

KRYSTYNA HUDZIK
Biblioteka Główna
Uniwersytet Marii Curie-Skłodowskiej w Lublinie
e-mail: k.hudzik@umcs.pl
ORCID: 0000-0002-1449-828X

DATA LIBRARIAN – NOWY PROFIL ZAWODOWY W BIBLIOTEKACH NAUKOWYCH



Krystyna Hudzik – dr, pracuje w Bibliotece Głównej UMCS w Lublinie. Autorka książki i kilkunastu artykułów dotyczących niemieckiego bibliotekarstwa. Swoje zainteresowania rozwijała dzięki stażom i praktykom odbytym w bibliotekach uniwersyteckich w Niemczech i w Austrii (Jena, Kolonia, Tybinga, Konstancja, Karlsruhe, Wiedeń) oraz w trakcie rocznego pobytu w Instytucie Nauki o Bibliotece i Informacji Uniwersytetu Humboldtów w Berlinie. Książka *Biblioteki uniwersyteckie w Niemczech w epoce cyfrowej: ciągłość i zmiana* otrzymała wyróżnienie Rektora Uniwersytetu Warszawskiego dla publikacji akademickiej w dziedzinie nauk społecznych i humanistycznych w konkursie na najlepszą książkę akademicką i naukową ACADEMIA 2018.

SŁOWA KLUCZOWE: Badacz danych. Bibliotekarz danych. Big Data. Dane badawcze. *Data librarian* – kompetencje. *Data librarian* – kształcenie. Kompetencje w zakresie danych.

ABSTRAKT: Teza/cel artykułu – *Data librarian* funkcjonuje w zagranicznych bibliotekach naukowych jako specjalista w zakresie danych badawczych. Przedstawiono kompetencje wymagane na tym stanowisku w akademickich bibliotekach w USA oraz program kształcenia akademickiego *data librarian* na przykładzie wybranej uczelni w Niemczech. Kształtowanie się nowej specjalności w zawodzie bibliotekarza omówiono w szerszym kontekście kultury cyfrowej, procesu datafikacji rzeczywistości, produkcji Big Data oraz kompetencji w zakresie danych. **Metody badań** – Analiza treści publikacji na temat *data librarian*, danych badawczych i zadań bibliotek naukowych. Uzupełnienie stanowią informacje na stronach internetowych wybranych uczelni, repozytoriów i bibliotek. Wykorzy-

stano także głosy w dyskursie prowadzonym wokół Big Data, datafikacji i przemian z nimi związanych. **Wyniki/Wnioski** – *Data librarian* jest specjalistą o wysokich kompetencjach w zakresie programów i sieci informatycznych oraz systemów informacyjnych. Zakres jego umiejętności w zarządzaniu danymi badawczymi odpowiada w dużym stopniu zadaniom realizowanym przez *data scientists*. Dalszy rozwój i stabilizacja tych stanowisk będzie uzależniona od formalnego określenia ram przypisanych im kompetencji i zadań oraz związanych z nimi programów kształcenia. Programy nowych kierunków studiów powinny uwzględniać ocenę zewnętrznego środowiska praktyków – w tym stowarzyszeń zawodowych bibliotekarzy i związków instytucjonalnych bibliotek.

WSTĘP

Artykuł dotyczy problemu kształtowania się nowej specjalności w zawodzie – bibliotekarza danych – w związku z rosnącym znaczeniem danych badawczych w procesie naukowym i zasadą otwartego do nich dostępu. W literaturze przedmiotu – anglojęzycznej, polskiej i niemieckiej – najczęściej używana jest jego oryginalna nazwa *data librarian*. Tekst składa się z czterech części: w pierwszej, przedstawiono ogólny kontekst przemian kulturowych pod wpływem masowej produkcji danych, tzw. datafikacji rzeczywistości oraz przetwarzania Big Data. Część druga, poświęcona jest procesom naukowym związanym z danymi – chodzi o powstanie nowej dyscypliny naukowej i kierunku studiów o nazwie nauka o danych (*data science*) oraz kształcenia kompetencji w zakresie danych (*data literacy*) i zawodu badacza danych (*data scientist*). W części trzeciej, skoncentrowano się na zagadnieniu danych badawczych – co to znaczy? – oraz związanych z nimi zadaniach bibliotek naukowych. Omówiono kwestię miejsca danych w procesie poznania naukowego w standardowym modelu hierarchii mądrości (ang. *wisdom hierarchy*). Ostatnia, czwarta, część artykułu stanowi próbę odpowiedzi na kluczowe pytania o zakres kompetencji i program kształcenia akademickiego *data librarian*.

BIG DATA I DATAIZM

Nauka i komunikacja naukowa to niejedynie dziedziny aktywności społecznej, które w ciągu ostatnich lat wykazują duże zainteresowanie danymi oraz kwestią ich gromadzenia, przechowywania i wykorzystania. Problem dotyczy niemal wszystkich aspektów naszego życia, ponieważ nieustannie rośnie ilość wytwarzanych danych w instytucjach publicznych, sektorze finansowym, handlu, medycynie i edukacji. Jednocześnie prawie każdy z nas ma swój osobisty i bezpośredni udział w generowaniu danych w związku z postępującą informatyzacją codzienności (ang. *everyday turn*). Korzystając z własnych komputerów, smartfonów, aparatów cyfrowych, kart płatniczych i innych urządzeń z dostępem do Internetu

staliśmy się producentami Big Data, wytwarzamy ogromną ilość różnorodnych danych, które w wielu przypadkach są przekazywane dalej zainteresowanym instytucjom, tak jak to czyni na przykład Google Maps, identyfikując w smartfonach miejsce naszego pobytu, czy też jak to się dzieje z danymi z opłat kartami bankowymi, które wykorzystywane są w interpretacji czasu i preferencji naszych zakupów. „Datafikacja” lub inaczej „danetyzacja” to nowo powstałe wyrazy, neologizmy zapożyczone od angielskiego *datafication*, które odnoszą się do „kwantyfikowania i przekładania wszelkich elementów rzeczywistości na dane, by móc je poddać agregowaniu i algorytmizacji”¹. To dzisiaj modne słowa-slogany oznaczające przemysł tworzenia (ang. *datify*) Big Data. Przemysł ten polega na przekształcaniu społecznych działań na dane liczbowe oraz opisywaniu ich za pomocą wskaźników i algorytmów. Kwantyfikacja, czyli ujęcie ilościowe, przedstawia określony proces lub działanie w formie danych, które mogą być monitorowane, analizowane, optymalizowane i wykorzystywane w różnych celach.

Temat danych znajduje się obecnie w centrum dyskusji o cyfrowości i zajmuje umysły nie tylko informatyków, menadżerów i ekonomistów, lecz także filozofów kultury i mediów oraz socjologów. Ich spojrzenia na zachodzące procesy są ważne, ponieważ pochodzą z szerszej humanistycznej i historycznej perspektywy. W zachodzących przemianach cywilizacyjnych usiłują oni odnaleźć człowieka jako indywidualium, pytają o znaczenie jego przeżyć i doświadczeń, wreszcie o jego miejsce w świecie. Tak należałoby rozumieć chociażby podejście medioznawcy Kazimierza Krzysztofka, który zaczyna jeden ze swoich artykułów od biblijnych słów z Księgi Daniela (5, 131): „*Mane-Tekel-Fares* (zważono, policzono, rozdzielono), które to słowa, po aramejsku (מנה, מנה, תקל, ופרסוי) niewidzialna ręka wypisała na ścianie, podczas uczyty Baltazara” (Krzysztofek, 2012). Nomen omen słowa te okazały się złowieszcze i tragiczne dla króla Baltazara – były zapowiedzią upadku jego królestwa. Autor opisuje, w jaki sposób narodziła się nowa antropometria, zdolna zmierzyć wiele z tego, co wcześniej było niepoliczalne i niemierzalne; dla której nie istnieje nic, co nie ma wartości rynkowej. Oto dalsze wyjaśnienie: „Na jakiejś skali muszą być mierzone wiedza, emocje, doznania, przeżycia, wrażenia estetyczne i artystyczne (np. gimnastyków, skoczków z wieży, łyżwiarzy figurowych), kwalifikacje, inteligencja (IQ). Na pytanie dlaczego, odpowiedź wydaje się dosyć prosta: bo to wszystko jest przedmiotem obrotu rynkowego, musi

