tworzeniem alternatywnych kont (lub nadawaniem sobie pseudonimów) w celu oszukania osób korzystających z danego portalu (strony dyskusyjnej czy bloga)[18].

Obecnie definicja *sock puppet* rozszerzyła się, gdyż celem działania może być również manipulowanie opinią publiczną czy ominięcie zakazu administratora co do wypowiadania się na określonej stronie internetowej. Dla porządku warto wskazać, że pacynki, jak czasem nazywane jest opisywane to działanie w języku polskim, nie zawsze powinny być traktowane jako zjawisko negatywne. Na stronie regulującej zasady Wikipedii możemy przeczytać, że na przykład „Posiadanie wielu kont pozwala też na ochronę prywatności. Ktoś, kto jest znany publicznie lub w pewnym kręgu osób, może zostać rozpoznany z np. jego zainteresowań i dokonywanych edycji. Rozdział tych edycji między różne konta może pozwo-

[16] W tym celu można by powołać się na art. 23 kodeksu cywilnego w brzmieniu: „Dobra osobiste człowieka, jak w szczególności zdrowie, wolność, cześć, swoboda sumienia, nazwisko lub pseudonim, wizerunek, tajemnica korespondencji, nietykalność mieszkania, twórczość naukowa, artystyczna, wynalazcza i racjonalizatorska, pozostają pod ochroną prawa cywilnego niezależnie od ochrony przewidzianej w innych przepisach". Ustawa z 23 kwietnia 1964 r. – Kodeks cywilny (t.j. Dz.U. z 2019 r. poz. 1145).

[17] People v. Golb 23 N.Y.3d 455 (2014).

[18] „Sock puppet is a fake persona used to discuss or comment on oneself or one's work, particularly in an online discussion group or the comments section of a blog", WordSpy.com, https://wordspy.com (dostęp: 6.08.2018).

lić zachować anonimowość. Niektórzy użytkownicy, szczególnie administratorzy i biurokraci, używają pacynek, kiedy pracują na nie swoim komputerze, w celu uniknięcia przejęcia swojego głównego konta (często dysponującego narzędziami administracyjnymi) przez kradzież hasła"[19]. W większości przypadków alternatywne konta uważane są jednak za niezgodne z polityką działania portali internetowych czy stron dopuszczających do dyskusji pomiędzy użytkownikami platformy. Jednakże także w Wikipedii, mimo dostrzeżenia pewnych korzyści z wykorzystywania pacynek, dotyczące tegoż zasady zaczynają się od stwierdzenia: „Odradza się używania pacynek, gdyż może to doprowadzić do naruszania zasad ustalonych w Wikipedii, takich jak jedna osoba–jeden głos"[20]. Popularny w Polsce portal społecznościowy Facebook również wskazuje, że „Podstawą naszej społeczności jest autentyczność. Uważamy, że ludzie czują się bardziej odpowiedzialni za swoje wypowiedzi i działania, gdy znana jest ich prawdziwa tożsamość. Dlatego wymagamy od użytkowników Facebooka używania imion i nazwisk, którymi posługują się na co dzień"[21]. W innym miejscu tego samego dokumentu możemy przeczytać, że jedną z czynności niedozwolonych jest zakładanie fałszywych kont oraz podszywanie się pod inne osoby[22].

Przytoczone polityki idą zatem dalej niż uczynił to polski ustawodawca, zakazując lub znacznie ograniczając wykorzystywanie *sock puppets* bez względu na zamiary, w jakich tworzy się alternatywne konta. Rozwiązanie takie wydaje się słuszne, gdy bierzemy pod uwagę cel, jakim jest tworzenie zaufania pomiędzy użytkownikami portali społecznościach czy projektów takich jak Wikipedia. W tym kontekście ciekawie przedstawia się amerykańska sprawa United States v. Drew[23], w której jednym z zarzutów postawionych oskarżonej było właśnie złamanie zasad statutu serwisu społecznościowego MySpace poprzez założenie fikcyjnego konta na portalu. Lori Drew pragnęła w ten sposób zawrzeć znajomość internetową z koleżanką swojej nastoletniej córki – Megan Meier. Dowiedziała się bowiem, że M. Meier niepochlebnie wypowiada się na temat jej dziecka. Oskarżona założyła więc fikcyjne konto, podając się

---

19 Hasło: *Pacynka*, Wikipedia Wolna Encyklopedia, https://pl.wikipedia.org/wiki/Wikipedia:Pacynka (dostęp: 6.08.2018).

20 Tamże.

21 Facebook, Standardy Społeczności, *Integralność i autentyczność*, https://pl-pl.facebook.com/communitystandards/integrity_authenticity (dostęp: 6.08.2018).

22 Facebook, Centrum Pomocy, *Podszywanie*, https://pl-pl.facebook.com/help/212826392083694?helpref=search&sr=1&query=podszywanie (dostęp: 6.08.2018).

23 U.S. v. Drew, 259 F.R.D. 449 (C.D. Cal. 2009).

za 16-letniego Josha Evansa i zaczęła korespondować z M. Meier. Po pewnym czasie rzekomy Josh zerwał znajomość, stwierdzając, że „[…] świat byłby lepszy bez Megan”[24]. Tego samego dnia M. Meier popełniła samobójstwo. Ważnym elementem procesu było rozstrzygnięcie kwestii, czy zachowanie L. Drew w postaci złamania zasad serwisu MySpace było przestępstwem federalnym w rozumieniu Computer Fraud and Abuse Act (CFAA)[25]. Ostatecznie sąd uznał, że w tym przypadku wyrok niekorzystny dla oskarżonej oznaczałby danie możliwości serwisom internetowym zakwalifikowania określonych zachowań do kategorii przestępstw. Byłoby to *de facto* przyznanie uprawnień ustawodawczych twórcom portali, co nie jest zgodne z postanowieniami amerykańskiej konstytucji.

Wszystkie opisane sprawy pokazują wyraźnie, że samo wykorzystanie *sock puppets* nie jest karane na podstawie prawa karnego powszechnie obowiązującego. Znaczenie ma w tym przypadku jednakże zamiar, w jakim ktoś używa pacynki, pozyskując lub manipulując informacją. Tylko bowiem w takim przypadku można zidentyfikować znamiona określonych przestępstw (na przykład oszustwa czy stalkingu). Brak regulacji prawnej w tym zakresie nie wyklucza jednakże wprowadzania zakazów stosowania *sock puppets* przez serwisy czy portale internetowe, które uważają takie zachowania za nieetyczne. Konsekwencją łamania zasad może być jednakże jedynie zablokowanie dostępu do określonych treści.

Dopuszczalność pozyskiwania i przetwarzania prywatnych danych

Kwestią wartą rozważenia w kontekście pozyskiwania danych w Internecie jest możliwość uzyskania informacji o pracowniku (lub kandydacie na pracownika) przez pracodawcę (lub potencjalnego pracodawcę). Nie jest bowiem tajemnicą, że firmy już podczas procesu rekrutacji coraz częściej poszukują dodatkowych informacji o osobie ubiegającej się o określone stanowisko zawodowe. Na pytanie o dopuszczalność sięgania po prywatne informacje o pracowniku (na przykład informacje dostępne na portalu społecznościowym) w dużej mierze odpowiada opinia tzw. Grupy

[24] U.S. v. Drew, 259 F.R.D. 449 (C.D. Cal. 2009), tłumaczenie własne.
[25] 18 U.S. Code § 1030.

Roboczej art. 29[26] na temat przetwarzania danych w miejscu pracy[27]. Poruszono w niej szereg kwestii związanych z przetwarzaniem danych osobowych pracownika oraz kandydata na pracownika (na przykład monitoring w miejscu pracy, monitorowanie aktywności pracownika w sieci itp.).

Z perspektywy niniejszego artykułu ciekawy wydaje się zwłaszcza problem możliwości wykorzystania informacji „białowywiadowczych" o kandydacie na pracownika (na przykład w sytuacji, gdy pracownik ma publiczny profil na portalu społecznościowym). Twórcy opinii wskazują, że sięganie w procesie rekrutacji po dodatkowe informacje, które nie zostały udzielone przez samego kandydata, jest dopuszczalne, ale jedynie wtedy, gdy spełnione zostaną określone przesłanki. Pierwszą jest poinformowanie kandydata w ogłoszeniu o pracę o podejmowaniu przez rekrutującego wskazanych działań – nie mogą one mieć zatem charakteru niejawnego. Po drugie, sięganie po omawiany rodzaj danych musi mieć podstawę prawną, a więc pracodawca wykazać powinien, że ma uzasadniony interes w tym, żeby pozyskiwać dodatkowe informacje o kandydacie. Przykładem takiego prawnie uzasadnionego interesu jest sprawdzenie przez pracodawcę, czy osoby nie obowiązuje zakaz konkurencji ze względu na to, że wcześniej zatrudniona była w określonej firmie (informacje o poprzednim zatrudnieniu mogą być zamieszczane na różnego rodzaju portalach zawodowo-biznesowych, jak na przykład LinkedIn czy Goldenline). Po trzecie, aby dopuszczalne było pozyskanie informacji z portali społecznościowych o kandydacie, konieczne jest ustalenie, w jakim zakresie profil danej osoby powiązany jest z celami biznesowymi. Po wypełnieniu powyższych wymagań i przeprowadzeniu odpowiednich kontroli możliwości pozyskania danych pracodawca może zbierać określone informacje znajdujące się na profilach publicznych kandydata, nie jest jednakże uprawniony do domagania się udostępnienia mu profilu rekrutowanej osoby, na przykład poprzez zaakceptowanie

[26] Grupa robocza do spraw ochrony osób fizycznych w zakresie przetwarzania danych osobowych – tzw. Grupa Robocza art. 29, była niezależnym europejskim organem doradczym opiniującym zagadnienia związane z ochroną prywatności i danych osobowych. Została powołana na mocy art. 29 dyrektywy nr 95/46 Parlamentu Europejskiego i Rady z 24 października 1995 r. w sprawie ochrony osób fizycznych w zakresie przetwarzania danych osobowych i swobodnego przepływu tych danych. Grupa ta została rozwiązana 25 maja 2018 r., a w jej miejsce została powołana Europejska Rada Ochrony Danych.

[27] Opinia Grupy Roboczej art. 29 nr 2/2017 z 8 czerwca 2017 r. na temat przetwarzania danych w miejscu pracy; opinia dostępna na stronie: https://uodo.gov.pl/data/filemanager_pl/18.pdf (dostęp: 15.11.2018).

zaproszenia do grona znajomych. Wskazane zalecenia pokazują zatem wyraźnie, że Grupa Robocza art. 29 zdawała sobie sprawę z faktycznego korzystania przez pracodawców z możliwości pozyskiwania informacji w Internecie i starała się wprowadzić obostrzenia w tym zakresie prowadzące do ochrony prywatności pracownika lub osoby kandydującej na określone stanowisko w firmie[28].

Metadane

Innym problemem związanym z pozyskiwaniem informacji w Internecie jest zagadnienie metadanych. Metadane to dane o danych, a więc szczegółowe informacje opisujące zasoby informacji[29]. Dla przeciętnego obywatela metadane wydają się abstrakcyjnym pojęciem i zdają się nie mieć wpływu na jego codzienne życie. Okazuje się jednak, że pozyskanie metadanych może dać wiele informacji na temat miejsca przebywania osoby, jej zainteresowań czy nawyków. Najprostszym przykładem są dane, jakie wyczytać można ze zdjęcia niewyczyszczonego z metadanych. Poza takimi wskazówkami, jak sposób wykonania zdjęcia (na przykład przysłona czy czas naświetlania), które mogą być przydatne dla osób zainteresowanych fotografią, z obrazu pozyskać można również informacje dotyczące daty i miejsca wykonania zdjęcia. Wykorzystanie wskazanych metadanych może mieć zarówno pozytywne, jak i negatywne skutki. Pozytywne dotyczą między innymi posłużenia się informacjami ze zdjęcia w celu ujęcia osoby podejrzanej o popełnienie przestępstwa. Przykładem takiej sprawy jest ujęcie Johna McAfee'ego poszukiwanego przez policję w związku z zabójstwem sąsiada. Jego miejsce przebywania zostało ujawnione dzięki zdjęciu, którym jedna z gazet zilustrowała wywiad ze zbiegiem. Dziennikarz wykonujący fotografię smartfonem nie wyłączył danych geolokalizacyjnych, co pozwoliło na ujęcie J. McAfee'ego i postawienie go przed sądem[30]. Pozyskiwanie metadanych może mieć jednak również i negatywne skutki. W 2007 roku doszło do zniszczenia helikopterów klasy Apache stacjonujących w amerykańskiej bazie w Iraku. Zlokalizowanie bazy stało się możliwe dzięki metadanym znajdującym

[28] Tamże.

[29] Hasło: *Metadane*, Encyklopedia Zarządzania, https://mfiles.pl/pl/index.php/Metadane (dostęp: 15.11.2018).

[30] B. Weitzenkorn, *McAfee's Rookie Mistake Gives Away His location*, Scientific American, 4.12.2012, https://www.scientificamerican.com/article/mcafees-rookie-mistake/ (dostęp: 9.08.2018).

się w zdjęciach helikopterów zamieszczonych przez żołnierzy w Internecie[31]. Obecnie na stronie internetowej amerykańskiej armii można przeczytać, w jaki sposób unikać między innymi ujawniania informacji geolokalizacyjnych znajdujących się w fotografii[32].

Metadane oczywiście nie dotyczą jedynie zdjęć. Sama informacja o kilkukrotnym kontaktowaniu się z lekarzem określonej specjalności może – nawet bez wiedzy o treści rozmowy – być cenną wskazówką dotyczącą stanu zdrowia osoby. Podobne znaczenie będzie miała na przykład informacja o dzwonieniu na linię zaufania dla ofiar przemocy rodzinnej czy dla samobójców. Oczywiście uzyskanie akurat takich informacji wymaga dostępu do danych telekomunikacyjnych, które ujawniane są przez operatorów jedynie określonym podmiotom (art. 180d prawa telekomunikacyjnego[33]). Pozyskanie powyższych danych przez przeciętnego obywatela nie jest natomiast zgodne z prawem.

Część ujawnianych metadanych może mieć znaczenie nie tylko dla prawa jednostki do prywatności, ale również dla bezpieczeństwa publicznego. Przykładem takiego działania jest wykorzystywanie aplikacji rejestrujących aktywność sportową swoich użytkowników. Dobrą ilustracją opisanej sytuacji jest ujawnienie przez aplikację Strava mapy wskazującej trasy biegowe osób z niej korzystających. Analiza map pozwoliła na zlokalizowanie potencjalnych tajnych baz wojskowych[34]. Jak zatem widać z powyższych przykładów, metadane pozyskane w drodze legalnego, białego wywiadu mogą dostarczyć wielu informacji na temat miejsca przebywania osoby, jej nawyków czy sposobu spędzania wolnego czasu.

Internetowy biały wywiad a RODO

Pozyskiwanie podstawowych informacji w Internecie w ramach białego wywiadu nie stanowi obecnie większego problemu. Każda bowiem

---

[31] *Insurgents Destroyed US Helicopters Found in Online Photos*, „Live Science", 16.03.2012, https://www.livescience.com/19114-military-social-media-geotags.html (dostęp: 9.08.2018).

[32] Ch. Rodewig, *Geotagging Poses Security Risks*, 7.03.2012, https://www.army.mil/article/75165/geotagging_poses_security_risks (dostęp: 9.08.2018).

[33] Ustawa z 16 lipca 2004 r. Prawo telekomunikacyjne (t.j. Dz.U. z 2018 r. poz. 1954 ze zm.

[34] A. Hern, *Fitness Tracking App Strave Gives Away Location of Secret US Army Bases*, „The Guardian", 28.01.2018, https://www.theguardian.com/world/2018/jan/28/fitness-tracking-app-gives-away-location-of-secret-us-army-bases (dostęp: 9.08.2018).

aktywność podjęta w sieci pozostawia po sobie ślad, który może być przydatny do stworzenia profilu zawodowego lub prywatnego określonej osoby. Przetwarzanie danych osobowych od dawna znajduje się w kręgu zainteresowania unijnych instytucji[35]. Szybki rozwój Internetu sprawił, że nie tylko stare problemy w tym zakresie stały się bardziej widoczne, ale pojawiły się nowe, związane z gromadzeniem danych osobowych. Dlatego też prawodawca unijny wprowadził mechanizm broniący przed nadmiernym przetwarzaniem danych. Tym mechanizmem jest rozporządzenie Parlamentu Europejskiego i Rady (UE) 2016/679 z 27 kwietnia 2016 roku w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych i w sprawie swobodnego przepływu takich danych oraz uchylenia dyrektywy 95/46/WE (ogólne rozporządzenie o ochronie danych)[36] (dalej jako: RODO). Artykuł 17 RODO wprowadził tzw. prawo do bycia zapomnianym, które polega na możliwości zwrócenia się przez określoną osobę do administratora danych z żądaniem usunięcia przetwarzanych informacji dotyczących jednostki. W RODO wymienione zostały przesłanki, które powodują, że administrator musi dostosować się do prośby wnoszącego wniosek o usunięcie danych. Są to między innymi:

- zasada celowości – oznaczająca, że dane mogą być przetwarzane tylko do momentu, do którego jest to niezbędne ze względu na istotę gromadzenia danych; w sytuacji gdy zbieranie określonych informacji przestaje mieć pierwotny cel, administrator zobowiązany jest usunąć takie dane;
- cofnięcie zgody na przetwarzanie danych – dotyczy na przykład sytuacji, w której osoba najpierw wyraziła zgodę na przesyłanie jej informacji marketingowych, a po jakimś czasie takich informacji otrzymywać już nie chce;
- przetwarzanie danych od początku było niezgodne z prawem.

Prawodawca unijny pomyślał również o sytuacji, w której administrator danych dokonał upublicznienia zbieranych informacji (art. 17 ust. 2 RODO). W takim przypadku, w sytuacji skorzystania przez jednostkę z prawa do żądania usunięcia danych, administrator musi postarać się

[35] Patrz np. Dyrektywa 95/46/WE Parlamentu Europejskiego i Rady z 24 października 1995 r. w sprawie ochrony osób fizycznych w zakresie przetwarzania danych osobowych i swobodnego przepływu tych danych, Dz.Urz. z 1995 r., L 281.

[36] Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2016/679 z 27 kwietnia 2016 r. w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych i w sprawie swobodnego przepływu takich danych oraz uchylenia dyrektywy 95/46/WE (ogólne rozporządzenie o ochronie danych), Dz. Urz. z 2016 r., L119.

również o usunięcie danych przez podmioty trzecie. W art. 17 została w tym zakresie zastosowana klauzula „podejmowania rozsądnych działań, w tym środków technicznych". Wydaje się, że dopiero konkretne sprawy toczące się w przyszłości pozwolą doprecyzować znaczenie powyższej klauzuli. Prawo do bycia zapominanym nie jest jednakże prawem absolutnym. Istnieją bowiem sytuacje, wobec których prawodawca uznał, że żądanie usunięcia danych nie będzie musiało zostać uwzględnione przez administratora danych (art. 17 ust. 3 RODO). Są to między innymi przypadki, gdy:

- upublicznienie informacji jest konieczne do wywiązania się z prawnego obowiązku ciążącego na administratorze,
- informacje przechowywane są do celów statystycznych, badań naukowych lub historycznych,
- przetwarzanie informacji wiąże się z korzystaniem z prawa do wolności wypowiedzi i informacji.

Warto zauważyć, że szczególnie w przypadku ostatniej przesłanki realizacja prawa do bycia zapomnianym może stanowić bardzo trudne zadanie. Rozważenie bowiem, czy w danym przypadku ważniejszy jest interes konkretnej osoby, czy też interes publiczny polegający na uzyskaniu określonej informacji, nie jest w praktyce łatwe.

Pierwsza istotna sprawa, która rozpatrywana była przed Trybunałem Sprawiedliwości UE w 2014 roku (a więc jeszcze przed wprowadzeniem przepisów RODO) w zakresie prawa do zapomnienia dotyczyła obywatela Hiszpanii Mario Gonzaleza (Sprawa Google Spain SL, Google Inc. v Agencia Española de Protección de Datos, Mario Costeja González[37]). Żądał on usunięcia nieaktualnych już danych na jego temat, czego przedsiębiorstwo Google nie chciało uczynić, gdyż nie czuło się odpowiedzialne za treści zamieszczane w wyszukiwarce przez osoby trzecie. Sprawa dotyczyła zatem dwóch ważnych aspektów prawa do bycia zapomnianym. Pierwszy wiązał się z samą możliwością złożenia wniosku o usunięcie danych, która w czasie rozpatrywania sprawy nie była jeszcze uregulowana prawnie. W tym zakresie Trybunał Sprawiedliwości stwierdził, że: „[…] należy w szczególności przeanalizować kwestię, czy osoba, której dotyczą dane, ma prawo do tego, aby dana dotycząca jej informacja nie była już, w aktualnym stanie rzeczy, powiązana z jej imieniem i nazwiskiem poprzez listę wyświetlającą wyniki wyszukiwania mającego za punkt wyjścia to imię i nazwisko, przy czym stwierdzenie, iż takie prawo przysługuje, pozostaje bez związku z tym, czy zawarcie

[37] Wyrok Trybunału Sprawiedliwości UE C131/12.

na tej liście wyników wyszukiwania danej informacji wyrządza szkodę tej osobie. Ponieważ osoba ta może, ze względu na przysługujące jej i przewidziane w art. 7 i 8 karty prawa podstawowe, zażądać, aby dana informacja nie była już podawana do wiadomości szerokiego kręgu odbiorców poprzez zawarcie jej na takiej liście wyników wyszukiwania, prawa te są co do zasady nadrzędne nie tylko wobec interesu gospodarczego operatora wyszukiwarki internetowej, lecz również wobec interesu, jaki ten krąg odbiorców może mieć w znalezieniu rzeczonej informacji w ramach wyszukiwania prowadzonego w przedmiocie imienia i nazwiska tej osoby. Taka sytuacja nie ma jednak miejsca, jeśli ze szczególnych powodów, takich jak rola odgrywana przez tę osobę w życiu publicznym, należałoby uznać, że ingerencja w prawa podstawowe tej osoby jest uzasadniona nadrzędnym interesem tego kręgu odbiorców polegającym na posiadaniu, dzięki temu zawarciu na liście, dostępu do danej informacji"[38]. Takie postawienie sprawy dało początek dyskusjom na temat prawa do bycia zapomnianym, a w konsekwencji doprowadziło do wprowadzenia do RODO omawianego wcześniej art. 17. Niestety zarówno wyrok, jak i przepisy RODO dały jedynie pobieżne odpowiedzi co do kryteriów odmówienia osobie usunięcia informacji w ramach realizacji prawa do zapomnienia. Jest to zatem nadal sytuacja, którą powinno rozważać się w odniesieniu do konkretnego przypadku.

Drugi istotny aspekt poruszony w przytaczanym wyroku dotyczył odpowiedzialności ponoszonej przez przedsiębiorstwo Google za treści zamieszczane w wyszukiwarce przez osoby trzecie. W swoim orzeczeniu Trybunał Sprawiedliwości uznał, że właściciel wyszukiwarki jest odpowiedzialny za przetwarzane informacje, a tym samym, że jednostka ma prawo żądać usunięcia danych bezpośrednio od operatora wyszukiwarki.

* * *

Internet jest obecnie kopalnią informacji, które odpowiednio zinterpretowane mogą dostarczyć wiedzy na każdy interesujący poszukującego temat. Ze źródła tego korzystają zarówno podmioty publiczne, jak i pracodawcy czy osoby prywatne. Granice legalnego pozyskiwania danych reguluje prawo krajowe. Ze względu na szybki postęp technologiczny na rynku pojawia się jednakże coraz więcej narzędzi pozwalających na pozyskiwanie w sieci bardziej lub mniej „ukrytych" informacji. Rodzi to kolejne problemy natury nie tylko prawnej, ale również etycznej odnoszące się do kwestii pozyskiwania danych w Internecie. Wydaje się jed-

[38] Tamże.

nak, że jest to materia, której na chwilę obecną nie da się uregulować całościowo.

Streszczenie

Artykuł dotyczy możliwości pozyskiwania informacji w Internecie z perspektywy przepisów prawnych obowiązujących na terenie Polski. Poza niewątpliwymi korzyściami, jakie daje wykorzystanie komputera w życiu codziennym, rozwój technologiczny rodzi określone zagrożenia. Jednym z nich jest możliwość zdobywania informacji o jednostce za pomocą legalnych oraz nielegalnych metod działania. Celem artykułu jest próba analizy regulacji prawnych związanych z możliwością zbierania danych w Internecie.

Magda Tomaszewska

Legal Aspects of Obtaining Information on the Internet

Article deals with legal possibilities of obtaining information on the Internet. It indicates methods of obtaining information legally (such as open source intelligence) and methods recognized in Polish law as unlawful (such as activities aimed at disrupting the operation of an IT system). The article also discusses the issues of phishing, the use of false identity in the Internet (so-called sock puppetry), as well as the information potential of metadata analysis.

Key words: *infobrokering legal aspects, internet information retrieval, phishing, sock puppetry*

Bibliografia

Demidova N., Shcherbakova T., Vergelis M., *Spam and Phishing in Q1 2018*, Kaspersky.com, 23.05.2018, https://securelist.com/spam-and-phishing-in-q1-2018/85650/ (dostęp: 12.11.2018).

Hern A., *Fitness Tracking App Strave Gives Away Location of Secret US Army Bases*, „The Guardian", 28.01.2018, https://www.theguardian.com/world/2018/jan/28/fitness-tracking-app-gives-away-location-of-secret-us-army-bases (dostęp: 9.08.2018).

Jasiołek K., *Spear phishing, czyli ataki spersonalizowane*, Komputer Świat, 13.08.2013, https://www.komputerswiat.pl/aktualnosci/bezpieczenstwo/spear-phishing-czyli-ataki-spersonalizowane/m5th9v9 (dostęp: 12.11.2018).

Mider D., Garlicki J., Mincewicz W., *Pozyskiwanie informacji w Internecie metodą Google Hacking – biały, szary czy czarny wywiad?*, „Przegląd Bezpieczeństwa Wewnętrznego" 2019, nr 20.

Rodewig Ch., *Geotagging Poses Security Risks*, 7.03.2012, https://www.army.mil/article/75165/geotagging_poses_security_risks (dostęp: 9.08.2018).

Stromczyński B., Waszkiewicz P., *Biały wywiad w praktyce pracy organów ścigania na przykładzie wykorzystania serwisów społecznościowych*, „Prokuratura i Prawo" 2014, nr 5.

Weitzenkorn B., *McAfee's Rookie Mistake Gives Away His location*, Scientific American, 4.12.2012, https://www.scientificamerican.com/article/mcafees-rookie-mistake/ (dostęp: 9.08.2018).

STUDIA
POLITOLOGICZNE
VOL. 54
DOI: 10.33896/SPolit.2019.54.6

Konrad Gałuszko
ORCID: 0000-0003-1298-7765
Joanna Lewczuk
ORCID: 0000-0003-4812-6537
Konrad Krystian Kuźma
ORCID: 0000-0001-7159-6903

Walidować? Weryfikować? Nie ruszać? O niestatystycznych, statystycznych i stochastycznych metodach oceny jakości danych ilościowych opowieść

Słowa kluczowe:
ocena jakości danych, metody statystyczne, kontrola logiczna, testy statystyczne, rozkład normalny

Wstęp

Celem artykułu jest zaprezentowanie możliwych metod weryfikacji danych ilościowych z wykorzystaniem narzędzi niestatystycznych, statystycznych oraz stochastycznych. I choć najprostszymi z nich są metody zdroworozsądkowe, oparte na pewnych założeniach logicznych, to jednak nie mają one umocowania w matematyce, a co za tym idzie – ich wyniki mogą być podważane przez naszych adwersarzy[1].

Nieco inaczej jest w przypadku metod statystycznych, które swoimi korzeniami sięgają starożytnego Egiptu. Już wtedy przeprowadzano pierwsze spisy powszechne[2], a ich wyniki prezentowano w wygodnym

[1] Nie oznacza to jednak, że są to metody nieużyteczne, każdą metodę bowiem można spróbować podważyć. Choć jest to trudniejsze w przypadku metod statystycznych i stochastycznych, nadal pozostaje możliwe.

[2] Zob. D. Valbelle, *Les recensements dans l'Egypte pharaonique des troisième et deuxième millénaires*, „CRIPEL" 1987, nr 9, s. 37–49; P. Cartledge, P. Garnsey, E.S. Gruen (red.), *Hel-*

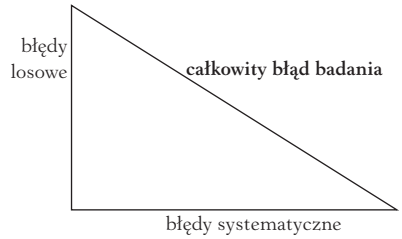
układzie tabelarycznym zawierającym podstawowe informacje o stanie państwa. Do czasów Pierre'a de Fermata oraz Blaise'a Pascala minęło wiele lat. W XVII wieku obaj badacze rozpoczęli analizę gier losowych, co pozwoliło na znaczne rozszerzenie dostępnego katalogu metod statystycznych. Jednak współczesną statystykę, w jej obecnym kształcie, zawdzięczamy w dużej mierze Andriejowi Kołmogorowowi, który przed II wojną światową dokonał aksjomatyzacji teorii prawdopodobieństwa.

Wraz z rozwojem metod statystycznych badacze zaczęli zastanawiać się, w jaki sposób można je ulepszyć. Wynikiem ich działań są metody stochastyczne, które oprócz czynników głównych dają również możliwość – w większym stopniu niż wspomniane wcześniej metody statystyczne – uwzględnienia czynników losowych w celu lepszego poznania badanych zależności[3].

Przedstawione w ramach artykułu egzemplifikacje są najbliższe sercu autorów, ale prezentowane metody są pewną prawdą uniwersalną i – jako takie – mogą być zastosowane do większości już istniejących (i nowo tworzonych) zbiorów danych ilościowych. Przedstawiane przez autorów metody są bardzo dobrze rozpoznane, przetestowane i mają bogatą literaturę opisującą ich działania. Jako takie – dysponują również licznymi narzędziami analitycznymi, a także są zaimplementowane w wielu dostępnych programach umożliwiających analizę statystyczną.

Najczęściej ocena jakości danych ilościowych jest częścią problemu związanego z błędami popełnianymi w trakcie realizacji badania. Na błąd całkowity składają się błędy losowe i błędy systematyczne. Relacje między tymi błędami przedstawia rysunek 1.

Rysunek 1. Zależności między elementami składowymi całkowitego błędu badania



Źródło: opracowanie własne na podstawie: G. Lissowski, *Z zagadnień doboru próby*, [w:] K. Szaniawski (red.), *Metody statystyczne w socjologii: wybrane zagadnienia*, Warszawa 1968, s. 68.

lenistic Constructs: Essays in Culture, History, and Historiography, Berkeley – Los Angeles – Londyn 1997, s. 242.

[3] F. Bławat, *Podstawy analizy ekonomicznej. Teorie, przykłady, zadania*, Warszawa 2011, s. 35.

Ujęcie tego problemu zaprezentowane na rysunku 1 po raz pierwszy przedstawiono na gruncie polskiej socjologii w dziele, na podstawie którego sporządzono rysunek[4]. Zgodnie z tym ujęciem błędy losowe dotyczą doboru próby i są całkowicie możliwe do obliczenia (skwantyfikowania). Błędy systematyczne mają z kolei dwoistą naturę: mogą wynikać zarówno z procedury wybierania próby (na przykład wadliwego operatu losowania), jak i elementów badania niezwiązanych z samą próbą. Klemens Szaniawski wymienia tu trzy źródła tego podtypu błędu systematycznego:

- występowanie jednostek niedostępnych[5];
- błąd mierzenia – chodzi tu głównie o błędy odpowiedzi (*errors in response*), których źródłem może być respondent lub ankieter;
- błędy proceduralne (kodowania, obliczania itp.)[6].

Tematyka podjęta w tym artykule odnosi się zatem do błędów systematycznych w badaniach, a uściślając: błędów mierzenia, których źródłem jest przede wszystkim ankieter, a także błędów proceduralnych. Błędy systematyczne, zgodnie z K. Szaniawskim, można ograniczyć między innymi „przez zastosowanie lepszej procedury badawczej: […] podniesienie jakości pracy terenowej czy też kodowania danych […]”[7].

---

[4] Nie są to jedyne koncepcje typologii błędów, ich źródeł i relacji między nimi. Więcej informacji na ten temat: P. Jabkowski, *Reprezentatywność badań reprezentatywnych. Analiza wybranych problemów metodologicznych oraz praktycznych w paradygmacie całkowitego błędu pomiaru*, Poznań 2015, s. 21–46, a także F. Sztabiński, *Ocena jakości danych w badaniach surveyowych*, Warszawa 2011, s. 53–60. Celem porównania warto dodać, że na gruncie analiz weryfikacyjnych Jan Lutyński używa takich pojęć, jak: błąd netto, błąd brutto i błąd wyrównany. Są one swego rodzaju odwzorowaniem schematu podanego w publikacji pod redakcją K. Szaniawskiego *Metody statystyczne w socjologii: wybrane zagadnienia*, Warszawa 1968, ale dotyczą tylko weryfikacji zewnętrznej, które to pojęcie jest wytłumaczone w niniejszym tekście. Więcej informacji: J. Lutyński, *Analizy weryfikacyjne w badaniach z zastosowaniem wywiadu kwestionariuszowego, ich rodzaje i możliwości*, [w:] Z. Gostkowski, J. Lutyński (red.), *Studia pilotażowe i analizy weryfikacyjne*, Analizy i Próby Technik Badawczych w Socjologii, t. V, Wrocław 1975, s. 338–339.

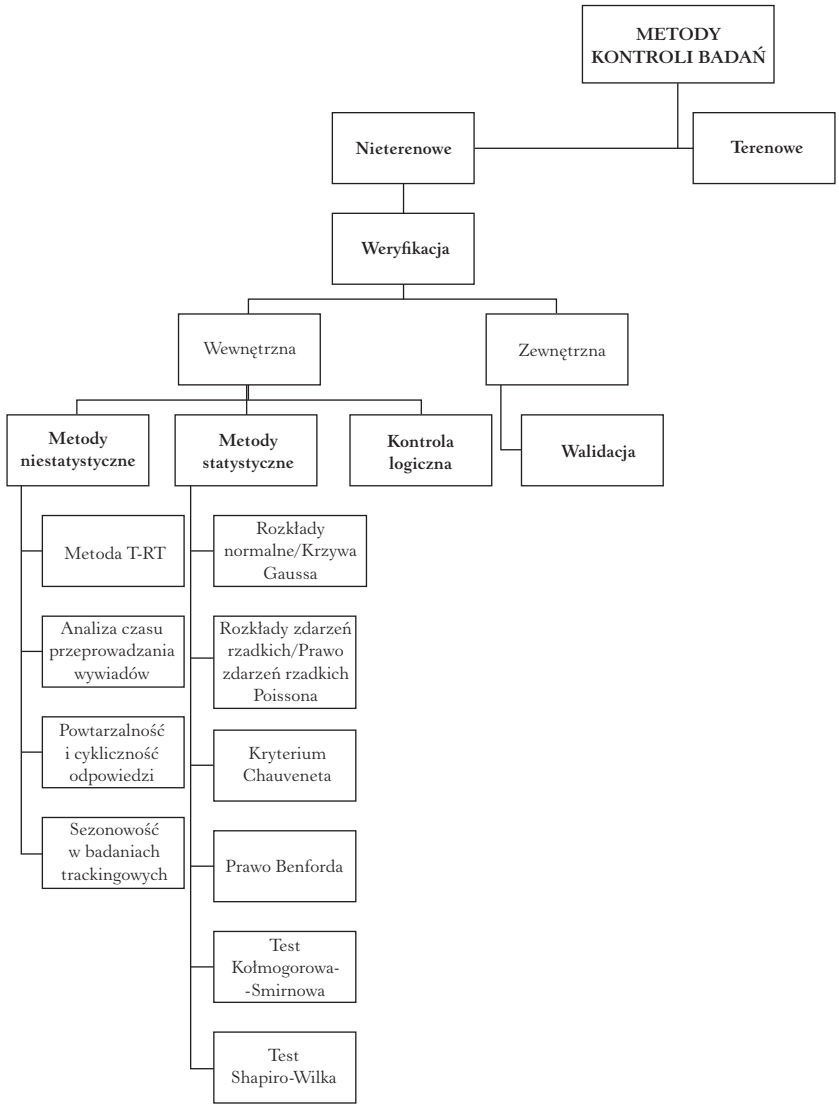
[5] Ten problem rozwinął w innej pracy pod redakcją K. Szaniawskiego socjolog Grzegorz Lissowski. Przy jego wyjaśnianiu posłużył się także tą samą klasyfikacją błędów co w *Metodach statystycznych w socjologii. Wybrane zagadnienia* (K. Szaniawski (red.), Warszawa 1968). Więcej informacji: G. Lissowski, *Problem jednostek niedostępnych w reprezentacyjnych badaniach socjologicznych*, [w:] K. Szaniawski (red.), *Metody matematyczne w socjologii. Zagadnienia wybrane*, Wrocław 1971, s. 7–34.

[6] K. Szaniawski (red.), *Metody statystyczne…*, s. 68. Z powyższego wynika, że błędu systematycznego – a co za tym idzie całkowitego błędu badania – nie można w pełni skwantyfikować.

[7] Tamże, s. 69.

Techniki oceny jakości danych przedstawione w tym artykule służą właśnie tym celom. Zależności między nimi przedstawia rysunek 2.

Rysunek 2. Zależności pomiędzy metodami kontroli badań (ocenami jakości danych)



Źródło: opracowanie własne na podstawie: F. Sztabiński, *Kontrola realizacji badania*, [w:] Z. Sawiński, F. Sztabiński, P.B. Sztabiński (red.), *Fieldwork jest sztuką. Jak dobrać respondenta, skłonić do udziału w wywiadzie, rzetelnie i sprawnie zrealizować badanie*, Warszawa 2005, s. 355–358; F. Sztabiński, *Ocena jakości danych w badaniach surveyowych*, Warszawa 2011, s. 97–113; H.M. Blalock, *Statystyka dla socjologów*, Warszawa 1977, s. 92–99, 232–234; M. Rószkiewicz, *Statystyka*, Warszawa 2002, s. 50–74.

Czym różni się walidacja od weryfikacji i kontroli logicznej? Te trzy pojęcia brzmią podobnie i osobie niewtajemniczonej w temat analizy wiarygodności danych może wydawać się, że mogą być stosowane wymiennie. Tak jednak nie jest, a wyjaśnienie niuansów między analizowanymi pojęciami będzie przedmiotem tej części artykułu. Istotne jest przy tym podkreślenie, że kontekstem omawianych treści, zarówno w tej części jak i następnej, będą wyłącznie dane ilościowe pozyskiwane w badaniach socjologicznych z użyciem wywiadów kwestionariuszowych[8].

Wszelkie metody kontroli wiarygodności danych opisane w części pierwszej i drugiej artykułu należą do tzw. metod nieterenowych. Ale metody statystyczne i stochastyczne[9] także zaliczają się do tej grupy. Kontrola nieterenowa[10] – w przeciwieństwie do terenowej – opiera się na

[8] Po 1989 roku, kiedy do Polski zaczęły docierać nowoczesne technologie służące obróbce danych ilościowych (sprzęt, ale przede wszystkim oprogramowanie różnego rodzaju), łatwiejsze stały się metody walidowania, weryfikowania i logicznej kontroli danych, ale przede wszystkim – bardziej precyzyjne. Można byłoby zatem założyć, że w polskich naukach społecznych (jak też na rynku badań ilościowych i jakościowych) dopiero od lat 90. XX w. poruszany jest problem walidacji, weryfikacji i kontroli logicznej. Jest to założenie błędne, gdyż pierwsze znaczące teksty i badania w tym zakresie powstawały już pod koniec lat 60. XX w. Szczególny wkład ma tutaj tzw. łódzka szkoła metodologiczna na czele z badaczami takimi jak Jan Lutyński, a w późniejszym okresie także Franciszek Sztabiński. Nie są to oczywiście wszyscy socjologowie z tej grupy – tutaj należy wymienić również Pawła Daniłowicza, Zygmunta Gostkowskiego, Jerzego Koniarka, Krystynę Lutyńską czy Pawła Sztabińskiego – natomiast to teksty przede wszystkim wcześniej wymienionych autorów są podstawą tej części artykułu. Więcej o łódzkiej szkole metodologicznej na stronie Instytutu Socjologii UŁ: Katedra Metod i Technik Badań Społecznych, *O katedrze*, http://instytutsocjologii.uni.lodz.pl/instytut/katedry-i-zaklady/katedra-metod-technik-badan-spolecznych/ [dostęp: 10.01.2019]. Z kolei najważniejsze teksty poruszające temat artykułu znajdują się w serii Analizy i Próby Technik Badawczych w Socjologii wydawanej przez Polską Akademię Nauk, Instytut Filozofii i Socjologii oraz Uniwersytet Łódzki, Instytut Socjologii, warto tu wspomnieć tom 13: *Analizy weryfikacyjne – przeszłe i obecne doświadczenia badawcze* (Łódź 2017) pod redakcją Katarzyny Grzeszkiewicz-Radulskiej i Anety Krzewińskiej; także w czasopiśmie „ASK. Research and Methods" (IFiS PAN); ponadto w publikacjach: Z. Sawiński, F. Sztabiński, P.B. Sztabiński (red.), *Fieldwork jest sztuką*, Warszawa 2005 oraz przywołanej już pracy: F. Sztabiński, *Ocena jakości danych w badaniach surveyowych*.

[9] Metody te, co więcej, zaliczają się do tzw. weryfikacji wewnętrznej, o czym będzie dalej mowa.

[10] Warto dodać, że metody kontroli służą także dwu celom: po pierwsze, poprawieniu danych lub ocenie jakości danych znajdujących się w bazie, ale także, po drugie, ocenie ankieterów uczestniczących w procesie pozyskiwania informacji. Artykuł ten skupia się jednak tylko na tym pierwszym kontekście, aczkolwiek warto podkreślić, że oba te cele się wzajemnie przenikają. Więcej informacji dotyczących drugiego celu można zna-

takiej analizie danych, która nie wymaga ponownego kontaktu z respondentem celem ustalenia, jakiej udzielił odpowiedzi. Ze względu na zakres omawianych pojęć weryfikacja jest zdecydowanie terminem najszerszym, dlatego od niej rozpoczniemy. Co więcej, walidację można uznać za specyficzną metodę tzw. weryfikacji zewnętrznej, a kontrolę logiczną – weryfikacji wewnętrznej.

Weryfikacja

Analizy weryfikacyjne – zgodnie z Janem Lutyńskim – można podzielić na dwa główne typy: wewnętrzną i zewnętrzną[11]. Typ pierwszy dostarcza wiedzy o procesach otrzymywania informacji za pomocą narzędzia badawczego. Innymi słowy w pierwszej kolejności ocenie podlega proces, a dopiero potem oceniane są same wyniki badania[12]. W przypadku drugiego typu weryfikacji kolejność jest odwrotna, to jest na podstawie danych uzyskanych inną techniką niż założona w badaniu (lub podczas innego pomiaru wykonywanego tą samą techniką) można ocenić, czy wybrana technika (pomiar) przyniosła wiarygodne wyniki. Warto przy tym nadmienić, że analizy weryfikacyjne oprócz informacji o metodologii mogą także pomóc przy ocenie wielkości błędów popełnionych w procesie pozyskiwania danych, a nawet poprawieniu tych błędów, to jest tzw. błędnych zaklasyfikowań[13].

Weryfikacja wewnętrzna jest metodą, którą trudniej stosować w ocenie jakości danych. Ponadto nie zawsze musi przynosić wyniki liczbowe (rozumiane jako liczba bądź odsetek błędnych zaklasyfikowań), a co za tym idzie – prowadzić do poprawek w zestawieniu danych. Nie oznacza to, że przy tym typie weryfikacji niemożliwe jest używanie technik

leźć szczególnie w publikacji: Z. Sawiński, F. Sztabiński, *Czy ankieterzy oszukują? Jak można to sprawdzić?*, [w:] Z. Sawiński, F. Sztabiński, P.B. Sztabiński (red.), *Fieldwork…*, s. 361–380.

11 J. Lutyński, *Analizy weryfikacyjne…*, s. 330. Należy dodać, że w zupełnie innym, znacznie szerszym kontekście użył tego terminu Stefan Nowak w swojej publikacji *Metody badań socjologicznych*. Zgodnie z nim celem badań weryfikacyjnych jest „sprawdzenie empiryczne, kontrola prawdziwości jednego bądź też całego zespołu ogólnych twierdzeń o związkach między pewnymi, ogólnie zdefiniowanymi klasami zjawisk. W badaniu takim rzeczywistość przebadana interesuje nas jedynie jako próba tej ogólnie zdefiniowanej klasy przedmiotów, do których odnieść się ma nasze twierdzenie". Więcej informacji: S. Nowak, *Metody badań socjologicznych*, Warszawa 1965, s. 191–193.

12 Tamże, s. 362; por. F. Sztabiński, *Ocena jakości…*, s. 79.

13 J. Lutyński, *Analizy weryfikacyjne…*, s. 340.

matematycznych, ponieważ te analizy weryfikacyjne dzielą się na dwa główne typy: niestatystyczne i statystyczne. O tych drugich będzie mowa w dalszych częściach niniejszego artykułu.

„Niestatystyczne metody weryfikacji wewnętrznej polegają […] na szacowaniu, na ile coś jest prawdopodobne”[14]. Pomocne przy takich weryfikacjach może być tworzenie tzw. normatywnych modeli formowania informacji[15]. Modele budować należy według tego, co J. Lutyński nazwał metodą „logiczną”[16]. Polega ona na określeniu przesłanek służących przyporządkowaniu danej jednostki badania do arbitralnie stworzonych przez badaczy (lub analityka) klas respondentów. Rysunek 3 przedstawia przykład zastosowania takiego uproszczonego modelu (pogrubieniem zaznaczono odpowiedzi respondenta).

Jak wykazuje rysunek 3, respondent udzielił odpowiedzi nietypowej, to jest niezgodnej z przewidywaniami badacza. Co bardzo ważne – jeśli założenie badacza nie zostało spełnione, nie oznacza to od razu, że w zbiorze danych pojawił się błąd. Oznacza to jedynie, że – nadal – korzystając z technik nieterenowych można sprawdzić, czy respondent ze względu na swoje cechy, najczęściej społeczno-demograficzne, udzielił odpowiedzi typowej dla respondentów zbliżonych parametrami. Metoda ta określana jest jako model najbliższego sąsiedztwa[17]. Jeśli natomiast wyniki takiej analizy wskażą, że odpowiedzi respondenta stanowią pewną anomalię to jest to wskazaniem do sprawdzenia reszty kwestionariuszy dostarczonych przez ankietera. W ostateczności konieczny może być ponowny kontakt z respondentem (telefoniczny lub inny, to jest wymagający skorzystania z terenowej metody kontroli[18]).

---

[14] W tym wypadku nie ma oczywiście mowy o prawdopodobieństwie matematycznym, a jedynie jego intuicyjnym rozumieniu. Por.: F. Sztabiński, *Ocena jakości…*, s. 79. Wynika z tego wniosek, że ten typ weryfikacji może być – lub wręcz powinien być – stosowany także przed fazą realizacji badania.

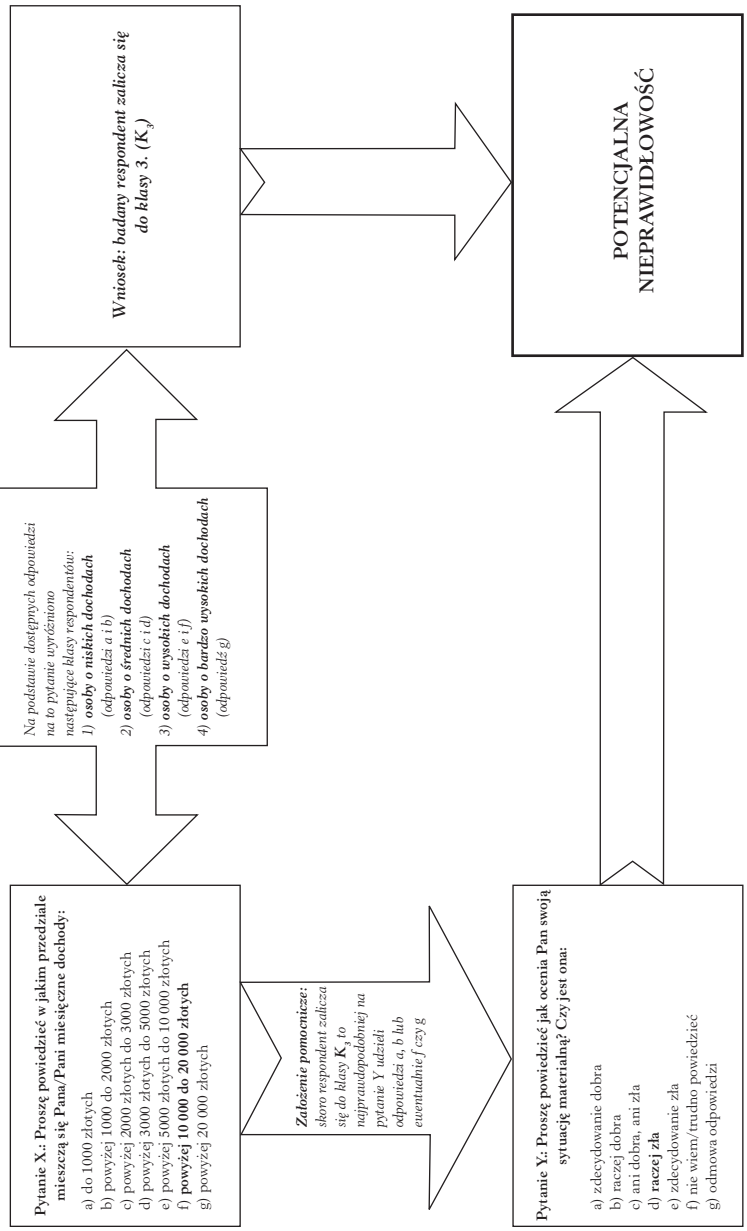
[15] J. Lutyński, *Analizy weryfikacyjne…*, s. 367–370.

[16] Tamże, s. 370–380.

[17] Model najbliższego sąsiedztwa jest modelem ekonometrycznym, który został stworzony przez Konrada Kuźmę w 2015 roku na podstawie – służącego do rozwiązywania problemu komiwojażera – algorytmu najbliższego sąsiedztwa. Model opiera się na analizie regresji. Więcej na ten temat można przeczytać np. tutaj: *StatSoft (2006). Elektroniczny Podręcznik Statystyki PL*, Kraków, https://www.statsoft.pl/textbook/stathome_stat.html (dostęp: 10.01.2019).

[18] Informacje na temat terenowych metod kontroli można znaleźć m.in. w: F. Sztabiński, *Kontrola realizacji badania*, [w:] Z. Sawiński, F. Sztabiński, P.B. Sztabiński (red.), *Fieldwork…*, s. 356–358 oraz: tenże, *Ocena jakości…*, s. 105–113.

Rysunek 3. Przykład normatywnego modelu formowania informacji
(Przegląd schematu należy rozpocząć od prostokąta z pytaniem X. Pogrubieniem zaznaczono odpowiedzi respondenta)



Źródło: opracowanie własne.

Weryfikacja zewnętrzna zakłada porównanie zbiorów danych, które różnią się od siebie znanym wcześniej (lub domniemanym) stopniem pewności (wiarygodności). Można pod tym względem wyróżnić trzy typy takich analiz[19]:

- wyniki ze źródła, do którego porównujemy inne rezultaty, są praktycznie pewne (przykładem może być porównanie deklarowanego wieku respondentów z ich danymi zapisanymi w bazie PESEL, a uzyskanymi podczas losowania próby na podstawie tego operatu);
- wyniki ze źródła, do którego porównujemy, są bardziej pewne niż wyniki z wywiadu[20] (taka analiza mogłaby się opierać na porównaniu deklaracji pracowników przedsiębiorstwa na temat spóźnień i wywiadów z ich kierownikami (bez użycia tzw. danych dokumentalnych[21], czyli w tym wypadku na przykład rejestru godzin wejść i wyjść);
- nie jest wiadome, jak pewne są rezultaty z obu źródeł[22] (ta analiza z kolei mogłaby dotyczyć preferencji partyjnych wyborców, przy czym jedno badanie byłoby oparte na próbie kwotowej, a inne na losowej).

Analizy weryfikacyjne można podzielić także w sposób przedstawiony w tabeli 1. W tym miejscu warto objaśnić pojęcia składowe w niej użyte. Metoda nie jest w tym wypadku równoznaczna z pomiarem, ponieważ analizy weryfikacyjne pierwszego i drugiego typu można wykonywać z użyciem tego samego narzędzia badawczego (por. z dalej opisaną metodą T-RT). Inaczej jest w przypadku typu trzeciego, którego istotą jest obliczanie korelacji[23] między wynikami z różnych źródeł[24]. Ważne jest

---

[19] Opracowanie własne na podstawie: J. Lutyński, *Analizy weryfikacyjne…*, s. 337.

[20] Konieczne jest w tym przypadku, jak i poprzednim, uzasadnienie, dlaczego jedno ze źródeł jest bardziej pewne od drugiego.

[21] Więcej na ten temat czytaj np. w: F. Sztabiński, *Ocena jakości…*, s. 89.

[22] Więcej informacji na temat porównywania wyników niepewnych: J. Lutyński, *Analizy weryfikacyjne…*, s. 346–358.

[23] Pomimo że J. Lutyński proponuje w swoim rozwiązaniu obliczanie korelacji, nie jest to jedyna metoda. Możliwym rozwiązaniem byłaby również regresja lub analiza wariancji z powtórzeniami. Więcej na ten temat można znaleźć np. w: *StatSoft (2006). Elektroniczny Podręcznik…*

[24] W tym wypadku dobrym przykładem są tzw. sondaże *exit poll*, czyli badanie kwestionariuszowe przeprowadzane z losowo dobranymi wyborcami w dniu wyborów. Badanie takie nie dotyczy deklaracji, jak respondent zamierza zagłosować (preferencji wyborczych), lecz tego, jak już zagłosował. Jego celem jest oszacowanie wyników głosowania i wyniki takiego badania są z reguły upubliczniane (np. przez stacje telewizyjne) tuż po zamknięciu lokali wyborczych, kiedy jeszcze nie są możliwe do obliczenia i rozpowszechnienia rzeczywiste rezultaty elekcji. Więcej informacji: R. Hilmer, *Exit Polls – A Lot More than Just a Tool for Election Forecasts*, [w:] M. Carballo, U. Hjelmar (red.), *Public Opinion Polling in a Globalized World*, Berlin 2008, s. 95. Jeśli wyniki takiego sondażu pokrywają

Tabela 1. Podział weryfikacyjnych analiz zewnętrznych bez uwzględnienia kryterium pewności źródeł

Numer typu	**Metody**[a]	**Powtórzony pomiar z użyciem tego samego narzędzia**	**Zbiorowość**	**Możliwość bezpośredniej identyfikacji respondentów**	**Próba**	**Dopuszczalne wnioski**
1a	Mogą być różne	Możliwy	Ta sama	**Wymagana**	**Musi być taka sama**	**Liczba błędnych zaklasyfikowań z możliwością ich identyfikacji**
1b				**Wykluczona**		**Liczba błędnych zaklasyfikowań**
2				Opcjonalna	**Ten sam typ (np. losowa), ale różne jednostki**	
3	**Muszą być różne**	**Niemożliwy**			Może być różna	**Tylko dotyczące pewności obu źródeł danych** (jedno z nich jest pewniejsze, a drugie mniej pewne)

[a] Opracowanie dotyczy tylko danych ilościowych (zgodnie z kontekstem opisanym w: J. Lutyński, *Analizy weryfikacyjne…*, s. 319–325), ale należy mieć na uwadze, że dane weryfikuje się także w badaniach jakościowych.

Źródło: opracowanie własne na podstawie J. Lutyński, *Analizy weryfikacyjne…*, s. 333–337.

tu ukryte założenie, że korelacja między różnymi pomiarami z użyciem tego samego narzędzia byłaby bardzo wysoka, o ile nie równa jeden[25].

W przypadku analiz typu pierwszego chodzi o co najmniej dwukrotny kontakt z dokładnie takimi samymi jednostkami badania. Interesujący jest tutaj problem identyfikacji respondentów, ponieważ przy założeniu, że ich identyfikacja jest zapewniona (1a), możliwa jest dokładna weryfikacja danych co do każdego przypadku (respondenta)[26]. Przy braku identyfikacji (1b) wiadomo jedynie, na ile różnią się pierwotne wyniki od powtórzonych, ale nie jest możliwe określenie, których badanych należy zaklasyfikować inaczej[27]. Natomiast w przypadku analizy drugiego typu jej istotą jest porównanie wyników uzyskanych z całkowicie różnych prób, dlatego, nawet jeśli możliwa jest identyfikacja respondentów, to wykluczony jest ponowny wywiad z nimi[28].

Opisywane trzy typy analiz dają pole do różnych wniosków. W najwyższym stopniu mogą podnieść jakość danych analizy typu 1a, ponieważ oprócz określenia skali rozbieżności (w liczbach czy procentach) możliwe jest – dzięki identyfikacji respondentów – poprawienie błędnych zaklasyfikowań z pierwszego pomiaru. Natomiast analizy typu 1b i 2 dostarczą tylko informacji, jaka jest wielkość niewłaściwych wskazań, ale nie będzie uprawnione na ich podstawie korygowanie danych w zbiorze. Najmniej użyteczne pod tym względem są analizy korelacyjne (typ 3), ponieważ na ich podstawie można jedynie ocenić, która z przyjętych

się w dużej mierze (czyli korelacja jest wysoka) z faktycznymi rezultatami wyborów, oznacza to, że rezultaty badania są wiarygodne. Jak pokazują sondaże *exit poll* prowadzone przez zespół Pawła Predki w TNS OBOP, a obecnie IPSOS Polska, wyniki są, szczególnie w przypadku wyborów prezydenckich, prawie identyczne jak wyniki PKW. Najbardziej widoczne było to w czasie przyspieszonych wyborów prezydenckich 2010 r., kiedy to różnica – w II turze – pomiędzy wynikami *exit poll* a PKW wynosiła 0,01%.

[25] Nie znaczy to jednak, że tak jest zawsze, o czym mowa w dalszej części tekstu. Co więcej, dokładnie to samo założenie stoi za analizami weryfikacyjnymi typu 1. i 2. (oczywiście tylko w wypadku użycia identycznego narzędzia).

[26] Choć jak pokazuje praktyka badawcza, nie zawsze tak się dzieje.

[27] Przykładem analizy opartej na próbie umożliwiającej identyfikację respondentów jest np. zapytanie o znajomość pisarza, a następnie o wymienienie kilku jego publikacji. W drugim przypadku można wyobrazić sobie wywiad kwestionariuszowy z pytaniem o znajomość pisarza, a następnie (kolejny pomiar) – podczas w pełni anonimowej ankiety audytoryjnej – pytanie o wymienienie kilku dzieł. W obu przypadkach podstawą badania są ci sami respondenci.

[28] Przykładem takiej analizy jest wykonanie w tym samym czasie badania dotyczącego znajomości marek produktów, opartego na dwóch różnych próbach losowych (to jest dwóch wylosowanych w ten sam sposób próbach, przy tym żaden z respondentów nie znajduje się w obu próbach jednocześnie).

metod jest mniej (lub bardziej) pewna oraz jaka jest skala i kierunek zmian w zbiorze danych w porównaniu jednej metody z drugą[29].

Na koniec należy powiedzieć, że wyniki z weryfikacji wewnętrznej i zewnętrznej nie muszą się ze sobą pokrywać[30]. Jeśli tak nie jest, oznacza to, że rezultaty badania należy dokładniej sprawdzić, używając innych sposobów.

Walidacja

Walidacja, jak wskazuje Franciszek Sztabiński za Hughiem Parrym i Helen Crossley, może być uznawana za formę weryfikacji wyników badania. W tym wypadku byłaby to weryfikacja dokumentalna[31], czyli porównanie danych faktualnych z oficjalnymi rejestrami urzędowymi uznawanymi za całkowicie pewne źródło informacji[32]. Innymi słowy, walidacja jest specjalnym typem weryfikacji zewnętrznej. W wymienionej definicji zawarte jest także istotne ograniczenie walidacji: tylko niektóre typy danych mogą podlegać tej metodzie[33]. Przykładowo, do pytań dotyczących postaw czy opinii nie będzie można jej zastosować.

Warto podkreślić, że w innej swojej publikacji F. Sztabiński definiuje weryfikację bez rozróżniania na podtypy, w sposób tożsamy z walidacją[34]. Ponadto w tej samej publikacji wyróżnia funkcję walidacyjną[35] jako jedną z funkcji kontroli realizacji badania. Jest to „ocena badania w kategoriach rzetelności i wiarygodności jego wyników"[36].

---

[29] F. Sztabiński, *Ocena jakości*…, s. 83–84.

[30] Dobrym zilustrowaniem takiej sytuacji jest dowolne badanie przeprowadzane na studentach, w którym muszą oni podać swój rok urodzenia. Można założyć, że jeśli ktoś jest studentem, to nie może być starszy niż x lat, ale nie może być młodszy niż y lat (x > y; weryfikacja wewnętrzna). Jeśli w zbiorze danych pojawi się nietypowy rezultat, to sięgając do oficjalnych rejestrów udostępnionych przez uczelnię można dokonać weryfikacji zewnętrznej i pozostawić albo usunąć daną obserwację z bazy. Jeśli natomiast konstrukcja próby umożliwiałaby ponowny bezpośredni kontakt z takim respondentem, to sięgając do terenowych metod kontroli, można byłoby także w ten sposób zweryfikować ów wynik. Więcej informacji: J. Lutyński, *Analizy weryfikacyjne*…, s. 387–392.

[31] F. Sztabiński, *Ocena jakości*…, s. 78; por. tamże, s. 85.

[32] Tamże, s. 78.

[33] Tamże, s. 85–86.

[34] F. Sztabiński, *Kontrola realizacji*…, s. 356.

[35] Tamże, s. 350–351.

[36] Tamże, s. 350. Rzetelność może być uznana za cechę jakości pomiaru i „mówi o tym, z jaką dokładnością zmierzyliśmy to, co zmierzyliśmy, lub dokładność z jaką test (bada-

Ten sam autor wskazuje, że istnieje także możliwość walidowania danych na podstawie rzeczywistego zachowania badanych[37]. Niemniej, celem uporządkowania pojęć, w niniejszym artykule przyjęto definicję walidacji zaproponowaną przez H. Parry'ego i H. Crossley. Schematyczny, uproszczony przykład procedury walidowania danych pokazują tabele 2 i 3.

Tabela 2. Pierwszy etap walidacji danych – przykładowe pytanie kontrolne

Pytanie Z: Przed nami już ostatnie pytanie. Jest to pytanie służące wyłącznie kontroli mojej pracy przez przełożonych, dlatego byłbym wdzięczny za udzielenie odpowiedzi. Proszę o podanie swojej dokładnej daty urodzenia i miejscowości, w której się Pan/Pani urodził(-a):	
Dzień (DD):	13
Miesiąc (MM):	07
Rok (RRRR):	87
Nazwa miejscowości urodzenia:	gdansk

Źródło: opracowanie własne.

Tabela 3. Drugi etap walidacji danych – porównanie danych wpisanych przez ankietera z danymi z operatu losowania

Rodzaj danych	**Metryczka respondenta nr xxx (z próby)**	**Metryczka kontrolna wpisana przez ankietera**
Dzień (DD):	13	13
Miesiąc (MM):	08	07
Rok (RRRR):	1987	87
Nazwa miejscowości urodzenia:	Gdańsk	gdansk

Źródło: opracowanie własne.

nie, narzędzie) mierzy to, co mierzy"; F. Sztabiński, *Ocena jakości*…, s. 71. Z kolei z pojęciem rzetelności jest mocno związane pojęcie trafności, które także może być kryterium jakości pomiaru: „mówi o tym, czy narzędzie mierzy to, co ma mierzyć lub z jaką dokładnością zmierzyliśmy to, co chcieliśmy zmierzyć"; tamże, s. 62. Wynika z tego, że trafność jest pojęciem szerszym niż rzetelność. Więcej informacji: tamże, s. 61–76.

37 Tamże, s. 89–90.

Tabela 4. Trzeci etap walidacji danych. Poprawienie danych wpisanych przez ankietera

ID respondenta	Dzień	Miesiąc	Rok	Nazwa miejscowości urodzenia
XXX	13	08	1987	Gdańsk

Źródło: opracowanie własne.

Jak widać na podanym przykładzie, ankieter, prócz pomijalnych błędów, dobrze wpisał dane respondenta, co każe sądzić, że – przynajmniej w rozmowie z tym badanym – rzetelnie wykonał swoją pracę.

Kontrola logiczna

Kontrola logiczna jest specyficznym typem weryfikacji wewnętrznej[38] i odnosi się do merytorycznej analizy wypełnionych kwestionariuszy[39]. Dalsza część opracowana została na podstawie materiałów wewnętrznych firm CBM INDICATOR oraz TNS OBOP. Kontrola ta dotyczy dwóch aspektów odpowiedzi respondentów, które – na podstawie naszego doświadczenia – nazwaliśmy[40]:

- aspektem formalnym: sprawdzanie wewnętrznej logicznej spójności odpowiedzi (na przykład ich zgodności z filtrami), jak również na przykład ustalenie, czy ankieter zadał wszystkie pytania zawarte w kwestionariuszu;
- aspektem nieformalnym: weryfikacja „zależności między odpowiedziami na różne pytania, które w mniejszym lub większym stopniu związane są z kryteriami losowania lub doboru"[41]. W tym przypadku ten proces jest niemal tożsamy z normatywnymi modelami formowania informacji.

Warto podkreślić, że pilotaż badania odpowiednio przeprowadzony przed fazą realizacji pomaga uniknąć lub znacznie ograniczyć później-

[38] Świadczy o tym przede wszystkim fakt, że nie korzysta się podczas niej z żadnych danych zewnętrznych.

[39] F. Sztabiński, *Ocena jakości*…, s. 101.

[40] Autorem nazw jest Konrad Gałuszko, pracownik firmy CBM Indicator i współautor artykułu. Nazwy zostały nadane, ponieważ w materiałach CBM Indicator nie określono ich dla tych dwóch aspektów kontroli logicznej.

