

Anna Kugler
anna.kugler@bsbmuenchen.de
Tobias Beinert
tobias.beinert@bsbmuenchen.de
Astrid Schoger
astrid.schoger@bsbmuenchen.de
Munich Digitization Center/Digital Library
Bavarian State Library

Archiwizacja internetu jako usługa naukowa

Streszczenie: Gromadzenie i archiwizowanie stron internetowych istotnych dla nauki to jak dotąd bardzo zaniedbana sfera aktywności bibliotek niemieckich. Aby zapobiec groźnym stratom oraz zapewnić pracownikom naukowym stały dostęp do stron internetowych ponad dwa lata temu Bavarian State Library (BSB) stworzyła system archiwizacji stron internetowych. Głównym celem projektu zaakceptowanym przez German Research Foundation (DFG) był rozwój i realizacja kooperacyjnego modelu usługowego. Usługa ta ma wspierać inne instytucje dziedzictwa kulturowego w ich aktywności archiwizacyjnej i ułatwiać budowanie rozproszonego niemieckiego systemu archiwizacji naukowych stron internetowych. Dzięki temu projektowi biblioteka bawarska chce poprawić zarówno ilość, jak i jakość zarchiwizowanych treści oraz promować ich wykorzystanie w obszarze nauki.

Słowa kluczowe: archiwizacja internetu; Bavarian State Library; projekty biblioteczne; informacja naukowa; zabezpieczenie zasobów cyfrowych; pobieranie i przetwarzanie danych; nauki humanistyczne

1. Wstęp

Komunikacja cyfrowa będzie determinowała przyszłość humanistyki. Takie stwierdzenie padło podczas konferencji „Recenzowanie – komentowanie – blogowanie : Jak będą się komunikowali humaniści w cyfrowej przyszłości?“, która została zorganizowana przez [Bavarian State Library](#) na początku roku 2013. Konferencja odbyła się z okazji 2. rocznicy istnienia platformy European History, zwanej [recensio.net](#), a jej uczestnicy zastanawiali się, jak dalece sieć internetowa jako platforma publikowania treści naukowych może sprostać wymaganiom jakościowym szkolnictwa wyższego [1]. Prezentacje i dyskusje przy okrągłym stole jasno pokazały, że należy rozwijać nowe instrumenty, usługi i infrastrukturę w celu publikowania, oceny i wreszcie zachowania nowych rodzajów źródeł naukowych. Jednym z nowych typów zasobów cyfrowych, które były *niemal całkowicie pomijane*, jak to ujął Nils Brügger, *nawet jeśli można je uznać za jeden z najważniejszych współczesnych zasobów dziedzictwa kulturowego ludzkości* [2], są strony internetowe.

Archiwa internetowe, które zabezpieczają przechwycone strony internetowe i umożliwiają stały dostęp do nich, dla wielu naukowców są jeszcze nieznanne lub nietypowe jako narzędzia badawcze. W porównaniu do bieżącej sieci istnieje kilka charakterystycznych cech archiwów internetowych, które stanowią o ich potrzebie:

1. Archiwa webowe zawierają treści, które znikają z bieżącej sieci. Szacunki dotyczące średniej długości życia stron różnią się bardzo, ale to, co pokazują,

świadczy o tym, że [...] *web jest postrzegany jako magazyn pamięci naszej cywilizacji, który nie zabezpiecza sam siebie na przyszłość – stare strony zwykle są nie do odnalezienia* [3].

2. Obecnie powszechne jest cytowanie stron internetowych. Naukowcy coraz częściej odnoszą się do tych źródeł, ale znikające lub zmieniające się adresy URL często utrudniają bądź wręcz uniemożliwiają dostęp do cytowanych źródeł.
3. Historia sieci ilustruje także ważną część naszej kultury. Okresowe przechwytywanie stron internetowych pokazują nie tylko ewolucję i zmiany technologii oraz projektowania stron internetowych, ale i zmiany dyskursów politycznych czy naukowych.
4. W bardziej technicznym kontekście może zdarzyć się w przyszłości, że niektórych dokumentów nie będzie można oddzielić od narzędzi i platform, przy pomocy których je wygenerowano. Z tego powodu archiwizacja stron internetowych ma zupełnie odmienne zastosowanie niż archiwizacja książek drukowanych i jest w tym samym czasie dużo bardziej wymagająca [4].
5. Internetowe archiwa oferują tematyczne kolekcje danych, które mogą być analizowane za pomocą nowych metod masowej eksploracji danych. W kontekście wspierania e-humanistów mogą być oferowane bardzo zaawansowane możliwości dostępu do zasobów.

