

AGNIESZKA ROSA

Uniwersytet Mikołaja Kopernika

**HUMAN TRACE ON THE INTERNET – THE ISSUE
OF ARCHIVING THE WEB FROM THE POINT OF VIEW
OF ANTHROPOLOGY-ORIENTED ARCHIVAL SCIENCE**

<http://dx.doi.org/10.12775/AKZ.2015.006>

Słowa kluczowe

archiwizacja Internetu; strategie archiwizacyjne; antropologizowanie archiwistyki; Fundacja Internet Archive; zasoby cyfrowe;

Keywords

archiving the Internet; archiving strategies; anthropology-oriented archival science; the Internet Archive Foundation; digital holdings

Streszczenie

W artykule podjęto próbę odpowiedzi na pytanie dotyczące obecności i wytworów człowieka w sieci Internet z punktu widzenia archiwistyki odpowiadającej na wyzwania współczesnej humanistyki. Przyjrzano się działaniom podejmowanym przez człowieka w Internecie w kontekście oferowanych możliwości. Zwrócono także uwagę na obecnie stosowane sposoby archiwizacji zasobów Internetu, co pozwoliło na wyciągnięcie ogólnych wniosków na temat czynników, które najczęściej decydują o archiwizacji zasobów Internetu. Podejście do archiwizacji Internetu od strony archiwistyki zantropologizowanej wymaga zastosowania odpowiednich strategii archiwiz-

acyjnych. Podstawą w budowaniu tych strategii powinna z kolei być analiza działań i wytworów ludzkich, zwrócenie uwagi na aktywności człowieka w Internecie i jego wytwory – w sieci zamieszczane i w sieci powstające.

In the article I make an attempt to answer the question about human presence and products on the Internet from the point of view of archival science meeting the challenge of contemporary humanities. I examined actions taken by humans on the Web in the context of all offered possibilities. I also paid attention to currently used ways of archiving the Internet; thanks to that I managed to draw some general conclusions concerning factors determining archiving Web sources. Approaching the issue of archiving the Internet from the point of view of anthropology-oriented archival science requires using proper archiving strategies. A base for building those strategies should be analysis of human actions and products, paying attention to human activity on the Web and its fruits – put on the Internet and created on the Internet.

On the one hand, the Internet is a network of computers connected with each other; on the other, there are also resources accessible through the network, put on the Web and created on the Web. The Internet exists thanks to people and for them; it enables communication, publishing various content, providing access to services, it allows information flow, keeping in contact, but also creating ones identity. It makes learning easier and offers many forms of entertainment; it enables on-line shopping and financial transfers, as well as getting to know new people; it also can influence people's behavior and their choices. It is a huge gathering of information from practically every existing topic, all parts human life; it is a part of human life and people actively influence the shape of the Internet – by creating or using various websites, applications or services.

While building strategy for archiving the Web sources one has to take into consideration the human presence on the Internet, consisting both actions and products. After all, the overriding goal of archiving is preserving human activity in particular place and time, through people's products. Following types of human activity on the Internet can be mentioned: learning (life-long learning, professional training, parallel learning), cultivating one's interest (pursuing hobby, indulging passion), entertainment (playing, meeting people) and work. I intentionally pass over dividing people into those creating and those using resources (what is more, the division might not be so

simple due to Internet users participation in creating its content). But one has to consider the fact, if people act officially or unofficially, and what is one's role – is he/she a private person, an institution/enterprise employee, or a representative of a social group. Also phenomena opposing, or coming across, these divisions must be mentioned: blogosphere,¹ social networks, e-mail, chatsand other communicators, as well as the phenomena of hate,² trolling,³ or even the TOR.⁴

¹ Blog, blogger, blogosphere – terms used in everyday language to describe a phenomenon of creating web logs. Blogs are a specific form of publication – they are websites containing separate, usually chronologically arranged entries that can be commented by users. These entries contain personal thoughts and comments, but also drawings, audio and video recordings that present author's world view. At present we recognize text blogs (culinary, fashion, legal, feature, scientific and others), but also video blogs and photoblogs on various subjects. These sites enable categorizing, tagging and archiving posts. Blogs' authors, bloggers, created a specific culture and a social phenomenon called blogosphere (the world of blogs). It is a network of blogs operating as an entity, because authors usually follow other blogs, create links to them and contact with their authors. Look: M. Jeleński, *Definicja i rodzaje blogów*, e-Redaktor.pl, March 25, 2008, <http://eredaktor.pl/teoria/definicja-i-rodzaje-blogow/> [accessed: June 18, 2015].

