

Magdalena Wójcik

Instytut Informacji Naukowej i Bibliotekoznawstwa
Wydział Zarządzania i Komunikacji Społecznej
Uniwersytet Jagielloński

***Big data* w zarządzaniu informacją – przegląd wybranych zagadnień**

Big data in information management –
a review of selected issues

Słowa kluczowe: biblioteki akademickie, biblioteki publiczne, *big data*, wielkie dane, zarządzanie informacją

Keywords: academic libraries, big data, information management, public libraries

Abstrakt

Przedmiot artykułu stanowi zagadnienie wykorzystania tzw. wielkich danych (ang. *big data*) w procesach zarządzania informacją. Celem jest określenie potencjalnego zakresu i form wykorzystania *big data* w bibliotekach oraz konsekwencji wynikających z ich przetwarzania. Bazując na systematycznym przeglądzie piśmiennictwa naukowego i profesjonalnego z lat 2010-2015, dokonano analizy wybranych zagadnień związanych z przetwarzaniem *big data*. Zwrócono uwagę na cechy i specyfikę tego rodzaju danych, ich potencjał dla zarządzania informacją oraz źródła ewentualnych trudności w ich wykorzystaniu w pracy bibliotek publicznych i akademickich. W rezultacie sformułowano wnioski dotyczące potencjału wykorzystania *big data* w procesach zarządzania informacją w bibliotekach.

Abstract

The subject of this article is using big data in the information management processes. The aim is to determine the potential scope and forms of big data uses in libraries, and examine the consequences resulting from their processing. Basing on a systematic review of scholarly and professional literature from years 2010-2015 an analysis of selected topics related to processing of big data was conducted. Attention was drawn to the characteristics and specifics of this kind of data, their potential for information management and sources of possible difficulties in their use in the work of public and academic libraries. As a result, conclusions were drawn regarding the potential use of big data in the information management processes in libraries.

Wprowadzenie

Sfera zarządzania informacją podlega ciągłym przeobrażeniom: zmieniają się narzędzia i strategie, pojawiają się nowe problemy i wyzwania (Galliers, Leidner, 2014). Przyczyną wielu spośród obserwowanych zmian jest dynamiczny rozwój technologii informacyjno-komunikacyjnych, a co za tym idzie globalny wzrost produkcji danych (Gathegi i in., 2014; Ordóñez de Pablos, 2012). Wymusza to na bibliotekarzach, info-brokerach i specjalistach w zakresie zarządzania informacją projektowanie nowych rozwiązań dotyczących gromadzenia i przetwarzania informacji (Burton, Petrie, 2013). Jednym z problemów, które warto w tym kontekście rozważyć jest zarządzanie tzw. wielkimi danymi (ang. *big data*)³.

Przedmiot i cel

Przedmiot artykułu stanowi zagadnienie wykorzystania wielkich danych w procesach zarządzania informacją. Celem jest określenie potencjalnego zakresu i form wykorzystania *big data* oraz konsekwencji wynikających z ich przetwarzania. Do celów szczegółowych można zaliczyć:

- określenie cech i specyfiki wielkich danych,
- zdiagnozowanie potencjału *big data* dla zarządzania informacją,
- wskazanie źródeł ewentualnych trudności w wykorzystaniu wielkich danych w pracy bibliotek publicznych i akademickich,
- sformułowanie wniosków dotyczących potencjału wykorzystania *big data* w procesach zarządzania informacją w bibliotekach.

³ W dalszej części tekstu nazwy wielkie dane i *big data* będą używane zamiennie.

Metoda

Bazując na systematycznym przeglądzie piśmiennictwa naukowego z lat 2010-2015, dokonano analizy wybranych zagadnień związanych z przetwarzaniem wielkich danych. Pod uwagę wzięto wyłącznie prace opublikowane w językach polskim i angielskim. Publikacje wytypowano w oparciu o wyszukiwanie prowadzone w bazach Biblioteki Narodowej (części: *Przewodnik Bibliograficzny*, *Bibliografia Zawartości Czasopism*, *Polska Bibliografia Bibliologiczna*). Przeprowadzono również wyszukiwanie za pomocą narzędzia *Google Scholar*, indeksującego zasoby opublikowane w wielu bazach naukowych (np. *Scopus*, *Web of Science*) i dostarczone przez różnych wydawców (np. *Wiley*, *Elsevier*). Zgromadzony materiał – pełne teksty i/lub abstrakty 50 najbardziej relevantnych prac – poddano kategoryzacji, zwracając szczególną uwagę na wątki takie jak: cechy i specyfika wielkich danych, ich potencjał dla zarządzania informacją oraz źródła ewentualnych trudności w ich wykorzystaniu w pracy bibliotek publicznych i akademickich.