Pomimo tego raporty z działalności archiwów internetowych działających już od dłuższego czasu, takich jak np. UK Web Archive, pokazują, że nadal jest *mało dowodów na ich wykorzystanie w nauce* [5]. Ponadto wielu uczonych nie wie o istnieniu archiwów internetowych, jednak aktywni użytkownicy doceniają przyrost treści naukowych i dostrzegają potrzebę ulepszenia narzędzi do eksploracji danych [6]. A zatem celem całej społeczności archiwizującej internet, którą tworzą przede wszystkim biblioteki narodowe i większe biblioteki regionalne, musi być poprawa jakości i liczby naukowych archiwów internetowych oraz promowanie ich wykorzystywania w obszarze nauki [7].

2. Czy rozproszona niemiecka biblioteka naukowa ostatecznie stanie się rozproszonym niemieckim naukowym archiwum webowym?

W Niemczech stan aktywności w zakresie internetowej archiwizacji wygląda trochę niby dywan utkany z małych kawałków inicjatyw zajmujących się gromadzeniem stron internetowych. Niektóre biblioteki podejmują się archiwizacji stron ze względu na regionalny egzemplarz obowiązkowy (EO) (np. Badenia-Württembergisches Online-Archiv, <http://www.boa-bw.de/> lub *edoweb*, cyfrowe archiwum Rheinland-Pfalz, <http://www.lbz-rlp.de/cms/landeskunde/edoweb/>), lecz pozostałe biblioteki regionalne w ogóle nie archiwizują stron internetowych. Niemiecka Biblioteka Narodowa, która gromadzi EO dla Niemiec, w roku 2012 rozpoczęła gromadzenie najważniejszych dla niemieckiej społeczności stron internetowych z zakresu historii, polityki, ekonomii, kultury i zamierza udostępnić je w swoich czytelniach [8]. Podsumowując, jak do tej pory niemieckie biblioteki nie zaspokajają potrzeb społeczności naukowej w zakresie archiwizowania stron

internetowych. Częściowo wina leży po stronie federalnie ukształtowanej struktury bibliotek niemieckich.

W Niemczech z powodu sytuacji politycznej i historycznej nigdy nie została uchwalona centralna biblioteka naukowa. W roku 1949 uruchomiono unikalny w skali międzynarodowej specjalny system współpracy, odpowiedzialny za dostarczanie ponadregionalnej literatury naukowej i edukacyjnej. Za organizację systemu odpowiada niemiecka fundacja German Research Foundation (Deutsche Forschungsgemeinschaft – DFG). Dwadzieścia trzy państwowe i akademickie biblioteki oraz niektóre biblioteki specjalistyczne biorą udział w systemie współpracy, w którym każda biblioteka jest odpowiedzialna za jedną lub większą liczbę dziedzin nauki (Sondersammelgebiete, SSGS). Wszystkie biblioteki uczestniczące zobowiązują się do tego, że ich kolekcje będą udostępniane dla całego kraju. Punkty dostępu do zbiorów umożliwiają tzw. biblioteki wirtualne, które oferują zróżnicowane możliwości wyszukiwawcze w ramach danego interfejsu użytkownika (*one-stop-shop*) [9].