[41] F. Sztabiński, *Ocena jakości*…, s.101.

sze błędy w zbiorze danych[42]. Należy zatem uznać, że pilotaż przed badaniem będzie dotyczyć przede wszystkim samego narzędzia badawczego, a kontrola logiczna po zebraniu danych – ocenie jakości danych czy ankieterów. Inaczej mówiąc, kontrola logiczna jest przydatna do wskazania, które podmioty badania – badacze, ankieterzy czy respondenci – są źródłem błędów. Jeśli badacze nie zauważyli błędów w kwestionariuszu podczas jego wieloetapowej analizy oraz pilotażu, oznacza to, że za późniejsze niezgodności w zbiorze danych są odpowiedzialni przede wszystkim badacze. W przypadku poprawnej konstrukcji, jeśli stwierdzono błędy (już po zakończeniu realizacji badania), oznacza to, że najprawdopodobniej ankieterzy – celowo lub przypadkowo – dokonali pomyłki. Ponadto, zdaniem niektórych badaczy, to na ankieterze spoczywa obowiązek stałego nadzorowania przebiegu wywiadu i sprawdzania, czy sam respondent udziela spójnych logicznie, najbardziej prawdopodobnych odpowiedzi. Aczkolwiek na przestrzeni lat założenie to uległo zmianie i obecnie w większości prowadzonych projektów odpowiedzi respondentów pozostawia się nienaruszone. Jeśli zatem kwestionariusz jest poprawny, a ankieter wiarygodny, oznacza to, że źródłem błędu jest najprawdopodobniej sam respondent[43].

Obecnie dane z badań są przechowywane głównie w formie elektronicznych baz danych, dlatego należy dodać, że kontrola logiczna jest związana z tzw. redundancją danych. Pojęcie to pochodzi z zakresu informatyki i oznacza „nadmiarowość; praktykę przechowywania więcej niż jednego elementu tych samych danych (w kilku kopiach)"[44]. Aby ograniczyć redundancję, stosuje się tzw. *data reduction plan*, czyli „plan redukcji danych; zestaw instrukcji dla edycji i kodowania kwestionariuszy i ankiet oraz analizy i weryfikacji udzielanych odpowiedzi i uzyskiwanych danych"[45]. Kontrola logiczna, jako działanie redukujące redundancję, pozwala

[42] Oczywiście obecnie programy i komputery uniemożliwiają w dużym stopniu popełnienie błędów formalnych – zarówno przed, jak i po realizacji badania. Należy zatem przyjąć, że w niniejszym tekście rozważania dotyczące formalnej kontroli logicznej dotyczą szczególnie badań PAPI.

[43] T. Pawłowski, *Logiczne podstawy weryfikacji wewnętrznej badań kwestionariuszowych*, [w:] Z. Gostkowski, J. Lutyński (red.), *Wywiad kwestionariuszowy w świetle badań metodologicznych*, Analizy i Próby Technik Badawczych w Socjologii, t. 4, Wrocław 1972, s. 277–278. Więcej informacji na ten temat także w: C. Blattman i in., *Measuring the Measurement Error: A Method to Qualitatively Validate Survey Data*, https://ore.exeter.ac.uk/repository/bitstream/handle/10871/31847/Measuring%20the%20measurement%20error.pdf?sequence=2&isAllowed=y (dostęp: 20.02.2019).

[44] M. Trojański, *Dictionary of Applied Informatics*, Warszawa 2007, s. 372.

[45] Tamże, s. 149.

po zakończeniu badania uzupełnić pewne braki lub usunąć nieuzasadnione nadmiary danych. Odwołując się do wcześniejszego przykładu z rysunku 1, badacze podczas weryfikacji wewnętrznej przed badaniem mogą uznać prawdopodobieństwo takiego układu odpowiedzi za niskie. Niemniej respondent miał prawo odpowiedzieć w taki sposób, tak więc decyzja o usunięciu takiej odpowiedzi lub całej obserwacji będzie efektem kontroli logicznej drugiego typu. Natomiast jeśli ankietowany odpowiedział na pytanie kwestionariusza, w sytuacji gdy jeden z filtrów nakazywał ominięcie następnego pytania, a mimo to respondent zaznaczył jakąś odpowiedź (nadmiar danych), jej usunięcie nie wpłynie na późniejsze rezultaty, a nawet poprawi ich wiarygodność (por. z częścią o powtarzalności i cykliczności odpowiedzi). Można wyobrazić sobie też sytuację, w której respondent w jednym miejscu kwestionariusza podaje swoją pełną datę urodzenia, a w innym miejscu pomija pytanie kontrolne o wiek w liczbach. Podczas kontroli danych można taki brak danych uzupełnić[46].

Metody niestatystyczne

Jak pisaliśmy wcześniej w części poświęconej weryfikacji, metody niestatystyczne bazują na intuicyjnej ocenie prawdopodobieństwa wystąpienia danego układu czy rozkładu odpowiedzi w całym badaniu i/lub w kwestionariuszach dostarczonych przez określonego ankietera[47]. W niniejszej części artykułu będą opisane następujące metody: T-RT, analiza czasu przeprowadzania wywiadów, powtarzalności i cykliczności odpowiedzi, a także sezonowości w badaniach powtarzalnych i ciągłych (w praktyce badawczej nazywanych trackingowymi). Należy mieć na uwadze, że metody statystyczne i niestatystyczne niekiedy się ze sobą

[46] Przy kontroli logicznej pomocne jest – oprócz normatywnych modeli formowania informacji – zdefiniowanie związków logicznych między poszczególnymi pytaniami i ich kafeteriami. T. Pawłowski wyróżnia następujące związki: 1. czyste: wynikanie logiczne, równoważność logiczna, sprzeczność, przeciwieństwo, podprzeciwieństwo; 2. eliptyczne: wynikanie eliptyczne, równoważność eliptyczna, sprzeczność eliptyczna. Niestety ze względu na ograniczoną objętość tekstu nie jest możliwe dokładne ich omówienie. Należy dodać, że przy określaniu związków logicznych pomocne mogą być także tabele zestawiające relacje między poszczególnymi pytaniami. Więcej informacji: T. Pawłowski, *Logiczne podstawy weryfikacji wewnętrznej badań kwestionariuszowych*, [w:] Z. Gostkowski, J. Lutyński (red.), *Wywiad kwestionariuszowy w świetle badań metodologicznych*, Analizy i Próby Technik Badawczych w Socjologii, t. 4, Wrocław 1972, s. 279–290.

[47] F. Sztabiński, *Ocena jakości*…, s.79; Z. Sawiński, F. Sztabiński, *Czy ankieterzy*…, s. 361–380.

zazębiają – na przykład aby wydać werdykt dotyczący jakości danych za pomocą techniki niestatystycznej, należy zastosować metodę statystyczną, ponieważ czasem zwykły „rzut oka" nie będzie wystarczający. W psychologii – skąd pochodzi to pojęcie – jest to działanie nazywane kwantyfikacją wyników. Dobrym przykładem mogą tu być analizy powtarzalności i cykliczności odpowiedzi, a także sezonowości w badaniach powtarzalnych i ciągłych.

METODA T-RT

Nazwa metody pochodzi od angielskiego skrótu oznaczającego *test-retest*[48]. Wymiennym określeniem jest „powtórzony pomiar". T-RT jest wskaźnikiem oceniającym między innymi rzetelność pomiaru[49], ale także narzędzia, za pomocą którego go dokonano, lub elementu tego narzędzia (na przykład skali)[50]. Korzenie tej metody można odnaleźć w naukach przyrodniczych, w których powtarzalność rezultatów jest jedną z wymaganych cech pomiaru lub narzędzia[51].

Istotą metody powtórzonego pomiaru jest stabilność wyników w czasie[52]. Należy podkreślić przy tym, że oba pomiary muszą być dokonane przy użyciu identycznego narzędzia badawczego. Jeśli wyniki uzyskane za jego pomocą są takie same lub zbliżone, oznacza to, że pomiar i/lub narzędzie są wiarygodne[53].

Metoda ta może jednak przynosić błędne, nieuzasadnione konkluzje, ponieważ jest wrażliwa na trzy następujące czynniki: upływ czasu między pomiarami, warunki zewnętrzne wpływające na respondentów[54], a także

---

48 Takie rozwinięcie tego skrótu pada m.in. w: H. Domański, A. Dukaczewska, *Stabilność odpowiedzi w badaniach socjologicznych*, „ASK. Research and Methods" 1996, nr 1, s. 74, a także w: F. Sztabiński, *Ocena jakości…*, s. 72.

49 F. Sztabiński, *Ocena jakości…*, s. 72.

50 H. Domański, A. Dukaczewska, *Stabilność odpowiedzi…*, s. 71.

51 F. Sztabiński, *Ocena jakości…*, s. 72.

52 W przypadku małych prób i/lub zjawisk o niskim rozpowszechnieniu stabilność ta może być zaburzona, dlatego zbadanie np. długofalowych trendów może przynieść błędne wnioski. Aby tego uniknąć, należy – oprócz zachowania odpowiednio długiego okresu kontynuacji badania (np. trackingu) – dokonać zabiegu tzw. agregacji podłużnej, zwanego też rolowaniem danych. Polega on na łączeniu ze sobą danych pochodzących z kilku kolejnych transz (punktów pomiarowych). Z. Sawiński, *Badania trackingowe*, [w:] D. Maison, A. Noga-Bogomilski, *Badania marketingowe. Od teorii do praktyki*, Gdańsk 2007, s. 112–114.

53 H. Domański, A. Dukaczewska, *Stabilność odpowiedzi…*, s. 71.

54 Dobrym przykładem w badaniach opinii może tu być tzw. zjawisko *agenda-setting*. Autorami tego pojęcia, jak i pierwszych studiów na ten temat są Marshall McCombs

tzw. zjawisko *pre-testu* i *post-testu*[55]. Z tej też przyczyny T-RT może być zastosowany tylko przy zachowaniu warunków zdefiniowanych przez Marka Stycznia, które w swojej pracy przytacza F. Sztabiński:

- „oba pomiary winny być dokonane w takim odstępie czasu, aby mierzone własności nie mogły ulec zmianie;
- dystans czasowy między tymi dwoma pomiarami winien być na tyle długi, aby na pomiar nie wpływały te same, niekontrolowane czynniki;

---

i Donald Shaw. Zgodnie z hipotezą *agenda-setting*, „media mają zdolność narzucania opinii publicznej przekonania co do ważności określonych tematów – kwestii, problemów, a to sprawia, że obywatele uważają za ważne te same tematy, o których najczęściej mówi się w mediach"; E. Nowak, *Teoria „agenda setting" a nowe media*, „Studia Medioznawcze" 2016, nr 3(66), s. 13.

55 H. Domański, A. Dukaczewska, *Stabilność odpowiedzi*…, s. 73–74. Zjawisko to oznacza wpływ samego badania, a zwłaszcza pierwszego pomiaru (pretestu) – a nie czynników zewnętrznych czy innych – na uzyskiwane odpowiedzi w dalszych pomiarach (posttestach). Innymi słowy na stabilność lub zmienność wyników w czasie może wpływać podczas procedury T-RT to, że respondenci „aktywizują się poznawczo: zwracają uwagę na pewne sprawy, skłaniają się do szukania nowych o nich [przedmiocie badania] informacji […]"; A. Sułek, *Eksperyment w badaniach społecznych*, Warszawa 1979, s. 76. Dobrym przykładem ilustrującym wpływ pretestu jest tzw. efekt panelowy albo „uczenie się" odpowiedzi. Pierwsze zjawisko przedstawił już w 1966 r. na gruncie socjologii polskiej Z. Gostkowski w pracy *Analiza „efektu panelowego" w badaniach wyborczych w Łodzi w r. 1961*. Badacz ten udowodnił, że udział respondenta w badaniu wpłynął na poziom zainteresowania wyborami, a także zmiany postaw, co udowodniły następne pomiary. Więcej informacji: Z. Gostkowski, *Analiza „efektu panelowego" w badaniach wyborczych w Łodzi w 1961 r.*, [w:] tegoż (red.), *Analizy i próby technik badawczych w socjologii*, t. 1, Wrocław – Warszawa – Kraków 1966, s. 282–310. Drugi efekt polega na tym, że respondenci celowo deklarują pewne istotne zmiany (choć w rzeczywistości one nie zaszły lub zajść w ogóle nie mogły – dotyczy to szczególnie danych dokumentalnych) albo – odwrotnie – dążą do zachowania spójności i stabilności swoich odpowiedzi (choć tu zmiana zaszła – np. w opiniach, postawach). Więcej informacji: A. Dyjas-Pokorska, *Badania trackingowe i panelowe*, [w:] Z. Sawiński, F. Sztabiński, P.B. Sztabiński (red.), *Fieldwork*…, s. 245–246. Na zaburzenie następnych pomiarów może wpłynąć dodatkowo sytuacja, gdy badani znają rezultaty pierwszego pomiaru: „Mogą one stanowić dla nich bardzo ważną informację, uświadamiającą im stopień rozbieżności między rzeczywistym stanem rzeczy a ich standardami i skłaniającą ich do działań zmierzających do ich uzgodnienia"; A. Sułek, *Eksperyment*…, s. 76–77. Więcej informacji na temat zjawiska pretestu i posttestu także w: F. Sztabiński, *Ocena jakości*…, s. 72; J. Brzeziński, *Metodologia badań psychologicznych*, Warszawa 2005, s. 314–316. Celem porównania warto także dodać, że w branży reklamowej zjawisko pretestu i posttestu podczas badań jest pożądane i jest źródłem cennych opinii dotyczących kampanii reklamowej. Więcej informacji: D. Maison, *O badaniach reklamy, czyli jak na podstawie badań przewidzieć skuteczność reklamy i ocenić skuteczność przeprowadzonej kampanii*, [w:] D. Maison, A. Noga-Bogomilski, *Badania marketingowe*…, s. 153–177.

- niezależność pomiaru, to znaczy pierwszy pomiar nie może wpływać na przebieg i wynik drugiego"[56].

ANALIZA CZASU PRZEPROWADZANIA WYWIADÓW

Analiza czasu przeprowadzania wywiadów jest, zgodnie ze stanowiskiem F. Sztabińskiego, jednym z elementów oceny pracy ankieterów i dotyczy poprawności realizacji wywiadu. Niemniej, jak wskazano wcześniej, sama kontrola ich pracy jest jedną z metod oceny wiarygodności wyników badania i jest jednym z czynników, które pomagają podjąć decyzję, czy informacje zebrane przez danego ankietera mogą się znaleźć w zbiorze, czy też powinny być usunięte. Należy wspomnieć, że stwierdzenie ewentualnych anomalii jest jedynie rekomendacją do głębszej kontroli, która pozwoli podjąć ostateczną decyzję. W obowiązującej praktyce badawczej, zgodnie z materiałami wewnętrznymi firm CBM Indicator i TNS OBOP, dodatkowymi elementami mogą być między innymi: braki numerów telefonów (zbierane do celów kontrolnych); odsetek poprawnych numerów telefonów (za niepoprawne przyjmuje się również te numery, które są cały czas wyłączone lub brak jest kompetentnego informatora); powtarzalność numerów telefonów[57] (pomiędzy różnymi badaniami u tego samego ankietera lub w tym samym regionie koordynatorskim[58]).

Przy analizie czasu trwania badania bardzo pomagają programy komputerowe tworzące tzw. historię pracy ankietera[59]. Wśród porównywanych parametrów, które pomagają w ocenie wiarygodności wyników, mogą się znaleźć: wielkość i typ miejscowości, w której był przeprowadzany wywiad, data jego przeprowadzenia, dzień tygodnia, godziny realizacji wywiadu i czas trwania oraz odstępy między wywiadami (dla badań CAPI czy PAPI)[60]. Otrzymane rezultaty z dwu ostatnich parametrów porów-

---

[56] F. Sztabiński, *Ocena jakości…*, s. 72–73.

[57] Sprawdzane wyłącznie wtedy, gdy firma ma zgodę na przechowywanie numerów telefonów w dłuższej perspektywie czasowej. Zwykle numer telefonu jest bowiem usuwany bezpośrednio po zakończeniu projektu.

[58] Badania społeczne, ewaluacyjne i marketingowe realizowane przez wyspecjalizowane firmy – agencje badawcze – opierają się na współpracy z tzw. koordynatorami regionalnymi. Jest to tzw. model dwustopniowy – badanie zlecane jest koordynatorom, a dopiero oni przydzielają je ankieterom (mało popularny jest model jednostopniowy, w którym pracownik agencji badawczej kontaktuje się bezpośrednio z ankieterem z pominięciem koordynatora). Dzięki temu można w sposób istotny ograniczyć zatrudnienie w działach realizacji badań w agencjach.

[59] F. Sztabiński, *Kontrola realizacji…*, s. 356, 373.

[60] Opracowanie własne na podstawie: tamże, s. 373. W przypadku badań CATI porównywanych parametrów jest znacznie więcej, są to także parametry innego typu. Tutaj

nuje się do ogólnej średniej dla badania, a także do historii pracy innych ankieterów i na tej podstawie formułuje się wnioski dotyczące jakości danych[61]. Do ilustracji tej procedury posłuży tabela 5. Dla uproszczenia przyjęto, że przykład dotyczy badania CAPI realizowanego w piątek. Data nie ma w tym wypadku znaczenia, ale założono, że tabela odnosi się do zwykłego okresu w roku, to jest przed piątkiem były cztery dni pracujące, a po tym dniu następują dwa dni wolne.

Tabela 5. Przykładowa historia pracy ankietera[a]

Numer wywiadu	Typ miejscowości	Wielkość miejscowości (przedział liczby mieszkańców w tys.)	Godzina rozpoczęcia wywiadu	Godzina zakończenia wywiadu	Czas trwania rozmowy (w min.)	Odstęp między poszczególnymi wywiadami (w min.)
1	Wieś	–	9:12	9:37	25	–
2	Wieś	–	10:43	11:11	28	66
3	Wieś	–	13:07	13:45	38	116
4	Miasto	100–499	15:17	15:28	11	92
5	Miasto	100–499	15:41	16:09	28	13
6	Miasto	100–499	16:23	16:50	27	14
7	Miasto	100–499	17:14	17:39	25	24
8	Miasto	powyżej 500	19:21	19:43	22	102
9	Miasto	powyżej 500	20:14	20:47	33	31
10	Miasto	powyżej 500	21:19	21:48	29	32

[a] Średnia długości przeprowadzanego wywiadu wyniosła 26 min. i 36 sek. (średnia dla całego badania: 31 min. 13 sek.), a średni odstęp między wywiadami – 54 min. 27 sek. (średnia dla całego badania podczas tego samego dnia: 61 min. 44 sek.).

Źródło: opracowanie własne na podstawie: Z. Sawiński, F. Sztabiński, P.B. Sztabiński (red.), *Fieldwork…*, s. 374.

można wymienić np. liczbę wykonanych połączeń podczas sesji w ogóle, liczbę wywiadów przerwanych, liczbę wywiadów przełożonych, średni czas trwania wywiadu, czas trwania wywiadu najkrótszego i najdłuższego. (Opracowanie własne na podstawie materiałów wewnętrznych firmy CBM Indicator). Badania CATI mają także tę przewagę, że w tym wypadku znacznie łatwiejsza jest możliwa bezpośrednia kontrola działań ankieterów w czasie rzeczywistym. Kontrolerzy mogą bowiem podsłuchiwać wywiady w trakcie. Więcej informacji: Z. Sawiński, *Badania CATI i wywiady przez telefon*, [w:] Z. Sawiński, F. Sztabiński, P.B. Sztabiński (red.), *Fieldwork…*, s. 344–346.

61 Tamże, s. 373–375.

Historia pracy przedstawiona w tabeli 5 nakazuje bliżej przyjrzeć się pracy tego ankietera. O ile możliwe jest, że w trakcie swojego dnia pracy ankieter wykonał podróż ze wsi do dużego miasta, a średnie nie wzbudzają podejrzeń, o tyle zastanawiające są inne parametry jego pracy. Trudno uwierzyć, aby w średniej wielkości mieście (wywiady numer od 4 do 7) w porach dnia, gdy wiele osób jest niedostępnych, ponieważ pracują, ankieter mógł znaleźć respondentów w przedziale czasu tak krótkim w porównaniu do czasu na wsi. Ponadto budzą wątpliwości (choć nie tak duże) godziny zrealizowanych z sukcesem wywiadów w piątek w mieście liczącym powyżej 500 tys. mieszkańców.

POWTARZALNOŚĆ I CYKLICZNOŚĆ ODPOWIEDZI

Analiza powtarzalności i cykliczności odpowiedzi zalicza się także do nieterenowych metod kontroli[62]. F. Sztabiński zalicza ją też do kontroli wewnętrznej opierającej się na określeniu „częstotliwości pojawiania się określonych sekwencji odpowiedzi w kwestionariuszach ankieterów"[63], ale również na „porównywaniu prawdopodobieństw występowania określonych układów odpowiedzi"[64]. Warto podkreślić, że tutaj metody statystyczne także mają zastosowanie (na przykład *data mining*[65]). Z kolei patrząc od strony oszustw popełnianych przez ankieterów, dziwne wzorce odpowiedzi[66] mogą dotyczyć zarówno fałszerstw związanych z doborem respondenta[67], jak i nierzetelnego przeprowadzenia wywiadu (czego jednym z przejawów jest między innymi celowe ominięcie części pytań zawartych w kwestionariuszu[68]).

W tabeli 6 przedstawiono różnicę między powtarzalnością a cyklicznością odpowiedzi. Oczywiście należy mieć na uwadze, że ten schemat jest bardzo uproszczony i dotyczy raczej badań PAPI.

---

62 Warto dodać, że F. Sztabiński w swojej publikacji *Ocena jakości…*, powołując się na Paula Lavrakasa, nazywa techniki nieterenowe *data analytic methods*. Oprócz wcześniej wspomnianej merytorycznej analizy wypełnionych kwestionariuszy do tej grupy zaliczają się też analizy opisu sytuacji wywiadu. Patrz: tenże, *Ocena jakości…*, s. 101.

63 Tenże, *Kontrola realizacji…*, s. 356.

64 Tenże, *Ocena jakości…*, s. 101.

65 Z. Sawiński, F. Sztabiński, *Czy ankieterzy…*, s. 371.

66 Tamże.

67 Tamże, s. 367.

68 Tamże, s. 368–373.

Tabela 6. Przykładowe rozkłady odpowiedzi na pytania dotyczące posiadania prawa jazdy i zachowań na drodze – różnica między powtarzalnością a cyklicznością (pogrubieniem zaznaczono odpowiedzi respondentów wprowadzone przez ankieterów, instrukcje dla ankietera – kursywą)

Pytania	Powtarzalność odpowiedzi			Cykliczność odpowiedzi		
1. Wpisz rok urodzenia respondenta (RRRR):	2003			1978		
2. Czy posiada Pan/Pani prawo jazdy? *(ANKIETER: jeśli „tak" – pytanie 3. pozostaw puste i przejdź do pytania 4.)*	**tak**	nie	odmowa	tak	**nie**	odmowa
3. A czy zamierza Pan/Pani zrobić je w ciągu najbliższego roku? *(ANKIETER: w przypadku wątpliwości respondenta dodaj, że chodzi o sytuację, gdy respondent planuje zacząć kurs, ale też gdy jest jego uczestnikiem lub go zakończył i oczekuje na egzamin) (ANKIETER: jeśli „nie" – przejdź do następnej części badania, a kolejne pytania pozostaw puste)*	**tak**	nie	odmowa	tak	nie	**odmowa**
4. Czy kiedykolwiek prowadził(-a) Pan/Pani samochód pod wpływem takiej ilości alkoholu, że Pan/Pani i/lub Pana/Pani otoczenie uznawało, że jest Pan/Pani pijany(-a)?	**tak**	nie	odmowa	**tak**	nie	odmowa
5. Czy kiedykolwiek prowadził(-a) Pan/Pani auto z prędkością wyższą o co najmniej 10 km/h niż to było dozwolone?	**tak**	nie	odmowa	tak	**nie**	odmowa
6. Czy kiedykolwiek otrzymał(-a) Pan/Pani mandat drogowy?	**tak**	nie	odmowa	tak	nie	**odmowa**
7. Czy kiedykolwiek otrzymał(-a) Pan/Pani przynajmniej jeden punkt karny?	**tak**	nie	odmowa	**tak**	nie	odmowa

Źródło: opracowanie własne.

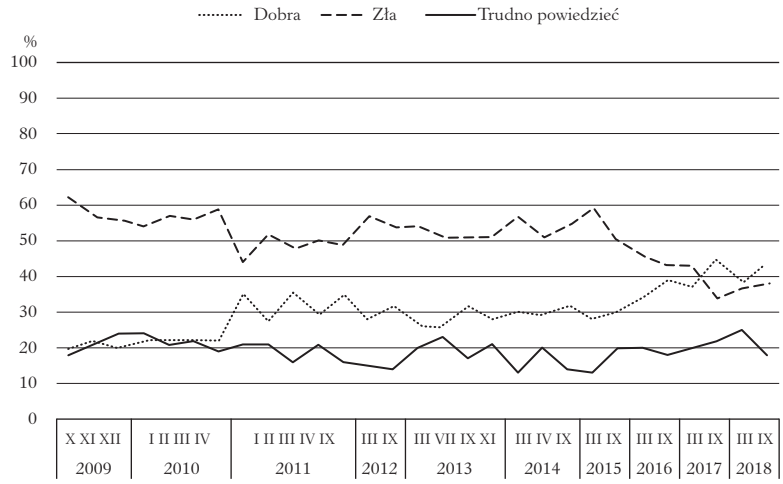
Gdyby odpowiedzi wprowadzone przez ankieterów (z wyjątkiem roku urodzenia) zamienić na liczby, w pierwszym przypadku (powtarzalności) byłby to układ: 1–1–1–1–1–1, a w drugim przypadku (cykliczności): 2–3–1–2–3–1. Powtarzalność cechuje zatem występowanie tych samych odpowiedzi, a cykliczność regularne występowanie tego samego układu (w omawianym przykładzie: 2–3–1). O fałszerstwie, oprócz zaobserwowanych anomalii, świadczy także brak wewnętrznej logicznej spójności odpowiedzi.

SEZONOWOŚĆ W BADANIACH POWTARZALNYCH I CIĄGŁYCH

Badania powtarzalne i ciągłe są szczególnym typem wywiadu ilościowego, który realizowany jest na dużej próbie, a sam pomiar realizowany jest wielokrotnie przy użyciu takiego samego lub nieco modyfikowanego narzędzia. Przy każdym pomiarze próba jest inna, ale dobierana w ten sam sposób i reprezentuje tę samą populację generalną[69]. Jak pisze Anna Dyjas-Pokorska: „trackingi [badania powtarzalne i ciągłe – aut.] służą do badania dynamiki zachowań, postaw, preferencji i wyborów konsumenckich lub społecznych i do śledzenia trendów, a więc przewidywania, jakie zmiany w mierzonych zjawiskach mogą nastąpić w najbliższej i dalszej perspektywie”[70].

Ze względu na powtarzalność pomiaru możliwe jest zaobserwowanie, czy oraz kiedy (w jakich warunkach) występują regularne, powtarzalne zmiany dotyczące badanego zjawiska. Za przykład niech posłuży wykres opracowany na podstawie badań statutowych Centrum Badania Opinii Społecznej (rys. 4).

Rysunek 4. Oceny działalności Zakładu Ubezpieczeń Społecznych w latach 2009–2018[a]



[a] Połączenia między kolejnymi punktami pomiarowymi służą jedynie zilustrowaniu wzrostów lub spadków ocen.

Źródło: opracowanie własne na podstawie raportów Centrum Badania Opinii Społecznej: *Opinie o działalności prezydenta, parlamentu, ZUS, ABW i CBA*, komunikat z badań BSS/166/2009; *Oceny instytucji publicznych*, komunikat z badań BS/128/2012; *Oceny działalności instytucji publicznych*, komunikat z badań nr 121/2018.

69 A. Dyjas-Pokorska, *Badania trackingowe…*, s. 237–238.

70 Tamże.

Jak widać na wykresie, w latach 2015–2018 oceny działalności ZUS we wrześniu ulegają względem marca łagodnej poprawie. Jednym z uzasadnień występowania sezonowości może być fakt, że pomiędzy marcem a kwietniem do emerytów i rencistów docierają informacje o rewaloryzacji rent i emerytur[71], a we wrześniu respondenci mają już za sobą kilka wypłat świadczenia w nowej wysokości[72]. Ponadto pod koniec każdego roku i na początku roku następnego przedsiębiorcy (oraz osoby pracujące) dowiadują się o wysokości składek wpłacanych do Funduszu Ubezpieczeń Społecznych[73]. Może to dodatkowo tłumaczyć łagodne spadki między wrześniem a marcem następnego roku (wyjątek stanowi przełom roku 2015 i 2016, kiedy nastąpiło zaburzenie tego trendu). Należy także zwrócić uwagę na rok 2012, w którym pomiary zostały dokonane w takim samym interwale czasowym, wspomniane zaś wcześniej zmiany również są widoczne. Sezonowość widać również w 2011 roku. Jednakże, ze względu na nierównomierność rozmieszczenia punktów pomiarowych w czasie, wnioskowanie nie jest dokładne. Warto przy tym pamiętać, że im więcej punktów pomiarowych, tym dokładniejszy będzie wyznaczony trend.

Weryfikacja danych – kilka praktycznych przykładów

Analiza zjawisk otaczającego świata może zostać dokonana z użyciem wielu różnych metod statystycznych. Niektóre z nich dostarczają wartościowej poznawczo informacji i nie wymagają korzystania z zaawansowanych instrumentów. Istnieją narzędzia, które w prosty sposób umożliwiają sprawdzenie, czy zbiór danych, który jest w naszym posiadaniu, bądź jakieś jego elementy nie zostały sfałszowane. Do pierwszej grupy należy rozkład normalny Carla Friedricha Gaussa oraz rozkład zdarzeń

[71] Zgodnie z art. 88 ust. 1 ustawy z 17 grudnia 1998 r. o emeryturach i rentach z Funduszu Ubezpieczeń Społecznych (tekst jedn. Dz.U. 2018, poz. 1270, z późn. zm.) – stan prawny na dzień 8.01.2019.

[72] Z kolei jednym z argumentów przeczących temu wyjaśnieniu jest fakt, że w połowie każdego roku Polacy otrzymują od ZUS prognozę swojej przyszłej emerytury. Są one zauważalnie niższe względem nawet obecnych świadczeń emerytalnych.

[73] *Najniższa podstawa wymiaru składek oraz kwoty składek na ubezpieczenia społeczne w 2019 r.*, http://www.zus.pl/o-zus/komunikaty/-/publisher/komunikat/1/najnizsza-podstawa-wymiaru-skladek-oraz-kwoty-skladek-na-ubezpieczenia-spoleczne-w-2019-r_/2136916 (dostęp: 10.01.2019); *Wysokość składek na ubezpieczenia*, http://www.zus.pl/baza-wiedzy/skladki-wskazniki-odsetki/skladki/wysokosc-skladek-na-ubezpieczenia-spoleczne (dostęp: 10.01.2019).

rzadkich Siméona Denisa Poissona. Weryfikacja elementów wchodzących w skład zbioru danych może odbyć się z użyciem testu Ulricha Grafa, kryterium Williama Chauveneta bądź z zastosowaniem prawa Franka Benforda.

WERYFIKACJA ZBIORÓW DANYCH – ROZKŁAD GAUSSA I ROZKŁAD ZDARZEŃ RZADKICH POISSONA

Rozkład normalny, choć powszechnie używany w nauce, pozostaje niedoceniony w pracy analityków informacji. Nazwę „rozkład normalny" wprowadził w 1889 roku brytyjski antropolog, genetyk i statystyk Francis Galton. Niepośledni wkład w badania nad tym fenomenem wniósł niemiecki matematyk, fizyk i astronom Carl Friedrich Gauss – to od jego nazwiska właśnie wzięła się nazwa krzywej dzwonowatej, opisującej między innymi liczne zjawiska występujące w przyrodzie czy cechy charakteryzujące organizmy, a także, co bardzo istotne, rozkład błędów pomiaru.

Prace C.F. Gaussa w tej materii zostały zapomniane przez naukowców na długie lata. Popularność i wszechstronność zastosowań rozkładu normalnego przywrócił francuski matematyk, astronom, geodeta i fizyk Pierre Simon de Laplace[74]. To właśnie on zetknął się z pracami C.F. Gaussa w 1810 roku, kiedy przedstawił Akademii Nauk pracę, w której udowodnił tak zwane centralne twierdzenie graniczne, zgodnie z którym prawdopodobieństwo, iż suma dużej liczby niezależnych czynników losowych przyjmie wartość z danego zakresu, jest takie jak przewiduje rozkład normalny. Czytając prace Gaussa, Laplace zdał sobie sprawę, że mógłby z nich skorzystać i poprawić własne wyniki oraz że jego uzupełniona w ten sposób praca dostarcza lepszych argumentów niż te użyte przez samego Gaussa na poparcie tezy, że rozkład normalny stanowi tzw. uniwersalne prawo błędu. P.S. Laplace szybko opublikował krótki dodatek do swojej wcześniejszej pracy. Centralne twierdzenie graniczne oraz prawo wielkich liczb są do dnia dzisiejszego dwoma najsłynniejszymi twierdzeniami w rachunku prawdopodobieństwa.

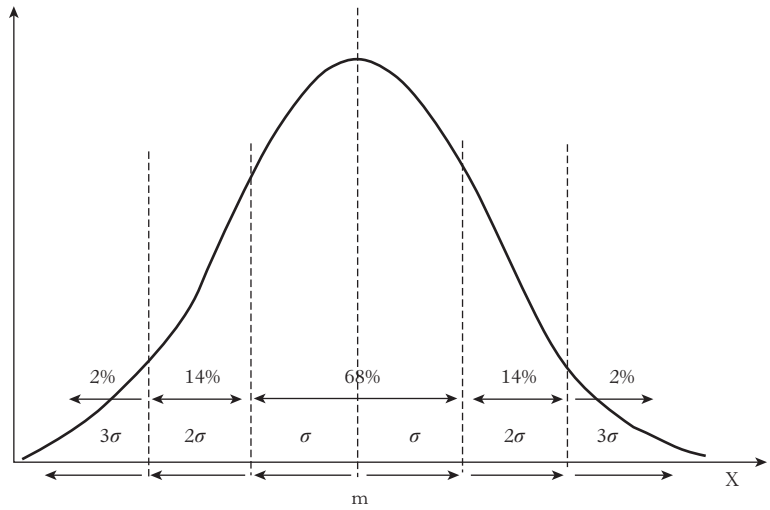
Wśród prekursorów współczesnej statystyki bardzo często wspominany jest żyjący w XIX wieku belgijski astronom, matematyk, meteorolog, socjolog i kryminolog Lambert Adolphe Jacques Quételet, który przeprowadził obszerne badania, aby wykazać statystyczną regularność pewnych cech czy faktów (jak liczba urodzeń, zgonów czy popełnianych

[74] L. Mlodinow, *Matematyka niepewności. Jak przypadki wpływają na nasz los*, przekł. P. Strzelecki, Warszawa 2009, s. 168–170.

przestępstw). W poszukiwaniu danych potwierdzających występowanie rozkładu normalnego pomogło mu pewne szkockie czasopismo, które opublikowało dane dotyczące obwodu klatki piersiowej oraz wzrostu ponad 5000 żołnierzy odbywających służbę w różnych szkockich pułkach. Dane te pozwoliły mu wykazać, że zmienność cech żołnierzy jest tego samego rodzaju co opisywana rozkładem normalnym. Dalsze prace uczonego w tej materii zaowocowały opracowaniem wykorzystywanego do dnia dzisiejszego wskaźnika masy ciała, zwanego wskaźnikiem Quételeta lub BMI (*Body Mass Index*).

Krzywa Gaussa pokazuje, z jakim prawdopodobieństwem w dowolnej populacji, w tym ludzkiej, występują poszczególne wartości danej cechy. Mogą to być na przykład: iloraz inteligencji, wzrost, zarobki itd. Wierzchołek krzywej opisuje średnią wartość występowania cechy, a wielkość oznaczona symbolem greckim σ – odchylenie standardowe. Ta ostatnia informuje o tym, jak szeroko wartości badanej cechy rozrzucone są wokół średniej. Kształt krzywej zmienia się w zależności od wartości odchylenia standardowego – im jest ono mniejsze, tym krzywa jest bardziej smukła. Dla dużych wartości odchylenia standardowego kształt krzywej zmienia się na bardziej rozłożysty. W przypadku rozkładu normalnego 68,2% wartości cechy leży w odległości mniejszej lub równej σ od wartości średniej (patrz rys. 5 – pas $(\bar{x} - \sigma; \bar{x} + \sigma)$), 95,4% wartości cechy leży w odległości 2σ od średniej, a 99,6% w odległości 3σ.

Rysunek 5. Parametry rozkładu normalnego



Źródło: opracowanie własne.

Praktyczną analityczną przydatność rozkładu normalnego prześledzono w następującej egzemplifikacji. Krzywa normalna między innymi opisuje rozkład ilorazu inteligencji w populacji mieszkańców Europy. Standardowo przyjmuje się, że średnia wartość tej cechy wynosi 100 IQ, a odchylenie standardowe – 15. Zatem 68% populacji to osoby mające średni iloraz inteligencji (pas pomiędzy 85 a 115 IQ). Ludzie o inteligencji wyższej niż przeciętna – mający IQ powyżej 116, ale poniżej 130 – stanowią blisko 14% populacji. Tyle samo jest osób mało inteligentnych mających IQ między 70 a 84. Skrajne wartości cechy znajdujące się po obu stronach wartości średniej, a oddalone od niej o odległość 3σ, to ludzie uznawani za upośledzonych umysłowo – IQ poniżej 70 (lewa strona krzywej) lub ponadprzeciętnie inteligentnych – IQ powyżej 130 (prawa strona krzywej). Podsumowując: na 100 Europejczyków dwóch będzie upośledzonych umysłowo, 14 – mało inteligentnych, 68 – średnio inteligentnych, 14 – o ponad przeciętnej inteligencji i 2 geniuszy.

Kolejnym jakże istotnym, z punktu widzenia statystyki, jest rozkład Poissona, opisujący tzw. rozkład zdarzeń rzadkich, znany także jako prawo małych liczb Poissona. Jest wykorzystywany wszędzie tam, gdzie prawdopodobieństwo wystąpienia danego zjawiska jest bardzo małe, niemalże równe zeru, natomiast liczba jednostek, którym można je przypisać, jest bardzo duża, zbiegająca nawet do nieskończoności. Twórcą tego rozkładu jest Siméon Denis Poisson, francuski matematyk i fizyk, który w pracach nad badaniem prawdopodobieństwa wydania danego wyroku sądowego w sprawach cywilnych i karnych wprowadził teorię pozwalającą przewidzieć występowanie określonego zdarzenia w wyznaczonym przedziale czasu[75]. Badaczem, który wniósł duży wkład w rozwój tej teorii, był rosyjski matematyk polskiego pochodzenia Władysław Bortkiewicz. Stąd niekiedy w literaturze można się spotkać z inną nazwą teorii – prawem małych liczb Poissona–Bortkiewicza. W. Bortkiewicz znalazł zastosowanie rozkładu Poissona w szacowaniu wśród kawalerzystów pruskich korpusów liczby zgonów spowodowanych kopnięciem konia[76]. Rozkład ten jest bardzo często wykorzystywany w medycynie, finansach oraz ubezpieczeniach do przewidywania liczby zachorowań na bardzo rzadką chorobę, wykrycia systemowych błędów pojawiających się podczas księgowania transakcji finansowych czy do oszacowania prawdopodobieństwa, że na

[75] D. Mider, A. Marcinkowska, *Analiza danych ilościowych dla politologów. Praktyczne wprowadzenie z wykorzystaniem programu GNU PSPP*, Warszawa 2013, s. 268–269.

[76] Więcej na temat W. Bortkiewicza oraz jego dokonań w: J. Schumpeter, *Ladislaus von Bortkiewicz*, „Economic Journal" 1932, nr 42, s. 338–340; P.A. Samuelson, *Resolving a Historical Confusion in Population Analysis*, „Human Biology" 1976, nr 48, s. 559–580.

początku trwania ubezpieczenia osoba, która dokonała zakupu polisy, dozna trwałego uszczerbku na zdrowiu.

Rozkład Poissona jest przykładem rozkładu dyskretnego (skokowego) i jest szczególnym przypadkiem rozkładu Jakoba Bernoulliego, definiowanego za pomocą dwóch parametrów: λ oraz k, gdzie $\lambda = n * p$ (n – liczba obserwacji, p – prawdopodobieństwo wystąpienia zdarzenia), a k jest liczbą sukcesów wystąpienia zdarzenia o prawdopodobieństwie p. Rozkład Poissona określony jest następującym wzorem:

$$p_k = \frac{\lambda^k}{k!} e^{-\lambda}$$

gdzie e jest liczbą Nepera, równą w przybliżeniu 2,72.

Ciekawą własnością tego rozkładu jest to, że wartość oczekiwana jest równa wariancji i wynosi λ. W przypadku, kiedy λ przybiera bardzo duże wartości, rozkład Poissona jest z dobrą dokładnością przybliżany przez rozkład Gaussa z tą samą wartością średnią i odchyleniem standardowym. Nie jest więc błędem stosowanie rozkładu Gaussa zamiast rozkładu Poissona w przypadkach, kiedy iloczyn liczby obserwacji i prawdopodobieństwa wystąpienia danego zdarzenia przyjmuje dużą wartość.

W statystyce istnieje wiele testów sprawdzających normalność rozkładu. Badanie, czy rozkład zmiennej jest zbliżony do rozkładu normalnego, jest podstawą do stosowania dalszych metod statystycznych, takich jak: analiza wariancji, korelacja r-Pearsona, regresja wieloraka itd. W dalszej części skupimy się na jednym z najczęściej stosowanych testów opracowanym w 1965 roku przez dwóch statystyków: Samuela Shapiro i Martina Wilka, który – jak wykazała analiza porównawcza metodą Monte Carlo – ma największą moc spośród innych testów sprawdzających normalność rozkładu. Żeby zapoznać czytelnika z łatwością stosowania oraz interpretacją tego testu (bez znajomości zaawansowanej wiedzy statystycznej czy matematycznej), posłużmy się następującym przykładem. Załóżmy, że w pewnym bardzo małym mieście urząd skarbowy pozyskał informacje dotyczące wysokości zarobków swoich mieszkańców z deklaracji podatkowych, które złożyli. Naczelnik urzędu jest osobą bardzo podejrzliwą i chce sprawdzić testem Shapiro–Wilka, czy uzyskane dane nie zostały przez podatników zafałszowane. Wprowadził wyniki do programu SPSS[77] i wykonał kolejno następujące polecenia:

[77] SPSS (Statistical Package for the Social Sciences) – płatne oprogramowanie do statystycznej analizy danych. Najczęściej wykorzystywane jest w badaniach naukowych, rynku i opinii, badaniach epidemiologicznych. Typowa praca z SPSS jest pracą z oknami

w menu wybrał Analiza → Opis statystyczny → Eksploracja, w oknie dialogowym w miejscu Zmienne zależne wprowadził zmienną „zarobki", a w kolejnym kroku w menu Wykresy zaznaczył opcję Wykresy normalności z testami oraz Histogram. W efekcie otrzymał różne tabelki i wykresy. Najważniejszą z nich jest ta o nazwie Testy normalności rozkładu, z niej można odczytać, czy podejrzenie o zafałszowanie danych jest słuszne. W tabeli 7 prezentujemy wynik obliczeń, jaki otrzymał nasz naczelnik.

Tabela 7. Testy normalności rozkładu (obrazujące przykładowe działania weryfikacji poprawności danych)

	Kołmogorow-Smirnow[a]			**Shapiro-Wilk**		
	Statystyka	**df**	**Istotność**	**Statystyka**	**df**	**Istotność**
zarobki	,081	77	,200*	,964	77	,027

* dolna granica rzeczywistej istotności

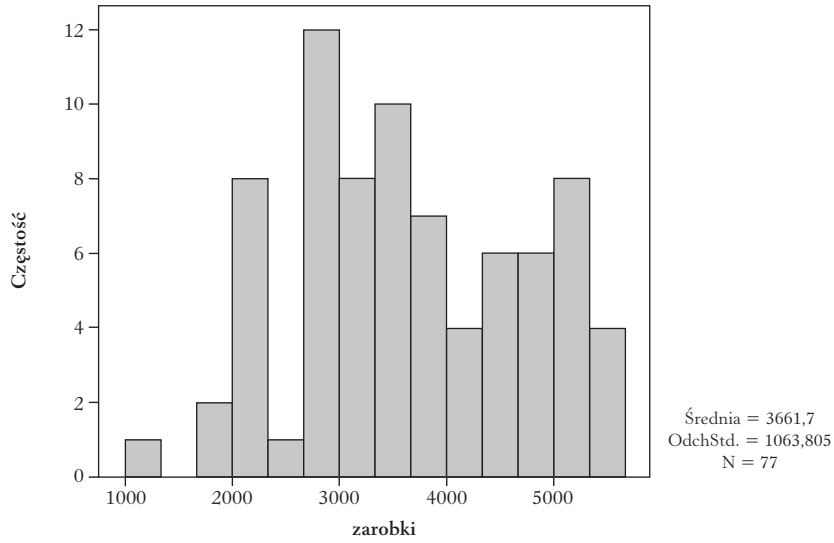
[a] z poprawką istotności Lillieforsa

Źródło: opracowanie własne.

Tabela zawiera wyniki dwóch testów, ale naczelnika interesuje tylko część prezentująca wynik testu Shapiro–Wilka. Kluczową informacją jest wartość współczynnika istotności. Jeśli jest ona niższa niż 0,05, to możemy sądzić, że dane, które poddaliśmy testom, nie spełniają warunku normalności rozkładu. W naszym przykładzie wartość ta wynosi 0,027 i jest mniejsza niż 0,05, a zatem podejrzenie o zafałszowaniu zarobków mieszkańców miasta zostało potwierdzone, a ujawnia ten fakt otrzymany histogram. Już „na oko" widać, że gdybyśmy połączyli środki każdego z wierzchołków słupków, to linia je łącząca nie będzie w żaden sposób przypominała krzywej Gaussa.

dialogowymi i kreatorami graficznymi. Od 1968 r. w sposób ciągły jest rozbudowywane o nowe moduły i metody analityczne. Darmową alternatywą dla SPSS jest GNU PSPP, w którym stosowany język dąży do zgodności ze swoim płatnym odpowiednikiem. Obecnie na rynku jest wiele programów wspomagających przeprowadzanie analiz statystycznych. Autorzy proponują wykorzystanie GNU PSPP nie tylko ze względu na to, że jest bezpłatny, ale również ze względu na wygodę i intuicyjne jego wykorzystanie.

Rysunek 6. Przykładowy histogram prezentujący rozkład zmiennej



Źródło: opracowanie własne.

Równie szybkie i proste jest sprawdzenie, czy uzyskane w toku pomiarów dane przyjmują rozkład Poissona. Wykorzystamy w tym celu test Kołmogorowa–Smirnowa, którego nazwa pochodzi od nazwisk dwóch słynnych rosyjskich matematyków: Andrieja Kołmogorowa oraz Nikołaja Smirnowa. Podobnie jak przy rozkładzie normalnym posłużmy się następującym przykładem. Pewien pracownik dużej firmy otrzymuje dodatkową premię – za każdego klienta, z którym podpisze długotrwałą umowę na wykorzystanie pewnego sprzętu. Z informacji uzyskanych od analityków tej firmy wynika, że podczas jednego dnia pracy podpisywane są średnio cztery umowy z nowymi klientami. Pracownik złożył kwartalny raport z wyników swojej pracy. Jego przełożony przypuszcza, że pracownik zawyżył liczbę umów, które podpisał w tym czasie, celem wymuszenia na pracodawcy wyższej premii. Wykaz dokładnej liczby codziennych sprzedaży został poddany testowi Kołmogorowa–Smirnowa. Wykonano następujące polecenia w programie SPSS: Analiza → Testy nieparametryczne → Testy tradycyjne → K-S dla jednej próby, następnie wybrano rodzaj testu → Poissona. Otrzymano dwie tabele, z których najważniejsza to test Kołomogorowa–Smirnowa dla jednej próby. Podobnie jak w poprzednim przykładzie, informacją potwierdzającą lub zaprzeczającą przypuszczeniom zafałszowania liczby umów jest wartość współczynnika istotności asymptotycznej. Jeśli jest on większy od 0,05, to nasza zmienna przyjmuje rozkład Poissona, w przeciwnym wypadku możemy

stwierdzić, że dane takiego rozkładu nie przypominają. W naszym przykładzie wartość ta wynosi 0,037, jest zatem mniejsza niż 0,05. Przełożony może zatem śmiało stwierdzić, że jego podwładny dokonał zafałszowania wyników swojej pracy.

Rysunek 7. Test Kołmogorowa–Smirnowa dla jednej próby

		VAR00002
N		92
Parametr rozkładu	Średnia	5,6413
Największe różnice	Wartość bezwzględna	,147
	Dodatnia	,147
	Ujemna	-,047
Z Kołmogorowa–Smirnowa		1,413
Istotność asymptotyczna (dwustronna		,037

Źródło: opracowanie własne.

WERYFIKACJA POSZCZEGÓLNYCH ELEMENTÓW ZBIORU DANYCH – KRYTERIUM CHAUVENETA, TEST GRAFA, PRAWO BENFORDA

Wymienione tu testy umożliwiają identyfikację odstających obserwacji w zbiorze danych. Czy odbiegająca wartość jest spowodowana błędem pomiaru i należy ją odrzucić, czy też uzyskany wynik może być odzwierciedleniem jakiegoś ważnego efektu, a jego pojawienie się nie jest żadną pomyłką? Zwykle niemożliwe jest ustalenie zewnętrznych przyczyn powstania anomalnego wyniku, a decyzja o pozostawieniu go w zbiorze danych lub usunięciu należy do badacza. Rozstrzygnięcie dylematu nie jest sprawą prostą, zwłaszcza w sytuacjach, kiedy w grę wchodzi bardzo precyzyjne wykonanie analiz, na podstawie których mają być podjęte ważkie lub strategiczne decyzje, na przykład o wprowadzeniu nowego leku czy w sprawie produkcji urządzeń wysokiej precyzji.

Opracowano kilka testów bazujących na porównaniach różnic od wartości średniej i odchylenia standardowego, które pozwalają na wykrycie odstających obserwacji przy założeniu, że pomiar wielkości ma rozkład normalny. Do najważniejszych z nich i najczęściej stosowanych należą: kryterium Williama Chauveneta, test Ulricha Grafa, test Franka Grubbsa oraz kryterium Benjamina Peirce'a. W literaturze przedmiotu wszystkie należą do grupy tzw. testów na błąd gruby[78]. W niniejszym artykule

[78] W literaturze możemy spotkać się z podziałem błędów pomiaru na trzy kategorie: błędy przypadkowe (niepowtarzające się, przyjmujące wartości całkowite, wynikające z nie-

szczegółowo zostaną omówione pierwsze dwa. Przed przystąpieniem do realizacji testu z wykorzystaniem kryterium Chauveneta należy mieć pewność, że każdy z następujących trzech warunków został spełniony: wartości zmiennej muszą mieć rozkład normalny; tylko jeden z wyników pomiaru może znacznie odbiegać od pozostałych; mamy niemalże pewność, że podejrzany wynik jest przejawem błędu, a nie odzwierciedleniem jakiegoś ważnego efektu.

Obliczeń dokonuje się według następującego algorytmu[79]:

- obliczamy średnią $\bar{x}$ oraz wariancję s z całości próby;
- liczymy odchylenie otrzymanego wyniku od wartości średniej ze wzoru: $t = \frac{|x - \bar{x}|}{s}$ gdzie x, to nasza odstająca wartość;
- liczymy prawdopodobieństwo, że nasz podejrzany wynik będzie oddalony od wartości średniej o nie mniej niż $(t * s)$ za pomocą wzoru: $p = 2 - 2\varphi(t)$, gdzie $\varphi(t)$ jest wartością dystrybuanty rozkładu normalnego w punkcie t (odczytujemy ją z tablic rozkładu normalnego);
- obliczamy iloczyn $(n * p)$;
- sprawdzamy, czy $(n + p) < 0{,}5$. Jeśli tak, podejrzany wynik należy odrzucić, jeśli nie, to należy pozostawić go w zbiorze wyników.

Jak już zostało wspomniane, kryterium Chauveneta stosujemy w przypadku, gdy mamy do czynienia tylko z jednym odstającym wynikiem pomiaru. Co zatem należy zrobić, gdy jest ich więcej? Załóżmy na początek, że mamy dwa podejrzane wyniki: x_1 oraz x_2, przy czym x_2 jest bardziej odległe od średniej niż x_1. W pierwszej kolejności powinniśmy zastosować kryterium Chauveneta do x_1. Jeśli spodziewana liczba pomiarów równie odległych jest mniejsza niż jeden, powinniśmy odrzucić obie wartości. Jeśli przeciwnie – liczba ta jest większa niż jeden, to stosujemy kryterium używając x_2. W przypadku, kiedy uzyskamy liczbę pomiarów równie odległych od średniej mniejszą niż 0,5, to x_2 należy odrzucić. Po odrzuceniu dowolnego pomiaru, który nie spełnia kryterium, należy ponownie obliczyć $\bar{x}$ i s, używając tylko pozostałych danych. Nowa wartość s będzie mniejsza niż pierwotna i dla niej więcej pomiarów mogłoby

kontrolowanych podczas pomiaru czynników – np. zmiany napięcia sieci elektrycznej czy też ograniczonej dokładności obserwacji), błędy grube (duże błędy spowodowane nieuwagą osoby dokonującej pomiaru) oraz błędy systematyczne (powtarzające się, powodowane przez czynniki działające w jednakowy sposób, w czasie wielokrotnego powtarzania tego samego pomiaru).

79 Patrz także: J.R. Taylor, *Wstęp do analizy błędu pomiarowego*, przekł., A. Babiński, R. Bożek, Warszawa 2018, s. 200–202.

nie spełniać kryterium Chauveneta. Wielu naukowców podchodzi jednak bardzo sceptycznie do stosowania kryterium z przeliczonymi wartościami średniej oraz wariancji. Dlatego warto zastosować inne metody wykrywania odstających obserwacji, kiedy w wyniku pomiaru pojawiło się ich więcej niż dwie.

Prostszym i szybszym w obliczeniach algorytmem oceny, czy daną wartość należy pozostawić w zbiorze, czy odrzucić, jest test Grafa. Należy jednak pamiętać, że test ten jest mniej „czuły"[80] na wyszukiwanie danych odbiegających niż kryterium Chauveneta, co związane jest w wykorzystaniem w przypadku tego drugiego precyzyjniejszych narządzi matematycznych. Zaleca się, aby po realizacji testu Grafa otrzymane wyniki dodatkowo poddać analizie kryterium Chauveneta.

Test Grafa przeprowadza się według następującego algorytmu[81]:

- liczymy wartość średnią $\bar{x}$ oraz wariancję s z wyłączeniem podejrzanej obserwacji;
- wyliczone wartości podstawiamy do wzoru $(\bar{x} - 4s; \bar{x} + 4s)$ opisującego przedział dopuszczalnych wyników;
- sprawdzamy, czy nasza podejrzana wartość mieści się w tym przedziale – jeśli tak, to należy ją pozostawić w zbiorze, a w przeciwnym przypadku należy ją odrzucić.

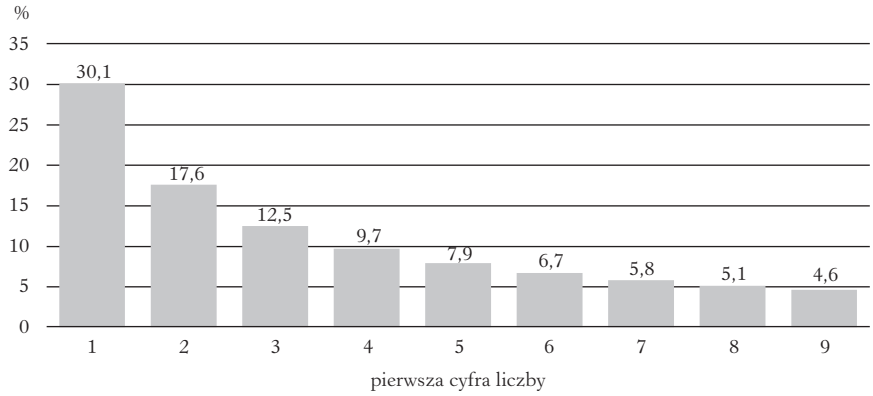
Kolejny rodzaj testu to prawo Benforda. Jednym z pierwszych odkrywców ciekawej zależności, jaką wykazują liczby wiodące, to jest pierwsze cyfry występujące w wielu zbiorach danych liczbowych, był astronom i matematyk Simon Newcomb. Szukając w bibliotece United States Naval Observatory materiałów do swojej nowej pracy, zauważył, że strony ksiąg z tablicami logarytmicznymi są bardziej zużyte na początkowych niż na końcowych kartach. Przyczyną był fakt zapotrzebowania przez badaczy częściej na logarytmy liczb zaczynających się od 1 niż innych. W 1881 roku S. Newcomb sformułował i opublikował w „American Journal of Mathematics" hipotezę, że częstość występowania cyfr wiodących ma rozkład procentowy taki, jak na rysunku 8.

---

[80] M. Słowik, M. Bartkowiak, *Elementy statystycznej analizy wyników pomiarów na przykładzie badań wybranych cech mieszanek mineralno-asfaltowych*, „Drogownictwo" 2016, nr 7–8, s. 247–253.

[81] Patrz także: J.R. Taylor, *Wstęp do analizy…*, s. 200–202.

Rysunek 8. Rozkład częstości występowania liczb znaczących



Źródło: opracowanie własne.

Liczba 1 występuje w 30,1% przypadków, 2 – w 17,6%, 3 – w 12,5%. Spadek częstości występowania jest tak gwałtowny, że jedynki są niemalże siedem razy bardziej prawdopodobne od dziewiątek.

Rozkłady procentowe S. Newcomba oparte są na logarytmach[82]. Wywnioskował on, że prawdopodobieństwo tego, iż dana liczba zaczyna się na cyfrę a, wynosi [log(a+1) – log a]. Jego odkrycie pozostało jednak bez echa najprawdopodobniej dlatego, że nie przedstawił ścisłego dowodu, co w świecie matematyków pozwala je traktować jedynie jako ciekawostkę.

Niemalże pół wieku później, w 1938 roku, tego samego odkrycia (bez znajomości wcześniejszej pracy S. Newcomba), również na podstawie tablic logarytmicznych, dokonał Frank Benford, fizyk zatrudniony w General Electric w Nowym Jorku. F. Benford nie ograniczył się jednak do tablic logarytmicznych. Przeanalizował występowanie cyfr wiodących pochodzących między innymi z tabel populacji amerykańskich miast, adresów pierwszych kilkuset osób, których dane zawarte były w American Men of Science, tablic ciężarów atomowych pierwiastków, zestawień powierzchni rzek i danych dotyczących wyników meczy baseballowych. Dla niemalże wszystkich zestawów liczbowych rozkład liczb wiodących był bliski oczekiwanego, z błędami na poziomie kilku dziesiątych procenta. F. Benford postawił hipotezę, że stanowi to przejaw uniwersalnego

[82] A. Bellos, *Alex po drugiej stronie lustra. Jak liczby odzwierciedlają życie, a życie odzwierciedla liczby*, przekł. M. Krośniak, Warszawa 2018, s. 52–53.

prawa, które nazwał prawem anomalnych liczb[83]. Nazwa ta się jednak nie przyjęła, obecnie prawidłowość ta zwana jest prawem Benforda.

Metoda sprawdzania zgodności pierwszych cyfr z prawem Benforda wykorzystywana jest obecnie do wykrywania manipulacji liczbami wszędzie tam, gdzie sensowne jest założenie, że prawo to powinno być spełnione.

Przykłady praktycznego zastosowania prawa Benforda są liczne. Do najsłynniejszych należy sprawa „kreatywnej księgowości" z 1992 roku. Postawiono wówczas w akt oskarżenia Jamesa Nelsona, głównego księgowego Arizona State Treasurer, zarzucając mu defraudację niemalże 2 mln dolarów. W poczet dowodów przeciwko oskarżonemu została zaliczona analiza finansów z wykorzystaniem prawa Benforda, która stanowiła kluczowy argument pozwalający na wydanie skazującego wyroku[84]. Walter Mebane, politolog z University of Michigan, użył prawa Benforda do udowodnienia manipulacji wynikami wyborów prezydenckich w Iranie w 2009 roku. Analizy wykazały, że głosy oddane na urzędującego prezydenta Mahmuda Ahmadineżada wyraźnie odbiegały od testu cyfr wiodących, a oddane na jego przeciwnika, Mira Hosseina Musawiego, rozbieżne nie były[85]. Naukowcy wykorzystują również omawiane prawo jako narzędzie diagnostyczne. Malcolm Sambridge z Australian National University przeanalizował zapisy dwóch różnych sejsmografów, które zarejestrowały trzęsienie ziemi w Indonezji w 2004 roku – z Peru i z Australii. Wyniki podawane przez pierwszy z wymienionych były zgodne z prawem Benforda, drugiego zaś nie. M. Sambridge wyciągnął wniosek, że te drugie dane te musiały być zaburzone przez wstrząsy sejsmiczne o niewielkiej sile w rejonie Canberry. W tym przypadku test cyfr wiodących ujawnił trzęsienie ziemi, które w innym przypadku pozostałoby niezauważone[86].

Dowód prawa Benforda nie jest matematycznie prosty, ale zastosowanie ma wręcz banalne. Aby zrozumieć dowód, należy znać teorię

---

[83] F. Benford, *The Law of Anomalous Numbers*, „Proceedings of the American Philosophical Society" 1938, nr 78(4), s. 551–572.

[84] S.A. Barnes, *Identifying Fraud Can Be as Easy as 1, 2, 3: Applying Benford's Law to Forensic Analyses and Investigations*, https://files.constantcontact.com/85e69272601/6ec4247f-14c0-49cf-92a6-674a5a4c37c2.pdf (dostęp: 19.01.2019).

[85] W.R. Mebane, Jr., K. Kalinin, *Comparative Election Fraud Detection*, APSA 2009 Toronto Meeting Paper, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1450078, 2009 (dostęp: 19.01.2019).

[86] M. Sambridge, H. Tkalčić, P. Arroucau, *Benford's Law of First Digits: From Mathematical Curiosity to Change Detector*, „Asia Pacific Mathematics Newsletter" 2011, nr 1, s. 1–5.

ergodyczną, łączącą teorię prawdopodobieństwa z fizyką statystyczną. Ciekawą rzeczą jest natomiast właściwość niezmienniczości skalowania tego prawa. Jeśli jakieś dane finansowe wyrażone w złotówkach spełniają test cyfr wiodących, to będzie on również spełniony po przeliczeniu złotówek na dolary czy euro. Aby samemu się przekonać o tym, że niemalże wszystkie zjawiska, jakie zachodzą w naszym otoczeniu, spełniają prawo Benforda, proponujemy wziąć zwykłą gazetę. Policzywszy wiodące cyfry, każdy jest w stanie się przekonać, że ich rozkład ułoży się wzdłuż opadającej skali z jedynką znacznie przeważającą pod względem częstości wobec pozostałych cyfr. Prawo Benforda zapoczątkowało odkrycie serii innych praw i zależności, takich jak prawo George'a Kingsley'a Zipfa oraz prawo Vilfreda Pareto.

Opinie ekspertów

W artykule przedstawiono wiele różnych metod weryfikacji danych. Metod zarówno niestatystycznych, zdroworozsądkowych, jak i tych wywodzących swoje korzenie z matematyki i mających mocne oparcie w nauce. Warto zastanowić się jednak, z czego na co dzień korzystają osoby zajmujące się weryfikacją jakości danych. Przygotowując niniejszy artykuł, rozmawialiśmy z kilkoma z nich na ten temat.

W ramach badawczej praktyki stosowanej w komercyjnych instytutach badawczych[87] stosuje się metody najszybsze, wskazane w części poświęconej metodom niestatystycznym. Taką procedurą może być – i najczęściej jest – kontrola logiczna odbywająca się zgodnie z określonym schematem, różnym dla różnych firm. Na przykład w firmie respondenta A jako obserwacje odstające traktuje się wyłącznie górny 1% wszystkich wartości na danej zmiennej, gdzie standardem jest sformułowanie +/–n% oznaczające zarówno górny, jak i dolny odsetek. Również kontrola logiczna opisana w początkowej części artykułu prowadzona jest w sposób uproszczony. Zwykle już po zakończeniu realizacji danego badania sprawdza się wyłącznie rozkłady odpowiedzi oraz wybrane zależności pomiędzy określonymi zmiennymi.

[87] Respondent A: mężczyzna, lat 41, wykształcenie wyższe, kierownik zespołu analitycznego w komercyjnym instytucie badań rynkowych od 7 lat, całkowity staż pracy na stanowiskach związanych z kontrolą i analizą danych: 13 lat. Respondent nie wyraził zgody na udostępnienie danych osobowych w postaci imienia i nazwiska oraz nazwy firmy.

Nieco inaczej kontrola logiczna przebiega w przypadku dużych badań naukowych[88], takich jak na przykład Polski Generalny Sondaż Społeczny czy Europejski Sondaż Społeczny. Przede wszystkim cykl badawczy takich badań jest wydłużony. Podczas gdy cały proces badania prowadzonego przez komercyjny instytut zwykle trwa około dwóch miesięcy – od przygotowania niezbędnych narzędzi do zaprezentowania wyników, to w przypadku badań naukowych, szczególnie wymienionych wyżej, mowa najczęściej o co najmniej dwóch latach. Kontrola takich badań odbywa się jeszcze często z wykorzystaniem papierowych kwestionariuszy, na których doskonale widać wszystkie próby zafałszowania[89] dokonywane przez ankietera. Taka kontrola jest jednym z pierwszych kroków. Drugim jest kontrola logiczna – na przykład przejść pomiędzy pytaniami, kompletności wypełnienia, powtarzalności odpowiedzi[90], jakości pytań otwartych. Dopiero później socjologowie sięgają po metody i testy statystyczne. Respondent B poleca jak najczęstsze korzystanie właśnie z tych ostatnich. Szczególnie w sytuacji, gdy badacz nie ma zbyt dużego doświadczenia, a uzyskanie wsparcia takiego specjalisty może być problematyczne. Ponadto respondent B zaleca zaplanowanie badania w taki sposób, żeby pomiędzy poszczególnymi czynnościami pozostały wolne okienka czasowe.

Zupełnie inne podejście reprezentuje respondent C[91]. Instytucja, w której on pracuje, podlega jednemu z resortów siłowych, a duża część danych pozyskiwana jest w wyniku pracy operacyjnej. Pierwszym krokiem, który wykonuje jego zespół, jest próba zdobycia tych samych danych z innego, niepowiązanego i zweryfikowanego źródła. Jeśli to zawiedzie, następuje sięganie po modele niepoprawnie nazywane w jego

---

88 Respondent B: kobieta, lat 62, samodzielny pracownik naukowy, kierownik instytutu, uczelnia państwowa, jako pracownik naukowy związana z uczelniami od ponad 35 lat. Respondentka nie wyraziła zgody na udostępnienie danych osobowych w postaci imienia i nazwiska oraz nazwy uczelni.