Przeprowadzona analiza pozwoliła ustalić główne kierunki rozważań tej problematyki. Należą do nich tematy, takie jak:

- znaczenie wielkich danych w działalności komercyjnej – wykorzystanie *big data* w promocji, reklamie, PR, handlu i usługach (Kachniewska, 2014; Płoszajski, 2013; Schmarzo, 2013; Chluski, Ziara, 2015),
- logistyczne i techniczne aspekty gromadzenia i przetwarzania wielkich danych (Fan, Bifet, 2013; Wu i in., 2014),
- znaczenie wielkich danych w procesach decyzyjnych przedsiębiorstw i instytucji (LaValle i in., 2011; Liebowitz, 2013; Provost, Fawcett, 2013),
- perspektywy rozwoju wielkich danych (Assunção i in., 2015; Kaisler i in., 2014; Mayer-Schönberger, 2014).

Problematyka wykorzystania wielkich danych była tematem chętnie i często poruszonym na gruncie wielu dziedzin, w tym: informatyki, zarządzania, ekonomii, marketingu, medycyny, bankowości czy architektury. Trudno jednak odnaleźć prace poświęcone wykorzystaniu wielkich danych w bibliotekach oraz publikacje omawiające konsekwencje *big data* dla procesów zarządzania informacją pisane z perspektywy bibliologii i informatologii. Do nielicznych prac reprezentujących ten nurt rozważań można zaliczyć: *Information Management and Big Data: a Framework for Success* (Upton, 2013), *Specialty Construction of the Information Management and Information System Department in View of Big Data* (Na, 2013), *Big Data, Big Libraries, Big Problems?* (Rinne, 2014), *Talking about the Document Delivery Service of Public Library in the Era of Big Data* (Yuxiang, 2015) oraz *The use of self-quantification systems for personal health information: big data management activities and prospects* (Almalki, Gray, Sanchez, 2015).

Warto w tym miejscu zauważyć, że choć problematyka wykorzystania wielkich danych w bibliotekach nie jest w dużym stopniu reprezentowana w literaturze przedmiotu, to jednak jest przedmiotem zainteresowania stowarzyszeń bibliotekarskich i organizacji profesjonalnych (Bieraugel, 2016), jak również jest szeroko komentowana w portalach branżowych i blogosferze (Matthews, 2013; Wilson, 2014).

Cechy i specyfika *big data*

Terminu *big data* używa się najczęściej na oznaczenie dużych, nieuporządkowanych i dynamicznie zmieniających się zbiorów danych. Specyfika Wielkich Danych opisywana jest często w modelu 3V lub 4V. Model 3V (ang. *Volume, Velocity, Variety*) został opracowany w 2001 roku przez firmę analityczną Gartner (wcześniej Meta Group) w celu pokazania głównych cech charakteryzujących *big data* – ich wielkości, różnorodności i zmienności (Laney, 2001). W nowszych interpretacjach, takich jak model 4V, dodaje się jeszcze jedną istotną cechę – wartość (ang. *Value*), rozumianą głównie jako użyteczność dla optymalizacji działań biznesowych (Kachniewska, 2014). W niektórych interpretacjach dodaje się jeszcze jeden element – *Veracity*, czyli wiarygodność surowych danych (Czajkowski i in., 2014).

Ze względu na swoją specyfikę wielkie dane są trudne w przetwarzaniu – wymagają wypracowania nowych rozwiązań zarówno na poziomie logistyki, jak i technologii. Zbiory *big data* są zbyt duże, różnorodne i nieuporządkowane, by można je było przetwarzać za pomocą zwykłych baz danych. Mimo tych trudności coraz więcej instytucji decyduje się zainwestować w rozwiązania umożliwiające zarządzanie wielkimi danymi, biorąc pod uwagę ich potencjalną wartość dla działań biznesowych (Kachniewska, 2014). *Big data* mogą pokazać nowe związki między danymi, ujawnić niewidziane wcześniej tendencje i przyczynić się do powstania nowej wiedzy, która może następnie zostać wykorzystana w celu zwiększania efektywności działań i poprawy rentowności firmy, co w dłuższej perspektywie rekompensuje koszty związane z zakupem specjalistycznego oprogramowania i zatrudnieniem specjalistów (LaValle i in., 2011; Liebowitz, 2013; Provost, Fawcett, 2013).