² Hate, hater, to hate – Internet hating consists in criticism, anger, aggression, negative views, in which a hater (usually an anonymous user of the Web) finds pleasure; a specific Internet hate speech in the form of words or images (statements, posts and tweets, but also films and graphic images) that usually refers to politicians, celebrities and journalists, but also particular institutions, events, processes, phenomena, inconvenient laws, and is aimed at defamation, offense, abuse, humiliation. Hate can be seen the most frequently on discussion forums and beneath entries on feature and tabloid portals. A research found out, that hate entries under Internet publications have more considerable effect on reception of these publications, that their content itself. The same study by American researchers shows, that most hated are publications about politics. Look: P. Stasiak, *Naukowcy zbadali hejt internetowy. Co piąty komentarz jest obelżywy, zaś trolle...*, Newsweek Polska, August 8, 2014, <http://nauka.newsweek.pl/naukowcy-zbadali-hejt-internetowy-co-piaty-komentarz-jest-obelzywy,-artykuly,345268,1.html> [accessed: June 17, 2015]. According to feature writers to journal "Newsweek Polska" Bronisław Komorowski probably would not have lost presidential elections, if not hate in on-line statements, posts <http://polka.newsweek.pl/tomasz-lis-o-nienawisci-w-sieci-hejterzy-wyniki-wyborow,artykuly,364605,1.html> [accessed: June 17, 2015], and tweets; T. Lis, Komorowski prawdopodobnie nie przegrałby wyborów, gdyby nie „hejt industry”, June 8, 2015, <http://polka.newsweek.pl/tomasz-lis-o-nienawisci-w-sieci-hejterzy-wyniki-wyborow,artykuly,364605,1.html> [accessed: June 17, 2015].

³ Troll, trolling – provoking, intentional actions consisting in on-line publishing of hostile, offensive, controversial or untrue information in order to provoke arguments, annoy, irritate, insult, frighten, mislead, or lead to content that is unlawful, offensive,

Products of human activity on the Internet, which are network resources (apart from the infrastructure – also built by humans), can be divided into two categories: resources published on the Internet and resources created on the Internet. Those first have their equivalent outside the Web, often they are copies of digitized or original digital texts/documents, photographs, films. Materials created on the Internet usually exist on-line exclusively, their producers often do not have full control on them, especially in case of sources based on outside applications and, services;⁵ not always can they print these resources, and only selected applications and services offer data export (e.g.

false or undesirable. A troll intentionally misleads and irritates other Internet users; he/she usually publishes entries on various forums, discussion groups, chats; a troll succeeds, when is able to mislead, offend or expose other users, change subject of discussion, settle arguments between users or cause closing a conversationthread. Internet users are familiar with a phrase “Don’t feed the trolls” – when nobody reacts to troll’s comments, a troll becomes discouraged and leaves. Look: D. Jachyra, *Trollowanie – antyspołeczne zachowania w Internecie, sposoby wykrywania i obrony*, Zeszyty Naukowe Uniwersytetu Szczecińskiego. Studia Informatica, no. 28, 2011, p. 253–261, http://wneiz.pl/nauka_wneiz/studia_inf/28-2011/si-28-253.pdf [accessed: June 17, 2015]; M. Wawrzyn, *Psychologia trolla. Czemu niektórzy internauci tak lubią wkurzać innych*, <http://gadzetomania.pl/3968,psychologia-trolla-czemu-niektorzy-internauci-tak-lubia-wkurzac-innych> [accessed: June 17, 2015].

⁴ Tor, The Onion Router – the anonymity network popularized in 2004 that runs simultaneously to the Internet, using specific software and onion routing (which provide multiple messages coding and sending it through numerous nodes called onion routers); it allows anonymous sharing of content and services, as well as anonymous access to the Internet resources; thanks to that users can still use the same websites anonymously, but also gain access to materials available only to the Tor users. The Tor network hides location of its user, but also enables access to hidden services. It is slower than the regular Internet because our IP address goes through several servers while connecting to Tor, but thanks to that it is more difficult to track a user. Look: B. Świdzki, *Najciemniejszy zakątek internetu naprawdę istnieje. Ukryta sieć TOR: „Lewe” papiery, pedofilia, przekręty i narkotyki*, Na temat.pl, <http://natemat.pl/32267,najciemniejszy-zakatek-internetu-naprawde-istnieje-ukryta-siec-tor-lewe-papiery-pedofilia-przekrety-i-narkotyki> [accessed: June 18, 2015].

