



## Egodokumentalne ślady człowieka w Internecie i ich archiwizacja

Bartłomiej Konopa

Naczelna Dyrekcja Archiwów Państwowych  
bartlomiejkonopa@gmail.com, ORCID 0000-0001-9843-5552

Agnieszka Rosa

Uniwersytet Mikołaja Kopernika w Toruniu  
arosa@umk.pl, ORCID 0000-0002-8560-1387

### STRESZCZENIE

Nadrzędnym celem artykułu jest charakterystyka cyfrowych źródeł born digital, które mają cechy gatunkowych egodokumentów, oraz problematyki ich archiwizacji. Źródła te zawierają ślady związane z obecnością i działaniami człowieka w postaci tekstu, lub tekstu z obrazem, i dają badaczowi możliwość poznania emocji, doświadczeń, wrażliwości czy życia osobistego ich twórcy. Autorzy artykułu poszukują odpowiedzi na pytania, czy człowiek zostawia w Sieci w postaci swoich śladów/wytworów egodokumenty, a także jakiego typu są to źródła. Czy możliwa jest, i jak przebiega, ich archiwizacja, czy zarchiwizowane źródła dostępne są dla badaczy, a następnie, czy możliwe jest ich wykorzystanie w badaniach egodokumentalnych. Ze względu na przyczynkowy charakter artykułu, autorzy podjęli decyzję, że pod uwagę wzięte zostaną dwa rodzaje wytworów człowieka w Internecie, tj. osobiste witryny domowe i blogi. Wskazano na ich egodokumentalny charakter. Zwrócono także uwagę na procesy archiwizacji tej części Webu i problemy z nimi związane, w szczególności metody i strategie archiwizacji, ograniczenia prawne, techniczne czy etyczne. Następnie omówiono kwestie związane z dostępem do zarchiwizowanych zasobów, gdzie znaczącą rolę odgrywają dwa zagadnienia: otwartość archiwum oraz metody przeszukiwania jego zasobu. Autorzy przywołali także wyniki wybranych badań naukowych, mając na celu ukazanie możliwości wykorzystania zarchiwizowanych egodokumentów Webowych w dalszych badaniach. W artykule wykorzystano analizę tekstów i badań dotyczących teorii egodokumentu oraz archiwizacji Webu, przede wszystkim jednak metodę porównawczą, rekonceptualizację definicji egodokumentu, a także definicji śladów cyfrowych dookreślonych jako całości egodokumentalne – czyli źródła born digital, które zarchiwizowane umożliwiają badaczom przeszłości za pomocą analizy tekstu (a nie kodów binarnych) poznanie osobowości i życia twórcy.

### SŁOWA KLUCZOWE

egodokument,  
egodokumentalność,  
archiwizacja Webu,  
blog, blogosfera,  
osobista witryna  
domowa, cyfrowe  
źródła historyczne,  
materiały  
archiwalne,  
internetowy  
pamiętnik,  
ślady człowieka  
w Internecie

### Ego-documentary human traces on the Internet and their archiving

#### ABSTRACT

The primary goal of the article is to present the issues of born digital sources, which have the characteristics of generic egodocuments, and the issues of their archiving. These sources contain traces related to human

#### KEY WORDS

egodocument,  
egodocumentality,  
Web archiving,

presence and actions in the form of text or text with images and give the researcher the opportunity to learn about the emotions, experiences, sensitivity, and personal life of their creator. The authors of the article are therefore looking for answers to the questions of whether people leave ego-documents on the Internet in the form of their traces or products, what type of sources they are, whether their archiving is possible and how it is carried out, whether archived sources are available to researchers, and then whether it is possible to use them in egodocumentary research. Ultimately, the authors decided that, due to the contributing nature of the article, two types of human creations on the Internet would be taken into account, i.e., personal home websites and blogs. Their ego-documentary nature was pointed out. Attention was also paid to the archiving processes of this part of the Web and the problems associated with them, in particular archiving methods and strategies and legal, technical, and ethical limitations. Then, issues related to access to archived resources were discussed, where two issues play a significant role: the openness of the archive and methods of searching its resources. The authors also cited selected scientific research, aiming to show the possibility of using archived Web egodocuments in research. The article uses the analysis of texts and research on the theory of egodocument and Web archiving, but above all, the comparative method, reconceptualization of the definition of egodocument, as well as the definition of digital traces defined as egodocumentary wholes – that is, born digital sources that, when archived, enable researchers of the past using text analysis (and not binary codes) to learn about the creator's personality and life.

blog, blogosphere, personal webpage, digital historical sources, archival materials, online diary, human trace on the Internet

## Wstęp

Autorzy artykułu prowadzą na co dzień badania naukowe mieszczące się w zakresie nauk historycznych, jednak odległe od siebie tematycznie. Pierwsze dotyczą teorii egodokumentu, jej ewolucji i badań nad egodokumentalnością źródeł – nie tylko ich określonych gatunków, takich jak pamiętniki czy testamenty, ale także zarchiwizowanych całości dokumentacyjnych na różnym poziomie opisu archiwalnego<sup>1</sup> (np. jednostek archiwalnych, serii akt, zespołów archiwalnych, spuścizn itp.). Drugie to badania nad archiwizacją Webu: metodami jego archiwizacji<sup>2</sup>, projektami archiwizacyjnymi, ale także zasobami archiwów Webowych<sup>3</sup>, możliwościami ich przeszukiwania i wykorzystania w badaniach przez history-

<sup>1</sup> A. Rosa, *Egodokumentalność polskiego zasobu archiwalnego*, [w:] *Archiwistyka między różnorodnością a standaryzacją*, Toruńskie Konfrontacje Archiwalne, t. 3, red. W. Chorążyczewski, A. Rosa, Toruń 2013, s. 103–112; eadem, *Egodokumenty a działalność archiwów*, [w:] *Zatrzymać przeszłość, dogonić przyszłość. Pamiętnik VI Powszechnego Zjazdu Archiwistów Polskich, Wrocław, 5–7 września 2012 r.*, red. W. Chorążyczewski, K. Strykowski, Toruń 2013, s. 271–275.

<sup>2</sup> B. Konopa, *Strategia selektywna jako narzędzie w archiwizacji Webu. Analiza wybranych przykładów*, „Archiwa – Kancelarie – Zbiory” 2020, nr 11(13), s. 97–118.

<sup>3</sup> Idem, *Archiwizacja Webu w Europie–narodowe archiwa Sieci*, „Archeion” 2020, t. 121, DOI: 10.4467/26581264ARC.20.016.12973, s. 445–465.

ków i badaczy innych dziedzin. Podjęta w artykule tematyka – egodokumenty born digital i proces ich archiwizacji – wymagała współpracy specjalistów zajmujących się dwoma wyżej wymienionymi zagadnieniami.

Połączenie badań egodokumentalnych i nad archiwizacją webową dały interesujące wyniki i nowe koncepcje. Nie ulega wątpliwości, że Internet i jego zasoby wymagają archiwizacji. Archiwizacja ta jest przeprowadzana przez różne instytucje, organizacje i stowarzyszenia. Tak jak i w przypadku dokumentacji analogowej, nie obejmuje ona jednak całego spektrum zasobów webowych. Autorzy artykułu pragną zwrócić uwagę na tę ich część, którą można określać i badać jako egodokumenty oraz na proces ich archiwizacji. Archiwizacja w tym przypadku rozumiana jest zgodnie z teorią archiwalną jako cały proces zarządzania zarchiwizowanymi zasobami, włączając w to zarówno ich gromadzenie, opis i przechowywanie, jak i udostępnianie oraz możliwości wykorzystania<sup>4</sup>.

Nadrzędnym celem artykułu jest ukazanie problematyki występowania w Sieci cyfrowych źródeł born digital, zawierających ślady związane z obecnością i działaniami człowieka w Internecie, utrwalonych w postaci tekstu lub tekstu z obrazem, mających cechy gatunkowych egodokumentów. Pozwalają one badaczom na poznanie emocji, doświadczeń, wrażliwości czy życia osobistego twórcy. Celem opracowania jest także przedstawienie zagadnienia ich archiwizacji i możliwości wykorzystania w badaniach. Autorzy artykułu poszukują odpowiedzi na pytania, czy człowiek tworzy w Sieci – w postaci swoich śladów (wytworów) – egodokumenty; jakiego typu są to źródła; czy możliwa jest i jak przebiega ich archiwizacja; czy zarchiwizowane źródła dostępne są dla badaczy, a także czy możliwe jest ich wykorzystanie w badaniach egodokumentalnych, i nie tylko.

Główną tezę artykułu jest założenie, że w Sieci powstają egodokumenty, które warto archiwizować i przechowywać wiecześnie. Pisząc o egodokumentach, autorzy odwołują się do definicji przedstawionej przez Jacoba Pressera<sup>5</sup>, rozszerzonej następnie przez Winfrieda Schulze'a (Selbstzeugnis)<sup>6</sup>. Wskazuje ona na określone teksty źródłowe, z których badacz może czerpać informacje o ich autorze (w przy-

<sup>4</sup> W ten sposób archiwizację Webu rozumie m.in. International Internet Preservation Consortium zrzeszające instytucje archiwizujące zasoby WWW: International Internet Preservation Consortium. Web Archiving, <https://netpreserve.org/web-archiving/> [dostęp: 20.08.2023].

<sup>5</sup> R.M. Dekker, *Ego-Documents in the Netherlands 1500-1814*, „Dutch Crossing” 1989, vol. 39, s. 61–72.

<sup>6</sup> W. Schulze, *Ego-Dokumente. Annäherung an den Menschen in der Geschichte? Vorüberlegungen für die Tagung „Ego-Dokumente“*, [w:] *Ego-Dokumente. Annäherung an den Menschen in der Geschichte*, ed. W. Schulze, Berlin 1996, s. 11–30.

padku np. listów, dzienników, wspomnień) lub bohaterze (w przypadku np. testamentów urzędowych, suplik chłopskich, podań i życiorysów). Informacje te to dane dotyczące osobowości, doświadczeń, światopoglądu, emocji czy przeżyć autora lub bohatera tego źródła.