Komercyjne wykorzystanie *big data*

Wielkie dane są w coraz szerszym zakresie wykorzystywane w sektorze komercyjnym. Łączenie ze sobą danych z wielu źródeł pozwala usprawnić procesy decyzyjne i zarządcze w wielu sektorach, szczególnie w telekomunikacji, turystyce, bankowości, ubezpieczeniach i marketingu (Chmielowski, Przybyszewski; Kachniewska, 2014; EY, 2014). Analiza wielkich danych pozwala lepiej szacować preferencje konsumentów i personalizować kierowaną do nich ofertę, przewidywać ich przyszłe decyzje i szacować wiarygodność klientów (np. kredytową), staje się także nowym narzędziem w procesach pozyskiwania lojalności, zarówno klientów, jak i pracowników firmy (Paharia, 2013). Skuteczne przetwarzanie wielkich danych staje się więc dla przedsiębiorstw źródłem przewagi konkurencyjnej (LaValle i in., 2011). Przedsiębiorstwa, które będą w stanie lepiej wykorzystać wnioski z analizy wielkich danych mogą spodziewać się większych zysków będących rezultatem podejmowania trafnych decyzji (Liebowitz, 2013; Provost, Fawcett, 2013).

Wykorzystanie *big data* w sektorze komercyjnym może mieć szerokie konsekwencje – zarówno pozytywne, jak i negatywne – nie tylko dla firm, ale także dla kon-

sumentów. Z jednej strony przedsiębiorstwa będą w stanie precyzyjniej przewidywać potrzeby klientów i personalizować ofertę, co powinno przyczynić się do polepszenia jakości świadczonych usług, z drugiej pojawiają się obawy o naruszanie prywatności klientów i wysnuwanie daleko idących wniosków (np. dotyczących sytuacji finansowej bądź zdrowotnej) na podstawie danych, których interpretacja nie zawsze bywa trafna. Przemyślenia wymagają zatem kwestie związane z aspektami etycznymi i prawnymi wykorzystania wielkich danych w działalności komercyjnej, tak by wyciąganie wniosków na podstawie analizy *big data* nie naruszało prywatności konsumentów i nie działało na ich szkodę.

Big data w zarządzaniu informacją

Trend polegający na coraz szerszym wykorzystaniu wielkich danych w procesach decyzyjnych dużych przedsiębiorstw ma szerokie konsekwencje dla procesów zarządzania informacją, tak w instytucjach komercyjnych, jak i publicznych (Chluski, Ziora, 2015; Kachniewska, 2014; Płoszajski, 2013; Schmarzo, 2013). Wymusza na specjalistach informacji projektowanie coraz to nowych rozwiązań – na poziomie logistycznym, technicznym i organizacyjnym – umożliwiających efektywne operowanie trudnym i zmiennym materiałem informacyjnym, jaki stanowią wielkie dane. Co istotne, wypracowane do tej pory rozwiązania w zakresie gromadzenia, przetwarzania i udostępniania *big data* muszą ulegać ciągłym modyfikacjom i – podobnie jak same dane – charakteryzować się dużą otwartością i elastycznością.

Procesy zarządzania *big data*, ze względu na swoją specyfikę, wymagają operowania specyficznym zestawem kompetencji, łączącym umiejętności techniczne z wiedzą z zakresu architektury informacji oraz szerokimi kompetencjami analitycznymi. Jedną z najważniejszych konsekwencji upowszechniania koncepcji wielkich danych dla zarządzania informacją jest wykształcenie się nowych zawodów bazujących na klasycznej wiedzy bibliologicznej i informatologicznej, poszerzonej jednak o wysoko-specjalistyczne kompetencje informatyczne i analityczne niezbędne dla organizacji, wyszukiwania i interpretacji wielkich danych. Przykładem mogą być zawody takie jak: *Big Data Analyst*, *Big Data Scientist* czy *Big Data Architect* (Granville, 2014; IBM; Voulgaris, 2014). Na uwagę zasługuje fakt, że zdobycie kompetencji w zakresie obsługi informacyjnej wielkich danych nie jest w Polsce łatwe. W programach studiów kluczowych polskich ośrodków kształcących w zakresie zarządzania informacją, informacji naukowej i bibliotekoznawstwa lub bibliologii zagadnienia związane z *big data* pojawiają się – jeśli w ogóle – jako element wykładów i konwersatoriów⁴. Trudno znaleźć zajęcia praktyczne przygotowujące studentów do podjęcia pracy w nowych zawodach

