

KOMPETENCJE KIEROWNIKÓW PROJEKTÓW – ANALIZA TEXT MINING

Marcin WYSKWARSKI

Wydział Organizacji i Zarządzania, Politechnika Śląska, Zabrze; marcin.wyskwarski@polsl.pl

Streszczenie: W artykule podjęto próbę identyfikacji oczekiwanych przez pracodawców kompetencji kierowników projektów. W tym celu przeanalizowano treść ofert pracy z wykorzystaniem analizy text mining. Stosownie do przyjętego celu ustalono strukturę pracy. W punkcie drugim przedstawiono najważniejsze informacje dotyczące kompetencji kierowników projektów. Wymieniono cztery modele kompetencji, przedstawiono czym one są, na jakie grupy można je podzielić oraz, które z nich według wybranych autorów są najistotniejsze. Punkt trzeci podzielono na trzy części. W pierwszej z nich przedstawiono sposób zebrania danych do analizy. Kolejna część opisuje procedurę przeprowadzonej analizy text mining. Ostatnia część prezentuje wyniki przeprowadzonej analizy. Ustalono jakie słowa w analizowanych ofertach pojawiały się najczęściej, oraz jak była korelacja tych słów z innymi słowami.

Słowa kluczowe: text mining, kompetencje, kierownik projektu, chmura słów.

PROJECT MANAGERS' COMPETENCES – TEXT MINING ANALYSIS

Abstract: This article presents a proposal of identification of project managers' competences in order to employers' requirements. For this purpose the article presents the analysis of job advertisements with use of text analysis. According to that point of view the structure of the paper was created. The second part containing description of competences of project managers'. It shows the definition of competence, how it could be divided and which are the most important according to the authors. The third part of the article is divided into three pieces. First shows how to collect data to analysis. Second part shows procedure of text mining analysis, and the last part shows the results of such kind of analysis. It contains the words which appeared most frequently in the job offers and the correlation of those words with other words in text.

Keywords: text mining, competences, project manager, world cloud.

1. Wprowadzanie

Jednym z istotnych czynników powodzenia projektu są wiedza, umiejętności oraz postawa osoby która nim zarządza. Sprawia to, że kompetencje kierowników projektów stają się tematem wielu badań. Są one również istotne dla przyszłych pracodawców, którzy starają się pozyskać kierowników projektów z odpowiednimi kompetencjami. Pożądane na rynku kompetencje mogą być również wskazówką dla obecnych i przyszłych kierowników projektów.

Celem artykułu była próba identyfikacji oczekiwanych przez pracodawców kompetencji kierowników projektów z wykorzystaniem analizy text mining ofert pracy.

Text mining jest stosunkowo młodą i interdyscyplinarną dziedziną wywodzącą się m.in. z data mining, wyszukiwania informacji, kategoryzacji tekstu i modelowania probabilistycznego (Kao and Poteet, 2007). To proces, którego celem jest wydobywanie z zasobów tekstowych nieznanych wcześniej informacji (Hearst, 1999). Ten sposób analizy informacji, przechowywanych w formie tekstu pisanego, pozwala całkowicie lub częściowo zautomatyzować proces przetwarzania i analizowania tekstu. Jest to szczególnie istotne, gdy zachodzi konieczność zapoznania się z treścią bardzo dużej liczby dokumentów tekstowych.

2. Kompetencje kierownika projektu

Na kompetencje danej osoby składają się cechy osobowości, motywacja, umiejętności, samoocena związana z funkcjonowaniem w grupie, oraz przyswojona i wykorzystywana wiedza (Gunz, 1983). Kompetencje kierowników projektów przedstawia się często jako zestaw wiedzy, umiejętności, cech osobowości i doświadczenie (Pawlak, 2006).

Kompetencje można podzielić na dwie podstawowe grupy tj. kompetencje twarde i miękkie. Twarde dotyczą konkretnego stanowiska pracy. Określa się je mianem kompetencji fachowych, technicznych, zawodowych, merytorycznych, funkcjonalnych. Kompetencje miękkie to natomiast cechy osobowości wnoszone do danych ról zawodowych (behawioralne, społeczne, interpersonalne) (Armstrong et al., 2016). Kierownik projektu powinien posiadać fachową wiedzę z obszaru, w którym realizowany jest projekt, oraz być ekspertem w obszarze metod i technik zarządzania projektem (Pawlak, 2006).

Według Wachowiaka umiejętności kierownika projektu można podzielić na następujące grupy (Wachowiak et al., 2004):

- techniczne - pozwalają zrozumieć istotę projektu i wykonać zadania,
- interpersonalne - to umiejętność nawiązywania i utrzymywania kontaktów między ludźmi,
- koncepcyjne - pozwalają na twórcze rozwiązywanie problemów,

- diagnostyczne i analityczne - umożliwiają dokonanie diagnozy problemów,
- polityczne - umożliwiają skuteczne oddziaływanie na środowisko projektu.