Do analizy i opisu wytypowano dwa rodzaje śladów, rozumianych jako wytwory człowieka w Internecie, tj. osobiste witryny domowe i blogi, przyjmując je za najbardziej reprezentatywne dla tej grupy źródeł, i określając jako egodokumentalne wytwory/ślady człowieka w Internecie. Zgłębiając problematykę działalności człowieka w Sieci, autorzy natrafili na koncepcję „cyfrowych śladów człowieka” stosowaną w badaniach psychometrycznych i kryminalistycznych, popularyzowaną i rozwijaną zwłaszcza przez psychologa Michała Kosińskiego<sup>7</sup>, specjalistę z zakresu kryminologii Wojciecha Kasprzaka<sup>8</sup> oraz informatologów<sup>9</sup>. Ślady te, w postaci kodów binarnych, dzielone są ze względu na informacje, które niosą, na cztery grupy: demograficzne, geograficzne, technologiczne i behawioralne<sup>10</sup>.

Koncepcja ta odbiega jednak od rozumienia cyfrowych śladów człowieka przyjętych na potrzeby tego artykułu. Pojmujemy je jako webowe wytwory człowieka w postaci tekstu lub tekstu z obrazem albo inną treścią, które można poddać analizie egodokumentalnej<sup>11</sup>, a nie wyłącznie jako kody binarne będące odzwierciedleniem aktywności człowieka w Sieci, na podstawie których można poznać preferencje użytkownika lub jego aktywność.

Psychometryczna koncepcja cyfrowych śladów człowieka i możliwości jej wykorzystania w badaniach są jednak w pewnym zakresie bliskie koncepcji badań egodokumentalnych. Obie pragną poznać człowieka, jego osobowość, emocje czy preferencje. Odróżnia je jednak cel badań. W przeciwieństwie do badań psycho-

<sup>7</sup> M. Kosinski, D.J. Stillwell, T. Graepel, *Private traits and attributes are predictable from digital records of human behavior*, „Proceedings of the National Academy of Sciences of the United States of America” 2013, vol. 110(15), DOI: 10.1073/pnas.1218772110, s. 5802–5805.

<sup>8</sup> W.A. Kasprzak, *Ślady cyfrowe. Studium prawno-kryminalistyczne*, Warszawa 2015.

<sup>9</sup> M. Żytomirski, *Określanie osobowości użytkowników Internetu poprzez analizowanie ich cyfrowych śladów w świetle wybranych badań dr. Michała Kosińskiego*, „Acta Universitatis Lodzianensis Folia Librorum” 2012, t. 1(32), DOI: 10.18778/0860-7435.32.05, s. 99–113.

<sup>10</sup> Idem, R. Żulicki, *Cyfrowe ślady a banki filtrujące – prezentacja informacyjna*, 15 August 2020, [https://www.researchgate.net/publication/343678108\\_Cyfrowe\\_slady\\_a\\_banki\\_filtrujace\\_-\\_prezentacja\\_informacyjna](https://www.researchgate.net/publication/343678108_Cyfrowe_slady_a_banki_filtrujace_-_prezentacja_informacyjna) [dostęp: 6.08. 2023].

<sup>11</sup> A. Rosa, K. Siuda, *Analiza egodokumentalna pamiętników archiwistów jako przyczynek do poznania mentalności określonej grupy zawodowej*, [w:] *Pogranicza archiwistyki*, Toruńskie Konfrontacje Archiwalne, t. 6, red. W. Chorążyczewski, A. Rosa, Toruń 2019, s. 85–100.

metrycznych, badania egodokumentalne poprzez analizę tekstu mają na celu poznanie człowieka – takiego, jakim on sam się kreuje, jego subiektywnego oglądu świata. Nie skupia się zatem ona na tym, co można za pomocą analiz ilościowych odczytać o człowieku z jego śladów cyfrowych pozostawionych w postaci kodów binarnych, lecz zakłada możliwość analizy śladów człowieka pozostawionych w postaci tekstów (też z obrazem lub inną treścią), obojętnie w jakim środowisku (cyfrowym lub niecyfrowym) autor sam o sobie opowiada. Na podstawie tych śladów badacz przeszłości, za pomocą analizy egodokumentalnej, może poznawać osobowość i świat wartości ludzi, którzy te teksty tworzyli lub są ich bohaterami. Zbadać to, co sami o sobie piszą, jak się postrzegają i jak chcą być postrzegani oraz co nimi kieruje. Nie zawsze przedstawione fakty będą zgodne z obiektywną prawdą. Celem badań egodokumentalnych nie jest jednak dążenie do prawdy obiektywnej o człowieku, lecz poznanie jego subiektywnego świata.

Dane do artykułu zebrano za pomocą analizy tekstów i badań dotyczących teorii egodokumentu oraz archiwizacji Webu. Zastosowano także metodę porównawczą, rekonceptualizację definicji egodokumentu (jej rozszerzenie na wybrane źródła webowe), a także definicji śladów cyfrowych (dookreślonych jako całości egodokumentalne – czyli źródła born digital, które zarchiwizowane umożliwiają badaczom, za pomocą analizy tych wytworów, poznanie subiektywnego świata, osobowości i życia twórcy).

Ostatecznie autorzy podjęli decyzję, że ze względu na przyczynkowy charakter artykułu – dotyczy on bowiem kwestii archiwizacji specyficznych zasobów webowych, dotychczas nie badanych w kategoriach egodokumentu – pod uwagę wzięte zostaną dwa rodzaje wytworów człowieka w Internecie, tj. osobiste witryny domowe i blogi. W tekście wskazano na ich egodokumentalny charakter. Zwrócono także uwagę na procesy archiwizacji tej części Webu i problemy z nimi związane, w szczególności metody i strategie archiwizacji, ograniczenia prawne, techniczne czy etyczne. Następnie omówiono kwestie związane z dostępem do zarchiwizowanych zasobów, gdzie znaczącą rolę odgrywają dwa zagadnienia: otwartość archiwum oraz metody przeszukiwania jego zasobu. Autorzy, mając na celu ukazanie możliwości wykorzystania zarchiwizowanych egodokumentów Webowych w badaniach, przywołali także wybrane prace naukowe z tego zakresu.

W artykule odwołano się do szeregu opracowań z zakresu teorii egodokumentu oraz dotyczących badań przeszłości prowadzonych w oparciu o te źródła. Istotne dla autorów artykułu były zwłaszcza prace badaczy wywodzących się z toruńskiego ośrodka badań nad egodokumentami: Waldemara Chorążyczewskiego,

Agnieszki Rosy czy Stanisława Roszaka. Odniesiono się także do literatury prezentującej badania poświęcone osobistym witrynom domowym i blogom, w tym powodom ich zakładania oraz charakterystyce publikowanych na nich treści, a także do ogólnych opracowań dotyczących archiwizacji zasobów WWW.

## Egodokumentalność zasobów WWW

Internet umożliwia komunikację, publikację treści, udostępnianie usług, oferuje wiele form rozrywki, daje możliwość dokonywania zakupów oraz wykonywania operacji finansowych, a przez to znacząco wpływa na zachowania ludzi i wybory przez nich dokonywane. Pozwala użytkownikom na kreowanie swojej tożsamości i daje możliwość wyrażania siebie, ukazania swoich spraw, doświadczeń, problemów czy zainteresowań. Człowiek, w zależności od przyjętej roli, aktywnie wpływa na kształt Sieci, tworząc lub korzystając z różnych witryn, aplikacji i usług. Poprzez swoje działania buduje jej zasoby, czasem mniej, czasem bardziej świadome, pozostawia też w Sieci swoje ślady (wytwory)<sup>12</sup>. Te z kolei – nie tylko współcześnie, ale też w przyszłości – mogą być wykorzystywane jako źródła<sup>13</sup>, których analiza przybliży badacza i czytelników do subiektywnego świata twórcy i społeczności, w której funkcjonował: jego spraw, działań, być może także upodobań, problemów, życia osobistego, emocji czy doświadczeń. Wyzwaniem teraźniejszości jest zatem odpowiednie zabezpieczenie tego typu śladów.

Sieć jest archiwizowana od 1996 r., w różnej skali i natężeniu, z różną dokładnością i skutkiem<sup>14</sup>. Zarchiwizowane materiały gromadzone są w archiwach webowych, a trafiające do nich zasoby to nie tylko rządowe czy oficjalne witryny, ale także treści osobiste, tworzone przez osoby prywatne, w tym także te, które badacz tych źródeł może wykorzystać jako egodokumenty<sup>15</sup>. Są to teksty oraz inne

<sup>12</sup> Zob. A. Rosa, *Human trace on the Internet – the issue of archiving the Web from the point of view of anthropology-oriented archival science*, „Archiwa – Kancelarie – Zbiory” 2015, nr 6(8), DOI: 10.12775/AKZ.2015.006, s. 193–205.

<sup>13</sup> Zob. M. Wilkowski, *Strona internetowa jako źródło historyczne*, „Historyka. Studia Metodologiczne” 2021, t. 51, DOI: 10.24425/hsm.2021.138366, s. 83–110.