⁴ Wniosek został wyciągnięty na podstawie przeglądu programów kształcenia i wybranych sylabusów udostępnianych przez wytypowane ośrodki (Uniwersytet Jagielloński, Uniwersytet Pedagogiczny w Krakowie, Uniwersytet Warszawski, Uniwersytet Łódzki, Uniwersytet Śląski) kształcące w zakresie bibliotekoznawstwa, informacji naukowej, bibliologii lub zarządzania informacją.

informacyjnych związanych z wielkimi danymi. Upowszechnienie wiedzy o wielkich danych w środowisku komercyjnym może wpłynąć na oczekiwania klientów z tego sektora wobec infobrokerów, bibliotekarzy i innych specjalistów informacji. Warto byłoby więc zapewnić absolwentom studiów z zakresu zarządzania informacją i bibliotekoznawstwa możliwości zdobycia wiedzy w tym zakresie, czy to w toku studiów I i II stopnia, czy też w formie dodatkowych kursów, szkoleń i studiów podyplomowych.

Przetwarzanie wielkich danych ma wpływ nie tylko na sferę praktycznego zarządzania informacją, zarówno na poziomie indywidualnym, jak i instytucjonalnym, ale wpływa także na sferę metodologii i badań naukowych. Konieczne wydaje się stworzenie teoretycznych modeli przetwarzania wielkich danych i symulowanie – w kontrolowanym środowisku – rozwiązań organizacyjnych i logistycznych związanych z gromadzeniem, przetwarzaniem, opracowaniem, archiwizacją i udostępnianiem *big data*. Przydatne wydaje się również przemyślenie metodologicznych aspektów badania procesów informacyjnych zachodzących w środowisku wielkich danych, w tym metod i technik oceny jakości, użyteczności i funkcjonalności zbiorów wielkich danych oraz wpływu ich wykorzystania na stopień satysfakcji użytkowników. Problematyka *big data* – obecna już w anglojęzycznej literaturze przedmiotu, choć wciąż reprezentowana w niedostatecznym stopniu – stwarza nowe pola badań dla bibliologii i informatologii i wymaga podjęcia zintensyfikowanych prac badawczych w polskim środowisku naukowym.

***Big data* – potencjał dla bibliotek**

Wdrażanie rozwiązań z zakresu gromadzenia i przetwarzania Wielkich Danych wiąże się najczęściej z wysokimi nakładami finansowymi, co może skutecznie zniechęcać instytucje publiczne, w tym biblioteki, do zainteresowania się tymi zagadnieniami. Można już jednak zaobserwować pierwsze inicjatywy w tym zakresie, podejmowane najczęściej przez duże biblioteki. Przykładem może być projekt realizowany przez British Library (BL) w porozumieniu z University College London. Celem projektu, realizowanego z wykorzystaniem oprogramowania *Microsoft Azure*, jest usprawnienie przepływu informacji w naukach humanistycznych. Efektem projektu ma być utworzenie platformy ułatwiającej dostęp do rozproszonych zasobów cyfrowych BL pochodzących głównie z domeny publicznej (Baker, 2014). Podobną inicjatywę realizuje także Library of Congress (LC). W projekcie realizowanym przez LC, przetwarzanie Wielkich Danych ma pomóc m.in. w opracowaniu i udostępnianiu danych pochodzących z serwisów społecznościowych (Zimmer, 2015). W tym kontekście pojawiają się jednak opinie, że wykorzystanie *big data* stwarza możliwości przede wszystkim dla dużych bibliotek, stawiając jednocześnie małe biblioteki publiczne w niekorzystnej sytuacji (Matthews, 2013). Kimberly Matthews zwraca uwagę, że może to doprowadzić do znacznego rozdzwieńku w poziomach usług świadczonych przez biblioteki, które mają dostęp do oprogramowania umożliwiającego przetwarzanie *big data*, a tymi które nie

mogą sobie na to pozwolić (Matthews, 2013). Podobne obawy przedstawia Nathan Rinne, rozważając z kolei etyczne i prawne aspekty przetwarzania wielkich danych (Rinne, 2014).