⁵ For example the Endomondo application for joggers and sportspersons collects information on sport activity of its users, but also about routes, weight, height, burnt calories and others. About the Internet and its resources also in: 20 things I learned about browsers and the Web, The Google Chrome Team, November, 2010, <http://www.20thingsilearned.com/en-US/home> [accessed: June 21, 2015].

some social networks and e-mail servers administrators).⁶ These materials are ephemeral and changeable.

How much does the Internet influence the way people's actions are recorded? It can be seen by examining common Internet applications like Gmail, Facebook, YouTube, Twitter – software that uses Internet search engines to access our data stored on the Web. Thanks to these applications no matter where we are (unlike in case of traditional desktop applications that store data on hard drives of our computers) we can access our data, read our messages and answer them, watch and comment photos and films, talk to our friends and do some everyday activities. It is frequent that we do not think about saving our data or consolidating it outside the Web, and who can access our data, apart from us, and when.

Internet resources can be categorized according to the place and way they were created primarily – into those that came into existence outside of the Web (and they are stored physically or on digital carriers, and the Internet contains only their copies) and those materials/resources that were created on the Web (they exist here as an original version, and their prints, hard drive or cloud records are just copies – and it is not always possible to create them). Websites of public offices, institutions and enterprises publish materials that in most cases are also available outside these websites (documents, photographs, films and others), but the structure and visualization of these sites is present only on the Internet. Materials published on scientific portals (also Internet publications), hobby portals and various types of blogs (connected in a blogosphere) not always have their place outside the Web. Only on-line exists their inseparable element – users' comments and interactive links/citations/references. It also may be addressed to resources from informational portals and services. Internet forums, discussion groups with their

⁶ A future personal archive of a researcher who uses the Internet might consist of not only paper and electronic documentation, but also his/her "web products". It may contain personal website, scientific or other topically based blog (with relation to other blogs), social networks profiles (like Google Scholar or Academia.edu) and works, opinions or reviews published there, as well as actions taken to ease access to his/her works by creating Internet bibliography and publishing these works on-line in various types of repositories, but also activity on social and hobby networks and virtual mail. Some part, if not most, of these data will be available also outside of the Web, but some part will be accessible only on-line. Surely, omitting this important sphere of human activity may limit the whole picture of scientific activity.

structures, topics, posts and reactions to them, as well as social networks and their users' profiles and other tools of entertainment also exist on-line exclusively. It also concerns electronic mail with whole systems of contacts and interactions between senders and addressees.

How are the Internet resources archived at present? Observations of the Web and examination of literature concerning the issue of archiving⁷ enable listing several strategies of archiving the Internet, according to who is the subject archiving:

- 1) Archiving made by national/government institutions (especially national libraries – for example cases of the Czech Republic⁸ and the United Kingdom⁹); algorithm: resources in particular language, particular domain, on particular topic; archiving made automatically, based on existing algorithms evaluating value of materials; archiving aided by human (for example possibility to nominate sites (that have particular background) to be archived in the UK Web Archive¹⁰);
- 2) Archiving made by global non-profit organizations, that store materials generally (international consortia like the Internet Archive www.archive.org); algorithm: resources available and visible on-line, frequ-

⁷ Look: F. Kłębczyk, *Archiwizacja zasobów Internetu – kierunki i wyzwania*, Archiwista Polski 2012, no. 3, p. 105–112; A. Sobczak, *Internet jako globalne archiwum społeczne – rozważania na temat roli Internetu w dokumentowaniu dziejów ludzkości*, [in:] Toruńskie Konfrontacje Archiwalne, vol. 4, *Nowa archiwistyka – archiwa i archiwistyka w późnonowoczesnym kontekście kulturowym*, ed. W. Chorążyczewski, W. Piasek, A. Rosa, Toruń 2014, p. 237–247; M. Wilkowski, *Od osobistej archiwistyki cyfrowej do edukacji medialnej*, EBiB, no. 6(151), 2014, <http://open.ebib.pl/ojs/index.php/ebib/article/view/274/436> [accessed June 21, 2015]; *Nowe media = nowa partycypacja*, ed. P. Celiński, Lublin 2014, p. 12–21, http://kulturacyfrowa.org/wp-content/uploads/2014/12/IKC_book_v11.pdf [accessed: June 21, 2015]; M. Halawa, *Nowe media i archiwizacja życia codziennego*, Kultura współczesna, no. 4, 2011, Powrót do archiwów, p. 27–41, <http://kulturawspolczesna.pl/readpdf/1271/Nowe%20media%20i%20archiwizacja%20%C5%BCycia%20codziennego> [accessed: June 21, 2015].