¹ Słownik Języka Polskiego PWN. *Datafikacja czy danetyzacja*. [online], [dostęp: 13.04.2020]. Dostępny w WWW: <https://sjp.pwn.pl/poradnia/haslo/Datafikacja-czy-danetyzacja;19546.html>.

mieć swą liczbową wielkość, niezbędną do finansowego oszacowania. Dziś jest to szczególnie istotne, gdy gamifikacja, ukonkursovienie wkracza do wszystkich dziedzin życia: korporacji, polityki, nauki itp.” (Krzysztofek, 2012, cz. I). Według logiki dzisiejszej cywilizacji numerycznej wszelkie wartości: ekonomiczne, poznawcze, moralne, społeczne – podlegają algorytmizacji, zostają zamienione na informacje według binarnego kodu, a każde ludzkie działanie można zautomatyzować zgodnie z przyjętymi regułami/wskaźnikami racjonalności i efektywności. Retoryczne pozostaje pytanie, czy taki projekt wszechdatafikacji jest podszyty pewną ideologią i czy nie wróży mu to tego samego losu, który spotkał króla Baltazara...

Zasadne stają się wobec tego nasze obawy o przyszłość – zdanie „Jutro będę algorytmem” coraz mniej brzmi nam jak cytata z filmu *science fiction* (Wilk, 2017). To tytuł jednej z recenzji książki pt. *Homo deus: krótka historia jutra* – jej autorem jest Yuval Noah Harari, historyk i futurysta izraelski, odnoszący duże sukcesy w USA jako ikona popularnonaukowych teorii przyszłości. Upowszechnił on pojęcie dataizmu jako wiary w dane oraz wyraz przekonania, że ludzie nie radzą sobie z ich przetwarzaniem, powinni więc przekazać kluczowe kompetencje sztucznej inteligencji i – mówiąc metaforycznie – oddać swój los w jej ręce. W tym znaczeniu dataizm staje się pojęciem otwartym, o bardzo rozległej i niejednolitej klasie przedmiotów/zjawisk, do których się odnosi – oznacza ideologię, pewien sposób myślenia, a nawet nową formę religii lub teorię transhumanistyczną (Harari, 2018, s. 467-504).

DATA SCIENCE, DATA LITERACY, DATA SCIENTIST: DYSCYPLINA NAUKOWA, KIERUNEK STUDIÓW, KOMPETENCJE I ZAWÓD

Datafikacja i dataizm to terminy, które pojawiają się w kontekście zmian cywilizacyjnych diagnozowanych jako początek przejścia od wieku informacji do ery *Big Data*, w czasie, w którym wzrost ilości danych jest tak duży, jak nigdy dotąd. W rezultacie tych zmian powstaje konkurencyjna do nauki o informacji dyscyplina o nazwie nauka o danych (ang. *data science*). Ma ona znamiona nauki otwartej, w sensie rozwijającej się, *in statu nascendi* – wyrasta bowiem z podążających za stale zmieniającymi się technologiami informatycznymi praktyk postępowania z danymi (podobnie jak przed kilkoma dekadami nauka o informacji z działalności informacyjnej); ma charakter interdyscyplinarny, mocno osadzona jest w informatyce, korzysta z matematyki, fizyki, statystyki, programowania komputerowego, eksploracji danych i analityki predykcyjnej; zajmuje się procesami pozyskiwania danych, ich wykorzystania, wizualizacji i wnioskowania na ich podstawie. Używając jeszcze bardziej ogólnych formuł, można powiedzieć, że *nauka o danych* zajmuje się metodami, procesami,

algorytmami i systemami służącymi ekstrakcji wiedzy z danych. (Data, 2018). Termin *data science* pojawił się już w latach 90. XX w., a w kwietniu 2002 r. ukazał się pierwszy numer czasopisma poświęconego nowej dyscyplinie². Szybko też stała się ona przedmiotem nauczania, początkowo na uczelniach amerykańskich, gdzie cieszy się dużym zainteresowaniem do dziś, a nieco później także w Europie. W Niemczech studia licencjackie i magisterskie z jej zakresu prowadzone są obecnie w kilkunastu uniwersytetach technicznych i szkołach wyższych (Data, 2018, s. 26).

Obok nowej dyscypliny powstaje także pokrewny i młodszy od niej przedmiot kształcenia *data literacy*, który nabiera coraz większego znaczenia w dwóch różnych aspektach: zawodowego kształcenia informatycznego oraz ogólnego kształcenia w zakresie korzystania i postępowania z danymi (Data, 2018, s. 6-13). I właśnie ten drugi wymiar, edukacyjny, nabywania kompetencji w zakresie danych w nauczaniu szkolnym i akademickim staje się ważny z punktu widzenia działalności bibliotek. Nowy klucz kompetencji *data literacy*, paralelny w nazwie do kompetencji informacyjnych (*information literacy*), należałoby postrzegać nie jako konkurencyjny wobec tych ostatnich, lecz jako bardziej adekwatny do obecnego stanu cyfryzacji oraz masowej produkcji danych, baz danych i systemów baz danych. *Data literacy* odnosi się do zdolności czytania i rozumienia danych jako informacji, jej sformułowania i komunikowania oraz racjonalnego wykorzystania danych w określonym kontekście. Postępowanie z danymi zgodnie z planem obejmuje ich opracowanie, zarządzanie, przechowywanie, analizowanie, wizualizację, interpretację, kontekstualizację, ocenę i zastosowanie. Przede wszystkim chodzi o kształtowanie nawyków rutynowego myślenia, zgodnie z którym dane należy gromadzić, udostępniać i przekazywać dalej w celu ich wykorzystania. Ważne są także praktyczne umiejętności analizy danych tworzonych przez media, wielkich i małych przedsiębiorców, publiczne instytucje państwowe i społeczne. Edukacja w tym zakresie wspiera rozwój cyfryzacji i globalnego społeczeństwa wiedzy w różnych dziedzinach życia (Wolff, 2018, s. 225-227).

Christian Wolff, profesor informatyki mediów na Uniwersytecie w Ratyźbonie, rozważa kompetencje informacyjne oraz kompetencje w zakresie danych pod wieloma względami, porównuje je ze sobą i wskazuje ich tendencje rozwojowe. W podsumowaniu swoich analiz stwierdza, że *data science* i *data literacy* uzupełniają dotychczasowe modele kompetencji przez to, że ukazują specyfikę problematyki danych i udzielają odpowiedzi na nowe wyzwania stawiane przez Big Data. Według Wolffa *data literacy* jest kolejną ważną kompetencją w kontekście cyfryzacji, nie może być ona jednak traktowana w niej jako centralna, ponieważ oznaczałoby

² *Data Science Journal* vol. 1 (2002). [online], [dostęp: 18.04.2020]. Dostępny w WWW: https://www.jstage.jst.go.jp/browse/dsj/1/0/_contents.

to redukcję całej problematyki cyfryzacji tylko do wymiaru danych. Informatyk mediów uważa oczywiście, że szerszego wyjaśnienia *data literacy* można udzielić z perspektywy takich dyscyplin, jak informatyka oraz nauka o informacji (obydwie zawierają w nazwie pojęcie informacji). Podobnie *data science*, jego zdaniem, ma przyczynić się do większych efektów w edukacji i nabywaniu kompetencji. Ma ona również wyjaśnić problem nierozdzielności obszarów danych i informacji oraz przejścia między danymi i informacją z punktu widzenia tego, kiedy, pod jakimi warunkami i w jakich celach do tego dochodzi? Z tego względu wydaje się uzasadnione tworzenie nadrzędnych koncepcji określanych jako metakompetencje (*metaliteracy*) lub kompetencje cyfrowe (*digital literacy*).