Do najważniejszych kompetencji kierowników projektów K. Piwowar-Sulej zalicza (Piwowar-Sulej, 2013):

- wysokie kwalifikacje zawodowe, wiedza techniczna związana z przedmiotem i zakresem projektu, znajomość metodyk zarządzania (metodyk tradycyjnych i nowoczesnych),
- umiejętność wyznaczania celów oraz organizacji pracy zespołu projektowego,
- niezależność w ocenie faktów,
- otwartość na niestandardowe metody pracy,
- dobrze rozwinięte umiejętności społeczne (m.in. zdolności negocjacyjne, dyplomatyczne, tolerancja dla innego punktu widzenia członków zespołu, marketingowe podejście do klienta).

Opracowana przez A. Musioł-Urbańczyk lista kompetencji kierownika projektu obejmuje 46 pozycji, które zostały podzielone na cztery grupy kompetencji tj. kompetencje profesjonalne (19 kompetencji), społeczne (9 kompetencji), osobiste (14 kompetencji) i biznesowe (4 kompetencje). Z przeprowadzonych przez autorkę badań w zakresie wpływu kompetencji na skuteczność działań kierownika projektu wynika, że kluczowe kompetencje to kolejno: umiejętność komunikowania się, umiejętność podejmowania decyzji, przywództwo, umiejętność motywowania członków zespołu, umiejętność budowania zespołu, umiejętność zarządzania komunikacją w projekcie, praca zespołowa, umiejętność negocjowania, lojalność, umiejętność zarządzania zakresem projektu, elastyczność (Musioł-Urbańczyk, 2010).

Dużym powodzeniem, wśród profesjonalistów z obszaru zarządzania projektami, znajdują także następujące cztery modele kompetencji (Wyrozębski, 2009):

- IPMA Competency Baseline - model kompetencji kierownika projektu opracowany przez International Project Management Association,
- Project Manager Competency Development Framework - model kompetencji autorstwa amerykańskiego Project Management Institute,
- National Occupational Standards for Project Management - model kompetencji opracowany przez brytyjską organizację Engineering Construction Industry Training Board,
- Professional Competency Standards for Project Management – australijski model kompetencji projektowych, stworzony przez Australian Institute for Project Management.

3. Identyfikacja oczekiwanych kompetencji kierownika projektu

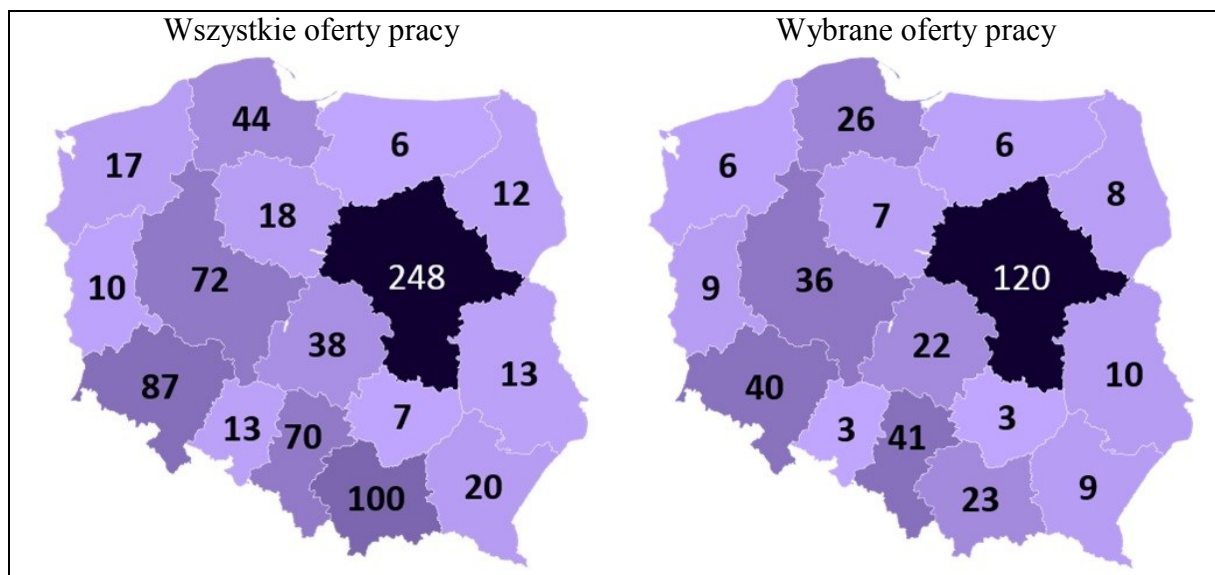
3.1. Źródło danych

Źródłem danych w przeprowadzonym badaniu były oferty pracy zamieszczone w serwisie www.pracuj.pl. Dane zebrano w terminie 09-11 kwiecień 2018. Analizie poddano oferty (ogłoszenia), które pojawiły się w wyniku wpisania frazy „kierownik projektu” w serwisie www.pracuj.pl. Po wstępnej, wrywkowej analizie treści ofert okazało się, że wśród 775 przedstawionych przez serwis www.pracuj.pl ogłoszeń znalazły się również takie, w których użyto innej nazwy niż „kierownik projektu”, oraz takie których treść była napisana w języku angielskim.