⁸ Webarchiv.cz, <http://www.webarchiv.cz/cs>; also: *Czeski projekt archiwizacji Internetu Presentation Transcript*, January 9, 2010, <http://www.slideshare.net/p3l3k/czeski-projekt-archiwizacji-internetu> [accessed: March 19, 2014].

⁹ UK Web Archive, <http://www.webarchive.org.uk/ukwa/>; look also: M. Wilkowski, *Archiwum brytyjskiego WWW: UK Web Archive*, November 9, 2011, Historia Media, <http://historiaimmedia.org/2011/11/09/archiwum-brytyjskiego-www-uk-web-archive/> [accessed: June 24, 2015].

¹⁰ Nominate a Site, <http://www.webarchive.org.uk/ukwa/info/nominate> [accessed: June 22, 2015].

- ency of changes and visits; archiving made automatically, based on existing algorithms evaluating value of materials;
- 3) Archiving made by local non-profit organizations, that store materials on specific topics (for example Internet Memory Foundation/former European Archive, www.europarchive.org, www.internetmemory.org); algorithm: resources in particular language, particular domain, on particular topic; archiving made automatically, based on existing algorithms evaluating value of materials;
 - 4) Archiving made by non-profit organizations/associations/institutions, that store materials on specific topics (for example the CyberCemetery:¹¹ an archive of inactive governmental websites in the USA); algorithm: resources on particular topic, selected automatically but also by human beings;
 - 5) Archiving made by Internet users for their own needs; groups/associations/institutions; algorithm: interests and current needs of users; archiving according to particular needs using available software; one archives what he/she wants for his/her own (for example using the Archive-It service by ReedTech¹² that through a special application allows archiving and accessing resources stored on data centers of the Internet Archive, and the Httrack Website Copier¹³ – an independent software for self-reliant web-archiving that allows to download whole content of World Wide Web pages on a local directory and use them later offline);
 - 6) Auto-archiving made by portals/services (social networks, hobby and information portals, on-line shops, blogs and other) focused on their own resources (posts, products, services); algorithm: archiving resources of particular portal/service; archiving made automatically, based on existing algorithms saving resources of the portal/service; data also archived for commercial use – like preparing offers of services or products;

Thus, following factors (and at the same time – qualities of network resources) decide about scope and frequency of archiving: availability – sites

¹¹ CyberCemetery, <http://govinfo.library.unt.edu/> [accessed: June 22, 2015].

¹² Website and Social Media Archiving, <http://www.reedtech.com/products-services/website-and-social-media-archiving/> [accessed: June 22, 2015].

¹³ Httrack Website Copier, <https://www.httrack.com/> [accessed November28, 2015].

indexed and accessible by the Web are archived (the Deep Web or the TOR network are not preserved, as well as sites not allowed to be indexed/archived by their creators/owners); also digital and multimedia sources put on-line under free licenses are archived (films, recordings, books, publications, scans of photographs and documents); frequency of changes; number of visits (interest in a site); location/topic – archiving is determined by site's address in particular domain/prefix, as well as by language of the site and its content, and if it relates to particular country, region, nation, events.¹⁴ On the background lies a factor that might determine seeing web resources as archival materials – their content, also as regards their originality/value, if they are primary or secondary sources.