W uzupełnieniu przeglądu przemian w zakresie badań i kształcenia związanych z Big Data jeszcze krótka wzmianka o nowym zawodzie – badacz danych (ang. *data scientist*), który ukształtował się w praktyce, szybko stając się cenionym i poszukiwanym fachowcem na rynku pracy na całym świecie. *Data scientists* stanowią grupę specjalistów zajmujących się analizą danych w dużych, niespójnych i różnorodnych bazach (Big Data), niemających uporządkowanej struktury. Posiadają oni wprawdzie wiele cech wspólnych ze znanym od dawna, szczególnie w ekonomii i biznesie, zawodem analityka danych, jednak ich umiejętności są szersze niż tylko stosowanie metod statystycznych i obejmują także znajomość najnowszych technologii zarządzania danymi. Nazwa zawodu *data scientist* pojawiła się w przestrzeni publicznej przed kilkoma laty za sprawą artykułu pod dość przewrotnym i zarazem ekscytującym tytułem – *Data Scientist: The Sexiest Job of the 21st Century (Badacz danych: najbardziej seksowny zawód XXI wieku)* – w „Harvard Business Review” (Davenport Thomas H., Patil D. J., 2012).

Na koniec tych uwag warto przyjrzeć się temu, jak wygląda u nas – w Polsce – zapotrzebowanie na specjalistów w zakresie *data science* oraz jakie kierunki studiów są im oferowane przez nasze uczelnie wyższe. Na stronach internetowych wybranych trzech prestiżowych uczelni technicznych widnieją następujące kierunki:

1. Data Science/Inżynieria i analiza danych – Politechnika Warszawska, Wydział Matematyki i Nauk Informacyjnych. Studia stacjonarne II stopnia 2020/2021³.

2. Analiza danych – Data Science – Akademia Górniczo-Hutnicza w Krakowie, Wydział Informatyki, Elektroniki i Telekomunikacji. Oferta studiów podyplomowych na rok akademicki 2019/2020 AGH⁴.

³ Politechnika Warszawska. [online], [dostęp: 13.04.2020]. Dostępny w WWW: <https://www.pw.edu.pl/Rekrutacja/Studia-I-i-II-stopnia/Studia-stacjonarne-II-stopnia/Oferta-studiow-stacjonarnych-2020-2021>.

⁴ Akademia Górniczo-Hutnicza w Krakowie. [online], [dostęp: 13.04.2020]. Dostępny w WWW: <https://www.agh.edu.pl/ksztalcenie/oferta-ksztalcenia/studia-podyplomowe-kursy-doksztalcajace-i-szkolenia/analiza-danych-data-science/>.

3. Big Data Analytics – Politechnika Wrocławska, Wydział Podstawowych Problemów Techniki. Studia stacjonarne w języku angielskim⁵.

Wymienione kierunki są studiami typowo inżynierskimi dla informatyków, którzy raczej nie trafiają do uniwersyteckich bibliotek. Warto jednak zapoznać się z programem studiów i kompetencjami absolwentów z uwagi na przyszłe kształcenie *data librarian*. Wybrany przykład stanowi opis kierunku Analiza danych – Data Science zamieszczony na stronie WWW AGH w Krakowie. Cytat: „Oferta studiów kierowana jest do wszystkich, którzy będą zajmowali się szeroko rozumianą analizą oraz przetwarzaniem, często bardzo dużych zbiorów różnorodnych danych. Pojęcia «Analiza danych» oraz «Data Science» łączą kompetencje z wielu dziedzin, specjalności takich jak: matematyka, statystyka, bazy danych, programowanie. Posiadanie takiej wiedzy oraz kompetencji i umiejętności jest coraz częściej pożądaną na rynku IT”⁶. Wgląd w pełny program studiów daje lista przedmiotów: 1) Wprowadzenie do analizy danych i data science, 2) Statystyka, 3) Bazy danych, 4) Programowanie w języku Python, 5) Ekstrakcja danych ze źródeł internetowych, 6) Hurtownie danych, 7) Analiza dużych zbiorów danych w środowisku Spark, 8) Uczenie maszynowe, 9) Eksploracja danych, 10) Analiza danych tekstowych, 11) Sieci społeczne, 12) Analiza danych przestrzennych, 13) Wizualizacja dużych zbiorów danych, 14) Prawne aspekty analizy danych.

Program studiów przyszłych menadżerów danych nasuwa pytania o kształcenie i kompetencje *data librarians* (bibliotekarzy danych). Jakie powinni mieć kwalifikacje? Czy wiedza z wyżej wymienionych przedmiotów byłaby również dla nich pożyteczna i niezbędna do profesjonalnego zarządzania danymi badawczymi, na przykład w dziedzinie chemii czy nauk o Ziemi? Na tego rodzaju pytania nie mamy, jak na razie, odpowiedzi. Są one tym trudniejsze w sytuacji, gdy ostatnio coraz głośniejszą rolę w naszym kraju odgrywa kryzys studiów bibliotekoznawczych w polskich ośrodkach akademickich (Grygowski, 2020). Takie diagnozy uświadamiają nam, że stoimy właściwie przed koniecznością zmierzenia się ze znacznie bardziej podstawowym problemem – czy w ogóle potrzebny jest nam dzisiaj profesjonalny bibliotekarz po deregulacji zawodu bibliotekarza wprowadzonej w roku 2013. Na to pytanie muszą najpierw odpowiedzieć decydenci – zarówno ci polityczni, jak i merytoryczni – rektorzy uczelni i dyrektorzy bibliotek.

⁵ Politechnika Wrocławska. [online], [dostęp: 13.04.2020]. Dostępny w WWW: <http://rekrutacja.pwr.edu.pl/wyszukiwarka-kierunkow-studiow/big-data-analytics-w11-ii-st-10481.html>.

⁶ Analiza danych – Data Science. Charakterystyka studiów. [online], [dostęp: 11.05.2020]. Dostępny w WWW: <https://www.agh.edu.pl/ksztalcenie/oferta-ksztalcenia/studia-podyplomowe-kursy-dokszalcajace-i-szkolenia/analiza-danych-data-science/>.

DANE BADAWCZE, PIRAMIDA WIEDZY I BIBLIOTEKI NAUKOWE

Wymienione tematy są obecne w rodzimej literaturze i dyskusjach specjalistów, z wyraźną przewagą danych badawczych z tej racji, że zasada ich gromadzenia, przechowywania i otwartego dostępu została już sformalizowana, objęta zaleceniami i stała się obowiązkowa. Zawdzięczamy to wytycznym Unii Europejskiej, programowi Horyzont 2020 (Wałek, 2019) oraz Narodowemu Centrum Nauki, które wprowadziło w 2019 r. wymóg opracowywania planu zarządzania danymi (PZD) w składanych wnioskach grantowych. Plan jako integralna część wniosku jest oceniany przez Zespół Ekspertów zarówno na etapie wstępnej oceny wniosku, jak i końcowego raportu z realizacji projektu badawczego⁷.