Do dalszej analizy wybrano tylko te ogłoszenia, które spełniały następujące dwa warunki:

- dotyczyły stanowiska „kierownik projektu” (stanowisko to mogło być również nazwane inaczej),
- były napisane w języku polskim.

Gdy wykorzystana w ogłoszeniu nazwa stanowiska budziła wątpliwość czy oferta na pewno dotyczy stanowiska „kierownika projektu” (np. „koordynator projektu”, „project coordinator”, „program manager”) analizowano treść ogłoszenia (szczególnie zakres obowiązków) i podejmowano decyzję o odrzuceniu lub akceptacji ogłoszenia. Z 775 wyszukanych w serwisie www.pracuj.pl ofert do dalszej analizy wybrano 369. Liczbę tych ogłoszeń w poszczególnych województwach przedstawia rysunek 1.



Rysunek 1. Liczba ofert pracy w poszczególnych województwach.

W wybranych 369 ofertach użyto następujących określeń dla stanowiska „kierownik projektu”: project manager (142 ogłoszenia), kierownik projektu (128 ogłoszeń), koordynator projektu (54 ogłoszenia), lider projektu (11 ogłoszeń), manager projektu (11 ogłoszeń),

młodszy kierownik projektu (11 ogłoszeń), junior project manager (5 ogłoszeń), senior project manager (3 ogłoszenia), program manager (2 ogłoszenia), project coordinator (1 ogłoszenie) oraz szef projektu (1 ogłoszenie).

Dla każdego z zakwalifikowanych ogłoszeń utworzono plik z rozszerzeniem txt, zawierający fragment ogłoszenia opisujący wymagane przez pracodawcę kompetencje. Fragment ten był najczęściej zatytułowany słowem „wymagania”. Inne stosowane dla tego fragmentu określenia, które pojawiały się w ogłoszeniach to m.in. takie frazy i słowa jak: „profil kandydata”, „profil idealnego kandydata”, „czego oczekujemy od Ciebie”, „szukamy kandydata który...”, „poszukujemy kandydata który...”, „u kandydatów zwracamy uwagę na”, „poszukiwane kompetencje”, „musthave”, „nasze oczekiwania”, „oczekiwania”. Pliki umieszczono w 16 folderach, po jednym folderze dla każdego z województw.

3.2. Etapy analizy text mining

Analiza text mining objęła następujące trzy etapy¹:

- wstępne przetwarzanie tekstu,
- utworzenie korpusów analizowanych dokumentów, budowa macierzy częstości występowania słów,
- wykorzystanie klasycznych metod pochodzących z Data Mining.

Wstępne przetwarzanie plików tekstowych przeprowadzono korzystając z aplikacji Notepad++ v.7.3.3 oraz RStudio v.1.0.136. W ramach tego etapu dla każdego z analizowanych dokumentów tekstowych utworzono tzw. worek słów (ang. bag of words). W tym celu z dokumentów usunięto wszystkie znaki z wyjątkiem liter, zamieniono duże litery na małe, korzystając z samodzielnie utworzonej tzw. stop-listy (ang. stop-words) usunięto słowa uznawane za nieprzydatne (spójniki, przyimki itp.), przekształcono słowa do ich wersji uznawanej za podstawową. Na końcu z plików usunięto słowa niebędące rzeczownikami oraz przymiotnikami, a każdy wyraz został umieszczony w osobnej linii².

Do przekształcenia słów do postaci podstawowej oraz pozostawienia w plikach tylko przymiotników i rzeczowników wykorzystano słownik morfosyntaktyczny „polimorfologik 2.1” udostępniony na portalu Github³. W momencie korzystania ze słownika miał on 4 811 854 linii tekstu. Ma on formę pliku tekstowego w kodowaniu UTF-8. Słownik zaimportowano do programu RStudio jako tabelę składającą się trzech kolumn: forma podstawowa, forma odmieniona, znaczniki gramatyczne. Przekształcenie słowa do postaci podstawowej polegało na jego odszukaniu w kolumnie „forma odmieniona” i zamianie na słowo znajdujące się w tym samym wierszu, w kolumnie „forma podstawowa”. Jeżeli dane słowo nie zostało odnalezione w kolumnie „Forma odmieniona” to pozostawało

¹ Uproszczony podział zaproponowany przez A. Gładysz (Gładysz, 2012).

² Plik zawierający np. 34 słowa ma 34 linii tekstu.