89 W przypadku osób, które – podobnie jak autor (K.K. Kuźma) – zajmują się kontrolą danych od kilkunastu lat, jest to możliwe do wychwycenia również w elektronicznej bazie danych.

90 Na odbywającym się w Szczecinie w 2013 r. Zjeździe Socjologicznym jeden z profesorów zajmujących się na co dzień metodologią i metodami nauk społecznych mówił o tym, że ankieter jest w stanie jedną ankietę jeszcze dobrze sfałszować, przy dwóch zaczyna się robić trudno, a przy trzech lub więcej można to łatwo wychwycić.

91 Respondent C: mężczyzna, lat 39, doktor, kapitan, kierownik zespołu odpowiedzialnego za weryfikację danych operacyjnych. Ze wspomnianym resortem związany od blisko 20 lat. Respondent nie wyraził zgody na udostępnienie danych osobowych w postaci imienia i nazwiska oraz nazwy organizacji.

organizacji ekonometrycznymi. Model ekonometryczny jest bowiem modelem matematyczno-statystycznym, a sama ekonometria jest bardzo często mylona ze statystyką. W rzeczywistości są to właśnie modele statystyczne, w dużej mierze korzystające z wnioskowania indukcyjnego, głównie kanonów Milla[92].

W podstawowej, najczęściej wykorzystywanej wersji postępowania następuje zebranie wszystkich (dostępnych) możliwych przyczyn i/lub skutków badanego zjawiska. Po ich zebraniu, korzystając z kanonów Milla (kanon jedynej zgodności[93], kanon jedynej różnicy[94], kanon zmian współtowarzyszących[95], kanon połączonej metody zgodności i kanon różnicy reszt), przeprowadza się dalszą analizę, zgodnie z algorytmem kanonów Milla, co pozwala z dużym prawdopodobieństwem[96] przewidzieć zależności pomiędzy przyczynami/skutkami, co dalej przekłada się na jakość pozyskanych w sposób operacyjny danych.

Podsumowanie

Zarówno autorzy, jak i eksperci, z którymi rozmawiano podczas przygotowania tego artykułu, solidarnie uważają, że omówione metody mają szeroki zakres zastosowania. Najwygodniejszymi i najprostszymi w użyciu będą omawiane w artykule metody statystyczne jako – z jednej strony – dające gwarancję bezpieczeństwa (poprzez swoje osadzenie w matematyce są one trudniejsze do podważenia), a z drugiej – dające względnie jednoznaczny wynik.

Jak pokazuje przykład respondenta C, metody statystyczne – jak choćby omówione kanony Milla – można wykorzystać do oceny jakości danych zebranych w toku pracy operacyjnej. W podobny sposób możemy wykorzystać rozkłady Gaussa i Poissona. Jeśli nasze dane – i w tym miejscu nie jest istotny sposób ich pozyskania – spełniają którykolwiek z nich, to prawdopodobieństwo tego, że są one dobre (poprawne) rośnie skokowo.

---

92 Kanony zostały sformułowane przez Johna Stuarta Milla w 1843 r. jako podstawowe – podówczas – schematy wnioskowania indukcyjnego, to jest takiego, którego celem jest wnioskowanie o prawdziwości na podstawie posiadanych przesłanek. Indukcja eliminacyjna Milla pozwala na wyszukiwanie związków przyczynowych pomiędzy zjawiskami.

93 Dotyczy związków pomiędzy przyczyną a skutkiem zjawiska.

94 Występuje, gdy możemy wskazać warunki niezbędne do zaistnienia określonej sytuacji.

95 Możemy zastosować wówczas, kiedy możemy obserwować zmiany w natężeniu badanego zjawiska w zależności od różnych sytuacji towarzyszących.

96 Zgodnie z informacjami od respondenta C – sprawdzalność przewidywań z wykorzystaniem kanonów Milla sięga około 85%.

Analogicznie jest w przypadku omawianych testów statystycznych. Jednym z przykładów omawianych w artykule jest liczba podpisanych umów raportowana przez pracownika. Test pozwoli wykryć, czy liczba umów nie jest sztucznie zawyżona. Może on także posłużyć do sprawdzenia wielkości wydatków czy – na przykład – weryfikacji liczby kontaktów osoby, której badaniem się zajmujemy.

Podsumowując, każda z metod ma swoje zastosowanie. A raczej wiele możliwych zastosowań. Staraliśmy się w artykule przedstawić w miarę szeroki katalog możliwości. Zdajemy sobie jednak sprawę, że temat nie został wyczerpany i pozostawia szerokie pole możliwości jego kontynuacji w kolejnych publikacjach.

Streszczenie

Artykuł porusza problem jakości danych otrzymywanych w procesie badań ilościowych z wykorzystaniem kwestionariuszy. Autorzy postanowili poruszyć ten temat, ponieważ w polskiej politologii – w opozycji do socjologii – zajmuje on stosunkowo mało miejsca. Tekst powstał głównie na bazie przeglądu literatury poświęconej poszczególnym poruszanym w nim problemom i stanowi jej syntezę. W trakcie prac odkryto, że dzięki stosunkowo prostym narzędziom, a także wykorzystaniu programów (na przykład Excel czy SPSS) można w łatwy i precyzyjny sposób sprawdzić jakość zbioru danych.

Konrad Gałuszko, Joanna Lewczuk, Konrad Krystian Kuźma

Validate? Verify? Do Not Move? About Non-Statistical, Scholastic and Quantitative Methods of Quantitative Data Storytelling

The article deals with the quality of data obtained in the quantitative research process. The authors decided to raise the subject, because in the Polish political science – in the opposition to sociology – it's not well described. The text was made mainly on the basis of a literature review devoted to particular parts of the article and is its synthesis. During the work it was discovered that thanks to relatively simple tools, as well as the use of some programs the quality of the data set can be checked in a simple and precise way.

Key words: *data quality assessment, statistical methods, logical control, statistical tests, normal distribution*

Bibliografia

Barnes S.A., *Identifying Fraud Can Be as Easy as 1, 2, 3: Applying Benford's Law to Forensic Analyses and Investigations*, https://files.constantcontact.com/85e69272601/6ec4247f-14c0-49cf-92a6-674a5a4c37c2.pdf (dostęp: 19.01.2019).

Benford F., *The Law of Anomalous Numbers*, „Proceedings of the American Philosophical Society" 1938, nr 78(4).

Blattman C. i in., *Measuring the Measurement Error. A Method to Qualitatively Validate Survey Data*, https://ore.exeter.ac.uk/repository/bitstream/handle/10871/31847/Measuring%20the%20measurement%20error.pdf?sequence=2&isAllowed=y (dostęp: 20.02.2019).

Bławat F., *Podstawy analizy ekonomicznej. Teorie, przykłady, zadania*, Warszawa 2011.

Brzeziński J., *Metodologia badań psychologicznych*, Warszawa 2005.

Carballo M., Hjelmar U. (red.), *Public Opinion Polling in a Globalized World*, Berlin 2008.

Cartledge P., Garnsey P., Gruen E.S. (red.), *Hellenistic Constructs: Essays in Culture, History, and Historiography*, Berkeley – Los Angeles – Londyn 1997.

Domański H., Dukaczewska A., *Stabilność odpowiedzi w badaniach socjologicznych*, „ASK. Research and Methods" 1996, nr 1.

Gostkowski Z., *Analiza „efektu panelowego" w badaniach wyborczych w Łodzi w 1961 r.*, [w:] tegoż (red.), *Analizy i próby technik badawczych w socjologii*, t. 1, Wrocław – Warszawa – Kraków 1966.

Jabkowski P., *Reprezentatywność badań reprezentatywnych. Analiza wybranych problemów metodologicznych oraz praktycznych w paradygmacie całkowitego błędu pomiaru*, Poznań 2015.

Mebane W.R. Jr., Kalinin K., *Comparative Election Fraud Detection*, APSA 2009 Toronto Meeting Paper, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1450078 2009 (dostęp: 19.01.2019).

Mider D., Marcinkowska A., *Analiza danych ilościowych dla politologów. Praktyczne wprowadzenie z wykorzystaniem programu GNU PSPP*, Warszawa 2013.

Mlodinow L., *Matematyka niepewności. Jak przypadki wpływają na nasz los*, przekł. P. Strzelecki, Warszawa 2009.

Nowak E., *Teoria „agenda setting" a nowe media*, „Studia Medioznawcze" 2016, nr 3(66).

Pawłowski T., *Logiczne podstawy weryfikacji wewnętrznej badań kwestionariuszowych*, [w:] Z. Gostkowski, J. Lutyński (red.), *Wywiad kwestionariuszowy w świetle badań metodologicznych*, Analizy i Próby Technik Badawczych w Socjologii, t. 4, Wrocław 1972.

Sambridge M., Tkalčić H., Arroucau P., *Benford's Law of First Digits: From Mathematical Curiosity to Change Detector*, „Asia Pacific Mathematics Newsletter" 2011, nr 1.

Sztabiński F., *Ocena jakości danych w badaniach surveyowych*, Warszawa 2011.

Taylor J.R., *Wstęp do analizy błędu pomiarowego*, przekł., A. Babiński, R. Bożek, Warszawa 2018.

STUDIA POLITOLOGICZNE VOL. 54
DOI: 10.33896/SPolit.2019.54.7

Patrycja Hrabiec-Hojda
ORCID: 0000-0001-7893-811X
Justyna Trzeciakowska
ORCID: 0000-0003-4917-4493

Techniki wyszukiwania informacji w mediach społecznościowych dla celów białego wywiadu

Słowa kluczowe:
media społecznościowe, SOCMINT, Social Media Intelligence, OSINT, biały wywiad

Wstęp

Media społecznościowe w 2019 roku będą miały ponad 2,77 mld użytkowników na całym świecie, z czego większość będzie korzystała z social mediów tylko przez telefon komórkowy[1]. Najpopularniejszą platformą społecznościową od lat pozostaje Facebook. W Polsce jest to trzecia najpopularniejsza strona, zaraz za Google i YouTube[2].

Największym wyzwaniem społeczeństwa informacyjnego już teraz jest opanowanie rosnącej wykładniczo liczby informacji oraz znalezienie sposobów na jej efektywne przetwarzanie i analizowanie. Media społecznościowe przyczyniają się do jeszcze szybszego wzrostu informacji i wzmagają informacyjną opresję. Przeciętny użytkownik Facebooka publikuje dziennie nawet 10 postów[3]. Wyobraźmy sobie, jaka liczba informacji

[1] *Number of Social Media Users Worldwide from 2010 to 2021 (in Billions)*, https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/ (dostęp: 26.01.2019).

[2] Według rankingu Alexa z 30.01.2019.

[3] *Digital 2018 Q4 Global Digital Statshot*, https://www.slideshare.net/DataReportal/digital-2018-q4-global-digital-statshot-october-2018-v2 (dostęp: 27.01.2019).

pojawia się tylko na jednej platformie społecznościowej dziennie, jeśli aż 26% całej ludzkości z niej korzysta. Taki przyrost informacji wymaga coraz lepszych narzędzi i technik ich przetwarzania. Szczególnie, że jak pokażemy w kolejnej części artykułu, większość publikowanych treści pozostaje poza zasięgiem głównego narzędzia do pozyskiwania informacji, czyli wyszukiwarki.

Poniższy artykuł traktować będzie o metodach wyszukiwania i pozyskiwania informacji w źródłach, jakimi są portale społecznościowe. Proponowane techniki wyszukiwawcze wynikają z wieloletniej praktyki autorek w obszarze infobrokeringu oraz białego wywiadu na potrzeby biznesu. Opiszemy trzy możliwe podejścia do pozyskiwania informacji: bezpośrednią, z wykorzystaniem narzędzi oraz z wykorzystaniem wyszukiwarki. Narzędzia opisane w artykule są na chwilę powstawania tego tekstu bezpłatne dla użytkowników. Nie będziemy odnosić się do narzędzi proponowanych przez duże firmy i będących rozwiązaniami płatnymi. Nie będziemy również poruszać tematu analizy i weryfikacji pozyskiwanych danych.

W 2012 roku po raz pierwszy użyte zostało określenie SOCMINT (*Social Media Intelligence*), czyli wykorzystanie mediów społecznościowych do pozyskania informacji w ramach tzw. białego wywiadu (OSINT). W niniejszym artykule skupimy się na technikach i narzędziach, jakie wykorzystuje się w praktyce researchu. „Pozyskiwanie danych ze źródeł otwartych nie opiera się jednak na zdobywaniu dostępu do nich kanałami oficjalnymi ani na gromadzeniu informacji, do których dostęp jest przez użytkowników zastrzeżony. Opiera się na założeniu, że wiele przydatnych i znaczących informacji jest dostępnych publicznie"[4]. Informacje dostępne na platformach społecznościowych są umieszczane dobrowolnie, z dużą samoekspresją użytkowników i często jeszcze większym brakiem rozwagi. Internauci publikują wiele informacji na swój temat, których w realnym świecie by tak łatwo nie ujawnili. Należy nadmienić, że rośnie również zagrożenie związane z celowo tworzonymi fałszywymi kontami, których celem jest między innymi szerzenie nieprawdziwych informacji, tzw. dezinformacji. Szerzej temat ten poruszono na końcu artykułu.

[4] P. Karasek, *Analiza informacji z mediów społecznościowych jako narzędzie wspierające kontrolę bezpieczeństwa w procedurach migracyjnych*, „Przegląd Bezpieczeństwa Wewnętrznego" 2018, nr 19, s. 200.

Czym są media społecznościowe i co czyni je fenomenem?

Początki mediów społecznościowych szacuje się na 1997 rok. Powstał wtedy amerykański portal pod nazwą SixDegrees. Jako pierwszy wprowadził model „kręgów znajomych" i pozwalał na podtrzymywanie relacji ze znajomymi i rodziną. Kolejnym portalem był MySpace w 2000 roku. Dopiero cztery lata później powstał dzisiejszy lider, czyli Facebook, a zaraz za nim Twitter[5]. W Polsce, zanim prym objął globalny Facebook, mieliśmy własne platformy, jak najpopularniejsza Nasza Klasa, o której wciąż warto pamiętać pod kątem pozyskiwania informacji, oraz bardziej niszowe, mniej znane Grono.net. Wielu użytkowników tych platform dziś nawet nie pamięta, że posiadali tam konta.

Minęło ponad 20 lat od powstania pierwszych mediów społecznościowych, a można przypuszczać, że nadal nie każdy rozumie wagę pozyskiwania informacji za ich pośrednictwem. O popularności social mediów jako źródeł informacji w pewnych kręgach może świadczyć fakt, że informacja szybciej roznosi się na Twitterze lub portalu Wykop.pl niż ukazuje się w portalach informacyjnych. Do grup użytkowników świadomie pozyskujących informacje z mediów społecznościowych należą na pewno dziennikarze, politycy i coraz częściej osoby związane z mediami i marketingiem. Pomimo stale rosnącego zainteresowania i promowania social mediów w biznesie można się spotkać z opinią, jakoby korzystanie z nich było „zbyt młodzieżowe", „zbyt niepoważne".

Ponieważ publikowanie w mediach społecznościowych niektórym jeszcze kojarzy się zbyt młodzieżowo, nie zawsze dopuszcza się myśl o pozyskiwaniu informacji z tego zasobu jako źródła godnego uwagi. O ile niektóre dyscypliny naukowe, takie jak antropologia, socjologia, marketing czy politologia, zaczęły badać przekazy publikowane w social mediach oraz ich wpływ na decyzje zakupowe, o tyle temat SOCMINT-u czy samych metod pozyskiwania informacji jest poruszany bardzo rzadko lub prawie w ogóle. W polskiej literaturze ukazują się publikacje, które zarysowują istnienie takiej metody pozyskiwania danych. Temat „białego wywiadu" również nie należy do zagadnień powszechnie opisanych ani zbadanych w polskim piśmiennictwie naukowym. W literaturze obcojęzycznej temat jest lepiej opracowany, głównie pod kątem użycia SOCMINT-u w bezpieczeństwie międzynarodowym i walce z terroryzmem.

[5] K. Merrell, *The History of Social Media: Social Networking Evolution!*, https://historycooperative.org/the-history-of-social-media/ (dostęp: 27.01.2019).

Nieliczne artykuły pokazują, jak można pozyskać informacje o konkurencji, preferencjach klientów czy o poglądach politycznych obywateli.

W mediach społecznościowych można znaleźć informacje praktycznie na każdy temat. Najogólniej podzielić można je na kilka grup:

1. Informacje o osobach (użytkownikach) – celem jest ustalenie jak największej liczby danych o konkretnej osobie: od statusu jej związku, pracy, zainteresowań, po relacje z innymi i poglądy[6].
2. Informacje o firmach/organizacjach:
   - informacje publikowane przez firmy/organizacje w określonych celach;
   - informacje publikowane przez osoby postronne, na przykład klientów na temat współpracy z firmą;
   - informacje publikowane przez osoby związane obecnie lub kiedyś z daną firmą;
   - informacje o produktach, markach, usługach;
   - informacje pozwalające rozwiązać problem osobisty, rodzinny lub związany z życiem zawodowym – informacje poradnikowe;
   - informacje związane z bieżącą sytuacją w kraju, gospodarce czy innym interesującym użytkownika obszarem;
   - informacje o zjawiskach, zagadnieniach, tematach interesujących dla użytkownika.

Media społecznościowe na przestrzeni lat intensywnie się zmieniały. Ewoluowały poszczególne platformy, ich popularność, grupy użytkowników, dochodziło też do przejęć, które intensywnie wpłynęły na zmianę globalnej mapy mediów społecznościowych. W 2011 roku autor nieaktualizowanego już dzisiaj bloga Oxyweb opublikował interaktywną mapę mediów społecznościowych, która pokazuje zmiany, jakie zaszły w strukturze platform społecznościowych na świecie w latach 2008–2011. Nie trzeba zaawansowanej analizy, by zauważyć, że zniknęło w tym czasie sporo lokalnych mniejszych platform, przejętych w części krajów przez najbardziej ekspansywnego gracza – Facebooka. Dziś najpopularniejsze obok niego platformy to Instagram[7], Twitter ze swoją specyfiką restrykcyjnego limitowania długości wpisów, LinkedIn dedykowany profesjonalistom na skalę niemal globalną.

Oprócz tego moglibyśmy wskazać na obecność mniej lub bardziej popularnych platform społecznościowych, które wyróżniają się także ze

[6] Należy wziąć pod uwagę, że informacje udostępniane są przez użytkowników deklaratywnie, a czasem część informacji świadomie przedstawiana jest w sposób nieprawdziwy.

[7] Właścicielem portalu Instagram jest firma Facebook.

względu na specyficzną grupę użytkowników, jak choćby ResearchGate dedykowany osobom związanym ze światem nauki. Warto też zwrócić uwagę na platformy przygotowane przede wszystkim dla użytkowników z określonych krajów: rosyjski odpowiednik Facebooka – Vkontakte, działające w Chinach WeChat – najpopularniejszy komunikator oraz ChinaWebo – odpowiednik Twittera. W krajach niemieckojęzycznych popularniejszą od LinkedIn siecią dla profesjonalistów jest Xing. Na uwagę zasługują także platformy społecznościowe, które dziś nie cieszą się już największym uznaniem internautów, ale wciąż stanowią repozytorium danych, których na próżno szukać w najpopularniejszych sieciach. W przypadku Polski można wskazać na przykład portal nk.pl (dawna Nasza Klasa) lub GoldenLine.

Nie jest naszym celem wyliczanie kolejnych platform społecznościowych i podkreślanie faktu ich istnienia. Chcemy jednak zwrócić uwagę na to, że przygotowując się do wyszukiwania informacji w mediach społecznościowych, należy zastanowić się nie tylko nad tym, jakich informacji, o kim i o czym potrzebujemy, ale także nad tym, które media społecznościowe powinniśmy przeszukać, aby zwiększyć swoje szanse na dotarcie do kluczowych dla nas informacji. Przy tym wszystkim trzeba zachować równowagę, bo nie zawsze więcej znaczy lepiej.

Żyjemy w czasach mediów społecznościowych, ponad 1/4 ludzkości komunikuje się za ich pomocą. Co takiego dają media społecznościowe, że internauci chcą z nich korzystać? Fenomen kryje się w łatwości dostępu do informacji o innych, poczuciu bycia na bieżąco, bezpieczeństwie komunikacji i w końcu łatwości w kreowaniu swojego wizerunku. Media społecznościowe dają też poczucie bycia w grupie, dla części osób jest to również sposób na podniesienie własnej pewności siebie.

Kluczem do wyszukiwania jest zrozumienie social mediów

Dobry research zaczyna się od zrozumienia tematu, na jaki będziemy szukać informacji, ustalenia słów kluczowych i początkowych źródeł. W przypadku mediów społecznościowych dochodzi jeszcze zrozumienie, jak one funkcjonują, jakie funkcjonalności działają na poszczególnych portalach, jakie obowiązują zasady itd.

Pierwszym ważnym krokiem jest zrozumienie, czym różni się hashtag od słowa kluczowego. Hashtag to słowo lub zlepek słów poprzedzony znakiem # – na przykład #artykuł. Jest to ważne z punktu widzenia funkcjonowania wyszukiwarek. W niektórych platformach społecznościo-

wych szukanie według hashtagów daje inne wyniki niż szukanie według słów kluczowych. Hashtagi są też istotne ze względu na ustalanie grup powiązanych ze sobą na przykład zainteresowaniem, uczestniczeniem w wydarzeniu czy manifestacją określonych poglądów lub treści celowo przez autorów wiązanych w cykle tematyczne.

Kolejną istotną kwestią w zrozumieniu mediów społecznościowych jest rozróżnienie, na jakich platformach jakiego rodzaju informacje możemy znaleźć. Jeśli chcemy sprawdzić osobę, z którą mamy nawiązać biznes, to w pierwszej kolejności sprawdzamy ją na platformie biznesowej, na przykład LinkedIn. Jednak ważne jest też, aby sprawdzić, co jest publikowane w kanałach bardziej prywatnych, na przykład na Facebooku, bo może okazać się, że wzorowy wizerunek z LinkedIn nie pokrywa się z wypowiedziami w innych portalach. W social mediach kreowanie własnego wizerunku jest bardzo łatwe i może zwodzić też na manowce.

Nie wszystko, co jest publikowane w social mediach, zostaje zaindeksowane w Google lub innej wyszukiwarce. Wynika to z technicznych zabezpieczeń mediów społecznościowych. Z jednej strony są to ustawienia prywatności na profilach użytkowników, z drugiej są to na przykład informacje, do których dostęp z założenia ma być ograniczony (przykładowo tylko dla członków określonej grupy).

W celu pozyskania informacji z mediów społecznościowych można użyć jednej z poniżej opisanych technik:

1. szukanie bezpośrednio w mediach społecznościowych za pomocą wbudowanej wewnętrznej wyszukiwarki oraz opcji wyszukiwania zaawansowanego;
2. szukanie z wykorzystaniem zewnętrznych narzędzi, które poprzez wpięcie API do mediów społecznościowych pozwalają na pobieranie danych i ich ustrukturyzowanie, lub narzędzi, których zadaniem jest stworzenie zaawansowanej kwerendy wyszukiwawczej do Google.
3. szukanie w wyszukiwarce z wykorzystaniem zaawansowanych operatorów i techniki *Boolean string*.

Bezpośrednie techniki wyszukiwania

Technikę wyszukiwania bezpośrednio w social mediach omówimy na przykładzie Facebooka i Twittera, ze względu na ich popularne użytkowanie w Polsce.

SZUKANIE NA FACEBOOKU

Chcąc wyszukiwać informacje na Facebooku o jego użytkownikach, musimy zacząć od właściwego skonfigurowania konta. Chodzi tutaj przede wszystkim o zmianę ustawień językowych tak, aby jako domyślny język Facebooka wskazany został angielski amerykański (*English US*). Pozwoli to na skorzystanie z wyszukiwarki Facebook Graph Search i na wyszukiwanie informacji o użytkownikach Facebooka nawet w warunkach, kiedy nie mamy kompletu informacji na ich temat. Przydaje się to szczególnie ze względu na fakt, że coraz więcej osób korzystających z Facebooka modyfikuje swoje imiona i nazwiska w taki sposób, by ich wyszukiwanie po podstawowych parametrach było trudniejsze. Aby skorzystać z tych możliwości wyszukiwawczych, konieczne będzie konstruowanie kwerend w języku angielskim.

Przykładowe kwerendy:

- *People named Anna Kowalska* – osoby, które nazywają się Anna Kowalska;
- *People who live in Warsaw* – osoby, które mieszkają w Warszawie;
- *People who work at XYZ* – osoby, które pracują w XYZ;
- *People who live in Warsaw and work at XYZ* – osoby, które mieszkają w Warszawie oraz pracują w XYZ;
- *People who are interested in marketing* – osoby, które interesują się marketingiem;
- *People who checked in Hotel X* – osoby, które oznaczyły swoją lokalizację na Facebooku w hotelu X.

W przypadku wszystkich tych informacji pamiętać należy w pierwszej kolejności, że każda informacja podawana na Facebooku (a także w innych mediach społecznościowych) ma tylko charakter deklaratywny i nie musi mieć wiele wspólnego z rzeczywistością.

Druga istotna sprawa dla nas, użytkowników korzystających z mediów społecznościowych jako narzędzi do wyszukiwania informacji, to fakt, że barierę, która będzie nas ograniczać w dostępie do informacji, stanowią ustawienia prywatności. Niewiele więc zdziałamy, jeśli chcemy wykorzystać Facebooka do profilowania osoby, której w gronie znajomych nie mamy, a która zadbała o ustawienia prywatności. Chodzi o ustawienie prywatności na konkretnych poziomach: jakie informacje mają być widoczne „tylko dla mnie", a jakie informacje mogą zobaczyć „znajomi moich znajomych" lub „wszyscy".

Jeszcze jeden fakt wpływający na dostępność informacji, z którego należy zdawać sobie sprawę w kontekście korzystania z Facebooka, to potęga grup. Grup, których zadaniem jest integrowanie użytkowników

o wspólnych potrzebach, zainteresowaniach itp. Trzeba mieć świadomość tego, że jeśli mamy do czynienia z grupą zamkniętą, nie będziemy mieć dostępu do jej zawartości, w tym aktywności jej członków, dopóki sami jej członkami nie zostaniemy. Tym samym nie poznamy pełnego spektrum zainteresowań/aktywności w mediach społecznościowych profilowanej osoby, a właściwie powinniśmy powiedzieć – konta. To ostatnie wynika bowiem z faktu, że nigdy nie mamy gwarancji, że osoba, której działania nas interesują, korzysta z mediów społecznościowych używając tylko konta założonego na swoje nazwisko. Dość powszechną praktyką jest korzystanie z tzw. avatarów, to jest kont założonych na niekiedy zupełnie fikcyjne dane.

Mówiąc o pozyskiwaniu treści z mediów społecznościowych, warto poruszyć także wątek ich regularnego pozyskiwania, w ramach monitoringu. Popularne, przede wszystkim wśród specjalistów marketingu, są narzędzia do monitoringu Internetu, w tym mediów społecznościowych. Wykorzystywane przede wszystkim do monitorowania, kto, gdzie i jak intensywnie dyskutuje o firmie czy marce, mogą kusić też specjalistów OSINT-u. Sam pomysł nie jest zły, jednak trzeba tutaj zwrócić uwagę na jedno istotne ograniczenie w działaniu tych narzędzi: nie mogą one monitorować zawartości grup działających na Facebooku. Nie dlatego, że ich twórcy nie potrafią opracować takich algorytmów. Z technicznego punktu widzenia jak najbardziej jest to możliwe, jednak w polityce Facebooka zawartość grup ma pozostawać treścią poniekąd ekskluzywną, dostępną jedynie dla członków danej grupy. Na etapie budowania strategii pozyskiwania informacji i jej wdrażania trzeba ten element koniecznie wziąć pod uwagę i albo zainwestować więcej czasu w przygotowanie warsztatu do pozyskiwania danych z Facebooka, albo pogodzić się z niedostępnością części treści.

SZUKANIE NA TWITTERZE

Twitter jest jednym z niewielu mediów społecznościowych, w których nie trzeba mieć konta, aby móc w nim szukać. W tym celu należy w dowolną wyszukiwarkę, choćby Google, wpisać kwerendę składającą się z nazwiska znanej osoby, która używa Twittera – na przykład „Robert Lewandowski Twitter". Wyszukiwarka zwróci nam linki do profili na portalu. Jeśli wyświetlimy czyjeś konto w portalu, to w prawym górnym rogu pojawia się pole wyszukiwania. Można rozpocząć wyszukiwanie. Wpisujemy poszukiwane hasło. Po ukazaniu się wyników wyszukiwania należy wybrać odpowiednią z dostępnych zakładek:

- najlepsze – wyniki posortowane przez wewnętrzny algorytm portalu;

- najnowsze – czyli tweety (posty) zawierające szukane słowo;
- użytkownik – czyli konta mające w nazwie lub w opisie konta szukane słowo;
- zdjęcia opatrzone szukanym słowem;
- filmy podpisane szukanym hasłem;
- news – posty podające linki do newsów zawierających szukane słowo;
- transmisje – posty zawierające nagrania wideo relacjonowane bezpośrednio na Twitterze, również nagrania archiwalne.

Po lewej stronie od wyników wyszukiwania dostępne jest pole o nazwie Dostępne filtry. W filtrach możemy ustalić parametr, taki jak język postów. Jest to szczególnie przydatne, gdy chcemy sprawdzić, co piszą media w innych państwach na temat na przykład polskich polityków czy sytuacji w naszym kraju. W tym polu dostępny będzie również link Wyszukiwanie zaawansowane. Tabela wyszukiwania zaawansowanego pozwala na (opis według kolejności pól):

1. szukanie pojedynczego słowa lub kilku słów;
2. szukanie dokładnego wyrażenia, na przykład „polityka zagraniczna";
3. wykluczenie słów – aby ta funkcja działała, należy wpisać słowo w pozycji pierwszej – na przykład wpisz w niej słowo polityka, a w pozycji trzeciej słowo zagraniczna; co oznacza: „pokaż wyniki, w których jest słowo polityka, ale nie ma słowa zagraniczna";
4. wyszukiwanie po hashtagu – bardzo ważne w przypadku szukania w mediach społecznościowych (jak wspomniano, słowo kluczowe i hashtag to nie to samo) – na przykład post o treści: „Jadę pociągiem na weekend #PKP #kochamkolej" – znajdziesz pod słowem pociąg i pod hashtagiem #PKP, ale pod słowem PKP już go nie będzie;
5. określenie języka, w jakim mają być posty – jest to powtórzenie z dostępnych filtrów, jednak daje możliwość stworzenia bardziej skomplikowanej kwerendy;
6. określenie miejsca, z którego został wysłany post – lokalizacja postów ustalana jest na podstawie adresów IP komputerów, z których zostały wysłane, w przypadku urządzeń mobilnych – z pomocą lokalizacji GPS;
7. określenie, jakie ma być zabarwienie emocjonalne wypowiedzi, której szukamy.

Każdy z parametrów dostępnych w tabeli ma również odpowiednik w komendach, które można wpisać bezpośrednio w pole wyszukiwania. Komenda near: pozwala na znalezienie postów ze wskazanej lokalizacji. Komenda -rt pozwala wykluczyć tak zwane retweety, czyli posty udostępnione od innego użytkownika.

Wyszukiwanie informacji na temat osób na Twitterze jest w porównaniu z Facebookiem o tyle bardziej problematyczne, że twórcy portalu nie przewidzieli różnych kont personalnych i – nazwijmy to – firmowych. Finalnie więc na etapie wyszukiwania informacji będziemy musieli zmierzyć się z koniecznością samodzielnego odfiltrowania kont personalnych od tych, które mogą należeć na przykład do firm, czasopism itd.

Narzędzia ułatwiające zdobywanie i analizę informacji w social mediach

Social media generują ogromne liczby danych, dlatego coraz częściej powstają specjalistyczne aplikacje lub programy komputerowe, których zadaniem jest monitorowanie lub agregowanie danych ze wskazanych platform. „Uzyskiwanie dostępu do informacji oraz ich zdobywanie może następować kompleksowo lub częściowo, w formie procesu zautomatyzowanego, bez konieczności uciążliwego ręcznego »klikania« w treści na portalu społecznościowym”[8], dlatego dalej prezentujemy kilka wybranych narzędzi ułatwiających pozyskiwanie informacji z mediów społecznościowych. Do opisu wybrano narzędzia dostępne dla każdego i do łatwego znalezienia w sieci.

STALKSCAN

Wyszukiwanie informacji na Facebooku można uczynić skuteczniejszym poprzez wykorzystanie kwerend w Facebook Graph Search. Gdy już przejdziemy do analizy profilu danej osoby, może jednak okazać się, że mamy do czynienia z jednostką mocno ekstrawertyczną, publikującą sporo treści, bardzo aktywną na Facebooku. Może to spowodować, że analiza profilu takiego użytkownika okaże się dla nas mocno uciążliwa, dlatego z pomocą przyjdzie nam narzędzie StalkScan.

Jest to bardzo prosta w obsłudze aplikacja, której zadaniem jest pokazanie w kawałkach profilu użytkownika Facebooka. Wklejamy link do analizowanego profilu lub podajemy numer użytkownika Facebooka (tzw. ID). Jeśli tylko mamy do czynienia z profilem, którego właściciel nie zadbał wystarczająco o ustawienia prywatności, będziemy mogli otwierać sobie wybrane „kawałki” analizowanego profilu. Jeśli klikniemy na przykład w ikonę *Pictures*, narzędzie przeniesie nas do zdjęć opublikowanych na analizowanym profilu, ikona *All* w sekcji *Places* pozwoli nam podejrzeć, w jakich miejscach sprawdzana osoba się oznaczyła.

[8] P. Karasek, *Analiza informacji z mediów społecznościowych…*, s. 201.

StalkScan – i inne podobne jej aplikacje – nie łamie ustawień prywatności, dlatego dobra konfiguracja tychże będzie zawsze stanowić barierę w pozyskiwaniu informacji z profili użytkowników.

Niektóre z tego rodzaju narzędzi ułatwiających pracę specjalisty OSINT-u będą wymagały podania tylko i wyłącznie numeru ID konta na Facebooku. W jaki sposób go pozyskać? Sposoby są dwa. Pierwszy sposób, to umiejętne analizowanie i przyglądanie się temu, co dzieje się w trakcie korzystania z narzędzi automatyzujących naszą pracę. Kiedy już korzystamy ze StalkScan i wygenerujemy sobie na przykład czyjeś *Pictures*, automatycznie w adresie URL ląduje numer ID konta, które jest poddawane analizie. Będzie to ciąg 15 cyfr jednoznacznie identyfikujący konkretne konto.

Drugi sposób to skorzystanie z narzędzia dedykowanego do odczytywania numeru ID, mianowicie aplikacji Find My Fb ID (www.findmyfbid.com). Mechanizm postępowania jest podobny jak w przypadku SkalkScan: wklejamy link do profilu użytkownika, a jako rezultat otrzymujemy ID konta, który możemy wykorzystać z powodzeniem w dalszych analizach.

SEARCH IS BACK

Search Is Back (searchisback.com) to jeszcze jedna aplikacja, której zadaniem jest ułatwianie nam wyszukiwania treści pochodzących z Facebooka. Praca na tym narzędziu jest bardzo intuicyjna, a ono samo wykorzystuje znane nam już metody tworzenia zapytań z wykorzystaniem Facebook Graph Search. Tyle tylko, że teraz możemy tę pracę zautomatyzować.

Na początek będziemy musieli zdefiniować, czego mają dotyczyć wyszukiwane przez nas informacje. Do wyboru mamy: osoby, zdjęcia, wydarzenia i posty. Jeśli zdecydujemy się na wyszukiwanie informacji o użytkownikach Facebooka, będziemy musieli w kolejnym kroku wybrać zbiorowość poddawaną analizie: wszyscy użytkownicy, tylko nasi znajomi, znajomi znajomych lub osoby spoza kręgu naszych facebookowych znajomych.

Na kolejnym etapie pozostanie nam – korzystając z dostępnych filtrów – definiować kryteria wyszukiwania osób. Do wyboru będziemy mieć parametry odwołujące się do deklarowanego na Facebooku miejsca pracy, miejsca zamieszkania, płci czy wreszcie imienia i nazwiska. Dodatkowo dla parametrów zmiennych, jak miejsce zamieszkania bądź zatrudnienie, mamy możliwość zdefiniowania, czy chodzi nam o stan obecny, o przeszłość, czy o oba te warianty.

Po zdefiniowaniu wymaganych parametrów wyszukiwania narzędzie wywoła nam już na Facebooku wyniki kwerendy. Podobny mechanizm działań będzie do wykorzystania podczas wyszukiwania treści innych niż profile użytkowników Facebooka.

RECRUITIN.NET

recruitin.net to portal typu *x-ray search*. To anglojęzyczne sformułowanie oznacza wyszukiwanie informacji z wykorzystaniem zaawansowanych komend wyszukiwawczych, o których opowiadamy w dalszej części niniejszego artykułu (*Google hacking*). Bardzo często o wyszukiwaniu *x-ray* mówi się w kontekście branży HR i wyszukiwania kandydatów do pracy oraz weryfikacji ich CV.

Recruitin to nic innego jak narzędzie, które stanowi swego rodzaju nakładkę na Google, ułatwiającą nam pisanie zaawansowanych kwerend w wyszukiwarce koncentrujących się na wybranym portalu: LinkedIn, Xing, Dribble, Twitter, GitHub, StackOverflow.

Za pomocą Recruitin możemy zdefiniować kryteria wyszukiwawcze, które, na przykładzie LinkedIn, będą dotyczyły miejsca zatrudnienia, zajmowanego stanowiska, kraju, w którym ma pracować nasz kandydat. Jeśli szukamy konkretnej osoby, jej imię i nazwisko trzeba będzie potraktować po prostu jako słowa kluczowe wymuszane w kwerendzie. Dla równowagi Recruitin pozwala nam też na zdefiniowanie słów, które wykluczamy z naszych wyników wyszukiwania.

Być może oczekiwalibyśmy, że od razu po uruchomieniu wyszukiwania Recruitin przeniesie nas do przeszukiwanego portalu, by wyświetlić nam wyniki wyszukiwania. Tymczasem jako rezultat kwerendy otrzymujemy tzw. *string* do Google, a więc kwerendę, którą możemy w Google uruchomić. Jako rezultat naszego przeszukiwania LinkedIn za pomocą Recruitin otrzymamy listę kont użytkowników LinkedIn spełniających nasze kryteria. Częstokroć profile te będziemy mogli zobaczyć w pełnej wersji, bez logowania do LinkedIn. Możliwości te zostaną ograniczone wtedy, gdy będziemy chcieli podejrzeć zbyt wiele kont lub wcześniej logowaliśmy się z tego samego komputera i IP do naszego konta w LinkedIn. Dostęp, przynajmniej częściowo anonimowy, do treści w Internecie to jednak osobny wątek i temat na odrębny artykuł.

ONEMILLIONTWEETMAP.COM

Onemilliontweetmap to aplikacja łącząca w sobie dwa intensywne trendy w podejściu do informacji: nieustannego monitoringu oraz wizualizacji. Na interaktywnej mapie otrzymujemy wizualizację intensywności

życia Twittera w czasie rzeczywistym, na bieżąco. Możemy prześledzić, w których częściach świata użytkownicy Twittera są najbardziej aktywni. Co więcej, narzędzie pozwala na analizowanie tego ruchu. Możemy wykorzystać w tym celu operatory, które zostały opisane we fragmencie dotyczącym wyszukiwania treści bezpośrednio na Twitterze.

Interesującą funkcjonalnością narzędzia jest możliwość wyfiltrowania tweetów opublikowanych w danej lokalizacji z możliwością ich czytania bezpośrednio na portalu, bez przechodzenia do Twittera i logowania się w portalu.

Możemy wykorzystać także mapę ciepła – popularnej w tego rodzaju narzędziach funkcjonalności, odzwierciedlającej zaangażowanie użytkowników Twittera.

Możliwości wykorzystania tego narzędzia do analizy ruchu na Twitterze są bardzo duże i tak naprawdę ograniczane tylko i wyłącznie wyobraźnią analityka. Nie jest to jednak jedyne warte uwagi narzędzie do analizy tego, co na Twitterze się dzieje, a dobór aplikacji będzie też finalnie zależał od subiektywnych upodobań i potrzeb poszukującego informacji.

TWITONOMY

By pokazać możliwości różnego podejścia do analizowania i przetwarzania treści pochodzących z Twittera, poświęcimy nieco uwagi narzędziu Twitonomy. To aplikacja, która – w odróżnieniu od Onemillion-tweetmap – każdorazowo będzie wymagać od użytkownika zalogowania i „wpięcia" jej do Twittera. Po zalogowaniu będziemy mogli przystąpić do wykorzystania możliwości aplikacji. Jest to narzędzie, dzięki któremu będziemy mogli monitorować wybrane konta na Twitterze, monitorować hashtagi i odczytywać pojawiające się w poszczególnych wątkach tweety bezpośrednio z poziomu Twitonomy, a nie Twittera. Będzie to oczywiście monitoring dokonywany w czasie rzeczywistym, a więc archiwum może obejmować, w zależności od intensywności dyskusji w danym wątku, kilka do kilkudziesięciu godzin.

Przez specjalistów marketingu narzędzie może być wykorzystywane także do analizy skuteczności działań na Twitterze, ale to oczywiście osobny wątek. Warto jednak zwrócić uwagę na fakt, że wielokrotnie, zajmując się OSINT-em, docieramy do miejsc, w których wykorzystujemy narzędzia pierwotnie przygotowane z myślą o grupach zawodowych zajmujących się innymi obszarami. W tym przypadku mówimy o marketingu, który jak najbardziej korzysta z dobrodziejstw OSINT-u, nie zawsze nazywając to wprost.

Podsumowując, przedstawiłyśmy kilka propozycji narzędzi, które dedykowane są przeszukiwaniu treści z platform społecznościowych. Jest to ledwie sygnalizacja możliwości, a temat do całościowego opisania jest tym trudniejszy, że same narzędzia, o których mowa, często się zmieniają. Pojawiają się, działają, po czym znikają i zastępowane są przez kolejne aplikacje. Wszystko to prowadzi to tego, że chcąc być na bieżąco z tematem, będziemy musieli testować różne narzędzia i uważnie się przyglądać ich możliwościom i ograniczeniom. Niewątpliwie w pewnym momencie natkniemy się też na barierę związaną z koniecznością uiszczenia opłaty za korzystanie z poszczególnych aplikacji. Coraz więcej narzędzi jest udostępnianych w modelu tzw. *freemium*, pozwalającym na korzystanie bez opłat do pewnego momentu, ograniczonego na przykład czasem lub liczbą zrealizowanych kwerend. Po wykorzystaniu limitu konieczne jest jednak opłacenie subskrypcji. Niektóre z narzędzi, w szczególności te najbardziej zaawansowane, w ogóle nie są dostępne do wytestowania bez ponoszenia opłat. Przykład to choćby Wynyard Social Media Analyzer – aplikacja stworzona przede wszystkim z myślą o prewencji incydentów kryminalnych i nacjonalistycznych. Nie jest to też narzędzie, na którego zakup będzie mogło pozwolić sobie mikroprzedsiębiorstwo z niewielkim budżetem.

Wyszukiwanie z wykorzystaniem technik *Google hacking*

Znajomość narzędzi do wyszukiwania treści w mediach społecznościowych niewątpliwie może ułatwiać życie. Warto jednak nie tylko dobrze rozumieć mechanizmy działania tych narzędzi, ich wady i zalety, możliwości i ograniczenia, ale też umieć niekiedy obejść się po prostu bez nich. Dobrą strategią będzie znajomość wspomnianych już technik *x-ray search*, a więc budowania kwerend wyszukiwawczych z wykorzystaniem komend, na przykład w wyszukiwarce Google.

W przypadku wyszukiwania informacji w Google przydatna jest znajomość składni adresu URL strony, którą chcemy przeszukać. Komend wyszukiwania zaawansowanego w wyszukiwarkach jest kilkanaście, jednak przy mediach społecznościowych najbardziej przydadzą nam się cztery z nich:

1. komenda site: na przykład „anna kowalska" site:facebook.com – pozwoli na wyszukiwanie osób o imieniu i nazwisku „Anna Kowalska" na portalu Facebook;
2. komenda inurl: na przykład inurl:polityka site:facebook.com/pages – pozwoli na wyszukanie stron na Facebooku, w których adresie www

pojawia się słowo polityka (bo tylko te są indeksowane w Google), postów na Facebooku mających w swojej treści słowo polityka, przykładowo: https://www.facebook.com/pages/biz/Polityka-w-Europie-1239066012883016/;

3. filetype: na przykład biotechnologia filetype:pdf site:researchgate.net – wymusi wyszukiwanie plików pdf ze słowem kluczowym „biotechnologia" w serwisie ResearchGate;
4. intitle: na przykład intitle: „Jan Kowalski" site:linkedin.com – Google wyszuka strony, które zawierają w tytule słowo „Jan Kowalski" w obrębie portalu LinkedIn.

Znajomość i umiejętność wykorzystania w praktyce operatorów wyszukiwania zaawansowanego daje dużo elastyczności w budowaniu kwerend. Pozwala też zrozumieć mechanizmy stojące za wyszukiwaniem treści, a finalnie także przygotować własne narzędzia automatyzujące kwerendy.

Podsumowanie

Media społecznościowe wywierają realny wpływ na życie polityczne i społeczne. Zbyt duża liczba informacji pojawiająca się w nich sprawia, że coraz więcej osób ma problem z ich weryfikacją. Stąd pojawiające się nierzadko doniesienia dotyczące manipulacji, prowokacji lub celowych działań dezinformacyjnych. Najnowsze dane pokazują, że prawie połowa kont w social mediach może być fałszywa. Zatem umiejętność skutecznego wyszukiwania informacji w tych zasobach staje się coraz istotniejsza. Opisane metody i narzędzia pomocne będą zarówno dla osób zajmujących się researchem, jak i dla analityków bezpieczeństwa, marketerów, handlowców czy managerów. Research w mediach społecznościowych pozwala ustrzec się przed pułapką *fake news* i innych nieprawdziwych informacji. Rośnie liczba osób, dla których opisywany kanał komunikacji staje się jedynym sposobem na pozyskiwanie informacji bieżącej o świecie. Jednak nie idzie za tym rosnąca świadomość zjawiska, jakim jest *bubble filters* (po polsku bańka informacyjna). O personalizacji wyników przez wyszukiwarkę wie już prawie każdy użytkownik Google. Jednak mało kto wie, że to, co ukazuje nam się w mediach społecznościowych, nie jest informacją obiektywną, ale dostosowaną do naszych preferencji, jakie ustalił algorytm na podstawie zgromadzonych o nas danych. W celu wyjścia z bańki informacyjnej konieczne jest weryfikowanie informacji, a tego nie da się zrobić bez rozwijania kompetencji cyfrowych w zakresie wyszukiwania informacji.

Streszczenie

Media społecznościowe stały się popularnym kanałem informacji dla dużej części społeczeństwa. Liczba informacji umieszczanych tam każdego dnia pozostaje w większości poza zasięgiem wyszukiwarek. Celem tego artykułu jest pokazanie technik i narzędzi, które można wykorzystać do przeszukiwania zasobów social mediów. Zaprezentowano tylko te narzędzia, które działają na rynku polskim i są bezpłatne przynajmniej w wersji podstawowej.

Patrycja Hrabiec-Hojda, Justyna Trzeciakowska

Information Retrieval Techniques in Social Media for Open Source Intelligence Purpose

Social media has become a popular information channel for a large part of society. The amount of information that is placed there every day remains mostly beyond the reach of search engines. The purpose of this article is to show the techniques and tools that can be used to search and explore social media. Only those tools that operate on the Polish market and are free at least in the basic version are presented.

Key words: *social media, SOCMINT, Social Media Intelligence, OSINT, information retrieval*

Bibliografia

Bartosik-Purgat M., *Media społecznościowe jako źródło informacji o produktach w świetle badań międzykulturowych – przykład Facebooka*, „Handel Wewnętrzny" 2016, nr 6.

Dzikowski J., *Wyszukiwanie danych osobowych w internecie dla celów informatyki śledczej*, „Studia Oeconomica Posnaniensia" 2013, nr 1.

Karasek P., *Analiza informacji z mediów społecznościowych jako narzędzie wspierające kontrolę bezpieczeństwa w procedurach migracyjnych*, „Przegląd Bezpieczeństwa Wewnętrznego" 2018, nr 19.

Merrell K., *The History of Social Media: Social Networking Evolution!,* https://historycooperative.org/the-history-of-social-media/ (dostęp: 27.01.2019).

Omad D., Bartlett J., Miller C., *Introducing Social Media Intelligence (SOCMINT)*, „Intelligence and National Security" 2012, nr 27 (6).

STUDIA POLITOLOGICZNE VOL. 54
DOI: 10.33896/SPolit.2019.54.8

Daniel Mider
ORCID: 0000-0003-2223-5997

Sztuka wyszukiwania w Internecie – autorski przegląd wybranych technik i narzędzi

Słowa kluczowe:
społeczeństwo informacyjne, biały wywiad, wywiad jawnoźródłowy, infobrokering

Wprowadzenie

Internet ma dualny charakter – można go rozpatrywać jako medium komunikacji[1] lub jako zbiornik danych[2], co wyznacza odrębne podejścia w zakresie pozyskiwania informacji[3]. W pierwszym wariancie źródło informacji stanowią osoby lub grupy osób, a platformy ich wyszukiwania to zasadniczo media społecznościowe oraz fora dyskusyjne. W takim ujęciu techniki eksploracji będą zogniskowane na kompetencjach efektywnej komunikacji interpersonalnej[4], mniejsze zaś znaczenie mają umiejętności

[1] Takie podejście ilustrowane jest np. przez: O.P. Ohiagu, *The Internet: The Medium of the Mass Media*, „Kiabara Journal of Humanities" 2011, nr 16(2).

[2] M. Bazzell, *Open Source Intelligence Techniques. Resources for Searching and Analyzing Online Information*, 6th ed., Charleston 2018.

[3] Bywają próby łączenia obu pojęć, jednakże w ograniczonym pod względem kanałów komunikacyjnych zakresie: F. Giglietto, L. Rossi, D. Bennato, *The Open Laboratory: Limits and Possibilities of Using Facebook, Twitter, and YouTube as a Research Data Source*, „Journal of Technology in Human Services" 2012, nr 30(3–4).

[4] Elementy werbalne, w tym znajomość – charakterystycznych dla poszczególnych grup – socjolektów, aspekty niewerbalne, w tym proksemiczne, kompetencje w zakresie stwarzania komfortu podczas rozmowy, a nawet umiejętności socjotechniczne.

techniczne, jednak i one odgrywają pewną rolę[5]. Niniejszy tekst zawiera analizę i próbę usystematyzowania narzędzi i technik eksploracji charakterystycznych dla drugiego z wymienionych sposobów podejścia do zjawiska – Internetu jako zbiornika danych. W tym kontekście kompetencje ogniskują się na informatyczno-technicznych elementach: biegłości w posługiwaniu się narzędziami wyszukiwawczymi (oprogramowanie), technikach eksploracji (szczegółowe sposoby formułowania zapytań) oraz taktykach (ogólne procedury wyszukiwania). Ich determinantami są kanały komunikacyjne (rozeznanie w różnorakich obszarach Internetu). Wysiłki poznawcze skupiono na tak zwanym powierzchniowym Internecie (*clearnet*)[6], a strukturę analiz wyznaczyły poszczególne techniki i narzędzia wyszukiwawcze.

Przegląd wybranych technik eksploracji Internetu zogniskowanych na wyszukiwarkach internetowych

Skuteczna procedura wyszukiwania w wyszukiwarkach każdorazowo wymaga opracowania i wdrożenia dwóch następujących elementów: haseł wyszukiwawczych oraz operatorów[7]. Hasła odnoszą się do wyszukiwanych treści, operatory zaś do zakresu zapytania (modyfikują je poprzez uszczegółowienie i dookreślenie).

Odnośnie do tworzenia haseł wyszukiwawczych można na podstawie doświadczeń własnych sformułować kilka ogólnych zasad. Pierwszorzędne znaczenie ma aspekt merytoryczny. Przede wszystkim należy utworzyć liczne, lecz adekwatne wersje wyszukiwanych haseł. W tym celu niezbędne jest wstępne rozeznanie w odniesieniu do wyszukiwanych treści. W toku tworzenia odrębnych, testowanych wedle przemyślanej kolejności haseł konieczne wydaje się uwzględnienie następujących elementów merytorycznych. Tworzenie haseł powinno się odbywać z uwzględnieniem refleksji o charakterze semantycznym, przede wszystkim sprawdzenia

[5] J.A. Benfield, W.J. Szlemko, *Internet-based Data Collection: Promises and Realities*, „Journal of Research Practice" 2006, nr 2(2).

[6] Pojęcie powierzchniowego (*Profound Web*, *Clearnet*) Internetu w zestawieniu z jego dopełnieniem logicznym, to jest Internetem głębokim/ukrytym (*Deep Web*, *Hidden Web*), zostało szerzej omówione w: D. Mider, *Mappa Mundi ukrytego Internetu. Próba kategoryzacji kanałów komunikacji i treści*, „EduAkcja. Magazyn Edukacji Elektronicznej" 2015, nr 2(10).

[7] W tym miejscu należałoby zaznaczyć istnienie jeszcze jednej istotnej dla wyszukiwania zmiennej – oceny jakości i wiarygodności źródeł informacji. Jest to jednak temat tak obszerny, iż wymaga odrębnego tekstu.

tożsamych nazw w różnych (najlepiej znanych wyszukującemu[8]) językach. Pomimo iż język angielski stanowi *lingua franca* Internetu, to liczne zasoby pozostają nietłumaczone na język angielski: z własnych doświadczeń wartymi polecenia wydają się treści publikowane w języku rosyjskim i ukraińskim, szczególnie, iż standardy i zasady obowiązywania własności intelektualnej różnią się w tych przypadkach od zachodnich. Konieczne wydaje się również przestudiowanie sposobów nazywania wyszukiwanych treści w socjolektach różnych grup społecznych. W Internecie rozmaite kanały komunikacyjne wymusiły specyficzne sposoby komunikacji (na przykład Snapchat, Twitter), powstały także społeczności wirtualne (subkultury) posługujące się rozmaitymi socjolektami (na przykład socjolekt używany przez użytkowników 4chan, 8chan czy grup JBWA w serwisie Facebook). Na te nowe podziały i sposoby komunikacji nakładają się odmienne w istotny sposób socjolekty różnych grup zawodowych czy zwolenników określonych opcji politycznych (różnice językowe wyznacza przede wszystkim oś prawica *versus* lewica). Różnice dotyczą nazewnictwa i odmiennego definiowania pojęć, uwzględniania kontekstów (sub) kulturowych i prowadzenia na ich podstawie werbalno-wizualnych gier słownych. Odnoszą się one również do sposobu kodowania informacji: zwielokrotniania, redukcji, zmiany znaków interpunkcyjnych i liter, stosowania lub nie zasad poprawnej pisowni, w tym znaków diakrytycznych danego języka, tworzenia form hybrydowych łączących tekst i grafikę (memy) oraz pisemnego oznaczania reakcji niewerbalnych i użycia emotikonów. Opracowanie listy haseł powinno również zawierać słownikowe ćwiczenie – zarówno odwołanie się do wokabularza synonimów i wyrazów bliskoznacznych, jak również uwzględnianie nazw pojęć w języku potocznym, publicystycznym i akademickim. Dla tworzenia listy haseł wyszukiwawczych istotna wydaje się również warstwa syntaktyczna, a więc szyk wyrazów w zdaniu (co warto każdorazowo testować *in vivo* w wyszukiwarkach), jak również alternatywne odmiany wyszukiwanych pojęć (o ile język, w którym wyszukujemy, je uwzględnia). Ważny okazuje się także techniczny aspekt tworzenia haseł. Niech egzemplifikacją będzie w tym zakresie wyszukiwarka Google. Wyszukuje ona maksymalnie 32 słowa w każdym zapytaniu, nadliczbowe słowa są przez nią ignorowane[9]. Wyszukiwarka Google nie odnotowuje różnic pomiędzy

[8] Niekiedy warto nawet skorzystać z automatycznego tłumaczenia treści z nieznanych wyszukującemu języków, gdyż treści takie mogą potencjalnie dostarczyć wartościowej informacji.

[9] Każda z wyszukiwarek ma odmienne charakterystyki.

słowami zapytań pisanymi wersalikami (majuskułą) a pismem zwykłym. Pozostaje również względnie niewrażliwa (własne testy ujawniają niewielkie różnice w wynikach wyszukiwania) na znaki diakrytyczne w językach narodowych, a także ignoruje szereg słów uznanych za zbyt krótkie, aby poddać je wyszukiwaniu – są to tak zwane *stop words*. Tym mianem określa się słowa uzupełniające tekst, lecz niemające samodzielnego sensu przenoszącego informację. Należą do nich przede wszystkim spójniki, zaimki, przyimki, jak również sformułowania typu: co to jest, jak rozumieć, na czym polega, jak definiować itd.[10] Warto również – szczególnie gdy przedmiotem poszukiwań są słowa używane potocznie lub gdy treści, w których wyszukujemy, są tworzone oddolnie – uwzględnić typowe błędy składniowe i językowe popełniane przez publikujących treści. Stworzoną listę haseł wyszukiwawczych wykorzystujemy w przemyślanej kolejności, dobrym zabiegiem wydaje się stosowanie „techniki lejka", według której podążamy od pojęć najogólniejszych do najbardziej wąskich.

Pracując z wyszukiwarką, należy zatem tworzyć kilka alternatywnych wersji hasła, biorąc pod uwagę zarówno warstwę semantyczną (nazwy w różnych językach, nazwy używane w socjolektach różnych grup społecznych, synonimy, uwzględnienie nazewnictwa potocznego, publicystycznego i akademickiego), jak i syntaktyczną (szyk wyrazów w zdaniu) oraz odmianę wyszukiwanych nazw. Niezbędna jest również opracowana strategia wyszukiwawcza i jednocześnie konsekwentne jej wdrażanie.

Kluczowe znaczenie w procesie wspomagania wyszukiwania mają operatory. Jest to pojęcie polisemiczne, które zakotwiczyło się również w informatyce, przynależąc do rodziny terminów z zakresu języków programowania oraz innych sposobów komunikowania się z komputerami, w szczególności języków zapytań (*query languages*). Operator jest to taka konstrukcja logiczna, której zadaniem jest zwracanie określonej wartości (wyniku działania, to jest transformaty) po wykonaniu działania na argumencie operatora (operandzie). W wyszukiwarkach internetowych najpowszechniej występują operatory przedrostkowe (prefiksowe), to jest takie, w których operand poprzedzany jest przez operator. W wyszukiwarkach internetowych operatory nie należą do złożonych, choć są liczne, różnorodne i niestandaryzowane. Charakteryzować je można przede wszystkim poprzez wykonywane działanie (ich funkcję). Ta właściwość stała się zatem przesłanką próby ich uporządkowania. Studium funkcjonowania operatorów w wyszukiwarkach internetowych stwarza przesłanki

[10] Z kompletną listą *stop words* w językach narodowych można zapoznać się w: *Stopwords*, https://www.ranks.nl/stopwords/ (dostęp: 24.01.2019).

do wyodrębnienia następujących klas (kryterium podziału są wykonywane przez nie działania): operatory logiczne, operatory lokalizacyjne i operatory kanałów komunikacyjnych, operatory chronometryczne, operatory eksploracji treści witryny oraz operatory wyszukiwania określonych typów treści.

Operatory logiczne służą do działań dokonywanych bezpośrednio na treści operandów. Rudyment i powszechnik w językach zapytań stanowią następujące trzy (spośród pięciu) elementy algebry stworzonej przez brytyjskiego matematyka George'a Boole'a: jednoargumentowy operator negacji (dopełnienia, „nie", zaprzeczenia logicznego), dwuargumentowa koniunkcja (iloczyn, logiczne „i") oraz alternatywa (suma, logiczne „lub").

Operator negacji działa tak, iż poprzedzone nim słowa nie pojawią się w wynikach wyszukiwania. Do zapisu negacji w wyszukiwarce Google (G), Yahoo! (Y!), Yandex (Y), Bing (B) i DuckDuckGo (DDG) służy dywiz/minus (znak: -). W ostatnich dwóch z wymienionych można alternatywnie stosować słowo „NOT". Operator ten występuje powszechnie w wyszukiwarkach oraz innych programach wykorzystujących języki zapytania. Sugeruje się, by używać go w formule „lejka", to znaczy dodawać kolejne słowa poprzedzone tym operatorem zapoznawszy się już z wynikami wyszukiwania i sekwencyjnie zawężać obszar wyszukiwań. Na przykład zapytanie „daniel mider" skierowane do wyszukiwarki Bing generuje w pierwszej dziesiątce również wyniki dla amerykańskiego operatora filmowego Daniela Richarda Modera. W celu eliminacji tych wyników zapytanie należy sformułować następująco: daniel mider -moder.

Zastosowanie operatora koniunkcji, zwanego również operatorem włączania, powoduje, iż każdy poprzedzony nim operand jest wymagany w wynikach wyszukiwania, a wyniki częściowo spełniające kryterium nie pojawią się. Najpopularniejszym zapisem tego operatora jest znak dodawania: +. W Bing można zastępować go znakiem et („etką", handlowym „i"), to jest: &, w Google zaś stosować słowo „AND". Operator ten może się pojawiać przed jednym, wieloma lub wszystkimi elementami sformułowanego zapytania. Zastosować można go jak następuje: +mider +cyberterroryzm, by wyświetlić wyniki wyszukiwania łączące nazwisko Mider z pojęciem cyberterroryzmu.

Znajdującym najmniej zastosowań w praktycznych wyszukiwaniach jest operator alternatywy. W logice dwuwartościowych predykatów jego zastosowanie zwraca „prawda", jeśli co najmniej jedna z operand jest prawdą. A zatem służy on do wyszukiwania co najmniej jednego z operandów zawartych w zapytaniu do wyszukiwarki. Zapisuje się go w postaci znaku pisarskiego kreski pionowej (*pipe*) w Bing (w Bing także podwój-

nej) i Yahoo! albo słowa OR w Bing, Google, Yahoo! i DuckDuckGo: | (kod ASCII: 124, dostępny po naciśnięciu na klawiaturze klawiszy Shift + \)[11]. Nie występuje on w Yandex. Wpisanie: cyberterroryzm | cyberprzestępczość | cyberzagrożenia sprawi, iż wyszukane zostaną te strony, gdzie znajduje się choć jedno z trzech wymienionych pojęć.

Funkcjonują dwa operatory zastępowania znaków – operator zastępowania ciągu znaków reprezentowany przez asterysk (znak: *, uzyskiwany po wciśnięciu Shift + 8) oraz operator zastępowania pojedynczego znaku zapisywany jako kropka (znak: .). Operatory te pozwalają na wyszukiwanie pojęć lub fraz, których dokładnego (poprawnego) zapisu nie znamy. Przykładowy zapis: m.der winien zwrócić wszystkie słowa zawierające wskazane litery oraz dowolny znak pomiędzy pierwszą a pozostałymi (w tym na przykład spację lub kropkę). Zastosowanie asterysku wydłuża dowolnie odległość pomiędzy wpisanymi znakami, a zatem ten operator nadaje się raczej do fraz. Praktyka zastosowania powyższych operatorów dowodzi, iż w różnych wyszukiwarkach mają one umiarkowane wyniki wyszukiwania, ponadto są one wrażliwe na spacje. Operator zastępowania ciągu znaków jest powszechniejszy (B, DDG, G, Y, Y!) niż operator zastępowania pojedynczego znaku (tylko G i deklaratywnie w Y!).

Wyszukiwarki (B, DDG, Y, Y!, w mniejszym stopniu G) umożliwiają również wyszukiwanie danych numerycznych według podawanych przez użytkownika zakresów, na przykład cen czy rozmiarów. Służy do tego operator zakresu zapisywany za pomocą dwóch kropek (następująco: ..). Zapis wyszukiwania: +BMW +650GS +cc600..cc800 lub +BMW +650GS +cc +600..800 dla hipotetycznej giełdy motocykli zwróci jako wynik wskazane modele motocykli, lecz o pojemności zawartej pomiędzy 600 a 800 cm^3. Zapis cen oznaczamy jak następuje: $100..$400. Operator ten działa niesatysfakcjonująco, jego funkcjonowanie polepsza użycie go w parze z operatorem site:.

Zaimplementowane w wyszukiwarkach, a przede wszystkim w Google, mechanizmy ułatwiające wyszukiwanie poprzez rozszerzenie jego zakresu i zasugerowanie użytkownikowi „właściwego", to jest najczęściej wyszukiwanego wyniku lub mechanizmy sztucznej inteligencji rozpoznające składnię i znaczenie zapytań (na przykład Koliber w Google) paradoksalnie utrudniają wyszukiwanie. Istnieje jednak w nielicznych wyszukiwarkach (DDG, Y) mechanizm wyłączający działanie powyższych algorytmów i umożliwiający użytkownikowi wyszukanie dokładne – co

[11] Niekiedy na klawiaturze reprezentowany jest przez złamaną (przerywaną) pionową kreskę: ¦.

do znaku i bez zmian rozszerzających. Operator ten oznaczany jest za pomocą wykrzyknika (znak: !), a zapytanie: „!daniela !midera" oznacza, by wyszukać nazwisko dokładnie w takiej formie: „Daniela Midera" (nie jest to pomyłka). DuckDuckGo umożliwia stosowanie tego operatora dla innych wyszukiwarek jako pośrednik. W menu Google można zamiast powyższego operatora użyć zakładki Narzędzia → Dokładnie.

Występują dwa typy operatorów grupowania. Pierwszy z nich umożliwia wyszukiwanie dokładnie jak wprowadzono (we wpisanej kolejności) i nazywany jest operatorem grupowania uporządkowanego. Do jego zapisania używamy cudzysłowu (znaki: „"). Zapytanie: „volenti non fit iniuria" zwróci wyłącznie tak zapisaną łacińską sentencję (kolejność słów jak wprowadzono). Operator ten działa wadliwie lub nie w pełni (B, DDG, Y!). Drugi – operator grupowania nieuporządkowanego – reprezentowany jest przez parę nawiasów (znaki: ()). Wyszukuje wyrazy umieszczone w nawiasie, jednakże mogą one występować w kolejności dowolnej.

Operatory lokalizacyjne służą do filtrowania wyników wyszukiwania wedle całości lub części adresu internetowego (dowolnie zapisanego: w postaci adresu IP, nazwy domeny, URL, nazwy DNS, ale także lokalizacji geograficznej lub geograficzno-językowej), jest ich kilkanaście.