Ważną krajową inicjatywą w tym zakresie jest tworzenie Repozytorium Otwartych Danych (RepOD) przez Interdyscyplinarne Centrum Modelowania Matematycznego i Komputerowego Uniwersytetu Warszawskiego (ICM UW) w ramach działań Platformy Otwartej Nauki. W założeniu jest ono przeznaczone dla wszystkich naukowców w Polsce i umożliwia otwarte udostępnianie danych badawczych, szczególnie tzw. małych danych, tworzonych w trakcie prac prowadzonych przez pojedynczych badaczy oraz niewielkie zespoły naukowe⁸. Statystyki repozytorium prezentowane na stronie internetowej w dniu 16.04.2020 r. są następujące: zbiory danych – 129; użytkownicy – 242; ilość danych – 854,2 GB.

RepOD wykazuje baza indeksująca 2450 instytucjonalnych, dziedzinowych i tematycznych repozytoriów danych tworzonych przez uniwersytety i instytuty badawcze na świecie pod nazwą re3data.org (*Registry of Research Data Repositories*). Realizacja projektu rozpoczęła się w 2012 r. dzięki funduszom Niemieckiej Wspólnoty Badawczej (*Deutsche Forschungsgemeinschaft – DFG*), która ma kontynuować finansowanie portalu re3data od stycznia 2020 r. przez kolejne trzy lata w ramach grantu przeznaczonego na rozwój schematu metadanych, unowocześnienie serwisu oraz poprawę cytowania danych badawczych w publikacjach. W przedsięwzięciu uczestniczą: Instytut Technologiczny w Karlsruhe, Centrum Wspólnoty Helmholtza w Poczdamie – Niemieckie Centrum Badań Ziemi, Instytut Nauk o Bibliotece i Informacji Uniwersytetu Humboldtów w Berlinie oraz DataCite⁹.

⁷ Wytyczne dla wnioskodawców do uzupełnienia PLANU ZARZĄDZANIA DANymi w projekcie badawczym. [online], [dostęp: 14.04.2020]. Dostępny w WWW: https://www.ncn.gov.pl/sites/default/files/pliki/regulaminy/wytyczne_zarzadzanie_danymi.pdf; Narodowe Centrum Nauki. Posiedzenie Rady NCN 17-18 kwietnia 2019 r. [online], [dostęp: 14.04.2020]. Dostępny w WWW: <https://www.ncn.gov.pl/aktualnosci/2019-05-07-posiedzenie-rady-kwiecien-2019>.

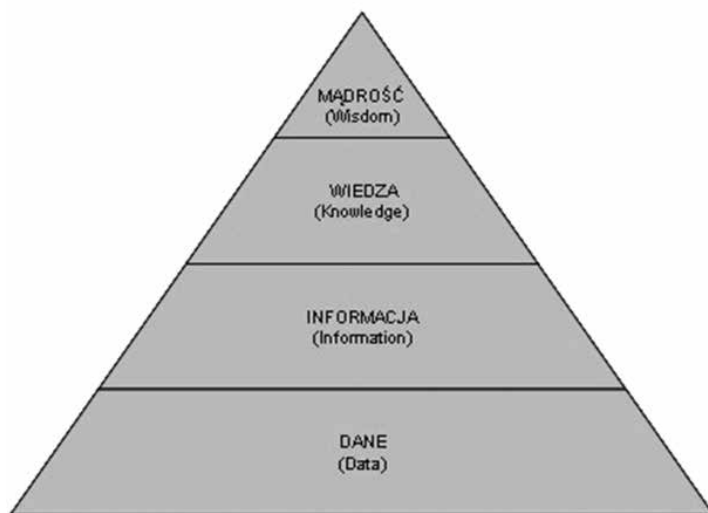
⁸ Repozytorium Otwartych Danych – RepOD. [online], [dostęp: 15.04.2020]. Dostępny w WWW: <https://reporod.pon.edu.pl/pl/about>.

⁹ German Research Foundation to fund new services of re3data. [online], [dostęp: 19.04.

Dane badawcze tradycyjnie stanowią część pracy naukowej, która w przeszłości pozostawała w cieniu jej opublikowanych efektów. W wyniku ilościowego wzrostu danych w procesie badawczym pojawiła się w latach 90. minionego wieku dyskusja o potrzebie zabezpieczenia danych stanowiących podstawę publikacji, co skutkowało zaleceniami o ich przechowywaniu na trwałych i bezpiecznych nośnikach w instytucji, w której powstają (np. w Niemczech było to zalecenie DFG o zasadach dobrej praktyki naukowej z 1998 r.). Zmiana paradygmatu polegająca na postawieniu kwestii danych w centrum pracy naukowej, a nie jak dotychczas na pozycji „szarych” materiałów, danych „skoroszytowych”, jak określało je w przeszłości, przybrała na intensywności w kontekście rozwijanej idei otwartego dostępu do publikacji oraz otwartej nauki. Analogicznie powstała koncepcja otwartej dostępności również danych badawczych, dzielenia się nimi, ponownego ich wykorzystania, łączenia różnych danych i tekstów w celu prowadzenia zaawansowanych analiz. Wymienione zadania stały się kluczowe dla rozwoju cyfrowego otoczenia nauki i wyznaczyły praktykę budowania infrastruktury dla danych badawczych umożliwiającą ich publikowanie (repozytoria), przechowywanie, powtórne wykorzystanie i długoterminową archiwizację. Zasada otwartego dostępu stała się sprawą nadrzędną, bo dotycząca finansowania badań z pieniędzy publicznych i wydatkowania na ten cel wysokich kwot przez poszczególne państwa, dlatego pojawił się postulat – pod hasłem: dane badawcze „jedne dla wszystkich” – by raz zgromadzone dane były udostępniane innym naukowcom i instytucjom.

Rozważając krótko kwestię miejsca danych w procesie poznania naukowego, warto odwołać się do standardowego modelu hierarchii mądrości (ang. *wisdom hierarchy*) przedstawianego w kształcie piramidy opisanej za pomocą czterech terminów: dane – informacja – wiedza – mądrość. Model nosi skrótową nazwę DIKW (ang. *Data – Information – Knowledge – Wisdom*), (Frické, 2018). To dobrze znana, powielana w literaturze przedmiotu wizualizacja.

Cztery pojęcia tworzą oś konceptualną i wyznaczają perspektywy poznania rzeczywistości. Obecnie punkt ciężkości tej osi leży u samej podstawy piramidy, którą stanowią dane; poprzednio zainteresowanie ogniskowało się wokół informacji. Uwagę zwraca dysproporcja między rozległą przestrzenią danych a małym obszarem ilustrującym wymiar mądrości, ulokowanym na samym wierzchołku trójkąta. Wzajemne relacje między poszczególnymi elementami hierarchii są różnie interpretowane i dlatego warto przypomnieć niektóre z nich (Grabowski, Zajac, 2009). Często cytowana jest koncepcja semiotyczna, na której oparł swoje po-



Rys. 1. Piramida wiedzy (DIKW). (Wikimedia Commons, CC-BY-SA)

głądy amerykański filozof Charles Morris, przyporządkowując dane do struktur, czyli do syntaksy, wiedzę do rozumienia i semantyki, a informację do pragmatyki, czyli użyteczności działań. Z tego punktu widzenia dane są przetwarzane, wiedza reprezentowana, a informacja opracowywana (Morris, 1966, pp. 6-20). Historyczne uwarunkowania i współzależności elementów hierarchii wyjaśnił David Weinberger w nawiązaniu do badań teoretyka organizacji Russela Ackoffa (Weinberger, 2010; Ackhoff, 1989). Zróżnicowanie między danymi i informacją miało dla niego charakter nie strukturalny, lecz funkcjonalny. Dane bez kontekstu nie mogą służyć jako informacje, a z kolei informacja jest wyprowadzana i powstaje z danych. Dopiero w opracowanych informacjach dane są poprawnie sformułowane i ukształtowane. Wyraźnie widać, że proponowany model zakłada brak możliwości dokładnego oddzielenia danych od informacji i odwrotnie, oraz wskazuje na to, że jednostronna koncentracja tylko na danych jest równie krótkowzroczna, jak rozważania o informacji bez jej zestawienia z materiają danych. Podobna sytuacja występuje w przypadku kontyngencji – współzależności – informacji i wiedzy.