³ <https://github.com/morfologik/polimorfologik/releases/tag/2.1>

w dokumencie w niezmienionej formie - w takim przypadku nie uwzględniano także formy gramatycznej danego słowa.

W kolejnym etapie korzystając z tzw. modelu przestrzeni wektorowej (ang. Vector Space Model)⁴ utworzono cztery korpusy, dla których wykonano macierze dokumentów-wyrażeń (ang. document - term matrix) z częstościową reprezentacją występowania wyrażeń (ang. term frequency - TF). Utworzono korpusy dla województw, w których minimalna liczba ogłoszeń wynosiła 40 czyli dla województwa dolnośląskiego (korpus nr 1), mazowieckiego (korpus nr 2) oraz śląskiego (korpus nr 3). Czwarty korpus objął ogłoszenia pochodzące ze wszystkich województw wszystkie. Rysunek 2 przedstawia fragment macierzy dokumentów-wyrażeń⁵ utworzonej dla korpusu z ofertami z województwa dolnośląskiego.

Docs	Terms			
	dobry	dotatkowy	doświadczenie	działanie
w_10.txt	0	0	1	0
w_11.txt	1	1	1	0

Rysunek 2. Fragment macierzy dokumentów - wyrażeń dla województwa dolnośląskiego.

W ostatnim etapie dla każdego korpusu obliczono liczbę najczęściej występujących słów oraz ich korelację z innymi słowami. Wyniki zaprezentowano na rysunkach 4, 5, 6, 7.

Docs	Terms					Docs	Terms				
	angielski	dobry	praca	projekt	zdolność		angielski	dobry	praca	projekt	zdolność
1.txt	0	0	0	0	0	1.txt	0	0	0	0	0
2.txt	1	0	0	0	0	2.txt	1	1	0	0	0
3.txt	1	1	0	0	0	3.txt	3	3	0	0	0
4.txt	1	1	1	0	0	4.txt	1	1	1	0	0
5.txt	1	1	1	1	1	5.txt	4	4	3	1	0
6.txt	1	1	1	1	1	6.txt	2	2	2	1	2

<code>> findAssocs(dtm, "angielski", 0)</code>						<code>> findAssocs(dtm, "angielski", 0)</code>								
\$angielski						\$angielski								
	dobry	praca	projekt	zdolność		dobry	praca	projekt	zdolność		dobry	praca	projekt	zdolność
	0.63	0.45	0.32	0.20		1.00	0.64	0.61	0.06		1.00	0.64	0.61	0.06

Rysunek 3. Przykładowe macierz dokumentów-wyrażeń oraz wartość korelacji dla słowa „angielski”.

Korelację pomiędzy słowami obliczono z wykorzystaniem funkcji `findAssocs()`, która bazuje na standardowej funkcji `cor()` dostępnej w pakiecie statystycznym języka R. Rysunek nr 3. przedstawia dwie przykładowe macierze dokumentów-wyrażeń oraz obliczoną wartość korelacji dla słowa „angielski”. W dokumentach tworzących macierz dokumentów-wyrażeń przedstawioną po lewej stronie rysunku 3, słowa (angielski, dobry, praca, projekt, zdolność) wystąpiły co najwyżej raz w każdym z dokumentów. W dokumentach tworzących macierz dokumentów-wyrażeń zaprezentowaną po prawej stronie rys. 3, słowa te wystąpiły częściej niż jeden raz (np. słowa „angielski” oraz „dobry” wystąpiły po 4 razy w dokumencie nr 5. Jak widać różna liczba słów w przykładowych macierzach dokumentów-wyrażeń sprawiła, że korelacja słowa „angielski” z innymi słowami ma inną wartość. Korelacji równa 1 oznacza, że dane dwa słowa zawsze występują razem (w takiej samej ilości) w dokumentach. Wartość 0 oznacza, że słowa nigdy nie wystąpiły razem.

⁴ Model opisano m.in. w (Gładysz, 2012; Mirończuk, 2012).

⁵ Macierz ma rozmiary 40 x 78 (40-dokumentów, 78 - wyrażeń (słów)).