<sup>14</sup> B. Konopa, *Periodyzacja archiwizacji Webu*, <https://webarch.uw.edu.pl/zasoby/periodyzacja-archiwizacji-webu/> [dostęp: 28.08.2021]; idem, *Archiwizacja Webu w Europie...*

<sup>15</sup> Na potrzeby artykułu przyjmujemy definicję egodokumentu podającą, iż jest to każdy dokument poświadczający życie osobiste i doświadczenie osoby, który został wykorzystany przez badacza w celu odnalezienia umysłowości i tożsamości autora, za: W. Chorążyczewski, A. Rosa, *Egodokumenty – egodokumentalność – analiza egodokumentalna – spuścizna egodokumentalna*, [w:]

treści, głównie o charakterze autobiograficznym, mające jednak inną formę niż te zgromadzone w tradycyjnych archiwach i bibliotekach. Ich cyfrowość sprawia, że archiwizuje się je w odmienny sposób, inaczej też zabezpiecza i udostępnia. Badacz ponadto musi być wyposażony w umiejętności (także w zakresie krytyki źródeł) korzystania z tego typu archiwów i zgromadzonych w nich źródeł<sup>16</sup>.

Sama problematyka egodokumentów znajdujących się w zasobach tradycyjnych archiwów i bibliotek jest niezbyt często podejmowana w rozważaniach naukowych. Zazwyczaj są to dociekania osób zajmujących się badaniem przeszłości. Ponadto, z uwagi na specyfikę danych, które wzbudzają zainteresowanie badaczy egodokumentów, zagadnienie to nadal bywa dyskusyjne<sup>17</sup>. To, co dla badaczy egodokumentów jest ich istotą – afektywność, dla innych staje się często elementem pozwalającym na podważenie wiarygodności źródła i jego odrzucenie, w części lub w całości, w badaniach.

Zdając sobie sprawę ze specyfiki badań egodokumentalnych, pragniemy zwrócić uwagę na zasoby znajdujące się w Sieci, które – choć bardziej efemeryczne – często są następcami/odpowiednikami/nowymi formami tradycyjnych pamiętników, dzienników i innych tekstów autobiograficznych. Egodokumentalna perspektywa badawcza pozwala na odmienne od dotychczasowego spojrzenie na internetowe źródła oraz traktowanie ich jako swoistych śladów pozostawionych przez współczesnego człowieka, które mogą dać możliwość poznania subiektywnych doświadczeń zarówno pojedynczego człowieka, jak i zbiorowego obrazu umysłowości współczesnych grup, pokolenia czy – w przyszłości – kultury.

Rozważając problematykę zachowania śladów człowieka w Sieci, jak już zapowiedzieliśmy, szczegółowo zajmiemy się jego dwiema konkretnymi formami, które do tej pory udało się w różnej skali, lecz skutecznie, archiwizować, i których przykłady możemy już znaleźć w zasobie inicjatyw gromadzących zasoby Webu<sup>18</sup>. W odróżnieniu od pojedynczych komentarzy (np. tweet), uwagę skupiliśmy na

---

*Egodokumenty, tradycje historiograficzne i perspektywy badawcze*, red. W. Chorążyczewski, A. Pacевичius, S. Roszak, Toruń 2015, s. 11–22.

<sup>16</sup> I. Milligan, *History in the Age of Abundance? How the web is transforming historical research*, Montreal 2021, s. 106–142.

<sup>17</sup> Zob. np. *Egodokumenty, tradycje historiograficzne...*; S. Roszak, *Egodokumenty – kilka uwag o polskich i europejskich doświadczeniach historiograficznych oraz metodologicznych*, „Biuletyn Polskiej Misji Historycznej” 2013, t. 8, DOI: 10.12775/BPMH.2013.001, s. 27–42.

<sup>18</sup> Kolekcja blogów prawniczych w zasobie Biblioteki Kongresu USA: Library of Congress. Legal Blawgs Web Archive, <https://www.loc.gov/collections/legal-blawgs-web-archive/about-this-collection/> [dostęp: 28.05.2021]; Kolekcja brytyjskich blogów w zasobie UK Web Archive: UK Web Archive. Blogs, <https://www.webarchive.org.uk/en/ukwa/collection/44> [dostęp:

cyfrowych wytworach człowieka, często budowanych latami, które zazwyczaj są odpowiednikami tradycyjnych form, takich jak pamiętniki, dzienniki czy życiorysy. Niejako intencjonalnie mają one charakter egodokumentów. Ludzie tworzą je, żeby móc wyrazić siebie, być zapamiętanym, zrozumianym, odkrytym, pozostawić po sobie ślad czy opowiedzieć swoją historię. Jest to ich celowe i świadome działanie, raczej nie poparte jedynie chwilowym impulsem. Prywatne witryny domowe (ang. *homepage*) oraz blogi, zwłaszcza te o charakterze pamiętnikarskim, są często źródłami informacji na temat drugiego człowieka mającymi bardzo bezpośredni i osobisty charakter. Oba te rodzaje witryn internetowych można potraktować jako egodokumenty, czego przesłanki zostaną zaprezentowane w następujących akapitach.

Z omówienia zostały włączone media społecznościowe, które również są oczywistym, bardziej współczesnym i mocno osobistym, zbiorem egodokumentalnych śladów człowieka w Sieci. Zdecydowano tak, ponieważ ich archiwizacja jest utrudniona ze względu na kwestie techniczne, etyczne, prawne, a także politykę właścicieli i administratorów tych zasobów. Pomimo trudności, próby ich zabezpieczenia są podejmowane przez różne instytucje<sup>19</sup>.

## Egodokumentalność witryn domowych

Jako pierwsze zdecydowaliśmy się omówić witryny domowe, ponieważ są chronologicznie wcześniejsze od blogów i funkcjonują niemalże od samego początku rozpowszechnienia się World Wide Web (były to lata 1994–1999). Przykładem może być tu serwis GeoCities, którego początki sięgają końca 1994 r., pozwalający użytkownikom Sieci na łatwe i darmowe prowadzenie własnej witryny internetowej<sup>20</sup>. Witryna domowa lub witryna osobista jest prowadzona przez jed-

---

28.05.2021]; O archiwizacji GeoCities przez ArchiveTeam: GeoCities, <https://wiki.archiveteam.org/index.php/GeoCities> [dostęp: 28.05.2021].

<sup>19</sup> Zob.: C.C. Marshall, F.M. Shipman, *On the Institutional Archiving of Social Media*, [w:] JCDL '12: *Proceedings of the 12<sup>th</sup> ACM/IEEE-CS joint conference on Digital Libraries*, New York 2012, DOI: 10.1145/2232817.2232819, s. 1–10; H. Hockx-Yu, *Archiving Social Media in the Context of Non-print Legal Deposit*, presented at IFLA WLIC 2014 in Lyon, France, 16–22 August 2014, <http://library.ifla.org/id/eprint/999> [dostęp: 6.07.2021].

<sup>20</sup> Na temat historii GeoCities zob.: I. Milligan, *Welcome to the web. The online community of GeoCities during the early years of the World Wide Web*, [w:] *The Web as History. Using Web Archives to Understand the Past and the Present*, ed. N. Brügger, R. Schroeder, UCL Press 2017, DOI: 10.2307/j.ctt1mtz55k.13, s. 137–140.



ną osobę (rzadziej grupę, np. rodzinę) oraz wypełniana treścią, którą postanowi ona tam zamieścić<sup>21</sup>. Taką witrynę można rozumieć jako pewnego rodzaju wyraz autoekspresji i autopromocji, stanowiący hipertekstowy amalgamat życiorysu, CV, wizytówki, albumu z fotografiami, wpisami i podobnych form. Wytwory te można też traktować jako sieciowy odpowiednik *silva rerum*<sup>22</sup>. Dzięki udostępnieniu łatwych w obsłudze edytorów stron internetowych oraz systemów zarządzania treścią (*content management system* – CMS), Internauci zyskali nie tylko możliwość multimedialnego wyrażania siebie, ale także sposobność łatwego publikowania swoich treści z pominięciem wysoce selektywnego procesu wydawniczego<sup>23</sup>. Warto zauważyć, że do kreowanego w ten sposób wizerunku, używają często ustrukturyzowanych informacji<sup>24</sup>.

Obecnie, w czasie dominacji mediów społecznościowych, popularność witryn domowych spadła, ale w pierwszych dwóch dekadach istnienia WWW były one dość popularne. Dowodem tego może być przywoływane już GeoCities, które w 1997 liczyło około miliona użytkowników i było jednym z popularniejszych miejsc w Sieci<sup>25</sup>. Przeprowadzony przez Nicolę Döring przegląd badań naukowych dotyczących prywatnych witryn domowych wykazał, że dotychczas analizowane były one z perspektyw teorii tożsamości oraz autoprezentacji. Właściciele osobistych witryn internetowych często wykorzystują je jako narzędzie do tworzenia i rozpowszechniania własnego wizerunku<sup>26</sup>, m.in. poprzez zamieszczanie na nich danych na swój temat, przez co możliwe jest ich autobiograficzne odczytanie<sup>27</sup>. Poza treścią, za istotną można uznać warstwę wizualną witryny, zarówno jej część widoczną (układ elementów, zastosowane kolory i czcionki) jak i niewidoczną (kod źródłowy)<sup>28</sup>. Informacje na temat autora witryny mogą zdradzać również

<sup>21</sup> N. Döring, *Personal Home Pages on the Web. A Review of Research*, „Journal of Computer-Mediated Communication” 2002, vol. 7, no 3, DOI:10.1111/j.1083-6101.2002.tb00152.x.

<sup>22</sup> Zob. np. S. Roszak, *Archiwa sarmackiej pamięci. Funkcje i znaczenie rękopiśmiennych ksiąg typu silva rerum w kulturze Rzeczypospolitej XVIII wieku*, Toruń 2004.

<sup>23</sup> J.D. Bolter, *Przestrzeń pisma. Komputery, hipertekst i remediacja druku*, tłum. A. Małecka, M. Tabaczyński, Kraków–Bydgoszcz 2014, s. 152–154.

<sup>24</sup> T. Erickson, *The World-Wide-Web as social hypertext*, „Communications of the ACM” 1996, vol. 39, no 1, DOI:10.1145/234173.234174, s. 16.

<sup>25</sup> I. Milligan, *Welcome to the web*, s. 138–139.