Przywołany powyżej Martin Frické, amerykański logik i bibliotekoznawca, dokonał ostatnio krytyki „hierarchii wiedzy” stanowiącej część kanonu nauki o informacji i zarządzaniu. Przedstawił argumenty, według których hierarchia ta oparta jest na błędnych przesłankach i przestarzałych stanowiskach filozoficznych operacjonalizmu (wedle którego pojęcia empiryczne definiuje się przez podanie operacji określających ich zastosowanie) i induktywizmu (przekonania, że z koniunkcji zdań szczegółowych logicznie wynika zdanie ogólne). Frické wyjaśnił naturę elementów hierarchii DIKW w następujący sposób: dane są czymś dającym się zapi-

sać/ zarejestrować semantycznie i pragmatycznie; informacja jest tym, co w literaturze znane jest jako „wiedza słaba”, czyli oparta na niepewnych i/lub niepełnych danych; wiedza nie ogranicza się więc tylko do wiedzy „ściślej”, „naukowej”, należy bowiem do niej także „słaba wiedza”, a mądrość oznacza posiadanie i używanie, jeśli to konieczne, szerokiej praktycznej wiedzy przez podmiot, który docenia zawodną i omylną naturę wiedzy (Frické, 2009).

Powracając do pytania o rolę bibliotek naukowych i zawodu bibliotekarza wobec rosnącego znaczenia danych badawczych, należy stwierdzić, że są one obecne w działalności wielu bibliotek. W ostatnich latach zostały wypracowane procedury związane z gromadzeniem, opracowaniem i archiwizowaniem danych, które stanowią nowy rodzaj zbiorów i obejmują zróżnicowane materiały takie jak: dokumenty tekstowe, dane liczbowe, notatki, kwestionariusze, ankiety, wyniki badań ankietowych, nagrania audio i wideo, fotografie, próbki, protokoły laboratoryjne, artefakty, modele matematyczne, oprogramowanie, wyniki symulacji komputerowych, obiekty i archiwa mediów społecznościowych. Mamy tu do czynienia z różnorodnymi danymi w zależności od dyscyplin i dziedzin wiedzy oraz stosowanych metod badawczych – od archeologii, medycyny, geofizyki, socjologii, budownictwa, po muzykologię, itd. Możliwe sfery aktywności bibliotek pojawiły się w trzech grupach zadaniowych: doradztwa i planowania, opracowania i wyszukiwania oraz publikacji (Bertelmann, Pfeiffenberger, 2016). Pierwsze zadanie jest oczywiste, drugie domaga się już wyjaśnienia: oznacza tworzenie metadanych o wysokich standardach, wykazywanie opisów danych badawczych w katalogach i wyszukiwarkach, zapewnienie wymiany danych między repozytoriami/bazami danych badawczych i bibliotecznymi katalogami oraz wyszukiwanie ich w systemie zarządzania publikacjami w celu generowania raportów i wykazów. Trzecia grupa zadań związana jest z publikowaniem danych w repozytorium i zapewnieniem powiązania opisów (rekordów) prac naukowych z opisami należących do tych prac danych badawczych. Dzięki temu zostanie zapewniony jednoczesny dostęp i możliwość korzystania z pracy/publikacji naukowej oraz danych badawczych powstających w trakcie jej powstawania.

Biblioteki stały się także biorcą licencji na identyfikatory DOI (Digital Object Identifier), co zapewnia publikowanym danym badawczym jednoznaczną identyfikację. Jest to ważne w cytowaniu danych i powoływaniu się na nie w przypadku powtórnego ich wykorzystania. W opracowaniu danych muszą być stosowane standardowe metody ich prezentacji i zachowana wysoka jakość rekordów, po to, żeby mogły stanowić integralną część *Records of Science* na równi z publikacjami naukowymi. Kolejne czynności wiążą się z przechowywaniem i archiwizacją danych. Udział bibliotek w publikowaniu danych odbywa się najczęściej we współpracy

z innymi jednostkami naukowymi uczelni oraz centrum komputerowym. Ważny jest także udział samych naukowców i specjalistów z konkretnych dziedzin w ocenie, klasyfikacji i selekcji danych. Niektóre dane ze względu na znikomą wartość nie muszą być publikowane i archiwizowane. Omawiane praktyki wchodzą dopiero w życie – nie tylko u nas. Podobnie rzecz ma się w Niemczech – w tamtejszym dyskursie o bibliotekach (według informacji z 2016 r.) postępowanie z danymi badawczymi jest także tematem nowym, a praktyczne osiągnięcia bibliotek naukowych w tym zakresie znajdują się w początkowym stadium. Z tymi problemami zdążyły zmierzyć się przede wszystkim biblioteki uniwersytetów z wysoko zaawansowanymi badaniami w dziedzinie medycyny, nauk przyrodniczych, szczególnie nauk o Ziemi, oraz uczelnie techniczne i instytuty badawcze¹⁰.

DATA LIBRARIAN – KOMPETENCJE I KSZTAŁCENIE

W odpowiedzi na podjęte przez biblioteki zadania w odniesieniu do danych badawczych kształtowały się nowe umiejętności bibliotekarzy. Na tym gruncie zaczęła się rozwijać nowa specjalizacja w zawodzie określana jako bibliotekarz danych (ang. *data librarian*). Funkcjonuje również nazwa kustosz danych (ang. *data steward*), odpowiadający za „cykl życia danych”, koordynator prac w procesie tworzenia i zarządzania danymi na wydziałach i w zespołach badawczych (Wałek, 2019, s. 505-507). Wariantów nazwy nowej specjalizacji zawierających słowa kluczowe *data librarian* jest znacznie więcej. Można je prześledzić w publikacji amerykańskich autorów (referat wygłoszony na 84. Światowym Kongresie IFLA w Kuala Lumpur) pod tytułem, który zaczyna się od pytania „kim jest bibliotekarz danych?” (Khan, Du, 2018). Odpowiedzi na nie autorzy postanowili szukać na podstawie analizy treści ogłoszeń o pracę na stanowisku *data librarian* zamieszczonych online przez biblioteki akademickie w Stanach Zjednoczonych. Badacze sami przyznają, że wybrana przez nich metoda ma swoje słabe strony, ponieważ zajmuje się opisem „pustych” stanowisk za pomocą słów i fraz właściwych ofertom pracy, pozwala ona jednak zebrać wartościowy materiał prezentujący oczekiwania i wymagania pracodawców. Pracodawcami, autorami ogłoszeń są dyrektorzy bibliotek akademickich, jak również pracujący na danym uniwersytecie badacze zainteresowani zatrudnieniem w swoich zespołach dobrze przygotowanych specjalistów. W badanej grupie ogłoszeń zawód bibliotekarza wy-

¹⁰ Na wielu uczelniach działają specjalne zespoły zajmujące się danymi badawczymi. Takie rozwiązanie funkcjonuje w Instytucie Technologicznym w Karlsruhe, który jest typowym uniwersytetem badawczym. Więcej: Research Data Management (RDA) am KIT. <http://www.rdm.kit.edu/index.php>.

stępował z następującymi określeniami (bibliotekarz – liczba ogłoszeń): usług w zakresie danych (*data services librarian*) – 14; zarządzania danymi badawczymi (*research data management librarian*) – 11; danych badawczych (*research data librarian*) – 7; przechowywania danych (*data curation librarian*) – 7; wizualizacji danych (*data visualization librarian*) – 3; danych i informacji rządowych (*data and government information librarian*) – 1; kompetencji w zakresie danych (*data literacy librarian*) – 1.