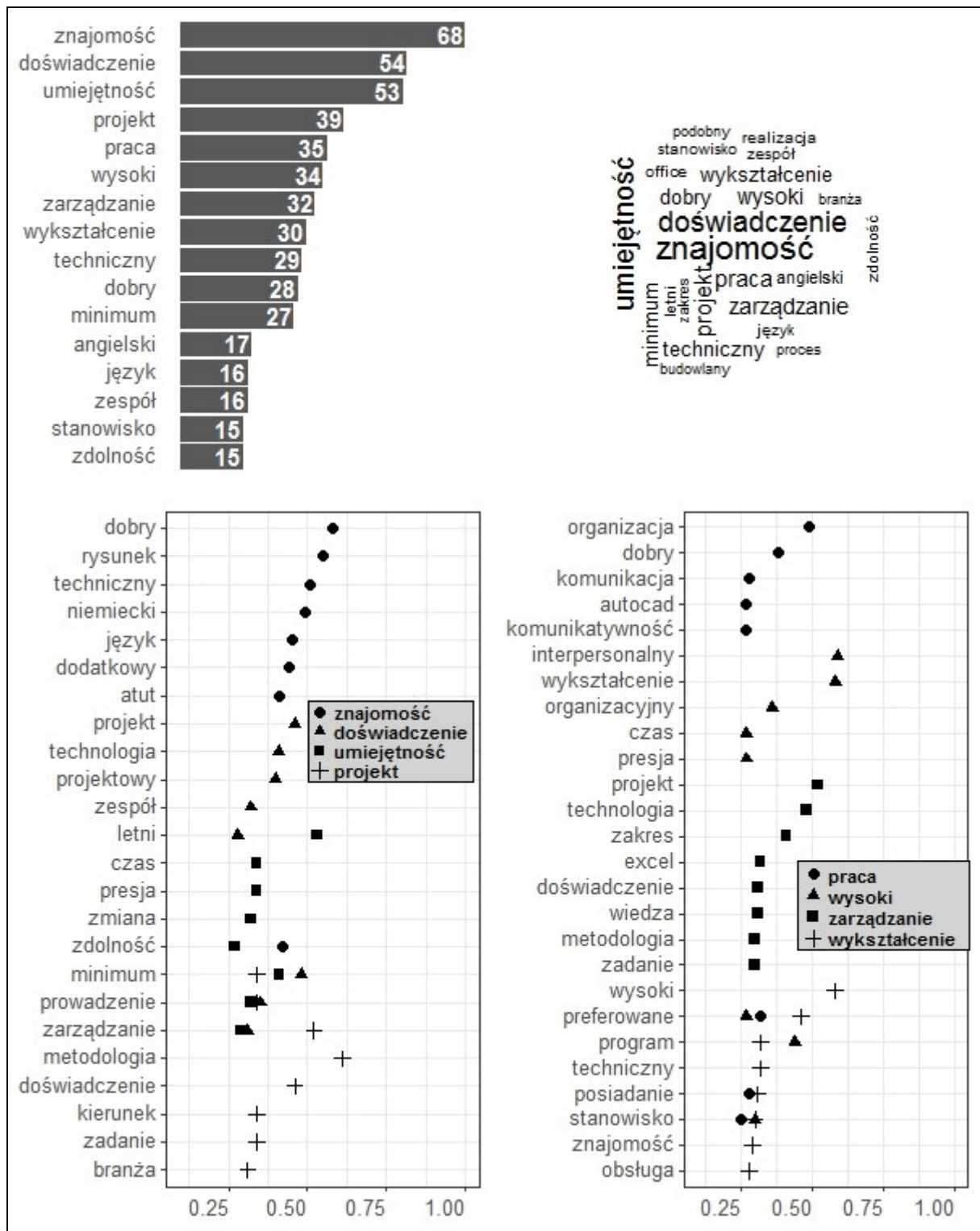
3.3. Wyniki przeprowadzonej analizy

Lewa górna część rysunków nr 4, 5, 6 i 7 przedstawia szesnaście najczęściej występujących w danym korpusie słów. Słowa te zostały uporządkowane malejąco według liczby wystąpień. Najczęściej używane słowa zaprezentowano także w postaci tzw. chmury słów⁶ (prawa górna część rysunków nr 4, 5, 6 i 7). Wielkość słów w chmurze zależy od liczby ich wystąpień - występujące najczęściej są największe a najrzadziej najmniejsze. Widać to na przykładzie m.in. chmury słów wykonanej dla wszystkich analizowanych ofert pracy (rys.7), na której dwa największe słowa „znajomość” i „doświadczenie”, użyto odpowiednio 645 i 536 razy. Patrząc na chmury słów przedstawione na rysunkach nr 4, 5, 6 i 7 widać, że mają one różne rozmiary. Dla korpusów składających się z większej liczby ogłoszeń (korpus 2 - rys. 5, korpus 4 - rys. 7) chmury słów zajmują większą powierzchnię. Jest to spowodowane tym, że został ustalone dwa kryteria liczbowe wyboru słów przy tworzeniu chmury tj. maksymalna liczba słów w chmurze, oraz minimalna wymagana liczba wystąpień słowa w korpusie.

Zgodnie z informacjami przedstawionymi na rys. 7 najczęściej występującym słowem we wszystkich 369 ogłoszeniach - korpus 4 - było słowo „znajomość”. Słowo to było również najczęściej używanym słowem w pozostałych trzech korpusach (rys. nr 4, 5, 6). Kolejnymi najczęściej występującymi słowami w analizowanych korpusach są słowa „doświadczenie”, „umiejętność”, „projekt”, „wysoki”. Ich kolejność jest różna w poszczególnych korpusach.

