

Wojciech Myszka*

Książki elektroniczne (stan technologii)

1. Trochę historii

W 1971 roku Michael Hart otrzymał na komputerze Xerox Sigma V będącym własnością Laboratorium Badań Materiałów Uniwersytetu Illinois konto komputerowe z prawem wykorzystania czasu procesora o wartości 100 000 000 \$. Były to czasy, kiedy łatwiej było uzyskać dostęp do komputera niż znaleźć ludzi, którzy wiedzieli, jak go wykorzystać. Około 100 minut po tym niespodziewanym dla Harta wydarzeniu właśnie on wiedział, jak tę możliwość wykorzystać: nie do prowadzenia obliczeń, ale do przechowywania, odtwarzania i wyszukiwania informacji w zasobach, które normalnie znajdują się w bibliotekach.

Następnym krokiem było zdigitalizowanie (przepisanie) Deklaracji Niepodległości Stanów Zjednoczonych Ameryki Północnej. Rozpoczęła się era książek elektronicznych. Kolejnymi digitalizowanymi książkami były Konstytucja Stanów Zjednoczonych wraz z Kartą Praw Stanów Zjednoczonych Ameryki, Biblia i dzieła Szekspira [6]. Wszystkie te książki zostały udostępnione publicznie jako Projekt Gutenberg, który w 1997 roku oferował 1000 tytułów, w kwietniu 2002 — 5000, w październiku 2003 — 10000, w styczniu 2005 — 15000, w kwietniu 2008 — 25000 książek. Obecnie jest ich ponad 40000, każdego miesiąca przybywa ponad 300 książek [7].

Projekt Gutenberg zaproponował (i praktycznie rozwiązał) wszystkie problemy pojawiające się podczas posługiwania się treściami elektronicznymi:

- format udostępniania informacji (początkowo jedynie pliki tekstowe kodowane jako „vanilla ASCII”),
- zdobywanie treści (przepisywanie ręczne książek, skanowanie, OCR i sprawdzanie poprawności),
- prawa autorskie — udostępniane są jedynie książki, do których prawa autorskie wygasły.

Początkowo głównym problemem było czytanie tych książek (czy raczej brak powszechnego dostępu do komputera). Zmieniało się to bardzo szybko, gdy

* dr inż. Wojciech Myszka, Dolnośląska Wyższa Szkoła Przedsiębiorczości i Techniki w Polkowicach, Politechnika Wrocławska, Instytut Materiałoznawstwa i Mechaniki Technicznej.

do powszechnego użytku weszły komputery PC. Ale dopiero pojawienie się urządzeń przenośnych zmieniło sytuację radykalnie.

Pierwszy pomysł urządzenia przenośnego to urządzenie „Newton” firmy Apple dostępne na rynku w latach 1993–1998. Kolejnymi urządzeniami przenośnymi, które odniosły sukces na rynku, były komputery przenośne z linii „Palm” (od 1996 roku) produkowane przez firmę Sony. Mniej więcej w tym samym czasie zaczęły być dostępne komputerki przenośne z systemem pochodzącym od firmy Microsoft: „Palm-size PC” zastąpione później przez „Pocket PC”.

Praktycznie wszystkie z tych urządzeń przenośnych posiadały możliwość wyświetlania książek elektronicznych. Projekt Gutenberg był jednym z głównych dostawców darmowych treści.

Firma Google w 2004 roku podpisała porozumienie z kilkoma dużymi (i bogatymi) bibliotekami uniwersyteckimi. Celem tego porozumienia było zdigitalizowanie (a następnie udostępnienie w wersji cyfrowej) zawartości tych bibliotek. Pierwotnie zawartość udostępniana była w postaci rastrowych skanów stron książek. Nieco później Google uruchomiło proces rozpoznawania tekstu. Niestety nie zostały wykorzystane doświadczenia Projektu Gutenberg: rozpoznawanie jest automatyczne i pominięto proces sprawdzania. Pozwoliło to na stosunkowo szybkie zgromadzenie bogatej biblioteki cyfrowej zawartości.

Więcej informacji na temat historii książek elektronicznych zawiera interesująca publikacja elektroniczna [8].

2. E-ink — elektroniczny atrament

Kolejnym przełomem w historii książek elektronicznych było pojawienie się elektronicznej wersji atramentu. Podobnie jak w przypadku książek elektronicznych, historia jest długa. W roku 1970 firma Xerox PARC opracowała prototyp książki elektronicznej. Anegdota głosi, że „wyświetlacz” składał się z ogromnej liczby miniaturowych magnesów, których bieguny były pomalowane na czarno i białą zamocowanych w bardzo cienkiej i miękkiej gumie. Zmiana pola magnetycznego (wywoływana prądem elektrycznym) powodowała odpowiednie zmiany obrazu na „wyświetlaczu”. Prototyp nigdy nie został skomercjalizowany: był bardzo duży i bardzo niewygodny w użyciu. Faktem jest, że już w latach siedemdziesiątych Xerox PARC opracowała technologię papieru elektronicznego zbliżoną do używanej dziś powszechnie (dwukolorowe, polimerowe dipole) i w pewnym sensie ją zaniedbała, koncentrując się na technologii urządzeń drukujących i kopiujących.