Wyraźnie widać, że zdecydowanie przeważają dwie pierwsze nazwy zawierające zadania kompleksowych usług w zakresie danych oraz zarządzania danymi badawczymi. Ciekawie wypadły dalsze analizy zawartych w ogłoszeniach wymagań stawianych kandydatom ubiegającym się o zatrudnienie. Zestawiono je w dwóch grupach w oddzielnych tabelach jako umiejętności wymagane i preferowane wraz z ich procentowym udziałem w ogłoszeniach. Oto umiejętności określone w ogłoszeniach jako konieczne:

- asystent badacza – 100%
- myślenie krytyczne/zdolność rozwiązywania problemów – 89%
- wiedza na temat federalnych i zastrzeżonych źródeł danych w USA – 82%
- zastrzeżone (*proprietary*) lub otwarte źródła pakietów programów statystycznych (R, SPSS, Stata, SAS, Python, NVivo) – 70%
- aktywność i wiarygodność społeczna – 61%
- wizualizacja – 25%
- ilościowa i jakościowa analiza danych – 20%
- nauczanie – 14%
- wiedza o standardach i formatach metadanych – 8%
- kompetencje w zakresie informacji o danych – 6%.

Spośród umiejętności preferowanych i dobrze widzianych na stanowisku *data librarian* znalazły się:

- języki programowania – 60%
- współczesne technologie – 45%
- zarządzanie danymi i ich przechowywanie – 30%
- myślenie krytyczne/zdolność rozwiązywania problemów – 28%
- repozytoria instytucjonalne i metadane – 26%
- pakiety statystyczne (R, SPSS, Stata, SAS, Python) – 24%
- geoprzestrzenne dane i technologie – 20%
- zarządzanie projektami – 14%
- dane – kompletowanie, kodowanie, ponowne wykorzystanie – 7%
- pisanie grantów – 4%.

Podsumowując przegląd kompetencji oczekiwanych przez pracodawców, można wyciągnąć wniosek, że nowa specjalizacja bibliotekarza da-

nych zakłada wysokie kwalifikacje informatyczne, takie jak znajomość języków programowania, posługiwanie się programami statystycznymi, programami wizualizacji danych czy korzystanie z nowoczesnych technologii. Stwarza ona również nowe pole współpracy z naukowcami, którzy widzą bibliotekarza w roli asystenta udzielającego im fachowego wsparcia w formie szkoleń, instrukcji i praktycznego postępowania z danymi badawczymi. Cytowany artykuł kończą dwie sugestie. Pierwsza wskazuje na pilną potrzebę określenia dla nowego profilu zawodowego *data librarian* podstawowych ram kompetencji przez Stowarzyszenie Bibliotek Amerykańskich (*American Library Association* – ALA) oraz konieczność akredytacji kierunków studiów magisterskich w dziedzinie nauki o bibliotece i informacji, kształcących specjalistów w zakresie danych. Drugi postulat dotyczy uwzględnienia w programach studiów wymagań pracodawców oraz realnych obowiązków przewidzianych na stanowisku *data librarian* przez menadżerów prowadzących rekrutację i zatrudnianie.

Przegląd praktycznych kwalifikacji zawodu *data librarian* w akademickich bibliotekach amerykańskich, gdzie on się ukształtował i obecnie funkcjonuje, skłania do pytania o program kształcenia w tym zakresie przedmiotowym. Oferty studiów szkół wyższych mogą realnie oddziaływać na jego stabilizację. Dla przykładu Instytut Nauki o Bibliotece i Informacji Uniwersytetu Humboldtów w Berlinie we współpracy ze Szkołą Wyższą w Poczdamie przygotował program studiów magisterskich o nazwie *Digitales Daten Management*, których pierwszy semestr rozpoczął się w kwietniu 2020 r.¹¹. Więcej na temat programów kształcenia można dowiedzieć się na podstawie doświadczeń Wyższej Szkoły Technicznej w Kolonii – zaproponowane przez nią rozwiązania można potraktować jako studium przypadku. Uczelnia ta kształci cenionych absolwentów w zakresie nauki o bibliotece, informacji, komunikacji i mediach, z których obecnie prowadzi 11 kierunków magisterskich i pięć licencjackich (TH Köln, 2020). Tu właśnie w latach 2017/2018 podjęto prace nad formalną akredytacją nowej specjalizacji *Data Librarian* w zespole prof. dr Simone Fühles-Ubach, dyrektorki Instytutu Nauki o Informacji (Fühles-Ubach, 2019). Postępowano zgodnie z obowiązującymi zasadami tworzenia programu studiów (*Curriculumwerkstatt*), które na pierwszym miejscu stawiają troskę o kompetencje absolwentów i możliwości ich zatrudnienia na rynku pracy (ang. *employability*).

Przygotowania odbyły się w dwóch etapach, pierwszy miał charakter wewnętrzny, uczestniczyli w nim pracownicy wymienionego instytutu, którzy dokonali krytycznej oceny treści i struktury dotychczasowych kie-

¹¹ Digitales Daten Management. [online], [dostęp: 30.04.2020]. Dostępny w WWW: <https://www.ibi.hu-berlin.de/de/studium/studiengaenge/ddm-master>; <https://www.ddm-master.de/>.

runków, wskazali ich słabe i mocne strony (analiza SWOT), a następnie wypracowali założenia zmian. W tym celu zapoznali się z danymi rynku pracy, analizowali rozpisane konkursy na stanowiska, a także aktualne dokumenty instytucji gospodarczych oraz stowarzyszeń bibliotecznych (m.in. Niemieckiego Związku Bibliotek i IFLA), po to, by trafnie ocenić warunki i przyszłe oczekiwania rynku pracy. Przeprowadzony został także benchmarking pokrewnych kierunków studiów na innych uczelniach w Niemczech. Drugi etap polegał na działaniach skierowanych do środowiska zewnętrznego w celu uzyskania opinii o kierunkach aktualnie prowadzonych i tych planowanych. Był to cykl spotkań z około 20 przedstawicielami bibliotek publicznych, naukowych i specjalnych różnej wielkości z regionu Nadrenii Północnej-Westfalii, które reprezentowali ich dyrektorzy, kierownicy oddziałów technologii informacyjnych oraz osoby odpowiedzialne za kształcenie i szkolenia bibliotekarzy. Odrębną grupę odniesienia stanowili absolwenci, którzy już „zderzyli” się z rynkiem pracy i mogli podzielić się swymi doświadczeniami.

W efekcie rozmów podjęto istotne decyzje o zróżnicowaniu programów studiów dla pracowników bibliotek publicznych i naukowych oraz o rozwijaniu nowego profilu *Data Librarian* jako specjalisty w zakresie technologii informacyjnych w takich obszarach jak dane badawcze, repozytoria i e-nauka. Przygotowano programy dwóch kierunków i czterech specjalności studiów licencjackich trwających siedem semestrów:

1. Biblioteka i cyfrowa komunikacja – specjalności: Biblioteki publiczne (*Public Education and Community Communication*); Biblioteki naukowe (*Digital Scholarship and Scholarly Communication*).

2. Dane i nauka o informacji (*Data and Information Science*) – specjalności: *Data Librarian* (nauka, badania, edukacja); *Data Analyst* (gospodarka).