Kolekcje tych wirtualnych bibliotek dziedzinowych obejmują swobodnie dostępne treści o naukowym znaczeniu, które z zasady znajdują się na stronach web. Każda strona jest selekcyonowana merytorycznie przez ekspertów dziedzinowych. Dużo wysiłku pochłania proces katalogowania. Kilka bibliotek postanowiło współpracować w tym zakresie; stworzyły one system Academic Link Share (ALS) (<http://www.academic-linkshare.de>). Baza danych zawiera już ponad 110 000¹ wpisów. Nie tylko sama treść, lecz także praca, która jest czasochłonna, może być utracona, gdy znika odpowiednia strona internetowa, a archiwizacja ciągle jeszcze nie jest obowiązkowa dla bibliotek dziedzinowych.

W rezultacie dokładnej oceny tego systemu dziedzinowego, zleconego przez DFG w 2010 r. w celu sprawdzenia jego przydatności dla nauki, opracowano kilka rekomendacji: ukierunkowanie w przyszłości na gromadzenie zasobów elektronicznych, silniejsze skupienie się na tworzeniu kolekcji zgodnie z potrzebami naukowymi oraz ogólne nastawienie na naukowość. Dla zapewnienia trwałości, długoterminowego zabezpieczenia i stałego dostępu do danych cyfrowych należy opracowywać nowe modele usług [10].

W tym kontekście BSB aplikowała do fundacji DFG z projektem, który miał skupić się na rozwijaniu kooperatywnego modelu usługowego archiwizującego strony webowe. Będzie on budowany na infrastrukturze BSB i czerpał z doświadczeń, które ta zyskała w tworzeniu naukowego archiwum stron internetowych i w ten sposób wesprze inne instytucje zajmujące się zachowaniem dziedzictwa kulturowego w ich działalności archiwizacyjnej. Ważnym celem jest bliższa współpraca ze środowiskiem naukowym. Wyniki projektu mogą stać się załącznikiem dla rozproszonego systemu German Research Web Archive. Projekt rozpoczęto w roku 2013.

¹ Od tłumacza i redakcji: dane poprawiono w dniu 10 marca 2017 r. po sprawdzeniu na wskazanej w tekście stronie internetowej ALS.

3. Archiwizowanie stron internetowych w Bavarian State Library

W 2010 r. biblioteczne centrum digitalizacji Munich Digitization Center/Digital Library (MDZ) (<http://www.digitale-sammlungen.de>) rozpoczęło fazę pilotażową kolekcjonowania i archiwizowania stron internetowych, a od początku 2012 r. archiwum internetowe efektywnie się powiększało (<http://www.babsmuenchen.de>). Koncentrowano się na stronach wcześniej wybranych i skatalogowanych przez biblioteki dziedzinowe (Virtuelle Fachbibliotheken, ViFas). BSB była odpowiedzialna za pięć tematów odnoszących się do następujących obszarów:

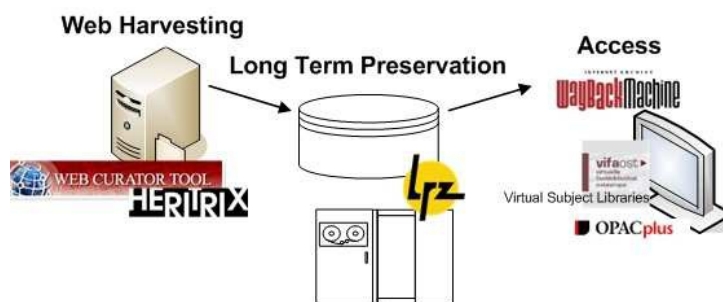
- historia (www.propylaeum.de, www.historicum.net),
- muzykologia (www.vifamusik.de),
- Europa Wschodnia (www.vifaost.de),
- romanistyka (www.vifarom.de),
- bibliotekoznawstwo i informacja naukowa (www.b2i.de).

Wirtualne biblioteki dziedzinowe zawierają już około 110 000 opisów witryn w bazie danych ALS. Ponieważ obowiązujące przepisy prawne Bawarii dotyczące egzemplarza obowiązkowego pozwalają zbierać i archiwizować tylko strony internetowe władz bawarskich, a nie witryny naukowe w kontekście międzynarodowym, do pobierania, archiwizowania i udostępniania tych stron potrzebna jest osobna zgoda. Z tego powodu do każdego właściciela strony internetowej wysyła się mailem bardzo szczegółowe zapytanie o pozwolenie, w którym opisane są prawa do pobierania i sekwencyjnego archiwizowania strony internetowej oraz udostępniania ich w internecie, jeśli nie koliduje to z prawami osób trzecich. Ponadto we wniosku podkreśla się, że ma tu zastosowanie niemieckie prawo autorskie. Pozytywne odpowiedzi na listy stanowią około 20% do 30%.