Kluczowym operatorem jest operator site: zawężający wyniki wyszukiwania do danej strony (i podstron). Zapytanie: mider site:www.inp.uw.edu.pl wyszukuje nazwisko Mider wyłącznie w witrynie Instytutu Nauk Politycznych UW. Operatora tego można również używać jako samodzielnego – wówczas efektem jego działania jest enumeratywna lista podstron danej witryny (prezentowana jednak w sposób nieuporządkowany z punktu widzenia jej struktury).

Odmienne jest działanie operatora URL[12]: wyszukującego adres danej strony, gdziekolwiek się on znajduje. Prawidłowy zapis zapytania jest następujący: url:http://www.inp.uw.edu.pl/. Można zastąpić go innymi operatorami, jego istnienie nie jest niezbędne. Jego pochodną są operatory inurl:, allinurl:, url: (działają tylko dla G i DDG, ten ostatni dla Y). Wyszukują one słowo/słowa użyte w adresie URL danej witryny. Na przykład zapytanie allinurl:inp uw spowoduje, iż wyszukane zostaną dokumenty mające w adresie URL zarówno słowo „inp", jak i „uw". Operatory te reagują wyłącznie na słowa, nie zaś na składniki adresu URL (na

[12] Nazwa jest abrewiaturą i w pełnej wersji brzmi: *Uniform Resource Locator*. Oznacza standaryzowany na podstawie dokumentu RFC 1738 format adresowania zasobów w sieci globalnej i sieciach lokalnych. Składa się z trzech elementów: protokołu (np. http, https, ftp, telnet, nntp, mailto), adresu serwera oraz (opcjonalnie) ścieżki do zasobu.

przykład kropka, dwukropek, prawy ukośnik) i nie istnieje możliwość ominięcia tego ograniczenia. Różnica pomiędzy inurl: oraz allinurl: jest następująca: allinurl:mider daniel spowoduje wyszukanie takich adresów, które zawierają słowa zarówno Daniel, jak i Mider. Z kolei inurl:mider daniel wyszuka adresy URL ze słowem Mider oraz słowem Daniel w dowolnym miejscu strony. Zapytania: inurl:mider inurl:daniel oraz allinurl:mider daniel są tożsame. Dobrym sposobem wyszukiwania subdomen (poddomen) przypisanych do danej domeny jest operator domain: (B, DDG, Y!). Następujące sformułowanie: domain: inp.uw.edu.pl lokalizuje przykładowo subdomeny takie jak gpss.inp.uw.edu.pl czy podyplomowe.inp.uw.edu.pl. Operator ip: (wyłącznie dla: B, DDG, Y!) zwraca adres DNS (*Domain Name System*), czyli nazwę mnemoniczną, łatwą do posługiwania się w komunikacji międzyludzkiej. Zapis: ip:86.111.240.162 spowoduje wyświetlenie się wyników dla adresu www.inp.uw.edu.pl. Operatory host: i rhost: działają w Yandex i wydają się działać w Bing, choć nie w pełni prawidłowo. Wyszukują one podstrony witryny, jednak wyłącznie w ramach danego hosta, to jest maszyny, na której zamieszczone są zasoby.

Możliwe jest również wyszukiwanie zasobów wedle lokalizacji stron (serwerów, na których się znajdują). Służą do tego celu location: i loc: dla Bing, region: i r: dla DuckDuckGo oraz cat: dla Yandex. Wymienione operatory zawężają wyniki wyszukiwania do stron znajdujących się w określonej lokalizacji geograficznej. Bing akceptuje dwuliterowe kody normy ISO 3166-1 (tzw. kod alfa-2)[13]. Wyszukiwarka Yandex miała możliwość wyszukiwania identyfikatorów tematycznych za pomocą operatora cat:[14]. Wyszukiwanie odbywało się w katalogu Yandex. Obecnie nie funkcjonuje. Google oferuje tę usługę w Wyszukiwaniu zaawansowanym. Wyszukiwarki Yandex i Bing wprowadziły dodatkowo możliwość wyszukiwania jednocześnie lokalizacji języka danej strony. Operandy powinny być formułowane (prawdopodobnie!) według normy ISO 639-1. Na przykład zapytanie: institute altloc:pl-en wyszukuje polskie strony (lokalizacja) w języku angielskim zawierające słowo „institute".

Blisko spokrewnione z wyżej analizowanymi są operatory kanałów komunikacyjnych w mediach społecznościowych – umożliwiają ograniczenie wyszukiwania do wybranych obszarów mediów społecznościowych. Hasztag (*hashtag*, w skrócie tag) jest to pojedyncze słowo lub

[13] *Lista kodów dla Bing*, https://pl.wikipedia.org/wiki/ISO_3166-1 (dostęp: 12.02.2019); *Lista kodów dla DuckDuckGo*, https://duckduckgo.com/params (dostęp: 12.02.2019).

[14] *Lista kodów regionalnych Yandex*, *Lista kodów tematycznych Yandex*, http://search.yaca.yandex.ru/cat.c2n (dostęp: 12.02.2019).

fraza nierozdzielona spacjami, poprzedzone symbolem # (*hash*, kratka, krzyżyk lub płotek). Jest to forma znacznika umożliwiająca w mediach społecznościowych oddolne, niehierarchiczne grupowanie wiadomości[15]. Usługa wyszukiwania w hasztagach mediów społecznościowych funkcjonuje w Google i Yahoo! oraz w mniejszym stopniu w DuckDuckGo i Bing (wadliwie). Wpisanie: #covfefe w wymienionych wyszukiwarkach spowoduje odszukanie w mediach społecznościowych wiadomości oznaczonych tym właśnie tagiem. Z kolei poprzedzenie nazwy własnej użytkownika znakiem @ (*at, commercial at*, małpa) pozwala na wyszukiwanie profili w social mediach (B, DDG, G, Y!). Operator blogurl: wyszukujący blogi w określonej domenie funkcjonuje wyłącznie w Google. Zapytanie: blogurl:inp.uw.edu.pl spowoduje wyszukanie blogów w domenie Instytutu Nauk Politycznych UW. Jedynie w Bing można skorzystać z operatora feed: odnajdującego kanały RSS (*Really Simple Sindication*) / Atom mające wskazaną nazwę. Przydatnym rozwiązaniem może okazać się również hasfeed: pozwalający na sprawdzanie, czy w danej domenie ulokowano kanały RSS. Operatory kanałów komunikacyjnych mediów społecznościowych to dość siermiężne narzędzie – istnieją zautomatyzowane, profesjonalne programy pozwalające na wielokrotnie bardziej efektywne i precyzyjne wyszukiwanie w social mediach (na przykład mechanizm wyszukiwania dla mikrobloga Twitter wbudowany w program Maltego).

Kluczowe znaczenie mają operatory chronometryczne pozwalające na zawężenie wyników wyszukiwania do określonych przedziałów lub punktów czasu. Możliwość takiego wyszukiwania jest ze względu na wygodę użytkowników realizowana w innej formule – wyszukiwania okienkowego i z użyciem menu. Operator date: pozwalający na wyszukiwanie punktowe (dla dnia, miesiąca, roku) pozostawiono wyłącznie w wyszukiwarce Yandex. W Google, a do niedawna i w Yandex, funkcjonuje operator daterange: oferujący wyszukiwanie w zakresach dat. W Google wymagał użycia daty juliańskiej[16] oraz rozdzielenia wskazywanych przez użytkownika dat dywizem, a w Yandex – dwiema kropkami. Na przykład 4 kwietnia 2018 to wedle zapisu daterange:2458212.500000. Operatory te są na tyle ważne, iż zaimplementowano je w trybie okienkowym (kalendarzowym), a w trybie linii poleceń – zmarginalizowano. Funkcjonującym i przydatnym spośród operatorów chronometrycznych jest cache: poka-

[15] Przede wszystkim dotyczy to Facebooka, Instagrama i Twittera.

[16] Dla ułatwienia zamiany daty gregoriańskiej na juliańską można było skorzystać z następującego konwertera: *Julian Date Converter*, http://aa.usno.navy.mil/data/docs/JulianDate.php (dostęp: 12.02.2019).

zujący wersję strony z pamięci podręcznej wyszukiwarki, a więc stronę (najprawdopodobniej) archiwalną. Funkcjonuje on tylko w Google, z kolei w Bing umożliwia wyszukiwanie poprzez funkcję Zbuforowano (strzałka skierowana w dół przy odnośniku).

Operatory wyszukiwania w treści strony umożliwiają systematyczne przeszukiwanie witryn internetowych pod kątem określonych wartości w podziale na tytuł witryny, treść witryny, jej metaznaczniki i elementy informacyjne oraz inne. Operatory wyszukiwania w treści strony to operatory elementarne, istnieją stosunkowo długo, powstały wraz z Web 1.0.

Wyszukiwanie odnośników w tekście strony jest możliwe za pomocą operatora inanchor:. W HTML *anchor* oznacza treść umieszczaną pomiędzy znacznikami <a> oraz . Taka oto treść jest wyświetlana na stronie www odnośnik (link):

Instytut Nauk Politycznych UW

Aby wyszukać taką treść na przykład na stronie Wydziału Nauk Politycznych i Studiów Międzynarodowych wpisujemy: site:wnpism.uw.edu.pl inanchor:instytut – uzyskujemy zwrot takich stron, dla których link stanowi słowo „instytut".

Potrzebne, lecz niedziałające są operatory linkfromdomain: (B) i link: (G, Y!) – służą uwidacznianiu URL, do których zawierają odnośniki. A zatem potencjalnie można byłoby odnajdywać wszystkie te strony, które zawierają odnośniki, linkują stronę będącą przedmiotem wyszukiwania.

Istnieje szereg operatorów umożliwiających precyzyjne, ukierunkowane wyszukiwanie zarówno w nagłówku strony www (informacje zawarte pomiędzy znacznikami nagłówka <head></head>), jak i jej treści (tzw. body, to jest zawartości z ulokowanej między znacznikami HTML <body></body>).

Stronę internetową opisują w sposób najbardziej ogólny tak zwane metaznaczniki informujące ogólnie o jej treści, standardach językowych i formatowaniu. Znacznik meta: pozwala na wyszukiwanie słów zamieszczonych w nagłówku strony w metaznaczniku „keywords", który zawiera ustalone przez autora strony rozdzielone przecinkami słowa lub frazy opisujące treść strony. W przypadku strony Instytutu Nauk Politycznych UW ów fragment kodu jest następujący:

<meta name="Keywords" content="nauki polityczne, politologia, bezpieczeństwo wewnętrzne, administracja rządowa, administracja publiczna, studia podyplomowe, studia bezpieczeństwa, zarządzanie systemami bezpieczeństwa, bezpieczeństwo narodowe, europeistyka, marketing polityczny, studia dzienne, studia wieczorowe, studia zaoczne,

studia licencjackie, studia magisterskie, rekrutacja uw, rekrutacja 2011, licencjat na uw, studia na uw, magister na uw" />

W celu wyszukania słowa kluczowego politologia w wyszukiwarce Bing (w innych operator nie działa, w Bing funkcjonuje – już lub jeszcze – wadliwie) użyjemy następująco sformułowanego zapisu: meta:politologia. Bing wprowadził analogiczny, lecz jeszcze bardziej restrykcyjny operator literalmeta:, jednak w chwili obecnej nie wydaje się on działać zgodnie z przeznaczeniem.

Informacje identyfikujące stronę www mogą być również przez autorów strony umieszczone w metaznaczniku „description":

<meta name="Description" content="Politologia i bezpieczeństwo wewnętrzne w Instytucie Nauk Politycznych Uniwersytetu Warszawskiego. Studia licencjackie i studia magisterskie." />

Wyszukiwanie w ramach wymienionych treści odbywa się za sprawą operatora info: (wyłącznie w Google). Jako operand wystąpić może zarówno słowo, jak i fraza. W drugim z przypadków należy ująć ją w cudzysłów. Ze względu na mechanizmy rozszerzające wyszukiwanie operator ten działa w Google niesatysfakcjonująco. Analogicznie funkcjonuje para operatorów intitle: (title: dla Yandex) i allintitle:. Operatory ograniczają wyniki do stron zawierających wszystkie wyrazy zapytania w tytule strony. Tytuł strony znajduje się w nagłówku źródła strony pomiędzy znacznikami HTML <title> i </title>. Zamieszczane są tam przez autorów/programistów stron internetowych. Z opcji tej można także skorzystać w Google na stronie Szukanie zaawansowane. Różnica pomiędzy rozważanymi operatorami jest następująca: intitle: jest operatorem jednowyrazowym, a allintitle: umieszczamy przed frazami. Są to operatory uniwersalne, dostępne we wszystkich pięciu analizowanych wyszukiwarkach.

W wyszukiwarkach Bing i Yahoo! i w nieco mniejszym stopniu w Yandex można wyszukiwać według języka strony www zadeklarowanego w nagłówku HTML. Przykładowy kod zawierający informację o języku strony wygląda jak następuje:

<html class="…" lang="en">

<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="pl">

Wyszukiwania dokonuje się za pomocą operatora language: (B, Y!) lub lang: (Y). Przykładowe zapytanie dla Bing: „politologia" lang:en. Kod języka pozyskuje się z predefiniowanych list[17].

[17] *Lista kodów języka dla Bing*, https://msdn.microsoft.com/en-us/library/dd250941.aspx (dostęp: 12.02.2019); *Lista kodów dla Yandex*, https://tech.yandex.com/translate/doc/dg/

Wyszukiwanie treści na stronach internetowych umożliwiają następujące operatory: intext: w DDG, allintext: w Google, Yahoo! i Yandex oraz inbody: w Bing. Najczęściej łączone są z operatorem site:.

Metaoperatory keyword: i instreamset: (Y, B) służą do wyszukiwania w ramach innych operatorów. Następujący rozkaz: keyword:(intitle inbody)infobrokering pozwoli wyszukać słowo „infobrokering" jednocześnie w tytule i w tekście dokumentu.

Przydatne rozwiązanie stanowi operator prefer:. W zamierzeniach twórców ma podkreślać dane słowo pośród innych wyszukiwanych, zwiększając umieszczenie na wyższych pozycjach w wyszukiwaniu stron zawierających to słowo. Działa wadliwie. Lepsze efekty uzyskujemy (sposób dla Google i Bing), jeśli wielokrotne powtórzymy dane słowo w zapytaniu.

Operatory near: oraz around: (tylko B i G, lecz i tam działają niesatysfakcjonująco) określają maksymalną odległość wyszukiwanych od siebie słów. Na przykład zapytanie w Google: daniel around(2) mider będzie oznaczało, że wyszukane mają zostać strony, na których słowa „daniel" i „mider" znajdują się maksymalnie w odległości dwóch wyrazów (oddzielone są maksymalnie dwoma wyrazami).

Operatory wyszukiwania określonych typów treści zapewniają odnajdywanie ściśle określonych informacji, na przykład ściśle zdefiniowanych typów plików (Word, Excel), informacji pogodowych czy definicji słownikowych i encyklopedycznych. Jest to zróżnicowana tematycznie grupa operatorów, jednak największa ich liczba występuje w Google.

Najbardziej przydatna i występująca we wszystkich wyszukiwarkach jest możliwość dookreślenia typu wyszukiwanych zasobów poprzez rozszerzenie pliku. Odbywa się to z użyciem operatora filetype: (G, DDG, Y!), mime: (Y), ext: (B – obecnie przestało działać). Operator wyszukuje określone typy plików, na przykład arkusze kalkulacyjne (.xls, .xlsx), dokumenty tekstowe (.doc, .docx, .odt) itd. Wstawiamy je bez kropki. A zatem zapytanie: site:inp.uw.edu.pl filetype:docx zwróci listę plików Word (Office 2007) dla www Instytutu Nauk Politycznych. W Bing i Yandex obecny jest operator contains:. Za jego pomocą odnajdujemy strony (a nie same dokumenty jak wyżej) zawierające linki do dokumentów o rozszerzeniach określonych w danym operatorze. Na przykład: contains:doc site:inp.uw.edu.pl zwraca wszystkie strony w INP, na których znalazły się odnośniki do dokumentów tekstowych edytora.

concepts/api-overview-docpage/ (dostęp: 12.02.2019); *Lista kodów dla Yahoo!*, https://developer.yahoo.com/search/languages.html (dostęp: 12.02.2019).

Google oraz Yahoo! umożliwiają przeszukiwanie treści encyklopedii, leksykonów i słowników. Operator define: umożliwia wyszukiwanie pośród tych zasobów według słowa lub frazy. Z kolei operator related: powoduje wyświetlenie listy stron „podobnych" do określonej strony internetowej. Na przykład zapytanie related:www.inp.uw.edu.pl spowoduje wyświetlenie stron internetowych, które są podobne do strony głównej Instytutu Nauk Politycznych UW. Z opcji tej można także skorzystać, wybierając Podobne strony na głównej stronie z wynikami wyszukiwania Google oraz na stronie Szukanie zaawansowane w sekcji Informacje o danej stronie internetowej → Podobne do. Z kolei znak ~ (tylda) to operator wyszukiwania synonimów. Działa jednak wadliwie lub nie działa wcale. Z wyszukiwarki Google można skorzystać również do pozyskania informacji liczbowych. Służy do tego operator convert: oferujący przeliczanie według kursów walut, a także rozmaitych miar (wagi, odległości)[18].

Możliwe jest również wyszukiwanie treści multimedialnych z użyciem operatorów, choć odchodzi się już od tego rozwiązania na rzecz prostszego, bardziej przejrzystego trybu okienkowego. Operator imagesize: umożliwia określenie wielkości wyszukiwanego obrazu. Do dyspozycji pozostają trzy wielkości: mała (*small*), średnia (*medium*) oraz duża (*large*). Zapytanie formułuje się następująco: „daniel mider" imagesize:small. Wbrew deklaracjom liczby (na przykład 600 dpi) w Bing nie działają.

W Bing funkcjonował operator msite:, ogniskujący wyniki wyszukiwania na stronach multimedialnych (fotografie i filmy). Przykładowe zapytanie: msite:mider.

Większość analizowanych wyszukiwarek (B, G, Y oraz do pewnego stopnia Y!) zapewnia możliwość wyszukiwania w mapach za pomocą operatora maps:. Zwraca uwagę fakt, iż wynik wyszukiwania jest odmienny od tego, jaki uzyskujemy otwierając zakładkę Mapy wyszukiwarki. Yahoo! ma wbudowaną funkcję wyszukiwania map we frazach (teren USA), jeśli padanie słowo *map*, na przykład: map of New York.

Wyszukiwarka Google oferuje możliwość wyszukiwania treści publikacji wedle tytułu i autora. Operator book: pozwala na zaawansowane wyszukiwanie książek według słów kluczowych zawartych w tytule. Przeszukuje bazę Google Books[19]. Z kolei operator author: wyszukuje autorów tekstów (artykułów).

---

[18] *Lista miar podlegających konwersji*, http://searchcommands.com/convert/ (dostęp: 12.02.2019). Obecnie składnia wymaga jedynie wpisania np.: 1 inch to cm.

[19] Wyszukiwanie można przeprowadzić również: https://books.google.com/advanced_book_search (dostęp: 12.02.2019).

W Google, dla terenu Stanów Zjednoczonych, przez krótki czas funkcjonowała usługa wyszukiwania numerów telefonów (operatory phonebook:, bphonebook:, rphonebook:). Usługa ta została zlikwidowana ze względu na zbyt duże zainteresowanie.

Warto zwrócić uwagę na operatory informacyjne. Operator movie: serwuje repertuar kin, informacje o filmach i recenzje. Przykładowe użycie: movie:Warszawa. Z kolei operator weather: – analogicznie – podaje informacje dotyczące aury.

Przegląd wybranych narzędzi eksploracji Internetu – wyszukiwarek internetowych

Przegląd nie ma charakteru wyczerpującego ani pogłębionego, służy raczej wstępnej orientacji zainteresowanych w uniwersum wyszukiwarek internetowych. Wskazano i przeanalizowano między innymi wyszukiwarki globalne, wyszukiwarki zogniskowane na prywatności użytkownika, meta- i multiwyszukiwarki, wyszukiwarki i katalogi lokalne, wyszukiwarki ludzi, wyszukiwarki szarej literatury i wyszukiwarki naukowe, a także wyszukiwanie w archiwach Internetu.

W tekście zrezygnowano z licznych wątków tematycznych – zabrakło na przykład wyszukiwarek multimediów oraz wyszukiwarek mediów społecznościowych, a także wielu branżowych wyszukiwarek. Pominięte zostały również wyszukiwarki i paradygmaty wyszukiwania w innych niż powierzchniowych obszarach Internetu, między innymi Usenecie, File Transfer Protocol, The Onion Router, Invisible Internet Project (I2P), OpenNIC, CesidianRoot, Freenet. Uzasadnieniem takiego działania jest ograniczona objętość tekstu oraz fakt, iż treści te mogą zostać wprowadzone po opanowaniu przedstawianych w tekście informacji, a także – co istotniejsze – dotarcie do wymienionych wyżej zasobów Internetu wymaga więcej niż podstawowej wiedzy na temat jego topografii i solidnego instruktażu związanego zarówno z instalacją i konfiguracją, jak również użyciem narzędzi.

WYSZUKIWARKI GLOBALNE[20]

Do globalnych wyszukiwarek internetowych zaliczymy „wielką trójkę": Google, Bing, Yahoo!, jednak wyliczenie takie ma charakter umowny: Google wyraźnie dystansuje pozostałe wymienione narzędzia. Jest to witryna najczęściej odwiedzana na świecie – dziennie odnotowuje ponad miliard unikatowych użytkowników (1 022 345 310), którzy otwierają ją ponad osiem miliardów razy (8 178 762 480)[21]. Wyszukiwarka Google to aż 90,28% wszystkich wyszukiwań, podczas gdy pozostałe wyszukiwarki globalne stanowią margines: Bing – 3,82% i Yahoo! – 2,76%[22]. Wyniki te od 2010 roku zmieniły się niewiele – Google stale utrzymuje przewagę, jedyna zmiana dotyczy popularności Bing, która od 2010 roku wzrosła aż dwukrotnie[23]. W przypadku dostępu za pomocą urządzeń mobilnych przewaga ta sięga aż 94,15%[24]. W Polsce wyszukiwarka Google uzyskała miażdżącą przewagę – 98,73%, podczas gdy Yahoo! – 0,56%, a Bing – 0,44%[25]. Nie wszędzie jednak Google ma dominującą pozycję – jest niepopularne w Chinach, gdzie na przykład w grudniu 2018 roku odwoływano się do niej zaledwie w 2,57% wyszukiwań. Na pierwszym miejscu lokuje się tam Baidu (z wynikiem 70,3%), Shenma (15,62%), Sogou (4,74%), Haosou (4,54%). Druga z wymienionych wyszukiwarek stara się konkurować z wiodącą Baidu – w lutym 2018 roku udział Baidu spadł do 58,55%, a w tym samym czasie wzrosło istotnie zainteresowanie wyszukiwarką Shenma – do 28,76%. W Federacji Rosyjskiej również rodzima wyszukiwarka ma znaczną przewagę – Yandex w grudniu 2018 miał 54,27% udziału w rynku wyszukiwarek, a Google – 42,42%. Na trzecim miejscu ulokowała się Mail.

---

[20] Pojęcie „globalne" zostało użyte w znaczeniu komercyjnym, to jest pokrycia rynku wyszukiwarek. Z akademickiego punktu widzenia słuszne jest rozróżnienie Sabiny Cisek na wyszukiwarki globalne (horyzontalne, uniwersalne) oraz wyszukiwarki specjalistyczne (wertykalne): S. Cisek, *Wyszukiwarki specjalistyczne*, http://sabinacisek.blogspot.com/2012/11/wyszukiwarki-specjalistyczne.html (dostęp: 12.02.2019).

[21] *Web Analysis and Statistics*, https://web2stat.com/w/google.com (dostęp: 24.01.2019).

[22] Są to najnowsze (na czas powstania tekstu) dostępne dane, za październik 2018 r.: *Worldwide Desktop Market Share of Leading Search Engines from January 2010 to October 2018, Statista. The Statistics Portal*, Statista, https://www.statista.com/statistics/216573/worldwide-market-share-of-search-engines/ (dostęp: 24.01.2019).

[23] C. Mangles, *Search Engine Statistics 2018*, SmartInsights, 30.01.2018, http://www.smartinsights.com/search-engine-marketing/search-engine-statistics/ (dostęp: 24.01.2019).

[24] *Mobile Search Engine Market Share Worldwide*, StatCounter, http://gs.statcounter.com/search-engine-market-share/mobile/worldwide (dostęp: 24.01.2019).

[25] *Search Engine Market Share Poland*, StatCounter, http://gs.statcounter.com/search-engine-market-share/mobile/worldwide (dostęp: 24.01.2019).

ru (z wynikiem 2,24%)[26]. Przewaga Google jest – paradoksalnie – nieco mniejsza w Stanach Zjednoczonych Ameryki Północnej, Google ma tam 86,91% udziału, Yahoo! – 6,31%, Bing – 5,56%, a DuckDuckGo – 0,9%. Notowane są tam również: MSN – 0,16% i Baidu – 0,04%[27].

Kluczową informacją wydaje się nie częstość używania, lecz odpowiedź na pytanie o liczbę stron zindeksowanych przez każdą z wyszukiwarek. Poznanie, choć przybliżone, całkowitej liczby zindeksowanych przez wyszukiwarki stron możliwe jest dzięki stylistyce kwantytatywnej – dyscyplinie naukowej z pogranicza stylistyki i matematyki. Konkretnie, dokonanie pomiaru umożliwia prawidłowość odkryta przez Jeana-Baptiste'a Estoupa i George'a Kingsleya Zipfa (prawo Estoupa–Zipfa). Określenie wielkości zindeksowanej sieci opiera się na przybliżonych rachunkach wolumenu Google, Bing i Yahoo! Search, uwzględniając wzajemne nakładanie się wyników, co prowadzi do przeszacowania, odpowiednio pomniejszając uzyskane sumy[28]. Wielkość sumaryczną indeksu określa się następująco. Po pierwsze, konieczny jest zbiór referencyjny (wzorzec, zbiór odniesienia, korpus) dostępny offline (w praktyce jego zawartość stanowi milion stron internetowych z katalogu DMOZ, co może być – potencjalnie – uważane za reprezentatywną próbkę World Wide Web). Po wtóre, niezbędne jest sekwencyjne użycie głównych wyszukiwarek. Obliczenie wolumenu stron odbywa się na podstawie porównania proporcji słów wykazywanych przez korpus z wykazywanymi przez wyszukiwarki. Przykładowo: jeśli słowo x występuje w korpusie w 75% dokumentów, to jeśli zostało ono wykazane online w 15 mld dokumentów, wówczas orzekamy, iż istnieje 20 mld stron. W praktyce każdego dnia 50 reprezentatywnych słów (rozlokowanych równomiernie w logarytmicznych odstępach), których częstotliwość została obliczona w korpusie, jest wysyłanych do poddawanych pomiarowi wyszukiwarek. Liczba stron znalezionych dla tych słów jest rejestrowana i porównywana z ich względnymi częstotliwościami w korpusie[29]. Na tej podstawie oce-

[26] *Search Engine Market Share Russian Federation*, StatCounter, http://gs.statcounter.com/search-engine-market-share/all/russian-federation (dostęp: 24.01.2019).

[27] *Search Engine Market Share United States of America*, StatCounter, http://gs.statcounter.com/search-engine-market-share/all/united-states-of-america (dostęp: 24.01.2019).

[28] Rachunek nakładających się wyników obliczany jest w sekwencji, począwszy od jednej z czterech wyszukiwarek, a zatem możliwych jest kilka porządków, co prowadzi do różnych całkowitych sum oszacowań. Ważne jest również to, że taki algorytm obliczeń sprawia, że wolumen stron jest niedoszacowany.

[29] A. van den Bosch, T. Bogers, M. de Kunder, *Estimating Search Engine Index Size Variability: A 9-year Longitudinal Study*, http://www.dekunder.nl/Media/10.1007_s11192-016-

nia się, iż Google zindeksowało 62 mld stron (sekwencja Google–Bing) lub nieco ponad 6 mld stron (sekwencja Bing–Google)[30]. Obecnie nie podaje się wyników dla dwóch pozostałych, pierwotnie uczestniczących w pomiarze wyszukiwarek – Yahoo! Search oraz Ask, ponieważ ich właściciele zrezygnowali z informowania użytkowników o liczebności wyników wyszukiwania.

WYSZUKIWARKI ZOGNISKOWANE NA PRYWATNOŚCI UŻYTKOWNIKA

Wyszukiwarki zogniskowane na prywatności użytkownika powstały w odpowiedzi na nieobecność na rynku wyszukiwarek nieprofilujących, a więc takich, które nie dokonują rozpoznawania, analizowania i segmentacji użytkowników pod kątem płci, wieku, lokalizacji, preferencji politycznych czy zainteresowań w celu dostosowania treści reklamowych przez użytkownika (warto podkreślić, iż nawet tryb incognito w Google nie chroni użytkownika przed tak zdefiniowanym profilowaniem)[31]. Obok naruszenia prywatności, stanowiącej dobro samo w sobie, działania Google prowadzą do utrudnień w korzystaniu z wyszukiwarki, czyniąc jej użytkowanie ze względu na obecność treści reklamowych uciążliwe, bo nieprzejrzyste.

Spośród wyszukiwarek szanujących prywatność użytkownika na pierwszym miejscu należy bezwzględnie wymienić wyszukiwarkę DuckDuckGo (nazwa pochodzi od dziecięcej zabawy *Duck, duck, goose*) dostępną pod adresem https://duckduckgo.com lub z przekierowaniem https://duck.com. Narzędzie to powstało w 2008 roku i obecnie występuje w wersji zarówno desktopowej, jak i dla urządzeń mobilnych. DDG została zaimplementowana do licznych przeglądarek internetowych (przede wszystkim Firefox 33.1. oraz Tor Browser 6.0), a obecnie (dane za styczeń 2019) odnotowuje około 30 mln wyszukiwań w ciągu doby (194. pozycja w rankingu Alexa[32]).

DDG chroni prywatność użytkowników, nie zbierając o nich informacji w celu profilowania (w efekcie nie serwuje żadnych treści marketingowych), a także umożliwia korzystanie z narzędzi wyszukiwawczych przez

1863-z.pdf (dostęp: 24.01.2019).

30 *The Size of the World Wide Web (The Internet)*, WorldWideWebSize.com, http://www.worldwidewebsize.com (dostęp: 24.01.2019).

31 L. Vaas, *Google's Private Browsing Doesn't Keep Your Searches Anonymous*, 6.12.2018, https://nakedsecurity.sophos.com/2018/12/06/googles-private-browsing-doesnt-keep-your-searches-anonymous/ (dostęp: 15.02.2019).

32 *Duckduckgo.com Traffic Statistics*, http://www.alexa.com/siteinfo/duckduckgo.com (dostęp: 12.02.2019).

system anonimizujący TOR. *Votum separatum* jednego z użytkowników wskazywało na to, że jednak DDG dokonuje ograniczonego trackingu danych socjo-psycho-demograficznych[33], co DDG zdementowała, wyjaśniając sposób i zakres działania jednego z aktywowanych przez wyszukiwarkę elementów przeglądarki Firefox. Popularność DDG wzrosła istotnie po ujawnieniu przez Edwarda Snowdena szczegółów projektu PRISM amerykańskiej Agencji Bezpieczeństwa Narodowego (National Security Agency, NSA) oraz zaimplementowania DDG w produktach Apple. Polityka nieprofilowania uprawiana przez DDG nie tylko chroni użytkowników przez treściami reklamowymi i utratą prywatności, lecz przede wszystkim zapobiega najpoważniejszemu problemowi wyszukujących – bańce filtrującej (*filter bubble*) – a więc zjawisku często nieuświadamianej i niechcianej preselekcji informacji wyszukiwanych przez użytkownika pod kątem jego wcześniejszych wyszukiwań[34].

Indeksowane przez DDG dane pochodzą z licznych źródeł: ponad 400 źródeł indywidualnych (na przykład Search Boss, Wolfram Alpha), w tym crowdsourcingowych (na przykład Wikipedia), zoptymalizowanych wyników wyszukiwań pochodzących z Bing, Yahoo! oraz Yandex, a także działań własnego webcrawlera DuckDuckBot.

Drugą z wyszukiwarek szanujących prywatność jest Startpage.com (istniejąca od 1998 roku), dawniej IxQuick.com (do 2016 roku), reklamująca się jako „Najbardziej prywatna wyszukiwarka na świecie". Jej zasada działania jest prosta – jest ona pośrednikiem (*proxy*) między Google a użytkownikiem końcowym. W imieniu użytkownika składa zapytania serwerom Google, a wyniki, bez trackingu i gromadzenia danych osobowych, przekazuje użytkownikowi końcowemu. Warto zwrócić uwagę, iż wyszukiwarka ta nie notuje i nie przechowuje jakichkolwiek danych użytkownika, w tym jego adresu IP.

Kolejną wyszukiwarką przyjazną anonimowości użytkowników jest Searx (http://searx.me) – metawyszukiwarka, która korzysta z innych wyszukiwarek, nie udostępniając jakichkolwiek danych składającego zapytanie. Ponadto sama nie zapisuje żadnych danych wyszukującego – zapewnia użytkownikowi wyszukiwanie poprzez mechanizm HTTP

[33] D. Parrack, *DuckDuckGo Denies Using Browser Fingerprinting*, http://www.makeuseof.com/tag/duckduckgo-denies-browser-fingerprinting/ (dostęp: 12.02.2019).

[34] Więcej na ten temat: E. Pariser, *The Filter Bubble: How the New Personalized Web Is Changing What We Read and How We Think*, Londyn 2012, a poglądy kwestionujące istnienie bańki filtrującej odnajdziemy w: P. Boutin, *Your Results May Vary*, http://web.archive.org/web/20151214060050/http://www.wsj.com/articles/SB10001424052748703421204576327414266287254, strona obecnie dostępna *via* IWM.

POST, co uniemożliwia zapis w logach serwera. Każdy z wyszukanych odnośników jest bezpośredni (a nie jak w Google – podawany w postaci linku przekierowania i śledzony). Wyszukane strony można przeglądać jako zbuforowane, a więc nie pozostawiając swojego cyfrowego odcisku palca na ich serwerach.

Do wyszukiwarek szanujących prywatność użytkownika zaliczana jest również Qwant (http://www.qwant.com), której twórcy podkreślają (niezgodnie z prawdą), że jest jedyną z europejskich wyszukiwarek. Deklarują również, iż nie zapisuje ona historii wyszukiwania użytkownika oraz plików *cookies*.

METAWYSZUKIWARKI I MULTIWYSZUKIWARKI

Pojęcia metawyszukiwarki i multiwyszukiwarki są zazwyczaj utożsamiane[35]. Na ogół rozumie się je jako narzędzia (programy) działające jednocześnie na wielu serwisach wyszukiwawczych. Na potrzeby dydaktyczne można zaproponować, by pojęciem multiwyszukiwarki określać te narzędzia, które wyłącznie grupują wyniki wyszukiwania wielu wyszukiwarek, a pojęciem metawyszukiwarki takie narzędzia, które dodatkowo zawierają algorytm indeksowania i rangowania, a w efekcie prezentowania użytkownikowi odnalezionych zasobów. Takie rozróżnienie wydaje się mieć istotne znaczenie w ocenie i walidacji pozyskiwanych informacji[36]. Przesłanką korzystania z więcej niż jednej wyszukiwarki jest fakt, że żadna z istniejących nie indeksuje zasobów w identyczny sposób (istnieje kilka algorytmów indeksowania, na przykład binarny, PageRank, klikohit), w tym samym czasie (obieg robota sieciowego wyszukiwarki trwa kilka–kilkanaście dni) i w efekcie nie pokrywa całości zasobów.

Klasyczną multiwyszukiwarką w zdefiniowanym wyżej znaczeniu był Bjorgul (http://www.bjorgul.com) umożliwiający wyszukiwanie w każdej z parudziesięciu wyszukiwarek razem i osobno. Obecnie (tymczasowo

[35] Patrz na przykład: S. Cisek, *Warsztat infobrokera – poszukiwanie informacji*, http://www.academia.edu/32396257/Warsztat_infobrokera_-_poszukiwanie_informacji (dostęp: 12.02.2019).

[36] Jeszcze inne rozumienie pojęcia „metawyszukiwarka" proponuje Dominika Paleczna – uznaje ona, że metawyszukiwarki to takie systemy wyszukiwań, które odpytują systemy zdalne w czasie rzeczywistym, co w zwykłym użytkowaniu przekłada się na długi czas oczekiwania na wynik. Zalicza do nich m.in. Katalog Rozproszony Bibliotek Polskich (KaRo), Bazy Biblioteki Narodowej oraz Mazowiecki System Informacji Bibliotecznej (EHIS). Takie rozumienie nie koliduje jednak z przyjętym – rozważania autorki dotyczą specjalistycznych systemów bibliotecznych, nie zaś wyszukiwarek globalnych. D. Paleczna, *Systemy discovery vs. metawyszukiwarki*, http://nowetrendy.bibliosfera.net/2014/08.systemy_discovery.pdf (dostęp: 12.02.2019).

lub stale) nie funkcjonuje. Z kolei Etools (http://www.etools.ch/) pełni jednocześnie funkcję metawyszukiwarki i multiwyszukiwarki. Umożliwia jednocześnie wyszukiwanie zsyntetyzowane, jak również eksplorację wedle pojedynczych wyszukiwarek na jednej stronie. Wyszukiwarka ta konsumuje wyniki 17 innych, a algorytm syntetycznego wyszukiwania pozostaje nieujawniony. Podany ranking narzędzi daje jednak orientację o rangach nadawanych poszczególnym indeksom cząstkowym wyszukiwarek[37]. Wyszukiwarka Izito (http://www.izito.com) jest metawyszukiwarką, korzysta bowiem jednocześnie z Yahoo!, Bing, zasobów – jak Wikipedia, a także YouTube oraz Entireweb, a algorytmy selekcji i rangowania pozostają nieujawniane przez twórców. Również metaEureka (https://www.metaeureka.com/) może być uznana za metawyszukiwarkę, stanowiąc wyjątek pod tym względem, że jej twórcy opisują indeks punktowy przyznawany określonym zasobom[38]. Wartościowa wydaje się wyszukiwarka Dogpile, pomimo faktu, że algorytm, którym się posługuje, pozostaje tajemnicą. Jej twórcy wdrożyli ją na podstawie interesujących studiów odnoszących się do pokrycia obszaru wyszukiwań przez poszczególne wiodące wyszukiwarki. Okazuje się, że poszczególne wyniki wyszukiwań w różnych wyszukiwarkach pokrywają się w mniej niż jednym procencie[39]. Warto również wspomnieć o innych wyszukiwarkach funkcjonujących długo, lecz nieujawniających swoich algorytmów działania, na przykład Entireweb (http://www.entireweb.com/), Gigablast (http://www.gigablast.com/), Lycos (http://www.lycos.com/).

WYSZUKIWARKI OFERUJĄCE AGREGOWANIE TREŚCI (TEMATYCZNIE, CHRONOMETRYCZNIE, LOKALIZACYJNIE)

Wyniki wyszukiwania we wszystkich większych wyszukiwarkach podawane są w postaci jednolitego, ciągłego rankingu, którego porządek wyznaczany jest przez liczbę punktów przyznanych według ustalonych algorytmów. Wyszukiwarki takie jak Carrot2 oraz Yippy (http://yippy.com/search) umożliwiają pozyskiwanie treści w postaci zagregowanej.

Wyszukiwarka Yippy umożliwia wyszukiwanie według źródeł, to jest miejsca zmieszczenia informacji (wówczas treści prezentowane są na przykład w podziale na te zamieszczone w prasie, zamieszczone na stro-

[37] *How Does eTools.ch Work?*, http://www.etools.ch/searchInfo.do (dostęp: 12.02.2019).

[38] *How is this Working?*, https://www.metaeureka.com/help.shtml (dostęp: 12.02.2019).

[39] *Different Engines, Different Results Web Searchers Not Always Finding What They're Looking for Online A Research Study by Dogpile.com*, 2007, http://cdn1.inspsearchapi.com/dogpile/11.6.0.452/content/downloads/overlap-differentenginesdifferentresults.pdf (dostęp: 12.02.2019).

nach agencji informacyjnych, stacji telewizyjnych itd.), według domen (wówczas możemy spodziewać się jako wyników wyszukiwania katalogów z nazwami domen, jak .com, .org, .edu, .net), czasu (zakreślane przez użytkownika wyszukiwarki interwałowo), tematyki (algorytm ustalany oddolnie, na podstawie analizy treści stron internetowych). Z kolei Carrot2 oferuje przeszukiwanie w podziale na eksplorację stron internetowych z użyciem metawyszukiwarki eTools, przeszukiwanie Wikipedii, wyszukiwanie w PubMed (baza danych obejmująca artykuły z dziedziny medycyny i nauk biologicznych, zawiera ponad 26 mln rekordów, w tym publikacji pełnotekstowych w wolnym dostępie) oraz PUT (narzędzie Politechniki Poznańskiej wykorzystujące eTools). Wyniki wyszukiwania prezentowane są w katalogach tematycznych. Niekiedy element usługi świadczonej przez wyszukiwarkę stanowi również wizualizacja treści, jak w przypadku Carrot2.

WYSZUKIWARKI I KATALOGI LOKALNE

Do najbardziej udanych projektów wyszukiwarek lokalnych należą rosyjski (obecnie rosyjsko-holenderski) Yandex[40] oraz czeski Seznam. Wyszukiwarka Yandex wdrożona została w 1997 roku jako wyszukiwarka internetowa, obecnie znajduje się globalnie w pierwszej dziesiątce wyszukiwarek, a w Rosji kontroluje ponad połowę rynku wyszukiwarek. Oferuje – oprócz wyszukiwania – ponad 70 różnego rodzaju usług online, między innymi nawigacji (Yandex.Navigator), tłumaczenia (Yandex. Translate), zamawiania taksówek (Yandex.Taxi), poczty elektronicznej (YandexMail), dysku w chmurze (Yandex.Disk). Wyszukiwarka Seznam[41] według rankingu Alexa zajmuje w Czechach trzecie miejsce spośród najczęściej otwieranych stron, mając aż 62% udziału w rynku wyszukiwarek – dystansuje Google z zaledwie 29% udziału w rynku[42]. Katalogi, zarówno globalne (Jasmine Directory, https://www.jasminedirectory.com) jak i lokalne, są technologią schyłkową, wyparły je bowiem wyszukiwarki. Pierwotnie zasoby Internetu usiłowano katalogować, lecz wraz z lawinowym wzrostem tych zasobów taki sposób agregowania treści stał się nieopłacalny i niemożliwy (na przykład Yahoo!, a i wiele innych wyszukiwarek, pierwotnie powstawały jako katalogi). Aktualnie sens funkcjonowaniu katalogów nadaje pozycjonowanie stron (notabene wiele katalogów

[40] *Yandex*, http://www.yandex.ru, www.yandex.com (dostęp: 12.02.2019).

[41] *Seznam*, http://seznam.cz (dostęp: 12.02.2019).

[42] G. Marczak, *Czeska wyszukiwarka seznam warta miliard dolarów!*, Antyweb, 18.08.2008, http://antyweb.pl/czeska-wyszukiwakra-seznam-warta-miliard-dolarow/ (dostęp: 12.02.2019).

wprost nawiązuje do tego procederu, na przykład Katalog SEO, https://katalogseo.net.pl). Wyczerpujący przegląd katalogów lokalnych zawiera OpenKontakt (http://www.openkontakt.com/pl/wyszukiwarki).

WYSZUKIWARKI LUDZI

Wyszukiwarki ludzi, w szczególności w wolnym dostępie, nie są zbyt powszechne. Do najbardziej udanych wydaje się należeć wyszukiwarka Yasni wyświetlająca na jednej stronie wszystkie ogólnodostępne informacje oraz pogrupowane wyniki wyszukiwana dla wybranego nazwiska: teksty, fotografie, inne dane, artykuły w mediach, profile społecznościowe, wypowiedzi na forach. Po stworzeniu konta i zalogowaniu się można doprecyzować informacje na swój temat[43]. Kolejna wyszukiwarka – Pipl jest przedsięwzięciem komercyjnym, jednak dostarczającym nieodpłatnie usługi informacyjne w wersji podstawowej. Jej twórcy twierdzą, iż jest to największa na świecie wyszukiwarka osób łącząca publicznie dostępne informacje online i offline z wielu źródeł[44]. Do innych tego typu wyszukiwarek należą między innymi Snitch (http://snitch.name) oraz PeekYou (https://www.peekyou.com/). Należy podkreślić, że wyszukiwarki ludzi nie stanowią narzędzia pierwszego wyboru przy ich poszukiwaniu. Istnieją liczne, bardziej profesjonalne i skuteczne metody, zarówno zautomatyzowane (vide: Maltego, Oryon OSINT Browser), jak i ręczne (na przykład algorytmy przeszukiwania mediów społecznościowych)[45].

Od 2016 roku oferowane są w Internecie nieodpłatne usługi wyszukiwania imion i nazwisk na podstawie numeru telefonicznego. Dostępne są one pod adresami: http://www.truecaller.com oraz http://www.sync.com. Warunek skorzystania z usługi stanowi zgoda na pobranie całości swojej książki adresowej z konta Google. Książka teleadresowa Google stanowi źródło informacji o numerach telefonicznych. Ominięcie tej niedogodności jest banalne i polega na założeniu nowego konta, które nie zawiera tego typu informacji. Pierwsza z wymienionych baz zawiera jakoby ponad trzy miliardy numerów telefonicznych, druga zaś – prawie miliard[46].

---

[43] *Yasni*, http://www.yasni.com (dostęp: 12.02.2019).

[44] *Pipl*, https://pipl.com/ (dostęp: 12.02.2019).

[45] Warto zdecydowanie polecić zestawienia Marcusa P. Zillmana uzupełniane, uaktualniane i publikowane przezeń systematycznie: M.P. Zillman, *Finding People Resources and Sites 2019*, http://whitepapers.virtualprivatelibrary.net/Finding%20People.pdf (dostęp: 12.02.2019).

[46] A. Haertle, *Jak ustalić nazwisko posiadacza numeru telefonu – Twoje pewnie też*, Zaufana Trzecia Strona, 26.11.2016, https://zaufanatrzeciastrona.pl/post/jak-ustalic-nazwisko-posiadacza-numeru-telefonu-twoje-pewnie-tez/ (dostęp: 20.01.2019).

Istnieją liczne możliwości wyszukiwania użytkowników (za pomocą ich nazw własnych, to jest *nicknames*) w mediach społecznościowych, na przykład CheckUsernames (https://checkusernames.com/), UserSherlock (http://www.usersherlock.com/), KnowEm? (https://knowem.com/) oraz UserSearch (https://usersearch.org). Istotnym identyfikatorem użytkownika w sieci jest adres poczty elektronicznej. Powstały liczne usługi walidacji i permutacji adresów poczty elektronicznej. Funkcjonowanie serwisów walidujących (na przykład Bulk Validation – http://leopathu.com/verif-email, Hunter – http://hunter.io/email-verifier) polega na weryfikacji adresu poczty elektronicznej (prawidłowość formatu adresu, istnienie i odpowiadanie serwera w danej domenie, akceptacja danego adresu e-mail przez ten serwer). Działanie permutatorów polega na systematycznym generowaniu i sprawdzaniu potencjalnych adresów poczty elektronicznej po wprowadzeniu danych wejściowych (na przykład imienia, nazwiska, potencjalnego *nickname*, ewentualnie domeny itd.). Tego typu programy istnieją zarówno online (na przykład http://inteltechniques.com/OSINT/email.html), jak i w postaci oprogramowania służącego do tego celu (na przykład narzędzie Querytool w Oryon OSINT Browser).

Odrębną kategorię stanowią agregatory treści zamieszczanych w mediach społecznościowych – umożliwiające wyszukiwanie treści zamieszczanych przez użytkowników w przystępnej, przejrzystej formule. Do tego typu narzędzi należy StalkScan (https://stalkscan.com). Za jego pomocą można przeglądać zasoby każdego z użytkowników medium społecznościowego Facebook w podziale na znajomych, fotografie, zainteresowania i inne. Informacje te nie różnią się od tych możliwych do pozyskania przy bezpośrednim przeglądaniu profilu, jednak przewagą tego narzędzia jest wygoda jego użytkowania.

Warto wskazać wyszukiwarkę zawierającą konta, do których hasła wyciekły (są to głównie dane kont usług poczty elektronicznej, dostępu do serwisów społecznościowych oraz sklepów internetowych). Usługa została nosi nazwę HaveIBeenPwned? i udostępniono ją pod adresem https://haveibeenpwned.com. Jest to serwis prowadzony przez pracownika Microsoft Troya Hunta. Umożliwia ocenę bezpieczeństwa własnych kont na podstawie pokaźnej bazy zawierającej blisko 7 mld rekordów. CERT Polska oraz Orange Polska przygotowali polską wersję tej strony: https://www.cert.orange.pl/haveibeenpwned. Ze znacznie szerszego zakresu usług możemy skorzystać w serwisie WeLeakInfo (https://weleakinfo.com/) zawierającego blisko 9 mld rekordów. Wyszukiwać wycieki możemy według następujących kryteriów: nazwy użytkownika, adresu

poczty elektronicznej, hasła, adresu IP, a nawet nazwiska lub numeru telefonu. Serwis świadczy również odpłatną usługę udostępniania danych (na przykład dostęp do całodziennego wyszukiwania to koszt zaledwie 2 dolarów, natomiast uiszczenie 666 dolarów pozwala uzyskać dostęp dożywotni). Odpłatną usługę wyszukiwania świadczy również serwis Citadel (http://citadel.pw) funkcjonujący od 2017 roku i zawierający – według szczegółowej deklaracji – ponad 9,6 mld rekordów. Z kolei serwis LeakedSource (http://www.leakedsource.ru, dawniej znajdujący się w domenie .com) oferuje możliwość sprawdzenia, czy konkretny adres e-mail znajduje się w wyciekach informacji, ogniskując się na serwisach Tumblr i LinkedIn. Usługi dostępne są w wariancie darmowym i odpłatnym – rozszerzonym. Aktualne wycieki (nazwę strony/usługi, datę oraz zakres wycieku) można zweryfikować w Hacked Emails (https://hacked–emails.com/) dokumentującym wycieki 14 mld rekordów.

WYSZUKIWARKI OPARTE NA MAPACH I GEOLOKALIZACYJNE

Właściwości, możliwości i ograniczenia map takich jak Google (http://maps.google.com), Bing (https://www.bing.com/maps) czy OpenStreetMap (http://www.openstreetmap.org/) są raczej dobrze rozpoznawane. Warto omówić kilka mniej znanych.

Wikimapia (http://wikimapia.org/) to serwis umożliwiający oddolne, zbiorowe opisywanie obiektów geograficznych. Zawiera wiele nie tylko ciekawych, lecz również praktycznych informacji umożliwiających dobrą orientację w terenie. Informacje nie są standaryzowane, zróżnicowana jest ich liczba, jakość i stopień szczegółowości w poszczególnych miejscach mapy. Serwis ten zawiera liczne informacje niedostępne na innych mapach: w tym elementy mikrotopografii, nazw lokalnych – nawet mikrotoponimów, ciekawostki historyczne i współczesne.

Serwis Mapillary (http://www.mapillary.com/) jest usługą tworzoną oddolnie, przez użytkowników. Zawiera geotagowane i naniesione na mapę fotografie. Omija ograniczenia Google, które zamazuje między innymi numery rejestracyjne pojazdów czy twarze uwiecznione na fotografiach. Fotografie wyposażone w znaczniki czasowe często zamieszczane są w dłuższych seriach. Z tego powodu (dla najbardziej uczęszczanych miejsc) usługa ta wydaje się przydatna jako narzędzie śledcze odsłaniające, choć wybiórczo, *status quo ante*. Analogiczne informacje pozyskamy również dzięki OpenStreetCam (https://openstreetcam.org).

Przydatne mogą okazać się mapy dostarczające odwzorowań historycznych – na przykład Historic Aerials (http://historicaerials.com),

która zawiera mapy z lat 1957–2015, jednak nie mają one charakteru globalnego i wyczerpującego. Podobne możliwości wyszukiwania dostępne są w serwisach: Terra Server (http://terraserver.com) – od 1997 roku oraz Land Viewer (http://eos.com) – fotografie od 1982 roku.

Infobroker Michael Bazzell stworzył funkcjonalność agregującą różnorakie usługi związane z mapami na jednej stronie – multiwyszukiwarkę geolokalizacyjną[47]. Jest to wartościowe narzędzie – niemal pełny zbiór usług związanych z mapami umożliwiający bezpośrednie w nich wyszukiwanie. Znajdują się tam między innymi: komplety map Bing, Google i Yandex, a także liczne mapy satelitarne (między innymi Zoom Earth Satellite, Land Viewer Satellite, Descartes Satellite).

Wyszukiwarki geolokalizacyjne przeznaczone są do ekstrakcji metadanych znajdujących się w serwisach społecznościowych w celu wydobycia geotagowania zamieszczanych wpisów, fotografii, filmów i innych materiałów. Zazwyczaj usługi tego typu są odpłatne, wymagając comiesięcznej subskrypcji (na przykład https://app.echosec.net).

Wyszukiwarka GeoSocialFootprint umożliwia pozyskiwanie metadanych geolokalizacyjnych pochodzących z wpisów w mikroblogu Twitter (http://geosocialfootprint.com/). Użyteczne informacje na temat różnego rodzaju zdarzeń nadzwyczajnych wraz z ich lokalizacją i krótkim opisem można pozyskać z mapy GlobalIncidentMap (http://www.globalincidentmap.com). Analogiczne narzędzie stanowi Keitharm (https://keitharm.me/). Są to następujące zróżnicowane kategorie zdarzeń zarówno endo-, jak i egzogennych: pożary lasów, epidemie, aktywność grup przestępczych (gangów), incydenty graniczne, akty terrorystyczne, trzęsienia ziemi, incydenty na statkach powietrznych niestanowiące aktów terrorystycznych, incydenty w obszarze medycyny i żywności, handel ludźmi.

WYSZUKIWARKI „SZAREJ LITERATURY" (*GREY LITERATURE*) I WYSZUKIWARKI NAUKOWE

Szara literatura to pewna forma magazynowania wiedzy akademickiej lokująca się między literaturą białą, to jest co najmniej opublikowanymi, a najlepiej zrecenzowanymi wedle określonych standardów książkami, artykułami w czasopismach i publikacjach pokonferencyjnych, a tzw. czarną literaturą rozumianą jako idee, pomysły, myśli, czyli zmaterializowaną i upowszechnioną w stopniu nikłym. Do szarej literatury zalicza się preprinty, wydania elektroniczne, raporty techniczne, treść

[47] Została przezeń udostępniona pod adresem: https://inteltechniques.com/osint/maps.html (dostęp: 11.02.2019).

wykładów, zbiory danych, multimedia (nagrania dźwiękowe i wizualne, fotografie), a także niektóre zasoby Internetu, jak zasoby blogów i mikroblogów, forów, podcasty, wiki i inne media społecznościowe. Należy podkreślić, że drogi pozyskiwania szarej literatury są liczne – odnajdywać ją można w bazach tradycyjnych, w tym archiwach, bazach elektronicznych (na przykład w repozytoriach prac promocyjnych, zasobach bibliotecznych, zasobach przedsiębiorstw i organizacji), pozyskiwać z komunikacji osobistej lub zdalnej[48]. Zalety i wady posługiwania się takimi źródłami w pracy naukowej widoczne są natychmiast: zyskujemy skrócenie obiegu myśli naukowej (o miesiące, a niekiedy o lata, tyle bowiem może trwać cykl publikacyjny), z drugiej strony nakładany jest na badacza bardziej rygorystyczny obowiązek weryfikacji źródłoznawczej pozyskanego materiału. Dostrzegalne wydają się tendencje zmierzające ku centralizacji tych zasobów – w sensie dyscyplinowym, narodowym, a nawet globalnym. Rozmaite organizacje gromadzą i udostępniają na różnych warunkach drukowaną i cyfrową szarą literaturę, jednakże kosztochłonność odnalezienia i skatalogowania sprawia, iż nie są to zbiory obszerne. Zasoby te jednak wciąż pozostają względnie rozproszone, a ich zbiory wymagają podjęcia prób dalszej eksploracji i usystematyzowania. Rozpoznawalną globalną inicjatywą jest projekt OpenGrey[49] zawierający wyszukiwarkę (choć jeszcze niezbyt zasobną) szarej literatury (na razie głównie europejskiej i w języku angielskim) oraz założony w 1992 roku Grey Literature Network Service (GreyNet)[50], którego celem jest animacja dialogu, badań i komunikacji między osobami i organizacjami w dziedzinie szarej literatury. Spośród narodowych zbiorów pokaźnym zasobem szarej literatury dysponuje The British Library[51], która zainteresowała się tą formą źródeł tuż po II wojnie światowej[52]. Odnajdziemy również na przykład

[48] Szerzej na temat kategoryzacji biała–szara–czarna literatura oraz subtypów szarej: D. Giustini, *Finding the Hard to Finds: Searching for Grey Literature*, 2010, http://www.slideshare.net/giustinid/finding-the-hard-to-findssearching-for-grey-gray-literature-2010 (dostęp: 12.02.2019); *Document Types in Grey Literature*, http://www.greynet.org/greysourceindex/documenttypes.html (dostęp: 12.02.2019). Systematyczny i – wydaje się – wyczerpujący przegląd typów szarej literatury zawiera: *Vocabulary of the Types Of Grey Literature*, http://repozitar.techlib.cz/record/3/files/idr-3_1.pdf (dostęp: 12.02.2019).

[49] *OpenGrey*, http://www.opengrey.eu/ (dostęp: 12.02.2019).

[50] *GreyNet*, http://www.greynet.org/home.html (dostęp: 12.02.2019).

[51] Wyszukiwarka znajduje się tutaj: https://ondemand.bl.uk/onDemand/search/index (dostęp: 12.02.2019). Należy nadmienić, iż większość zbiorów udostępnia się odpłatnie.

[52] S. Tillett, E. Newbold, *Grey Literature at The British Library: Revealing a Hidden Resource*, „Interlending & Document Supply" 2006, nr 34.

zasoby włoskie[53] i holenderskie[54]. Wiele zasobów narodowych czy dyscyplinowych zostało zamieszczonych na liście GreyNetu[55] – najbardziej rozpoznawalnej organizacji zajmującej się szarą literaturą oraz zostało zindeksowane przez Wikipedię[56]. Ostatnia z wymienionych list zawiera również wyszukiwarki naukowe o różnych zasadach dostępu. Korzystającą z licznych źródeł darmową wyszukiwarką literatury akademickiej, głównie białej, jest Bielfeld Academic Search Engine (BASE)[57]. Spośród baz zasobów naukowych spełniających warunki multidyscyplinarności (różnych dyscyplin, w szczególności społecznych i humanistycznych) i wolnego (nieodpłatnego) dostępu wymienić można:

- WorldWideScience.org – obejmuje dzięki wielostronnemu partnerstwu krajowe naukowe bazy danych i portale w ponad 70 krajach świata[58];
- GoogleScholar – funkcjonuje już od 2004 roku i obejmuje blisko 400 mln dokumentów[59];
- Microsoft Academic – darmowa publiczna wyszukiwarka internetowa dla publikacji naukowych i literatury opracowana przez Microsoft Research, uruchomiona ponownie w 2016 roku i wykorzystująca technologie semantycznego wyszukiwania. Jak deklarują jej twórcy, posiada ponad 375 mln dokumentów, z czego 170 mln to dokumenty naukowe, grupuje ponad ćwierć miliona autorów, blisko 50 tys. czasopism i ponad 25 tys. instytucji[60];
- ScienceOpen – stanowi nie tylko repozytorium zasobów, lecz również platformę wspierającą publikację, recenzowanie i dyskusję nad zamieszczanymi tekstami. Posiada blisko 40 mln artykułów, a dla każdego z tekstów umożliwiono śledzenie metadanych dotyczących recenzji, zmian i cytowań[61];

---

53 *Consiglio Nazionale delle Ricerche*, http://polarcnr.area.ge.cnr.it/cataloghi/bice/index.php?type=Grigia (dostęp: 12.02.2019).

54 *Grijze Literaturr in Nederland (GLIN)*, http://www.publiekwijzer.nl/bestanden.php?id=zoeknaar&db=3.2 (dostęp: 12.02.2019).

55 *GreySource A Selection of Web-based Resources in Grey Literature*, http://www.greynet.org/greysourceindex.html (dostęp: 12.02.2019).

56 *List of Academic Databases and Search Engines*, https://en.wikipedia.org/wiki/List_of_academic_databases_and_search_engines (dostęp: 12.02.2019).

57 *BASE*, http://www.base-search.net (dostęp: 12.02.2019).

58 *WorldWideScience. The Global Science Gateway*, https://worldwidescience.org (dostęp: 12.02.2019).

59 *GoogleScholar*, https://scholar.google.pl (dostęp: 12.02.2019).

60 *Microsoft Academic*, http:// academic.microsoft.com (dostęp: 12.02.2019).

61 *ScienceOpen*, https://www.scienceopen.com/ (dostęp: 12.02.2019).

- OAIster – jest to katalog z wyszukiwarką zawierający blisko 50 mln rekordów zasobów otwartego dostępu[62];
- Paperity – to agregator wolnodostępowych czasopism i dokumentów, który w swoich zasobach posiada blisko 2 mln dokumentów i ponad 4 tys. zarejestrowanych czasopism[63];
- SSRN (*Social Science Research Network*) – to repozytorium zogniskowane na naukach społecznych, humanistycznych, a także – od niedawna – biologii, chemii, inżynierii, medycynie, informatyce i innych. Autorzy mogą swobodnie zamieszczać i hostować swoje teksty, użytkownicy zaś mogą subskrybować spersonalizowane e-maile podawane w postaci abstraktów wraz z odnośnikami. Do pewnego stopnia można uznać portal za źródło szarej literatury, ponieważ publikowane są tam często materiały przed drukiem i w celu przedyskutowania ze społecznością użytkowników[64];
- Jurn – stanowi bezpłatne narzędzie do wyszukiwania pełnotekstowych prac naukowych w zakresie nauk społecznych, humanistycznych i ścisłych, współpracuje z licznymi bibliotekami akademickimi i rządowymi, w tym Centralną Biblioteką Komisji Europejskiej, University of Cambridge, University of California i Princeton University Library[65].

ARCHIWA INTERNETU

Archiwa Internetu powszechnie utożsamiane są z usługą Internet Wayback Machine[66] znajdującą się pod adresem http://archive.org/web/web.php, a funkcjonującą od 1996 roku. Celem powstania tej organizacji była próba zapobieżenia bezpośredniemu ulotowi treści stron internetowych, które z czasem są modyfikowane lub zamykane. Dostęp do archiwum jest bezpłatny, a historia witryn sięga 1996 roku (indeksowane z różnymi częstotliwościami, w zależności od ich popularności). Strony indeksowane zapisywane są nie tylko w sposób automatyczny – IWBM współpracuje z ponad 450 bibliotekami i innymi instytucjami za pośrednictwem pro-

[62] *OAIster*, https://www.oclc.org/en/oaister.html (dostęp: 12.02.2019) oraz wyszukiwarka: https://oaister.worldcat.org (dostęp: 12.02.2019).

[63] *Paperity*, http://paperity.org/ (dostęp: 18.02.2019).

[64] SSRN, https://www.ssrn.com/index.cfm/en/ (dostęp: 12.02.2019).

[65] *Jurn*, http://www.jurn.org (dostęp: 12.02.2019).

[66] Na temat tego przedsięwzięcia powstały ponad trzy setki artykułów naukowych, głównie w dyscyplinach takich jak nauki społeczne, bibliotekoznawstwo, technologie informacyjne. S.K. Arora, Y. Li, J. Youtie, P. Shapira, *Using the Wayback Machine to Mine Websites in the Social Sciences: A Methodological Resource*, „Journal of the Association for Information Science and Technology" 2015, nr 67.

gramu Archive-It, by zweryfikować ważne strony internetowe. Jak podają twórcy, obecnie archiwum zawiera 330 mld stron internetowych, 20 mln książek[67] i tekstów, blisko 5 mln nagrań audio (w tym 180 tys. koncertów na żywo), 4 mln filmów (w tym ponad 1,5 mln programów telewizyjnych[68]), 3 mln obrazów oraz 200 tys. programów komputerowych[69]. Po założeniu bezpłatnego konta można przesyłać własne multimedia do archiwum internetowego. Aktualnie archiwum zajmuje ponad 30 petabajtów. Od 2004 roku istnieje również brytyjska inicjatywa stworzenia archiwum Internetu – UK Web Archive (UKWA7, http://www.webarchive.org.uk/ukwa/), wspierana przez takie instytucje, jak Bodleian Libraries, Oxford University, British Library, Cambridge University Libraries, National Library of Scotland, National Library of Wales, Trinity College, Dublin, Bodleian Libraries. Inicjatywa ograniczona jest do brytyjskich stron WWW i zakłada kopiowanie ich co najmniej raz do roku. Zbieranie stron odbywa się automatycznie. Ważne strony internetowe (za takie zostały uznane głównie strony informacyjne) indeksowane są częściej (nawet codziennie). Uzupełnieniem automatycznego indeksowania jest praca ekspertów systematycznie zbierających strony internetowe[70].

Znacznie skromniejszym i pełniącym inne funkcje archiwum Internetu jest CachedPages (http://www.cachedpages.com/). Umożliwia korzystanie z wersji stron zapisywanych przez serwery jako kopie zapasowe. Usługa ta jest szczególnie przydatna, gdy aktualnie znajdująca się w Internecie wersja strony internetowej została zmodyfikowana lub gdy strona została usunięta albo gdy serwer, na której się znajduje, jest przeciążony lub doznał awarii. Buforowania stron internetowych dokonują między innymi Google i Coral, także CachedPages korzysta z nich. Ponadto interfejs wyszukiwarki umożliwia wyszukiwanie również w Archive.org. Warto podkreślić różnice pomiędzy buforowaniem w Google, które przechowuje najnowszą kopię strony sprzed od 1 do

---

[67] Program digitalizacji książek został rozpoczęty przez IWBM w 2005 r., a obecnie skanowanych jest tysiąc książek dziennie w 28 lokalizacjach na całym świecie. Książki wydane przed 1923 r. są dostępne do pobrania, a te później wydane można wypożyczać za pośrednictwem witryny Open Library.

[68] Archiwizację programów telewizyjnych rozpoczęto pod koniec 2000 r., a w 2009 r. rozpoczęto tworzenie wybranych wiadomości telewizyjnych w USA, które można było wyszukać przez napisy w archiwum TV News. Usługa ta pozwala badaczom używać telewizji jako cytowalnego i udostępnianego odniesienia.

[69] *About the Internet Archive*, https://archive.org/about/ (dostęp: 12.02.2019).