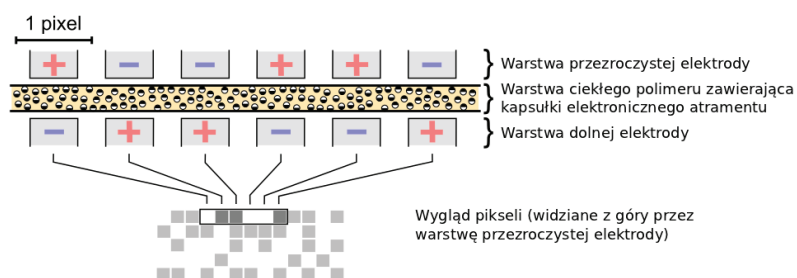
W 1993 roku Joe Jacobson z MIT rozpoczął prace nad wyświetlaczem działającym poprawnie przy dużym oświetleniu zewnętrznym, w 1996 zgłosił urządzenie w urzędzie patentowym, a w roku 2000 patent otrzymał. Jego urządzenie składało się z mikro-kulek polimerowych, których każda z połówek miała inny kolor i ładunek elektryczny, zamocowanych w cienkiej warstwie. Obrót kulek

i w efekcie tworzenie obrazu wywoływane było zmianą ładunku elektrycznego.

Wyświetlacz składa się z trzech warstw :

- zewnętrznej, przezroczystej warstwy przewodzącej,
- warstwy ciekłego polimeru zawierającej kapsułki e-ink,
- dolnej warstwy przewodzącej.

Dostarczany ładunek elektryczny obraca kapsułki z elektronicznym atramentem, formując punkty obrazu.



Rys. 1. Wyświetlacz e-ink (źródło: Wikipedia [13])

Jacobson powołał przedsiębiorstwo E Ink Corporation, które skomercjalizowało jego pomysł i zastosowało go w mnogich urządzeniach, we współpracy z wieloma firmami:

- Sony Reader z firmą Sony,
- telefon Motofone F3 z firmą Motorola,
- czytnik książek Nook z księgarnią Barnes & Noble,
- czytnik książek Kindle z firmą Amazon.

Obecnie na rynku jest bardzo wiele produktów dostarczanych przez wiele firm wykorzystujących własne rozwiązania technologiczne, o podobnych właściwościach użytkowych, różniące się głównie zastosowanym oprogramowaniem.

Główną zaletą wyświetlaczy działających na opisanej zasadzie jest ich bardzo dobra czytelność w pełnym świetle (nawet przy ostrym świetle słonecznym). Podobnie jak „zwykły” papier nie nadają się do używania w złych warunkach oświetleniowych. Powszechnie używane technologie nie nadają się również do wyświetlania obrazów ruchomych: wyświetlacz dosyć wolno reaguje na zmiany ładunku elektrycznego. Wyświetlacze pozwalają na odwzorowanie od czterech do szesnastu poziomów szarości.

Informacja wyświetlana na wyświetlaczu e-ink nie potrzebuje odświeżania, a energia zużywana jest tylko podczas „tworzenia” obrazu. Pozwala to wykorzystywać papier elektroniczny jako nośnik różnego rodzaju etykiet (na przykład adresowych).

Trwają bardzo intensywne prace nad stworzeniem wyświetlaczy wykorzystujących technologię papieru elektronicznego, ale pozwalających na wyświetlanie informacji w kolorze. Pierwsze wyniki nie są zbyt zachęcające: nasycenie kolorów

jest bardzo niskie, podobnie jak kontrast.

Trwają również poszukiwania bardziej tradycyjnych rozwiązań (zbliżonych od technologii LCD), ale zużywających znacznie mniej energii niż klasyczne wyświetlacze LCD i dobrze widocznych w warunkach pełnego oświetlenia słonecznego. Bardzo interesujące osiągnięcia ma firma Pixel Qi pracująca od około dziesięciu lat nad technologią wyświetlaczy LCD pracujących w świetle odbitym (transflective LCD [15]). W roku 2011 zapowiedziano dostępność pierwszego ich produktu: wyświetlacza 10,1" o rozdzielczości 1024 x 600 punktów. Zużycie energii jest około pięć razy mniejsze niż klasycznego wyświetlacza LCD. Wyświetlacz może pracować w trzech trybach:

- transmisyjny (światło przechodzące, jak klasyczny wyświetlacz LCD podświetlany diodami LED) — kolorowy,
- światło odbite (podświetlenie jest wyłączone, wyświetlacz pracuje w trybie z odcieniami szarości, zużywając najmniej energii),
- transflective — podświetlenie jest zredukowane, dostępne oświetlenie zewnętrzne jest wykorzystywane do podświetlenia, kolory są zredukowane, zużycie energii jest znacznie mniejsze niż w trybie klasycznym.

Tryb może być wybierany automatycznie, w zależności od wymagań użytkownika i oświetlenia zewnętrznego.