¹⁴ Also articles in Biuletyn EBiB, no. 1(128), 2012, http://www.ebib.pl/?page_id=380 [accessed: June 22, 2015]; also: T. Baliszewski, *Sieć dla potomnych. Biblioteka Brytyjska archiwizuje internet*, NaTemat.pl, February 2013, <http://natemat.pl/57431,biblioteka-brytyjska-archiwizuje-internet> [accessed: March 19, 2014]; G. Gmiterek, *Archiwum internetowe – czy możliwa jest archiwizacja zasobów sieci?*, August 25, 2010, <http://www.biblioteki.org/pl/wiadomosci/czytaj/786> [accessed: March 19, 2014]; K. Kopańko, *Archiwizacja Internetu dla przyszłych pokoleń*, April 9, 2013, Antyweb.pl, <http://antyweb.pl/archiwizacja-internetu-dla-przyszlych-pokolen/> [accessed March 19, 2014]; P. Michalak, K. Lange, *Archiwizacja stron internetowych szkół wyższych, instytucji naukowych i kulturalnych*, <http://archiwum.amu.edu.pl/publikacje/17.pdf> [accessed: June 22, 2015]; J. Pasztaleniec-Jarzyńska, *Polityczna historia archiwizacji sieci: Międzynarodowe Konsorcjum Archiwizacji Internetu*, June 13, 2012, 14:35, [in:] *Babin Bibliografia Analytyczna Bibliotekoznawstwa i Informacji Naukowej*, <http://babin.bn.org.pl/?p=1158> [accessed: March 19, 2014]; M. Wilkowski, *Internetowe archiwum 10 kwietnia*, April 20, 2010, Historia i Media, <http://historiaimedia.org/2010/04/20/internetowe-archiwum-10-kwietnia/> [accessed: March 19, 2014]; M. Wilkowski, *Web Archives: The Future(s) – trzy scenariusze rozwoju archiwistyki internetowej*, August 10, 2011, Historia i Media, <http://historiaimedia.org/2011/08/10/web-archives-the-futures-trzy-scenariusze-rozwoju-archiwistyki-internetowej/> [accessed: March 19, 2014]; M. Wilkowski, *Trzy argumenty przeciwko archiwizowaniu Internetu*, October 4, 2011, Historia i Media, <http://historiaimedia.org/2011/10/04/trzy-argumenty-przeciwko-archiwizowaniu-internetu/> [accessed: March 19, 2014]; M. Wilkowski, *Dostępne dla ludu – Brewster Kahle o Internet Archive*, July 19, 2011, Historia i Media, <http://historiaimedia.org/2011/07/19/dostepne-dla-ludu-brewster-kahle-o-internet-archive/> [accessed: March 19, 2014]; M. Wilkowski, *Biuletyn EBiB o archiwizacji Internetu*, February 25, 2012, Historia i Media, <http://historiaimedia.org/2012/02/25/biuletyn-ebib-o-archiwizacji-internetu/> [accessed: March 19, 2014]; M. Wilkowski, *Archiwum Twittera w Bibliotece Kongresu*, June 14, 2011, Historia i Media, <http://historiaimedia.org/2011/06/14/archiwum-twittera-w-bibliotece-kongresu/> [accessed: March 19, 2014]; M. Wilkowski, *Archiwum brytyjskiego WWW: UK Web Archive*, November 9, 2011, Historia i Media, <http://historiaimedia.org/2011/11/09/archiwum-brytyjskiego-www-uk-web-archive/> [accessed: March 19, 2014].

These deliberations lead to conclusion, that archiving all Internet resources is not necessary. From the point of view of archival science meeting the challenge of contemporary humanities the most important part is archiving what is unique on the Internet, what is original, what was created as primary information, and not secondary publication. Naturally, it is possible and justified to preserve resources created outside the Web, but published on the Web, that on-line are an original, unique information entity (for example Public Information Bulletin – Biuletyn Informacji Publicznej), and in general – saving traces of on-line human activity.

Maybe challenges of contemporary humanities might be met not by comprehensive Internet preservation, which is very costly, but by a broadly conceived problem (topic) archiving. It can be implemented both by community archives according to their goals (and chosen topic of collections), and by public archives, using tools designed for supporting this type of preserving: collecting projects and documentation strategies.¹⁵

A collecting project consists in collecting, and at the same time preserving existing documents that contain information on a particular event, phenomenon or process, which chronological frames can be outlined. But the process of gathering the documentation can be permanent.¹⁶ Thanks to a documentation strategy, on the other hand, one can choose from contemporary produced documents those that will be archived. By that a complex strategy is designed, that enables continuous collecting and preservation of information on an interesting event, phenomenon, process or area, that is produced in many various places.¹⁷ Both strategies are pursued according to similar formula. First an event, phenomenon or process needs to be selected. Then, identification and selection of producers of the chosen documentation (interesting according to our strategy) needs to be pursued, after that the documentation should be archived – after taking care of all legal, technical and substantive issues.¹⁸ Implementing these two strategies is a part of archives' activity apart from classical archival control over currently operating institu-

¹⁵ Look: R. Degen, M. Jabłońska, *Collecting projects i documentation strategies jako narzędzia kształtowania zasobu archiwalnego i budowania wizerunku archiwów*, [w:] Archiwa organizacji pozarządowych w Polsce, red. T. Czarnota, M. Konstankiewicz,, NDAP 2015, s. 167 – 173.