[70] Więcej na ten temat: K. Gmerek, *Archiwa internetowe po obu stronach Atlantyku – Internet Archive, Wayback Machine oraz UK Web Archive*, „Biuletyn EBIB" 2012, nr 1(128).

15 dni, a Coral, gdzie przechowywane strony są starsze (rzadziej bowiem indeksowane).

Podsumowanie

Należy podkreślić, iż ujawnienie poszukiwanych materiałów w Internecie stanowi preludium – nieodzowna jest wielostronna ewaluacja znaleziska (w zależności od rangi danych i wymaganej precyzji). W pierwszej kolejności prowadzi się krytykę zewnętrzną źródła (niższą). Jest to badanie jego cech zewnętrznych z wyłączeniem treści tego źródła. Istotą tego etapu jest ustalenie autentyczności źródła, to jest wykrycie potencjalnego falsyfikatu. W tym celu należy ustalić czas, miejsce pochodzenia (w szerokim sensie, w tym instytucjonalne) oraz autorstwo źródła. Dopełnieniem tego procesu jest wewnętrzna krytyka (wyższa) zmierzająca do ustalenia stopnia wiarygodności autora źródła lub samego źródła. Aby osiągnąć ten cel, należy dokonać interpretacji źródła, co czyni się w następujących wymiarach: syntaktycznym (formy językowe, struktura tekstu), semantycznym (znaczenie tekstu, rozumienie go) oraz pragmatycznym (podmiotowy sens tekstu, a więc interesy i poglądy autora, wpływ środowiska oraz możliwości i ograniczeń autora na ostateczny kształt źródła).

Wyłącznie techniczna, mechaniczna znajomość operatorów i wyszukiwarek analizowanych w artykule wydaje się całkowicie niewystarczająca, jeśli ich stosowanie nie podlega przemyślanym i wdrażanym w systematyczny sposób regułom heurystycznym. Jest to postępowanie żmudne, czasochłonne i powtarzalne, zatem w drodze oddolnych inicjatyw powstały liczne rozwiązania – mniej i bardziej zaawansowane pod względem informatycznym, a służące do automatyzacji wyżej analizowanych zapytań. Istnieje szereg narzędzi specjalistycznych, których element stanowi również możliwość zautomatyzowanego i uproszczonego wyszukiwania.

Jednym z najprostszych narzędzi oferujących na wpół zautomatyzowane wspomaganie wyszukiwania w Google i Bing jest program dostępny wyłącznie online o nazwie Advangle (http://advangle.com). Pozwala na wybór operatorów podanych w formie listy (na przykład *Page text*, *Domain*, *Country*, *Language*) z przejrzystego menu, a następnie zaopatrzenie ich w operandy z użyciem wygodnego okna. Efekty pracy otrzymujemy w postaci sformułowanego zapytania (operatory i operandy zestawione wedle reguł formułowania zapytań), gotowego do użycia w wymienionych wyszukiwarkach poprzez przycisk umożliwiający przekierowanie.

Istnieje również szereg bardziej zaawansowanych narzędzi dedykowanych szeroko pojętemu wywiadowi – nie tylko pozyskujemy za ich pomocą informacje odnośnie do osób, grup czy organizacji, ale przede wszystkim umożliwiają one badanie innych zasobów i typów struktur, ogniskując się na aspektach informatycznych (służyć mogą do testów penetracyjnych stron walidujących cyberbezpieczeństwo określonych podmiotów). Zaliczymy do nich narzędzia takie jak: Oryon OSINT Browser, Maltego SpiderFoot oraz FOCA (*Fingerprinting Organizations with Collected Archives*)[71].

Oryon OSINT Browser (dawniej: Oryon C Portable, Oryon Environment) stanowi aktualnie najpotężniejsze i najlepsze nieodpłatne narzędzie do prowadzenia wywiadu jawnoźródłowego. W pakiecie znajduje się 18 grup narzędzi, ponad 70 specjalistycznych programów przydatnych do codziennej pracy specjalisty pozyskującego dane w Internecie (odzyskiwanie danych, informatyka śledcza, *metadata harvesting* i inne), więcej niż 600 linków do rozmaitych wyszukiwarek i innych narzędzi. Wyposażony jest między innymi w rozwiązania zapewniające anonimowość (anonimowe surfowanie po Internecie, anonimowe maile i komunikatory, generatory fałszywych tożsamości), narzędzia analizy domen, wyszukiwarki osób i firm (rozmaite kryteria, wyszukiwanie w mediach społecznościowych), wyszukiwarki w DeepWeb, DarkWeb oraz HistoricalWeb, programy do analizy i wyszukiwania fotografii i filmów oraz wyszukiwania i analizy metadanych, mechanizmy Fake News Detection (ciekawe, lecz jeszcze słabo funkcjonujące rozwiązania). Dysponuje także Query Tool – narzędziem wspomagającym wyszukiwanie autorstwa polskiego infobrokera Marcina Mellera. Jego działanie polega na agregacji i automatyzacji wyników wyszukiwań różnych wyszukiwarek. Zawiera moduł ochrony anonimowości oraz komfortu użytkownika (adblock, proxy, brak rejestracji zapytań oraz cookies, ssl). Działa w systemie operacyjnym Windows, Mac, ponadto można je uruchomić w systemach Linux, wykorzystując na przykład Wine (umożliwiający zaimplementowanie środowiska Windows w systemach Linux).

Maltego, produkt informatyków z Republiki Południowej Afryki, pozwala na agregację informacji zamieszczanych w Internecie, ich przystępną i atrakcyjną wizualizację, przydatną również do wygodnej pracy z informacją. Jest narzędziem odpłatnym, choć darmowa (ograniczona czasowo) wersja programu dostępna jest pod adresem: https://www.

[71] Zastosowanie i reguły działania narzędzia zostały przybliżone przez Wojciecha Mincewicza w niniejszym tomie, dlatego tu zrezygnowano z opisu tego programu.

paterva.com/web7/downloads.php. Stanowi potężne narzędzie służące do pozyskiwania informacji metodą tzw. białego wywiadu (OSINT), choć do pewnego stopnia jest również narzędziem autoinwigilacji – przetwarzanie zapytań odbywa się (poza niektórymi wysokopłatnymi ofertami) wyłącznie na serwerach producenta programu.

SpiderFoot (aktualnie w wersji 2.12) jest jawnoźródłowym narzędziem stworzonym przez Steve'a Micallefa. Został zaprojektowany tak, aby był łatwy w użyciu, szybki i rozszerzalny. Jest to narzędzie służące do zdalnego rekonesansu (również w opcji pasywnej – niezauważalnej dla rozpoznawanego podmiotu). Prowadzi automatycznie wyszukiwanie, wykorzystując ponad setkę publicznych źródeł danych w celu zebrania informacji, między innymi na temat adresów IP, nazw domen, adresów e-mail i innych elementów.

Skuteczne użycie wymienionych narzędzi jest pozornie łatwe (i tak jest faktycznie na płaszczyźnie ich obsługi). Jednakże bez opanowania klasycznego wyszukiwania z użyciem operatorów oraz pozyskania wiedzy z zakresu działania i struktury zasobów Internetu ich wykorzystanie będzie nie w pełni efektywne, a nawet całkowicie nieefektywne. Sztuka i nauka wyszukiwania, ewaluacji, analizy i przetwarzania informacji pozyskiwanej z Internetu wciąż jest doskonalona, niemniej efekty tych udoskonaleń nie są znane ani powszechnie używane, dlatego niniejszy tekst zawiera silny rys dydaktyczny, mający na celu również popularyzację technik i narzędzi wyszukiwawczych.

Streszczenie

Tekst składa się z dwóch części tematycznych. W pierwszej przeanalizowano techniki wyszukiwania, w sensie ogólnym, to jest heurystyki, oraz szczegółowym, czyli konkretne techniki należące do rodziny języków zapytań – operatory (w podziale na operatory logiczne, lokalizacyjne, wyszukiwania kanałów komunikacyjnych w mediach społecznościowych, chronometryczne, wyszukiwania w treści strony www oraz operatory wyszukiwania określonych typów treści). Ich zasadniczą funkcję stanowi doprecyzowanie zapytań dla wyszukiwarek. Druga część tekstu zawiera przegląd i analizę wybranych narzędzi eksploracji Internetu – wyszukiwarek internetowych (wyszukiwarek globalnych, wyszukiwarek zogniskowanych na prywatności użytkownika, metawyszukiwarek i multiwyszukiwarek, wyszukiwarek i katalogów lokalnych, wyszukiwarek ludzi, wyszukiwarek szarej literatury i wyszukiwarek naukowych, archiwów Internetu). Dokonany przegląd nie ma

charakteru wyczerpującego, jest autorski i służy raczej do wstępnej orientacji zainteresowanym w uniwersum wyszukiwarek internetowych.

Daniel Mider

**THE ART OF SEARCHINIG ON THE INTERNET.
REVIEW OF SELECTED TECHNIQUES AND TOOLS**

The text consists of two parts. The first analyzed the internet search techniques – in a general heuristic sense and detailed i.e. specific techniques belonging to the family of query languages – so called operators (logical operators, localization operators, operators for communication channels in social media, chronometric operators, search operators in the content of the www and search operators for specific types of content). Their main function is to clarify search queries. The second part of the text contains a review and analysis of selected internet exploration tools – search engines (global search engines, search engines focusing on user privacy, metasearch engines and multisearch engines, regional search engines and catalogues, people search engines, search engines of gray literature and internet archives). Preview is not exhaustive or deepened, it serves rather the initial orientation of those interested in the search engine universe.

KEY WORDS: *information society, open source intelligence, infobrokering, search engine hacking*

Załącznik

Zestawienie operatorów dla wyszukiwarek Bing, DuckDuckGo, Google, Yahoo! i Yandex

Typ operatora	Operator	Google	DuckDuckGo	Yahoo!	Yandex	Bing
Operatory logiczne (w tym G. Boole'a)	- Google, Yahoo!, Bing, DDG, Yandex NOT Bing, DDG	+	+	+	+	+
	+ Google, Yahoo!, Yandex, DDG & Bing && Bing AND Bing, Google	+/-	+	+	+	+
	| Bing, Yahoo! || Bing OR Google, Bing, DDG, Yahoo!	+	+	+	- (?)	+
	*	+	+	+	+	+
	.	+	-	-/+	-	-
	..	+/-	+	+	+	+
	!	-	+	-	+	-
	„"	+	-/+	-/+	+	-/+
	()	+	+	+	+	+
Operatory lokalizacyjne	\	-	+	-	-	-
	site:[url]	+	+	+	+	+
	url:[url]	+/-	+	+	+	+

Typ operatora	Operator	Google	DuckDuckGo	Yahoo!	Yandex	Bing
Operatory lokalizacyjne	inurl:[tekst] allinurl:[tekst] url:[tekst] dla Yandex	+	+	-	+/-	-
	location:[kod iso] dla Bing loc:[kod iso] dla Bing region:[kod regionu] dla DDG r:[kod regionu] dla DDG cat:[kod regionu] dla Yandex	-	+	-	-/+	+
	altloc:[iso code]	-	-	-	+	+
	domain:[url]	-	+	+	-	+
	ip:[adres IP]	-	+	+	-	+
	host:[url dokładnie]	-	-	-	+	-/+
	rhost:[odwrócony url + operator zastępowania ciągu]	-	-	-	+	-/+
Operatory kanałów komunikacyjnych w mediach społecznościowych	#	+	+/-	+	-	+/-
	@	+	+	+	-	+
	feed:[nazwa kanału]	-	-	-	-	+
	blogurl:	+	-	-	-	-
	hasfeed:[nazwa kanału]	-	-	-	+	+
Operatory chronometryczne	cache:[url]	+	-	-	-	+/-
	date:				+	
	daterange:[data juliańska]-[data juliańska] dla Google daterange:[data]..[data] dla Yandex	-/+	-	-	-/+	-

Typ operatora	Operator	Google	DuckDuckGo	Yahoo!	Yandex	Bing
Operatory wyszukiwania w treści strony	linkfromdomain:[url] dla Bing link:[url] dla Google, Yahoo!	–/+	–	–/+	–	–/+
	inanchor:[tekst]	+/–	–	–	+	+/–
	info:[url]	+	–	–	–	–
	intitle:[tekst] allintitle:[tekst] title:[tekst] dla Yandex	+	+	+	+	+
	meta:[tekst]	–	–	–	–	–/+
	intext:[tekst] dla DDG allintext:[tekst] dla Google inbody:[tekst] dla Bing	+	+	+	+	+
	keyword:[tekst] instreamset:[tekst]	–	–	–	–/+	+/–
	language:[kod języka] dla Bing lang:[kod języka] dla Yandex	–	–	+	+/–	+
	prefer:[tekst]	–	–	–	–	–/+
	literalmeta:[tekst]	–	–	–	–	+
	near:[liczba – maksimum] around([liczba – maksimum])	+/–	–	–	–	+/–
Operatory wyszukiwania określonych typów treści	~	+/–	–	–	–	–/+
	msite:[tekst]	–	–	–	–	–/+

Typ operatora	Operator	Google	DuckDuckGo	Yahoo!	Yandex	Bing
Operatory wyszukiwania określonych typów treści	filetype:[rozszerzenie pliku] mime:[rozszerzenie pliku] dla Yandex ext: dla Bing	+	+	+	+	+
	contains:[rozszerzenie pliku]	–	–	–	–/?	+
	imagesize:[słowa: small, medium lub large]	–	–	–	–	+
	related:[url]	+	–	–	–	–
	book:[tytuł]	+	–	–	–	–
	author:[nazwisko]	+	–	–	–	–
	maps:[lokalizacja]	+	–	–/+	+	+
	define: [tekst]	+	–	+	–	–
	movie:[tekst]	+/–	+/–	+/–	–	+/–
	weather:[tekst]	+/–	+/–	+/–	–	+/–
	phonebook:[tekst] bphonebook:[tekst] rphonebook:[tekst]	–/+	–	–	–	–
	convert:	+	–	–	–	–

Legenda:
+ Operator występuje i funkcjonuje w danej wyszukiwarce
– Operator nie występuje i nie funkcjonuje w danej wyszukiwarce
+/– Operator występuje i z pewnymi ograniczeniami i błędami (wyszukiwania) funkcjonuje w danej wyszukiwarce
–/+ Operator występuje, jednak ze znacznymi ograniczeniami i błędami (wyszukiwania) funkcjonuje w danej wyszukiwarce
? Brak jednoznacznego potwierdzenia prawidłowego funkcjonowania operatora
[] W nawiasach kwadratowych znajdują się sugestie odnośnie do formatu operandu

Źródło: opracowanie własne.

Bibliografia

Arora S.K., Li Y., Youtie J., Shapira P., *Using the Wayback Machine to Mine Websites in the Social Sciences: A Methodological Resource*, „Journal of the Association for Information Science and Technology" 2015, nr 67.

Baran M., Cichocka E., Maranowski P., Pander W., *Cybernauci – diagnoza wiedzy, umiejętności i kompetencji dzieci i młodzieży, rodziców i opiekunów oraz nauczycieli w zakresie bezpiecznego korzystania z internetu. Raport podsumowujący badanie ex-ante*, Warszawa 2016, http://cybernauci.edu.pl/wp-content/uploads/2016/06/Cybernauci-diagnoza-wiedzy-umiejętności-i-kompetencji.-Raport.pdf (dostęp: 24.01.2019).

Bazzell M., *Open Source Intelligence Techniques. Resources for Searching and Analyzing Online Information*, 6th ed., Charleston 2018.

Benfield J.A., Szlemko W.J., *Internet-based Data Collection: Promises and Realities*, „Journal of Research Practice" 2006, nr 2(2).

Bosch A. van den, Bogers T., Kunder M. de, *Estimating Search Engine Index Size Variability: A 9-year Longitudinal Study*, http://www.dekunder.nl/Media/10.1007_s11192-016-1863-z.pdf (dostęp: 24.01.2019).

Boutin P., *Your Results May Vary*, http://web.archive.org/web/20151214060050/http://www.wsj.com/articles/SB10001424052748703421204576327414266287254, strona obecnie niedostępna.

Cisek S., *Warsztat infobrokera – poszukiwanie informacji*, http://www.academia.edu/32396257/Warsztat_infobrokera_-_poszukiwanie_informacji (dostęp: 12.02.2019).

Cisek S., *Wyszukiwarki specjalistyczne*, http://sabinacisek.blogspot.com/2012/11/wyszukiwarki-specjalistyczne.html (dostęp: 12.02.2019).

Giglietto F., Rossi L., Bennato D., *The Open Laboratory: Limits and Possibilities of Using Facebook, Twitter, and YouTube as a Research Data Source*, „Journal of Technology in Human Services" 2012, nr 30(3-4).

Giustini D., *Finding the Hard to Finds: Searching for Grey Literature*, 2010, http://www.slideshare.net/giustinid/finding-the-hard-to-findssearching-for-grey-gray-literature-2010 (dostęp: 12.02.2019).

Gmerek K., *Archiwa internetowe po obu stronach Atlantyku – Internet Archive, Wayback Machine oraz UK Web Archive*, „Biuletyn EBIB" 2012, nr 1(128).

Mangles C., *Search Engine Statistics 2018*, SmartInsights, 30.01.2018, http://www.smartinsights.com/search-engine-marketing/search-engine-statistics/ (dostęp: 24.01.2019).

Marczak G., *Czeska wyszukiwarka seznam warta miliard dolarów!*, Antyweb, 18.08.2008, http://antyweb.pl/czeska-wyszukiwakra-seznam-warta-miliard-dolarow/ (dostęp: 12.02.2019).

Mider D., *Mappa Mundi ukrytego Internetu. Próba kategoryzacji kanałów komunikacji i treści*, „EduAkcja. Magazyn Edukacji Elektronicznej" 2015, nr 2(10).

Ohiagu O.P., *The Internet: The Medium of the Mass Media*, „Kiabara Journal of Humanities" 2011, nr 16(2).

Paleczna D., *Systemy discovery vs. metawyszukiwarki*, http://nowetrendy.bibliosfera.net/2014/08.systemy_discovery.pdf (dostęp: 12.02.2019).

Pariser E., *The Filter Bubble: How the New Personalized Web Is Changing What We Read and How We Think*, Londyn 2012.

Parrack D., *DuckDuckGo Denies Using Browser Fingerprinting*, http://www.makeuseof.com/tag/duckduckgo-denies-browser-fingerprinting/ (dostęp: 12.02.2019).

Shenk D., *Data Smog. Survuving the Information Glut*, Nowy Jork 1998.

Tillett S., Newbold E., *Grey literature at The British Library: revealing a hidden resource,* „Interlending & Document Supply" 2006, nr 34.

Vaas L., *Google's Private Browsing Doesn't Keep your Searches Anonymous*, https://nakedsecurity.sophos.com/2018/12/06/googles-private-browsing-doesnt-keep-your-searches-anonymous/ (dostęp: 15.02.2019).

Zillman M.P., *Finding People Resources and Sites 2019*, http://whitepapers.virtualprivatelibrary.net/Finding%20People.pdf (dostęp: 12.02.2019).

VOL. 54 STUDIA POLITOLOGICZNE
DOI: 10.33896/SPolit.2019.54.9

Wojciech Mincewicz
ORCID: 0000-0003-0460-9158

Metadane – cichy zabójca prywatności

Słowa kluczowe:
metadane, anonimowość, bezpieczeństwo informacji, FOCA, nagłówek poczty elektronicznej

STUDIA I ANALIZY

Wprowadzenie

W ciągu minuty na świecie wysłanych zostaje 156 mln e-maili[1], które zawierają tysiące terabajtów danych: zdjęć, dokumentów tekstowych i innych plików cyfrowych. Tylko promil użytkowników Internetu wie, że udostępniając plik tekstowy czy też zdjęcie dostarcza swojemu interlokutorowi informację na przykład o: dacie utworzenia pliku, nazwie komputera, imionach i nazwiskach twórców, informacje o współautorach, a w przypadku zdjęć na przykład informację o typie urządzenia rejestrującego obraz czy o miejscu, gdzie został uwieczniony obraz. Wszystkie te dane nazywane są metadanymi.

Pojęcie metadanych było tradycyjnie używane w katalogach bibliotecznych do lat 80. XX wieku, gdy biblioteki przekształciły swoje dane katalogowe w cyfrowe bazy danych. Metadane definiowane są w sposób opisowy jako dane o danych[2]. Są to więc dane z przedrostkiem meta,

[1] Za. B. Marr, *Every Day? The Mind-Blowing Stats Everyone Should Read*, Forbes, 21.05.2018, https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/#4ffe1f6360ba (dostęp: 11.01.2019).

[2] Definicje taką odnajdujemy np. w: *Słownik encyklopedyczny informacji, języków i systemów informacyjno-naukowych*, B. Bojar (opr.), Warszawa 2002, s. 153; lub w: M.L. Zeng,

który swój źródłosłów bierze z języka greckiego i może oznaczać następstwo lub zmienność: poza, po, pod, łącznie, wśród, według, prze-[3]. Jako jedni z pierwszych temat danych o danych poruszali David Griffel i Stuart McIntosh, którzy w 1967 roku mówili o istnieniu tematycznych opisów danych oraz metajęzyku opisującym relacje danych[4]. W literaturze przedmiotu metadane definiowane są jako informacje strukturalne, które opisują, wyjaśniają, lokalizują lub w inny sposób ułatwiają pobieranie i wykorzystanie informacji[5], dane połączone z opisywanymi obiektami, odciążające ich potencjalnych użytkowników od konieczności posiadania pełnej apriorycznej wiedzy o ich istnieniu lub charakterystykach[6]. Oryginalne spojrzenie na metadane przedstawia na swoim blogu Lorcan Dempsey, który pisze, że lubi myśleć o metadanych jako danych, które usuwają z użytkownika potrzebę posiadania pełnej wiedzy na temat istnienia lub cech potencjalnie interesujących środowisko[7]. Inna definicja dostępna w literaturze amerykańskiej podpowiada, że metadane oznaczają uporządkowane informacje o zasobie informacyjnym w dowolnym formacie[8].

W literaturze polskiej kwestia metadanych zaczęła się pojawiać pod koniec XX wieku. Za prekursora badań uznawany jest Zdzisław Płoski, który metadane określa jako dane służące do opisu innych danych[9]. Inny z uznanych autorów Marek Nahotko rozszerza tę definicję i metadane określa jako: ustrukturyzowane, czytelne maszynowo dane, zawierające charakterystykę cyfrowych obiektów informacyjnych, które służą ich efek-

---

Metadata Basics, http://marciazeng.slis.kent.edu/metadatabasics/cover.htm (dostęp: 12.12.2018).

3 Za. *Słownik języka polskiego PWN*, M. Szymczak (red.), Warszawa 1995, s. 134. Zobacz także np.: *Słownik języka polskiego*, [b.a], Warszawa 1962, s. 577.

4 Zob. D. Griffel, S. McIntosh, *ADMINS – A Progres Raport*, MIT 1967, s. 27 i n., https://dspace.mit.edu/bitstream/handle/1721.1/82974/09487802.pdf (dostęp: 12.09.2018).

5 National Information Standards Organization, *Understanding Metadata*, Bethesda 2004, https://web.archive.org/web/20141107022958/http://www.niso.org/publications/press/UnderstandingMetadata.pdf [dostęp: 12.10.2018]. (Na temat NISO zob. przypis 20).

6 D. Campbell, *Dublin Core Metadata and the Australian Metaweb Project*, 10th National Library Technicians' Conference, Fremantle, 8–10.09.1999, https://www.nla.gov.au/bla/staffpaper/dcampvell1.html (dostęp: 15.10.2018).

7 L. Dempsey, *Registries: The Intelligence in Network*, 20.08.2006, http://orweblog.oclc.org/Registries-the-intelligence-in-the-network/ (dostęp: 12.11.2018).

8 M. Foulonneau, J. Riley, *Metadata for Digital Resources. Implementation, Systems Design and Interoperability*, Oxford 2008, s. 3, za. P. Caplan, *Metadata Fundamentals for All Librarians*, Chicago 2003, s. 3.

9 Zob. Z. Płoski, *Informatyka: słownik encyklopedyczny*, Wrocław 1999, s. 56.

tywnemu oraz trafnemu wyszukiwaniu, szczególnie w wielkich zasobach informacji w Internecie[10]. W innych definicjach autorzy kładą z kolei nacisk na funkcję metadanych i twierdzą, że za ich pomocą opisywane są dokumenty cyfrowe, w szczególności dokumenty dostępne poprzez sieci komputerowe. Wśród przykładowych danych wymienia: pliki, katalogi, datę utworzenia czy informacje o modyfikacji[11]. Nie jest to jednak zbiór zamknięty w zakresie funkcji, jakie spełniają metadane, albowiem wydaje się, że przede wszystkim służą one opisowi dokumentu, ale także ułatwiają wyszukiwanie danych, ukazują informacje o historii danych, ułatwiają wyszukiwanie określonych dokumentów, a także mogą ułatwić ocenę przydatności konkretnych danych.

W socjolekcie prawniczym za metadane uważane są logicznie powiązane z dokumentem elektronicznym usystematyzowane informacje opisujące ten dokument, które ułatwiają jego wyszukiwanie, kontrolę, zrozumienie i długotrwałe przechowywanie oraz zarządzanie[12].

Przytoczone rozumienia metadanych sprawiają, że dla dalszych rozważań niezbędne jest zaproponowanie definicji regulującej, która uściśli znaczenie kluczowego dla całego artykułu pojęcia i będzie definicją operacyjną dla studium. W tym celu posłużono się ustaleniami zawartymi w dokumencie *Standardy metadanych e-PL*, w którym przyjęto, że metadane to „wszelkie dane o dokumentach lub zbiorach dokumentów odnoszące się do ich treści, parametrów technicznych i fizycznych. Metadane mogą się odnosić do wszelkich dokumentów bez względu na sposób ich wytworzenia lub zapisu, w tym także do dokumentów elektronicznych. Metadane mogą określać istotne elementy takie jak tematykę dokumentów lub zbiorów dokumentów, osoby lub instytucji odpowiedzialnych za powstanie, czas wytworzenia, sposób zapisu, zasady dostępu itd."[13].

---

[10] M. Nahotko, *Metadane – sposób na uporządkowanie Internetu*, Kraków 2004, s. 15. Bliźniaczą definicję odnajdujemy w naukach bibliotekoznawczych, gdzie metadane definiowane są jako ustrukturyzowane, zazwyczaj czytelne maszynowo dane, zawierające charakterystykę dokumentów służących ich efektywnemu oraz trafnemu wyszukiwaniu, zarządzaniu nimi i ich wartościowaniu. Zob. G. Czapnik, Z. Gruszka (red.), *Podręczny słownik bibliotekarza*, Warszawa 2011, s. 197.

[11] Zob. M. Niebrzydowska, R. Kotowicz, *Wstęp do informatyki śledczej*, „Przegląd Bezpieczeństwa Wewnętrznego" 2012, nr 6, s. 66.

[12] Zob. Rozporządzenie Ministra Spraw Wewnętrznych i Administracji z 30 października 2006 r. w sprawie niezbędnych elementów struktury dokumentów elektronicznych (Dz.U. Nr 206, poz. 1517).

[13] Naczelna Dyrekcja Archiwów Państwowych, *Standardy metadanych, e-PL, wersja 0.1*, Warszawa 2005, s. 7, https://www.archiwa.gov.pl/images/docs/e-PL-0.1-2.pdf (dostęp: 15.10.2018).

Definicja ta, chociaż jest wystarczająca dla zrozumienia problematyki, wymaga jednak doprecyzowania, albowiem w ocenie autora metadane należy rozumieć przede wszystkim jako ustrukturyzowane dane, zapisane w logiczny sposób, które w jakikolwiek sposób identyfikują i charakteryzują dokument elektroniczny. Jest to tzw. szerokie rozumienie problematyki, mające charakter prymarny wobec pozostałych przymiotów, które zależne będą od tego, z jakim typem metadanych mamy do czynienia. Wówczas mowa jest o zawężaniu definicji do konkretnej kategorii.

Typologie metadanych

W literaturze przedmiotu występuje co najmniej kilka propozycji klasyfikacji i systematyzacji metadanych. Aby usystematyzować pojęcie, należy rozróżnić co najmniej trzy podejścia do zagadnienia. Pierwsze jest charakterystyczne dla bibliotekoznawstwa i opiera się na wykorzystaniu metadanych do tworzenia opisów bibliograficznych materiałów bibliotecznych oraz ich wyszukiwaniu. Drugie podejście ma charakter bardziej informacyjny i pozwala na wykorzystanie metadanych dla zarządzania danymi, kładąc nacisk na ich funkcje administracyjne i zarządzające zasobami[14]. W tym przypadku podstawowym zadaniem metadanych jest dostarczenie uporządkowanej, logicznie spójnej dokumentacji, która opisuje konstrukcję, powstanie i sposób wykorzystania danych w określonym systemie komputerowym. Trzecie podejście, infobrokerskie, znajduje zastosowanie w codziennej pracy brokera informacji czy informatyce śledczej. Należy je sytuować pomiędzy podejściem bibliotekoznawczym i informacyjnym. W tym przypadku metadane zarówno będą służyły do tworzenia opisów plików, jak i znajdą zastosowanie w czynnościach operacyjnych, albowiem dane zdobyte w trakcie działań mogą stać się dowodem cyfrowym. Są to dane ukryte dla zwykłego użytkownika, zawarte w tworzonych w plikach, dokumentach elektronicznych. Mogą zostać pozyskane przy użyciu technik charakterystycznych dla białego wywiadu[15]. W takich cyfrowych „śmieciach" pozostawionych często nieświadomie przez użytkownika infobroker może odnaleźć interesujące go informacje, takie jak na przykład:

[14] Por. M. Nahotko, *Metadane…*, s. 15.

[15] Więcej na temat białego, szarego i czarnego wywiadu czytaj: J. Garlicki, D. Mider, W. Mincewicz, *Pozyskiwanie informacji w Internecie metodą Google Hacking – biały, szary czy czarny wywiad?*, „Przegląd Bezpieczeństwa Wewnętrznego" 2019, nr 20, s. 68-91.

- datę i czas utworzenia pliku;
- adres lub położenie geograficzne miejsca utworzenia pliku;
- imię i nazwisko, nazwę firmy, nazwę komputera lub adres IP;
- nazwy wszystkich współtwórców dokumentu lub dodane komentarze;
- typ użytego aparatu i jego ustawienia podczas robienia zdjęcia;
- typ użytego urządzenia rejestrującego audio lub wideo i jego ustawienia podczas nagrywania;
- informacje o marce lub modelu smartfona[16].

Próba typologizacji problematyki metadanych również nie jest zadaniem łatwym. Na przykład Francis Patton Bretherton i Paul T. Singley rozróżniają dwie odrębne klasy metadanych: strukturalne i przewodnie. Metadane strukturalne opisują strukturę obiektów baz danych, takich jak tabele, kolumny i indeksy. Metadane przewodnie pomagają znaleźć określone przedmioty i są zwykle wyrażane jako zestaw słów kluczowych w języku naturalnym[17]. Rozszerzenie i ustrukturyzowanie tej typologizacji stanowi podejście różnicowania poprzez kryterium „typu danych". Na tej podstawie można wyróżnić metadane dla danych pierwotnych i metadane przetwarzania danych. Dane pierwotne to wszystkie te, które zarządzane są przez źródła danych, zasoby danych, giełdy danych oraz aplikacje. Odpowiednie metadane zawierają informacje związane ze strukturą źródła danych i zasobów danych. Metadane przetwarzania to informacje związane z przetwarzaniem danych: dotyczące procesów ładowania i modyfikacji danych, ich analizy oraz administrowania[18]. Podziału metadanych można także dokonać z uwzględnieniem kryterium czasu, w którym zostały wytworzone. W tym przypadku możemy mówić o następujących danych o danych:

- zbieranych w czasie projektu;

---

[16] Należy zdawać sobie sprawę, że metadane są zjawiskiem multidyscyplinarnym. Z uwagi na wymogi redakcyjne w artykule zaprezentowane zostały i zdefiniowane trzy podejścia do problematyki metadanych. W literaturze naukowej definiowane są one i opisywane m.in. w ekonomii, administracji czy też geolokalizacji. Zob. więcej odnośnie do np. administracji: M. Ganczar, *Informatyzacja administracji publicznej*, Warszawa 2009, s. 72–73; do ekonomii: S. Wrycza, *Informacja ekonomiczna. Podręcznik akademicki*, Warszawa 2010, s. 412 lub J. Zawiła-Niedźwiecki, K. Rostek, A. Gąsiorkiewicz (red.), *Informacja gospodarcza 4*, Warszawa 2010, s. 289. Metadane znajdują także zastosowanie choćby w geolokalizacji – zob. L. Litwin, M. Rossa, *Metadane geoinformacyjne w INSPIRE i SDI*, Gliwice 2010.

[17] Zob. F.P. Bretherton, P.T. Singley, *A User's View*, [w:] *Proceedings of the 7th International Working Conference on Scientific and Statistical Database Management*, Charlottesville 1994, s. 166–174.

[18] M. Nahotko, *Metadane…*, s. 23.

- tworzonych w czasie powstawania systemu;
- tworzonych podczas uruchamiania systemu;
- statycznych;
- dynamicznych;
- krótkoterminowych;
- długoterminowych[19].

Inne podejście do klasyfikacji metadanych odnajdujemy w zestawieniu przygotowanym przez National Information Standards Organization (NISO)[20]. Wyróżniono tam trzy typy:

- metadane opisowe – służą do odnajdywanie i identyfikacji kluczowych informacji, które umożliwiają lokalizację obiektu. Za przykład mogą służyć informacje o autorze, słowach kluczowych, wydawnictwie czy tytule;
- metadane strukturalne – opisują strukturę danego obiektu. W przypadku zbiorów bibliotecznych, które stanowiły *ad initium* dla zestawienia NISO, była to na przykład liczba rozdziałów, stron;
- metadane administracyjne – co odnosi się do informacji technicznych, w których zawarte są na przykład informacje o czasie i sposobie utworzenia pliku[21].

Typologia proponowana w niniejszym artykule została przyjęta w sposób aprioryczny i stanowi punkt odniesienia do dalszych rozważań, w których dokonano jej egzemplifikacji na przykładzie dwóch najpopularniejszych formatów plików: 1. JPEG – Joint Photographic Experts Group[22], charakterystycznego dla obrazów graficznych; 2. rozszerzenia

---

[19] Typologia za: *Podstawowe funkcje metadanych*, https://www.heuristic.pl/blog/internet/Co-to-sa-metadane;129.html (dostęp: 11.11.2018).

[20] National Information Standards Organization to amerykańskie stowarzyszenie non-profit, które opracowuje normy w zakresie usług informacyjnych. NISO jest akredytowane przy amerykańskiej organizacji normalizacyjnej – American National Standards Institute. NISO opracowuje normy, raporty techniczne, zalecenia i inne dokumenty dla tradycyjnych procesów informacyjnych i nowych technologii, czyli w zakresie wszystkiego, co jest w jakikolwiek sposób związane z wyszukiwaniem, przetwarzaniem i ochroną informacji oraz metadanymi. Zostało założone w 1939 r. Informacja za oficjalną witryną: https://groups.niso.org/home (dostęp: 11.11.2018).

[21] National Information Standards Organization, *Understanding…*; inne źródło, w którym przytoczono typologie: M.L. Zeng, *Metadata…*

[22] Najpopularniejszym formatem zapisu metadanych plików graficznych są tagi Exchangeable Image File Format (EXIF). W przypadku aparatu cyfrowego podstawowymi metadanymi przechowywanymi na karcie pamięci są daty utworzenia pliku, ostatniej modyfikacji oraz ostatniego użycia. Sam aparat cyfrowy w momencie wykonania zdjęcia zapisuje w EXIF nazwę producenta urządzenia, model czy datę utworzenia pliku

.doc lub .docx, charakterystycznego dla plików tekstowych[23]. Ekstrahowanie metadanych z plików i graficznych, i tekstowych możliwe jest za pomocą darmowych narzędzi dostępnych w przeglądarce Google, jak również ręcznie na komputerach osobistych.

Pracy nad artykułem autorowi przyświecały dwa cele. Pierwszy związany jest z uświadomieniem czytelnikowi potencjalnych zagrożeń i możliwości związanych z wykorzystaniem metadanych. Po drugie, chodziło o zawarcie takich elementów praktycznych, które pozwolą na samodzielne wykorzystanie technik defensywnych umożliwiających usunięcie metadanych z plików udostępnianych publicznie. Artykuł ma charakter autorskiego wyboru i opisuje najważniejsze zagadnienia, które umożliwiają zrozumienie problematyki opracowania.

Typologizacja metadanych w artykule odbywać się będzie według gradacyjnego ujęcia poszczególnych typów metadanych. Najpierw scharakteryzowane zostaną metadane opisowe, które są najszerszą kategorią identyfikującą autora pliku, następnie metadane strukturalne, które wprost opisują obiekt cyfrowy, a na końcu metadane administracyjne – zawierające dane techniczne pliku.

Metadane opisowe

Pierwszy typ metadanych wyróżnionych w toku powyższych rozważań to metadane opisowe (ang. *descriptive metadata*), które stanowią podstawę każdego dokumentu cyfrowego. Obejmują całą wiedzę, która gromadzona jest wokół obiektu. Mogą zawierać podstawowe informacje ewidencyjne, ale też szczegóły dotyczące historii czy przeznaczenia obiektu. Metadane

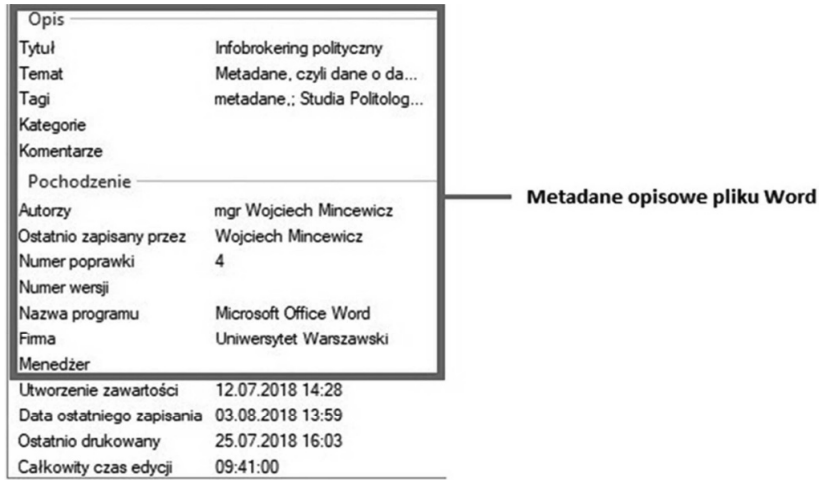
dla czasu ustawionego w aparacie cyfrowym. W metadanych pliku graficznego zlokalizowana jest także np. informacja o ustawieniach aparatu czy też pozycji GPS. Inne formaty zapisu metadanych plików graficznych to IPCT czy XMP. Zob. *Zanim wgrasz wakacyjne zdjęcie do sieci…*, Niebezpiecznik.pl, 2.07.2015, https://niebezpiecznik.pl/post/zanim-wgrasz-wakacyjne-zdjecia-do-sieci/ 2 lipca 2015 (dostęp: 12.11.2018).

[23] W przypadku dokumentów tekstowych plik zawiera – oprócz samego tekstu wytworzonego przez użytkowania – dodatkowe różnorodne niejawne informacje dotyczące tego pliku. System operacyjny Windows przekazuje do metadanych informacje np. o datach: utworzenia, ostatniej modyfikacji i użycia pliku. Program MS Word wytwarza informacje np. o autorze, osobach, które edytowały plik, dacie ostatniego zapisania, ostatniego drukowania czy czasie edycji. Za: P. Wichań, *Informatyka śledcza*, post FB z 16.11.2015, https://www.facebook.com/iswichran/posts/metadane-w-informatyce-czyli-dane-niejawne-ukryte-dla-zwyk%C5%82ego-u%C5%BCytkownika-zawar/1713057508925147/ (dostęp: 11.12.2018).

opisowe służą do opisu treści oraz elementu zbiorów, pomagają w łączeniu autorów ze zbiorem i zapewniają ważny kontekst dotyczący zasobu po jego wykryciu. Ten typ metadanych daje możliwość wyszukiwania, przeglądania, sortowania i filtrowania informacji. Metadane opisowe mogą zawierać takie elementy, jak tytuł, opis, dane autorów. Jest to najczęstszy typ metadanych, znany większości twórców treści cyfrowych. Służą one do identyfikowania oraz opisywania zbiorów. Odgrywają szczególną rolę przy sortowaniu zbiorów bibliotecznych, w których poszczególne pozycje identyfikowane są najczęściej poprzez autora lub tytuł.

Metadane opisowe plików z rozszerzeniem .doc i .docx tworzone są przez edytor Microsoft Word. Zgodnie z przytoczoną definicją metadane opisowe mają za zadanie w pierwszej kolejności pomóc wyszukać dany plik dzięki tytułowi, autorowi. Metadane opisowe w plikach Word w darmowy sposób wyodrębnić można na co najmniej dwa sposoby. Pierwszy – poprzez wybranie prawego przycisku myszy → właściwości → szczegóły. Wówczas odnajdziemy informacje, w których zawarty jest tytuł i temat naszego pliku, tagi, kategorie i komentarze w opisie, a także dane o autorze, osobach, które modyfikowały dokument, liczbie modyfikacji, a także potencjalnie instytucji, z którą związany jest autor dokumentu (rys. 1).

Rysunek 1. Metadane opisowe wyodrębnione z pliku o rozszerzenie .docx



Źródło: opracowanie własne.

Nie jest to jednak jedyny sposób wyodrębniania metadanych opisowych z dokumentu tekstowego. Możliwe jest także skorzystanie z dowol-

nego darmowego generatora online (rys. 2). Wówczas uzyskujemy podobny zestaw metadanych opisowych. Generatory dostarczają informacje w sposób mniej usystematyzowany, bardziej chaotyczny, jednakże zaletą jest to, że w przypadku dokumentów o rozszerzeniu .doc (starsza wersja pakietu Office) uzyskujemy więcej informacji.

Rysunek 2. Metadane opisowe wyodrębnione online z pliku o rozszerzenie .doc

File Type	DOC
File Type Extension	doc
Mime Type	application/msword
Title	Temat: Metadane, czyli dane o danych
Author	Wojciech Mincewicz
Template	Normal
Last Modified By	WOJTEK

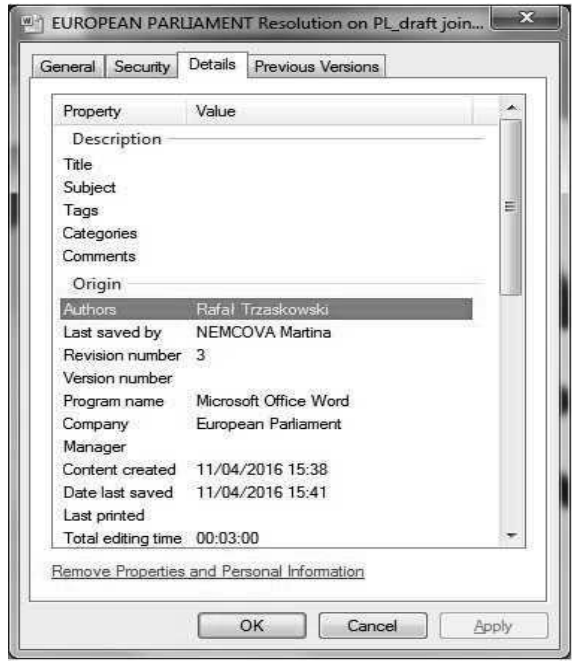
Źródło: opracowanie własne.

Metadane opisowe stanowią grupę danych, które zostały zaimplementowane do świata informatyki z bibliotekoznawstwa, gdzie spełniają porządkową funkcję i są jak najbardziej pożądane. W przypadku plików tekstowych, które przechowujemy na prywatnym dysku i które nie są nigdzie udostępniane, pełnią one podobną funkcję i są jak najbardziej właściwe. Sytuacja zmienia się jednak, gdy dokument, którego jesteśmy autorami, ma zostać udostępniony lub też umieszczony w Internecie. Wówczas warto zastanowić się, czy metadane opisowe nie mogą być źródłem problemów? Boleśnie przekonał się o tym Rafał Trzaskowki w kwietniu 2016 roku. Wówczas w Parlamencie Europejskim trwały prace nad rezolucją potępiającą zmiany w wymiarze sprawiedliwości[24]. Po kilku dniach Tomasz Poręba, jeden z parlamentarzystów Prawa i Sprawiedliwości ujawnił na jednym z portali społecznościowych, że nad rezolucją, pracował Rafał Trzaskowski (rys. 3). Polityk Platformy Obywatelskiej, mimo początkowego dementi, przyznał się w jednym z programów publicystycznych do prac nad dokumentem, gdy jego interlokutor Adam Bielan powiedział o metadanych[25].

[24] Zob. np.: js.ika, *Europarlament przyjął rezolucję. Jakie będą kolejne kroki*, TVN24, PAP, 13.04.2016, https://www.tvn24.pl/wiadomosci-z-kraju,3/europarlament-przyjal-rezolucje-ws-polski-jakie-beda-kolejne-kroki,635357.html (dostęp: 11.11.2018).

[25] *Bielan: To są metadane, pan jest autorem. Trzaskowski: nie wstydzę się tego*, TVN24, 15.04.2016, https://www.tvn24.pl/kropka-nad-i,3,m/kropka-nad-i-bielan-to-sa-metadane-pan-jest-autorem,636737.html (dostęp: 11.12.2018).

Rysunek 3. Analiza metadanych projektu rezolucji przyjętej przez Parlament Europejski w kwietniu 2018 roku



Źródło: Twitter @TomaszPoreba, https://twitter.com/tomaszporeba/status/720559877115011072 [dostęp: 24.11.2018].

Metadane opisowe plików z rozszerzeniem JPEG dostarczają przede wszystkim informacji o sprzęcie, na którym został uwieczniony obraz. Z metadanych opisowych zawartych w EXIF-ach odczytamy informację o samym aparacie, nazwie pliku (domyślnie zlokalizowana tam będzie data wykonania zdjęcia), obiektywie, ekspozycji (czyli ilości światła padającego na element światłoczuły), o lampie błyskowej czy skupieniu. Są więc one tworzone przez sam sprzęt, którym wykonujemy fotografię. W przypadku plików JPEG w ocenie autora o wiele bardziej przydatne do odczytania metadanych są narzędzia, które robią to w sposób automatyczny, a są dostępne w darmowej wersji. Ręczne odczytywanie metadanych (poprzez opcję właściwości → szczegóły) dostarcza dużo mniej informacji.

Rysunek 4. Metadane opisowe wyekstrahowane z tagów EXIF dla pliku z rozszerzeniem JPEG

Artist:	**Picasa**
Camera:	Nikon D60
Lens:	AF-S DX Zoom-Nikkor 18-55mm f/3.5-5.6G ED II Shot at 42 mm
Exposure:	Auto exposure, Not Defined, 1/60 sec, f/5.3, ISO 200
Flash:	Auto, Fired, Return detected
Focus:	AF-A, Mid-right, at 2.7m, with a depth of field of about 86cm, (from about 36cm before the focus point to about 50cm after) AF Area Mode: Dynamic Area (closest subject)
Date:	**June 25, 2014** 6:35:51PM (timezone not specified) (4 years, 5 months, 20 days, 5 hours, 8 minutes, 39 seconds ago, assuming image timezone of 1 hour ahead of GMT)
Time Zone Offset:	+01:00
File:	**3,872** × **2,592** JPEG (**10.0** megapixels) 3,623,255 bytes (3.5 megabytes)
Color Encoding:	WARNING: Color space tagged as sRGB, without an embedded color profile. **Windows and Mac browsers and apps treat the colors randomly**. Images for the web are most widely viewable when in the sRGB color space and with an embedded color profile. See my Introduction to Digital-Image Color Spaces for more information.

Źródło: opracowanie własne.

Metadane strukturalne

Metadane strukturalne (ang. *structural metadata*) to najbardziej niezrozumiana forma metadanych. Jest powszechnie ignorowana, nawet wśród osób, które pracują z metadanymi. Kiedy jest omawiana, może być mylona z innymi typami. Nawet osoby, które poprawnie rozumieją metadane strukturalne, nie zawsze doceniają ich pełny potencjał. Metadane strukturalne to dane dotyczące struktury treści. Pod pewnymi względami wcale nie jest to tajemnicze[26]. Ułatwiają one nawigację i prezentację zasobów elektronicznych. Dostarczają informacji na temat wewnętrznej struktury zasobów, znaczników struktury, takich jak: nazwa strony tytułowej, informacje o spisie treści, tytułach rozdziałów i podrozdziałów. W przypadku obrazów metadane strukturalne identyfikować będą sam obraz, a nie jak wcześniej sprzęt, którym został on uwieczniony. Zawarte

[26] M. Andrews, *Structural Metadata. Key to Structured Content*, 11.10.2017, https://storyneedle.com/structural-metadata-key-to-structured-content/ (dostęp: 12.12.2018).

tam będą dane między innymi o wymiarach obrazu czy jego rozdzielczości.

Metadane strukturalne plików z rozszerzeniem .doc i .docx powstają z automatu i są tworzone przez edytor Microsoft Word. Zgodnie z powyższą definicją dostarczają informacji o strukturze pliku. W przypadku plików z tym rozszerzeniem analizie podlega przede wszystkim zawartość pliku, to jest: liczba stron, statystyka wyrazów, liczba znaków, liczba wierszy, akapitów (rys. 5). Metadane strukturalne, podobnie jak opisowe, w plikach Word w darmowy sposób wyodrębnić można na co najmniej dwa sposoby ręcznie, jak również za pomocą generatorów online.

Rysunek 5. Przykładowe metadane strukturalne wyodrębnione z pliku .doc

Stan zawartości	
Typ zawartości	application/vnd.openxmlfor...
Strony	28
Statystyka wyrazów	4568
Liczba znaków	27408
Liczba wierszy	228
Liczba akapitów	63
Szablon	Normal
Skala	Nie
Linki brudne?	Nie
Język	

Źródło: opracowanie własne.

Generatory online, jak już sygnalizowano wcześniej, dostarczają informacje w sposób mniej ustrukturyzowany, ale są dokładniejsze, albowiem możemy tam odnaleźć na przykład informacje o spisie treści, czego nie zrobimy, pracując ręcznie (rys. 6).

Rysunek 6. Spis treści (dane zostały celowo usunięte ze względu na wrażliwe treści)

HYPERLINKS	
0	#_Toc516470996
1	#_Toc516470995
2	#_Toc516470994
3	#_Toc516470993
4	#_Toc516470992
5	#_Toc516470991
6	#_Toc516470990
7	#_Toc516470989
8	#_Toc516470988
9	#_Toc516470987
10	#_Toc516470983
11	#_Toc516470982
12	#_Toc516470981
13	#_Toc516470980
14	#_Toc516470979
15	#_Toc516470978
16	#_Toc516470977
17	#_Toc516470976
18	#_Toc516470975
19	#_Toc516470974
20	#_Toc516470973

Źródło: opracowanie własne.

Analiza metadanych strukturalnych fotografii zapisanej w formacie JPEG dostarcza brokerowi wiedzy na temat samego zdjęcia. Jest to informacja na temat oryginalnego rozszerzenia, wymiarów (szerokości, wysokości) czy rozdzielczości, a także szereg innych specjalistycznych informacji dotyczących samego zdjęcia związanych z kolorami czy jednostkami rozdzielczości.

Rysunek 7. Metadane strukturalne dla pliku zapisanego w formacie JPEG

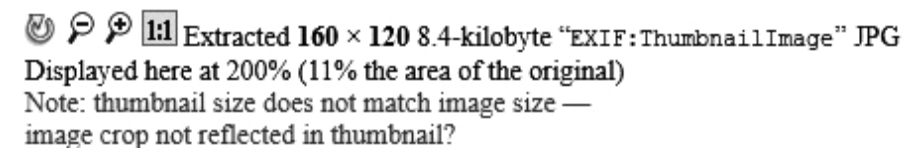
Composite
This block of data is computed based upon other items. Some of it may be wildly incorrect, especially if the image has been resized.

Aperture	5.30
Lens ID	AF-S DX Zoom-Nikkor 18-55mm f/3.5-5.6G ED II
Lens Spec	18-55mm f/3.5-5.6 G
Megapixels	10.0
Shutter Speed	1/60
Create Date	**2014:06:25** 18:35:51.70 4 years, 5 months, 19 days, 20 hours, 8 minutes, 39 seconds ago
Date/Time Original	**2014:06:25** 18:35:51.70 4 years, 5 months, 19 days, 20 hours, 8 minutes, 39 seconds ago
Modify Date	**2014:09:23** 22:22:16.70 4 years, 2 months, 21 days, 16 hours, 22 minutes, 14 seconds ago
Blue Balance	1.226563
Light Value	9.7
Red Balance	1.960938
Scale Factor To 35 mm Equivalent	1.5
Circle Of Confusion	0.020 mm
Depth Of Field	0.86 m (2.30 - 3.16 m)
Field Of View	31.4 deg (1.50 m)
Focal Length	42.0 mm (35 mm equivalent: 63.0 mm)
Hyperfocal Distance	16.62 m

Źródło: opracowanie własne.

Metadane strukturalne to także analiza miniatury zdjęcia, która pozwala stwierdzić, czy zdjęcie jest oryginalne, czy zostało poddane retuszowi, kadrowaniu lub innej obróbce. Dzięki informacji, która umieszczona jest na rysunku 8, można stwierdzić, że analizowane na potrzeby artykułu zdjęcie było edytowane, ponieważ kadrowanie obrazu nie odzwierciedla tego, co jest zawarte w miniaturze zdjęcia.

Rysunek 8. Metadane strukturalne miniatury pliku graficznego



Źródło: opracowanie własne.

Telefony i aparaty zapisują miniaturkę oryginalnego zdjęcia w tagach EXIF, ponieważ standard ten służy do wyświetlenia zdjęcia na widoku archiwum wykonanych zdjęć danego urządzenia. Dzieje się tak, gdyż aparaty czy telefony nie dysponują mocnymi procesorami, a miniaturka wyświetlana jest po prostu szybciej niż oryginalne zdjęcie wykonane w dużej rozdzielczości.

O miniaturkach zdjęć do końca życia będzie zapewne pamiętała amerykańska dziennikarka Catherine Schwartz. W 2003 roku opublikowała ona na swoim blogu kilka retuszowanych i przyciętych zdjęć. Jak się okazało, dzięki miniaturze fotografii, na niektórych z nich Schwartz była naga.

Metadane administracyjne

Metadane administracyjne (ang. *administrative metadata*) to dane techniczne, które powstają zazwyczaj automatycznie przy tworzeniu zasobu cyfrowego (zdjęcia lub pliku), a zawarte są w nich takie informacje, jak na przykład daty utworzenia, modyfikacji i użycia pliku, rozdzielczość, typ oraz model aparatu, przestrzeń kolorów, format pliku, kompresja, źródło światła, właściciel itp. W przypadku zdjęć cyfrowych – takie dane zapisane są bezpośrednio w pliku odwzorowania (na przykład w formacie EXIF) i nie powinno się ich usuwać – to cenne źródło informacji na temat procesu tworzenia zdjęcia i stanowią informację choćby o prawach autorskich.

Metadane administracyjne plików z rozszerzeniem .doc i .docx są generowane przez system operacyjny Windows, nie przez program. Zawierają podstawowe informacje o pliku, takie jak: rozmiar, data utworzenia, daty ostatniej modyfikacji oraz ostatniego użycia. Stanowią grupę metadanych, które identyfikują nasz dokument. Ich wyodrębnienie, podobnie jak w przypadku metadanych opisowych i strukturalnych, możliwe jest na dwa sposoby – w sposób ręczny, a także za pomocą darmowych generatorów metadanych. Dane te w przypadku metadanych administracyjnych mają podobny kształt, jeżeli chodzi o strukturę (rys. 9 zawiera metadane wyodrębnione ręcznie, a rys. 10 – dane wyodrębnione za pomocą darmowego generatora).

Rysunek 9. Metadane administracyjne pliku tekstowego

Źródło: opracowanie własne.

Rysunek 10. Metadane administracyjne pliku tekstowego, generator online

Software	Microsoft Office Word
Total Edit Time	1.7 hours
Last Printed	2018:06:11 07:08:00
Create Date	2018:06:11 04:55:00
Modify Date	2018:06:11 07:09:00
Pages	148
Words	41172
Characters	247034
Security	None
Company	Hewlett-Packard
Lines	2058
Paragraphs	575

Źródło: opracowanie własne.

Metadane administracyjne plików z rozszerzeniem JPEG stanowią zdecydowanie największą i najciekawszą z perspektywy infobrokera grupę danych, jakie przechowywane są w EXIF-ach. Zlokalizowane są tam

bowiem informacje identyfikujące sam plik, czyli w pierwszej kolejności o dacie powstania i modyfikacji, co umożliwia sprawdzenie, czy plik był w jakikolwiek sposób zmieniany.

Rysunek 11. Metadane administracyjne dla pliku graficznego

Composite
This block of data is computed based upon other items. Some of it may be wildly incorrect, especially if the image has been resized.

Aperture	5.30
Lens ID	AF-S DX Zoom-Nikkor 18-55mm f/3.5-5.6G ED II
Lens Spec	18-55mm f/3.5-5.6 G
Megapixels	10.0
Shutter Speed	1/60
Create Date	**2014:06:25** 18:35:51.70 4 years, 5 months, 19 days, 20 hours, 8 minutes, 39 seconds ago
Date/Time Original	**2014:06:25** 18:35:51.70 4 years, 5 months, 19 days, 20 hours, 8 minutes, 39 seconds ago
Modify Date	**2014:09:23** 22:22:16.70 4 years, 2 months, 21 days, 16 hours, 22 minutes, 14 seconds ago
Blue Balance	1.226563
Light Value	9.7
Red Balance	1.960938
Scale Factor To 35 mm Equivalent	1.5
Circle Of Confusion	0.020 mm
Depth Of Field	0.86 m (2.30 - 3.16 m)
Field Of View	31.4 deg (1.50 m)
Focal Length	42.0 mm (35 mm equivalent: 63.0 mm)
Hyperfocal Distance	16.62 m

Źródło: opracowanie własne.

W metadanych administracyjnych plików JPEG przy odrobinie szczęścia i nieuwadze autora zdjęcia odnajdziemy także koordynaty GPS. Większość urządzeń mobilnych automatycznie generuje lokalizację, co niekiedy bywa zgubne. Boleśnie przekonał się o tym Higinio O. Ochoa III, jeden z członków grupy Anonymous o nazwie Cabin Cr3w, który wykradł i upublicznił w Internecie dane setki policjantów. Funkcjonariusze bez problemu byli w stanie zlokalizować hakera, a stało się to dzięki koordynatom GPS zdjęcia, które pozostały na stronie internetowej policji[27].

Inny osobliwy przykład tego samego typu stanowi John McAfee, programista i twórca firmy McAfee Associates, który popadł w konflikt

[27] Za. S. Hang, *The Life of an Ex-Hacker Who Is Now Banned from Using the Internet*, 25.04.2015, https://gizmodo.com/the-life-of-an-ex-hacker-who-is-now-banned-from-using-t-1700074684 (dostęp: 12.12.2018). (Higinio O. Ochoa został skazany przez sąd na zakaz korzystania z Internetu – nie może dotknąć niczego, co jest do niego podłączone. Gdy chce obejrzeć film, musi to czynić w towarzystwie żony).

z prawem. W 2012 roku w Belize został posądzony o zabójstwo swojego sąsiada. Tuż po przesłuchaniu w tej sprawie przez tamtejszy sąd uciekł. W ręce organów ścigania wpadł w Gwatemali, gdzie się ukrywał. Do swojej kryjówki zaprosił dziennikarzy magazynu „Vice", którzy dokumentowali całość spotkania fotografiami, które opublikowali, nie usuwając przed tym koordynatów GPS[28].

Rysunek 12. Metadane administracyjne zawierające koordynaty GPS

Basic Image Information

Target image: https://christianheilmann.com/wp-content/uploads/2014/10/me.jpg

Camera:	Lge Nexus 5
Lens:	1.2 mm
Exposure:	1/40 sec, f/2.9, ISO 102
Flash:	none
Date:	**October 19, 2014** 4:06:20PM (timezone not specified) (5 years, 24 days, 20 hours, 34 minutes, 54 seconds ago, assuming image timezone of 1 hour ahead of GMT)
Location:	Latitude/longitude: **59**° 19′ 6.8" North, **18**° 3′ 35.5" East (59.318554, 18.059870) Map via embedded coordinates at: Google, Yahoo, WikiMapia, OpenStreetMap, Bing (also see the Google Maps pane below) Altitude: 0 meters (0 feet) Camera Pointing: East-southeast Timezone guess from earthtools.org: 1 hour ahead of GMT
File:	**1,280** × **960** JPEG (**1.2** megapixels) 425,979 bytes (416 kilobytes)
Color Encoding:	**WARNING:** Color space tagged as sRGB, without an embedded color profile. **Windows and Mac browsers and apps treat the colors randomly**. Images for the web are most widely viewable when in the sRGB color space and with an embedded color profile. See my Introduction to Digital-Image Color Spaces for more information.

Źródło: opracowanie własne.

Sposoby usuwania metadanych

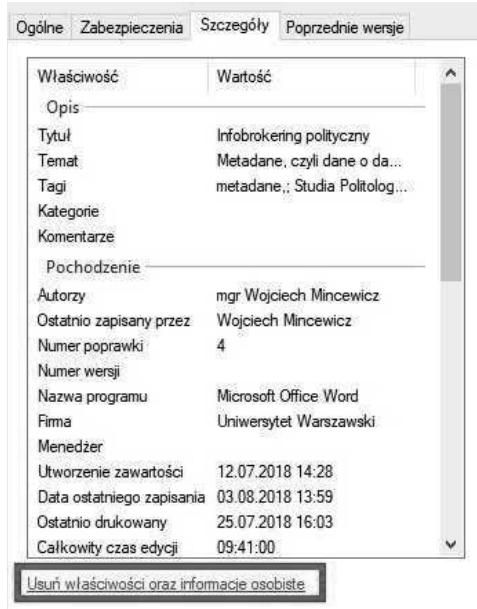
Znając już typy metadanych, można udzielić odpowiedzi na pytanie, które metadane plików tekstowych i graficznych usuwać i jak to robić tak, aby udostępniając dokument zyskać pewność, że czynimy to bez szkody dla naszej prywatności. Warto podkreślić, że ta część rozważań stanowi autorskie przemyślenia i nie powinna być w żadnym stopniu uznana za rozstrzygającą. Nie ma także jednej unikalnej metody, która zagwarantuje

[28] C.G. Weissman, *The Insane Life of Former Fugitive and Eccentric Cybersecurity Legend John McAfee*, 23.06.2015, https://www.businessinsider.com/the-insane-life-of-john-mcafee-2015-7?IR=T (dostęp: 13.12.2018).

użytkownikom całkowite bezpieczeństwo. Znając *case study* przytoczone w artykule oraz typologie przedstawioną powyżej, każdy sam powinien odpowiedzieć na pytanie, które dane usunąć i kiedy. Warto pamiętać, jaką rolę odgrywają poszczególne typy i jakie dane są tam przechowywane.

Metadane opisowe w pierwszej kolejności odgrywają rolę pomocniczą. Mają nam – użytkownikom ułatwić wyszukiwanie treści i spełniają pożyteczną funkcję, gdy są przechowywane w plikach na własnym komputerze. Przytoczony w artykule przykład dobitnie pokazuje, że mogą być jednak źródłem problemów, dlatego w ocenie autora usuwanie tych danych powinno być złotą zasadą postępowania z metadanymi opisowymi, gdy udostępniamy dokument. Jak to uczynić? W przypadku dokumentów z rozszerzeniem .docx metadane opisowe można usunąć ręcznie, wybierając usuń właściwości → informacje osobiste (zob. rys. 13). Wówczas w kolejnym oknie usuwamy: tytuł, temat, tagi oraz dane autora. Dane te możemy także usuwać ręcznie, poprzez ich edytowanie.

Rysunek 13. Ręczne usuwanie metadanych z pliku tekstowego

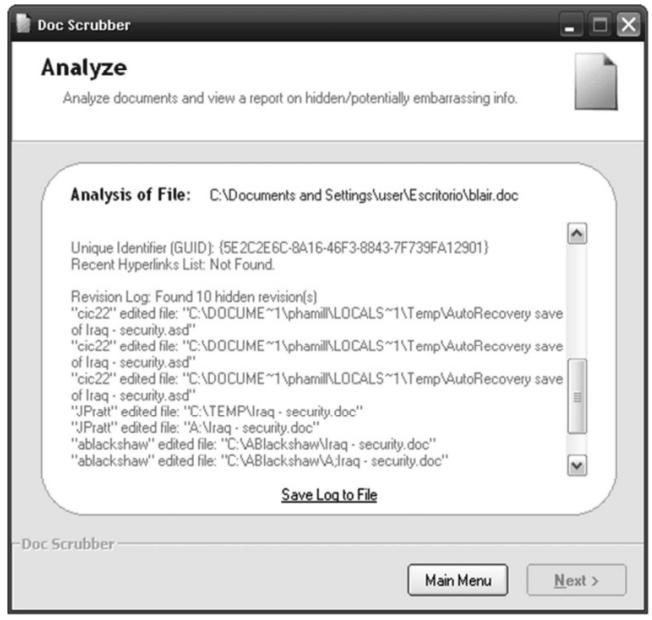


Źródło: opracowanie własne.

Inna metoda to mechaniczne usunięcie danych ukrytych w dokumencie, dzięki darmowemu programowi na przykład Doc Scrubber (rys. 14). Jest to jedno z narzędzi firmy Javacool Software, znanej między innymi ze Spyware Blastera, służące do przeglądania i usuwania ukrytych danych

zaszytych. Jest dedykowane szczególnie dla starszych wersji plików z rozszerzeniem .doc, których nie możemy usunąć ręcznie na naszym komputerze osobistym[29]. Bliźniacze rozwiązania w zakresie mechanicznego usuwania metadanych zapisanych w tagach EXIF jest Easy Exif Delete.

Rysunek 14. Widok z programu DOC Scrubber



Źródło: opracowanie własne.

W przypadku dokumentów tekstowych, jak również zdjęć dobrym i skutecznym sposobem jest zastanowienie się, czy wszystkie formaty zapisu przechowują metadane. W artykule zostały omówione najpopularniejsze formaty zapisów plików tekstowych i graficznych, to jest .doc/.docx i JPEG. Można na przykład, zamiast rozpowszechniania dokumentu programu Word, przekonwertować dokument na format .rtf, .txt lub do klasycznego pdf. W przypadku obrazów zamiast formatu JPEG stosować format PNG. Formaty te przechowują małe ilości meta znaczników lub nie przechowują ich wcale.