3. Formaty książek elektronicznych

Najstarszym formatem stosowanym w książkach elektronicznych był zwykły tekst (.txt), kodowany jako tak zwane *plain ASCII* — bez użycia jakichkolwiek znaków dodatkowych („akcentowanych”). Z oczywistych powodów format nie był specjalnie istotny: nie było zbyt wiele urządzeń mogących służyć do czytania książek elektronicznych. (Pierwszym urządzeniem był komputer „Dynabook” opracowany przez PARC — oddział badawczy firmy Xerox).

Podstawową zaletą tego formatu jest prostota i uniwersalność, a główną wadą możliwość obrazowania jedynie 26 liter (i pewnej liczby znaków dodatkowych, cyfr, znaków interpunkcyjnych z zakresu standardowego zestawu ASCII).

Dopiero rozpowszechnienie się Internetu (przełom lat osiemdziesiątych i dziewięćdziesiątych) oraz upowszechnienie formatu HTML (warto pamiętać, że pierwsze wersje tego języka oraz pierwsze wersje przeglądarek nie oferowały żadnej metody oznaczenia stosowanego kodowania zawartości ani nie dopuszczały znaków innych niż ASCII) doprowadziło do spopularyzowania i ustandaryzowania narzędzi służących do tworzenia i interpretacji tekstów zawierających znaki narodowe objęte standardem Unicode.

Upowszechnienie się poczty elektronicznej i formatu MIME (1992) pozwoliło na stosowanie szerszego zestawu znaków niż *plain ASCII*, ale pamiętać trzeba że alternatywą dla ASCII były wówczas alternatywne, ośmiobitowe „strony kodowe” ISO-8859-x. Stosowanie ich pozwalało, na przykład, przedstawiać treści

zapisane cyrylicą, ale nie było możliwości równoczesnego przedstawienia ich z treściami zapisanymi w języku polskim (w zakresie wykraczającym poza *plain ASCII*).

Pojawienie się niewielkich urządzeń przenośnych (komputerki „Palm” czy „Pocket PC” czy niewielkie urządzenia z systemem Symbian) dało możliwość dostępu do treści w praktycznie dowolnych warunkach i wygenerowało potrzebę jakiegoś formatu służącego do wygodnego zapisu treści, która miała być prezentowana na niewielkim ekranie urządzenia przenośnego.

Urządzenia te wprowadzały szereg ograniczeń wynikających głównie z ograniczonej pojemności pamięci oraz niewielkich wymiarów ekranu. Użytkownicy tych urządzeń chcieli jednak korzystać z różnych krojów pisma, stylów lub wariantów fontu. Wyświetlany tekst nie powinien w żaden sposób nawiązywać do tradycyjnego układu tekstu na stronie książki: ze względu na rozmiar ekranu nie było to możliwe. Tekst powinien „wlewać” się w bardzo ograniczony rozmiar ekranu.

Powstało wówczas kilka formatów spełniających opisane założenia: wywodziły się one z formatu HTML (rezygnując często ze zbyt zaawansowanych i niepotrzebnych jego możliwości), operowały pojęciem ustępu (paragrafu, akapitu) tekstu, którego najmniejszą jednostką są słowa złożone z liter. Tak traktowany tekst znakomicie mógł być wyświetlany na małym ekranie. Kilka dodatkowych znaczników pozwalało na nadanie dokumentowi bardzo podstawowej struktury (listy) lub różnych form wyróżniania tekstu. Całość uzupełniał spis treści i jakiś sposób zagnieżdżenia ilustracji. Dodatkowo wszystkie informacje były kompresowane.

Nie bardzo mogły być tu stosowane typowe formaty używane do zapisu dokumentów biurowych (.doc, .rtf, ...), gdyż zbyt często dowoływały się do fizycznych rozmiarów strony. Również formaty typu PDF czy DjVu opracowywane specjalnie do przedstawiania złożonej lub zeskanowanej strony nie mogły być powszechnie stosowane.

Najpopularniejszymi formatami książek elektronicznych używanymi na początku XXI wieku były [2, 9]:

- tekst (.txt),
- hipertekst (.html, .htm),
- plucker,
- kompresowany HTML (.chm),
- eReader (.pdb),
- Microsoft LIT (.lit),
- Mobipocket (.prc, .mobi).

Wykorzystywane były również specjalnie przygotowane pliki w formacie PDF: pozwalały one na bardzo łatwe wyodrębnienie zawartości (tekstu) z pliku PDF i na ekranie prezentowany był wyłącznie tekst (po odrzuceniu znacznej części jego formatowania). Dzięki temu wypełniał on bardzo dobrze (*reflow*) niewielki ekran urządzenia przenośnego. Niestety niezbyt wiele narzędzi do tworzenia

plików PDF pozwalało na odpowiednie ich przygotowanie.

Swojego miejsca nie znalazł format DjVu zaprojektowany głównie do przechowywania zeskanowanych dokumentów. Co prawda posiada on możliwość przechowywania tekstu (na przykład OCR) powiązanego z obrazem, ale przeglądanie całych stron dokumentu na niewielkim ekranie urządzenia przenośnego jest problematyczne. Znacznie wygodniej wykorzystywać sam zeskanowany i przetworzony tekst.