<sup>26</sup> N. Döring, *Personal Home Pages*.

<sup>27</sup> L. Warley, *Reading the autobiographical in personal home pages*, [w:] *Tracing the autobiographical*, red. M. Kadar, S. Egan, L. Warley, J. Perreault, Ontario 2005, s. 26–28.

<sup>28</sup> R. Dopierała, *Internetowe manifestacje prywatności*, „Media – Kultura – Społeczeństwo” 2008, t. 1(3), s. 112.

zamieszczane na niej hiperłącza, które wskazują na treści przez niego lubiane lub takie, z którymi się identyfikuje lub interesuje<sup>29</sup>. W badaniach egodokumentalnych istotne jest zarówno to, co autor chce przekazać, jak i to, w jaki sposób planuje to zrobić. Liczy się zatem nie tylko sama treść, ale również forma i przekaz, a także to, czego nie ma, co autor ukrywa, pomija i przemilcza. W przypadku witryn domowych można zwrócić uwagę również na to, na ile autor korzystał z gotowych rozwiązań oraz jak bardzo chciał podkreślić, choćby poprzez formę, swój indywidualizm. Interesujące jest także to, w jakich sytuacjach, dlaczego, co i po co publikował, a także czy umożliwiał czytelnikom wchodzenie w interakcję poprzez dodanie opcji komentowania lub zadawania pytań za pomocą dostępnych narzędzi cyfrowych.

W celu pokazania możliwości egodokumentalnego odczytania witryny domowej postanowiliśmy poddać uproszczonej analizie witrynę należącą do Dragana Espenschieda<sup>30</sup>, artysty nowych mediów i badacza kultury cyfrowej, a także muzyka. Wraz z Olią Lialiną jest on współtwórcą projektu „One Terabyte of Kilobyte Age” oraz kierownikiem Programu Konserwacji Cyfrowej w Rhizome. Pierwszą rzeczą, która rzuca się w oczy, jest bardzo prosty design witryny, na który składają się białe tło, czarny kolor fontu i brak rozbudowanych elementów graficznych, co można odebrać jako odwołanie do witryn z początków WWW. Koresponduje to z zainteresowaniami D. Espenschieda i może stanowić sposób na ich wyrażenie. Zawartość strony głównej to niekompletny wykaz jego prac z lat 1991–2013 opatrzonych linkami – zarówno wewnętrznym, jaki i zewnętrznymi – odsyłającymi do miejsc, gdzie można się z nimi zapoznać. Oprócz krótkich opisów dołączono do nich słowa kluczowe, pośród których mocno wyróżnia się nazwisko O. Lialiny, ponieważ zostało ono podkreślone tęczowym gradientem. Taki zabieg może wskazywać na bliskie relacje i współpracę pomiędzy obojgiem artystów. Ponadto pod nagłówkiem znajduje się link do poprzedniego bloga artysty, który również zawiera informacje o jego projektach, archiwum krótkich wiadomości oraz inne informacje z nim związane. W sekcji „About”, oprócz danych kontaktowych, umieszczono biogram D. Espenschieda w dwóch postaciach, zapisanych za pomocą 222 i 91 słów, w których podkreśla on swoją działalność artystyczną oraz współpracę z O. Lialianą, w tym ich wspólny projekt „One Terabyte of Kilo-

<sup>29</sup> J.R. Dominick, *Who Do You Think You Are? Personal Home Pages and Self-Presentation on the World Wide Web*, „Journalism & Mass Communication Quarterly” 1999, vol. 76, no 4, DOI: 10.1177/107769909907600403, s. 646–647.

<sup>30</sup> Dragan Espenschied, <http://1x-upon.com/~despens/> [dostęp: 22.09.2023].

byte Age”. Na stronie znajduje się również funkcja pozwalająca na zmianę imienia oraz nazwiska artysty (np. na Dragon Espenbaum), co pokazuje świadomość ich obcego brzmienia oraz swobodne podejście do wynikających z tego trudności. Zamieszczono też na niej odnośnik do wykazu wydarzeń, w których D. Espenschied występował. Ujęto w nim m.in. wystawy, wykłady oraz koncerty z lat 1999–2014. Podsumowując charakterystykę witryny prowadzonej przez D. Espenschieda, zauważyć można, że jest to jego artystyczne portfolio, swoisty pamiętnik, podkreślający zainteresowanie sztuką wykorzystującą komputery (zwłaszcza sztuką Internetu), pozwalający poznać przebieg jego kariery i relacje z osobami, z którymi współpracował, jego stylistykę, umiejętności, sposób, w jaki chce być odbierany przez widownię, w jaki się kreuje, co mu się podoba, jak chce być zapamiętany. Co ciekawe, mimo, iż stworzył nową witrynę domową, nadal utrzymuje starą i umożliwia korzystanie z niej, traktując ją jako jedno ze swoich dzieł.

## Egodokumentalność blogów

Drugą formą osobistych cyfrowych śladów internautów są blogi, zwłaszcza te, które pełniły rolę internetowego pamiętnika. Podobnie jak prywatne witryny domowe wyłoniły się one w pierwszych latach istnienia Sieci, jednak częściej są utożsamiane z epoką Webu 2.0. Weblog, od którego pochodzi skrócona nazwa blog, to witryna WWW poddawana częstym aktualizacjom ze względu na dodawanie do niej nowych treści, które podzielone są na ułożone chronologicznie posty. Dzięki istnieniu łatwych w obsłudze narzędzi i serwisów, takich jak WordPress czy Blogger, powstały prawdopodobnie setki milionów blogów, dające internautom możliwość podzielenia się własnymi doświadczeniami, opiniami i przemyśleniami. Same blogi są bardzo zróżnicowane: obejmują zarówno amatorskie i intymne dzienniki, jak i profesjonalne blogi naukowe<sup>31</sup>. Można zaryzykować hipotezę, że większość z nich znajduje się bliżej pierwszego końca tego spektrum.

Tak jak w przypadku witryn domowych, blogi stworzyły przestrzeń do łatwego publikowania swoich spostrzeżeń, dając tym możliwość wypowiedzenia się

<sup>31</sup> M. Jones, I. Alony, *Blogs – The New Source of Data Analysis*, „Issues in Informing Science and Information Technology” 2008, vol. 5, DOI: 10.28945/1019, s. 433–435.

grupom wcześniej w źródłach historycznych raczej nieobecny, np. młodzieży<sup>32</sup> czy przedstawicielom różnych mniejszości. Ponadto, pomimo tego, że na pierwszy rzut oka Internet wolny jest od granic państwowych, sieciowe pamiętniki często przedstawiają różne wydarzenia i zjawiska z bardziej lokalnej perspektywy<sup>33</sup>. Osobiste blogi można traktować również jako indywidualny ślad człowieka pozostawiony w Sieci. Badaczka komunikacji komputerowej Erin E. Hollenbaugh w swoim badaniu dotyczącym motywów prowadzenia internetowego dziennika wyłoniła 7 głównych czynników, wśród których znalazły się: relacje społeczne, ekshibicjonizm, organizowanie własnych przemyśleń oraz otrzymywanie opinii<sup>34</sup>. Wydaje się, że do listy należałoby dodać również charakter terapeutyczny prowadzenia bloga<sup>35</sup>. Może stanowić on miejsce, w którym jego autor prezentuje własne zainteresowania, ale przede wszystkim przemyślenia i odczucia. Robi to, aby je zachować i uporządkować, a także podzielić się nimi z szerszą publiką. Poza elementami wskazanymi w przypadku witryn domowych (treść, wygląd, hiperłącza), na blogach bogate w informacje mogą być także takie elementy jak sekcja komentarzy oraz księga gości, umożliwiające wchodzenie w interakcję z czytelnikami. Oba komponenty umożliwiają poznanie i określenie relacji autora z odbiorcami jego dziennika. W pierwszym możemy spodziewać się dyskusji z nimi, w drugim zaś listy osób, u których wzbudził on zainteresowanie<sup>36</sup>.

Biorąc pod uwagę przedstawioną powyżej charakterystykę osobistych witryn domowych oraz blogów, a także ich popularność w Sieci, należy zaznaczyć że, poza mediami społecznościowymi, zasoby te wydają się posiadać najbardziej egodokumentalny charakter i – w przeciwieństwie do tych pierwszych – łatwiej jest je zachować. Częściej można też je znaleźć w zasobie inicjatyw archiwizujących WWW.

<sup>32</sup> D. Huffaker, *Teen Blogs Exposed: The Private Lives of Teens Made Public*, Presented at the American Association for the Advancement of Science (AAAS) in St. Louis, MO., February 16–19, 2006, s. 2–4, <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.556.1504&rep=rep1&type=pdf>, [dostęp: 6.06. 2021].

<sup>33</sup> Y. Ibrahim, *Weblogs as Personal Narratives. Displacing History and Temporality*, „M/C Journal” 2006, vol. 9, no 6, DOI:10.5204/mcj.2690.

<sup>34</sup> E.E. Hollenbaugh, *Motives for Maintaining Personal Journal Blogs*, „Cyberpsychology, Behavior, and Social Networking” 2011, vol. 14, no 1–2, DOI:10.1089/cyber.2009.0403, s. 16–18.

<sup>35</sup> J. Lent, *Journaling Enters the 21<sup>st</sup> Century. The Use of Therapeutic Blogs in Counseling*, „Journal of Creativity in Mental Health” 2009, vol. 4, s. 68–73.

<sup>36</sup> M. Olcoń, *Blog jako dokument osobisty. Specyfika dziennika prowadzonego w Internecie*, „Kultura i Społeczeństwo” 2003, t. 47, nr 2, s. 126–127.