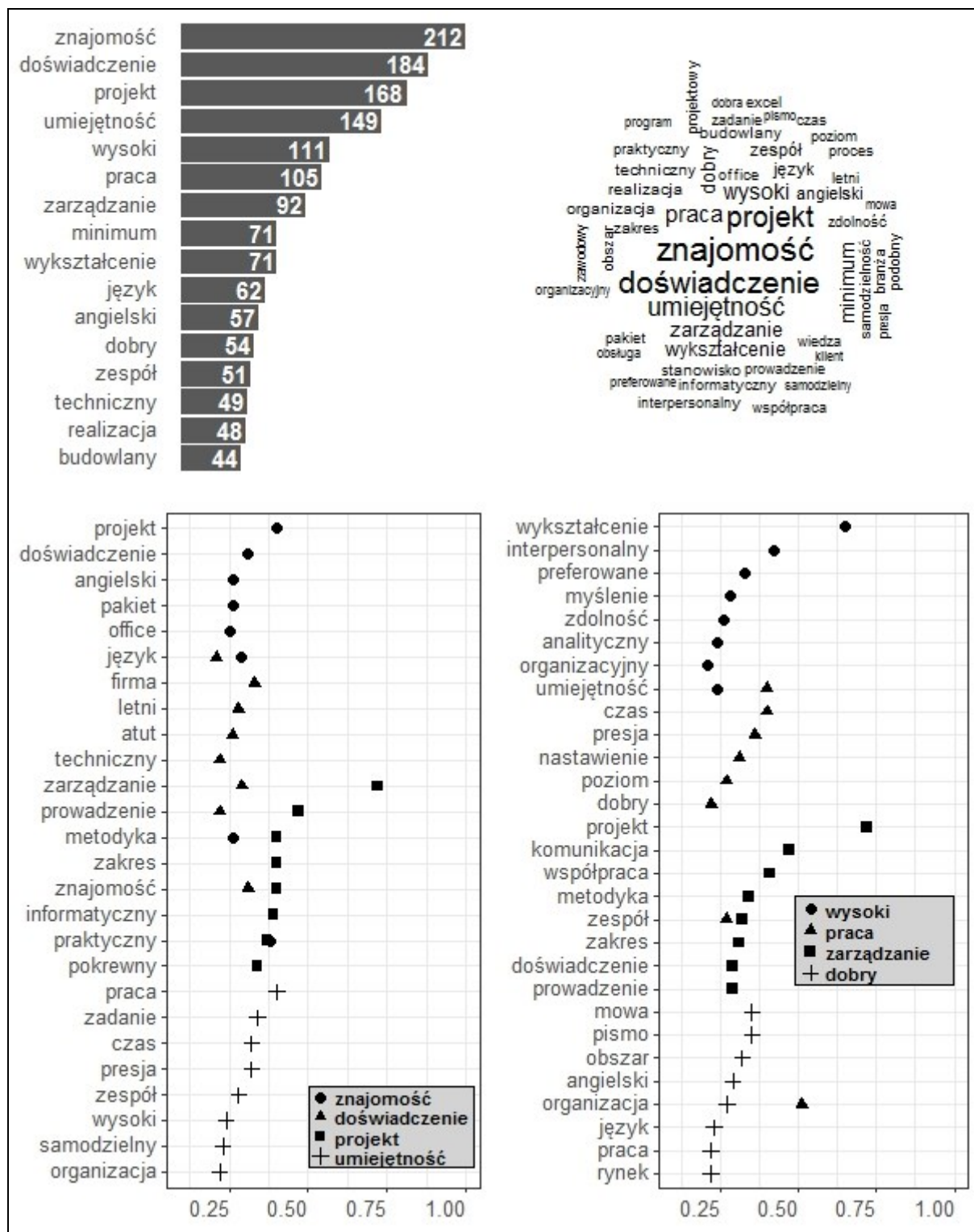
Dolna część rysunków nr 4, 5, 6 i 7 przedstawia dla ośmiu najczęściej używanych słów ich związek (na bazie obliczonego współczynnika korelacji) z innymi słowami występującymi w danym korpusie. Lewa strona prezentuje korelację dla pierwszych czterech, najczęściej występujących słów a prawa dla kolejnych czterech. Na przykład słowo „umiejętność” (149 użyć), czyli czwarte najczęściej używane słowo w ofertach pracy dla województwa mazowieckiego (rys. 5), miało największą korelację ze słowami „praca”, „zadanie”, „czas”, „presja”, „zespół” itd. Wartość korelacji dla słowa „umiejętność” przedstawia znak „+”, słowa „znajomość” znak „●”, słowa „doświadczenie” znak „▲” a słowa „projekt” znak „■”. W celu lepszej przejrzystości, dane na wykresie posortowano w porządku malejącym według liczby wystąpień danego słowa w korpusie, a następnie według wartości jego korelacji z innymi słowami.