Do pobierania stron BSB wykorzystuje narzędzie Web Curator Tool (WCT). Jest to oprogramowanie open source rozwijane wspólnie przez British Library i Bibliotekę Narodową Nowej Zelandii. Wybrano je, ponieważ umożliwia ono zintegrowanie procesów selektywnego zbierania stron internetowych, w tym administrowanie prośbami o zezwolenie, zbieranie witryn zgodnie z harmonogramem prac i częściową automatyzację kontroli jakości. Robotem pobierającym jest Heritrix, opracowany przez Internet Archive. Aby zapewnić dostęp do pobranych stron, konieczna była lokalna adaptacja Wayback Machine. Zebrane pliki WARC (Web Archive) są archiwizowane w systemie Rosetta, komercyjnym oprogramowaniu Ex Librisa przeznaczonym do długoterminowego cyfrowego zabezpieczenia plików. Gwarancje bitstream (sekwencji bajtów) zapewnia Leibniz Supercomputing Centre (LRZ), z którym BSB ściśle współpracuje od wielu lat.

Wszystkie zbierane i archiwizowane strony są swobodnie dostępne. Zasoby najpierw udostępniane były przez indeks wirtualnych bibliotek dziedzinowych, a potem za pośrednictwem katalogu lokalnej biblioteki. Pierwotny punkt dostępu do archiwalnych stron internetowych bazuje na indeksowanych metadanych wirtualnych bibliotek dziedzinowych. Każde źródło internetowe ma szczegółowy opis ALS zawierający: tytuł, adres URL, instytucję sprawczą, kontakt, słowa kluczowe, a nawet krótkie streszczenie. Obok linku działającego w danym momencie został dodany tzw. Archive URL, który jest powiązany

stabilnie z systemem archiwizacji Rosetta, gdzie są przechowywane stare strony internetowe. Jeśli użytkownik kliknie na link Archive URL, otworzy się lokalna adaptacja Wayback Machine i pokaże wszystkie zarchiwizowane przechwycone stare strony dla tej wskazanej. Dla celów cytowania przydziela się trwałe identyfikatory (Uniform Resource Names, urny) na poziomie strony lub nawet w trakcie przechwytywania. W systemie katalogowym BSB każda strona jest opisana na poziomie minimum. W chwili obecnej w systemie katalogowym BSB można znaleźć około 500 zarchiwizowanych stron internetowych, pochodzących z kilku procesów przechwytywania (dane z czerwca 2013 r.).



Il. 1. Procesy pracy BSB's Web Archiving.

4. Budowanie usługi na rzecz innych instytucji kulturalnych

Nowo zatwierdzony projekt DFG koncentruje się na tworzeniu archiwum stron internetowych dla innych (niemieckich) instytucji dziedzictwa kulturowego, podobnych do Web Archiving Service (WAS), który został stworzony przez Californian Digital Library (<http://webarchives.cdlib.org>)² lub Internet Archive (<http://www.archive-it.org>). Tworzenie tej usługi bazuje na kilku usprawnieniach i rozszerzeniach już istniejącej infrastruktury.