Przykład British Library i Library of Congress pokazuje, że biblioteki – szczególnie te większe – mogą i powinny zająć istotną pozycję na rynku zarządzania wielkimi danymi, świadcząc w tym zakresie zarówno usługi nieodpłatne dla szerokiej publiczności, jak i komercyjne dla sektora biznesu. Co za tym idzie, bibliotekarze muszą orientować się w problematyce zarządzania wielkimi danymi, choćby po to, by rozumieć procesy zachodzące obecnie na rynku gospodarczym i móc świadczyć, relewantne dla potrzeb biznesu, usługi informacyjne (Bieraugel, 2016). Biblioteki mogą nie tylko świadczyć usługi na zewnątrz, ale także same wykorzystywać efekty przetwarzania *big data* dla usprawniania pracy i lepszej, bardziej spersonalizowanej obsługi użytkowników. Biblioteki mogłyby wykorzystywać zgromadzone dane analogicznie jak firmy komercyjne – do przewidywania potrzeb i zachowań klientów oraz budowania ich lojalności. Bibliotekarze mogliby również prowadzić warsztaty i kursy w zakresie zarządzania wielkimi danymi, pod warunkiem oczywiście, że uprzednio sami zdobyliby wysokie kompetencje w tym zakresie, o co w Polsce trudno. Prowadzenie działalności szkoleniowej i edukacyjnej mogłoby przyczynić się do umocnienia wizerunku bibliotek jako nowoczesnych instytucji, które przodują we wdrażaniu innowacji.

Podsumowując, wydaje się, że biblioteki – przy odpowiednim wsparciu finansowym i merytorycznym – mogłyby zająć istotną pozycję jako podmioty koordynujące zarządzanie Wielkimi Danymi, tak na użytek innych instytucji publicznych i indywidualnych użytkowników, jak i na potrzeby firm komercyjnych. Proces ten nie będzie jednak łatwy – musi odbywać się wieloetapowo, z rozmysłem i uwzględniać warunki finansowe, kadrowe i sprzętowe polskich bibliotek.

Wnioski

Przetwarzanie wielkich danych jest zjawiskiem, które może w znacznym stopniu wpłynąć na sposób zarządzania informacją, tak w sektorze komercyjnym, jak i publicznym. Analiza *big data* może pomóc w przewidywaniu potrzeb i zachowań klientów oraz umożliwić przedsiębiorcom podejmowanie trafnych decyzji biznesowych. Biblioteki i ośrodki informacji będą wcześniej czy później zmuszone określić swoje miejsce w procesach zarządzania wielkimi danymi. Dla wizerunku i przyszłości bibliotek jest istotne, by zajęły one silną pozycję w obszarze zarządzania *big data*. Nie będzie to jednak możliwe bez zapewnienia bibliotekarzom możliwości specjalistycznego kształcenia w zakresie przetwarzania wielkich danych oraz zapewnienia bibliotekom finansowania umożliwiającego zakup nowoczesnego oprogramowania i sprzętu.