Na koniec tej części należy zwrócić uwagę na pewien wspólny problem witryn domowych i blogów, który wynika z zatarcia się granic pomiędzy tym co prywatne i publiczne. Treści na nich publikowane mają wysoce subiektywny charakter i powstają na styku tego, co autor chce przekazać, oraz tego, czego mogą oczekiwać jego odbiorcy. Demarkacja pomiędzy tym, co autentyczne, a tym, co stworzone na potrzeby publiki, zawsze jest trudna do przeprowadzenia i raczej pozostaje poza możliwościami badacza<sup>37</sup>. Jest to znana badaczom egodokumentów specyfika tego typu źródeł. Egodokumentalna perspektywa badawcza nie służy porządkowaniu faktów i odkrywaniu prawdy obiektywnej, umożliwia poznanie subiektywnego świata twórcy. Bazuje na emocjach, ukazuje doświadczenia i umysłowość osoby, grupy, pokolenia czy epoki. Jej celem jest przywracanie pamięci<sup>38</sup>.

Podobnie jak w przypadku witryn domowych, również w tym przypadku podaliśmy analizie egodokumentalnej, w ograniczonym zakresie, przykładowy blog. W tym celu wybraliśmy blog *Webstory*<sup>39</sup> należący do Petera Webstera, historyka zajmującego się dziejami chrześcijaństwa w dwudziestowiecznej Brytanii oraz historią cyfrową, w tym archiwizacją Webu. Omawiana witryna ma skromną szatę graficzną i została przygotowana z pomocą oprogramowania WordPress, a więc rozpowszechnionego i popularnego narzędzia CMS. Posty publikowane na blogu bardzo rzadko opatrzone są grafikami lub zdjęciami, widać więc, że autor skupia się na komunikowaniu się z odbiorcami za pomocą tekstu. Witryna ma charakterystyczną dla blogów budowę opartą o dwie kolumny: lewą – szerszą – zawierającą treść postów oraz prawą – z linkami do postów o powiązanej tematyce lub do zewnętrznych witryn, np. bazy ORCID. W górnej belce znajdują się zakładki prowadzące do 4 zasadniczych części: strony głównej bloga, sekcji o mnie, informacji o doradztwie oferowanym przez firmę autora oraz wykazu jego prac naukowych. Odnosząc się do samych treści publikowanych na blogu przez Webstera, należy zaznaczyć, że mają one ścisły związek z jego zainteresowaniami badawczymi. Znajdziemy tam fragmenty lub abstrakty jego prac naukowych, komentarze do ich recenzji czy nagrania z konferencji. Nie brakuje wśród nich także refleksji nad warsztatem historyka, archiwizacją Webu czy cyfrową humanistyką. Reasumując tę skróconą charakterystykę, zauważyć można, że blog prowadzony przez Webstera traktowany jest przez niego samego jak jego wizytówka, której celem jest umożliwienie poznania go jako profesjonalnego i zaangażowanego w swoją

<sup>37</sup> Ibidem, s. 141–143.

<sup>38</sup> Zob. S. Roszak, *Egodokumenty – kilka uwag o polskich...*, s. 27–42.

<sup>39</sup> *Webstory*. Peter Webster's blog, <https://peterwebster.me/> [dostęp: 27.10.2023].

pracę badawczą historyka. Chcąc być tak postrzeganym, dba o relacje z czytelnikami, na bieżąco komentuje różne sprawy i doniesienia. Witryna nieprzypadkowo jest przejrzysta, schludna, nowoczesna i aktualna.

## Archiwizacja witryn domowych i blogów

Zasoby WWW, w tym omówione powyżej witryny domowe i blogi, uznane za pozostawiony w cyberprzestrzeni ślad człowieka, który wymaga udokumentowania i zachowania (co ma już miejsce<sup>40</sup>), potrzebują rozwiązań to umożliwiających. Proces ten, a więc ich celowe gromadzenie i zabezpieczanie, nazwany jest archiwizacją Webu, zgodnie z rozumieniem terminu zaproponowanym przez Nielsa Brüggera<sup>41</sup>. Badacz wskazał na konkretne metody, które mogą być w tym celu zastosowane. Zaliczył do nich m.in. wykonywanie zrzutów i nagrań ekranu oraz pobierania stron WWW, a także metodę crawlingu oraz wykorzystywanie API udostępnianych przez media społecznościowe. Wymienione sposoby na zachowywanie zasobów sieciowych znacząco się między sobą różnią i generują odmienne jakościowo i ilościowo rezultaty, a w związku z tym, mogą być wykorzystywane w zależności od potrzeb badaczy lub archiwistów. Najbardziej rozpowszechnioną obecnie metodą, zwłaszcza pośród inicjatyw określanych jako archiwa Webu, jest gromadzenie z wykorzystaniem specjalnych robotów indeksujących nazywanych crawlerami<sup>42</sup>. Bazuje na niej wiele projektów, które – zależnie od swoich celów i możliwości – gromadzą różne fragmenty WWW<sup>43</sup>, w tym blogi, witryn domowe oraz pozostałe wytwory i ślady obecności i aktywności człowieka w Sieci.

Archiwizacja WWW z wykorzystaniem robotów indeksujących umożliwia prowadzenie tego zadania w sposób bardzo elastyczny. Crawling pozwala na gromadzenie pojedynczych stron (zawartość jednego adresu URL), witryn internetowych (zespół połączonych ze sobą stron), grupy witryn powiązanych tematycznie, a nawet całych domen krajowych najwyższego poziomu (ccTDL). Działanie robota polega na pobieraniu kodu źródłowego wskazanej strony oraz potrzeb-

<sup>40</sup> Zob. Charter on the Preservation of Digital Heritage, UNESCO 15.10.2003, [http://portal.unesco.org/en/ev.php-URL\\_ID=17721&URL\\_DO=DO\\_TOPIC&URL\\_SECTION=201.html](http://portal.unesco.org/en/ev.php-URL_ID=17721&URL_DO=DO_TOPIC&URL_SECTION=201.html) [dostęp: 20.08.2021].

<sup>41</sup> N. Brügger, *The Archived Web. Doing History in the Digital Age*, Cambridge MA 2018, s. 79–80.

<sup>42</sup> Ibidem, s. 81–82.

<sup>43</sup> Wikipedia. List of Web archiving initiatives, [https://en.wikipedia.org/wiki/List\\_of\\_Web\\_archiving\\_initiatives](https://en.wikipedia.org/wiki/List_of_Web_archiving_initiatives) [dostęp: 20.08.2021].

nych do jego działania plików, np. arkuszy stylu CSS czy osadzonych zdjęć lub grafik. Umożliwia to archiwizację nie tylko jej zawartości, ale też takich elementów jak hiperłącza prowadzące do innych obiektów w WWW. Ponadto możliwa jest archiwizacja wielu wersji jednej strony internetowej. Należy jednak podkreślić, iż wykonane w ten sposób kopie (snapshots) ulegają pewnym zmianom lub mogą być wykonane niepoprawnie. Dlatego też nie można traktować ich dokładnego odwzorowania jako tego, co znajdowało się kiedyś w Webie, a raczej jako nowe twory, które przywoływany wcześniej N. Brügger określa jako re-born digital<sup>44</sup>. Ponadto trzeba pamiętać, że ze względu na ograniczenia techniczne i prawne, a także kwestie merytoryczne, archiwa WWW nie są w stanie zachować każdej treści oraz wszystkich zmian w niej zachodzących.

Wpływ na zachowanie webowych źródeł, oprócz metody, posiada również zastosowana strategia. W kontekście działalności archiwów WWW najczęściej mówi się o dwóch podejściach: archiwizacji masowej oraz selektywnej, które mogą być również ze sobą łączone<sup>45</sup>. Największe nadzieje na zachowanie szeroko rozumianych egodokumentów rozproszonych w Webie (czyli takich jak wspomniane blogi i osobiste witryny domowe) wiązać można z gromadzeniem masowym. Jego zakres najczęściej obejmuje całą ccTDL lub, jak ma to miejsce w przypadku Internet Archive, całość dostępnych dla crawlerów zasobów WWW. Archiwizacja szeroka przeprowadzona jest przeważnie od jednego do czterech razy w roku, co pozwala zachować przekrojowy obraz Webu, obejmujący konkretny wycinek czasu, na który mogą składać się setki tysięcy stron internetowych, w tym właśnie witryny domowe i blogi.

Gromadzenie selektywne przynosi zdecydowanie mniej liczne zasoby i ma zdecydowanie bardziej dokumentacyjny charakter. Ze względu na swoje ograniczenia, wykorzystywane jest do archiwizowania źródeł powiązanych z konkretnym wydarzeniem (tzw. event harvesting), problemem<sup>46</sup> lub tematyką. Istotną cechą tej strategii jest dokonywanie przez archiwistę wyboru zasobów do zachowania

<sup>44</sup> N. Brügger, *When the present web is later the past. Web historiography, digital history and internet studies*, „Historical Social Research” 2012, vol. 37, no 4, DOI: 10.12759/hsr.37.2012.4.102-117, s. 107–109.

<sup>45</sup> F. Thouvenin, P. Hettich, H. Burkert, U. Gasser, 4 *Web Archives*, [w:] *Remembering and Forgetting in the Digital Age*, ed. F. Thouvenin, P. Hettich, H. Burkert, U. Gasser, Cham 2018, DOI: 10.1007/978-3-319-90230-2\_6, s. 94–96.

<sup>46</sup> Zob. UK Web Archive. Topics and Themes, <https://www.webarchive.org.uk/en/ukwa/collection> [dostęp: 20.08.2021]; Library of Congress. Web Archives, <https://www.loc.gov/programs/web-archiving/web-archives/> [dostęp: 20.08.2021].

na podstawie wcześniej obranych kryteriów, co obarcza ją znacznym subiektywizmem. Ważną rolę pełnią same kryteria, które zwracają uwagę na takie cechy jak popularność, reputacja autora lub znacząca wartość treści<sup>47</sup>. Mogą one sprawić, że w procesie selekcji zostaną pominięte egodokumenty dotyczące partykularnej codzienności, mniej popularne blogi czy portfolio nieznanych artystów.