Tabela 1

Program studiów licencjackich (BA – Bachelor) Dane i Nauka o Informacji (*Data and Information Science*) – specjalność: Bibliotekarz danych (*Data Librarian*)

Semestr 1	Programowanie Rozwój sieci	Opracowanie informacji Reprezentacja wiedzy	Cyfrowe społeczeństwo informacyjne Etyka informacji Rekonesans zawodu	Wizualizacja informacji	Informacja w przedsiębiorstwie
Semestr 2	Programowanie Rozwój programów (software)	Opracowanie informacji Struktura opisu dokumentów	Modelowanie danych	Analiza statystyczna danych	Źródła informacji Wyszukiwanie informacji

Semestr 3	Systemy informacyjne Systemy zarządzania treścią i dokumentami	Wyszukiwanie informacji (Information retrieval)	Systemy baz danych	Data Mining	Analiza informacji
Semestr 4	Moduł praktyczny	Moduł praktyczny	Faza praktyczna	Faza praktyczna	Faza praktyczna
Semestr 5	Technologia wyszukiwarek Technologie sieciowe	Praca nad projektem I <i>Data Librarian</i> Projekt interdyscyplinarny	Prawo informacyjne Prawo ochrony danych Praca naukowa	ASC Activities and Social Credit (Aktywność i wiarygodność społeczna) Portfolio studenta	Konsulting informacyjny Komunikacja naukowa i polityka naukowa <i>Data Librarian</i>
Semestr 6	Praca nad projektem II		Dane badawcze I. <i>Data Librarian</i> Zarządzanie licencjami, publikowanie elektroniczne i Open Access	Infometria, bibliometria, naukometria Metody badań empirycznych	Katalogowanie automatyczne
Semestr 7			Dane badawcze II. <i>Data Librarian</i> Digitalizacja i długoterminowa archiwizacja	Praca licencjacka Seminarium licencjackie	Praca licencjacka

Źródło: Fühles-Ubach, Simone; Seidler-de Alwis, Ragna. *Der Data Librarian als Antwort auf Veränderungen in wissenschaftlichen Bibliotheken*, s. 8. (tłum. Krystyna Hudzik). [online], [dostęp: 28.04.2020]. Dostępny w WWW: <https://www.zbw.eu/fileadmin/pdf/veranstaltungen/2019/2019-digitalitaet-2-3-data-librarian.pdf>.

W programie połączono zajęcia z nauki o danych i nauki o informacji oraz założono podstawową znajomość programowania na podstawie Pythona, administrowania programami i bazami, a także znajomość systemów zarządzania dokumentami i treścią. Widoczna jest wyraźna przewaga przedmiotów informatycznych i praktycznych. Aby zapewnić wysoką jakość kształcenia, powołano do Kolegium Instytutu osoby ze specjalizacją w następujących zakresach: wyszukiwanie informacji (informatyk); wyszukiwanie w sieci (fizyk); medioznawstwo i sieć; kompetencje w zakre-

sie danych i kompetencje informacyjne (bioinformatyk). Utworzono dwa nowe stanowiska profesorskie: E- nauka i zarządzanie danymi badawczymi oraz Open Access i zarządzanie e- źródłami. Zlikwidowano natomiast dwie dawne profesury – IT w bibliotekach oraz Usługi biblioteczne i komunikacja naukowa (Fühles-Ubach, Seidler-de Alwin, 2019).

Pierwszy semestr na kierunku Data and Information Science zaczął się w semestrze zimowym 2018/2019. Naukę rozpoczęło 64 studentów, z wyraźną przewagą mężczyzn (51). Studia mają profil praktyczny, podczas zajęć rozwijane są umiejętności ukierunkowane na określony typ biblioteki (medycznej, ekonomicznej, technicznej). Moduł praktyczny *Data Librarian* obejmuje prawie sześć miesięcy, co oznacza szkolenie przez cały semestr w bibliotece lub instytucie badawczym, którymi są m.in. tak znakomite jednostki jak: Centralna Biblioteka Medyczna w Kolonii i Bonn (ZB MED in Köln und Bonn), Centrum Badawcze w Jülich (Forschungszentrum Jülich) czy *GESIS* Instytut Leibniza Nauk Społecznych (*GESIS* Leibniz Institut für Sozialwissenschaften). Dzięki temu, że uczelnia uczestniczy w projektach w kooperacji z bibliotekami, biorą w nich udział także studenci. Przewidywane stanowiska pracy absolwentów kierunku *Data and Information Science* stanowią: *data librarian*, menadżer informacji i wiedzy, bibliotekarz systemowy w zakresie technologii informacyjnych (IT), analityk badań i rynku, analityk Big Data i badacz danych (*data scientist*).

PODSUMOWANIE

Jeśli zgodzić się z Henrykiem Hollendrem, że zadanie bibliotekarza w przyszłości będzie polegać na „rekonstrukcji obiegu informacji, wyszukiwaniu ‘brakujących ogniwi’, kojarzeniu odległych fraz, tekstów, zjawisk, ludzi” (Hollender, 2018), to należałoby także przyjąć, iż to właśnie te kryteria doskonale spełnia *data librarian*. Jego pojawienie się w bibliotekach świadczy o trwającej ewolucji zawodu bibliotekarza pod wpływem cyfrowych technik przekazu i komunikacji. W jej wyniku coraz bardziej musi on odchodzić od swych tradycyjnych umiejętności i szybko przyswajać sobie nowe, po to, żeby nadal mógł pośredniczyć w dostępie do wiedzy, informacji i kultury. Obecnie nowym wyzwaniem dla niego są dane i procesy fachowego postępowania z nimi. Wnioski nasuwające się z powyższych rozważań można zestawić w następujących punktach.

Stanowisko *data librarian* powstało w związku z rosnącym znaczeniem danych badawczych w procesie naukowym oraz zasadą otwartego do nich dostępu. Jego stabilizacja zależy od formalnego określenia kwalifikacji nowej specjalności zawodowej, a dalszy rozwój zależy od profilu – programów – studiów akademickich w tym zakresie.

Data librarian jest specjalistą o wysokich kompetencjach w zakresie programów i sieci informatycznych oraz systemów informacyjnych. Zakres

jego umiejętności w zarządzaniu danymi badawczymi odpowiada zadaniom realizowanym przez *data scientists*.

Włączenie do zadań bibliotek naukowych gromadzenia i archiwizowania danych, będących efektem wstępnej fazy procesu badawczego, zwiększy zróżnicowanie między kompetencjami pracowników bibliotek naukowych i publicznych. Te pierwsze stanowią element infrastruktury i uczestniczą w komunikacji naukowej, natomiast te drugie – biblioteki publiczne – budują komunikację społeczności lokalnych i swoje zadania podporządkowują temu celowi.

Dziś trudno jest jednoznacznie wskazać podstawowe kompetencje i odpowiedzieć na pytanie, co powinni bibliotekarze wiedzieć i umieć (a czego nie?), dlatego ogólne kształcenie w zakresie nauki o informacji i bibliotekoznawstwa zderza się z konkurencją, jaką stanowią dla bibliotekarza inne zawody zorientowane na przyszłość i oferujące specjalistyczne umiejętności w zakresie przetwarzania i przechowywania danych.

Programy nowych kierunków studiów nie powinny wyrażać punktu widzenia tylko prowadzących zajęcia i podlegać ocenie wyłącznie pracowników naukowych, z których składają się u nas komisje akredytacyjne według zasady: jedni profesorowie akredytują drugich. Równie ważne są opinie interesariuszy zewnętrznych – stowarzyszeń zawodowych bibliotekarzy i związków instytucjonalnych bibliotek. Ich obecność w tych komisjach powinna stać się obowiązkowa, tak jak to jest w przypadku akredytacji studiów w zakresie nauki o bibliotece i informacji prowadzonej przez Amerykańskie Stowarzyszenie Bibliotek. Inny sposób współpracy ze środowiskiem zewnętrznym stanowią spotkania z praktykami zawodu, dyrektorami bibliotek oraz absolwentami, których głosy wskażą kompetencje wymagane i oczekiwane przez pracodawców (doświadczenia TH w Kolonii).