[29] Leszek Ignatowicz sugeruje, że najlepszym sposobem na usunięcie metadanych z plików tekstowych zapisanych w formacie .docx jest użycie inspektora dokumentów. Zob. L. Ignatowicz, *Cyfrowe ślady mówią. Poradnik ochrony prywatności*, Warszawa 2015, s. 7. Jest to kwestia co najmniej dyskusyjna, albowiem w ten sposób usunięte zostaną tylko metadane opisowe pliku. Bez zmian pozostaną metadane administracyjne, które dostarczają informacji na temat daty powstania i modyfikacji dokumentu.

Rysunek 15. Analiza pliku graficznego zapisanego w formacie PNG[30]

Basic Image Information

Target file: hostessy.png

File:	**960 × 640** PNG 920,804 bytes (0.88 megabytes)
Color Encoding:	**WARNING:** No color-space metadata and no embedded color profile: **Windows and Mac web browsers treat colors randomly**. Images for the web are most widely viewable when in the sRGB color space and with an embedded color profile. See my Introduction to Digital-Image Color Spaces for more information.

Źródło: opracowanie własne.

Analiza skrajnych przypadków przytoczonych w artykule uzmysławia, że przed udostępnieniem pliku tekstowego lub cyfrowego warto usunąć z niego metadane opisowe, które pozwalają identyfikować twórcę dokumentu. Wraz z rozwojem Internetu zaczęły powstawać pierwsze programy, które służyły do masowego odnajdywania i analizowania metadanych zawartych w upublicznionych dokumentach (ang. *metadata harvesting*). Metodyka ich działania nakierowana jest na pozyskiwanie danych ukrytych pozostawionych na witrynach internetowych dokumentów. Za pomocą odpowiedniego zapytania skierowanego do internetowej wyszukiwarki lokalizowane są dokumenty opublikowane przez interesującą nas organizację. Po ich zlokalizowaniu program analizuje pozostawione przez nieświadomych użytkowników metadane.

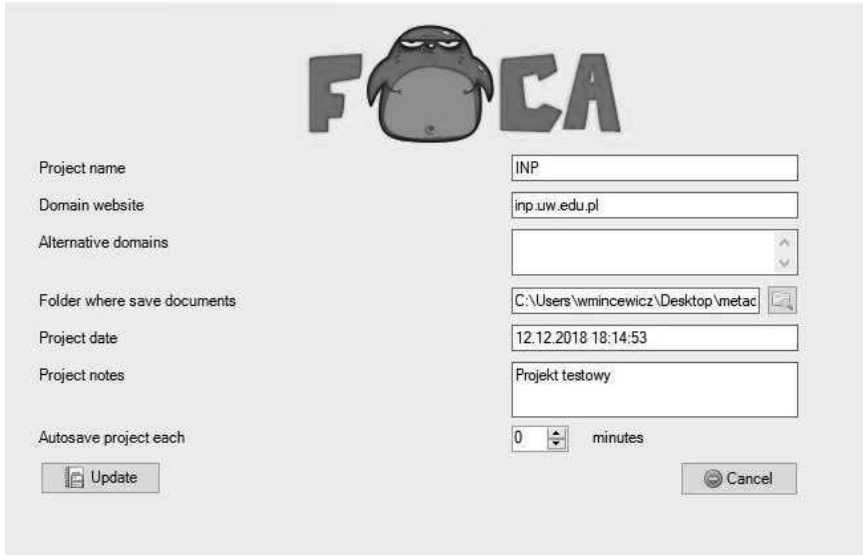
Jednym z najpopularniejszych programów przeznaczonych do *metadata harvesting* jest Fingerprinting Organizations with Collected Archives (FOCA). Program został stworzony przez firmę ElevenPaths w celu wyszukiwania, pobierania i analizowania dokumentów dla pozyskania cyfrowych informacji pozostawionych przez użytkowników świadomie bądź nie w zasobach Internetu[31]. Mechanizm działania FOCA jest bar-

[30] Format PNG przechowuje tylko szczątkowe metadane strukturalne dla plików graficznych (rozmiar), nie zachowuje natomiast metadanych opisowych i administracyjnych.

[31] Pozyskiwanie i wyodrębnianie metadanych nie jest jedynym zastosowaniem programu FOCA, ponieważ za jego pomocą można również m.in.: analizować budowę witryny, identyfikować Domain Name System, wyszukiwać kopie zapasowe strony, wykrywać katalogi lub luki w zabezpieczeniach. Jednak większość z tych funkcji dostępna jest w pełnej, płatnej wersji programu. Darmowa wersja dostępna obecnie (styczeń 2019) w wersji 3.4 umożliwia skorzystanie z podstawowych funkcji programu, wśród których znajduje się analiza metadanych dokumentów. Pełen opis wszystkich funkcji programu dostępny na stronie producenta: https://www.elevenpaths.com/labstools/foca/index.html (dostęp: 1.01.2019).

dzo prosty: za pomocą odpowiedniego zapytania kierowanego do wyszukiwarki – Google, Bing i Exalead, odnajdujemy dokumenty opublikowane przez interesującą nas organizację i analizuje zawarte w nich metadane.

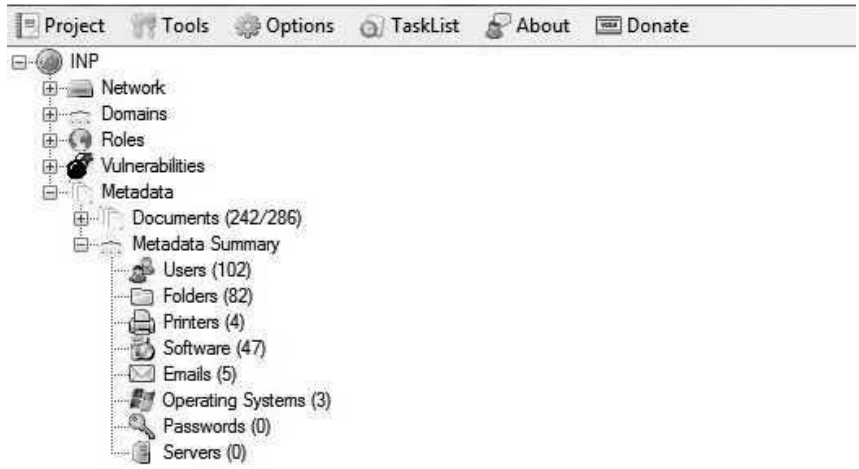
Rysunek 16. Widok programu FOCA po utworzeniu nowego projektu i wybraniu docelowej domeny



Źródło: opracowanie własne.

Po utworzeniu nowego projektu i wybraniu docelowej domeny (rys. 16) możemy od razu rozpocząć poszukiwanie publicznie dostępnych dokumentów. Program FOCA bez problemu odnajdzie dokumenty umieszczone na testowanej domenie. Po zakończeniu wyszukiwania i wybraniu opcji Download All można pobrać wszystkie odnalezione pliki, a dzięki Extract ALL Metadata / Analyze All Metadata otrzymujemy zbiór metadanych, które uzyskał dla nas FOCA.

Rysunek 17. Widok programu FOCA po wyodrębnieniu metadanych z plików odnalezionych na analizowanej domenie



Źródło: opracowanie własne.

Finalnym efektem pracy programu jest pozyskanie przez infobrokera w pierwszej kolejności metadanych opisowych plików umieszczonych na analizowanej domenie (102 pliki miały nazwy użytkowników). Inne pozyskane dane to informacja o ścieżce zapisu plików, o sprzęcie oraz oprogramowaniu, którym posługują się autorzy plików. Program pozyskał także pięć adresów e-mail oraz informacje o systemach operacyjnych. W skrajnie ekstremalnych przypadkach za pomocą narzędzi do zmechanizowanego pozyskiwania metadanych pentesterzy[32] pozyskują informacje o hasłach pozostawionych przez nieświadomych użytkowników. Metadane administracyjne oraz strukturalne plików, które zostały pozyskane z testowanej domeny, pozyskujemy w sposób „klasyczny" opisany w pierwszej części artykułu.

Jednym z najpopularniejszych sposobów przekazywania plików tekstowych i graficznych jest przesyłanie ich za pomocą poczty e-mail. Czy jeśli przed wysłaniem pliku w załączniku poczty elektronicznej usuniemy metadane opisowe, które identyfikują nasz dokument, to czy nasza przesyłka jest całkowicie pozbawiona meta znaczników? Absolutnie nie. Nawet przy zachowaniu wszelkich środków cyberhigieny szereg danych ukrytych zawiera sam nasz e-mail, a właściwie sam jego nagłówek. Dane

[32] Pentester – osoba przeprowadzająca kontrolowany atak na jakiś system teleinformatyczny w celu oceny bieżącego stanu zabezpieczenia tego systemu; www.nowewyrazy.uw.edu.pl/haslo/pentester.html?pdf=1 (dostęp: 19.06.2019).

te mogą służyć między innymi potwierdzeniu autentyczności wiadomości otrzymanej z nieznanego źródła czy zlokalizowaniu nadawcy.

Pierwszą grupę metadanych, które zawiera nagłówek e-mail, są metadane opisowe pozwalające zidentyfikować nadawcę wiadomości. Meta znaczniki tam zawarte pozwalają stwierdzić, kto jest nadawcą e-maila, kiedy została napisana wiadomość, jaki jest jej ID oraz adres serwera gmail, gdzie zapisana jest wiadomość oraz kto jest jej adresatem.

Rysunek 18. Metadane opisowe nagłówka e-mail[33]

Źródło: opracowanie własne.

Kolejną grupę metadanych zawartych w nagłówku e-mail stanowią te, które należy zaklasyfikować jako metadane administracyjne. W pierwszej kolejności będą to informacje, które identyfikują czas powstania e-maila: kiedy został wysłany przez autora, kiedy znalazł się na serwerze DNS i kiedy został dostarczony do adresata(ów) wiadomości. Czas ten w nagłówku e-maila zapisuje się według czasu pacyficznego standardowego (ang. *Philippine Standard Time*, PST)[34]. Widoczne na rysunku 19 pozostałe składowe metadanych administracyjnych nagłówka e-mail – *Authenticated Received Chain*, uwierzytelniają kolejne serwery DNS, przez które przesyłana jest wiadomość e-mail. ARC składa się z trzech kluczowych komponentów: *Authentication-Results*; *Seal*; *Message-Signature*. Z perspektywy infobrokera najistotniejszą rolę spełnia pierwszy z nich, odpowiadający za weryfikację autentyczności otrzymanej wiadomości. Jeżeli e-mail został pozytywnie zweryfikowany przez uwierzytelnianie DMARC (ang. *Domain-based Message Authentication, Reporting, and Conformance*), zostanie on przekazany do odczytu. Podczas sprawdzania autentyczności DMARC korzysta z rekordów SPF (ang. *Sender Policy Framework*) i DKIM (ang. *DomainKeys Identified Mail*). Wiadomość, która nie przejdzie weryfikacji SPF lub DKIM, uruchomi zasady DMARC

[33] Pozostali adresaci e-maila zostali w sposób celowy zanonimizowani.

[34] Autor w pracy nad artykułem posługuje się adresem e-mail w aplikacji gmail.com, której serwery zlokalizowane są w Stanach Zjednoczonych Ameryki, dlatego przy analizie nagłówka e-mail otrzymujemy informację o czasie strefy czasowej PST.

i najprawdopodobniej nie dotrze do adresata, ponieważ zostanie uznana za spam lub niebezpieczną przesyłkę[35].

Rysunek 19. Metadane administracyjne nagłówka e-mail

```
Delivered-To: w.mincewicz@student.uw.edu.pl
Received: by 2002:a17:90a:c482:0:0:0:0 with SMTP id j2-v6csp4181879pjt;
        Tue, 6 Nov 2018 07:49:33 -0800 (PST)
X-Received: by 2002:a17:902:a516:: with SMTP id s22-v6mr5214202plq.255.1541519366055;
        Tue, 06 Nov 2018 07:49:26 -0800 (PST)
ARC-Seal: i=1; a=rsa-sha256; t=1541519366; cv=none;
        d=google.com; s=arc-20160816;
        b=cy4c8pRAm5xw4kduJjzjdi3+tIiwohNpAtD09DUc5+LwL+Y7Kc3D8gYfRcsLIvh0PN
         1hgVJQ5EypU3MnfWXz+eYRnnicWZdScfcrXhscb7XvE/dtyuVHBvF6FRpWyUk7VhZGZu
         EW2Wj77wMVZNcPm/U31rbRkzJxqMb7zqHbfutEVYmUlEm+Qpx6fnixAawyZ1qUMvK2eH
         MWguUm7gw6o2XUKUfAlAjMqAclt2VvG70pxYZaCfPNPVnqUHQYyAA+lqce2Zn8UASdiB
         XpDM1zvgTfa/JAMG0aWUuKBhLukNMpa+BS4JWSovJGN3WfxNJrmnkpLA2OimQ8Q7KgLP
         iZJA==
ARC-Message-Signature: i=1; a=rsa-sha256; c=relaxed/relaxed; d=google.com; s=arc-20160816;
        h=to:subject:message-id:date:from:mime-version:dkim-signature;
        bh=s9doqhbBV+wP4/G+/aEryI2QB+QSFKm5Gkwk49wNCm4=;
        b=s/zqUG6BivwtTU1LcSwZJ5BgcQlj4RWtGU3kav2LOSCH6pnO/pn1s3LGnWlX175x2n
         6skEoCGGrz+Nt0E15ZVAQpNrVcXOfiV9TqpoZWjaCRr7F8CBHFygaVKdpKgkUp7fuiWh
         uc+J8PkmPbURMGlSTukObyP4ckrdEm9VLrFzlsIISiaSQiifvWzL0s9ww9pbNcnBv3Q0
         YG12HW+Skm/5tdfffFwnzu9ANW593w664M0DEqVoXI8c4jSA+jQzyR5qRAXYPddmFG3z
         PizDtSe6/MzEzXPwBCvzuC4CXgyCEPFuN601+guHGDP0UepZEgYiRax9wyyqxDs8MB8Z
         6lKA==
ARC-Authentication-Results: i=1; mx.google.com;
       dkim=pass header.i=@uw-edu-pl.20150623.gappssmtp.com header.s=20150623 header.b=q1LlFtpF;
       spf=pass (google.com: domain of d.mider@uw.edu.pl designates 209.85.220.41 as permitted sender)
smtp.mailfrom=d.mider@uw.edu.pl
Return-Path: <d.mider@uw.edu.pl>
Received: from mail-sor-f41.google.com (mail-sor-f41.google.com. [209.85.220.41])
        by mx.google.com with SMTPS id b17-v6sor44882111pls.11.2018.11.06.07.49.25
        for <w.mincewicz@student.uw.edu.pl>
        (Google Transport Security);
        Tue, 06 Nov 2018 07:49:25 -0800 (PST)
Received-SPF: pass (google.com: domain of d.mider@uw.edu.pl designates 209.85.220.41 as permitted sender) client-
ip=209.85.220.41;
```

Źródło: opracowanie własne.

Ostatni typ metadanych, który możemy zidentyfikować w nagłówku e-maila, to grupa metadanych strukturalnych, najtrudniej definiowalnych. W przypadku e-maila, który stanowi *case study* dla artykułu, metadanymi strukturalnymi będą dane o załączniku do e-maila. W innych przypadkach za metadane strukturalne uznawane będą na przykład informacje o czcionce, które odpowiedzialne są za formatowanie stopki e-maila.

[35] Więcej na ten temat: G. Colburn, *How to Explain Authenticated Received Chain (ARC) in Plain English*, 6.11.2015, https://blog.returnpath.com/how-to-explain-authenticated-received-chain-arc-in-plain-english-2/ (dostęp: 2.01.2019).

Rysunek 20. Metadane strukturalne nagłówka zawierające informacje o czcionkach w stopce e-maila

```
<html xmlns:v=3D"urn:schemas-microsoft-com:vml" xmlns:o=3D"urn:schemas-micr=
osoft-com:office:office" xmlns:w=3D"urn:schemas-microsoft-com:office:word" =
xmlns:m=3D"http://schemas.microsoft.com/office/2004/12/omml" xmlns=3D"http:=
//www.w3.org/TR/REC-html40">
<head>
<meta http-equiv=3D"Content-Type" content=3D"text/html; charset=3Diso-8859-=
2">
<meta name=3D"Generator" content=3D"Microsoft Word 15 (filtered medium)">
<style><!--
/* Font Definitions */
@font-face
=09{font-family:"Cambria Math";
=09panose-1:2 4 5 3 5 4 6 3 2 4;}
@font-face
=09{font-family:Calibri;
=09panose-1:2 15 5 2 2 2 4 3 2 4;}
/* Style Definitions */
p.MsoNormal, li.MsoNormal, div.MsoNormal
=09{margin:0cm;
=09margin-bottom:.0001pt;
=09font-size:11.0pt;
=09font-family:"Calibri",sans-serif;
=09mso-fareast-language:EN-US;}
a:link, span.MsoHyperlink
=09{mso-style-priority:99;
```

Źródło: opracowanie własne.

Podsumowanie

Rozwój społeczeństwa informacyjnego, które stanowi następstwo społeczeństwa przemysłowego, możliwy jest dzięki rozwojowi techniki. Ten z kolei napędzany jest przez rozwój Internetu. Świat w realiach XXI wieku staje się globalną wioską, w której kolejne obszary życia społeczeństw przenoszą się do wirtualnej przestrzeni. Transfer ten oraz rozwój technologii i coraz łatwiejszy do nich dostęp są na pewno szansą na dalszą prosperitę, ale stanowi także zagrożenie dla naszej prywatności i anonimowości. Organizacje zajmujące się bezpieczeństwem, korporacje czy inne podmioty zabiegają o pozyskanie jak najszerszej gamy informacji o każdym z nas. Motywacje ich działania nie zawsze są w pełni zgodne z wyobrażeniami użytkowników Internetu. Informacjami, które sami produkujemy, wytwarzając kolejne pliki tekstowe lub graficzne, są metadane, czyli ustrukturyzowane dane zapisane w logiczny sposób, które w różnorodny sposób identyfikują i charakteryzują dokument elektroniczny. Mogą w nich być zawarte informacje identyfikujące dokument, jak również samego autora pliku.

Większość metadanych nie jest sama w sobie szkodliwa, pełnią one wręcz w jakimś stopniu pożyteczną funkcję. Przykładem mogą być metadane opisowe, które służą do opisu treści oraz elementu dokumentu. Ten

typ metadanych daje możliwość wyszukiwania pliku w ogromnych zbiorach danych poprzez tytuł lub dane autora. Służą więc one do identyfikowania oraz opisywania zbiorów, zatem w wypadku, gdy nikt poza osobami uprawnionymi nie ma do nich dostępu, są pożądane i pełnią pomocniczą funkcję. Problem pojawia się jednak, gdy plik tekstowy zostanie udostępniony lub zamieszczony w ogólnodostępnym Internecie. Wówczas warto zastanowić się, czy chcemy, aby każdy z użytkowników Internetu był w stanie identyfikować nas jako twórców danego dokumentu. Aby nie było to możliwe, warto znać podstawowe techniki defensywne, które mają służyć usunięciu metadanych z pliku. W trosce o swoją anonimowość oraz prywatność warto z którejś z nich skorzystać, szczególnie w sytuacji, gdy wiemy, że nasz plik będzie dostępny publicznie i może zostać pozyskany przez każdego internautę.

Streszczenie

W artykule dokonano pogłębionej refleksji nad zagadnieniem metadanych, czyli danych, które definiują lub opisują inne dane. W warstwie teoretycznej wyodrębnione zostały trzy typy metadanych: opisowe, strukturalne i administracyjne. Metadane opisowe służą do odnajdywanie i identyfikacji kluczowych informacji, które umożliwiają lokalizację obiektu. Metadane strukturalne opisują strukturę wewnętrzną danego obiektu, metadane administracyjne odnoszą się do informacji technicznych, zawarte są tam informacje na przykład o czasie i sposobie utworzenia pliku. Celem publikacji jest dostarczenie wiedzy teoretycznej, jak również praktycznej. W drugiej części artykułu dokonana została egzemplifikacja pojęcia na przykładzie plików graficznych i tekstowych oraz wskazane zostały proste techniki samoobrony, które umożliwiają usuniecie metadanych przed udostępnieniem pliku. Uzupełnienie tekstu stanowi analiza możliwości pozyskiwania meta informacji za pomocą programu Fingerprinting Organizations with Collected Archives (FOCA), który służy do zmechanizowanego pozyskiwania metadanych, a także refleksja nad tym, jakie metadane zawiera nagłówek e-maila.

Wojciech Mincewicz

Metadata – a Silent Privacy Killer

The article has a deeper reflection on the issue of metadata, that is, data which are defined or describe other data. The theoretical layer extracted three types of

metadata: descriptive, structural, and administrative. Descriptive metadata is used to find and identify key information that allows the location of an object. Structured metadata describes the internal structure of the object, but administrative metadata refers to the technical information, where information is provided for example about the time and how the file was created. The purpose of the publication is to provide theoretical knowledge as well as practical. The second part of the article depicts the concepts of graphic and text files, and simple self-defense techniques are indicated, which allow you to remove metadata before sharing the file. The supplementing of article is: analysis the ability to extract meta information by Fingerprinting Organizations with Collected Archives (FOCA), which is used to mechanizedly extract metadata reflection on what the metadata includes the email header.

Key words: *metadata, anonymity, information security, FOCA, e-mail header*

Bibliografia

Andrews M., *Structural Metadata. Key to Structured Content*, 11.10.2017, https://storyneedle.com/structural-metadata-key-to-structured-content/ (dostęp: 14.12.2018).

Bretherton F.P., Singley P.T., *A User's View* [w:] *Proceedings of the 7th International Working Conference on Scientific and Statistical Database Management*, Charlottesville 1994.

Campbell D., *Dublin Core Metadata and the Australian Metaweb Project*, 10th National Library Technicians' Conference, Fremantle, 8–10.09.1999, https://www.nla.gov.au/bla/staffpaper/dcampvell1.html (dostęp: 15.10.2018).

Caplan P., *Metadata Fundamentals for All Librarians,* Chicago 2003.

Colburn G., *How to Explain Authenticated Received Chain (ARC) in Plain English*, 6.11.2015, https://blog.returnpath.com/how-to-explain-authenticated-received-chain-arc-in-plain-english-2/ (dostęp: 15.01.2019).

Dempsey L., *Registries: The Intelligence in Network*, 20.08.2006, http://orweblog.oclc.org/Registries-the-intelligence-in-the-network/ (dostęp: 11.11.2018).

Foulonneau M., Riley J., *Metadata for Digital Resources. Implementation, Systems Design and Interoperability*, Oxford 2008.

Garlicki J., Mider D., Mincewicz W., *Pozyskiwanie informacji w Internecie metodą Google Hacking – biały, szary czy czarny wywiad?*, „Przegląd Bezpieczeństwa Wewnętrznego" 2019, nr 20.

Griffel D., McIntosh S., *ADMINS – A Progres Raport*, MIT 1967, https://dspace.mit.edu/bitstream/handle/1721.1/82974/09487802.pdf (dostęp: 20.09.2018).

Hang S., *The Life of an Ex-Hacker Who Is Now Banned from Using the Internet*, 25.04.2015, https://gizmodo.com/the-life-of-an-ex-hacker-who-is-now-banned-from-using-t-1700074684 (dostęp: 14.12.2018).

Ignatowicz L., *Cyfrowe ślady mówią. Poradnik ochrony prywatności*, Warszawa 2015.

Litwin L., Rossa M., *Metadane geoinformacyjne w INSPIRE i SDI*, Gliwice 2010.

Marr. B, *Every Day? The Mind-Blowing Stats Everyone Should Read*, Forbes, 21.05.2018, https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-

every-day-the-mind-blowing-stats-everyone-should-read/#4ffe1f6360ba (dostęp: 15.01.2019).

Nahotko M., *Metadane – sposób na uporządkowanie Internetu*, Kraków 2004.

Niebrzydowska M., Kotowicz R., *Wstęp do informatyki śledczej*, „Przegląd Bezpieczeństwa Wewnętrznego" 2012, nr 6.

Weissman C.G., *The Insane Life of Former Fugitive and Eccentric Cybersecurity Legend John McAfee*, 23.06.2015, https://www.businessinsider.com/the-insane-life-of-john-mcafee-2015-7?IR=T (dostęp: 15.12.2018).

Zeng, M.L. *Metadata Basics*, http://marciazeng.slis.kent.edu/metadatabasics/cover.htm (dostęp: 14.11.2018).

VOL. 54 STUDIA POLITOLOGICZNE
DOI: 10.33896/SPolit.2019.54.10

Paweł Tomczyk
ORCID: 0000-0002-4983-5864
Daniel Mider
ORCID: 0000-0003-2223-5997
Józef Grzegorczyk
ORCID: 0000-0001-5348-3737

Inwigilacja elektroniczna jako metoda pozyskiwania informacji – ewaluacja i prognozy

Słowa kluczowe:
społeczeństwo nadzoru, informatyka społeczna, inwigilacja elektroniczna, infobrokering

STUDIA I ANALIZY

Wprowadzenie

W dyskursie filozoficznym wiedza bywa utożsamiana z władzą. Dla Michela Foucaulta te dwa pojęcia są kluczowe dla rozumienia rzeczywistości i nierozdzielne do tego stopnia, że ukuł pojęcie „wiedzy-władzy". Pojmował te kategorie jako nierozłączne: „Ani władza nie może być praktykowana bez wiedzy, ani wiedza nie może nie płodzić władzy"[1].

Pomimo że koncepcja ta została wyłożona na początku lat 70. ubiegłego wieku[2], to intensywny rozwój technologii informacyjnych, a w szczególności środków komunikowania elektronicznego sprawił, że refleksje M. Foucaulta pozostają w niekwestionowany sposób aktualne. Współczesny status informacji jest nie do przecenienia – jest to szczególne

[1] M. Foucault, *Gry władzy*, przekł. T. Komendant, „Literatura na Świecie" 1988, nr 6, s. 319.

[2] Tenże, *Historia seksualności*, przekł. B. Banasiak, T. Komendant, K. Matuszewski, Gdańsk 2010; tenże, *Nadzorować i karać*, przekł. T. Komendant, Warszawa 1998.

dobro niematerialne równoważne lub cenniejsze od dóbr materialnych, w myśl reguł wyłożonych w *Zmianie władzy* Alvina Tofflera[3]. Intensywny rozwój technologii informacyjnych doprowadził do ujawnienia się lub zintensyfikowania licznych negatywnych zjawisk, których induktorem jest łatwość pozyskiwania i dystrybucji informacji z użyciem urządzeń elektronicznych. Technologie informacyjne zamknęły współczesne społeczeństwa w swoistym więzieniu opisywanym wprost w koncepcji społeczeństwa nadzorowanego (*surveillance society*) lub alegorycznie w koncepcji Panoptykonu[4]. Taki status informacji we współczesnych społeczeństwach sprawia, że staje się ona centralną kategorią analityczną, bez której niemożliwe jest zrozumienie relacji społecznych, w tym relacji władzy.

Inwigilacja ma obecnie charakter powszechny, stała się nieodłącznym elementem krajobrazu współczesnych społeczeństw, co skłoniło do charakteryzowania ich jako *„eavesdropping societies"* („społeczeństwa podsłuchu")[5]. Niniejszy artykuł ogniskuje się na jednym z elementów przynależnych społeczeństwu nadzorowanemu – inwigilacji z użyciem narzędzi elektronicznych. Autorzy podejmują w nim próbę odpowiedzi na szereg następujących pytań. Po pierwsze, jakie typy negatywnych zjawisk są wytwarzane i intensyfikowane przez technologie inwigilacji elektronicznej? Po wtóre, jak głęboki jest stan „bezbronności inwigilacyjnej" współczesnych społeczeństw, to jest jakie są możliwości urządzeń służących inwigilacji? Po trzecie, czy istnieje możliwość praktycznego przeciwstawienia się im, a jeśli tak – w jaki sposób i jakie są tego granice? Po czwarte, jaka jest geneza tych zjawisk i jakie spodziewane scenariusze przyszłości można szkicować na podstawie antycypacji zaobserwowanych trendów? Tak zdefiniowany zbiór pytań badawczych wymaga oglądu zarazem z dwóch perspektyw: socjologicznej i technicznej.

---

[3] A. Toffler, *Zmiana władzy. Wiedza, bogactwo i przemoc u progu XXI stulecia*, przekł. P. Kwiatkowski, Poznań 2003.

[4] O. Gandy, *Data Mining and Surveillance In the Post – 9/11 Environment*, [w:] K. Ball, F. Webster (red.), *The Intensification of Surveillance. Crime, Terrorism and Warfare in the Information Era*, Londyn 2003; D. Lyon, *The Electronic Eye. The Rise of Surveillance Society*, Minneapolis 1994; G.T. Marx, *The Surveillance Society: the Threat of 1984-style Techniques*, „The Futurist" 1985, nr 6; B. Simon, *The Return of Panopticism: Supervision, Subjection and the New Surveillance*, „Surveillance & Society" 2005, nr 3(1).

[5] K. Reis, *The Eavesdropping Society. Electronic Surveillance and Information Brokering*, „Patents, Copyrights, Trademarks, and Literary Property", June 2001.

Rozważania definicyjne

Łaciński źródłosłów tytułowego pojęcia „inwigilacja" – *invigilare* – dosłownie oznacza „czuwanie nad czymś", lecz takie rozumienie wydaje się zbyt oględne, ogólnikowe i przez to nietrafne. Pojęcie inwigilacji wykazuje pokrewieństwo z pojęciem podsłuchu, przy czym jest odeń szersze. Podsłuch stanowi szczególną formę inwigilacji: akustyczną z użyciem urządzeń elektronicznych i ukrytą, to jest bez wiedzy i autoryzacji poddanych podsłuchowi. Z technicznego punktu widzenia używa się do tego celu specjalistycznych elektronicznych urządzeń, jak na przykład miniaturowych mikrofonów, mikrofonów kierunkowych, zazwyczaj połączonych z cyfrowymi rejestratorami, wzmacniaczami lub z nadajnikami radiowymi[6]. Uprawnienia do stosowania inwigilacji (tzw. kontroli operacyjnej) mają w Polsce służby dyspozycyjne cywilne i wojskowe[7], a jak pokazuje praktyka detektywistyczna, podsłuchy stosują powszechnie podmioty drugiego sektora, a także osoby prywatne. Rodzime prawodawstwo konstytuuje trzy kategorie podsłuchu: procesowy, operacyjny i prywatny[8]. Pierwszy z wymienionych – podsłuch procesowy regulowany jest przez kodeks postępowania karnego (roz. 26, art. 237–242, dalej kpk)[9], w którym określone zostały warunki kontroli i utrwalania rozmów – telefonicznych i innych – lub przekazów informacji, w tym przesyłanych drogą elektroniczną. Kontrolę taką ordynuje prokurator za zgodą sądu albo uprzednią, albo w przypadkach niecierpiących zwłoki – następczą (co oznacza konieczność wystąpienia do sądu o zatwierdzenie samodzielnej decyzji w ściśle określonym terminie trzech dni). Może być ona prowadzona do trzech miesięcy, z możliwością jej przedłużenia na kolejne trzy miesiące.

Prawodawstwo precyzyjnie definiuje przypadki, gdy taki środek pozyskiwania dowodów jest legalny. Kontroli rozmów podlega wyłącznie

---

6 M. Pečenka i in., *Encyklopedia szpiegostwa*, przekł. K. Wojciechowski, Warszawa 1995, s. 202.

7 Są to: Agencja Bezpieczeństwa Wewnętrznego, Agencja Wywiadu, Centralne Biuro Antykorupcyjne, Krajowa Administracja Skarbowa, Policja, Służba Kontrwywiadu Wojskowego, Służba Ochrony Rządu, Służba Wywiadu Wojskowego, Straż Graniczna, Żandarmeria Wojskowa.

8 Por. M. Rogalski, *Kontrola i utrwalanie rozmów w procesie karnym*, „Prokuratura i Prawo" 2017, nr 6; K. Marszał, *Podsłuch w polskim procesie karnym* de lege lata *i* de lege ferenda, [w:] *Problemy nauk penalnych. Prace ofiarowane Pani Profesor Oktawii Górniok*, Katowice 1996, s. 343.

9 Ustawa z 6 czerwca 1996 r. – Kodeks postępowania karnego (t.j. Dz.U. z 2018 r., poz. 1987, ze zm.).

podejrzany, oskarżony lub pokrzywdzony, gdy istnieje domniemanie, iż mogą kontaktować się z nim dwie pierwsze kategorie wymienionych. Podsłuchowi może również podlegać osoba mogąca mieć związek ze sprawcą lub z grożącym przestępstwem. Kontrola jest możliwa wyłącznie wówczas, gdy toczące się postępowanie bądź uzasadniona obawa popełnienia nowego przestępstwa dotyczy między innymi zabójstwa, narażenia na niebezpieczeństwo powszechne lub sprowadzenia katastrofy, zamachu stanu, handlu ludźmi, stręczycielstwa, kuplerstwa i sutenerstwa, płatnej protekcji, łapownictwa, wymuszenia rozbójniczego lub rozboju, szpiegostwa lub odnosi się do mienia znacznej wartości (art. 237 § 3 kpk). Warto podkreślić, iż choćby zebrano dowody dotyczące innych przestępstw lub osób niż wymienione, to w świetle prawa procesowego takie dowody są bezużyteczne. Analogicznie skonstruowane są procedury podjęcia podsłuchu operacyjnego, z tym że reguluje go inny akt prawny – ustawa o policji (art. 19)[10]. Zarządzić takie działania jest władny sąd okręgowy na wniosek Komendanta Głównego Policji / komendanta wojewódzkiego, który uprzednio uzyskał w tej sprawie zgodę Prokuratora Generalnego lub odpowiednio prokuratora okręgowego. Czynności operacyjno-rozpoznawcze można stosować w celach zapobiegawczych, wykrywania, ustalania sprawców oraz uzyskania i utrwalenia dowodów ściganych z oskarżenia publicznego. Trzecia z kategorii to podsłuch prywatny. Jest on legalny, jeśli rejestracji podlega rozmowa, w której uczestniczy prowadzący rejestrację rozmowy. Nielegalność orzeka się, gdy posługując się urządzeniem podsłuchowym stosuje się je dla pozyskania informacji, do której uzyskania nie jest się uprawnionym (art. 267 § 3 kodeksu karnego)[11]. Wykorzystanie dowodów pochodzących z podsłuchu prywatnego nie jest w polskim prawodawstwie wprost zakazane, a więc nie obowiązuje doktryna „owoców zatrutego drzewa"[12]. Niektóre jednak orzeczenia odnoszą się do tej zasady krytycznie. Wskazuje się, że użyty podstęp godzi w konstytucyjną zasadę swobody i ochrony komunikowania się, a dowody tak zebrane nie powinny być dopuszczane w postępowaniu cywilnym[13]. Warto też podkreślić stosowanie działań wywiadowczych bez dbania o wyżej wymienione reguły przez wywiady obcych państw oraz kategorię wywiadu gospodarczego.

---

[10] Ustawa z 6 kwietnia 1990 r. o policji (t.. Dz.U. z 2019 r. poz. 161, ze zm.).

[11] Ustawa z 6 czerwca 1997 r. – Kodeks karny (t.j. Dz.U. z 2018 r. poz. 1600, ze zm.).

[12] Zasada ta po raz pierwszy pojawiła się w orzecznictwie Sądu Najwyższego USA (sprawa Nardone v. USA z 1939 r.).

[13] Takie stanowisko zajął np. Sąd Apelacyjny w Poznaniu (wyrok z 10.01.2008 r., I ACa 1057/07) oraz Sąd Apelacyjny w Warszawie (orzeczenie z 6.07.1999 r., I ACa 380/99).

Takie rozumienie pojęcie, choć specjalistyczne, wydaje się nazbyt wąskie na potrzeby niniejszego tekstu, jednocześnie nie uwzględniając aktualnego stanu rozwoju techniki – w szczególności urządzeń rejestrujących obraz czy przechwytujących komunikację elektroniczną (*via* Internet). Zatem na potrzeby niniejszego tekstu pojęcie inwigilacji będzie rozumiane jako przechwytywanie i utrwalanie obrazu, dźwięku lub treści dokumentów elektronicznych w sposób ukryty, to jest bez wiedzy i/lub zgody podmiotu kontrolowanego. Drugi człon związku frazeologicznego – „elektroniczna" odnosi się do technik rejestracji za pomocą zarówno specjalistycznych, jak i amatorskich urządzeń elektronicznych.

Technologiczny potencjał inwigilacji elektronicznej – przegląd autorski

Analiza technicznych aspektów inwigilacji elektronicznej oraz zasad i sposobów realizacji podsłuchów ma kluczowe znaczenie: uświadamia jej współczesne możliwości oraz łatwość naruszania prywatności. Potencjał środków inwigilacji trafnie określają ich rozmaite typologie. Elementarną kategoryzację urządzeń służących inwigilacji można utworzyć na podstawie schematu typowego systemu podsłuchowego, składającego się z trzech głównych elementów: punktu nadawczego / rejestratora, linii przesyłowej oraz punktu odbiorczego.

PUNKT NADAWCZY / REJESTRATOR

Punkt nadawczy można scharakteryzować z użyciem następujących parametrów: typ rejestrowanego przekazu, rodzaj zastosowanego kamuflażu, sposób i miejsce umieszczenia urządzenia w miejscu poddawanym inwigilacji, czas działania oraz sposób aktywacji urządzenia.

- Typ rejestrowanego przekazu – wydaje się najistotniejszym parametrem charakteryzującym punkt nadawczy. Wyróżnić można: urządzenia rejestrujące dźwięk otoczenia (podsłuchy klasyczne – i jako takie – najliczniejsze); urządzenia przechwytujące obraz bądź w formie statycznej (fotografie – nazywamy je fotopułapkami), bądź jako obraz dynamiczny (nagranie filmowe); urządzenia (lub programy) przechwytujące aktywność na urządzeniach komputerowych (uderzenia klawiszy, obraz, w tym metadane, jak na przykład przebieg aktywności użytkownika lub ruch w sieci); urządzenia (lub programy) przechwytujące komunikację telefoniczną (zarówno telefonii stacjonarnej i mobilnej). Rejestrowane są także inne typy

przekazu. Na przykład wszystkie oferowane na rynku fotopułapki i wiele kamer służących do dyskretnej inwigilacji mają wbudowany czujnik ruchu (PIR – *Passive Infra Red*) o zasięgu do parudziesięciu metrów. Z kolei większość nowszych kamer służących inwigilacji zaopatrzona jest w oświetlacz/reflektor podczerwieni pozwalający na prowadzenie rejestracji w nocy lub w innych oświetleniowo niesprzyjających warunkach (oświetlacze IR wbudowane w kamery zazwyczaj umożliwiają obserwację w odległości do kilku–kilkunastu metrów, a zewnętrzne, profesjonalne reflektory IR pozwalają na zwiększenie tego dystansu nawet do jednego kilometra). Dodatkowym kryterium podziału w ramach analizowanego aspektu jest wydzielenie dwóch grup urządzeń: profesjonalnych i pozostałych (amatorskich), przy czym różnice te mają istotne znaczenie dla jakości rejestrowanego przekazu.

- Rodzaj kamuflażu – może być rozumiany jako wizualna charakterystyka urządzenia nadawczego uniemożliwiająca lub utrudniająca odkrycie jego przeznaczenia. Urządzenia mogą wymagać – lub nie – dokonania samodzielnych zabiegów ich ukrywania. W pierwszym przypadku konieczne jest umieszczanie ich w takiej lokalizacji, która umożliwia skuteczną inwigilację, a jednocześnie stanowi zabezpieczenie urządzenia przed wykryciem. Parametrem wspomagającym, to jest utrudniającym wykrycie, jest miniaturyzacja urządzenia. Rozwiązania i możliwości wyboru zakamuflowanych urządzeń są nader bogate. Przykładowo na cywilnym rynku dostępne są liczne modele urządzeń podsłuchowych ukrytych przykładowo w: pendrivie, długopisie, zapalniczce, żarówce, czujniku dymu, listwie przepięciowej (przedłużacz i rozgałęźnik), sieciowej ładowarce samochodowej, karcie płatniczej, płycie CD wraz z opakowaniem. Takie urządzenia bardzo łatwo pozostawić w inwigilowanym pomieszczeniu, nie wzbudzając niczyich podejrzeń – na przykład podmieniając oryginalne przedmioty, darowując je lub pozostawiając, rzekomo z roztargnienia. Ceny takich urządzeń wahają się od kilkuset do kilku tysięcy złotych. Tworzone są również bardziej profesjonalne rozwiązania – podsłuchy umiejscawiane w odpowiednich statuetkach lub innych przedmiotach pamiątkowych wręczanych osobie przewidzianej do inwigilowania. Z kolei kamery wraz z mikrofonami w gotowych produktach inwigilacyjnych przykładowo kamuflowane są jako: piloty TV, latarki, zasilacze sieciowe, czujniki ruchu czy przeciwpożarowe, okulary i inne części garderoby (guziki, krawaty). Stosunkowo dobrze rozwinięty jest pod tym względem rynek urządzeń komputero-

wych[14] – dostępne są keyloggery (to jest urządzenia zapisujące znaki wybierane na klawiaturze przez osobę inwigilowaną) w postaci gotowych klawiatur lub adapterów (przejściówek USB/PS2), a także urządzeń sczytujących obraz z ekranu monitora montowanych pomiędzy nim a jednostką centralną jako adapter DVI/HDMI/VGA (urządzenia takie określa się mianem *frame-grabber*, jednym z przykładów jest kosztujący kilkaset złotych VideoGhost). Typ kamuflażu może być również rozpatrywany jako charakterystyka techniczna urządzenia rozumiana jako ekranowanie elektromagnetyczne urządzenia. Zasadniczo w istotny sposób potencjał ulotu informacji w wyniku elektromagnetycznej emisji ujawniającej redukuje fakt umieszczenia urządzenia w innym urządzeniu elektronicznym.

- Sposób ulokowania – to kolejny parametr urządzenia inwigilującego, a determinowany jest przez liczne zmienne: posiadanie kamuflażu lub jego brak, sposób zasilania, sposób przesyłania informacji do punktu odbiorczego, zasięg jego działania, konieczność jego zabrania po wykonaniu zadania, możliwość celowego lub przypadkowego wykrycia bądź zniszczenia. Pierwszą z przesłanek stanowi konieczność zapewnienia odpowiednio wysokiej jakości zbieranej informacji, co wymusza umieszczenie urządzenia w bezpośrednim sąsiedztwie obiektu inwigilowanego (na przykład urządzenie podsłuchowe powinno znajdować się w jak najbliższej odległości od miejsca rozmów, a urządzenie rejestrujące obraz nigdy bezpośrednio naprzeciwko okna). Po wtóre, urządzenia należy lokować z dala od źródeł zakłóceń dźwiękowych i/lub wizualnych. Po trzecie, nie należy umieszczać urządzeń przekazujących informację drogą radiową w miejscach ekranowanych (na przykład aluminiowych obudowach lub stalowych szafkach).
- Czas działania urządzenia – jest zależny przede wszystkim od możliwości zapewnionego zasilania, wtórnie zaś od zapotrzebowania energetycznego wyznaczanego przez typ rejestrowanego przekazu oraz sposobu komunikacji urządzenia nadawczego z urządzeniem odbiorczym. Zasadniczo wyróżniamy urządzenia z zasilaniem własnym oraz zasilaniem zewnętrznym. Pierwszy typ ma ograniczony czas działania, drugi zaś – potencjalnie nieograniczony. Zasilanie własne wykorzystuje baterie lub akumulatory umożliwiające nieprzerwane

[14] Istnieją liczne możliwości inwigilacji komputerów bez użycia sprzętu wymagającego dostarczenia i instalacji w miejscu znajdowania się komputera. Zagadnienie to jest jednak zbyt obszerne jak na wymagania objętościowe niniejszego tekstu.

działanie od około kilku do nawet kilkudziesięciu godzin (dla podsłuchów). Z kolei niektóre nowe fotopułapki mogą działać w stanie uśpienia bezobsługowo nawet kilka miesięcy (na przykład model LTL Acorn 6511WMG). Za bezpieczną w kontrinwigilacyjnej praktyce przyjmuje się powszechnie półroczną cezurę – po upływie tego czasu uznaje się baterię w potencjalnym podsłuchu za rozładowaną. Stąd wynika reguła biznesowego bezpieczeństwa umieszczania wszelkich nowych przedmiotów w przestrzeni neutralnej (najlepiej w pomieszczeniu, gdzie nie prowadzi się poufnych rozmów, jak gablota znajdująca się w korytarzu) na wskazany półroczny okres. Urządzenia mające zewnętrzne źródło zasilania można podzielić na takie, do których zasilanie zostało doprowadzone za pomocą kabla (rzadziej stosowane, wymagające bezpośredniego dostępu do miejsca w celu przeprowadzenia montażu), wbudowane w urządzenie mające zasilanie (na przykład keylogger montowany w obudowie klawiatury) lub takie, które same stanowią źródła zasilania (podsłuchy montowane w gniazdach elektrycznych – w elektroinstalacyjnych puszkach) oraz dołączane do źródła zasilania (jak pendrive USB z podsłuchem, rozgałęźnik z gniazdami elektrycznymi, żarówka zawierająca podsłuch, odświeżacz powietrza wpinany do gniazda elektrycznego).

- Sposób działania urządzenia – podstawowe rodzaje to urządzenia działające w sposób ciągły oraz te wzbudzane impulsami: dźwiękowymi (rozmowa), świetlnymi (włączenie światła) lub innymi (na ogół jest to ruch, jak na przykład w rejestratorze PV-TM10FHD zawierającym czujnik ruchu – całość ukryta została w wielofunkcyjnym zegarku).

LINIA PRZESYŁOWA

Przegląd wariantów przesyłania informacji z urządzenia nadawczego do urządzenia odbiorczego ukazuje różnorodność i możliwości współczesnych urządzeń inwigilujących dostępnych na cywilnym rynku. Aktualnie stosowane są następujące sposoby transmisji sygnału z urządzenia inwigilującego: radiowe (w tym GSM, Wi-Fi, Bluetooth), elektromagnetyczne (w tym podczerwień bliska i średnia), przewodowe (w podziale na medium własne, medium obce), zdalne (laserowe, sejsmiczne/kontaktowe, mikrofony kierunkowe), sieć Internet.

Transmisja radiowa ma największy zasięg, jednak jest stosunkowo łatwa do wykrycia, ponadto wymaga ustawienia punktu odbiorczego w promieniu zasięgu nadajnika. Obecnie najpopularniejsza jest transmisja zgodna ze standardem telefonii komórkowej (GSM oraz trans-

misji danych)[15]. Urządzenie zaopatrzone jest w moduł GSM, działając analogicznie jak telefon komórkowy, wymagając karty SIM i korzystając z infrastruktury operatorów sieci telefonii komórkowej. Transmisja oparta na GSM ma największy zasięg. W innym wypadku wymaga ustawienia punktu odbiorczego w zasięgu działania urządzenia, przeważnie do kilkuset metrów. Pierwotnie urządzenia służące inwigilacji wykorzystywały zakres UKF (VHF, *Very High Frequency*, fale ultrakrótkie), jest to zakres dedykowany między innymi radiofonii oraz rozmaitym systemom łączności lokalnej (policja, radiotaxi). Do transmisji używane mogą być również standardy sieci bezprzewodowej Wi-Fi (2,4 GHz, 5 GHz) oraz Bluetooth (2,4 GHz), jednak na niewielkie odległości – bez zastosowania repeaterów (urządzeń wzmacniających sygnał). W przypadku urządzeń o standardowych parametrach jest to kilkanaście–kilkadziesiąt metrów dla sieci Wi-Fi i maksymalnie kilkanaście metrów dla sieci Bluetooth.

Nadajniki podsłuchowych urządzeń inwigilujących pracują w różnych zakresach, co łączy się z licznymi wadami i zaletami. Najniższy zakres – od około 30 do 50 MHz – ma bardzo dobrą siłę przenikania w środowisku miejskim, jednak zarówno odbiornik, jak i nadajnik wymagają instalacji stosunkowo długich anten. Sygnał nadajnika może odbijać się od jonosfery na długich odległościach w porze nocnej i docierać do odległych miejsc. Zaletą nadajników pracujących w zakresie od 88 do 130 MHz jest fakt ich stosunkowo niewielkich kosztów, są one jednak słabo zabezpieczone przed wyciekiem danych, ich pasmo bowiem pokrywa się z komercyjnym pasmem UKF[16], mogą być zatem odbierane za pomocą zwykłego odbiornika radiowego, a także zagłuszane przez stacje radiowe. Z kolei zakres pracy urządzenia od 130 do 180 MHz sugeruje profesjonalne urządzenie, mniejsza jest szansa przypadkowego wykrycia, anteny są krótkie, a urządzenia charakteryzują się wystarczającą czułością. Są to jednak urządzenia stosunkowo drogie, ponadto częstotliwości te są wykorzystywane do komunikacji przez polskie służby dyspozycyjne. Zakres pracy od 330 do 360 MHz zapewnia bardzo duży stopień bezpieczeństwa, gdyż niewielkie jest wykorzystywanie tej długości fal. Anteny mogą być krótkie, brak jest zakłóceń przez inne nadajniki i naturalne źródła. Są to najdroższe urządzenia.

[15] M. Pavithran, *Eavesdropping on GSM*, „International Journal of Engineering Research in Computer Science and Engineering" 2016, nr 3(9).

[16] Od 88 do 107 MHz – zwykle urządzenia podsłuchowe dostrojone są właśnie do takiej częstotliwości.

Istnieje utajniona do lat 80. ubiegłego wieku metoda rozpraszania widma w systemach szerokopasmowych – FHSS, *frequency-hopping spread spectrum*, co oznacza dosłownie w języku polskim skakanie sygnału po częstotliwościach w kolejnych odstępach czasu, w dostępnym widmie (paśmie). Metoda ta utrudnia wykrycie transmisji inwigilującego urządzenia nadawczego.

Eksperymentalne prace nad transmisją radiową w urządzeniach inwigilujących nie ustają. Na przykład uczeni z izraelskiego Uniwersytetu Ben Guriona stworzyli oprogramowanie komputerowe (robaka, wektor ataku – port USB), które zamienia kartę graficzną zainfekowanego komputera w radio. Technika ta, nazwana AirHopper, pozwala wykradać dane z komputerów znajdujących się w izolowanych sieciach (albo w ogóle niepodpiętych do sieci komputerowych). Do jej użycia wystarczy telefon komórkowy zaopatrzony w odbiornik radiowy FM. Po zainfekowaniu docelowej maszyny AirHopper wpływa na działanie kart graficznych tak, aby generowały swoją pracą odpowiednio silne fale radiowe, które będą mogły zostać podjęte przez odbiornik[17]. Transmisja danych następuje poprzez modulację fal radiowych generowanych przez wpływanie na pracę karty graficznej[18]. Z kolei tajemnicza Equation Group odkryła w 2015 roku, w jaki sposób za pomocą sprzętu o wartości nieco ponad tysiąca złotych można wydobyć między innymi klucze kryptograficzne z zainfekowanego uprzednio komputera, wykorzystując monitoring emisji fal radiowych z procesora[19].

Transmisja przewodowa to w podstawowym rozumieniu fizyczne połączenie mikrofonu lub obiektywu kamery z urządzeniem. Oprócz

---

[17] Możliwości generowania fal radiowych przez karty graficzne komputerów rozważano w literaturze przedmiotu od początku XXI w., zob. M.G. Kuhn, *Compromising Emanations: Eavesdropping Risks of Computer Displays*, „Computer Laboratory" 2003, nr 577; B. Kania, *VGASIG. FM Radio Transmitter Using VGA Graphics Card*, 2009, https://bk.gnarf.org/creativity/vgasig/vgasig.pdf (dostęp: 8.01.2019).

[18] Więcej na ten temat: igH, *AirHopper – narzędzie do wykradania danych z odizolowanych, odciętych od sieci komputerów*, Niebezpiecznik, 3.11.2014, http://niebezpiecznik.pl/post/airhopper-narzedzie-do-wykradania-danych-z-odizolowanych-odcietych-od-sieci-komputerow/ (dostęp: 20.12.2018); M. Guri, G. Kedma, A. Kachlon, Y. Elovici, *AirHopper. Bridging the Air-Gap between Isolated Networks and Mobile Phones using Radio Frequencies*, 2014, https://www.wired.com/wp-content/uploads/2014/11/air-hopper-malware-final-e-141029143252-conversion-gate01.pdf (dostęp: 8.01.2019).

[19] M. Błoński, *Najbardziej zaawansowana operacja hakerska w historii*, 17.02.2015, http://kopalniawiedzy.pl/Equation-Group-haker-szpiegostwo-NSA,21930 (dostęp: 20.12.2018); także portal Ars Technica, https://arstechnica.com/tag/equation-group/ (dostęp: 20.12.2018).

standardowych przewodów dostępne są również rozwiązania wykorzystujące rozmaite substancje przewodzące, jak na przykład farby. W tej kategorii mieszczą się także urządzenia korzystające z kabli energetycznych traktowanych jako medium transmisji sygnałów. Działają one podobnie do adapterów PowerLine służących do przesyłania sygnału internetowego w sieci energetycznej. Nadajnik może być podłączony do linii zasilającej w interesującym podsłuchującego pomieszczeniu, sygnał jest przesyłany „po kablu" do odbiornika podłączonego gdzieś w budynku, na tej samej fazie. Takie rozwiązanie ma przede wszystkim tę zaletę, że jest bardzo trudne do wykrycia. Nie występuje tu żadna transmisja w paśmie podczerwieni ani ultradźwięków. Jednym z nowszych rozwiązań tego typu jest podsłuch klawiatury przez gniazdko sieci elektrycznej. Uderzenia w klawisze wywołują zróżnicowanie szumu w linii naziemnej. Rejestrator działa na około 15 metrów, a jego koszt to około dwa tysiące złotych[20].

Transmisja bezkontaktowa niewymagająca bezpośredniego dostępu do inwigilowanego pomieszczenia obejmuje kilka technik: odczyt transmisji elektromagnetycznej, podsłuch laserowy oraz zastosowanie mikrofonów sejsmicznych i kierunkowych.

- Transmisja elektromagnetyczna – paradoksalnie nie jest tu konieczne urządzenie nadawcze – jest nim samo inwigilowane urządzenie, gdyż wszystkie urządzenia elektroniczne podczas przetwarzania sygnału elektrycznego generują promieniowanie elektromagnetyczne niezależnie od tego, czy przetwarzanie sygnału ma charakter cyfrowy czy analogowy. Urządzenia cyfrowe emitują promieniowanie elektromagnetyczne związane z dwustanowym charakterem tego sygnału zwykle w wysokim zakresie częstotliwości. Na przykład w komputerach stacjonarnych, laptopach, a w mniejszym stopniu w tabletach najsilniejszymi źródłami promieniowania elektromagnetycznego są: monitory LED/LCD/CRT, klawiatury, magistrale PCI, kontrolery SCSI/IDE, łącza RS-232 oraz łącza USB[21]. Ponadto nośnikami takich

[20] M. Vuagnoux, S. Pasini, *Compromising Electromagnetic Emanations of Wired Keyboards, 2007–2009 Security and Cryptography Laboratory – LASEC/EPFL*, 2009, https://lasec.epfl.ch/keyboard/ (dostęp: 20.12.2018).

[21] Na przykład możliwości zastosowań ulotu elektromagnetycznego monitorów analizuje M.G. Kuhn, *Electromagnetic Eavesdropping Risks of Flat-Panel Displays*, 2004, http://www.cl.cam.ac.uk/~mgk25/pet2004-fpd.pdf (dostęp: 20.12.2018). Z kolei pracownicy Security and Cryptography Lab at Switzerland's EPFL w 2008 r. opracowali procedurę odczytu nieszyfrowanych danych z łącz USB z wykorzystaniem ulotu elektromagnetycznego: P. Miller, *Keyboard „Eavesdropping" Just Got Way Easier, Thanks to Electromagnetic Emanations*, Engadget, 20.10.2008, http://www.engadget.com/2008/10/20/

sygnałów elektromagnetycznych umożliwiającymi propagowanie ich na większe odległości są sieci energetyczne. Tematyka ulotu elektromagnetycznego jest przedmiotem licznych dociekań oraz publikacji[22]. Jako pierwszy na cywilnym rynku[23] możliwości odczytu ulotu elektromagnetycznego w celu inwigilacji odkrył holenderski uczony Wim van Eck w latach 80. ubiegłego wieku[24]. Demonstrował on w praktyce i publicznie możliwość podsłuchu widma elektromagnetycznego monitorów CRT komputerów znajdujących się w londyńskiej dzielnicy biurowej. Urządzenia te powszechnie nazywane są „receptorami van Ecka". W 1985 roku we współpracy z British Broadcasting Corporation (BBC) stworzył dokument filmowy (zaprezentowany w programie *Tommorow's World*), w którym z użyciem furgonetki wyposażonej w 10-metrowy maszt z anteną UKF z powodzeniem odczytywano treści pojawiające się na monitorach komputerów znajdujących się „w dużej odległości". Skuteczny zasięg receptorów to według W. van Ecka od dziesięciu do kilkudziesięciu metrów, a koszt wykonania urządzenia podsłuchowego przekraczał zaledwie o parędziesiąt dolarów cenę telewizora i anteny UKF. Użycie tej metody nie pozostawia śladów, zatem brak jest wiarygodnych danych dotyczących skali tego typu wycieków. Współczesne monitory LCD są co najmniej tak samo narażone na wyciek danych drogą emisji elektromagnetycznej

keyboard-eavesdropping-just-got-way-easier-thanks-to-electrom/?guccounter=1 (dostęp: 20.12.2018); H.-J. Choi i in., *Reconstruction of Leaked Signal From USB Keyboards*, 2016, http://www.researchgate.net/publication/309327769_Reconstruction_of_leaked_signal_from_USB_keyboards (dostęp: 20.12.2018)..2018]. Przegląd współczesnej aparatury oraz metodyka dokonywania ataków tego typu została wyłożona w: F. Elibol, U. Sarac, I. Erer, *Realistic Eavesdropping Attacks on Computer Displays with Low-Cost and Mobile Receiver System* [w:] *Proceedings of the 20th European Signal Processing Conference (EUSIPCO)*, 2012, http://www.eurasip.org/Proceedings/Eusipco/Eusipco2012/Conference/papers/1569583239.pdf (dostęp: 20.12.2018).

[22] Por. R. Frankland, *Side Channels, Compromising Emanations and Surveillance. Current and Future Technologies*, Londyn 2011, http://pdfs.semanticscholar.org/87a4/182d66ab649a35eff0267c5e3a73bb2a5087.pdf (dostęp: 20.12.2018).

[23] Stany Zjednoczone od lat 60. XX w. prowadzą program „TEMPEST" badający potencjał ulotu i podsłuchu urządzeń elektronicznych (komputerów i innych urządzeń komunikacyjnych), opracowując standardy urządzeń ekranowanych. Więcej na ten temat: *The Complete, Unofficial TEMPEST Information Page*, http://cryptome.org/tip/tempestintro.html (dostęp: 20.12.2018).

[24] W. van Eck, *Electromagnetic Radiation from Video Display Units: An Eavesdropping Risk?*, „North-Holland Computers & Security" 1985, nr 4 (artykuł dostępny obecnie na stronach organizacji Cryptome: http://cryptome.org/emr.pdf (dostęp: 20.12.2018)).

(w tym przez nieekranowane kable VGA)[25]. Opracowano profesjonalne urządzenia służące monitorowaniu wycieku elektromagnetycznego, między innymi do rejestrowania uderzeń klawiszy (dotyczy zarówno klawiatur przewodowych, jak i bezprzewodowych, USB/PS2, w komputerach stacjonarnych i laptopach). Taki analizator widma elektromagnetycznego przechwytujący sygnał działa na odległość do 20 metrów, a jego koszt to około 20 tys. złotych[26].

- Jeśli chodzi o urządzenia laserowe umożliwiające prowadzenie inwigilacji z dużej odległości, to zasada ich działania polega na odbieraniu odbitej wiązki promieniowania laserowego. Promień lasera pada na powierzchnię, na przykład okna, ulegając częściowemu odbiciu. Precyzja działania urządzenia jest bardzo wysoka i możliwe jest wychwycenie wibracji okien, które wywoływane są przez dźwięki wewnątrz pomieszczenia. Odbity promień jest zmodulowany tymi wibracjami i urządzenie jest zdolne na tej podstawie odtworzyć treść rozmowy prowadzonej w podsłuchiwanym pomieszczeniu. Do zalet takiego urządzenia należy duży zasięg pracy nawet do 400 metrów, niska wykrywalność, rozdzielenie modułów nadajnika i odbiornika pozwalające na prowadzenie posłuchu, gdy nie jest możliwe prostopadłe ustawienie (wiązka lasera może być kierowana na inne niż szyby powierzchnie – na przykład zastawę szklaną, butelki z wodą itd.), zintegrowany cyfrowy rejestrator pozwalający na automatyczne archiwizowanie pozyskanych informacji. Wadą jest niewątpliwie cena i choć są one coraz tańsze, to profesjonalny można obecnie kupić za około 200 tys. złotych. Za pomocą urządzeń laserowych można podsłuchiwać nie tylko rozmowy, lecz również pracę klawiatury odległych komputerów, możliwa jest bowiem rejestracja wibracji obudowy laptopa / klawiatury komputera. Każdy klawisz generuje unikatowy wzór wibracji, który można odczytać, jeśli skieruje się wiązkę laserową na miejsce urządzenia, które dobrze odbija światło (na przykład logotyp producenta w laptopie).
- Mikrofony sejsmiczne/kontaktowe – są to urządzenia pozwalające na prowadzenie inwigilacji osób znajdujących się w pomieszczeniu obok, przez ścianę o grubości nawet do 50 centymetrów. Umożliwiają one wzmocnienie dźwięku nawet do 20 tys. razy, pozwalając na swobodny podsłuch przez ściany betonowe, drewniane, metalowe, ceglane czy szklane. Podsłuch może być prowadzony zarówno przez

[25] M.G. Kuhn, *Electromagnetic Eavesdropping Risks…*

[26] M. Vuagnoux, S. Pasini, *Compromising Electromagnetic…*

ściany boczne, jak również sufity i podłogi. Bardzo czuły mikrofon wyłapuje każdy wstrząs przeszkody wywołany przez falę akustyczną. Urządzenie ma algorytm korekcji błędów: zastosowanie zaawansowanej filtracji pasmowej powoduje wzmocnienie sygnałów w paśmie ludzkiej mowy i niweluje niepożądane dźwięki. Mikrofony takie są urządzeniami pasywnymi – nie emitują żadnych fal radiowych, co utrudnia ich zdemaskowanie. Mogą być zasilane z akumulatora lub z sieci i wyposażone w rejestrator. Instalacja urządzenia może przebiegać również w inny sposób, jeśli możliwy jest dostęp do pomieszczenia sąsiadującego z tym, które pragniemy poddać inwigilacji, i jeśli wyposażeni jesteśmy w specjalistyczny mikrofon. W ścianie wierci się otwór, wprowadza weń szklaną rurkę jak najbliżej pomieszczenia inwigilowanego. W rurce instalowany jest mikrofon podłączony do rejestratora. Dzięki temu, że jest on oddalony od podsłuchiwanego pomieszczenia, jest praktycznie nie do wykrycia, nawet przy zastosowaniu wykrywacza złącz nieliniowych.

- Mikrofony kierunkowe umożliwiają podsłuch z dużej odległości w otwartym terenie. Dystans, na jakim urządzenia są efektywnie wykorzystane, dochodzi do 500 metrów w warunkach testowych. Realnie, w warunkach miejskich, jest to około 200 metrów.

Nieprzerwanie odbywa się poszukiwanie nowych dróg emisji dla linii przesyłowych. Jako egzemplifikacje wymienić można podsłuch cieplny/podczerwieni, pierwsze próby z transmisją ultradźwiękową, a także emisją dźwiękową.

Transmisja cieplna stanowi swoistą egzotykę urządzeń inwigilacyjnych i prawdopodobnie znajduje się w fazie eksperymentalnej. Po raz pierwszy w ogólnodostępnej literaturze naukowej została opisana w 2015 roku[27]. Urządzenie zostało nazwane BitWhisper. Wymaga ono uprzedniego zainfekowania inwigilowanego komputera, a także umieszczenia obok komputera inwigilującego[28]. Bezwładność cieplna uniemożliwia wysyłanie dużej liczby danych w krótkim czasie i jest to zaledwie osiem bitów w ciągu godziny. Zmiany temperatury maszyny są trudne do zauważenia, a urządzenie jest wyposażone w korekcję błędów spowodowanych czynnikami obocznymi (jest odporne na zakłócenia w postaci zmian temperatury w otoczeniu). Transmisja w podczerwieni jest trudna do wykrycia, jednak

[27] M. Guri, M. Monitz, Y. Mirski, Y. Elovici, *BitWhisper. Covert Signaling Channel between Air-Gapped Computers using Thermal Manipulations*, „Cryptography & Security" 2015.

[28] Eksperyment prowadzony był z maszynami ulokowanymi w odległości 40 centymetrów, jednak należy przyjąć, że urządzenia mogą się znajdować w większej odległości, lecz pomiędzy nimi nie może być żadnych nieprzejrzystych przeszkód.

nadajnik i odbiornik muszą pozostawać nieoddzielone żadnymi przeszkodami, które ograniczałyby widoczność. Tą drogą może być prowadzony również klasyczny podsłuch.

Zastosowanie dla transmisji ultradźwiękowej w urządzeniach inwigilujących znalazł Tristan Lawry. Dostrzegł on, iż ultradźwięki (powyżej 20 kHz) mają tę właściwość, że przenikają przez bariery skutecznie blokujące promieniowanie elektromagnetyczne (ekrany stalowe, klatka Faradaya), a także przez ściany. Jego konstrukt składa się z dwóch urządzeń o następujących funkcjach. Urządzenie nadawcze ma charakter pasywny – nie ma zasilania, musi jednak zostać dostarczone do inwigilowanego pomieszczenia. Z kolei urządzenie odbiorcze zasila urządzenie nadawcze wykorzystując ultradźwięki i zapewniając jednocześnie względnie szybki przesył informacji tą drogą (rzędu około 12 MB/s)[29]. Rejestrowanie uderzeń klawiszy dokonywane jest przez akcelerometr znajdującego się obok smartfona[30] (obecnie technologia ta umożliwia rozpoznawanie spójnego tekstu, lecz nie jest w pełni skuteczna przy odczytywaniu haseł, ma około 80-procentową skuteczność rozpoznawania znaków). W 2017 roku izraelscy badacze zaprezentowali koncepcję złośliwego oprogramowania nazwanego Fansmitter, które używa wentylatorów chłodzących komputery lub napędy dysków twardych do przesyłania skradzionych danych – w postaci fal dźwiękowych wytwarzanych przez te wentylatory[31].