W omawianym kontekście warto zwrócić uwagę także na inicjatywy o charakterze oddolnym i społecznym, których wrażliwość jest odmienna od ich zinstytucjonalizowanych odpowiedników. Różnica ta wyraża się m.in. w filozofii oraz działaniach podejmowanych np. przez niezależny kolektyw „łobuzerskich archiwistów” ArchiveTeam. Ich głównym obiektem zainteresowania jest zawartość serwisów i portali wytworzona przez ich własnych użytkowników<sup>48</sup>, co widoczne jest chociażby na przykładzie zabezpieczenia przez nich serwisu GeoCities<sup>49</sup> oraz Google+<sup>50</sup>. Inną, wartą odnotowania inicjatywą jest grupa Documenting the Now, która przygotowała narzędzia pozwalające na archiwizowanie zasobów z Twittera z pominięciem ograniczeń nałożonych na API oraz podejmująca kwestie ich etycznego wykorzystania<sup>51</sup>.

## Korzystanie z zarchiwizowanych witryn domowych i blogów

Niezwykle ważną kwestią, na którą chcemy zwrócić uwagę, jest dostęp do zgromadzonych już źródeł, a przez nie – do śladów człowieka pozostawionych w Sieci i możliwości prowadzenia badań egodokumentalnych w oparciu o te zasoby. Na to zagadnienie składają się zwłaszcza dwa czynniki: otwartość archiwum Webowego oraz metody przeszukiwania jego zasobu, czyli też pośrednio zasto-

<sup>47</sup> Na temat stosowanych kryteriów zob.: S. Schostag, E. Fønss-Jørgensen, *Web Archiving. Legal Deposit of Internet in Denmark. A Curatorial Perspective*, „Microform & Digitization Review” 2012, vol. 41, no 3–4, DOI: 10.1515/mir-2012-0018, s. 110–115; K. Holub, I. Rudomino, *A decade of web archiving in the National and University Library in Zagreb*, paper presented at IFLA WLIC 2015 – Cape Town, South Africa, s. 3–5, <http://library.ifla.org/id/eprint/1092> [dostęp: 23.08.2021]; Library Of Congress Collections Policy Statements Supplementary Guidelines. Web Archiving, Library of Congress 2017, <https://www.loc.gov/acq/devpol/webarchive.pdf> [dostęp: 23.08.2021].

<sup>48</sup> J. Scott, *Philosophy*, ArchiveTeam, <https://wiki.archiveteam.org/index.php/Philosophy> [dostęp: 20.08.2021].

<sup>49</sup> ArchiveTeam. GeoCities.

<sup>50</sup> ArchiveTeam. Google+, <https://wiki.archiveteam.org/index.php/Google%2B> [dostęp: 20.08.2021].

<sup>51</sup> Documenting the Now, <https://www.docnow.io/> [dostęp: 20.08.2021].



sowany opis zarchiwizowanych materiałów oraz przygotowanie badacza do tych poszukiwań. Wszystkie te czynniki znacząco wpływają na możliwość dotarcia do poszukiwanych informacji. Problem ten jest istotny, ponieważ pomimo zachowania dużej liczby materiałów pochodzących z WWW, sposób dostępu do nich znacząco wpływa na ich widoczność i późniejsze wykorzystanie.

Jako pierwsze należy poruszyć zagadnienie otwartości. Dostęp do zachowanej części dawnego Webu może być otwarty (np. dostępny dla wszystkich, najczęściej poprzez witrynę internetową), ograniczony (np. dostęp wyłącznie stacjonarny ze stanowisk na terenie biblioteki) lub zamknięty (zgromadzone zasoby nie są udostępnianie)<sup>52</sup>. Stopień otwartości najczęściej zależy od ograniczeń prawnych i może być różny dla poszczególnych części zasobów, czego przykładem jest choćby UK Web Archive<sup>53</sup>. Oczywiście z perspektywy badacza najbardziej korzystny jest swobodny dostęp do źródeł, jednak nie zawsze w tym modelu uwzględnione zostają kwestie etyczne, które mogą towarzyszyć wykorzystaniu śladów człowieka w Sieci<sup>54</sup>.

Druga kwestia to metody przeszukiwania i odnajdywania informacji (*information retrieval*) w zasobach zarchiwizowanego WWW. Możliwe jest ono za pomocą dwóch metod wyszukiwania informacji w Internecie: wyszukiwania (*searching*) i przeszukiwania (*browsing*). Podstawowa i najstarsza metoda – wyszukiwanie (*searching*) jest w tym przypadku oparta o adresy URL, co wymaga ich wcześniejszego ustalenia i pozyskania z zewnętrznych źródeł, także analogowych. Pokazują to choćby badania opublikowane przez Anat Ben-David<sup>55</sup> czy Marcina Wilkowskiego<sup>56</sup>. Obecnie metoda ta jest uzupełniana przez wyszukiwanie pełnotekstowe z opcją wykorzystania dodatkowych filtrów, zapewniających możliwość doprecyzowania uzyskanych wyników. Korzystanie z tego rozwiązania może przysparzać

<sup>52</sup> F. Thouvenin i in., *4 Web Archives*, s. 96–97.

<sup>53</sup> British Library. Collection guides. UK Web Archive, <https://www.bl.uk/collection-guides/uk-web-archive> [dostęp: 20.08.2021].

<sup>54</sup> Zob. J. Lin, I. Milligan, D.W. Oard, N. Ruest, K. Shilton, *We Could, but Should We? Ethical Considerations for Providing Access to GeoCities and Other Historical Digital Collections*, [w:] *CHIIR '20. Proceedings of the 2020 Conference on Human Information Interaction and Retrieval*, ed. H. O'Brien, L. Freund, New York 2020, DOI: 10.1145/3343413.3377980, s. 134–144.

<sup>55</sup> A. Ben-David, *What does the Web remember of its deleted past? An archival reconstruction of the former Yugoslav top-level domain*, „New Media & Society” 2016, vol. 18, no 7, DOI: 10.1177/1461444816643790, s. 1103–1119.

<sup>56</sup> M. Wilkowski, *Polish Web resources described in the “Polish World” directory (1997). Characteristics of domains and their conservation state*, „Archiwa – Kancelarie – Zbiory” 2020, nr 11(13), DOI: 10.12775/AKZ.2020.005, s. 119–140.

jednak trudności. Badaczka Alison Key, poszukując wspomnień jeńców wojennych, miała na przykład ogromne problemy ze stworzeniem wartościowego korpusu, pomimo stosowania rozbudowanych fraz wyszukiwawczych<sup>57</sup>.

Inną formą odnajdywania informacji w zarchiwizowanym WWW jest przeszukiwanie (*browsing*), które może odbywać się poprzez kilkustopniowe katalogi oraz kolekcje tematyczne. Ich zaletą jest ustrukturyzowanie treści oraz głębia informacyjna, jednak obejmują one wyłącznie zasoby zgromadzone selektywnie, które są mniej liczne i mogą być uboższe w teksty uznane za egodokumenty lub pod tym kątem badane. Warto ponadto zwrócić uwagę na fakt, iż w archiwach Webu często nie ma informacji o proveniencji zasobu lub innych jego cechach. Są one też trudno dostępne, przez co znacznie utrudnione jest lepsze ich zrozumienie.

Podsumowując tę część, należy stwierdzić, że korzystanie ze śladów człowieka zachowanych w dawnym WWW wymagać będzie od badacza nie tylko zrozumienia omówionej wcześniej natury tych źródeł oraz etycznego ich wykorzystania. Równie ważna będzie umiejętność ich zlokalizowania i zidentyfikowania w ogromie informacji, jaki stanowi zasób archiwów Webu. Konieczna w tym celu może być nie tylko nauka posługiwania się wyszukiwarką pełnotekstową, ale także bardziej zaawansowanymi narzędziami informatycznymi, np. specjalistycznymi językami programowania lub API. To samo będzie się odnosić do ich późniejszego badania, m.in. do tworzenia wizualizacji sieci łączących ich hiperłączy lub analizy korpusu tekstów. Ten wysoki próg wejścia nie sprawia jednak, że badania dawnych witryn domowych, blogów oraz innych Webowych egodokumentów jest nieosiągalne i niemożliwe z wykorzystaniem bardziej tradycyjnych rozwiązań.

Zakres pytań badawczych, jakie można postawić tego rodzaju źródłom, jest – o ile nie nieskończony – to na pewno niezwykle szeroki. Lepsze zrozumienie współczesnego człowieka będzie wymagało od badaczy skonfrontowania się z zasobami pochodzącymi z dawnego Webu. Internet oraz WWW stanowią dla niego ważną przestrzeń aktywności, a w związku z tym cyberprzestrzeń jest pełna pozostawionych przez niego śladów<sup>58</sup>, również tych egodokumentalnych. Trudne do przewidzenia jest to, jak uczeni, nie tylko badacze przeszłości, będą

<sup>57</sup> A.C. Kay, *Capture, commemoration and the citizen historian. Digital shoebox archives relating to PoWs in the Second World War*, [w:] *Web Archives as Big Data*, 03.12.2014, Institute of Historical Research (Unpublished), <https://sas-space.sas.ac.uk/6248/> [dostęp: 22.08.2021].