⁶ Chmury słów dla wszystkich korpusów utworzono z użyciem tych samych parametrów.

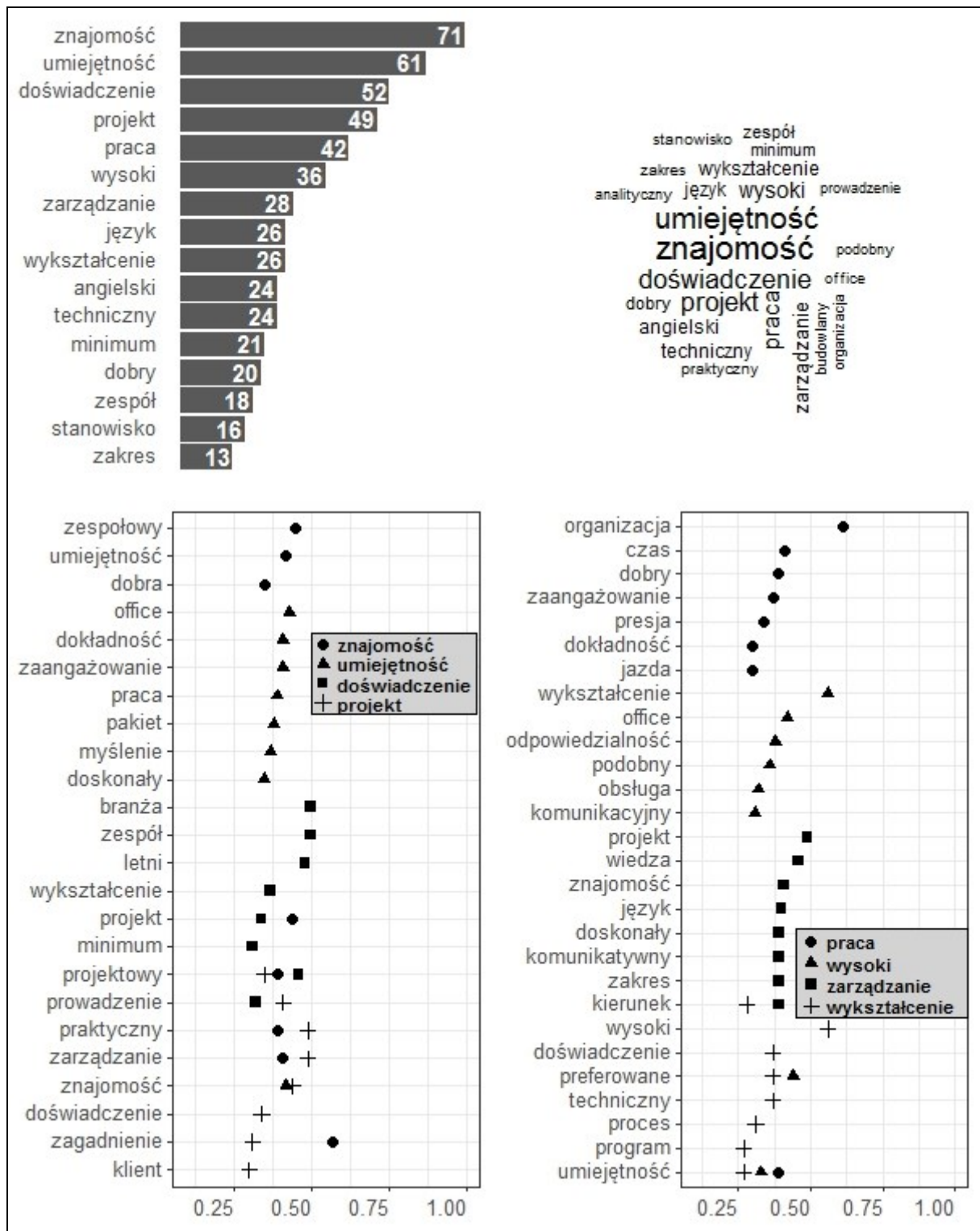


Rysunek 4. Analiza ofert pracy z województwa dolnośląskiego – korpus 1.

Z przedstawionych w tych przykładach słów bez problemu można ułożyć następujące frazy: „umiejętność pracy pod presją czasu”, „umiejętność pracy w zespole”, „znajomość języka angielskiego będzie dodatkowym atutem”, „wysokie zdolności interpersonalne”, „wykształcenie wyższe”, które mogły pojawić się w treści oferty o pracę.



Rysunek 5. Analiza ofert pracy z województwa mazowieckiego – korpus 2.



Rysunek 6. Analiza ofert pracy z województwa śląskiego – korpus 3.