URZĄDZENIA ODBIORCZE

Typologia urządzeń odbiorczych zamyka się w następujących dwóch klasach: urządzenia dedykowane odrębne od urządzeń nadawczych oraz scalone z nimi (jak w opisywanych wyżej przypadkach podsłuchu laserowego). Mają one znaczenie wtórne dla jakości zastosowanych urządzeń inwigilujących. Wśród urządzeń odrębnych możemy wyróżnić urządzenia dedykowane (stworzone wyłącznie na potrzeby odbierania transmisji z określonego urządzenia nadawczego) i uniwersalne (laptopy, tablety, smartfony), wymagające jedynie zainstalowania specjalnego oprogra-

[29] Autorzy podchodzą sceptycznie do podanego wolumenu przesyłu informacji. T. Lawry, *An Acoustic-Electric Bridge: Traversing Metal Barriers Using Ultrasound*, 2011, http://www.ttivanguard.com/ttivanguard_cfmfiles/pdf/miami11/miami11session7014.pdf (dostęp: 20.12.2018).

[30] [b.a.], *Hakerzy praw fizyki. Siły i fale w rękach włamywaczy*, http://mlodytechnik.pl/technika/29132-hakerzy-praw-fizyki (dostęp: 20.12.2018).

[31] M. Guri, Y. Solewicz, A. Daidakulov, Y. Elovici, *Fansmitter. Acoustic Data Exfiltration from (Speakerless) Air-Gapped Computers*, 2016, http://www.wired.com/wp-content/uploads/2016/06/Fansmitter-1.pdf (dostęp: 20.12.2018).

mowania. Producenci urządzeń oferują oprogramowanie dodatkowe, na przykład filtrujące i korygujące pozyskane zapisy.

Wybrane aspekty kontrinwigilacji

Rozwój metod wykrywania podsłuchów przypomina walkę pocisku z pancerzem. Urządzenia podsłuchowe wykorzystują różne sposoby przekazu zdobytych informacji: wysyłając sygnał w pasmie radiowym na częstotliwościach około 200–400 MHz do wykorzystujących pasmo GSM, Wi-Fi, przewody elektryczne, podczerwień, ultradźwięki czy nadawanie z przeskokiem częstotliwości. Istnieją również urządzenia pasywne – rejestratory.

Zasadniczo istnieją dwa sposoby zabezpieczenia się przed podsłuchem, stosowane jednocześnie. Po pierwsze, jest to właściwe zabezpieczenie osób (obejmujące wdrażanie do przestrzegania określonych procedur bezpieczeństwa, konfigurację oraz zabezpieczenie urządzeń osobistych, takich jak telefony i komputery osobiste) i pomieszczeń (ingerencja architektoniczna – ekranowanie, wdrożenie procedur bezpieczeństwa związanych z dostępem do pomieszczeń oraz instalacja urządzeń i oprogramowania zabezpieczającego pomieszczenia). Drugi filar kontrinwigilacji stanowi sprawdzanie pomieszczeń (cykliczne, regularne i/lub następcze dokonywane zazwyczaj po faktycznym lub domniemanym wystąpieniu incydentu bezpieczeństwa).

Wykrywanie podsłuchów w toku sprawdzania pomieszczeń wymaga z jednej strony przestrzegania przedstawionej niżej ścisłej procedury, a z drugiej strony – kreatywności i analitycznego myślenia. Procedura sprawdzenia obejmuje dwa następujące typy czynności: bezprzyrządowe oraz z użyciem specjalistycznej aparatury wykrywającej. Wymienione elementy są tylko pozornie odrębne, gdyż powinny one łączyć się wzajemnie i przenikać.

TECHNIKI BEZPRZYRZĄDOWE – WYWIAD ŹRÓDŁOWY I SPRAWDZENIE FIZYCZNE

Istotą technik bezprzyrządowych jest zogniskowanie się na czynniku ludzkim. Procedurę kontrinwigilacyjną nieodmiennie rozpoczyna wywiad źródłowy ze zleceniodawcą, a jeśli to możliwe, także z innymi osobami użytkującymi pomieszczenie. W toku wywiadu należy:

- ustalić, jakimi przesłankami kieruje się zleceniodawca, decydując o prowadzeniu działań kontrinwigilacyjnych (jakie wystąpiły incy-

denty naruszeń bezpieczeństwa informacyjnego: jakie informacje, kiedy, w jakich okolicznościach zostały ujawnione);
- przeanalizować luki bezpieczeństwa systemu zleceniodawcy (czynnik ludzki, czynnik proceduralny, czynnik techniczny);
- odtworzyć *status quo* – kto, na jakich warunkach i kiedy ma/miał dostęp do badanego pomieszczenia oraz jego otoczenia;
- zapoznać się z podejrzeniami zleceniodawcy odnośnie do osób lub organizacji zlecającej inwigilację.

Powyższe informacje w dużej mierze wyznaczają wektory poszukiwań, pozwalając między innymi na ewaluację potencjalnych środków przeznaczonych na inwigilację, a co za tym idzie – typ i sposób umieszczenia urządzeń inwigilacyjnych. Czynność wywiadu źródłowego nigdy, z przyczyn oczywistych, nie powinna odbywać się w miejscu poddawanym procedurze sprawdzenia. Warto podkreślić, że nie należy nigdy ograniczać się wyłącznie do wektorów poszukiwań wynikających z wywiadu.

Następnym elementem procedury jest wizualne i fizyczne sprawdzenie pomieszczenia. Jest ono, jak w klasycznym przeszukaniu, dzielone na sektory, które są następnie systematycznie rewidowane. Przesłanką podziału jest topografia pomieszczenia, ale także wiedza o typowych miejscach umieszczania urządzeń inwigilacyjnych. Inspekcji należy również poddać tzw. tunele akustyczne – umożliwiające ulot dźwięku, to jest sufity podwieszane, wspólne kanały centralnego ogrzewania itd. Inspekcja powinna obejmować umeblowanie pomieszczenia, wbudowane (architektoniczne) elementy wyposażenia, wszystkie urządzenia, okablowanie. Należy zwracać uwagę na potencjalne wskaźniki podejmowanej inwigilacji: naruszone gniazda śrub, plomby i inne elementy montażowe z widocznymi śladami ingerencji, starty kurz lub zabrudzone powierzchnie, świeże ślady farby lub szpachlówki, kawałki przewodów czy taśm, otwory itd. Tego typu sprawdzenia dokonuje się z użyciem latarki stanowiącej silne źródło światła, mającej możliwość przełączania w tryb ultrafioletu (UV)[32]. W toku sprawdzenia należy również poczynić ustalenia dotyczące pomieszczeń przylegających do badanego: typ, sposób użytkowania i zasady dostępu, a także ocenić możliwości prowadzenia inwigilacji z zewnątrz (podsłuch laserowy) oraz ustalić ze zleceniodawcą,

[32] Światło o wysokich częstotliwościach (nadfiolet) pozwala na ujawnienie śladów ingerencji w postaci świeżych farb, kitów, szpachlówek, a także rozmaitych zmian w strukturze, ukrytych przewodów oraz innych wskaźników ingerencji niewidocznych w świetle widzialnym.

czy i jakie w sprawdzanym pomieszczeniu pojawiły się nowe/nierozpoznawane przezeń elementy wyposażenia.

TECHNIKI Z UŻYCIEM SPECJALISTYCZNEJ APARATURY WYKRYWAJĄCEJ

Choć do wykrycia wszystkich rodzajów urządzeń inwigilacyjnych służy cała gama różnorakich przyrządów pomiarowych, to można wyróżnić dwie zasadnicze grupy aparatur: wykrywające transmisję linii przesyłowych oraz wykrywające obecność punktów nadawczych.

Wykrywanie transmisji linii przesyłowych obejmuje sprawdzenie pasma radiowego, propagacji w zakresie podczerwieni IR oraz sprawdzenie linii zasilających i innych przewodów. Do wykrycia transmisji mogą służyć wielofunkcyjne analizatory. Jednym z nich jest na przykład Piranha ST-031M. Koszt takiego urządzenia to około 30 tys. złotych. Umożliwia analizę emisji radiowej w zakresie częstotliwości od 140 MHz do 12 GHz, pozwalając na dynamiczne wyszukiwanie anomalii. Urządzenie wykrywa i identyfikuje sygnały GSM, UMTS, LTE i Wi-Fi. Zadaniem tego typu analizatorów jest ustalenie najsilniejszych sygnałów radiowych w pomieszczeniu. Zasięg ich działania wynosi od dziesięciu centymetrów do około jednego metra od źródła sygnału, co oznacza, że wyszukiwanie związane jest z systematycznym obchodem pomieszczenia, a w szczególności z oceną najbardziej podejrzanych miejsc. Urządzenie Piranha ST-031M umożliwia również detekcję w pasmie podczerwieni, ultrafioletu oraz ultradźwięków. Dodatkowo urządzenie umożliwia sprawdzenie linii zasilających oraz wszelkiego rodzaju innych przewodów (choć urządzeniem dedykowanym do sprawdzania przewodów jest analizator ST-300 SPIDER).

Z kolei wykrywanie obecności punktów nadawczych odbywa się z użyciem wykrywacza złącz nieliniowych, anteny elektromagnetycznej oraz wykrywaczy kamer.

Do wykrywania urządzeń pasywnych, w tym nieaktywnych (bez zasilania) w momencie sprawdzania, na przykład dyktafonów, służą wykrywacze złącz nieliniowych, czyli układów półprzewodnikowych. Najlepiej sprawdzają się one przy przeszukiwaniu miejsc, w których brak jest elektroniki i być jej nie powinno: drewnianych mebli, rzeźb, ścianek działowych. Zasada działania tych urządzeń jest homologiczna do zasady działania radaru. Urządzenie wysyła fale o określonej długości, następnie analizuje sygnał odbity, a dokładniej – składowe harmoniczne sygnału wejściowego (harmoniczna jest definiowana jako składowa przebiegu o częstotliwości będącej całkowitą krotnością częstotliwości podstawo-

wej). Pierwsza harmoniczna jest sygnałem o częstotliwości równej częstotliwości analizowanego sygnału okresowego (i z reguły pochodzi od złącz liniowych), a częstotliwości kolejnych składowych harmonicznych są wielokrotnościami tej częstotliwości[33]. Pojawienie się na wyjściu układu wyższych, parzystych harmonicznych przy pobudzaniu składową podstawową świadczy o nieliniowości tego układu (zniekształcenia nieliniowe), co sugeruje obecność układów półprzewodnikowych (diody, tranzystory, układy scalone). Największym problemem podczas używania tych urządzeń jest fakt, że tak zwane złącze m-o-m (*metal-oxide-metal*), czyli zwykła korozja, daje sygnał pseudoidentyczny jak urządzenie elektroniczne[34]. Może więc dojść do sytuacji, w której znajdujący się pod tynkiem kawałek zardzewiałego gwoździa zostanie zidentyfikowany jako urządzenie elektroniczne. Dlatego najlepszym rozwiązaniem jest stosowanie wykrywacza złącz nieliniowych, który potrafi odróżnić złącze m-o-m od rzeczywistego złącza nieliniowego. Do takich urządzeń należą między innymi wykrywacze z serii Cayman (ST-400, ST-401, ST-402, ST-403). Według analogicznej zasady działają anteny elektromagnetyczne – zaopatrzony w taki dodatek jest wykrywacz Piranha ST-031M.

Wykrywanie kamer może odbywać się między innymi z zastosowaniem urządzeń, których zasadą działania jest emitowanie światła lasera i obserwowanie przez specjalny okular światła odbitego od obiektywu. Urządzenie takie generuje promieniowanie podczerwone (na przykład z zastosowaniem systemu diod LED IR), okular urządzenia zaopatrzony jest w filtry przepuszczające tylko spolaryzowane światło, a takie właśnie jest odbijane od matrycy kamery. Przykładem takiego urządzenia jest Optic-2. Znacznie prostszy sposób stanowi wykorzystanie światła widzialnego – zwykłej latarki. Odpowiednio nakierowana wiązka światła sprawi, że odbije się ono od soczewki i matrycy kamery, ujawniając ją.

Przeprowadzona analiza potencjału urządzeń inwigilujących i kontrinwigilujących uwidacznia przewagę tych pierwszych – są one tańsze, a także liczne, co utrudnia, a niekiedy uniemożliwia przewidywanie wektorów ataku. Urządzenia kontrinwigilujące są kosztowne, wymagają wykwalifikowanego personelu i nie zawsze są całkowicie skuteczne (prawdopodobieństwo wykrycia podsłuchu wzrasta, jeśli zwiększymy czas wyszukiwania).

---

[33] Złącze nieliniowe daje parzyste harmoniczne, to jest drugą, czwartą i szóstą, liniowe zaś – harmoniczne nieparzyste – pierwszą, trzecią, piątą.

[34] Jest to sygnał podobny, jednak niejednorodny i niestabilny w czasie, szczególnie przy zakłóceniach mechanicznych.

Próba diagnozy

Dynamiczny rozwój technologiczny oraz liczne lokalne i globalne wydarzenia związane z inwigilacją (tzw. afery podsłuchowe) umożliwiają sformułowanie wniosków diagnostycznych oraz prognoz w zakresie antycypowanych kierunków rozwoju zjawiska. Argumentacja prezentowana jest na zasadzie kontrapunktu, przesłanką takiego zabiegu jest niemożność przeprowadzenia systematycznych badań. Diagnostyczno-prognostyczne refleksje formułowane są zatem na podstawie zestawienia historycznego i współczesnego oglądu danego aspektu. Dla uporządkowania wywodu opatrzono je następującymi etykietami: eskalacja, profesjonalizacja, instytucjonalizacja i normalizacja.

Najwyraźniej dostrzegalne zjawisko stanowi eskalacja zjawiska inwigilacji, polegająca na coraz łatwiejszym dostępie do urządzeń inwigilacyjnych oraz na stałym poszerzaniu się zakresu i treści informacji, jakie można pozyskiwać z ich użyciem. Rozwój technologii zwiększa zakres możliwej do pozyskania informacji co do jej ilości, a przede wszystkim jej rodzajów; aktualnie możliwy jest podsłuch rozmów, podgląd obrazu, ale także bieżący odsłuch i rejestracja rozmów telefonicznych, konwersacji poprzez komunikatory i czaty, poczty elektronicznej, innej prywatnej korespondencji w mediach społecznościowych, rejestracja miejsca pobytu oraz trasy przemieszczania się. Wejście w posiadanie informacji takich jak wymienione nie nastręcza obecnie trudności i nie wymaga istotnych nakładów finansowych. Tytułem przykładu – odczytywanie cudzej korespondencji prowadzonej z użyciem poczty elektronicznej jest możliwe z użyciem darmowego oprogramowania Social Engineering Toolkit dostępnego w nieodpłatnym systemie operacyjnym Kali Linux[35]. Przeszkolenie w użyciu tych narzędzi zajmuje około dwóch godzin, by przełamać typowe zabezpieczenia komputera potencjalnej osoby inwigilowanej.

Warto na zasadzie kontrapunktu wskazać, jak znaczne nakłady finansowe, czasowe i organizacyjne były niezbędne, by uzyskać informacje w przeszłości. Sun Tzu, autor *Sztuki wojny*, apoteozował w tym kontekście rolę wywiadowcy: „Spośród ludzi armii nikt nie jest tak bliski dowódcy jak poufny agent, ani też w nagrodach generał nie jest tak hojny wobec nikogo, jak wobec zaufanego informatora. Z wojskowych tajemnic żadne nie są tak dobrze strzeżone, jak te dotyczące tajnych planów”[36].

---

[35] https://www.kali.org (dostęp: 20.12.2018).

[36] Sun Tzu, *Sztuka wojny*, s. 81, https://www.lazarski.pl/fileadmin/user_upload/dokumenty/student/Sun_Tzu_sztuka_wojny.pdf (dostęp: 20.12.2018).

Pierwotnie podsłuch prowadzono na niewielkie odległości i zazwyczaj wymagał on ingerencji w konstrukcję budynku już na etapie jego wznoszenia, w postaci na przykład wbudowanych kanałów akustycznych, zapewniających gospodarzom podsłuch oddalonych pomieszczeń. W Malborku znajduje się miejsce, w którym Wielki Mistrz Zakonu mógł słuchać rozmów odbywających się w pokojach gościnnych[37]. Podobne miejsce zwane zakrystią akustyczną znajduje się w Archikatedrze pod wezwaniem Jana Chrzciciela i Jana Ewangelisty w Lublinie przy ulicy Królewskiej 14. Łukowate sklepienie pomieszczenia umożliwia propagację słów wypowiedzianych szeptem w jednym z jego rogów do rogu przeciwległego. Niemniej pouczająca jest współczesna egzemplifikacja, ukazująca, iż całkiem niedawno wdrożenie systemu podsłuchowego wymagało wiele pomysłowości, tyleż umiejętności technicznych, a także nakładów finansowych i organizacyjnych. Takimi cechami i możliwościami wykazali się sprawcy najgłośniejszej afery podsłuchowej z lat 40. XX wieku. Właścicielka ekskluzywnego domu publicznym w Berlinie, znanego pod nazwą „Salon Kitty", Katharina Zammit została zmuszona przez Gestapo szantażem do współpracy. Autorem pomysłu był ówczesny szef Policji Bezpieczeństwa (Sipo) Reinhard Heydrich. Stosowane wówczas mikrofony były duże, miały niewielki zasięg i wymagały podłączenia kabli. Zatem w każdym podsłuchiwanym pomieszczeniu należało przeprowadzić wysokonakładowe prace montażowe – umieścić po kilka mikrofonów (miejscem ich instalacji było głównie oświetlenie górne) oraz aparatów fotograficznych. W tym celu wynajęto ostatnie piętro budynku i Gestapo pod przykrywką prowadzonego remontu zainstalowało mikrofony w dogodnych miejscach pokoi. Kable z ostatniego piętra biegły do piwnicy, gdzie technicy rejestrowali rozmowy na magnetofonach. Przedsięwzięcie wymagało również zatrudnienia personelu (poddanego 7-tygodniowemu przeszkoleniu), w tym około 20 dodatkowych prostytutek. Plan wdrożono w życie w marcu 1940 roku. Podsłuchiwano wszystkich bywalców domu publicznego, w tym zagranicznych dyplomatów oraz niemieckich funkcjonariuszy wysokiego szczebla, w tym generała SS Seppa Dietricha, dowódcę dywizji Leibstandarte Adolf Hitler. System został przypadkowo wykryty przez brytyjskiego funkcjonariusza wywiadu Rogera Wilsona, działającego jako rumuński dyplomata Ljubo Kolczew. Do wykrytej wiązki kabli podłączył swoje, które zostały poprowadzone

[37] Notabene podczas ważnych rozmów w pomieszczeniach śpiewał chór, tworząc jeden z pierwszych systemów zagłuszających.

do sąsiedniego budynku, gdzie do operacji dołączyli pracownicy wywiadu brytyjskiego[38].

W taki oto zasobochłonny sposób prowadzano dawniej operacje inwigilacyjne, do czasu wystarczającej miniaturyzacji urządzeń podsłuchowych oraz zapewnienia im bezprzewodowej komunikacji na znaczne odległości. Przeważnie zainstalowanie podsłuchu wiązało się z wykonaniem instalacji w mieszkaniu figuranta lub w wynajmowanym mieszkaniu sąsiadującym. Była to operacja skomplikowana, do której wykonania potrzebne było zaangażowanie wielu osób, w tym stworzenie właściwej legendy dla podejmowanych i widocznych dla postronnych osób działań oraz zamaskowanie śladów po instalacji (niejednokrotnie istniała konieczność, by do zamaskowania śladów wysyłać pracownika o zdolnościach artystycznych, aby dokładnie dobrał właściwą kolorystykę farb).

Wskutek miniaturyzacji instalacja urządzeń inwigilacyjnych jest obecnie nieporównywalnie tańsza, prostsza i bardziej efektywna. Nie są konieczni wykwalifikowani pracownicy, a same urządzenia można często pozyskać za zdecydowanie niewielką kwotę. Taka sytuacja powoduje, że inwigilacja stała się szeroko stosowaną metodą zdobywania informacji już nie tylko gospodarczych, wojskowych i politycznych, jak w przypadkach historycznych, ale i prywatnych (służy na przykład pozyskiwaniu dowodów w sprawach rozwodowych, w których jeszcze kilka lat temu stosowanie podsłuchu byłoby nieopłacalne).

Lista odkrytych i nagłośnionych przypadków jest długa, trudno tu o zachowanie jakiejkolwiek systematyczności i wyczerpywalności. Tytułem przykładu – w 2016 roku w gabinecie przewodniczącego Rady Miejskiej Ostrowa Wielkopolskiego wykryto urządzenie podsłuchowe – umieszczono je pod fotelem urzędnika[39]. Podobne urządzenie odnaleziono w urzędzie miejskim w Karpaczu[40]. Z kolei w Strzegomiu ujawniono w sejfie Urzędu Miejskiego kilka mikrofonów podsłuchowych, odbiornik typu skaner szerokopasmowy, urządzenie odsłuchowe oraz prosty skaner radiowy mogący służyć do wykrywania podsłuchów, jak też odsłuchiwania

---

[38] Więcej na temat tej operacji: T. Crowdy, *Historia szpiegostwa i agentury*, przekł. J. Mikołajczyk, Warszawa 2010, s. 260; B. Wołoszański, *Wojna, miłość, zdrada*, Warszawa 2010, s. 80.

[39] Past, *„Urządzenie podsłuchowe" w Urzędzie Miejskim w Ostrowie Wielkopolskim. Sprawę bada policja*, Gazeta.pl, 14.07.2016, http://wiadomosci.gazeta.pl/wiadomosci/7,114883,20398093,urzadzenie-podsluchowe-w-urzedzie-miejskim-w-ostrowie-wielkopolskim.html (dostęp: 20.12.2018).

[40] P. Kołpajew, *Podsłuch w urzędzie w Karpaczu*, 12.09.2014, https://wroclaw.tvp.pl/16818908/podsluch-w-urzedzie-w-karpaczu#! (dostęp: 20.12.2018).

ogólnie dostępnych (niekodowanych) przekazów[41]. Przypadki te wskazują, jak powszechne jest stosowanie urządzeń inwigilacyjnych, co często czyni się w sposób nie tylko nieprofesjonalny, ale również nieprzemyślany.

Kolejna istotną cechą dotyczącą opisywanego zjawiska jest wyraźna profesjonalizacja urządzeń i usług inwigilacyjnych. Jest to zjawisko równoległe do opisanego wyżej upowszechniania urządzeń podsłuchowych. Wartościowe poznawczo wydaje się zestawienie dwóch przykładów – historycznego i współczesnego, wydobywających na zasadzie ostrego kontrastu postęp, jaki dokonał się w zakresie profesjonalizacji urządzeń i usług inwigilacyjnych. Jeszcze w połowie XX wieku nie istniał wyspecjalizowany rynek takich urządzeń. Wiele z nich tworzyli wynalazcy zatrudniani przez władze państwowe, czyniąc to na potrzeby konkretnych operacji. Szeroko dyskutowanym w literaturze przedmiotu jest rosyjski system podsłuchowy „Złotousty", pierwsze pasywne urządzenie podsłuchowe[42]. Jego autorem, a ściślej – wynalazcą, był genialny uczony Lew Termen. Autor licznych wynalazków – między innymi instrumentu muzycznego nazwanego od jego nazwiska theremin[43]. Pierwsze urządzenie podsłuchowe zostało przez L. Termena opracowane prawdopodobnie w 1943 roku, w tzw. szaraszce – radzieckim zamkniętym ośrodku naukowo-badawczym dla więźniów podległym NKWD[44]. Podsłuch – już po odkryciu przez służby amerykańskie w 1952 roku – został określony mianem *The Thing*. Po raz pierwszy wynalazek zastosowano w 1945 roku. Z okazji amerykańskiego Dnia Niepodległości wypadającego 4 lipca grupa pionierów wręczyła jako prezent ambasadorowi Stanów Zjednoczonych w Moskwie Williamowi A. Harrimanowi tzw. Wielką Pieczęć, czyli drewnianą płaskorzeźbę przedstawiającą bielika amerykańskiego z piersią zasłoniętą tarczą w kolorach amerykańskiej flagi, trzymającego 13 strzał i gałązkę oliwną. Anegdota głosi, iż gdy ambasador zadał retoryczne pytanie, gdzie powiesić rzeźbę, odpowiedziano mu, że najlepiej w jego gabinecie – by dopiec Anglikom. Naturalnie właściwe służby dokonały starannego sprawdzenia

---

41 M. Moczulska, *Strzegom: urzędnicy na podsłuchu*, „Gazeta Wrocławska", 26.04.2011, https://gazetawroclawska.pl/strzegom-urzednicy-na-podsluchu/ar/396470 (dostęp: 20.12.2018).

42 Obszerna analiza tego przypadku wraz z wyjątkami z materiałów źródłowych znajduje się w: K.D. Murray, *The Great Seal Bug*, http://counterespionage.com/great-seal-bug-part-1/ (dostęp: 20.12.2018).

43 Więcej na temat wynalazcy: P. Nikitin, *Leon Theremin (Lev Termen)*, „IEEE Antennas and Propagation Magazine" 2012, nr 54(5).

44 W marcu 1939 r. L. Termen został aresztowany i skazany za rzekome szpiegostwo i działalność wywrotową na osiem lat ciężkich robót.

prezentu, był on podobno również poddany kwarantannie – czyli przetrzymany przez kilka tygodni poza głównymi pomieszczeniami ambasady. W końcu jednak umieszczono go, zgodnie z sugestią darczyńców, w gabinecie ambasadora nad jego biurkiem. Dopiero po upływie siedmiu lat od tego wydarzenia Amerykanie przypadkowo zorientowali się, że na terenie ambasady jest podsłuch – operatorzy radiowi, nasłuchując rosyjskiej aktywności radiowej, natrafili na głos własnego ambasadora. Korzystając z trwającego remontu ambasady, rozpoczęto poszukiwania z pomocą specjalistów. Po wielu dniach stwierdzono, że sygnał z „pluskwy" dobiega ze ściany za biurkiem ambasadora. Zdjęto płaskorzeźbę i rozkuto ścianę. Nic nie znaleziono. Źródłem sygnału okazała się sama płaskorzeźba, ale po pierwsze była sprawdzana, po drugie wisiała już siedem lat, a nie była podłączona do żadnego źródła zasilania, więc wydawało się to niemożliwe. Mimo to została rozmontowana i w środku znaleziono zaledwie zwykły drut oraz membranę. Amerykańscy specjaliści nie potrafili wyjaśnić zasady działania podsłuchu i po kilku miesiącach zdecydowano się przekazać urządzenie w ręce brytyjskiego eksperta MI5 Petera Wrighta[45], któremu rozpracowanie urządzenia zajęło dwa miesiące. Okazało się, iż z domu naprzeciwko amerykańskiej ambasady wysyłano radiową wiązkę o częstotliwości 800 MHz, celując w godło w gabinecie ambasadora. Gdy ktoś rozmawiał, fale głosowe powodowały wibracje membrany. Drgania przenoszone do anteny długości niespełna 23 centymetrów zmieniały trafiający w nią sygnał radiowy, odbijany do stacji nasłuchowej. Urządzenie nie potrzebowało zasilania, tak jak lustro nie potrzebuje energii, żeby odbijać światło.

Podobny system zastosowali Rosjanie w latach 1976–1984. W maszynach do pisania IBM Selectric, w które była wyposażona ambasada amerykańska w Moskwie, umieścili pręty wyposażone w magnetometry. Mierzyły one zmiany pola magnetycznego związane z charakterystycznym dla każdego znaku ustawieniem metalowych ramion pozycjonujących głowicę drukującą. Odczyty były zamieniane na sygnał radiowy i przesyłane do stacji nasłuchowej w pakietach po osiem znaków. Amerykańska Agencja Bezpieczeństwa Narodowego odkryła to dopiero po przeprowadzeniu badań w Stanach Zjednoczonych[46].

---

[45] Polskiemu (i nie tylko) czytelnikowi znany z pracy: P. Wright, *Łowca szpiegów*, przekł. W. Kalinowski, M. Możejko, 1991.

[46] Dokonano tego w ramach operacji Gunman. Więcej na temat tej operacji oraz tzw. Selectric Bug na stronie CryptoMuseum: *IBM Selectric Bug. Operation GUNMAN – How the Soviets Bugged IBM Typewriters*, https://www.cryptomuseum.com/covert/bugs/selectric/index.htm (dostęp: 20.12.2018).

Sztandarowym przykładem rozwoju rynku produktów służących monitorowaniu komunikacji pojedynczych osób jest włoskie przedsiębiorstwo informatyczne Hacking Team powstałe w 2003 roku, które założyli David Vincenzetti i Valeriano Bedeschi[47]. Swoje główne zadanie przedsiębiorstwo definiuje jako wytwarzanie systemów zdalnej kontroli (*Remote Control Systems*), a produkty przeznaczone są wyłącznie dla podmiotów takich jak instytucje państwowe i niektóre korporacje[48]. W 2015 roku z witryny firmy wykradziono około 400 GB tajnych informacji[49]. Ujawniono poufną korespondencję między programistami i przedstawicielami ich klientów, hasła dostępowe do systemów tworzonych przez Hacking Team oraz umowy. Na ich podstawie został opracowany raport ujawniający, iż z usług firmy skorzystało ponad 70 agencji rządowych całego świata (w tym polskie Centralne Biuro Antykorupcyjne), a suma kwot wszystkich zakupów wynosi około 30 mln euro. Oferta Hacking Team obejmowała między innymi: niezauważalne dla zainfekowanego oprogramowaniem HT użytkownika przechwytywanie wiadomości e-mail, rejestrowanie uderzeń klawiszy (*keylogging*), przeszukiwanie zbiorów danych na komputerze, zapisywanie czatów głosowych i wideo, aktywowanie kamer i mikrofonów w komputerze lub smartfonie, śledzenie lokalizacji *via* GPS, ekstrahowanie haseł sieci Wi-Fi oraz możliwość śledzenia w sieci transakcji z użyciem kryptowaluty Bitcoin. Ponadto Hacking Team intensywnie pracował nad nowymi rozwiązaniami inwigilacyjnymi, między innymi prowadzono badania nad dronami zdolnymi do atakowania sieci Wi-Fi. W raporcie międzynarodowej organizacji pozarządowej globalnie monitorującej przejawy ograniczeń wolności słowa – Reporterzy bez Granic (ang. Reporters Without Borders) wymieniono obok Hacking Team jeszcze cztery inne organizacje (Amesys, Blue Coat, Gamma, Trovicor), nazywając je pięcioma największymi wrogami Internetu i piętnując tym samym wytwarzanie przez nie na potrzeby służb dyspozycyjnych

[47] Więcej na temat oferty Hacking Team: S. Špaček, P. Celeda, M. Drašar, M. Vizváry, *Analyzing an Off-the-Shelf Surveillance Software. Hacking Team Case Study*, 2.06.2017, http://spi.unob.cz/papers/spi2017.html https://is.muni.cz/repo/1382042/2017-SPI-hacking-team-case-study-presentation.pdf (dostęp: 20.12.2018)] (także na stronie przedsiębiorstwa http://www.hackingteam.it (dostęp: 20.12.2018).

[48] A. Batey, *The Spies Behind Your Screen*, „Telegraph", 24.10.2011, https://www.telegraph.co.uk/technology/8899353/The-spies-behind-your-screen.html (dostęp: 20.12.2018).

[49] Wykradzione informacje (część z nich zaopatrzono w wygodną wyszukiwarkę) znajdują się na stronach Transparency Toolkit: *Hacking Team Archive*, https://transparencytoolkit.org/project/hacking-team-archive/ (dostęp: 20.12.2018).

państw oprogramowania naruszającego prywatność komunikacji i zbiorów danych[50].

Trzeci aspekt to dostrzegalny wyraźnie trend instytucjonalizacji zjawiska inwigilacji, czyli powstawania wyspecjalizowanych agend i programów działania służących inwigilacji. Doskonałym tego przykładem są Stany Zjednoczone. Po II wojnie światowej rozpoczęto prace nad miniaturyzacją samych urządzeń podsłuchowych i ich źródeł zasilania oraz nad opracowywaniem innych niż „klasyczne" metod podsłuchowych. W tym celu w 1951 roku utworzono w ramach Centralnej Agencji Wywiadowczej (CIA) Technical Services Staff (od 1960 roku Technical Services Division) zajmującą się opracowywaniem i badaniem urządzeń inwigilacyjnych. W początkowym okresie liczyła ona zaledwie 50 pracowników, jednak już po dwóch latach działania liczba pracowników zwiększyła się pięciokrotnie. Intensywne prace, jak się współcześnie okazało, były kontynuowane również na innej płaszczyźnie: masowego podsłuchu połączeń telefonicznych.

Jednym z budzących największe kontrowersje przedsięwzięć jest projekt National Security Agency (NSA) noszący nazwę Echelon[51]. Prace nad nim trwały już od 1947 roku, jednak pierwsze wiarygodne dotyczące go informacje trafiły do opinii publicznej dopiero w 1988 roku. Stało się to za sprawą Margaret Newsham, administratorki systemów komputerowych w stacji nasłuchowej Echelon Menwith Hill położonej w hrabstwie Yorkshire[52]. Z systemu tego korzystają obecnie agencje wywiadowcze nie tylko Stanów Zjednoczonych Ameryki Północnej, ale także Australii, Kanady, Nowej Zelandii oraz Wielkiej Brytanii[53]. System Echelon składa się z dwóch modułów: nasłuchu oraz analizy. Pierwszy z modu-

---

[50] Najnowszy raport pochodzi z 2013 r.: Reporters Without Borders, *Special Report on Internet Surveillance, Focusing on 5 Governments and 5 Companies "Enemies of the Internet"*, 15.03.2013, https://rsf.org/en/news/special-report-internet-surveillance-focusing-5-governments-and-5-companies-enemies-internet (dostęp: 20.12.2018).

[51] European Parliament, *Report on the Existence of a Global System for the Interception of Private and Commercial Communications (ECHELON Interception System)*, Session document, FINAL A5-0264/2001, 11.07.2001, http://www.fas.org/irp/program/process/rapport_echelon_en.pdf (dostęp: 20.12.2018).

[52] D. Campbell, *Making History: The Original Source for the 1988 First Echelon Report Steps Forward*, 25.02.2000, cryptome.org/echelon-mndc.htm (dostęp: 24.12.2018); J. Rafa, *Złowrogi Echelon*, „PCKurier" 2000, nr 10, http://www.wsp.krakow.pl/papers/echelon.html (dostęp: 20.12.2018).

[53] H. Kozieł, *Systemy szpiegostwa elektronicznego*, 30.01.2006, http://www.psz.pl/124-polityka/hubert-koziel-systemy-szpiegostwa-elektronicznego http://www.pe24.pl/index2.php?option=com_content&do_pdf=1&id=2204 (dostęp: 20.12.2018).

łów składa się z co najmniej kilkunastu stacji nasłuchowych rozmieszczonych na całym świecie (Australia, Japonia, Kanada, Niemcy, Nowa Zelandia, Stany Zjednoczone, Wielka Brytania, prawdopodobnie Cypr) oraz stu parudziesięciu satelitów geostacjonarnych[54]. Przechwytywana jest komunikacja na wszystkich zakresach fal elektromagnetycznych: możliwy jest więc nasłuch radia, telefonii, w tym telefonii satelitarnej[55]. Mówi się, że system Echelon jest władny przechwycić nawet 90% komunikacji w ruchu międzynarodowym. Stacje nasłuchowe zaopatrzone są w oprogramowanie analityczne noszące nazwę Dictionary[56]. Umożliwia ono filtrowanie przekazów pod kątem określonych słów interesujących służby (tzw. *trigger words*, czyli słów-wyzwalaczy). Słowa te – podobno – automatycznie powodują uaktywnienie się nasłuchu systemu Echelon i odpowiednie raportowanie. Zautomatyzowane filtrowanie raportów oraz selekcja dokonywana przez analityków wyznaczają cele do dalszej obserwacji. Domyślać się można, jak wielkie środki finansowe i organizacyjne przeznaczane są na ten projekt.

Wymieniać można też inne liczne narzędzia służące współczesnym państwom do inwigilacji i nadzoru obywateli, jednak swoistym typem idealnym jawi się nowo ujawniony amerykański zestaw narzędzi noszący nazwę PRISM. Jest to program stworzony na zamówienie i używany przez NSA, a także CIA oraz inne służby. Jego istnienie ujawnił opinii publicznej Edward Joseph Snowden, były pracownik CIA, następnie zatrudniony w Booz Allen Hamilton mającej około 80 ośrodków na całym świecie i dostarczającej usługi informacyjne NSA. W lipcu 2013 roku E.J. Snowden ujawnił informacje, według których dobrowolnie lub pod naciskiem zgodziło się udostępniać dane o użytkownikach dziewięć następujących koncernów: Apple, America OnLine (AOL), Facebook, Google, Microsoft, Paltalk, Skype, Yahoo oraz YouTube. PRISM umożliwia odczytywanie następujących rodzajów wiadomości zdeponowanych przez użytkowników korzystających z usług wymienionych firm: e-maile, wiadomości z komunikatorów, filmy, zdjęcia, pliki przechowywane w chmurze, czaty głosowe, pliki przesyłane wewnątrz serwisów, wideokonferencje, czasy logowania, a także aktywność w profilach portali

[54] M. Assser, *Echelon: Big Brother without a Cause?*, BBC News, 6.06.2000, http://news.bbc.co.uk/2/hi/europe/820758.stm (dostęp: 20.12.2018).

[55] European Parliament, *Report on the Existence Of A Global System for the interception of private and commercial communications (ECHELON interception system)*, 11 lipca 2001, http://www.fas.org/irp/program/process/rapport_echelon_en.pdf (dostęp: 20.12.2018).

[56] Szerzej na ten temat: N. Hager, *Secret Power. New Zealand's Role in the International Spy Network*, Nelson 1996, s. 42–56.

społecznościowych. Jednocześnie E.J. Snowden ujawnił istnienie i wykorzystywanie innych programów szpiegowskich, między innymi oprogramowania Blarney służącego do monitorowania przepływu wiadomości e-mail i ruchu sieciowego. Zdemaskował również pozostającą na usługach NSA grupę hakerów noszącą nazwę Tailored Acess Operations zatrudniającą blisko 600 pracowników[57]. Ich główne zadanie – według E.J. Snowdena – polegało na infiltracji wskazanych komputerów i sieci na całym świecie. Istnienie PRISM, choć bez wskazania jego nazwy i z zastrzeżeniem, że programu używa się każdorazowo wyłącznie za wiedzą i zgodą sądu, potwierdził Dyrektor Wywiadu Narodowego (Director of National Intelligence)[58] James Robert Clapper[59]. Przykład Stanów Zjednoczonych ukazuje, jak w ciągu kilkudziesięciu lat powstały i funkcjonowały niezauważalnie potężne instytucje inwigilujące, których kadra liczy co najmniej kilka setek osób, zarówno urzędników państwowych i funkcjonariuszy, jak też przedstawicieli firm prywatnych.

Kolejny współczesny aspekt zjawiska inwigilacji to fakt, że konstytuuje się ono społecznie i że dostrzegalna jest swoista normalizacja. Kolejne afery podsłuchowe oraz ujawnione potężne narzędzia sprawiły, iż środki inwigilacji powszednieją; zjawisko można rozważać w kategoriach normalizacji dewiacji[60] bądź wręcz degradacji moralnej[61].

Rozważmy historyczny i współczesny przypadek społecznej reakcji na inwigilację. Na początku stycznia 1932 roku ujawniono w ulotkach krążących po Warszawie treść rozmowy telefonicznej ówczesnego premiera Kazimierza Bartla z prezydentem Ignacym Mościckim. Oznaczało to, iż linia telefoniczna pomiędzy Prezydium Rady Ministrów a pałacem w Spale musiała być na podsłuchu. Wyrazy oburzenia słyszano ze wszystkich stron sceny politycznej – na przykład zarówno endecka „Gazeta Warszawska", jak i prorządowy „Ilustrowany Kurier Codzienny"

---

[57] M.M. Aid, *Inside the NSA's Ultra-Secret China Hacking Group*, „Foreign Policy", 10.06.2013, http://www.foreignpolicy.com/articles/2013/06/10/inside_the_nsa_s_ultra_secret_china_hacking_group (dostęp: 20.12.2018).

[58] Dyrektor Wywiadu Narodowego (DNI) stoi na czele federacji 16 rządowych agencji wywiadowczych – Wspólnoty Wywiadowczej (Intelligence Community), jest głównym doradcą prezydenta i Rady Bezpieczeństwa Narodowego w sprawach wywiadowczych związanych z bezpieczeństwem narodowym.

[59] S. Waterman, *„ Prism" a Vital Program Used to Collect Personal Web Data*, „Washington Times", 7.06.2013, http://www.washingtontimes.com/news/2013/jun/7/prism-used-collect-personal-web-data-clapper-says/ (dostęp: 20.12.2018).

[60] A. Siemaszko, *Granice tolerancji. O teoriach zachowań dewiacyjnych*, Warszawa 1993.

[61] N. Dobos, *Two Concepts of Moral Injury: Moral Trauma and Moral Degradation*, [w:] T. Frame (red.), *Moral Injury: Unseen Wounds in an Age of Barbarism*, Sydney 2016.

potępiali zjawisko w kategoriach moralnych. Przeprowadzone śledztwo ujawniło sprawcę – był nim dziennikarz Jan Seinfeld, który wyspecjalizował się w zbieraniu poufnych informacji z urzędów centralnych. Pomimo znacznego oburzenia nie wyjaśniono nigdy sposobu, w jaki wszedł on w posiadanie informacji. Jedna z hipotez głosi, iż sanacyjne władze prowadziły inwigilację obywateli używając telefonicznych stacji podsłuchowych, a J. Seinfeld porozumiał się z funkcjonariuszami obsługującymi taką jednostkę, prawdopodobnie ich przekupując. Historia III Rzeczypospolitej w nieporównywalnie większym stopniu bogata jest w podsłuchy i również w sposób nieporównywalny – w osłabione reakcje na nie. W 2002 roku redaktor naczelny „Gazety Wyborczej" Adam Michnik zarejestrował w swoim gabinecie prywatną rozmowę między nim a znanym producentem filmowym Lwem Rywinem, który zaproponował zmiany w ustawie umożliwiające firmie Agora kupno telewizji Polsat w zamian za 17,5 mln dolarów na kampanię SLD i prezesurę Polsatu dla siebie. W 2006 roku posłanka Samoobrony Renata Beger nagrała spotkanie z czołowymi politykami Prawa i Sprawiedliwości – Adamem Lipińskim i Wojciechem Mojzesowiczem. W tym samym roku nagrany został premier Józef Oleksy przez Aleksandra Gudzowatego w siedzibie Bartimpeksu. J. Oleksy ujawnił liczne nieprawidłowości, do jakich dochodziło podczas rządów SLD oraz liczne informacje dotyczące prywatnego życia działaczy tej formacji. Rok później w toku dokonanej przez oficera Centralnego Biura Antykorupcyjnego działającego pod przykryciem (Tomasza Kaczmarka) kontrolowanej propozycji korupcyjnej zarejestrowano film, na którym posłanka Platformy Obywatelskiej Beata Sawicka przyjęła 50 tys. złotych w zamian za obietnicę wpłynięcia na przetarg publiczny. Z kolei w 2009 roku CBA ujawniło stenogramy z nagrań rozmów szefa klubu parlamentarnego Platformy Obywatelskiej Zbigniewa Chlebowskiego z Ryszardem Sobiesiakiem, biznesmenem działającym w branży hazardowej, zainteresowanym konkretnym kształtem ustawy o grach i zakładach. W latach 2014–2018 opublikowano szereg nagrań lub stenogramów z największej polskiej afery podsłuchowej – rozmowy zostały podsłuchane w restauracjach: Sowa & Przyjaciele, Amber Room oraz Osteria. Pierwszą reakcją obozu władzy było stanowcze oświadczenie, iż nikt nie zostanie zdymisjonowany, ponieważ byłoby to destabilizacją instytucji państwa[62]. Pierwsze dymisje nastąpiły dopiero rok później

[62] M. Mańkowski, *Donald Tusk: Dymisji i tak nie będzie. Polski rząd nie działa pod dyktando przestępców*, NaTemat.pl, 23.06.2014, https://natemat.pl/107269,donald-tusk-dymisji-i-tak-nie-bedzie-polski-rzad-nie-dziala-pod-dyktando-przestepcow (dostęp: 20.12.2018).

w rządzie Ewy Kopacz. Warto podkreślić, że treść rozmów, choć karygodna, nie wywołała szerszej reakcji społeczeństwa polskiego, ulicznych protestów podjęli się jedynie narodowcy[63], a w ogólnym społecznym czy medialnym odbiorze nie odnotowano oburzenia na akt podsłuchiwania, lecz raczej rodzaj *Schadenfreude*.

Zakończenie: antycypacje trendów

Przegląd historii techniki czyni dostrzegalnym kilka progów jakościowych, które pod względem łatwości użytkowania urządzeń i zakresu zbieranych informacji pokonały urządzenia inwigilacyjne. Uwidaczniają się trzy etapy już osiągnięte i jeden antycypowany, wyznaczając kolejne przełomy – „rewolucje inwigilacyjne". Pierwszy z przełomów to moment rozpoczęcia stosowania mikrofonu i przesyłu zdobytej informacji najpierw przewodami, a następnie drogą radiową, w miejsce podsłuchów „architektonicznych" lub – nazwijmy obrazowo – „analogowych", to jest ludzkich. Kolejną istotną zmianę wyznacza informatyzacja i powszechny dostęp do Internetu – zakres i kategorie informacji przesyłanych drogą elektroniczną zwiększyły się wówczas niepomiernie. Symbol kolejnego poziomu rozwoju, a zarazem nowej jakości w metodach inwigilacji stanowi smartfon – jako urządzenie osobiste, stale nam towarzyszące. Najbliższy z przewidywanych przełomów przyszłości wydaje się realny wraz z nastaniem tzw. Internetu rzeczy (*Internet of Things*), co umożliwi nieograniczoną obserwację i analizę nie tylko treści komunikacji ludzkiej, ale także dowolnych elementów behawioralnych – biologicznych i społecznych.

W perspektywie technologicznej i społecznej można przewidywać dwa zjawiska związane ze środkami inwigilacji elektronicznej. Po pierwsze, możliwą utratę kontroli nad systemami inwigilacyjnymi, co wiązać się może zarówno z autonomizacją tych systemów wskutek wdrażania algorytmów sztucznej inteligencji, jak również z czynnikiem ludzkim – rosnącą rolą swoistej technokracji, inżynierów i techników odpowiedzialnych za rozwój i obsługę tych systemów[64]. Rola, jaką odegrali

[63] IAR, *Protesty narodowców w całym kraju. „Rząd do dymisji"*, PolskieRadio.pl, 17.06.2014, https://www.polskieradio.pl/5/3/Artykul/1155094,Protesty-narodowcow-w-calym-kraju-Rzad-do-dymisji (dostęp: 20.12.2018).

[64] Zagadnienie autonomizacji i jej negatywnych skutków rozważa interesująco: K. Michalski, *Autonomizacja techniki i niepożądane skutki eliminowania człowieka jako źródła błędów*, „Filo-Sofija" 2017, nr 39(4/I).

Margaret Newsham, Edward Joseph Snowden i Chelsea [Bradley] Manning[65], wykazuje, że jest to żywioł nieprzewidywalny, ujawniane są bowiem najgłębiej skrywane inwigilacyjne tajemnice. Po wtóre, dostrzec można symptomy eskalacji wyścigu technicznego: środków podsłuchu i samoobrony, co w dalszej perspektywie może doprowadzić do ideologicznej i faktycznej negacji instytucji państwa przez grupy społeczne zjednoczone wokół idei wolności informacyjnej. Symptomy są już obecnie dostrzegalne – uformowało się silne środowisko skupione wokół technik zapewniających użytkownikom pełną prywatność i wyłączną kontrolę nad tworzonymi i przesyłanymi przez nich danymi (tzw. *Privacy Enhancing Technologies*, PET). Powstało już na przełomie lat 70. i 80. ubiegłego wieku, gdy dokonano technologicznego przełomu w ochronie prywatności elektronicznych danych. Whitfield Diffie i Martin Hellman opracowali metodę szyfrowania asymetrycznego (klucz publiczny, klucz prywatny), wówczas także stworzyli matematyczny koncept pierwszej kryptowaluty[66]. Jacob Appelbaum, współautor (wraz między innymi z Julianem Assange'm) książki *Cypherpunks. Freedom and the Future of the Internet*, niezależny dziennikarz i specjalista z zakresu cyberbezpieczeństwa, trafnie ujął znaczenie tego wynalazku: „Silna kryptografia może opierać się nieograniczonemu stosowaniu przemocy. Żadna siła przymusu nigdy nie rozwiąże problemu matematycznego"[67].

Amorficzny, lecz liczny i silny ruch na rzecz anonimowości rozwinął się istotnie od tego czasu, tworząc ideologiczno-doktrynalne treści (zarówno niezbyt obszerne, jak i niezbyt zaawansowane merytorycznie), lecz przede wszystkim silnie i konsekwentnie obudowując się w rozmaite „technologie wolności".

Jeśli chodzi o ideologiczno-doktrynalne treści, należy w pierwszej kolejności wymienić powstałą na przełomie lat 80. i 90. XX wieku grupę Cypherpunks (nazwa powstała *ad hoc*, w wyniku sytuacyjnego żartu) organizującą comiesięczne spotkania dotyczące prywatności oraz rządo-

[65] Informator WikiLeaks Chelsea [Bradley] Manning (postanowił zmienić płeć i teraz nosi imię Chelsea) miał stopień starszego szeregowego wojsk lądowych, został aresztowany w 2010 r. za przekazanie WikiLeaks ponad 700 tys. dokumentów wojskowych i dyplomatycznych, a także plików wideo dotyczących operacji wojskowej USA w Iraku.

[66] W. Gogłoza, *Cypherpunks, WikiLeaks i widmo kryptograficznej anarchii*, https://wgogloza.com/umcs/informatyka-prawnicza/cypherpunks/ (dostęp: 20.12.2018).

[67] W oryginale: „Strong cryptography can resist an unlimited application of violence. No amount of coercive force will ever solve a math problem". J. Assange, J. Appelbaum, A. Müller-Maguh, J. Zimmerman, *Cypherpunks. Freedom and the Future of the Internet*, Nowy Jork – Londyn 2012, s. 5.

wej i korporacyjnej kontroli informacji. Impuls do działania grupie nadał opublikowany w połowie lat 80. artykuł Davida Chauma *Security without Identification: Transaction Systems to Make Big Brother Obsolete*[68]. Istotne znaczenie należy przypisać treściwemu manifestowi Chucka Hammilla *Od kuszy do kryptografii, czyli psucie szyków państwa przy pomocy techniki* (*From Crossbows to Cryptography. Thwarting the State via Technology*), widzącego w nieskrępowanym dzieleniu się informacją akt oddziaływania na władze silniejszy niż przemoc[69]. Nazwę własną ruchu – kryptoanarchizm – wprowadził w 1992 roku Timothy C. May, amerykański pisarz, inżynier elektronik, dawny pracownik firmy Intel[70]. Jest on autorem innych ważkich dla kryptoanarchistycznego ruchu publikacji, jak *Cyphernomicon*[71], *True Nyms and Crypto Anarchy*[72]. Wartościowym i właściwie zamykającym listę lektur kryptoanarchistów wykładem wydaje się również praca zbiorowa *Crypto Anarchy, Cyberstates, and Pirate Utopias*[73].T.C. May jest autorem najbardziej nośnej koncepcji kryptoanarchizmu – tzw. rynku zabójców (*Assassination Market*), który ukazuje stopień determinacji w drastycznym ograniczeniu informacyjnych apetytów państw[74].

Rynek zabójców spopularyzowany został przez amerykańskiego naukowca, wynalazcę i dysydenta Jamesa Daltona Bella w klasycznym dla kryptoanarchistów 10-częściowym eseju zatytułowanym *Assassination Politics*. Rynek zabójców stanowi metaprojekt globalnego, trwałego wyeliminowania sfery politycznej ze stosunków międzyludzkich dzięki anonimowej komunikacji i anonimowemu, elektronicznemu pieniądzowi w Internecie. Na poziomie wdrożeniowym *Assassination market* to strona internetowa zaopatrzona w narzędzia zabezpieczające tożsamość użytkowników, na której przyjmowane są zakłady[75]. Przedmiotem zakładu jest

---

68 D. Chaum, *Security without Identification. Transaction Systems to Make Big Brother Obsolete*, „Communications of the ACM" 1985, nr 28(10).

69 Ch. Hammill, *Od kuszy do kryptografii, czyli psucie szyków państwa przy pomocy techniki*, przekł. J. Sierpiński, „Kultura i Historia" 2007, nr 11, http://www.kulturaihistoria.umcs.lublin.pl/archives/701 (dostęp: 20.12.2018).

70 T.C. May, *The Crypto Anarchist Manifesto*, Activism.net, 22.11.1992, https://www.activism.net/cypherpunk/crypto-anarchy.html (dostęp: 20.12.2018).

71 T.C. May, *Cyphernomicon. Cypherpunks FAQ and More, Version 0.666*, 10.09.1994, http://www.cypherpunks.to/faq/cyphernomicron/cyphernomicon.html (dostęp: 20.12.2018).

72 T.C. May, *True Nyms and Crypto Anarchy*, 2001, http://www.isfdb.org/cgi-bin/title.cgi?195636 (dostęp: 20.12.2018).

73 P. Ludlow (red.), *Crypto Anarchy, Cyberstates, and Pirate Utopias*, Cambridge 2001.

74 T.C. May, *Cyphernomicon*…

75 J. Bell, *Assassination Politics*, 1997, http://outpost-of-freedom.com/jimbellap.htm (dostęp: 20.12.2018).

śmierć osoby publicznej, uprzednio (po wniesieniu opłaty rejestracyjnej) zgłoszonej (tzw. *dead pool*). Każdy może zaproponować przewidywaną przez siebie datę, zawierając zakład i wpłacając kwotę, którą uzna za adekwatną. Wygrywa, otrzymując pulę wpłat, ten z uczestników, który najdokładniej przewidzi czas śmierci osoby, której zakład dotyczy. Jest to w istocie zakamuflowany tytułowy rynek zabójców: któż bowiem bardziej precyzyjnie przewidzi datę śmierci ofiary niż jej zabójca, a uczestnicy zakładu to w istocie zakamuflowani zbiorowi zleceniodawcy darzący daną osobę publiczną negatywnymi uczuciami. Im ich więcej, tym pula zakładu wyższa, w tym większym stopniu staje się finansowo opłacalne dla wystarczająco zdeterminowanej jednostki podjęcie zabójstwa politycznego. W założeniu globalność i sieciowość projektu zapewnią niemal każdorazowo opłacalność dokonania zabójstwa, rychło doprowadzając do permanentnej destabilizacji, a następnie, w dalszej perspektywie, całkowitej eliminacji sfery politycznej. Z technicznego punktu widzenia (anonimowa komunikacja oraz anonimowy pieniądz) obecnie nie istnieją przeszkody w realizacji takiego projektu. Przez pewien czas – od 2013 roku – w sieci TOR funkcjonował rynek zabójców pod adresem assmkedzgorodn7o.onion[76]. Pomysłodawcą, twórcą i administratorem serwisu była osoba nosząca pseudonim Kuwabatake Sanjuro[77], kryptoanarchista kierowany przesłankami ideologicznymi, głęboko przekonany, iż inicjatywa ta zapoczątkuje serie zabójstw polityków, doprowadzając do ziszczenia się anarchistycznego raju: zniesienia wszelkich form rządów na całym świecie.

Siłą kryptoanarchizmu są jednak wdrożenia technologiczne – powstały liczne produkty zapewniające bezpieczeństwo informacyjne użytkownikom Internetu – między innymi silnie anonimizujące użytkowników sieci The Onion Router wraz z dedykowaną przeglądarką internetową TOR Browser[78], Freenet[79] oraz Invisible Internet Project[80], niepozostawiający śladów użytkowania lokalnie i anonimizujący zdalnie system operacyjny Linux Tails[81] – dedykowany użytkownikom pragnącym zachować naj-

[76] Od 2015 r. strona nie funkcjonuje, jednak zdeponowane bitcoiny nie zostały podjęte przez właściciela witryny.

[77] Główna postać, samuraj, w japońskim filmie *Straż przyboczna* (reż. Akira Kurosawa, 1961).

[78] *TOR Project*, https://www.torproject.org (dostęp: 20.12.2018).

[79] *Freenet*, https://freenetproject.org/author/freenet-project-inc.html (dostęp: 20.12.2018).

[80] *Invisible Internet Project*, https://geti2p.net/pl/ (dostęp: 20.12.2018).

[81] *TAILS (The Amnesic Incognito Live System)*, https://tails.boum.org (dostęp: 20.12.2018).

wyższy stopień prywatności, VeraCrypt[82] (nierozwijany już TrueCrypt) umożliwiający szyfrowanie dowolnych danych silnymi algorytmami. Dostępnych jest także szereg mniej znanych produktów, jak pozwalający dowolnie zmieniać tożsamość w Internecie Advanced Onion Router[83], liczne usługi bezpiecznych, szyfrowanych czatów, komunikatorów oraz kont pocztowych[84].

Te dwa nurty: dynamiczny rozwój systemów inwigilacyjnych oraz równolegle technik samoobrony przed inwigilacją są nie tylko rywalizacją o charakterze technologicznym. W dalszej perspektywie sytuacja taka może doprowadzić do poważnej i nieodwracalnej erozji zaufania społecznego wobec instytucji państwa, skutkując postrzeganiem tych instytucji jako opresywnych i totalitarnych. Z kolei państwa mogą nasilać kulturowo-społeczną deprecjację prywatności jako wartości samej w sobie oraz penalizować akty anonimizacji działań i ukrywania tożsamości.

Streszczenie

Tekst ogniskuje się na rozmaitych aspektach inwigilacji z użyciem narzędzi elektronicznych. Autorzy poszukują odpowiedzi na szereg pytań. Po pierwsze, jakie typy negatywnych zjawisk generują i są intensyfikowane przez technologie inwigilacji elektronicznej? Po wtóre, jak głęboki jest stan „bezbronności inwigilacyjnej" współczesnych społeczeństw, to jest jakie są możliwości urządzeń służących inwigilacji? Po trzecie, czy istnieje możliwość praktycznego przeciwstawienia się im i – jeśli tak – w jaki sposób? Po czwarte, jaka jest geneza tych zjawisk i jakie spodziewane scenariusze przyszłości można szkicować na podstawie antycypacji zaobserwowanych trendów? Tak zdefiniowany zbiór pytań badawczych wymagał oglądu zarazem z dwóch perspektyw: socjologicznej i technicznej. Autorzy dostrzegają i analizują szereg negatywnych zjawisk związanych z inwigilacją elektroniczną – jej eskalację, profesjonalizację, instytucjonalizację i normalizację.

[82] *VeraCrypt*, https://www.veracrypt.fr/en/Downloads.html (dostęp: 20.12.2018).

[83] *Team Elite*, https://www.te-home.net/?do=work&id=advor (dostęp: 20.12.2018).

[84] Tę jakże obszerną grupę zagadnień autorzy poruszają koncepcyjnie i praktycznie w toku kursów prowadzonych przez nich na Uniwersytecie Otwartym Uniwersytetu Warszawskiego: *Nie daj się podsłuchać, nie daj się zhakować – warsztaty cybersamoobrony dla humanistów*.

Paweł Tomczyk, Daniel Mider, Józef Grzegorczyk

Electronic Surveillance as a Method ob Obtaining Information – Evaluation and Forecasts

The text focuses on one of the elements belonging to the surveillance society – surveillance with the use of electronic tools. The authors attempt to answer the following questions. What types of negative phenomena are produced and intensified by electronic surveillance technologies? How deep is the state of the „vulnerability" of modern societies, what are the possibilities of surveillance devices? Is it possible to practically oppose them, how and what are the limits? What is the genesis of these phenomena and what future scenarios can be sketched based on the anticipation of observed trends? A set of research questions defined in this way required both sociological and technical perspectives at the same time. The authors recognize the negative phenomena associated with electronic surveillance: escalation, professionalization, institutionalization and normalization.

Key words: *social informatics, surveillance society, electronic surveillance, infobrokering*

Bibliografia

Aid M.M., *Inside the NSA's Ultra-Secret China Hacking Group*, „Foreign Policy", 10.06.2013, http:// www.foreignpolicy.com/articles/2013/06/10/inside_the_nsa_s_ultra_secret_china_hacking_group (dostęp: 20.12.2018).

Assange J., Appelbaum J., Müller-Maguh A., Zimmerman J., *Cypherpunks. Freedom and the Future of the Internet*, Nowy Jork – Londyn 2012.

Assser M., *Echelon: Big Brother without a cause?*, BBC News, 6.06.2000, http://news.bbc.co.uk/2/hi/europe/820758.stm (dostęp: 20.12.2018).

Batey A., *The Spies Behind Your Screen*, „Telegraph", 24.10.2011, https://www.telegraph.co.uk/technology/8899353/The-spies-behind-your-screen.html (dostęp: 20.12.2018).

Bell J., *Assassination Politics*, 1997, http://outpost-of-freedom.com/jimbellap.htm (dostęp: 20.12.2018).

Błoński M., *Najbardziej zaawansowana operacja hakerska w historii*, http://kopalniawiedzy.pl/Equation-Group-haker-szpiegostwo-NSA,21930 (dostęp: 20.12.2018).

Campbell D., *Making History: The Original Source for the 1988 First Echelon Report Steps Forward*, 25.02.2000, cryptome.org/echelon-mndc.htm (dostęp: 20.12.2018).

Chaum D., *Security without Identification. Transaction Systems to Make Big Brother Obsolete*, „Communications of the ACM" 1985, nr 28(10).

Choi H.-J., i in., *Reconstruction of Leaked Signal From USB Keyboards*, 2016, http://www.researchgate.net/publication/309327769_Reconstruction_of_leaked_signal_from_USB_keyboards (dostęp: 20.12.2018).

Dobos N., *Two Concepts of Moral Injury: Moral Trauma and Moral Degradation*, [w:] T. Frame (red.), *Moral Injury: Unseen Wounds in an Age of Barbarism*, Sydney 2016.

Eck W. van, *Electromagnetic Radiation from Video Display Units: An Eavesdropping Risk?*, „North-Holland Computers & Security" 1985, nr 4.

Elibol F., Sarac U., Erer I., *Realistic Eavesdropping Attacks on Computer Displays with Low-Cost and Mobile Receiver System*, [w:] *Proceedings of the 20th European Signal Processing Conference (EUSIPCO)*, 2012, http://www.eurasip.org/Proceedings/Eusipco/Eusipco2012/Conference/papers/1569583239.pdf (dostęp: 20.12.2018).

European Parliament, *Report on the Existence of a Global System for the Interception of Private and Commercial Communications (ECHELON Interception System)*, Session document, FINAL A5-0264/2001, 11.07.2001, http://www.fas.org/irp/program/process/rapport_echelon_en.pdf (dostęp: 20.12.2018).

Frankland R., *Side Channels, Compromising Emanations and Surveillance. Current and Future Technologies*, Londyn 2011, http://pdfs.semanticscholar.org/87a4/182d66ab649a35eff0267c5e3a73bb2a5087.pdf (dostęp: 20.12.2018).

Gandy O., *Data Mining and Surveillance In the Post – 9/11 Environment* [w:] K. Ball, F. Webster (red.), *The Intensification of Surveillance. Crime, Terrorism and Warfare in the Information Era*, Londyn 2003.

Guri M., Kedma G., Kachlon A., Elovici Y., *AirHopper. Bridging the Air-Gap between Isolated Networks and Mobile Phones using Radio Frequencies*, 2014, https://www.wired.com/wp-content/uploads/2014/11/air-hopper-malware-final-e-141029143252-conversion-gate01.pdf (dostęp: 8.01.2019).

Guri M., Monitz M., Mirski Y., Elovici Y., *BitWhisper. Covert Signaling Channel between Air-Gapped Computers using Thermal Manipulations*, „Cryptography & Security" 2015.

Guri M., Solewicz Y., Daidakulov A., Elovici Y., *Fansmitter. Acoustic Data Exfiltration from (Speakerless) Air-Gapped Computers*, 2016, http://www.wired.com/wp-content/uploads/2016/06/Fansmitter-1.pdf (dostęp: 20.12.2018).

Hager N., *Secret Power. New Zealand's Role in the International Spy Network*, Nelson 1996.

Kania B., *VGASIG. FM Radio Transmitter Using VGA Graphics Card*, 2009, https://bk.gnarf.org/creativity/vgasig/vgasig.pdf (dostęp: 8.01.2019).

Kuhn M.G., *Compromising Emanations: Eavesdropping Risks of Computer Displays*, „Computer Laboratory" 2003, nr 577.

Kuhn M.G., *Electromagnetic Eavesdropping Risks of Flat-Panel Displays*, 2004, http://www.cl.cam.ac.uk/~mgk25/pet2004-fpd.pdf (dostęp: 20.12.2018).

Lawry T., *An Acoustic-Electric Bridge: Traversing Metal Barriers Using Ultrasound*, http://www.ttivanguard.com/ttivanguard_cfmfiles/pdf/miami11/miami11session7014.pdf (dostęp: 20.12.2018).

Ludlow P. (red.), *Crypto Anarchy, Cyberstates, and Pirate Utopias*, Cambridge 2001.

Lyon D., *The Electronic Eye. The Rise of Surveillance Society*, Minneapolis 1994.

May T.C., *The Crypto Anarchist Manifesto*, Activism.net, 22.11.1992, https://www.activism.net/cypherpunk/crypto-anarchy.html (dostęp: 20.12.2018).

May T.C., *Cyphernomicon. Cypherpunks FAQ and More, Version 0.666*, 10.09.1994, http://www.cypherpunks.to/faq/cyphernomicron/cyphernomicon.html (dostęp: 20.12.2018).

May T.C., *True Nyms and Crypto Anarchy*, 2001, http://www.isfdb.org/cgi-bin/title.cgi?195636 (dostęp: 20.12.2018).

Michalski K., *Autonomizacja techniki i niepożądane skutki eliminowania człowieka jako źródła błędów*, „Filo-Sofija" 2017, nr 39(4/I).

Miller P., *Keyboard "Eavesdropping" Just Got Way Easier, Thanks to Electromagnetic Emanations*, Engadget, 20.10.2008, http://www.engadget.com/2008/10/20/keyboard-eavesdropping-just-got-way-easier-thanks-to-electrom/?guccounter=1 (dostęp: 20.12.2018).

Murray K.D., *The Great Seal Bug*, http://counterespionage.com/great-seal-bug-part-1/ (dostęp: 20.12.2018).

Nikitin P., *Leon Theremin (Lev Termen)*, „IEEE Antennas and Propagation Magazine" 2012, nr 54(5).

Pavithran M., *Eavesdropping on GSM*, „International Journal of Engineering Research in Computer Science and Engineering" 2016, nr 3(9).