BIBLIOGRAFIA

- Ackoff, Russell L. (1989). From Data to Wisdom. *Journal of Applied Systems Analysis* vol. 16, iss. 1, pp. 3-9.
- Bertelmann, Roland u. Hans, Pfeiffenberger: Forschungsdaten und Bibliotheken. In: *Praxishandbuch Bibliotheksmanagement*, Bd. 2 / Rolf Griebel, Hildegard Schäffler, Konstanze Söllner (Hrsg.). – Berlin [u.a.]: De Gruyter; 2016. S. 642-649.
- Data Literacy und Data Science Education: Digitale Kompetenzen in der Hochschul-ausbildung. Policy Paper der Präsidiums-Task-Force „Data Science“ der Gesellschaft für Informatik e.V. in Zusammenarbeit mit Vertretern der Deutschen Mathematiker-Vereinigung e.V., der Deutschen Physikalischen Gesellschaft e.V. und der Gesellschaft Deutscher Chemiker e.V. Berlin: Gesellschaft für Informatik e.V., 2018 [online]. [dostęp: 19.04.2020.]. Dostępny w WWW: https://gi.de/fileadmin/GI/Hauptseite/Aktuelles/Aktionen/Data_Literacy/GI_DataScience_2018-04-20_FINAL.pdf.*
- Davenport, Thomas H., Patil D. J. (2012). Data Scientist: The Sexiest Job of the 21st

- Century. *Harvard Business Review* 2012 October [online], [dostęp: 24.04.2020]. Dostępny w WWW: <https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century>.
- Frické, Martin (2018). *Knowledge pyramid. The DIKW hierarchy. Encyclopedia of Knowledge Organization. International Society for Knowledge Organization* [online], [dostęp: 20.04.2020]. Dostępny w WWW: <https://www.isko.org/cyclo/dikw#refF>.
- Frické, Martin (2009). The Knowledge Pyramid: a Critique of the DIKW Hierarchy. *Journal of Information Science* 35, no. 2: 131-42. DOI: 10.1177/0165551508094050.
- Fühles-Ubach, Simone; Seidler-de Alwis, Ragna. (2019), *Der Data Librarian als Antwort auf Veränderungen in wissenschaftlichen Bibliotheken*. Innovations workshop 09.04.2019 Hamburg [online], [dostęp: 28.04.2020]. Dostępny w WWW: <https://www.zbw.eu/fileadmin/pdf/veranstaltungen/2019/2019-digitalitaet-2-3-data-librarian.pdf>.
- Fühles-Ubach (2019) Simone; Schaer, Philipp; Lepsky, Klaus; Seidler-de Alwis, Ragna: Data Librarian: ein neuer Studienschwerpunkt für wissenschaftliche Bibliotheken und Forschungseinrichtungen. *Bibliothek: Forschung und Praxis* (Preprint 29.4.2019) [online], [dostęp: 28.04.2020]. Dostępny w WWW: edoc.hu-berlin.de/handle/18452/20699 doi.org/10.18452/19901.
- Grabowski Mariusz, Zając Agnieszka (2009). *Dane, informacja, wiedza – próba definicji*. Kraków: Wydaw. Uniwersytetu Ekonomicznego. [online], [Dostęp: 21.04.2020.]. Dostępny w WWW: https://www.uci.agh.edu.pl/uczelnia/tad/APSI/cwiczenia/Dane_informacje_wiedza.pdf.
- Grygowski, Dariusz (2020). *Studia bibliotekoznawcze w krzysie. W: Interdyscyplinarium nauk o mediach i kulturze*. Pod red. M. Antczak i Z. Gruszki. Łódź: Wydaw. Uniwersytetu Łódzkiego, s. 145-176.
- Harari, Yuval Noah (2018). *Homo deus: krótka historia jutra*. Kraków: Wydaw. Literackie.
- Hollender, Henryk (2018), *Nauka jako mrowisko: miejsce biblioteki w sieciach współpracy badawczej* [online], [dostęp: 22.04.2020]. Dostępny w WWW: <http://repozytorium.wsb-nlu.edu.pl/handle/11199/10406?show=full>.
- Khan, Hammad Rauf; Du, Yunfei (2018) *What is a Data Librarian?: A Content Analysis of Job Advertisements for Data Librarians in the United States Academic Libraries* [online], [dostęp: 23.04.2020.]. Dostępny w WWW: <http://library.ifla.org/2255/1/139-khan-en.pdf>.
- Krzysztofek, Kazimierz (2012). BIG DATA SOCIETY. Technologie samoopis i samopokazu: ku humanistyce cyfrowej. *Historia i Kultura* 21. [online], [dostęp: 16.04.2020]. Dostępny w WWW: <https://www.kulturalihistoria.umcs.lublin.pl/archives/3626>.
- Morris, Charles, W. (1966). *Foundations of the theory of signs*. Chicago [u.a.]: University of Chicago Press.
- TH Köln (2020) Technische Hochschule Köln Technology Arts Sciences, [online], [dostęp: 28.04.2020]. Dostępny w WWW: https://www.th-koeln.de/studium/information-und-kommunikation_111.php.
- Wałek, Anna (2019). Data Librarian and Data Steward – New Tasks and Responsibilities of Academic Libraries in the Context of Open Research Data Implementation in Poland. *Przegląd Biblioteczny* 87 (2019), z. 4, s. 497-512.

- Weinberger, David (2010). The Problem with the Data-Information-Knowledge -Wisdom Hierarchy. *Harvard Business Review*, February 02. [online], [dostęp: 20.04.2020]. Dostępny w WWW: <https://hbr.org/2010/02/data-is-to-info-as-info-is-not>.
- Wilk, Paulina (2017). Jutro będę algorytmem. *Przekrój* z dnia 20.03. 2017. [online], [dostęp: 16.04.2020]. Dostępny w WWW: <https://przekroj.pl/artykuly/recenzje/jutro-bede-algorytmem>.
- Wolff, Christian (2018), *Informations- und Datenkompetenz: Gemeinsamkeiten, Unterschiede, Entwicklungstendenzen*. In: *Bibliotheksentwicklung im Netzwerk von Menschen, Informationstechnologie und Nachhaltigkeit: Festschrift für Achim Oßwald* / [Hrsg.] Simone Fühles-Ubach, Ursula Georgy. Bad Honnef: Bock + Herchen Verlag, 2019, S. 221- 235. [online], [dostęp: 27.04.2020]. Dostępny w WWW: https://www.th-koeln.de/mam/downloads/deutsch/studium/studiengaenge/f03/bib_inf_ma/festschrift_osswald.pdf.

Artykuł w wersji poprawionej wpłynął do Redakcji 22 czerwca 2020 r.

KRYSTYNA HUDZIK

Main Library

Maria Curie-Skłodowska University in Lublin

e-mail: k.hudzik@umcs.pl

ORCID 0000-0002-1449-828X

DATA LIBRARIAN – NEW PROFESSIONAL PROFILE IN RESEARCH LIBRARIES

KEYWORDS: Data librarian. Data librarian – competencies. Data librarian – education. Research data. Big Data. Data competencies. Data researcher.

ABSTRACT: **Thesis/Objective** – In foreign research libraries data librarians are professionals specialized in research data. The author presents data librarian competencies required in US academic libraries and the curriculum for the education of data librarians at selected German university. The development of a new specialization within the librarian profession is discussed in the broader context of digital culture, the process of reality datafication, the production of Big Data and the competencies related to data. **Research methods** – The research method used for this article was the analysis of publications on data librarians, research data and tasks of research libraries. The author enhanced it with information from the websites of selected universities, repositories and libraries as well as some opinions from the discussion on Big Data, datafication and related transformations. **Results/Conclusions** – Data librarians are highly qualified professionals in the field of computer software, networks and information systems. Their competencies as regards research data management largely reflect tasks of data scientists. Further development and stabilization of this profession would depend on a formal development of a framework for their competencies and tasks and related curricula. New university curricula should be based on the evaluation made by an external community of library professionals, including the professional librarian associations and institutional associations of libraries.