Współcześnie do najpopularniejszych formatów należą:

- Amazon Kindle (.azw),
- EPUB (.epub),
- Mobipocket (.mobi),
- Broadband eBooks (BBeB) (.lrf, .lrx),
- FictionBook (.fb2).

3.1. Format Amazon Kindle

Firma Amazon sprzedaje dziś najwięcej książek elektronicznych, więc używany przez nią format jest „najpopularniejszy”. Powstał on w wyniku rozwoju formatu Mobipocket. Amazon w roku 2005 wykupił firmę Mobipocket SA i rozpoczął bardzo intensywny rozwój tego formatu, dodając do niego różne rozszerzenia. Do najważniejszych należą własny system zabezpieczeń DRM oraz możliwość synchronizacji książek między różnymi urządzeniami (notatki, ostatnio otwarta strona, zakładki).

Wykorzystanie formatu, który w całości znajduje się pod kontrolą jednej firmy, pozwoliło Amazonowi zbudować bardzo sprawnie działający „ekosystem” składający się z:

- księgarni;
- kilku rodzajów czytników;
- oprogramowania do przeglądania książek działającego w najpopularniejszych środowiskach (MS Windows, Apple Mac OS, Apple iPhone, Apple iPad, RIM Blackberry, Android, MS Windows Phone). Na liście tej nie ma tylko systemu Linux;
- zestawu usług:
 - Amazon Whispernet — usługa pozwalająca na bezprzewodową dostawę (oraz synchronizację między urządzeniami) książek na urządzenia przenośne z wykorzystaniem sieci komórkowej 3G (praktycznie na całym świecie) lub WiFi;
 - konwersji zewnętrznych formatów do formatu .azw;
 - przekazywania dokumentów przesłanych pocztą elektroniczną na czytnik (każdy użytkownik otrzymuje specjalne konto pocztowe, dokumenty wysłane na to konto automatycznie przekazywane są na skojarzone z kontem urządzenia).

Czytniki Kindle potrafią obsłużyć oprócz formatu .azw jego „starsze wersje”: .mobi, .prc oraz pliki tekstowe (.txt) oraz .pdf. Inne formaty muszą być konwertowane.

3.2. Format EPUB

Format EPUB to otwarty i darmowy format do zapisu książek elektronicznych opracowany przez International Digital Publishing Forum (IDPF). Główne cechy formatu to:

- darmowość i otwartość,
- możliwość skalowania czcionki,
- możliwość wykorzystania grafiki wektorowej (SVG) i rastrowej,
- wbudowane metadane,
- możliwość zastosowania DRM,
- używanie CSS.

Obecnie powszechnie używana jest wersja 2 formatu [3]. Trwają intensywne prace nad opracowaniem wersji 3 [4].

Format EPUB swoją budową wewnętrzną przypomina pliki .docx lub .odt (wykorzystywane przez procesory tekstu Microsoft Word oraz LibreOffice (StarOffice, OpenOffice.org)). Jest to archiwum ZIP, o rozszerzeniu .epub, zawierające pliki:

- z zawartością „tekstową” (w formacie HTML/XHTML),
- pliki CSS opisujące sposób jej formatowania,
- ilustracje,
- fonty wykorzystane do wyświetlania treści;
- metadane:
 - spis treści oraz opis struktury,
 - typy mime poszczególnych dokumentów składowych,
 - manifest (zawierający spis wszystkich dokumentów składających się na książkę).

4. Czytniki książek elektronicznych

Na rynku polskim dostępnych jest kilka urządzeń powiązanych z usługą księgarni elektronicznej:

- Amazon Kindle,
- eClicto firmy Kolporter,
- różne urządzenia powiązane z księgarnią Empik.

Czytnik Kindle dostępny jest od pewnego czasu w Polsce i trzeba powiedzieć, że jest to dziś najtańsze urządzenie na rynku. W wersji 3G korzysta z (częściowo opłacanego przez Amazon) roamingu, dając bezprzewodowy (Whispernet) dostęp do księgarni (i w pewnym zakresie do Internetu) na całym świecie. Podstawową niedogodnością wykorzystania tego urządzenia jest bardzo ograniczona oferta książek w języku polskim.

Firma Kolporter jako pierwsza zdecydowała się na rozpoczęcie dystrybucji czytnika i usługi księgarni elektronicznej. Niestety, oferowane urządzenie jest dosyć przestarzałe (i dosyć drogie), natomiast księgarnia wykorzystuje swój własny sposób zabezpieczenia książek elektronicznych, w związku z tym użytkownik pozostaje przywiązany do księgarni Kolportera. Jej oferta nie należy do najbogatszych.