¹⁶ Ibidem.,

¹⁷ Ibidem.

¹⁸ Ibidem.

tions – which is a passive action, that values already produced documents, and does not provoke production of any other materials.

Archiving Internet resources using means that resemble collecting projects and documentation strategies is already done, successfully. An example of that is the mentioned above Cyber Cemetery – a topically based archiving of Web sources done as part of a federal program in the USA; it's goal is to provide access to websites of non-existing governmental agencies and committees. It is an archive of the US governmental web pages, that ceased to exist.¹⁹ Another example of this type of archiving is "Tamiment Library Web Archive: Labor and the Left", set up in 2007 at the New York University. One place stores information about topically based Internet archives concerning labor and left-oriented movements. These websites' topics are: New York, left-oriented movements, alternative mass-media, anarchism, animal rights, civil rights and liberties, communism, socialism, Trotskyism, ecology, feminism and women' rights movements, labor unions and organizations in the US.²⁰ Another examples are projects of Web-archiving devoted to the tragedy of September 11, 2001,²¹ described in the Polish literature by Marcin Wilkowski: "Understanding 9/11" created by the Internet Archive, Pew Internet & American Life Project and the Library of Congress²² or "September 11 Digital Archive", set up by the Center for History and New Media in cooperation with the Library of Congress.²³ Those projects consisted in gathering and preservation of Internet resources concerning the event. An interesting example is the "Hurricane Digital Memory Bank" run by the Roy Rosenzweig Center for History and New Media (CHNM) at George Mason University

¹⁹ Cyber Cemetery was set up in 1995 by the University of North Texas Libraries and the U.S. Government Printing Office, as part of the federal program the Depository Library. Cyber Cemetery, <http://govinfo.library.unt.edu/> [accessed: June 22, 2015].

²⁰ In the Tamiment Library Web Archive: Labor and the Left, websites are usually archived every three months, and access is provided after six months after archiving; Tamiment Library Web Archive, <http://www.nyu.edu/library/bobst/research/tam/wearchive.html> [accessed: June 22, 2015].

²¹ Look: M. Wilkowski, *Internetowe archiwum 10 kwietnia*, April 20, 2010, Historia i Media, <http://historiaimedia.org/2010/04/20/internetowe-archiwum-10-kwietnia/> [accessed: March 19, 2014].

²² Look: *Understanding 9/11*, september11.archive.org/, <https://archive.org/details/911> [accessed: June 21, 2015].

²³ Look: *September 11 Digital Archive*, <http://911digitalarchive.org/> [accessed: June 24, 2015].

and the University of New Orleans – collection and preservation of network information on Hurricanes Katrina and Rita.²⁴

To sum up, I would like to stress the fact, that collecting projects and documentation strategies can be also successfully implemented to collect resources originally produced on-line and existing only on the Internet – that are traces of human on-line activity (learning, playing, working, pursuing hobby).