4. Zakończenie

Panująca na rynku konkurencja oraz coraz częstsze stosowanie zarządzania przez projekty sprawia, że projekty mogą być coraz bardziej zróżnicowane a ich złożoność może wzrastać. Kierownicy projektów muszą uwzględniać m.in. dynamicznie zmieniające się czynniki zewnętrzne i wewnętrzne, interesy wielu interesariuszy, a także zmieniający się kontekst biznesowy realizowanego projektu. Jednym z krytycznych czynników sukcesu w zarządzaniu projektami jest kierownik projektu o odpowiednich kompetencjach, które umożliwią mu skuteczne kierowanie projektem.

Należy pamiętać że wykorzystane przez autora rozwiązanie text miningowe nie analizowało znaczenia wyrazów i zdań. Nie uwzględniło również faktu czy dane słowa występowały obok siebie w zadaniu. Ponadto część informacji została utracona na etapie wstępnego przetwarzania tekstu. W procesie tworzenia tzw. worka słów usunięte zostały wszystkie cyfry czyli np. informacja o wymaganym minimalnym doświadczeniu w latach, czy też liczbie zrealizowanych projektów. Wykorzystane rozwiązanie miało natomiast wykryć pewne reguły i prawidłowości dotyczące występowania określonych ciągów słów. Czyli pozyskać nowe wcześniej nieznanne informacje, tj. najczęściej używane słowa oraz ich korelacji z innymi słowami.

Patrząc na najczęściej wykorzystywane słowa można wysunąć następujące wnioski. Pracodawcom często zależało na tym aby pozyskać pracownika z odpowiednim doświadczeniem, znajomością języka obcego (szczególnie angielskiego), posiadającego wykształcenie wyższe. Szukano także pracowników charakteryzujących się dokładnością, zaangażowaniem, odpowiedzialnością czy też komunikatywnością.

Warto dodać, że małym nakładem pracy, analiza może zostać przeprowadzona dla pozostałych korpusów, lub zostać ponownie przeprowadzana w przyszłości.

Bibliografia

1. Armstrong, M. (1928-). et al. (2016). *Zarządzanie zasobami ludzkimi*. Warszawa: Wolters Kluwer.
2. Gładysz, A. (2012). Zastosowanie metod eksploracyjnej analizy tekstu w logistyce. *Logistyka*, nr 3, pp. 643–651. Available at: <http://yadda.icm.edu.pl/baztech/element/bwmeta1.element.baztech-article-BUS6-0042-0021> (Accessed: 28 June 2018).
3. Gunz, H. (1983). *The competent manager: A model for effective performance*, Richard E. Boyatzis, Wiley, New York, 1982. No. of pages: 308. ISBN 0–471–09031-X., *Strategic*

- Management Journal*. Wiley-Blackwell, 4(4), pp. 385-387. doi: 10.1002/smj.4250040413.
4. Hearst, M.A. (1999). Untangling text data mining, *ACL 99 Proceedings of the 37th annual meeting of the Association for Computational Linguistics on Computational Linguistics*. Morristown, NJ, USA: Association for Computational Linguistics, pp. 3–10. doi: 10.3115/1034678.1034679.
 5. Kao, A. Poteet, S.R. (2007). *Natural language processing and text mining*. *Natural Language Processing and Text Mining*. (Ed.), A. Kao and S.R. Poteet. London: Springer London. doi: 10.1007/978-1-84628-754-1.
 6. Mirończuk, M. (2012). Przegląd Metod I Technik Eksploracji Danych Tekstowych. *Studia i Materiały Informatyki Stosowanej*, 4(6), pp. 25-42. Available at: [https://repozytorium.ukw.edu.pl/bitstream/handle/item/3527/Przegląd metod i technik eksploracji danych tekstowych.pdf?sequence=1&isAllowed=y](https://repozytorium.ukw.edu.pl/bitstream/handle/item/3527/Przegl%C4%85d%20metod%20i%20technik%20eksploracji%20danych%20tekstowych.pdf?sequence=1&isAllowed=y).
 7. Musioł-Urbańczyk, A. (2010). Kluczowe kompetencje kierownika projektu, *Organizacja i Zarządzanie*, 10(2). Available at: http://delibra.bg.polsl.pl/Content/27305/BCPS_31016_-_Kluczowe-kompetencje_0000.pdf.
 8. Pawlak, M. (2006). *Zarządzanie projektami*. Warszawa: PWN.
 9. Piwowar-Sulej, K. (2013). Kierownik projektu – charakterystyka profesji. *Nauki Społeczne*. Wrocław, 1(7). Available at: <http://yadda.icm.edu.pl/yadda/element/bwmeta1.element.desklight-d1387a5e-f2ae-4e47-88ae-0a9ed2b54fe7> (Accessed: 4 July 2018).
 10. Wachowiak, P. et al. (2004). *Kierowanie zespołem projektowym*, Warszawa: Difin S.A..
 11. Wyrozębski, P. (2009). Modele kompetencji w zarządzaniu projektami, *E-Mentor*, 2(29), pp. 55–64. Available at: <http://www.e-mentor.edu.pl/artukul/index/numer/29/id/637>.