Kilka miesięcy po Kolporterze Empik wszedł na rynek z ofertą czytnika wykonanego na zamówienie międzynarodowego konsorcjum kilku księgarni elektronicznych (między innymi niemiecka Thalia). Z niezrozumiałych powodów firma stosunkowo szybko wycofała się z promowania tego produktu choć ciągle oferuje czytniki elektroniczne i książki elektroniczne. Niestety oferta (tworzona we współpracy z Virtualo) nie jest zbyt bogata. Brakuje w niej zwłaszcza nowości. Księgarnia korzysta ze standardu Adobe do ochrony książek elektronicznych, co pozwala odczytać każdą książkę zabezpieczoną tym systemem niezależnie od księgarni.

Na rynku dostępna jest też całkiem szeroka gama urządzeń niepowiązanych z żadną księgarnią. Jeżeli ich oprogramowanie wspiera Adept (system DRM od Adobe) — można na nich czytać książki kupione w każdej księgarni korzystającej z tego systemu zabezpieczeń.

Do ciekawszych urządzeń należy czytnik Onyx. Wyposażony jest on w oprogramowanie Ivona do odczytu tekstu, a firma Legimi próbuje stworzyć ekosystem zbliżony do tego oferowanego przez Amazon: dostarcza oprogramowanie do synchronizacji książek między różnymi urządzeniami użytkownika.

5. Oprogramowanie do odczytu książek elektronicznych

Wspominając o oprogramowaniu do czytania książek elektronicznych, mam na myśli aplikacje uruchamiane **poza** dedykowanymi urządzeniami do czytania. Urządzenia dedykowane, mimo że bardzo często pracujące pod kontrolą systemu operacyjnego Linux, mają zainstalowane aplikacje pochodzące od różnych dostawców. Ich kod jest zamknięty. Trwają co prawda prace nad stworzeniem opartego na otwartym kodzie oprogramowania urządzeń dedykowanych, ale ich zaawansowanie nie jest specjalnie duże [12].

Do najpopularniejszych programów należą [10]:

- Amazon Kindle
- Adobe Digital Editions,
- Microsoft Reader¹,
- MobiPocketReader,
- iSilo,
- Stanza,
- Azardi,

¹ 30 sierpnia 2012 r. firma Microsoft ogłosiła zamknięcie tego projektu.

- Calibre,
- FBReader,
- Adobe Acrobat Reader.

Istnieje też spora grupa narzędzi pozwalających na odczyt książek elektronicznych w Internecie (najczęściej w oknie przeglądarki WWW). Coraz popularniejsze stają się różnego rodzaju „aplikacje” na urządzenia przenośne (iPad lub inne tablety, smartfony). Aplikacje te dają dostęp do (bardzo często płatnych) treści w Internecie (gazety, serwisy WWW, internetowe wydania różnych dzienników czy tygodników).

Funkcjonalność oferowana przez programy jest różna, podobnie jak obsługiwane formaty. Niewątpliwie, najwięcej oferuje oprogramowanie oferowane przez Amazon. Bardzo ważnym narzędziem (mimo swojej ogromnej słabości) jest Adobe Digital Editions (ADE) dający dostęp do wszystkich treści, które wykorzystują system DRM od Adobe.

Ciekawą cechą bardzo wielu programów/systemów do udostępniania książek elektronicznych jest możliwość subskrypcji serwisów WWW (gazety, blogi). Specjalne oprogramowanie cyklicznie „ściąga” zawartość wskazanych stron WWW, konwertuje ją do formatu książki elektronicznej i wgrywa na urządzenie do czytania. Może być to usługa darmowa (jak oferowana, na przykład, przez księgarnie Legimi) lub płatna (jak oferowana przez Amazon). Sama usługa wywodzi się z czasów gdy urządzenia przenośne (komputery typu palmtop) nie miały dostępu do sieci. Powstało wówczas sporo serwisów i oprogramowań ściągających zawartość stron WWW i przystosowującego ją do wymogu urządzeń przenośnych. Zawartość była udostępniana jako HTML (rzadko) lub w jednym z wybranych formatów książek elektronicznych.

Oprogramowanie Calibre oferuje dziś każdemu jego użytkownikowi możliwość ściągania stron WWW i konwersji ich (we własnym zakresie) do wybranego formatu. Wykorzystuje się tu najczęściej skojarzony zwykle z blogiem lub zawartością portalu informacyjnego strumień RSS. Bardzo często można w ten sposób ściągnąć całą zawartość udostępnionych w Internecie archiwalnych numerów tygodników.

6. Oprogramowanie do tworzenia książek elektronicznych

Tworzenie książek elektronicznych przeznaczonych do czytania na urządzeniach przenośnych jest właściwie jedną z prostszych czynności. Oczywiście potrzebna jest sama książka (zawartość), ale wszystko później jest stosunkowo proste. Zwłaszcza gdy dysponujemy elektroniczną wersją dzieła. Gdy dysponujemy tylko papierowym oryginałem, dochodzi do tego szereg żmudnych czynności technicznych związanych ze skanowaniem, rozpoznawaniem tekstu (OCR), przetwarzaniem rozpoznanego tekstu (łączenie stron, podział na akapity,

likwidacja występujących w tekście „przeniesień”, eliminacja informacji znajdujących się w żywej paginie — numeracja stron, tytuły rozdziałów...) oraz wydobywania ilustracji.