Bibliografia

1. Almalki, Manal; Gray, Kathleen; Fernando, Martin Sanchez (2015). The use of self-quantification systems for personal health information: big data management activities and prospects. *Health Information Science and Systems*, vol. 3, suppl 1. <http://www.biomedcentral.com/content/pdf/2047-2501-3-S1-S1.pdf> (odczyt 20.11.2015).
2. Assunção, Marcos D. et al. (2015). Big data computing and clouds: trends and future directions. *Journal of Parallel and Distributed Computing*, vol. 79, pp. 3-15. <http://arxiv.org/pdf/1312.4722.pdf> (odczyt 4.11.2015).
3. Baker, James (2014). *The British Library big data experiment*. <http://britishlibrary.typepad.co.uk/digital-scholarship/2014/06/the-british-library-big-data-experiment.html> (odczyt 3.11.2015).
4. Bieraugel, Mark (2016). *Keeping up with... big data*. http://www.ala.org/acrl/publications/keeping_up_with/big_data (odczyt 5.04.2016).
5. Burton, Paul; Petrie, J. Howard (2013). *Information management technology: A librarian's guide*. Springer. https://books.google.pl/books?id=zhDyBwAAQBAJ&printsec=frontcover&dq=information+management&hl=pl&sa=X&redir_esc=y#v=onepage&q=information%20management&f=false (odczyt 5.11.2015).
6. Chlusiński, Leszek; Ziara, Andrzej (2015). The role of big data solutions in the management of organizations. Review of selected practical examples. *Procedia Computer Science*, vol. 65, pp. 1006-1012. DOI: S1877050915028896 (odczyt 19.11.2015).
7. Chmielowski, Grzegorz; Przybyszewski, Tomasz. *Raport: Praktyczne zastosowanie metod analitycznych big data*. http://www.hbrp.pl/nie_przeocz.php?id=1172&t=raport-praktyczne-zastosowanie-metod-analitycznych-big-data (odczyt 21.11.2015).
8. Czajkowski, Krzysztof et al. (2014). Tools and technologies in the big data environment. *Studia Informatica*, vol. 35, no. 2, pp. 29-40.
9. EY (2014). *Big data changing the way businesses compete and operate*. [http://www.ey.com/Publication/vwLUAssets/EY_-_Big_data:_changing_the_way_businesses_operate/\\$FILE/EY-Insights-on-GRC-Big-data.pdf](http://www.ey.com/Publication/vwLUAssets/EY_-_Big_data:_changing_the_way_businesses_operate/$FILE/EY-Insights-on-GRC-Big-data.pdf) (odczyt 05.04.2016).
10. Fan, Wei; Bifet, Albert (2013). Mining big data: current status, and forecast to the future. *ACM SIGKDD Explorations Newsletter*, vol. 14, no. 2, pp. 1-5. http://www.csd.uoc.gr/~hy562/local_copy/BigData/Large_Scale_Infrastructures_Big_Data_Analytics/MiningBigDataCurrentStatusAndForecast.pdf (odczyt 20.11.2015).
11. Galliers, Robert D.; Leidner, Dorothy E. (2014). *Strategic information management: challenges and strategies in managing information systems*. Routledge. https://books.google.pl/books?id=Y0LJAwAAQBAJ&printsec=frontcover&dq=information+management&hl=pl&sa=X&redir_esc=y#v=onepage&q=information%20management&f=false (odczyt 21.11.2015).
12. Gathegi, John N. et al. eds. (2014). *Challenges of information management beyond the cloud: 4th international symposium on information management in a changing world, IMCW 2013, Limerick, Ireland, September 4-6, 2013. Revised Selected Papers*, vol. 423. Berlin, Heidelberg: Springer. https://books.google.pl/books?hl=pl&lr=&id=ckVTBQAAQBAJ&oi=fnd&pg=PP5&dq=%22new+challenges%22+%22information+management%22&ots=tADGW0zatq&sig=Vb0S6T1nC0G6yIQsvGp9kOOTdEw&redir_esc=y#v=onepage&q=%22new%20challenges%22%20%22information%20management%22&f=false (odczyt 19.11.2015).