Pierwszy pakiet roboczy projektu miał na celu rozwój profilu gromadzenia i archiwizacji stron internetowych, który stanowiłby bardzo ważny zbiór wytycznych dla instytucji korzystających z nowej usługi archiwizacji internetowej. Profil kolekcji opiera się w dużej mierze na kryteriach tematycznych zbiorów naukowych zdefiniowanych przez DFG, ale dodatkowo zawiera kryteria, które wynikają z konkretnej polityki gromadzenia BSB jako biblioteki archiwizacyjnej Bawarii. Bardzo trudne jest określenie profilu archiwizacji, który ma obejmować kryteria treści, a także aspekty techniczne i czasową spójność. Z powodów technicznych dostępne roboty (harvesters) nie są w stanie indeksować luster stron internetowych, identycznych wobec pierwotnej wersji, zwłaszcza przy dynamicznych funkcjonalnościach, jakimi są np. zapytania (kwerendy), które już nie działają w archiwum internetowej bazy danych. Może tego typu adresy URL powinny być wykluczone z procesu archiwizacji, choć ich zawartość jest wysoko ceniona ze względu na profil kolekcji. Ponadto oczywiście istotne są czas trwania i częstotliwość indeksowania, gdyż odnoszą się do kompletności i spójności czasowej pobieranych stron [11]. Zatem zarchiwizowana strona internetowa może być zawsze tylko „rekonstrukcją” [2] tej żyjącej, która także

² Przepis tłumacza i redakcji: zmienił się tytuł i adres tej strony na Archive-it!: <https://archive-it.org/>.

powinna być odzwierciedlona w profilu archiwizacyjnym. Szczegółowy profil kolekcji i archiwizacji pomogą dobrze zaprojektować i zdefiniować zakres indeksowania oraz tym samym poprawić jakość archiwum internetowego.

Usprawnienie procesu zapewnienia jakości i zdefiniowanie kryteriów jakościowych jest kolejnym zadaniem w ramach nowego projektu, który wciąż stanowi wyzwanie dla wielu instytucji w kontekście archiwizacji internetowej [12]. W bibliotece BSB jakość nowych wychwyconych stron jest ciągle oceniana z autopsji przez porównanie strony istniejącej z wersją archiwalną oraz przez analizy rejestru zdarzeń (log files) robota. Jednak rosnąca liczba zrzutów stron internetowych sprawia, że rozwój bardziej zautomatyzowanych procesów pracy będzie nieunikniony. Tak więc definiując kryteria jakości oraz testując nowe narzędzia i metody już stosowane przez inne instytucje czy zalecane przez International Internet Preservation Consortium (IIPC, <http://netpreserve.org/>), możemy opracować (częściowo) zautomatyzowany proces pracy. Ponadto nowa wersja 1.6 WCT oferuje częściowo zautomatyzowany proces pracy dla zapewnienia jakości na podstawie technicznie przeglądanych wskaźników jakościowych i będzie ona implementowana tak szybko, jak to możliwe.

Trzeba również zbadać nowe sposoby dostępu, gdyż samo szukanie według adresów URL czy tytułów witryn często nie spełnia potrzeb naukowców [13]. Z tego powodu należy analizować i testować istniejące już technologie eksploracji danych, które mogą być odpowiednie dla archiwów internetowych. Szczególny nacisk powinien być położony na indeksowanie i przeszukiwanie pełnych tekstów, a jako możliwe rozwiązanie należy wziąć pod uwagę SOLR³ [14]. Innym sposobem na poprawę dostępności do treści archiwum internetowego może być praca kompatybilna z Memento, dzięki czemu historyczne wersje stron internetowych z różnych archiwów mogą być zintegrowane z bieżącymi za pomocą prostego rozszerzenia przeglądarki [15].

Kolejny pakiet roboczy dotyczy dostępnych długoterminowych środków zabezpieczenia archiwów internetowych. Różne narzędzia i metody nadające się do użycia w cyfrowej długoterminowej ochronie stron internetowych będą testowane i prawdopodobnie wdrożone. Wśród nich znajduje się JHOVE2, który służy do charakteryzacji formatów plików wewnątrz plików WARC. Na podstawie tej charakterystyki będzie realizowanych kilka testów migracji formatów. Ponadto częścią tego pakietu roboczego są testy deduplikacji, wykonywane w celu zdobycia większej wiedzy o możliwościach procesów związanych z redukcją składowania z Tivoli Storage Manager (TSM), który jest wykorzystywany w LRZ lub już w czasie pobierania zasobów (harvesting) z najnowszą wersją Heritrixa. System zachowania zasobów Rosetta oferuje pewne funkcjonalności oceny ryzyka i rozwiązania migracji formatów, które będą testowane ze stronami internetowymi w trakcie realizacji projektu.