Reis K., *The Eavesdropping Society. Electronic Surveillance and Information Brokering*, „Patents, Copyrights, Trademarks, and Literary Property", June 2001.

Simon B., *The Return of Panopticism: Supervision, Subjection and the New Surveillance*, „Surveillance & Society" 2005, nr 3(1).

Špaček S., Celeda P., Drašar M., Vizváry M., *Analyzing an Off-the-Shelf Surveillance Software. Hacking Team Case Study*, 2.06.2017, http://spi.unob.cz/papers/spi2017.html (dostęp: 20.12.2018).

Vuagnoux M., Pasini S., *Compromising Electromagnetic Emanations of Wired Keyboards, 2007–2009 Security and Cryptography Laboratory – LASEC/EPFL*, 2009, https://lasec.epfl.ch/keyboard/ (dostęp: 20.12.2018).

Waterman S., *„ Prism" a Vital Program Used to Collect Personal Web Data*, „Washington Times", 7.06.2013, http://www.washingtontimes.com/news/2013/jun/7/prism-used-collect-personal-web-data-clapper-says/ (dostęp: 20.12.2018).

STUDIA POLITOLOGICZNE VOL. 54
DOI: 10.33896/SPolit.2019.54.11

MICHAEL BAZZELL

Open Source Intelligence Techniques. Resources for Searching and Analyzing Online Information

Createspace Independent Publishing Platform, 6th ed., Charleston 2018, 461 s.

(Bartosz Biderman
ORCID: 0000-0002-8503-5207)

SŁOWA KLUCZOWE:
biały wywiad, wyszukiwanie informacji, infobrokering, wywiad jawnoźródłowy

Michael Bazzell jest postacią powszechnie rozpoznawaną, zarówno w środowisku osób zajmujących się białym, jak i szarym wywiadem internetowym. Sam siebie określa jako międzynarodowego konsultanta do spraw prywatności[1]. W branży białego wywiadu ma bardzo duże doświadczenie. Przez ponad 18 lat w imieniu rządu USA badał przestępstwa komputerowe. Większość tego czasu pracował dla Federalnego Biura Śledczego, gdzie był przydzielony do grupy zadaniowej Cyber Crimes Task Force[2]. Skupiał się tam na przeprowadzaniu badań online i gromadzeniu danych typu otwarte źródła informacji (ang. *open source intelligence*, OSINT). Jako

[1] Ang. *International Privacy Consultant*; zob. oficjalny profil autora publikacji: https://twitter.com/inteltechniques (dostęp: 16.12.2018).

[2] Wydział Cybernetyczny FBI, jednostka powstała w 2002 r. Zajmuje się cyberterroryzmem w czterech głównych dziedzinach: włamania komputerowe, kradzieże tożsamości, wykorzystywanie seksualne dzieci i poważne cyberoszustwa, w tym szpiegostwo. Ten pododdział FBI wykorzystuje informacje zebrane podczas śledztwa w celu informowania społeczeństwa o aktualnych tendencjach w cyberprzestępczości. *The Federal Bureau of Investigation's Ability to Address the National Security Cyber Intrusion Threat*, U.S. Department of Justice Office of the Inspector General Audit Division, April 2011, s. 2–4, https://itlaw.wikia.org/wiki/The_Federal_Bureau_of_Investigation%27s_Ability_to_Address_the_National_Security_Cyber_Intrusion_Threat (dostęp: 21.06.2019).

aktywny detektyw był zaangażowany w wiele poważnych śledztw kryminalnych, w tym prowadzonych w sprawie nagabywania dzieci przez Internet do czynności seksualnych, uprowadzeń dzieci, porwań, morderstw na zlecenie, gróźb terrorystycznych i włamań komputerowych.

Obecnie M. Bazzell od kilku lat zajmuje się przeprowadzaniem szkoleń z technik pozyskiwania i przetwarzania OSINT. W tym czasie wyszkolił tysiące osób (zarówno pracowników administracji państwowej, jak i w sektorze prywatnym), również w korzystaniu ze swoich autorskich rozwiązań i technik śledczych. Jako doradca techniczny wystąpił w pierwszym sezonie programu telewizyjnego *Mr. Robot*. Jego książki *Open Source Intelligence Techniques* (2012, szóste wydanie – 2018) oraz *Hiding from the Internet* (2012, czwarte wydanie – 2018) są najlepiej sprzedającymi się w Stanach Zjednoczonych i Europie pozycjami w zakresie OSINT[3].

Książka *Open Source Intelligence Techniques: Resources for Searching and Analyzing Online Information* (nr ISBN-13: 978-1984201577, nr ISBN-10: 1984201573) jest praktycznym wprowadzeniem do nauki pozyskiwania i przetwarzania informacji. W wyszukiwarce Google Scholar pod hasłem „OSINT" oprócz omawianej monografii odnaleźć można ponad pięć tysięcy innych druków zwartych i rozproszonych. Interesującym faktem jest, iż publikacja pierwszej edycji książki M. Bazzella pokrywa się z nagłym wzrostem popularności hasła „OSINT" w wyszukiwarce internetowej Google.com[4]. Od 2012 roku zarówno hasło „OSINT", jak i zwroty blisko z nim powiązane (między innymi nazwy oprogramowań: Maltego, FOCA) nieustannie zyskują na popularności. Wraz z kolejnymi edycjami książki zauważalne są skokowe wzrosty wyszukiwań (drugie wydanie – 2013, następne kolejno: 2014, 2015 i 2016). Na tej podstawie wnioskować można, iż autor wnosi znaczny wkład w popularyzację białego wywiadu oraz jego rozwój.

W omawianej monografii, oprócz przedstawienia powszechnie znanych w środowisku OSINT metod pozyskiwania informacji, M. Bazzell z dumą prezentuje również swoje autorskie rozwiązania mające znaczny wkład w rozwój cyberbezpieczeństwa. Jednym z nich jest stworzony w 2017 roku (we współpracy z Davidem Westcottem) Buscador Linux[5], któremu poświęcony został cały drugi rozdział oraz fragmenty w dalszych częściach publikacji. System ten ma za zadanie umożliwienie przeprowa-

[3] Oficjalna strona autora publikacji, gdzie streszczony jest jego życiorys: https://inteltechniques.com/live-keynotes.html (dostęp: 16.12.2018).

[4] Nagły, ponad pięciokrotny skok wyszukiwań można było zaobserwować w przeciągu dwóch miesięcy styczeń–luty 2012, https://trends.google.com (dostęp: 16.12.2018).

[5] System jest darmowy (niektóre narzędzia są dodatkowo płatne) i można go pobrać ze strony głównej autora. Najnowsza wersja 2.0 (ze stycznia 2019) dostępna pod adresem: https://inteltechniques.com/buscador/ (dostęp: 16.12.2018).

dzania badań online również osobom z niewielką wiedzą o obsłudze środowiska Linux. Środowisko pentesterów z dezaprobatą wyraża się o tym systemie, który *sensu stricto* jest czystą wersją Ubuntu Linux, urozmaiconą wbudowanymi dodatkowymi aplikacjami, takimi jak Maltego, Recon-ng, Creepy, Spiderfoot, TheHarvester czy Sublist3r. M. Bazzell w książce odpowiada jednak, iż takie było jego główne zadanie – utworzenie systemu lekkiego, przyjemnego w obsłudze, czyli systemu „dla każdego" oraz możliwego do postawienia na wirtualnej platformie[6]. Dodatkowy atut to fakt, że Buscador Linux jest jedną z niewielu platform, która bardzo płynnie współpracuje z komputerami Mac firmy Apple. Dokładny opis instalacji (zarówno wirtualnej maszyny, jak i stabilnego systemu lub bootowanego z pendriva) oraz opis obsługi systemu pozwalają na przeprowadzenie skomplikowanych operacji przez każdego, nawet z niewielkim doświadczeniem w informatyce.

Open Source Intelligence Techniques nie jest monografią *stricte* naukową. Brakuje w niej przypisów i odniesień do materiałów źródłowych. M. Bazzell nie jest jednak pracownikiem akademickim, lecz byłym pracownikiem służb, osobą z bardzo dużym doświadczeniem i chęcią podzielenia się nim. Prosty w użyciu język, duża liczba grafik oraz obszerne – czasami aż nazbyt – tłumaczenia problemów pozwalają traktować publikację jako poradnik w zakresie prowadzenia białego wywiadu. Autor po kolei pokazuje czytelnikowi, jak przejść całą drogę prawidłowo prowadzonego dochodzenia.

W rozdziale 1 znajdziemy więc wskazówki, jak bezpiecznie przygotować swój komputer przed przystąpieniem do pracy. Znajdują się tam szczegółowe techniki konfiguracji przeglądarek internetowych oraz proponowane ustawienia bezpieczeństwa sieci. Jest to element niezwykle ważny, gdyż umożliwiający zachowanie w miarę możliwości anonimowość w Internecie. W kolejnych rozdziałach umieszczone zostały szczegółowe opisy działania oraz metody użycia narzędzi śledczych, w zdecydowanej większości stworzonych przez autora. Umożliwiają one korzystanie z wielu aplikacji i stron służących do prowadzenia dochodzeń w usystematyzowany, łatwy sposób. Każdy rozdział poświęcony jest osobnemu źródłu informacji. Tak więc w czwartym znajduje się omówienie narzędzi śledczych służących do inwigilacji użytkowników Facebooka, a w piątym Twittera. Dalej opisane są między innymi narzędzia do przeszukiwań Internetu pod względem nazwy użytkownika (rozdz. 9), maila (rozdz. 8, 10) czy numeru telefonu[7] (rozdz. 11) lub adresu IP[8] (rozdz. 17). Rozdział 18 przybliża metody przeszu-

[6] Dzięki czemu z łatwością można kontrolować cały ruch wychodzący oraz przepuszczać go przez bramki proxy czy VPN.

[7] Działa wyłącznie z numerami zarejestrowanymi w krajach Ameryki Północnej.

[8] By odnaleźć omawiane narzędzia, należy na stronie https://inteltechniques.com kliknąć w zakładkę „Tools" (dostęp: 16.12.2018). Dostęp do nich

kiwania rządowych baz danych, jednak wyłącznie tworzonych i prowadzonych przez administrację USA. W rozdziale 19 autor omawia stosowane przez siebie inne aplikacje śledcze, w większości oparte na licencji *open source* i dostępne dla każdego[9]. Dalsza część powstała wraz z szóstym wydaniem i znajduje się w niej opis najnowszych metod śledczych stosowanych przez M. Bazzella, a często niesłusznie pomijanych w dochodzeniach ze względu na ich „młodzieżowy" charakter, pozorną nieprzydatność, jak i niedawne uruchomienie usług. W rozdziale 21 znajdziemy narzędzie wiążące nazwę użytkownika SnapChata z numerem telefonu czy wskazówki, do czego może się przydać aplikacja randkowa Tinder w zaawansowanych indagacjach śledczych. W rozdziale 24 autor porządkuje metody dochodzeniowe oraz proponuje gotowe schematy badań w zależności od posiadanych danych.

Głównym celem przyświecającym autorowi od pierwszego wydania pracy było – jak sam wskazuje we wstępie – umożliwienie przeprowadzania śledztw również osobom niespecjalizującym się w informatyce, takim jak pracownicy administracji publicznej czy prywatnych korporacji (uczestnicy jego szkoleń). Niewątpliwie ten cel został osiągnięty. Monografia w prosty sposób przybliża nawet bardziej zaawansowane metody śledcze i czyni je zrozumiałymi. Zastanawiający jest dla mnie jednak klucz doboru opisywanych narzędzi. W książce brakuje opisu działań aplikacji zawartych między innymi w autorskim systemie Buscador Linux, jak na przykład Maltego, ponadto należy wspomnieć praktyczne pominięcie opisu kombajnów takich jak SpiderFoot czy Harvester[10]. Wszystkie trzy służą do przeprowadzania ukierunkowanych na dany cel dochodzeń, zgodne są więc z założeniem przyjętym przez M. Bazzella, iż OSINT to zdobywanie wiedzy o osobie w realnym świecie przy pomocy danych pozostawionych w tym wirtualnym. Brak użycia tych aplikacji przyczynić się może do niekompletności wyników badań.

Sądzę jednak, iż powyższa publikacja w znacznym stopniu przyczynia się do rozwoju OSINT. Choć w Internecie znaleźć można dużo artykułów i dokumentów traktujących o otwartych źródłach informacji, mało który jest na tyle kompletny, by przeprowadzić czytelnika krok po kroku po drabinie dochodzeniowej. Ubolewać można, iż wśród światowych autorów brak jest polskich naukowców. Nie ma woli poszerzania wiedzy dotyczącej białego wywiadu w Internecie. Istniejące publikacje, jak ta Krzysztofa Liedela i Tomasza Serafina[11] oraz praca zbio-

jest w większości darmowy lub możliwe jest skorzystanie z wersji Trail.

9 Między innymi w te narzędzia wyposażony został Buscador Linux.

10 Usprawiedliwić autora może fakt, iż dwie ostatnie wymienione aplikacje w niektórych państwach, w pewnych konfiguracjach, przekraczają granice dopuszczone normami prawnymi. Przez to niedoświadczony użytkownik narazić się może na konsekwencje prawne.

11 K. Liedel, T. Serafin, *Otwarte źródła informacji w działalności wywiadowczej*, Warszawa 2011.

rowa pod redakcją Wojciecha Filipkowskiego i Wiesława Mądrzejowskiego[12], choć bardzo cenne, są już w dużej mierze nieaktualne i wymagają ponownej edycji. Nie zostały one również przetłumaczone na język angielski, przez co grupę odbiorców zawężono wyłącznie do polskich badaczy. *Open Source Intelligence Techniques* to odpowiednia publikacja zarówno dla osób zaczynających swoją przygodę z OSINT, jak i tych bardziej zaawansowanych. Z pewnością każdy znajdzie tu coś interesującego dla siebie.

Streszczenie

Informacja jest złotem XXI wieku. Czy Michael Bazzell w swojej książce *Open Source Intelligence Techniques. Resources for Searching and Analyzing Online Information* (2018) uczy, jak ją zdobywać? W niniejszej recenzji zawarto odpowiedź na to pytanie, wraz z obiektywnym spojrzeniem na całą publikację. We wstępie przedstawiona została biografia autora, wraz z jego doświadczeniem oraz wpływem na rozwój OSINT-u, czyli pozyskiwania informacji ze źródeł otwartych. W dalszej części recenzent opisuje rozwiązania proponowane swoim czytelnikom przez M. Bazzella. Pod koniec ocenie poddany został cel pracy, trafność argumentacji oraz orientacja autora w najnowszym dorobku reprezentowanej przez niego dziedziny, a także język i logika formułowanych wypowiedzi. Krytykę kończy syntetyczne podsumowanie omawianych wad i zalet oraz wyraźna opinia o recenzowanym dziele.

Bartosz Biderman

Michael Bazzell, Open Source Intelligence Techniques. Resources for Searching and Analyzing Online Information

In the 21st century, information is golden. The question is whether in his book entitled *Open Source Intelligence Techniques: Resources for Searching and Analyzing Online Information* Michael Bazzell teaches us how to get it? This review answers this question by providing an objective overview on the publication. The introduction contains certain biographical information on the author, including information on his experience and its impact on the development of the open-source intelligence (OSINT). Next, the reviewer goes on to describe the solutions that M. Bazzell proposes to his readers, followed by the assessment of the work's goal and reasoning, the author's familiarity with the recent studies in the field, as well as the language of and the logic behind the statements. The review ends with a summary of the strengths and weaknesses of the publication and a clear view on the work.

Key words: *OSINT, white hat, data broker, intelligence assessment*

[12] W. Filipowski, W. Mądrzejowski (red.), *Biały wywiad. Otwarte źródła informacji – wokół teorii i praktyki*, Warszawa 2012.

VOL. 54 STUDIA POLITOLOGICZNE
DOI: 10.33896/SPolit.2019.54.12

Christopher Hadnagy

Socjotechnika. Sztuka zdobywania władzy nad umysłami (tyt. oryg. *Social Engineering. The Art of Human Hacking*)

Wydawnictwo „Helion"
Gliwice 2017 (2 wyd.), 424 s.
tłumaczenie Magda Witkowska

(Anna Maria Złocka)
ORCID: 0000-0001-5979-494X)

Słowa kluczowe:
socjotechnika, manipulacja, informacja

Christopher Hadnagy uważany jest za specjalistę w dziedzinie tworzenia zabezpieczeń w sieci. Jednym z jego osiągnięć jest stworzenie pierwszego modelu socjotechniki. Swoim wieloletnim doświadczeniem w tej branży oraz wiedzą na temat IT dzieli się między innymi w jednej ze swoich książek: *Socjotechnika – sztuka zdobywania władzy nad umysłami*. Jego naukowy dorobek wzbogacają również takie pozycje (we współautorstwie), jak wydana na polskim rynku *Mroczne odmęty phishingu – nie daj się złowić!* (2016)[1] oraz *Unmasking the Social Engineer. The Human Element of Security*[2] (2014). Swoją wiedzę i umiejętności związane z socjotechniką Ch. Hadnagy pogłębiał na kursie w zakresie mikroekspresji (prowadzonym przez dr. Paula Ekmana)[3], który ukończył zdobywając poziom ekspercki. Ma rów-

[1] Ch. Hadnagy, M. Fincher, *Mroczne odmęty phishingu – nie daj się złowić!*, przekł. R. Ociepa, Gliwice 2017 (tyt. oryg. *Phishing Dark Waters. The Offensive and Defensive Sides of Malicious E-mails*, Indianapolis, Ind. 2015).

[2] Ch. Hadnagy, P. Ekman, *Unmasking the Social Engineer. The Human Element of Security*, Indianapolis, Ind. 2014).

[3] Zob. na temat kursu: https://www.ekmaninternational.com/blink-and-you-will-miss-them-microexpressions-the-window-to-your-emotions (dostęp: 22.06.2019).

nież certyfikaty OSCP (*Offensive Security Certified Professional*) oraz OSWP (*Offensive Security Wireless Professional*)[4]. Owocem jego pracy jest między innymi założenie strony internetowej www.social-engineer.org, na której – wraz z innymi inżynierami społecznymi, badaczami i psychologami – odkrywa tajemnice aspektów związanych z socjotechniką. Prowadzi wykłady nie tylko dla studentów, ale również dla przedstawicieli firm i korporacji. W swoich działaniach podkreśla istotność wiedzy w przeciwdziałaniu zagrożeniom związanym z socjotechniką. Jego zdaniem kluczem do bezpieczeństwa jest właśnie edukacja.

Książka *Socjotechnika – sztuka zdobywania władzy nad umysłami* swoją strukturą przypomina poradnik skonstruowany dla osób mogących mieć styczność z inżynierią społeczną[5]. Może być odczytana jako kompendium wiedzy na temat socjotechnicznych sztuczek i sposobów ich wykorzystywania w życiu. Autor – jako artysta obeznany ze sztuką zdobywania władzy nad umysłami – udziela w swojej książce licznych bezpośrednich wskazówek dla osób, które mogą być narażone na ataki socjotechniczne. W książce znajdują się przykłady i szczegółowe opisy sytuacji, w których użyte zostały techniki manipulacyjne. Czytelnik dowiaduje się, na jakie zachowania zwracać uwagę oraz w jakich momentach trzeba być szczególnie ostrożnym, żeby nie stać się ofiarą manipulacji. Oprócz profilaktycznych wskazówek i edukacyjnego wymiaru książka stanowi również praktyczne źródło wiedzy na temat doskonalenia umiejętności socjotechnicznych oraz metod ich wykorzystywania. Autor opisuje narzędzia i techniki, ale skupia się także na przedstawieniu psychologicznych aspektów związanych z korzystaniem z dorobku inżynierii społecznej.

Socjotechnika jest nauką, która swoim zakresem obejmuje szeroki wachlarz dziedzin. Tajemnice socjotechniki zgłębiają zarówno psycholodzy, jak i politycy, socjolodzy, informatycy. W zależności od kontekstu naukowego definicja socjotechniki oraz metody stosowanych w jej ramach działań ulegają modyfikacjom.

W swojej książce Ch. Hadnagy opisuje działania i metody wykorzystywane przez socjotechników, posługując się przykładami osób odpowiedzialnych za wyszukiwanie oraz zdobywanie informacji od firm i korporacji. Hakerskie techniki pozyskiwania danych w doskonały sposób obrazuje definicja działań socjotechnicznych sformułowana przez Marka K. Mlickiego. Według niej działania socjotechniczne to „[…] więcej niż dwupodmiotowe działanie, w którym podmiot (system sterujący) skłania przedmiot (system sterowany) do zachowań zgodnych z celem lub celami systemu sterującego. Skłanianie to odbywa się przez emocjonalne i/lub intelektualne oddziaływanie na postawy i zachowa-

[4] Certyfikaty wydawane przez The Offensive Security Team, Industry-Leading Online Penetration Testing Training and Certification for Information Security Professionals – zob. https://www.offensive-security.com/ (dostęp: 20.06.2019).

[5] Stosowane zamiennie z pojęciem „socjotechnika".

nia"[6]. Odziaływanie na przedmiot autor książki dzieli na kilka poszczególnych etapów. Pierwszy z nich polega na gromadzeniu informacji. Ch. Hadnagy uświadamia czytelnika, jak ważne jest segregowanie i porządkowanie zdobytych informacji na temat ofiary działań socjotechnicznych. Każde dane mogą okazać się w przyszłości istotne do zrealizowania celu podmiotu. Dlatego warto jest mieć je wszystkie skumulowane w jednym miejscu. Pomocnymi narzędziami, wymienionymi i opisanymi w książce, są aplikacje BasKet i Dradis. Drugi z wymienionych programów jest szczególnie przydatny w przypadku pracy więcej niż jednego podmiotu, ponieważ pozwala na swobodny dostęp i modyfikacje danych każdemu z zaangażowanych. Kolejnym elementem strategii inżyniera społecznego jest wywoływanie. Ta praktyczna technika pozwala na nakłanianie odbiorcy komunikatu do podejmowania z własnej woli określonych działań, na których zależy socjotechnikowi. Ważne, by praktykując tę metodę pamiętać o subtelności w nadawaniu komunikatu. Warto zadawać ofierze pytania, uwzględniając jednocześnie zasady uprzejmości i okazując zainteresowanie. Pomocną techniką jest również wchodzenie w rolę. Poprzez dostosowanie swojego zachowania do odpowiedniego scenariusza podmiot jest w stanie pozyskać potrzebne informacje. Jest to budowanie zupełnie nowej, zmyślonej tożsamości i odgrywanie pewnej roli, żeby nakłonić ofiarę do zwierzeń.

Czytając książkę *Socjotechnika – sztuka zdobywania władzy nad umysłami* odnosi się wrażanie przebywania na wykładzie, podczas którego autor zdradza sekrety korzystania z metod socjotechniki oraz przestrzega przez zgubnymi skutkami i zagrożeniami płynącymi z bycia potencjalną jej ofiarą. Sam Ch. Hadnagy określa swoje literackie dokonanie jako formę prezentacji zdobytej przez siebie wiedzy. Dzieli się informacjami na temat sztuczek i narzędzi wykorzystywanych przez socjotechników oraz radzi, jak zapobiegać negatywnym skutkom ataków. Książka ta jest z pewnością przydatnym źródłem wiedzy dla hakerów, testerów zabezpieczeń, szpiegów, złodziei tożsamości, pracowników i szefów firm, sprzedawców, polityków rządzących państwami oraz lekarzy, psychologów, prawników[7]. Wiedza w niej zgromadzona może również ułatwić niewymienionym tu grupom odbiorców wchodzenie w kontakt z innymi. Można ją skutecznie stosować w różnych sytuacjach życia codziennego. Będąc świadomym mechanizmów socjotechniki i związanych z nimi zagrożeń, odbiorca staje się bardziej ostrożny w udzielaniu informacji oraz ujawnianiu i udostępnianiu swoich danych.

Chociaż techniki i narzędzia opisane w książce mogą wydawać się brutalne oraz mogą wzbudzać wątpliwości natury etycznej, to nie da się zaprzeczyć stwierdzeniu, że współcześnie

[6] M.K. Mlicki, *Socjotechnika. Zagadnienia etyczne i prakseologiczne*, Wrocław – Warszawa – Kraków – Gdańsk – Łódź 1986.

[7] Ch. Hadnagy, *Socjotechnika – sztuka zdobywania…*, s. 35–37.

metody wykorzystywane przez socjotechników są stosowane powszechnie, a jednostka jest nieustannie narażona na atak ze strony inżynierów społecznych. Sposoby wpływania na innych i „naciągania" można dostrzec zarówno w przekazach medialnych, jak i w życiu codziennym. Strategie marketingowe i kampanie polityczne są doskonałym przykładem wykorzystywania narzędzi inżynierii społecznej w celu wywarcia odpowiedniego wpływu na jednostki i grupy. O regułach, na których się opierają, pisze również Robert B. Caldini w książce *Wywieranie wpływu na ludzi – teoria i praktyka*. Wymienia on sześć podstawowych reguł wywierania wpływu na ludzi. Są to: reguła wzajemności, konsekwencji, społecznego dowodu słuszności, lubienia, autorytetu i niedostępności[8]. Działanie zgodnie z nimi wywołuje efekt automatycznego i bezrefleksyjnego ulegania ludzi, którzy są im poddawani[9]. W książce Hadnagy'ego role opisane przez R.B. Caldiniego zyskują praktyczny wymiar przy zastosowaniu techniki wywoływania. Autor, omawiając taktyki wywierania wpływu na ludzi, pisze również o: zobowiązaniu, długu wdzięczności, daniu czegoś od siebie, przedstawieniu swojej prośby, zobowiązaniu i ustępstwach. Celem tych zabiegów ma być wywołanie u odbiorcy poczucia „bycia dłużnikiem", dzięki czemu ma on być bardziej skłonny do udostępniania cennych informacji, żeby poczuć, że się odwdzięczył[10].

Socjotechniki analizowane przez Christophera Hadnagy'ego, jak podkreśla sam autor, mają służyć jako źródło wiedzy na temat inżynierii społecznej. Świadomość mechanizmów wykorzystywanych do manipulacji jest kluczem do zapewnienia bezpieczeństwa. Posiadając wiedzę na ich temat, jesteśmy świadomi realnych zagrożeń oraz możemy przedsięwziąć odpowiednie kroki w celu zapobiegania negatywnym skutkom wykorzystywania informacji. W książce znajdziemy wiele elementów opisujących praktyczne porady o tym, jak zdobywać dane interesujących nas podmiotów. Możemy przeczytać o grzebaniu w śmietnikach w celu znalezienia dokumentów, oszustwach i manipulacjach, a nawet o sposobach otwierania zamków wytrychem. Te kontrowersyjne metody są narzędziami używanymi przez socjotechników. Jak jednak pisze Ch. Hadnagy: „Kluczowe znaczenie dla zapobiegania atakom socjotechnicznym i ograniczania szkód z nimi związanych ma świadomość możliwości zaistnienia takiego zdarzenia"[11].

8 R.B. Caldini, *Wywieranie wpływu na ludzi – teoria i praktyka*, Sopot 2016, s. 13.

9 Tamże, s. 13.

10 Ch. Hadnagy, *Socjotechnika – sztuka zdobywania…*, rozdz. *Wywieranie wpływu, czyli siła perswazji*.

11 Tamże, rozdz. *Zapobieganie atakom socjotechnicznym i ograniczanie ich skutków*, s. 394.

STRESZCZENIE

Social engineering to pojęcie tożsame z rodzimym pojęciem socjotechniki i związane z manipulacją i dezinformacją. Recenzja książki Christophera Hadnagy *Socjotechnika – sztuka zdobywania władzy nad umysłami* stanowi, zgodnie z zasadami sztuki, zestaw subiektywnych ocen autorki. Oceny te dotyczą w pierwszej kolejności wrażeń z lektury, a wsparte zostały własnym krytycznym obrazem świata, w którym inżynieria społeczna znajduje zastosowanie w codziennym życiu każdego z nas. Autorka recenzji zwraca uwagę, że pozycja ma także walor aplikacyjny, albowiem wyższy poziom świadomości i wiedzy na dany temat przekłada się w prosty sposób na bezpieczeństwo jednostki.

Anna Maria Złocka

CHRISTOPHER HADNAGY, SOCIAL ENGINEERING. THE ART OF HUMAN HACKING

Social engineering is the same concept as the concept of social engineering and related to manipulation and disinformation. Review of Christopher Hadnagy's book *Social engineering – the art of gaining power over minds* is, according to the principles of art, a set of subjective assessments by the author. These assessments relate first to the impressions of reading, and are supported by their own critical image of the world in which social engineering is applied in the everyday life of each of us. The author of the review points out that the position has an application value, because a higher level of awareness and knowledge on a given topic translates easily into the individual's safety.

KEY WORDS: *social engineering, manipulation, information*

STUDIA POLITOLOGICZNE VOL. 54
DOI: 10.33896/SPolit.2019.54.13

Питер фрейз (Peter frase)

Четыре сценария будущего: Жизнь после капитализма

(Cztery przyszłości. Wizje świata po kapitalizmie)

PWN
Warszawa 2018, 130 s.
tłumaczenie Maciej Szlinder

(Иванна Килюшик)
ORCID: 0000-0001-5347-649X)

Ключевые слова:
капитализм, автоматизация, будущее, утопия, антиутопия

Основным тезисом книги является утверждение, что капитализм такой, каковым он есть сейчас, закончится, основной причиной чего является массовая автоматизация, которая лишит большинство людей рабочих мест. Технологические изменения, которые мы можем наблюдать в настоящее время, направлены на полную автоматизацию. Автор принимает автоматизацию за постоянную, а экономический кризис и классовые отношения как переменную. Питер Фрейз утверждает, что автоматизация требует другой экономической системы, чем капитализм, потому что большинство профессий в недалеком будущем будут выполнять машины.

Таким образом, появляется ряд вопросов о том, как будет выглядеть будущее после автоматизации и климатических изменений? Кто выиграет и заработает на автоматизированной экономике? Как мы справимся с нынешними массовыми неровностями имущества, доходов и политической власти в мире? Как подойти к распределению благ, когда экологическая катастрофа может значительно ограничить их количество? Питер Фрейз пытается найти ответы именно на эти и другие вопросы об автоматизированном будущем перед лицом климатических изменений.

Чтобы рассмотреть тему будущего человечества, автор представляет четыре возможные сценария: комму-

низм – равенство и изобилие; рентизм – иерархия и изобилие; социализм – равенство и дефицит; экстремизм – иерархия и дефицит.

Все сценарии объединяет полная автоматизация. Из этих четырех сценариев будущего состоит конструкция книги, каждый сценарий автор представляет в отдельном разделе.

Итак, Питер Фрейз представляет читателю две утопии и две антиутопии. Автор умышленно выделяет такие, а не другие идеальные схемы и максимально все упрощает, как сам заявляет в книге, чтобы подчеркнуть возникающие в мире социальные тенденции. В некотором смысле его сценарии утопий и антиутопий – это история о виртуальном будущем, которое может ожидать человечество в недалеком времени.

Питер Фрейз делает умозаключение, что как бы не сложилась ситуация у мирового общества будет две возможности (сценария) для развития и исходя из этого улучшения своего существования и две возможности (сценария) для деградации и исходя из этого погружения в конфликты и разрушение. Все возможные пути, по которым пойдет человечество будут сопровождаться полной автоматизацией всех отраслей производства. То, что отличает эти пути между собой – это политический строй и соответственно классовые отношения (кто будет владеть новыми технологиями и распределять блага), а также климатические изменения, которые влияют на среду обитания человечества и запасы ресурсов необходимых для производства роботов.

В первом сценарии будущего автор представляет коммунизм. Чтобы понять, что такое коммунизм, Питер Фрейз предлагает не смотреть на современный Китай или Северную Корею, а обратится к первоначальному значению этого слова. С этой целью он ссылается на Маркса, который под коммунизмом подразумевал собственно не тоталитарный или авторитарный режим правления, а идеальную жизнь, которую мировое общество может достичь путем социальных и технологических изменений. Главной особенностью коммунистического общества, которое будет эгалитарным, является то, что никто не будет работать только для того чтобы выжить, как это сложилось сейчас. Это в свою очередь станет возможным при условии, что всю роботу будут выполнять роботы. Автор, для того чтобы подчеркнуть данную мысль, цитирует известное высказывание Маркса: „От каждого по способностям, каждому по потребностям”.

Питер Фрейз аргументирует, что коммунизм как идеальное будущее человечества вполне возможно, как раз благодаря массовой автоматизации производства, которое будет работать на чистых неограниченных источниках энергии. Человечество в таком будущем не будет зависеть от газа и нефти и потребности работать, чтобы существоватьНо здесь возникает другой, не менее сложный вопрос – как жить в мире, где нет необходимости в большом количестве сотрудников. Следовательно, как выражать свою идентичность во времена, когда не будет уже профессий? Питер Фрейз рассматривает от-

вет на этот вопрос в упрощенных категориях в духе самого Маркса: делать почти все то же, что и обычно, но вместо того, чтобы работать для выживания, работать бесплатно на благо всего общества – community work.

Автор отмечает, что переход к коммунизму будет очень сложным и даже невозможным, потому что мировая капиталистическая элита стремится сохранить свой капитал, мировое положение и образ жизни к которому она привыкла.

Вторая версия будущего Питера Фрейза – рентизм, который ожидает нас в случае, если капиталистическая элита сохранит свое доминирование в полностью автоматизированной экономике. Рентизм – это будущее человечества, в котором может присутствовать всеобщий достаток, но это будет зависеть от элиты, которая монополизирует способы производства роботов с помощью авторских прав. В этом варианте будущего почти всю роботу тоже будут выполнять роботы. Автор отмечает, что останется только несколько профессий, а именно в сфере инноваций и программирования, юридической (специалисты з интеллектуальной собственности), правоохранительные органы, которые будут контролировать использование роботов и защищать технологии перед бедными массами людей, которые не будут в состоянии за них заплатить. В целом, рентизм будет способствовать стагнации, поскольку экономика нуждается в потребителях, а безработные по определению не могут быть потребителями.

Следующим, третьим видением автора будущего является система, которая напоминает социализм. В этом будущем произведенные блага распределяются между всеми гражданами с помощью некого основного дохода. Проблема в том, что предоставление всем людям денег, позволяющих выжить, делает их переговорную силу на рынке труда определенно улучшенной (в нынешней ситуации большинство людей без работы не сможет накопить средств для выживания, поэтому дефакто работа является принуждением). По мнению Питера Фрейза, это не нравится элитам, которые проявляют сильное сопротивление таким идеям.

Последним возможным сценарием является экстремизм. В таком будущем человечество будет обладать богатством, вытекающим из всеобщей роботизации, но будет лишено эгалитаризма. Богатая элита будет жить в укрепленных анклавах, а другие будут жить в бедности (чтобы увидеть, как это будущее может выглядеть автор приводит в качестве примера фильм „Элизиум” из 2013 г.). В этом варианте будущего человечества, истощенные массы бедных людей будут постоянно представлять угрозу для небольшой группы элиты. Элита может поделиться с бедными частью своего богатства в обмен за спокойствие. Однако эти действия подвергнут их к бесконечным требованиям со стороны бедной части общества. Автор рассуждает, что элиты в какой-то момент могут прийти к выводу, что лучшим решением этой проблемы будет истребление этой части общества.

Питер Фрейз считает, что в случае полной автоматизации производства, наиболее вероятным будущем человечества является четвертый вариант, а именно экстремизм.

Следует отметить, что определение типа публикации составляет некую трудность, ведь ее наравне можно отнести так к научной литературе, как и к художественной. Объясняется это тем, что автор в книге с легкостью ссылается, с одной стороны, на Курта Воннегута и Кори Доктороу и такие фильмы, как „Элизиум" и „Игра Эндера", с другой же стороны, на теории политической экономики Джона Мейнарда Кейнса, Андре Горца, Карла Маркса, Василия Леонтьева. Питер Фрейз называет свою книгу своего рода социальной научной фантастикой, поскольку он использует инструменты социальных наук в сочетании с инструментами, используемыми художественной прозой. Автор утверждает, что использование такой комбинации инструментов позволят лучше исследовать пространство возможностей. Питер Фрейз также подчеркивает, что книга – это попытка представить возможности, а не вероятность будущего, потому что оценка возможностей ставит коллективное действие человечества в центр внимания, в то время как вероятности как определенные прогнозы поддерживают пассивность.

Может показаться, что книга слишком коротка, чтобы полностью рассмотреть тему посткапиталистического будущего, поскольку она имеет только около 130 страниц и все же данный труд исчерпывает тему. Хорошо, что автор не входит в подробности, описывающие то, почему в будущем мир ждет массовая автоматизация, потому что тема была подробно описана в других публикациях, достаточно упомянуть книги «Гонка с машиной» (Race Against the Machine) Эрика Бринолфссона и Эндрю Макафи или «Роботы наступают» (Rise of the Robots) Мартина Форда. В заключении, необходимо добавить, что сценарии, изложенные автором, хорошо аргументированные.

Недостатком книги является то, что автор, ссылаясь на футурологические книги и фильмы, делает ее менее достоверной в глазах читателя.

В целом, книга является ценным произведением, потому что обращает внимание на основные проблемы недалекого будущего человечества и является сигналом для размышления над его формой.

Резюме

Автор книги рассматривает проблемы быстрой автоматизации производства и кризиса капитализма. Он исследует пространство возможностей и размышляет о видении будущего, представляя две утопии и две антиутопии.

Иванна Килюшик

Peter Frase, Four Futures. Life After Capitalism

The author of the book raises the problem of fast automation of production and the crisis of capitalism. He analyzes the space of possibilities and wonders about the vision of the future, presenting two utopias and two anti-utopias.

Key words: *capitalism, automation, future, utopia, anti-utopia*

VOL. 54 STUDIA POLITOLOGICZNE
DOI: 10.33896/SPolit.2019.54.14

ANDRZEJ WIERZBICKI

Polish-Belarusian Relations, Between a Common Past and the Future

Nomos Verlagsgesellschaft
Baden-Baden, Germany 2018, 204 c.

(Яна Волчецкая)
ORCID: 0000-0003-4233-9849)

Ключевые слова:
Республика Польша, Беларусь, двусторонние отношения

Двусторонние отношения Республики Польша с Республикой Беларусь характеризуются относительно низким уровнем активности по сравнению с другими восточными соседями. Казалось бы, общее польско-белорусское историко-культурное наследие, начиная со времен Речи Посполитой, географическое положение, а также отсутствие значимых конфликтов на национальном уровне могли бы стать основой для стабильных и интенсивных отношений двух стран. На фоне гораздо более сложных российско-польских и украинско-польских отношений, обремененных историческим грузом и геополитическим балластом, двусторонние отношения между Республикой Беларусь и Республикой Польша не являются примером доброго сотрудничества стран, связанных общей историей и культурой. За последние десятилетия в отношениях двух стран не были зафиксированы значительные события, повлиявшие на улучшение сотрудничества между странами. Для восточной политики Республики Польша отношения с Республикой Беларусь остаются одним из самых больших вызовов современной политики.

Монография *Polish-Belarusian Relations, Between a Common Past and the Future* (*Польско-белорусские отношения: между общим прошлым и будущим*) Анджея Вержбицкого – известного польского ученого, исследователя, специалиста в области межэтнических отношений, профессор Института Политических Наук и Международных Исследований Варшавского университета была издана в Баден-Баден (Германия) в 2018 году. Монография проф. Анджея Вежбицкого является по-

пыткой понять двусторонние отношения между Республикой Беларусь и Польшей с учетом исторической и культурной перспективы, восточной политики, политических и экономических отношений, а также сложных аспектов, затрагивающих отношения между двумя странами.

Принятая автором научно-исследовательская концепция носит оригинальный характер и до сих пор не использовалась в польской научной литературе. Автор интерпретирует результаты проведенных исследований, принимая во внимание различные перспективы научных исследований и мнения различных сторон. Следует обратить особое внимание на источники, представленные автором, литературу российских, польских и иностранных авторов, многочисленные статистические данные, отчеты, экспертизы, документы, договоры и двусторонние соглашения. Представленная литература является авторитетной.

Монография состоит из введения, шести глав, заключения, а также списка использованных источников и литературы. Целью данной монографии является ответ на основной исследовательский вопрос, а именно, почему мы не можем считать, что нынешние польско-белорусские отношения являются образцовыми, напротив, они характеризуются взаимным подозрением и недоверием друг к другу в отношении истинных намерений. Следует согласиться с автором, что причины такого положения дел нужно искать в геополитическом контексте, который состоит из восточной политики реализуемой Польшей и Европейским Союзом в целом. Важным фактором также остаются тесные отношения Беларуси и России, которые рассматриваются Польшей в контексте исторической политики. Учитывая все факторы, стоит отметить, что несмотря на то, что отношения между двумя странами в ходе истории не были простыми, никогда не наблюдалась открытая враждебность.

В первой главе автор показывает исторические и культурные факторы повлиявшие на двусторонние отношения начиная с периода Великого Княжества Литовского до сегодняшнего дня. В главе в подробностях описаны важнейшие даты и события, в течение которых строились взаимоотношения. Тут можно обратить особое внимание на фактор идентичности белорусов и поляков, а также культурные различия между народами. Важнейшим фактором остается вопрос религиозной принадлежности: на региональном уровне распространено мнение, что белорус значит православный, а поляк – католик. Стоит учитывать, что в приграничных регионах Беларуси, в Брестской и Гродненской областях, процент населения, которое исповедуют католицизм достаточно высок, данный фактор связывает два народа и вносит свой вклад в отношения на местном уровне.

Вторая глава посвящена польско-белорусским отношениям с точки зрения восточной политики Польши. В данном разделе автор шаг за шагом показал, как проходила трансформация внешней политики Польши и Беларуси. В главе также представлены системные условия внешней политики обеих стран,

геополитическое положение. Стоит подчеркнуть, что именно геополитическое положение влияет на отношения двух стран. Польша является членом Европейского Союза, в то время как Беларусь активно участвует в интеграционных проектах, реализуемых Россией. В главе также представлена таблица, отражающая социально-экономический потенциал Польши и Беларуси.

В третьей главе «Политические отношения» автор показывает эволюцию польско-белорусских отношений, учитывая политический контекст. В моменте получения независимости в Республике Беларусь существовала реальная возможность к появлению стабильных и эффективных двусторонних отношений между равноправными странами. Несмотря на эти позитивные перспективы и отсутствие серьезных конфликтов в прошлом начало польско-белорусских отношений было сложным. Отношения проходили через период «критического диалога», определяемого внутренней и внешней политикой Беларуси, а также значимым фактором было вступление Польши в структуры ЕС. Монография была издана год назад и учитывает также события, имеющие место на Украине в последние годы.

Четвертая глава «Сложные вопросы» посвящена таким общим проблемам, как исторический диалог, национальные меньшинства – поляки в Беларуси и белорусы в Польше, пограничное и визовое движение. Глава была дополнена статистическими данными о численности поляков в Беларуси и белорусов в Польше, а также треугольником взаимодействия Брубейкера, анализ которого провел автор.

Пятая глава, посвященная экономическим отношениям, включает такие вопросы, как правовая и договорная основа экономического сотрудничества, торговля, инвестиции и сотрудничество в сфере капитала, трансграничное сотрудничество, энергетика, а также другие области и перспективы и препятствия для сотрудничества. Данная глава вносит большой вклад в монографию, потому что благодаря ей можно увидеть, что Беларусь и Польша несмотря на различные политические и социально-экономические системы могут сотрудничать на экономическом уровне. Автор особое внимание уделил взаимоотношениям стран в области торговли. Стоит подчеркнуть, что Беларусь является частью Евразийского экономического союза, в то время как Польша – Евросоюза, данный фактор создает все необходимые условия для взаимовыгодного сотрудничества.

Последняя, шестая глава «Культурное сотрудничество» касается взаимодействия и сотрудничества в области культуры и образования. Сотрудничество чаще всего осуществляется в приграничных регионах, как пример можно представить программу Белорусско-польского экономического форума «Добрососедство». В конце монографии представлены выводы и рекомендации.

Отношения между Беларусью и Польшей являются одними из наиболее важных в Восточной Европе, но при этом одними из наиболее сложных. Белорусско-польские от-

ношения прошедшие через периоды подъёма и упадка на современном этапе имеют большой потенциал. Монография, написанная профессором Анджеем Вежбицким, является комплексным анализом процессов, происходящих между двумя странами. В монографии рассматриваются наиболее важные аспекты, влияющие на двусторонние отношения. В польской и европейской научной литературе существует мало исследований на тему отношений между Республикой Беларусь и Польшей, поэтому для понимания современных отношений двух стран, с учетом исторической перспективы, данная публикация является обязательной к прочтению.

Резюме

Автор книги рассматривает проблемы двусторонних отношений Республики Польша с Республикой Беларусь. В нем представлен общий обзор истории двух стран, в том числе многовековые связи между народами, которые прошли различные этапы двусторонних отношений.

Yana Valcheyskaya

The monograph describes the stages of how bilateral relations were taking shape between the Republic of Poland and the Republic of Belarus. A general overview of the history of the two countries is offered, considering centuries-old ties between the nations that have gone through different stages in relations. Political and economic relations, the Eastern Policy of the Republic of Poland, trade and investment aspects, cultural cooperation and a number of other matters pertaining to the good-neighborly dialogue between the two countries were considered.

Key words: *Republic of Poland, Belarus, bilateral relations*

Autorzy

BARTOSZ BIDERMAN, (bartosz.biderman@protonmail.ch), absolwent Wydziału Nauk Politycznych i Studiów Międzynarodowych Uniwersytetu Warszawskiego. Publikuje w czasopismach naukowych z zakresu bezpieczeństwa, udziela się na konferencjach politologicznych. Zainteresowania badawcze – dezinformacja oraz wywiad jawnoźródłowy w Internecie.

PIOTR DELA, płk dr hab. inż., absolwent Wydziału Cybernetyki Wojskowej Akademii Technicznej w Warszawie, studiów podyplomowych w Akademii Obrony Narodowej w Warszawie i studiów podyplomowych w Instytucie Polityki Światowej w Waszyngtonie. Od początku pracy zawodowej związany z instytucjami centralnymi MON i szkolnictwem wojskowym. Od 1998 roku pracownik naukowo-dydaktyczny Akademii Obrony Narodowej w Warszawie (obecnie Akademii Sztuki Wojennej), gdzie w 2001 roku uzyskał stopień naukowy doktora, a następnie w 2012 roku stopień naukowy doktora habilitowanego. Autor ponad 100 artykułów, referatów, podręczników, opracowań naukowych i monografii. Obszar zainteresowań naukowych obejmuje: systemy wspomagania decyzji, systemy i sieci teleinformatyczne, bezpieczeństwo informacyjne, bezpieczeństwo cyberprzestrzeni.

KONRAD GAŁUSZKO, (k.galuszko@student.uw.edu.pl), pracuje w branży badawczej w Centrum Badań Marketingowych INDICATOR. Do swoich zainteresowań badawczych zalicza: metodologię badań socjologicznych, analizę danych.

JÓZEF GRZEGORCZYK, (jozef@sekkura.com.pl), absolwent Wyższej Oficerskiej Szkoły Radiotechnicznej w Jeleniej Górze (1974), Kursu Przeszkolenia Oficerów (1976), odbył służbę w Wojsku Polskim w latach 1970–2002 (w tym udział w misjach pokojowych: UNIFIL 1992, 1998, UNDOF 1994, SFOR 2000–2002). Pracownik Ministerstwa Finansów w latach 2002–2006. Obecnie pracuje w prywatnej spółce na rzecz obronności kraju. Zainteresowania: fotografia, wykrywanie śladowych ilości materiałów przy wykorzystaniu technologii IMS.

PATRYCJA HRABIEC-HOJDA, (patrycjahrabiec@gmail.com), doktorantka na Uniwersytecie Jagiellońskim w zakresie informatologii. Bada temat zjawiska infobrokeringu w różnych krajach. Na co dzień pracuje w agencji infobrokerskiej. Ekspert w zakresie kształcenia brokerów informacji i kompetencji

z zakresu OSINT. Prezes Zarządu Stowarzyszenia Profesjonalistów Informacji. Redaktor naczelna portalu Rynek informacji.

URSZULA KURCEWICZ, (uurban@uw.edu.pl), dr, pracownik Instytutu Nauk Politycznych Uniwersytetu Warszawskiego w Katedrze Historii Politycznej. Współtwórczyni specjalizacji Infobrokering polityczny. W latach 2007–2012 pracowała w charakterze eksperta w Polskiej Agencji Prasowej, uczestnicząc w międzynarodowych projektach ochrony dziedzictwa kulturowego. Autorka publikacji z zakresu najnowszej historii politycznej, procesów migracyjnych i komunikacji społecznej.

KONRAD KRYSTIAN KUŹMA, (kkuzma@aps.edu.pl), pracownik i doktorant Akademii Pedagogiki Specjalnej im. Marii Grzegorzewskiej w Warszawie. Od 2005 roku związany z badaniami marketingowymi, a od 2008 – społecznymi i ewaluacyjnymi. W swojej pracy zajmował się między innymi weryfikacją jakości danych, ewaluacją prowadzonych badań oraz jakością doboru prób badawczych. Jego zainteresowania badawcze, poza metodologią i metodami badań społecznych, obejmują socjologię kulturową, socjologię miasta oraz badanie jawnych i ukrytych społeczności internetowych.

JOANNA LEWCZUK, (jlewczuk@indicator.pl), absolwentka Wydziału Matematyki i Informatyki Uniwersytetu Warszawskiego oraz Uniwersytetu Łódzkiego. Z branżą badawczą związana od 2001 roku. Obecnie jako analityk specjalizuje się w realizacji projektów naukowych – od badań statutowych po granty badawcze finansowane przez Narodowe Centrum Nauki. Redaktorka statystyczna publikacji naukowych. Interesuje się metodyką nauczania matematyki, analizami statystycznymi oraz kulturą koreańską.

DANIEL MIDER, (d.mider@uw.edu.pl), dr hab. (2008 – doktorat *summa cum laude*, 2018 – habilitacja). Wykładowca na Wydziale Nauk Politycznych i Studiów Międzynarodowych Uniwersytetu Warszawskiego oraz w Szkole Głównej Handlowej, certyfikowany informatyk śledczy. Współtwórca i wykładowca specjalności *Infobrokering polityczny*. Autor licznych publikacji z zakresu socjologii Internetu, socjologii przemocy, metodologii badań oraz partycypacji politycznej.

WOJCIECH MINCEWICZ, (w.mincewicz@uw.edu.pl), mgr politolog, socjolog. Ukończył z wyróżnieniem studia na Wydziale Nauk Politycznych i Studiów Międzynarodowych Uniwersytetu Warszawskiego, ze specjalnością infobrokering polityczny oraz Wydziale Stosowanych Nauk Społecznych i Resocjalizacji – specjalizacja organizacje pozarządowe, współpraca międzysektorowa. Obecnie doktorant WNPiSM UW w dziedzinie nauki o polityce. Jego zainteresowania naukowe obejmują zagadnienia z zakresu socjologii, kultury i przywództwa politycznego, infobrokeringu oraz bezpieczeństwa w cyberprzestrzeni.

PRZEMYSŁAW POTOCKI, (p.potocki2@uw.edu.pl), dr, pracownik naukowo-dydaktyczny w Katedrze Ustroju Pracy i Rynku Pracy Wydziału Nauk Politycznych i Studiów Międzynarodowych Uniwersytetu Warszawskiego. Zainteresowania badawcze: marketing polityczny, społeczeństwo informacyjne, gospodarka cyfrowa, rynek pracy, zachowania wyborcze, komunikowanie publiczne.

PIOTR SOSNOWSKI, (p.sosnowski044@gmail.com), absolwent studiów cywilnych na Akademii Obrony Narodowej w Warszawie. Jego zainteresowania badawcze dotyczą kurdologii, bezpieczeństwa informacyjnego i polityki bezpieczeństwa państw.

MAGDALENA TOMASZEWSKA-MICHALAK, (tomaszewska.m@uw.edu.pl), dr, absolwentka Wydziału Prawa i Administracji Uniwersytetu Warszawskiego. Doktoryzowała się w międzywydziałowej jednostce Centrum Nauk Sądowych UW. Obecnie pracuje jako adiunkt w Katedrze Nauk o Bezpieczeństwie na Wydziale Nauk Politycznych i Studiów Międzynarodowych.

PAWEŁ TOMCZYK, (pawel@sekkura.com.pl), absolwent Wojskowej Akademii Technicznej i Akademii Sztabu Generalnego (kierunek: rozpoznanie), wyższy oficer Wojska Polskiego (przeniesiony do rezerwy), służbę odbywał w jednostkach liniowych wojsk zmechanizowanych, a następnie w Instytucjach Centralnych Ministerstwa Obrony Narodowej. Brał udział w misji SFOR w Bośni i Hercegowinie. Specjalista w zakresie ochrony informacji niejawnych. Instruktor sztuk walki, licencjonowany detektyw.

JUSTYNA TRZECIAKOWSKA, (jtrzeciakowska@infobrokerska.pl), ekspert w zakresie białego wywiadu, broker informacji z 10-letnim doświadczeniem. Na co dzień pracuje w agencji Infobrokerska.pl. Prezes Fundacji na rzecz rozwoju rynku informacji i społeczeństwa informacyjnego: Infobrokerska. Redaktor prowadząca portalu Rynek informacji. Autorka standardów zawodu specjalista analizy rynku.

ANNA MARIA ZŁOCKA, (mariazlo@o2.pl), członkini Uniwersytetu Zaangażowanego, a także uczestniczka programu Service Learning. Zainteresowania naukowe obejmują animację społeczną i animację kultury oraz antropologię kultury.

ЯНА ВОЛЬЧЕЦКАЯ, (valchetskayajana@gmail.com), магистр, Украина.

ИВАННА КИЛЮШИК, (ivannakyl@gmail.com), ввыпускница Восточноевропейского национального университета имени Леси Украинки, Украина.

STUDIA POLITOLOGICZNE
(wskazówki dla Autorów)

Forma przekazania tekstu: e-mailem, w edytorze Word (na adres sekretarza redakcji: zalesnyjacek@gmail.com).
Do tekstu dołącza się oświadczenie o oryginalności pracy oraz o tym, że aktualnie nie uczestniczy ona w innym postępowaniu wydawniczym.

Redakcja tekstu

Układ analizy:

Autor
Numer ORCID
Tytułu analizy w języku polskim
Abstrakt: w języku polskim do 600 znaków
Kluczowe słowa: 5 w języku polskim
Struktura analizy:
Wprowadzenie: uzasadnienie wyboru tematu i jego nowatorskość, cele analizy, hipotezy i tezy badawcze, zastosowane metody badawcze
– analiza
– konkluzje (wnioski)
Streszczenie w języku polskim
Tytułu analizy w języku angielskim
Abstrakt: w języku angielskim do 600 znaków
Kluczowe słowa: 5 w języku angielskim.
Bibliografia w alfabecie łacińskim
Nota o Autorze (w tym: nazwa instytucji, w której jest zatrudniony, tytuł naukowy, stopień naukowy, adres e-mailowy).
W pracy stosuje się śródtytuły

Czcionka: Times New Roman, 13.
Akapit: wyrównanie do prawej i lewej, wcięcie: 1,25 cm pierwszy wiersz, 1,5 odstępu między wierszami, brak dodatkowego odstępu po akapicie.
Przypisy polskie: na dole strony, numeracja ciągła, czcionka 10, według wzoru:

[1] S. Huntington, *Trzecia fala demokratyzacji*, Warszawa 1995, s. 206.
[1] Tamże, s. 27.
[1] M. Cichosz, *Transformacja demokratyczna – przyczyny, przebieg i efekty procesu*, [w:] A. Antoszewski (red.), *Systemy polityczne Europy Środkowo-Wschodniej*, Wrocław 2006, s. 52.
[1] M. Castells, *Społeczeństwo sieci*, przekł. M. Marody i in., Warszawa 2008, s. 15.
[1] S. Huntington, *Trzecia fala…*, s. 176.
[1] T. Kowalski, *Formy i przesłanki obecności kapitału zagranicznego w mediach drukowanych*, „Zeszyty Prasoznawcze" 1998, nr 1–2, s. 37–38.
[1] M. Górak, *Cyfrowa prasa: chwilowa moda czy przyszłość*, http://internetstandard.pl/artykuły/45301.html (dostęp: 6.12.2004).

Tekst podstawowy i przypisy: wyjustowane.
Ustawienia strony: standardowe.
Objętość: 25–35 tys. znaków (wraz ze spacjami).

W celu przeciwdziałania *„ghostwriting"* i *„guest authorship"* Redakcja „Studiów Politologicznych" wprowadziła procedury związane z zaporą *„ghostwriting"*.

Ghostwriting oraz *guest authorship* są przejawem nierzetelności naukowej. Wszelkie wykryte przypadki będą demaskowane, włącznie z powiadomieniem odpowiednich podmiotów (instytucje zatrudniające Autorów, towarzystwa naukowe, stowarzyszenia edytorów naukowych itp.).
Z *ghostwriting* mamy do czynienia wówczas, gdy ktoś wniósł istotny wkład w powstanie publikacji, bez ujawnienia swojego udziału jako jeden z Autorów lub bez wymienienia jego roli w podziękowaniach zamieszczonych w publikacji.
Z *guest authorship* (*honorary authorship*) mamy do czynienia wówczas, gdy udział Autora jest znikomy lub w ogóle nie miał miejsca, a pomimo to jest Autorem/współautorem publikacji.
Autor publikacji jest zobligowany poinformować o źródłach finansowania publikacji, wkładzie instytucji naukowo-badawczych, stowarzyszeń i innych podmiotów (*financial disclosure*).
Redakcja «Studiów Politologicznych» wymaga od Autorów publikacji ujawnienia wkładu poszczególnych Autorów w powstanie publikacji (z podaniem afiliacji oraz informacji, kto jest Autorem koncepcji, założeń, metod, protokołu itp. wykorzystywanych przy przygotowaniu publikacji). Autor ponosi odpowiedzialność za zgłoszoną publikację.
Redakcja «Studiów Politologicznych» dokumentuje wszelkie przejawy nierzetelności naukowej, zwłaszcza łamania i naruszania zasad etyki obowiązujących w nauce.

Teksty przekazywane do opublikowania w «Studiach Politologicznych» podlegają postępowaniu recenzyjnemu. W ciągu dwóch miesięcy od złożenia tekstu Autor jest informowany o zakwalifikowaniu go do postępowania recenzyjnego lub odrzuceniu ze względu na uchybienia formalne. Następnie każda praca (po nadaniu jej anonimowości) jest opiniowana przez jednego z Redaktorów „Studiów Politologicznych". Po uzyskaniu pozytywnej opinii tekst jest przekazywany dwóm recenzentom zewnętrznym, to jest spoza członków Redakcji. W przypadku uzyskania recenzji negatywnej informacja o tym fakcie jest podawana Autorowi, a postępowanie publikacyjne ulega zakończeniu ze skutkiem dlań negatywnym. W przypadku recenzji negatywnej Autor otrzymuje recenzję nadesłanego tekstu (po usunięciu personaliów recenzenta) oraz informację, że postępowanie publikacyjne uległo zakończeniu ze skutkiem negatywnym.

Redakcja nie zwraca tekstów niezamówionych oraz zastrzega sobie prawo do ich redagowania i skracania.

STUDIA POLITOLOGICZNE
(„ПОЛИТОЛОГИЧЕСКИЕ ИССЛЕДОВАНИЯ")

Указания для Авторов

Форма предоставления текстов (на русском языке): по электронной почте, в редакторе Word (на адрес секретаря «Политологических исследований»: zalesnyjacek@gmail.com).

К тексту прилагается заявление об оригинальности работы и о том, что на данное время она не заявлена в другие издания.

Редактирование текста

Схема статьи:

Автор
Номер ORCID
Название статьи на русском языке
Резюме: до 600 знаков на русском языке
Ключевые слова: 5 на русском языке
Текст статьи
Структура:
Введение: обоснование выбора темы и ее новизны, цели анализа, гипотезы и тезисы исследований, применяемые методы исследования
– анализ
– выводы
Название статьи на английском языке
Резюме: до 600 знаков на английском языке
Ключевые слова: 5 на английском языке
Библиография в латинском алфавите
Информация об авторе (наименование учреждения, в котором он работает, ученое звание, ученая степень).

Шрифт: Times New Roman «13»

Сноски: внизу страницы, непрерывная нумерация, шрифт «10», согласно образцу:

[1] И.В. Чубыкин, *Государственное управление стран ближнего зарубежья России*, Москва 2006, с. 99.
[1] Там же, с. 27.
[1] См.: Н. Дж. Мельвин, *Узбекистан: переход к авторитаризму на шелковом пути*, [в:] С.И. Кузнецова (ред.), *Страны Центральной Азии на рубеже XX–XXI веков: становление национальных государств*, Москва 2006, с. 78.
[1] А.С. Автономов, *Процесс становления парламентаризма в Казахстане*, «Представительная власть» 1995, № 2, с. 27.
[1] M. Górak, *Cyfrowa prasa: chwilowa moda czy przyszłość*, http://internetstandard.pl/artykuły/45301.html (дата обращения: 6.12.2004).

Параметры страницы: стандартные

Объем: 25–35 тыс. знаков с пробелами.

С целью противодействия «***ghostwriting***» и «***guest authorship***» редакция «Studiów Politologicznych» ввела процедуры, связанные с преградой «***ghostwriting***».
«***Ghostwriting***» и «***guest authorship***» являются проявлением научной недобросовестности. Все обнаруженные случаи будут разоблачены, включая уведомление соответственных субъектов (учреждений, в которых работают авторы, научные общества, сообщества научных редакторов и т.п.).
С «***ghostwriting***» имеем дело, когда кто-то внес весомый вклад в создание публикации, не сообщая о своем участии в роли соавтора либо без упоминания его роли в благодарностях, помещенных в публикации.
С «***guest authorship***» («***honorary authorship***») имеем дело, когда участие автора мизерно мало либо вообще отсутствует, и не смотря на это, он является автором/соавтором публикации.
Автор публикации обязан сообщить об источниках финансирования публикации, вкладе научно-исследовательских учреждений, обществ и других субъектов («*financial disclosure*»).
Редакция «Studiów Politologicznych» требует от авторов публикаций представления вклада всех конкретных авторов в создании публикации (с указанием аффилиации и данных, кто является автором концепции, основных тезисов, методов, протокола и т. п., использованных в подготовке публикации). Автор несет ответственность за заявленную публикацию.
Редакция «Studiów Politologicznych» документирует все проявления научной недобросовестности, в частности нарушения принципов этики, действующих в науке.

Тексты, направляемые для публикации в «Studiach Politologicznych», подлежат процессу рецензирования. В течение 2 месяцев с момента подачи текста автор уведомляется о том, что он допущен к процессу рецензирования либо не допущен в связи с формальными погрешностями. Далее каждая работа (после ее анонимизации) оценивается одним из редакторов «Studiów Politologicznych». После получения положительной оценки текст передается двум независимым рецензентам, не являющимся членами редакции. В случае отрицательной рецензии, данную информацию сообщают автору, а процесс публикации завершается с негативным результатом. В случае негативной рецензии автор получает рецензию на отправленный текст (после удаления имени рецензента) и информацию, что процесс публикации завершен с негативным для него результатом.

Редакция не возвращает не заказанных текстов и оставляет за собой право к их редактированию и сокращению.

STUDIA POLITOLOGICZNE
("POLITICAL SCIENCE STUDIES")

Instructions for Authors

Manuscripts should be submitted by email in Word format to the Secretary of "Political Science Studies": zalesnyjacek@gmail.com

A declaration confirming the original character of the paper and that it is not under consideration for publication elsewhere must be included.

Editing of the text

Structure of the paper:

A scheme of the analysis:

Author(s)
ORCID number
Manuscript title
Abstract (up to 600 characters)
Key words: up to 5
Body of the the manuscript
Introduction: justification of the research and its novelty, objectives of the analysis, hypotheses and research theses, applied research methods
– analysis
– conclusions
Bibliography
A short note about the Author(s) is also required (including the name of the institution where they are employed, the academic title and academic degree).

Font: 13-point font size (Times New Roman)

References: at the bottom of the page, continuous pagination, 10-point font size, according to the following model:

[1] F. Millard, *Elections, Parties and Representation in Post-Communist Europe*, Palgrave Macmillan 2004, p. 135.
[1] Ibidem, p. 27.
[1] T. Zittel, *Legislators and their representational roles: strategic choices or habits of the heart?*, [in:] M. Blomgren, O. Rozenberg (eds.), *Parliamentary Roles in Modern Legislatures*, Routledge 2012, p. 107.
[1] F. Millard, *Elections, Parties…*, p. 176.
[1] A. Grant, *The Politics of American Campaign Finance*, "Parliamentary Affairs" 1998, № 2, p. 227.
[1] M. Górak, *Cyfrowa prasa: chwilowa moda czy przyszłość*, http://internetstandard.pl/artykuły/45301.html (access: 6.12.2004).

Page setup: standard

Length: 25,000–35,000 characters (spaces included).

To counteract „***ghostwriting***" and „***guest authorship***", the Editorial Board of «Studia Politologiczne» has procedures to block „***ghostwriting***" and unacknowledged guest authorship. „***Ghostwriting***" and „***guest authorship***" are scientifically unreliable and dishonest. All detected cases will be disclosed, including notifying the proper entities (institutions employing the Authors, scientific societies, associations of scientific editors, etc.).
We are dealing with „***ghostwriting***" when a person who has made a significant contribution to the manuscript does not disclose their participation as one of the Authors or when their role is not mentioned in the Acknowledgements included in the publication.
We are dealing with „***guest authorship***" („***honorary authorship***") when the Author's participation is negligible or none despite the fact they are listed as an Author/co-author of the publication.
The Author(s) is obliged to disclose the sources of financing for the publication, such as research grants from scientific and research institutions, associations and other entities („*financial disclosure*").
The Editorial Board of "Studia Politologiczne" requires all Authors with reasonable claims to authorship to be named, including their institutional affiliations. The Authors should declare their contribution(s) to the manuscript including the development of the concept(s), assumption(s), method(s), protocols for data analysis, interpretation and conclusions. This information can be provided in a separate note (email) to the Secretary of "Political Science Studies".

The Editorial Board of "Studia Politologiczne" documents all signs of scientific unreliability, especially of breaking and infringing the principles of ethics binding in science. Manuscripts submitted for publication in "Studia Politologiczne" are subject to a double-blind review. Within two months from the time of submission, the Author is advised if their paper has been accepted for review or rejected due to formal faults. Next, each manuscript (after being anonymized) is assessed by one of the Editors of "Studia Politologiczne". After receiving a positive opinion, it is then passed on to two external reviewers, i.e. from outside the Editorial Board. In case of a negative review, that is, the manuscript being rejected for publication, the Author is advised accordingly and the paper, along with the anonymized reviewers' feedback, is returned to them.

The Editorial Board does not return the manuscripts which have not been requested and reserves the right to edit and abridge them.