<sup>58</sup> N. Brügger, *The Archived Web*, s. 1–10; R. Schroeder, N. Brügger, *Introduction. The web as history*, [w:] *The Web as History. Using Web Archives to Understand the Past and the Present*, ed. R. Schroeder, N. Brügger, London 2017, DOI: 10.2307/j.ctt1mzt55k.6, s. 1–20.

w przyszłości korzystać z archiwów Webu. Wstępne obserwacje w tym zakresie są możliwe dzięki już opublikowanym badaniom. Zaliczyć można do nich choćby artykuł Marty Więckiewicz-Archackiej. Autorka omawia pierwsze polskie, niekomercyjne platformy blogowe oraz ich wpływ na powstanie lokalnej blogosfery. Zwraca uwagę na komunikację oraz relacje pomiędzy użytkownikami, a także na treść publikowanych postów, które często miały charakter intymny<sup>59</sup>. Natomiast Esther Weltevrede oraz Anne Helmond, w oparciu o zasoby Internet Archive, przeprowadziły zbliżone badanie, w którym analizowały rozwój niderlandzkiej blogosfery w latach 2001–2009. Ma ono jednak ilościowy oraz strukturalny charakter i opiera się o takie elementy jak domena najwyższego poziomu, hiperłącza oraz stosowane przez autorów blogów oprogramowanie<sup>60</sup>. Nawet jeśli przywołane przykłady badań nie mają *stricte* egodokumentalnego charakteru, mogą stanowić punkt wyjścia do dalszych rozważań oraz pogłębionych analiz. Badanie przeprowadzone przez Saskię Huc-Hepher pokazuje natomiast, że możliwe jest zejście w głąb zasobów, do mniejszej skali, i skupienie się na konkretnym zagadnieniu. Autorka przeprowadzała wielopoziomową analizę wybranych zawartości blogów francuskich londyńczyków, zwłaszcza ich elementów wizualnych. W powiązaniu z jej wcześniejszymi badaniami pozwoliło to dostrzec, jak wyglądała aklimatyzacja mieszkańców Londynu i jak zmieniał się ich stosunek do miasta<sup>61</sup>.

## Podsumowanie

Sieć jest pełna egodokumentów, i to nie tylko tych scharakteryzowanych w powyższych rozważaniach. Oprócz blogów i witryn domowych, znamiona egodokumentów widzimy zarówno w tak oczywistych źródłach jak media społecznościowe, jak i tych nieoczywistych – witrynach urzędowych. Za decyzjami twórców dotyczącymi choćby szaty graficznej, ucieczki od konwencji lub uleganiu jej, sposobu zarządzania treścią, często stoi chęć wyrażenia siebie. Tym samym

<sup>59</sup> M. Więckiewicz-Archacka, *Nlog, Prawda.org, Blog.art.pl. The First Non-profit Blogging Platforms in Poland and Their Influence on the Network of Blogs' Development*, „Studia Medioznawcze” 2019, t. 20, nr 1 (76), DOI: 10.33077/uw.24511617.ms.2019.1.88, s. 70–81.

<sup>60</sup> E. Weltevrede, A. Helmond, *Where do bloggers blog? Platform transitions within the historical Dutch blogosphere*, „First Monday” 2012, vol. 17, no 2, DOI: 10.5210/fm.v17i2.3775.

<sup>61</sup> S. Huc-Hepher, *Searching for Home in the Historic Web. An Ethnosemiotic Study of London-French Habitus as Displayed in Blogs*, [w:] *Web Archives as Big Data*, 03.12.2014, Institute of Historical Research (Unpublished), <https://sas-space.sas.ac.uk/6252/> [dostęp: 22.08.2021].

materiały, które są w Sieci, bardzo często przedstawiają umysłowość i doświadczenia pojedynczych osób, ale też grupy, pokolenia, narodu, kultury. Możliwość egodokumentalnego odczytywania oraz potencjał badawczy tych zasobów pokazują przytoczone w materiale studia przypadku. Pozwalają one nie tylko prześledzić czyjąś karierę artystyczną czy pracę badawczą, ale także ukazać upodobania tej osoby, odpowiadając na pytania: z jakich gotowych narzędzi korzystała, i czy w ogóle; jak często aktualizowała swoją witrynę; jaką kolorystykę, układ graficzny, animacje lub inne dodatki uznała za odpowiednie; czy próbowała się w jakiś sposób kreować, podkreślać swoje dokonania, czy raczej je kamuflować lub ukrywać, itp.

Niezwykle ważnym zadaniem jest dążenie do zachowania zasobów sieci dla kolejnych pokoleń, w tym badaczy, którzy w oparciu o te źródła będą mogli odkrywać przeszłość i poznawać przejawy życia jej uczestników. Sieć jako całość jest obrazem kultury umysłowej naszej epoki. Jej poszczególne części są natomiast śladami pozostawionymi przez różne, partykularne osoby i grupy. Słusznie podejmowane są działania dążące do zachowania jak największej części tych zasobów i umożliwienia korzystania z nich. Wykorzystywane obecnie metody i narzędzia pozwalają na skuteczną archiwizację zasobów Webu, w tym licznie występujących w nich egodokumentów. Dzięki temu możliwe jest historyczne spojrzenie na te efemeryczne i zmienne źródła. Istotne jest m.in. zapisywanie ich różnych wersji, co pozwoli przyszłym badaczom na ich porównywanie i śledzenie wprowadzanych zmian. Nie są to jednak, co staraliśmy się pokazać, zadania łatwe – z uwagi na wiele ograniczeń prawnych, technicznych i innych. Z perspektywy potencjalnego badacza, jednym z istotniejszych w tym zakresie problemów jest dostęp i korzystanie ze zarchiwizowanych zasobów. Wymagać ono może odpowiednio przygotowanego warsztatu, w tym zaadaptowania bardziej zaawansowanych metod cyfrowych, które umożliwią właściwą analizę źródeł. Jednak – jak pokazują przytoczone w badaniach przykłady – już teraz możliwe jest wykorzystanie potencjału, który posiadają zachowane przez archiwa Webowe ślady człowieka.

## Bibliografia

- Ben-David A., *What does the Web remember of its deleted past? An archival reconstruction of the former Yugoslav top-level domain*, „New Media & Society” 2016, vol. 18, no 7, DOI: 10.1177/1461444816643790, s. 1103–1119.

- Bolter J.D., *Przestrzeń pisma. Komputery, hipertekst i remediacja druku*, tłum. Aleksandra Małecka, Michał Tabaczyński, Kraków–Bydgoszcz 2014.
- British Library. Collection guides. UK Web Archive, <https://www.bl.uk/collection-guides/uk-web-archive> [dostęp: 20.08.2021].
- Brügger N., *The Archived Web. Doing History in the Digital Age*, Cambridge MA 2018.
- Brügger N., *When the present web is later the past. Web historiography, digital history and internet studies*, "Historical Social Research" 2012, vol. 37, no 4, DOI: 10.12759/hsr.37.2012.4.102-117, s. 107–109.
- Charter on the Preservation of Digital Heritage, UNESCO 15.10.2003, [http://portal.unesco.org/en/ev.php-URL\\_ID=17721&URL\\_DO=DO\\_TOPIC&URL\\_SECTION=201.html](http://portal.unesco.org/en/ev.php-URL_ID=17721&URL_DO=DO_TOPIC&URL_SECTION=201.html) [dostęp: 20.08.2021].
- Chorążyczewski W., Rosa A., *Egodokumenty – egodokumentalność – analiza egodokumentalna – spuścizna egodokumentalna*, [w:] *Egodokumenty, tradycje historiograficzne i perspektywy badawcze*, red. W. Chorążyczewski, A. Pacevicius, S. Roszak, Toruń 2015, s. 11–22.
- Dekker R.M., *Ego-Documents in the Netherlands 1500-1814*, „Dutch Crossing” 1989, vol. 39, s. 61–72.
- Documenting the Now, <https://www.docnow.io/> [dostęp: 20.08.2021].
- Dominick J.R., *Who Do You Think You Are? Personal Home Pages and Self-Presentation on the World Wide Web*, „Journalism & Mass Communication Quarterly” 1999, vol. 76, no 4, DOI: 10.1177/107769909907600403, s. 646–658.
- Dopierała R., *Internetowe manifestacje prywatności*, „Media–Kultura–Społeczeństwo” 2008, nr 1(3), s. 105–127.
- Döring N., *Personal Home Pages on the Web. A Review of Research*, „Journal of Computer-Mediated Communication” 2002, vol. 7, no 3, DOI: 10.1111/j.1083-6101.2002.tb00152.x.
- Dragan Espenschied, <http://1x-upon.com/~despens/> [dostęp: 22.09.2023].
- Egodokumenty, tradycje historiograficzne i perspektywy badawcze*, red. W. Chorążyczewski, A. Pacevicius, S. Roszak, Toruń 2015.
- Erickson T., *The World-Wide-Web as social hypertext*, „Communications of the ACM” 1996, vol. 39, no 1, DOI: 10.1145/234173.234174, s. 15–17.
- GeoCities*, ArchiveTeam, <https://wiki.archiveteam.org/index.php/GeoCities> [dostęp: 20.08.2021].
- Google+*, ArchiveTeam, <https://wiki.archiveteam.org/index.php/Google%2B> [dostęp: 20.08.2021].
- Hockx-Yu H., *Archiving Social Media in the Context of Non-print Legal Deposit*, presented at IFLA WLIC 2014 in Lyon, France, 16–22 August 2014, <http://library.ifla.org/id/eprint/999> [dostęp: 6.06.2021].
- Hollenbaugh E.E., *Motives for Maintaining Personal Journal Blogs*, „Cyberpsychology, Behavior, and Social Networking” 2011, vol. 14, no 1–2, DOI: 10.1089/cyber.2009.0403, s. 13–20.