13. Granville, Vincent (2014). *Developing analytic talent: becoming a data scientist*. Indianapolis: John Wiley & Sons. https://books.google.pl/books?hl=pl&lr=&id=tp46AwAAQB AJ&oi=fnd&pg=PR21&dq=%22big+data+scientist%22&ots=e7Et9P1cgL&sig=vHIwWrecocsTUtuu9xc0Ae1LtDU&redir_esc=y#v=onepage&q=%22big%20data%20scientist%22&f=false (odczyt 7.11.2015).
14. IBM. *What is a data scientist?* <http://www-01.ibm.com/software/data/infosphere/data-scientist/> (odczyt 20.11.2015).
15. Kachniewska, Magdalena (2014). Big data analysis jako źródło przewagi konkurencyjnej przedsiębiorstw i regionów turystycznych. *Folia Turistica*, nr 32, s. 35-54. http://www.folia-turistica.pl/attachments/article/430/FT_32_2014.pdf#page=35 (odczyt 9.11.2015).
16. Kaisler, Stephen et al. (2013). Big data: issues and challenges moving forward. *System Sciences (HICSS), 46th Hawaii International Conference on System Sciences*. <http://www.computer.org/csdl/proceedings/hicss/2013/4892/00/4892a995.pdf> (odczyt 1.11.2015).
17. Laney, Doug (2001). *Application delivery strategies*. <http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf> (odczyt 21.11.2015).
18. LaValle, Steve (2011). Big data, analytics and the path from insights to value. *MIT Sloan Management Review*. <http://sloanreview.mit.edu/article/big-data-analytics-and-the-path-from-insights-to-value/> (odczyt 19.11.2015).
19. Liebowitz, Jay (2013). *Big data and business analytics*. Boca Raton FL [etc.]: CRC Press: Taylor & Francis Group.
20. Matthews, Kimberly (2013). *Big data vs. little library*. <http://21stcenturylibrary.com/2013/06/13/big-data-vs-little-library/> (odczyt 7.11.2015).
21. Mayer-Schönberger, Viktor (2014). *Big data: rewolucja, która zmieni nasze myślenie, pracę i życie*. Warszawa: MT Biznes.
22. Na, Yan (2013). Specialty construction of the information management and information system department in view of big data. *Research on Library Science*, vol. 11, pp. 9-12.
23. Ordóñez de Pablos, Patricia (2012). *Advancing information management through Semantic Web concepts and ontologies*. Hershey, PA: Information Science Reference. https://books.google.pl/books?id=-3cSldbU8b0C&printsec=frontcover&dq=information+management&hl=pl&sa=X&redir_esc=y#v=onepage&q=information%20management&f=false (odczyt 8.11.2015).
24. Paharia, Rajat (2013). *Loyalty 3.0: how big data and gamification are revolutionizing customer and employee engagement*. New York [etc.]: McGraw-Hill Education.
25. Płoszajski, Piotr (2013). Big data: nowe źródło przewag i wzrostu firm. *E-mentor*, nr 3 (50), s. 5-10.
26. Provost, Foster; Fawcett, Tom (2013). Data science and its relationship to big data and data-driven decision making. *Big Data*, vol. 1, no. 1 pp. 51-59. http://www.researchgate.net/profile/Tom_Fawcett/publication/256439081_Data_Science_and_its_relationship_to_Big_Data_and_data-driven_decision_making/links/02e7e5228cce561fd4000000.pdf (odczyt 3.11.2015).
27. Rinne, Nathan (2014). Big data, big libraries, big problems?: the 2014 LibTech Anti-talk? In: *2014 Library Technology Conference, Macalester College, St. Paul, Minnesota, March 19 & 20, 2014*. (Unpublished) [Conference paper]. <http://eprints.rclis.org/22750/1/LTC%202014%20final2.pdf> (odczyt 3.11.2015).

28. Schmarzo, Bill (2013). *Big Data: understanding how data powers big business*. Indianapolis: John Wiley & Sons.
29. Upton, Sheila (2013). Information management and big data: a framework for success. *IVI Executive Briefing Series*. <http://eprints.maynoothuniversity.ie/6393/1/IVI%20Exec%20Briefing%20-%20Big%20Data%20%2B%20EIM%20v1.0.pdf> (odczyt 19.11.2015).
30. Voulgaris, Zacharias (2014). *Data scientist: the definitive guide to becoming a data scientist*. Basking Ridge, NJ: Technics Publications.
31. Wilson, Christina (2014). *Why big data matters to libraries*. <http://inalj.com/?p=65209> (odczyt 9.11.2015).
32. Wu, Xindong et al. (2014). Data mining with big data, knowledge and data engineering. *IEEE Transactions on*, 26.1 (2014), p. 97-107. <http://lansainformatics.com/wp-content/plugins/project-mgt/file/upload/pdf/2440Data-mining-with-big-data-pdf.pdf> (odczyt 21.11.2015).
33. Yuxiang, Yang (2015). Talking about the document delivery service of public library in the era of big data. *Sci-Tech Information Development & Economy*, no. 1, pp. 8-10.
34. Zimmer, Michael (2015). The Twitter archive at the Library of Congress: challenges for information practice and information policy. *First Monday*, vol. 20, no 7. <http://firstmonday.org/ojs/index.php/fm/article/view/5619/4653> (odczyt 20.11.2015).