Na bazie tych doświadczeń i udoskonaleń fazy koncepcyjnej rozpocznie się budowanie wspólnego modelu usługi internetowej archiwizacji dla innych (niemieckich) instytucji

³ Wyjaśnienie od tłumacza i redakcji: Solr (czyt. solar) to serwer wyszukiwania pełnotekstowego bazujący na bibliotece Lucene. Rozwijany jest przez Apache Foundation i rozpowszechniany na licencji open source [Dostęp 6.03.2017]. Dostępny w: <https://wiki.apache.org/solr/>.

dziedzictwa kulturowego. Aby osiągnąć ten cel, BSB będzie współpracować z partnerami zewnętrznymi, poznawać ich wymagania dotyczące archiwizacji internetowej i projektować podstawową koncepcję serwisu z różnymi poziomami usług. Poziomy obsługi dotyczą przede wszystkim procesów selekcji i pobierania zasobów, które mogą być wykonane tak jak w BSB, w sposób scentralizowany lub zdecentralizowany przy użyciu pełnej nowej instalacji oprogramowania stworzonego dla instytucji partnerskiej. Po drugie należy omówić kwestię odpowiedzialności za archiwizację i zachowanie zasobu, co zostanie najprawdopodobniej przypisane centralnie BSB. Trzeci poziom usługi dotyczy dostępu, który w dużej mierze zależy od wymagań instytucji partnerskiej. Najważniejszym pytaniem jest, czy swobodny dostęp jest możliwy, czy też nie, i która z możliwości wyszukiwania jest preferowana.

Po zakończeniu prac koncepcyjnych powinna zostać zrealizowana implementacja technicznego modelu usługi. Obejmie to najprawdopodobniej dodatkowe instalacje WCT, utworzenie wymaganych interfejsów i ewentualnie własnych rozszerzeń technicznych. W ostatnim etapie, po fazie intensywnych testów, należy precyzyjnie obliczyć całkowite koszty procedury wspólnego archiwizowania internetu i zamieścić je w modelu biznesowym.

5. Wnioski

Opisany powyżej ambitny program działań ma na celu poprawę stanu prac archiwizacji stron internetowych w Niemczech. Dzięki rozbudowanej i wspólnej infrastrukturze istniejące możliwości doboru i katalogowania mogą zostać zintegrowane w bardziej zrównoważony proces, zawierający element archiwalny. Mamy nadzieję, że w ten sposób trwały dostęp do istotnych dla nauki stron internetowych może stać się normą, a nie wyjątkiem.

Bibliografia⁴:

1. Haslinger, P. „Block und Kugelschreiber oder Twitterwall? Thesen zur Zukunft der Wissenschaftskommunikation in den Geisteswissenschaften. Konferenzblog RKB 2013 [online]. [Dostęp 20.03.2017]. Dostępny w: <http://rkb.hypotheses.org/410>, <http://www.ahfmuenchen.de/Tagungsberichte/Berichte/pdf/2013/038-13.pdf>.
2. Brügger, N. & Finneemann, N.O., 2013. The Web and Digital Humanities: Theoretical and Methodological Concerns. In: Journal of Broadcasting & Electronic Media 57,1 (2013), p. 66-80, here p. 79.
3. Brügger, N. 2012. Web History and the Web as a Historical Source. In: Zeithistorische Forschungen / Studies in Contemporary History, Online Ausgabe [online]. 9 (2012), H. 2. [Dostęp 20.03.2017]. Dostępny w: <http://www.zeithistorische-forschungen.de/16126041-Bruegger-2-2012>.
4. van den Heuvel, Ch. 2010. Web Archiving in Research and Historical Global Collaboratories. In: Brügger, Niels (ed.). Web History. New York 2010, p. 279-303.
5. Hockx-Yu, H. 2013. Scholarly Use of Web Archives [online]. [Dostęp 20.03.2017]. Dostępny w: http://files.dnb.de/nesor/veranstaltungen/2013-02-27-scholarly-use-of-web-archives_public.pdf.