Na potrzeby Projektu Gutenberg (ale i pirackiego rynku) powstało szereg aplikacji ułatwiających ten proces. Oprogramowanie do rozpoznawania tekstu jest zazwyczaj standardową częścią oprogramowania kupowanego ze skanerem. Ogromnym ułatwieniem są skanery z podajnikiem (ale wówczas trzeba zniszczyć papierowy oryginał, odcinając grzbiet i rozdzielając go na pojedyncze kartki). Powstało też wiele amatorskich konstrukcji wykorzystujących cyfrowy aparat fotograficzny. Bardzo często są one na tyle sprytnie, że pozwalają stosunkowo szybko sfotografować zawartość książki bez niszczenia jej. Dzięki temu, że ceny cyfrowych aparatów potrafią być bardzo przystępne — można wykorzystać dwa aparaty, by równocześnie fotografować dwie strony.

Mówiąc o przygotowaniu publikacji elektronicznych, myślę o konwersji tekstu dostępnego już w postaci elektronicznej.

Również w tym przypadku dostępnych jest wiele narzędzi pozwalających na łatwą konwersję. W przypadku najpopularniejszych procesorów tekstu (MS Word, OpenOffice.org) sytuacja jest o tyle prosta, że format plików .docx i .odt jest zbliżony do jednego z najpopularniejszych formatów książek elektronicznych: EPUB. Praktycznie wszystkie formaty książek elektronicznych oparte są na języku HTML więc sama konwersja jest procesem prostym: wystarczy przekształcić dostępny tekst do postaci ustępów tekstu (elementów `<p>paragraph`), nadać tekstowi strukturę (wykorzystując elementy nagłówków `<h1>do <h6>`) i tak dalej, dodać podstawowe metadane (tytuł, autor) oraz zaprojektować atrakcyjny szablon CSS nadający całości ostateczny wygląd.

Podobnie jak w przypadku dobrych serwisów WWW ostatecznym wyglądem książki elektronicznej można sterować zmieniając szablon CSS. Wiele przykładów zawiera zawartość na specjalnie stworzonym serwisie [5].

6.1. Przegląd wybranych programów do tworzenia książek elektronicznych

W dalszej części opisane zostaną najpopularniejsze aplikacje wspomagające tworzenie książek elektronicznych. Podstawowy nacisk położony zostanie na oprogramowanie darmowe i/lub o otwartym kodzie źródłowym.

Adobe InDesign – Program InDesign uchodzi dziś za jedno z najważniejszych narzędzi używanych do składu publikacji i cieszy się zasłużoną sławą. Od pewnego czasu pozwala zapisać wynik pracy w formacie EPUB [1]. Adobe dostarcza również wskazówek na temat wykorzystania programu InDesign do przygotowywania książek elektronicznych przeznaczonych na urządzenia Kindle lub iPad.

Sigil – Program Sigil to stosunkowo prosta aplikacja do tworzenia książek elektronicznych. Swoją konstrukcją bardzo przypomina niezbyt rozbudowany

edytor tekstu WYSIWYG i może być wykorzystany do „ręcznego” wpisania zawartości książki elektronicznej [17].

Program może też być wykorzystany do dokonywania drobnych poprawek w istniejących już książkach elektronicznych w formacie EPUB.

eCub – eCub to stosunkowo prosta i darmowa aplikacja pozwalająca zaimportować tekst zawarty w plikach HTML lub .txt i przetworzyć je do formatu książki elektronicznej EPUB lub MobiPocket [16]. Znacznie bardziej zaawansowane, ale płatne narzędzie to Jutoh — program pochodzący od tego samego dostawcy. Program nie wymaga żadnej specjalnej wiedzy technicznej, może służyć osobom przyzwyczajonym do pracy ze zwykłym edytorem tekstu do tworzenia książek elektronicznych.

Calibre – Prawdziwym kombajnem służącym do zarządzania biblioteką książek elektronicznych oraz pozwalającym na konwersję między wielu ich formatami jest program Calibre. Może on również współpracować z wielu czytnikami książek elektronicznych jako menedżer zawartości, pozwalając pobierać książki z wielu źródeł, wgrywać je do urządzenia i kasować. Jedną z ciekawszych funkcji jest możliwość pobierania zawartości serwisów WWW i konwersji ich do formatu książki elektronicznej. Program obsługuje najpopularniejsze formaty: EPUB, MOBI, LRF, PDF, ODT, RTF, LIT, FB2.

LaTeX – System LaTeX może nie jest najpopularniejszym systemem do przygotowywania publikacji elektronicznych, ale ciągle jest bardzo często używany przez naukowców do profesjonalnego składu publikacji elektronicznych. LaTeX oprócz wykorzystywania swojego własnego formatu (DVI) może dodatkowo zapisywać książki w formatach PS, PDF oraz HTML². Jest to bardzo dobra podstawa do późniejszej konwersji na przykład do formatu EPUB. Do konwersji można użyć albo narzędzi pakietu Calibre albo programu eCub [11].