- Holub K., Rudomino I., *A decade of web archiving in the National and University Library in Zagreb*, paper presented at IFLA WLIC 2015 – Cape Town, South Africa, <http://library.ifla.org/id/eprint/1092> [dostęp: 23.08.2021].
- Huc-Hepher S., *Searching for Home in the Historic Web. An Ethnosemiotic Study of London-French Habitus as Displayed in Blogs*, [w:] *Web Archives as Big Data*, 03.12.2014, Institute of Historical Research (Unpublished), <https://sas-space.sas.ac.uk/6252/> [dostęp: 22.08.2021].
- Huffaker D., *Teen Blogs Exposed: The Private Lives of Teens Made Public*, Presented at the American Association for the Advancement of Science (AAAS) in St. Louis, MO., February 16–19, 2006, <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.556.1504&rep=rep1&type=pdf> [dostęp: 6.06.2021].
- Ibrahim Y., *Weblogs as Personal Narratives. Displacing History and Temporality*, „M/C Journal” 2006, vol. 9, no 6, DOI:10.5204/mcj.2690.
- International Internet Preservation Consortium. Web Archiving, <https://netpreserve.org/web-archiving/> [dostęp: 20.09.2023].
- Jones M., Alony I., *Blogs – The New Source of Data Analysis*, „Issues in Informing Science and Information Technology” 2008, vol. 5, DOI: 10.28945/1019, s. 433–446.
- Kasprzak W.A., *Ślady cyfrowe. Studium prawnokryminalistyczne*, Warszawa 2015.
- Kay A.C., *Capture, commemoration and the citizen historian. Digital shoebox archives relating to PoWs in the Second World War*, [w:] *Web Archives as Big Data*, 03.12.2014, Institute of Historical Research (Unpublished), <https://sas-space.sas.ac.uk/6248/> [dostęp: 22.08.2021].
- Konopa B., *Archiwizacja Webu w Europie – narodowe archiwa Sieci*, „Archeion” 2021, t. 121, DOI: 10.4467/26581264ARC.20.016.12973, s. 445–465.
- Konopa B., *Periodyzacja archiwizacji Webu*, <https://webarch.uw.edu.pl/zasoby/periodyzacja-archiwizacji-webu/> [dostęp: 28.08.2021].
- Konopa B., *Strategia selektywna jako narzędzie w archiwizacji Webu. Analiza wybranych przykładów*, „Archiwa – Kancelarie – Zbiory” 2020, nr 11(13), s. 97–118.
- Kosinski M., Stillwell D.J., Graepel T., *Private traits and attributes are predictable from digital records of human behavior*, „Proceedings of the National Academy of Sciences of the United States of America” 2013, vol. 110(15), DOI: 10.1073/pnas.1218772110, s. 5802–5805.
- Lent Jonathan, *Journaling Enters the 21<sup>st</sup> Century. The Use of Therapeutic Blogs in Counseling*, „Journal of Creativity in Mental Health” 2009, no 4, s. 68–73.
- Library Of Congress Collections Policy Statements Supplementary Guidelines. Web Archiving, Library of Congress 2017, <https://www.loc.gov/acq/devpol/webarchive.pdf> [dostęp: 23.08.2021].
- Library of Congress. Legal Blawgs Web Archive, <https://www.loc.gov/collections/legal-blawgs-web-archive/about-this-collection/> [dostęp: 28.05.2021].

- Lin J., Milligan I., Oard D.W., Ruest N., Shilton K., *We Could, but Should We? Ethical Considerations for Providing Access to GeoCities and Other Historical Digital Collections*, [w:] CHIR '20. *Proceedings of the 2020 Conference on Human Information Interaction and Retrieval*, ed. H. O'Brien, L. Freund, New York 2020, DOI: 10.1145/3343413.3377980, s. 134–144.
- Marshall C.C., Shipman M., *On the Institutional Archiving of Social Media*, [w:] JCDL '12: *Proceedings of the 12<sup>th</sup> ACM/IEEE-CS joint conference on Digital Libraries*, New York 2012, DOI: 10.1145/2232817.2232819, s. 1–10.
- Milligan I., *History in the Age of Abundance? How the web is transforming historical research*, Montreal 2021, s. 106–142.
- Milligan I., *Welcome to the web. The online community of GeoCities during the early years of the World Wide Web*, [w:] *The Web as History. Using Web Archives to Understand the Past and the Present*, ed. N. Brügger, R. Schroeder, UCL Press 2017, DOI: 10.2307/j.ctt1mtz55k.13, s. 137–158.
- Olcoń M., *Blog jako dokument osobisty. Specyfika dziennika prowadzonego w Internecie*, „Kultura i Społeczeństwo” 2003, t. 47, nr 2, s. 123–143.
- Rosa A., *Egodokumentalność polskiego zasobu archiwalnego*, [w:] *Archiwistyka między różnorodnością a standaryzacją*, Toruńskie Konfrontacje Archiwalne, t. 3, red. W. Chorążyczewski, A. Rosa, Toruń 2013, s. 103–112.
- Rosa A., *Egodokumenty a działalność archiwów*, [w:] *Zatrzymać przeszłość, dogonić przyszłość. Pamiętnik VI Powszechnego Zjazdu Archiwistów Polskich*, Wrocław, 5–7 września 2012 r., red. W. Chorążyczewski, K. Strykowski, Toruń 2013, s. 271–275.
- Rosa A., *Human trace on the Internet – the issue of archiving the Web from the point of view of anthropology-oriented archival science*, „Archiwa – Kancelarie – Zbiory” 2015, nr 6(8), DOI: 10.12775/AKZ.2015.006, s. 193–205.
- Rosa A., Siuda K., *Analiza egodokumentalna pamiętników archiwistów jako przyczynek do poznania mentalności określonej grupy zawodowej*, [w:] *Pogranicza archiwistyki*, Toruńskie Konfrontacje Archiwalne, t. 6, red. W. Chorążyczewski, A. Rosa, Toruń 2019, s. 85–100.
- Rozsak S., *Archiwa sarmackiej pamięci. Funkcje i znaczenie rękopiśmiennych ksiąg typu silva rerum w kulturze Rzeczypospolitej XVIII wieku*, Toruń 2004.
- Rozsak S., *Egodokumenty – kilka uwag o polskich i europejskich doświadczeniach historiograficznych oraz metodologicznych*, „Biuletyn Polskiej Misji Historycznej” 2013, nr 8, DOI: 10.12775/BPMH.2013.001, s. 27–42.
- Schostag S., Fønss-Jørgensen E., *Webarchiving. Legal Deposit of Internet in Denmark. A Curatorial Perspective*, „Microform & Digitization Review” 2012, vol. 41, no 3–4, DOI: 10.1515/mir-2012-0018, s. 110–120.

- Schroeder R., Brügger N., *Introduction. The web as history*, [w:] *The Web as History. Using Web Archives to Understand the Past and the Present*, ed. R. Schroeder, N. Brügger, London 2017, DOI: 10.2307/j.ctt1mtz55k.6, s. 1–20.
- Schulze W., *Ego-Dokumente. Annäherung an den Menschen in der Geschichte? Vorüberlegungen für die Tagung „Ego-Dokumente“*, [w:] *Ego-Dokumente. Annäherung an den Menschen in der Geschichte*, ed. W. Schulze, Berlin 1996, s. 11–30.
- Scott J., *Philosophy*, ArchiveTeam, <https://wiki.archiveteam.org/index.php/Philosophy> [dostęp: 20.08.2021].
- Thouvenin F., Hettich P., Burkert H., Gasser U., *4 Web Archives*, [w:] *Remembering and Forgetting in the Digital Age*, ed. F. Thouvenin, P. Hettich, H. Burkert, U. Gasser, Cham 2018, DOI: 10.1007/978-3-319-90230-2\_6, s. 84–101.
- UK Web Archive. Blogs, <https://www.webarchive.org.uk/en/ukwa/collection/44>, [dostęp: 28.05.2021].
- UK Web Archive. Topics and Themes, <https://www.webarchive.org.uk/en/ukwa/collection> [dostęp: 20.08.2021].
- Warley L., *Reading the autobiographical in personal home pages*, [w:] *Tracing the autobiographical*, ed. M. Kadar, S. Egan, L. Warley, J. Perreault, Ontario 2005, s. 25–42.
- Webstory. Peter Webster's blog, <https://peterwebster.me/> [dostęp: 27.10.2023].
- Weltevrede E., Helmond A., *Where do bloggers blog? Platform transitions within the historical Dutch blogosphere*, „First Monday” 2012, vol. 17, no 2, DOI: 10.5210/fm.v17i2.3775.
- Więckiewicz-Archacka M., *Nlog, Prawda.org, Blog.art.pl The First Non-profit Blogging Platforms in Poland and Their Influence on the Network of Blogs' Development*, „Studia Medioznawcze” 2019, t. 20, nr 1(76), DOI: 10.33077/uw.24511617.ms.2019.1.88, s. 70–81.
- Wikipedia. List of Web archiving initiatives, [https://en.wikipedia.org/wiki/List\\_of\\_Web\\_archiving\\_initiatives](https://en.wikipedia.org/wiki/List_of_Web_archiving_initiatives) [dostęp: 20.08.2021].
- Wilkowski M., *Polish Web resources described in the “Polish World” directory (1997). Characteristics of domains and their conservation state*, „Archiwa – Kancelarie – Zbiory” 2020, nr 11(13), DOI: 10.12775/AKZ.2020.005, s. 119–140.
- Wilkowski M., *Strona internetowa jako źródło historyczne*, „Historyka. Studia Metodologiczne” 2021, t. 51, DOI: 10.24425/hsm.2021.138366, s. 83–110.
- Żytomirski M., *Określanie osobowości użytkowników Internetu poprzez analizowanie ich cyfrowych śladów w świetle wybranych badań dr. Michała Kosińskiego*, „Acta Universitatis Lodziensis Folia Librorum” 2021, t. 1(32), DOI: 10.18778/0860-7435.32.05, s. 99–113.
- Żytomirski M., Żulicki R., *Cyfrowe ślady a banki filtrujące – prezentacja informacyjna*, 15 August 2020, [https://www.researchgate.net/publication/343678108\\_Cyfrowe\\_slady\\_a\\_banki\\_filtrujace\\_-\\_prezentacja\\_informacyjna](https://www.researchgate.net/publication/343678108_Cyfrowe_slady_a_banki_filtrujace_-_prezentacja_informacyjna) [dostęp: 6.08.2023].