4 Wyjaśnienie od tłumacza i redakcji: bibliografię uzupełniono o tytuły stron internetowych, jeśli w oryginale nie występowały. Zweryfikowano adresy i podano daty dostępu do wszystkich cytowanych zasobów sieciowych.

6. Meyer, E., Thomas, A. & Schroeder, R. 2011. Web Archives: The Future(s) [online]. [Dostęp 20.03.2017]. Dostępny w: <http://netpreserve.org/resources/web-archives-futures>.
7. Meyer, E. 2010. Researcher Engagement with Web Archives: Challenges and Opportunities [online]. [Dostęp 20.03.2017]. Dostępny w: http://repository.jisc.ac.uk/543/1/JISC%20DREWA_ChallengesandOpportunities_August2010.pdf.
8. Cremer, M. 2013. Providing Access to the DNB Web Archive [online]. [Dostęp 20.03.2017]. Dostępny w: <http://files.dnb.de/nesstor/veranstaltungen/2013-02-27-providing-access-to-the-DNB-web-archive.pdf>.
9. Webis. Wissenschaftliche Bibliotheken [online]. [Dostęp 20.03.2017]. Dostępny w: http://webis.sub.unihamburg.de/webis/index.php/Wissenschaftliche_Bibliotheken.
10. http://www.dfg.de/download/pdf/dfg_im_profil/evaluation_statistik/programm_evaluation/studie_evaluierung_sondersammelgebiete_empfehlungen.pdf. [Brak dostępu].
11. Spaniol, M., Mazeika, A., Denev, D. & Weikum, G. 2009. Catch me if you can: Visual Analysis of Coherence Defects in Web Archiving. In: The 9th International Web Archiving Workshop (IWAW 2009). Workshop Proceedings [online]. P. 27-38. [Dostęp 20.03.2017]. Dostępny w: <http://lekythos.library.ucy.ac.cy/handle/10797/14075>.
12. Gray, G., Scott, M. 2013. Choosing a Sustainable Web Archiving Method: A Comparison of Capture Quality. In: Dlib Magazine 19 [online]. Number 5/6 (2013). [Dostęp 20.03.2017]. Dostępny w: <http://dx.doi.org/10.1045/may2013-gray>.
13. Niu, J. 2012. Functionalities of Web Archives. In: D-Lib Magazine [online]. 18, Number 3/4 (2012). [Dostęp 20.03.2017]. Dostępny w: <http://dx.doi.org/10.1045/march2012-niu2>.
14. Pennock, M. 2013. Web-Archiving: DPC Technology Watch Report [online]. 13-01 March 2013, p. 23. [Dostęp 20.03.2017]. Dostępny w: <http://dx.doi.org/10.7207/twr13-01>.
15. Mementos, UK Web Archives [online]. [Dostęp 20.03.2017]. Dostępny w: <http://www.webarchive.org.uk/ukwa/info/mementos>.

Przekład z angielskiego Bożena Bednarek-Michalska

Artykuł został przetłumaczony za zgodą autorów. Tekst oryginału: Kugler, A., Beinert T., Schoger A., *Web Archiving as a Service for the Sciences*, Münchener Digitalisierungs Zentrum (MDZ) 2013. [online]. [Dostęp 20.02.2017]. Dostępny w: <https://www.digitale-sammlungen.de/content/dokumente/Web%20Archiving%20as%20a%20Service%20for%20the%20Sciences.pdf>.

Uzupełnienia redakcji zgodne z sugestią autorów. Artykuł został opublikowany w roku 2013, projekt jest rozwijany, a jego stan obecny opisany jest na stronie biblioteki: *Langzeitarchivierung von Websites durch Gedächtnisinstitutionen: Entwicklung eines Servicemodells auf Grundlage praktischer Erfahrungen* [online]. [Dostęp 20.02.2017]. Dostępny w: https://www.babs-muenchen.de/index.html?c=projekte_webarchivierung&.