System LaTeX bywa używany do składu czasopism naukowych, ponieważ dziś standardem jest udostępnianie publikacji w wersji elektronicznej, niektóre firmy podejmują bardzo intensywnie starania, aby cały proces maksymalnie zautomatyzować i dostarczać książki w możliwie wielu formatach elektronicznych (poza powszechnie używanym formatem PDF).

SISU – Prawdziwym kombajnem do przygotowywania publikacji elektronicznych jest system SISU [14]. Używa on własnego języka znaczników do opisu struktury i zawartości dokumentu elektronicznego i wykorzystuje różne narzędzia do przetwarzania zawartości do różnych formatów: TXT, HTML, EPUB, XML, ODF, LaTeX, PDF.

Serwisy Internetowe – Oprócz wielu aplikacji powstają serwisy oparte o specjalne oprogramowanie na serwerze WWW i przeglądarkę, pozwalające przygotować, opublikować oraz sprzedać książkę elektroniczną. Serwisów takich jest już całkiem sporo i tworzą one bardzo interesujący „ekosystem”: można przygotować książkę elektroniczną, skorzystać z usługi jej redakcji, skorzystać

² Dzięki dodatkowemu programowi TeX4ht.

z gotowej (lub zaprojektować własną) okładki, a następnie wystawić książkę na sprzedaż bądź w postaci elektronicznej lub skorzystać z jakiegoś serwisu „druku na żądanie” i w ten sposób ją udostępniać.

7. Prawa autorskie i DRM

Jednym z najciekawszych problemów związanych z rozpowszechnianiem książek elektronicznych jest kwestia „ochrony praw autorskich”, czyli ochrona przed nieograniczonym kopiowaniem książek elektronicznych.

Problem ochrony praw autorskich staje się coraz poważniejszym wyzwaniem dla wszystkich dostawców treści elektronicznych. Największy problem związany jest z pojęciem kopiowania. Nie da się skorzystać z jakiegokolwiek (elektronicznej) zawartości bez jej kopiowania. Po skopiowaniu strumienia bitów z jakiegoś serwera może być on „obejrzany” na komputerze odbiorcy. A jak już jest na tym komputerze w formie pozwalającej na jego odtworzenie — może być bez problemu skopiowany.

Stąd pojawiają się różnego rodzaju aplikacje (o zamkniętym kodzie), które wykorzystują jakąś formę „szyfrowania” zawartości na czas przesyłania z serwera na komputer użytkownika i przechowywania na tym komputerze. Z drugiej strony potrafią one zawartość odszyfrować i wyświetlić na ekranie. Z natury swej rzeczy systemy te podatne są na różnego rodzaju ataki: skoro i książka (zaszyfrowana), i program do jej odczytu są na prywatnym komputerze, jest tylko kwestią czasu i zawziętości złamanie systemu zabezpieczeń. I tak się dzieje.

Z drugiej strony skoro na rynku pojawiają się urządzenia do odczytu książek elektronicznych, a oferta księgarni nie jest zbyt bogata — pojawia się nielegalny rynek, na którym pojawiają się książki. Najciekawszym zagadnieniem jest sposób ich digitalizacji. Przypomina on ten opisany na samym początku, a wykorzystywany przez Projekt Gutenberg. Spore rzesze ochotników kupują książki, dzielą je na pojedyncze kartki i poddają żmudnemu procesowi rozpoznawania tekstu, a następnie poddają go redakcji. Otrzymane pliki (zazwyczaj w formatach, które potrafi zapisać popularne oprogramowanie OCR) umieszczane są na okrytych złą sławą serwerach udostępniania plików. W Polsce działała grupa ludzi, którzy w dosyć krótkim czasie zdigitalizowała całkiem pokaźny (ponad tysiąc pięćset tytułów) zestaw najpopularniejszych w kręgach młodzieżowych książek.

Natomiast trzeba też przyznać, że technologia stosowana przez firmy przygotowujące legalną zawartość do księgarni elektronicznych bardzo mocno przypomina tę nielegalną. Standardowym formatem używanym przez wszystkie drukarnie cyfrowe jest dziś PDF. Zatem wydawnictwa przygotowują pliki PDF i wysyłają je do druku. Tak przygotowany plik PDF może być przedmiotem oferty księgarni cyfrowej (oczywiście w pewnym uproszczeniu). Trzeba go tylko zaszyfrować. Jeżeli księgarnia chce udostępnić go w formacie EPUB (lub MOBI), to najczęściej stosowana w Polsce procedura sprowadza się do wydobycia z pli-

ku PDF (używanego do wydruku) tekstu i odpowiedniego sformatowania go i zaszyfrowania. Tak prymitywny proces ma spore konsekwencje: jakość wersji elektronicznej bardzo często jest niezadowalająca.

Tworzenie z jednej wersji źródłowej kilku wersji elektronicznych nie jest jeszcze częścią procesu wydawniczego. Natomiast wydawnictwa bardzo głośno walczą o prawa autorów (jak same mówią) i podejmują próby zatrzymania procesu digitalizacji treści. Nie tylko takie jak „amatorska” działalność czytelników. Również działalność Google (która rozpoczęła się od skanowania „białych kruków” w bibliotekach uniwersyteckich i bardzo szybko przerodziła się w komercyjne przedsięwzięcie sprzedaży książek elektronicznych) budzi bardzo silny sprzeciw wydawców.

Trzeba jednak przyznać, że są wydawcy, którzy podejmują ryzyko sprzedaży książek elektronicznych niezabezpieczonych DRM. Co więcej, tak formułują swoje cele biznesowe, że potrafią je zrealizować. Najbardziej prominentnym przykładem takiego postępowania jest wydawnictwo O’Reilly. Sprzedaje ono książki w kilku formatach (PDF, EPUB, MOBI) niezabezpieczonych żadną formą DRM. Kupując książkę elektroniczną, użytkownik dostaje dostęp do wszystkich formatów. Co więcej — pojawienie się nowej wersji (na przykład nowego wydania) daje użytkownikom automatyczny dostęp do niego. Przyjęty model pozwala na udostępnianie książek w wersji roboczej. Jest to bardzo istotne w przypadku opisywania najnowszych, zmieniających się bardzo szybko technologii. Czytelnik może mieć stały dostęp do wszystkich nowinek.

Firma stosuje różne formy promocji: posiadacze książek papierowych mają zniżkę na zakup elektronicznych. Podobna zniżka działa i w drugą stronę: po kupieniu książki elektronicznej papierowa jest tańsza. Stali klienci mają zniżki na zakup książek. Codziennie jest promocja jednej książki z oferty. Niestety, przykład tego wydawnictwa nie przemawia do wyobraźni innych wydawców.

Streszczenie

Książki elektroniczne (stan technologii)

Praca przedstawia książki elektroniczne oraz aktualny stan technologii, omawia zasadę działania wyświetlaczy stosowanych w najpopularniejszych urządzeniach konsumenckich do odczytu książek, omawia główne formaty przechowywania informacji stosowane w tych urządzeniach oraz omawia oprogramowanie używane do przygotowywania książek elektronicznych.

Summary

E-books (technology status)

This article presents electronic books (e-books) and the current state of technology, discusses the principle of the displays used in most popular consumer devices for reading

e-books, describes the major formats for storing information used in these devices and discusses the software used in the preparation of electronic books.

Literatura

- [1] *Eksport zawartości do formatu EPUB (CS 5.5)*, http://help.adobe.com/pl_PL/indesign/cs/using/WS8c5bc4f64c7a4a3d44f3b18d12dbcdf377a-8000.html#WS8c5bc4f64c7a4a3d-487d7b7912dbcdff154-8000.
- [2] *Comparison of e-book formats – Wikipedia, the free encyclopedia*, http://en.wikipedia.org/wiki/Comparison_of_e-book_formats, 2011.
- [3] *EPUB 2.0.1 | International Digital Publishing Forum*, <http://idpf.org/epub/201>, 2010.
- [4] *EPUB 3 overview*, <http://idpf.org/epub/30/spec/epub30-overview.html>, 2011.
- [5] *About ePub zen garden*, <http://epubzengarden.com/about/>, 2011.
- [6] Hart M.S., *History and philosophy of project gutenber*, 1992, http://www.gutenberg.org/wiki/Gutenberg:The_History_and_Philosophy_of_Project_Gutenberg_by_Michael_Hart.
- [7] Lebert M., *Project Gutenberg (1971-2008)*, 2008, <http://www.gutenberg.org/ebooks/27045>.
- [8] Lebert M., *A Short History of EBooks*, 2009, <http://www.gutenberg.org/ebooks/29801>.
- [9] *MobileRead wiki – e-book formats*, http://wiki.mobileread.com/wiki/E-book_formats, 2011.
- [10] *MobileRead wiki – e-book software*, http://wiki.mobileread.com/wiki/E-book_software, 2011.
- [11] Myszka W., *LaTeX to ePUB – the easy way*, <http://staryzgreed.blogspot.com/2010/10/latex-to-epub-easy-way.html>, 2010.
- [12] *OpenInkpot*, <http://openinkpot.org/>, 2011.
- [13] *Run!, File:E-ink.svg – wikimedia commons*, <http://commons.wikimedia.org/wiki/File:E-ink.svg>, 2006.
- [14] *Structured information, serialized units*, <http://www.sisudoc.org/>, 2011.
- [15] *TFT LCD Specification MODEL NO.: PQ 3Qi-01*, Raport techniczny, Pixel Qi, 2010.
- [16] Wimmer P., *Anthemion eCub – kompilator ebooków*, <http://poradnikwebmastera.blox.pl/2009/01/Anthemion-eCub-kompilator-ebookoacutew.html>, 2009.
- [17] Wimmer P., *Sigil – ePub lekko, łatwo i przyjemnie*, <http://poradnikwebmastera.blox.pl/2010/10/Sigil-ePub-lekko-latwo-i-przyjemnie.html>, 2010.