Bibliography

- 20 things I learned about browsers and the Web*, The Google Chrome Team, November 2010, <http://www.20thingsilearned.com/en-US/home> [accessed: June 21, 2015].
- Baliszewski T., *Sieć dla potomnych. Biblioteka Brytyjska archiwizuje internet*, NaTemat.pl, February 2013, <http://natemat.pl/57431,biblioteka-brytyjska-archiwizuje-internet> [accessed: March 19, 2014].
- Biuletyn EBiB, no. 1(128), 2012, http://www.ebib.pl/?page_id=380 [accessed: June 22, 2015].
- Czeski projekt archiwizacji Internetu*, Presentation Transcript, January 9, 2010, <http://www.slideshare.net/p3l3k/czeski-projekt-archiwizacji-internetu> [accessed: March 19, 2014].
- Degen R., Jabłońska M., Collecting projects i documentation strategies jako narzędzia kształtowania zasobu archiwalnego i budowania wizerunku archiwów, [w:] *Archiwa organizacji pozarządowych w Polsce*, red. T. Czarnota, M. Konstankiewicz, NDAP 2015, s. 167 – 173.
- Gmiterek G., *Archiwum internetowe – czy możliwa jest archiwizacja zasobów sieci?*, August 25, 2010, <http://www.biblioteki.org/pl/wiadomosci/czytaj/786> [accessed: March 19, 2014].
- Halawa M., *Nowe media i archiwizacja życia codziennego*, „Kultura współczesna”, no. 4, 2011 *Powrót do archiwów*, p. 27 – 41, <http://kulturawspolczesna.pl/readpdf/1271/Nowe%20media%20i%20archiwizacja%20%C5%BCycia%20codziennego> [accessed: June 21, 2015].
- Jachyra D., *Trollowanie – antyspołeczne zachowania w Internecie, sposoby wykrywania i obrony*, „Zeszyty Naukowe Uniwersytetu Szczecińskiego. Studia Informatica”, no. 28, 2011, p. 253–261, http://wneiz.pl/nauka_wneiz/studia_inf/28-2011/si-28-253.pdf [accessed: June 17, 2015].

²⁴ Look: *Hurricane Digital Memory Bank*, <http://hurricanearchive.org/> [accessed: June 24, 2015].

- Jeleśniański M., *Definicja i rodzaje blogów*, e-Redaktor.pl, March 25, 2008, <http://eredaktor.pl/teoria/definicja-i-rodzaje-blogow/> [accessed: June 18, 2015].
- Kłębczyk F., *Archiwizacja zasobów Internetu – kierunki i wyzwania*, „Archiwista Polski”, 2012, no. 3, p. 105–112.
- Kopańko K., *Archiwizacja Internetu dla przyszłych pokoleń*, April 9, 2013, Antyweb.pl, <http://antyweb.pl/archiwizacja-internetu-dla-przyszlych-pokolen/> [accessed: March 19, 2014].
- Lis T., *Komorowski prawdopodobnie nie przegratby wyborów, gdyby nie „hejt industry”*, June 8, 2015, <http://polska.newsweek.pl/tomasz-lis-o-nienawisci-w-sieci-hejterzy-wyniki-wyborow,artykuly,364605,1.html> [accessed: June 17, 2015].
- Michałak P., Lange K., *Archiwizacja stron internetowych szkół wyższych, instytucji naukowych i kulturalnych*, <http://archiwum.amu.edu.pl/publikacje/17.pdf> [accessed: June 22, 2015].
- Nowe media = nowa partycipacja*, ed. P. Celiński, Lublin 2014, p. 12–21, http://kulturacyfrowa.org/wp-content/uploads/2014/12/IKC_book_v11.pdf [accessed: June 21, 2015].
- Pasztaleniec-Jarzyńska J., *Polityczna historia archiwizacji sieci: Międzynarodowe Konsorcjum Archiwizacji Internetu*, June 13, 2012, 14:35, [in:] Babin Bibliografia Analityczna Bibliotekoznawstwa i Informacji Naukowej, <http://babin.bn.org.pl/?p=1158> [accessed: March 19, 2014].
- Sobczak A., *Internet jako globalne archiwum społeczne – rozważania na temat roli Internetu w dokumentowaniu dziejów ludzkości*, [in:] Toruńskie Konfrontacje Archiwalne, vol. 4, *Nowa archiwistyka – archiwa i archiwistyka w późnonowoczesnym kontekście kulturowym*, ed. W. Chorążyczewski, W. Piasek, A. Rosa, Toruń 2014, p. 237–247.
- Stasiak P., *Naukowcy zbadali hejt internetowy. Co piąty komentarz jest obelżywy, zaś trolle..., „Newsweek Polska”*, August 8, 2014, <http://nauka.newsweek.pl/naukowcy-zbadali-hejt-internetowy-co-piaty-komentarz-jest-obelzywy-,artykuly,345268,1.html> [accessed: June 17, 2015].
- Świderski B., *Najciemniejszy zakątek internetu naprawdę istnieje. Ukryta sieć TOR: „Lewe” papiery, pedofilia, przekręty i narkotyki*, Na temat.pl, <http://natemat.pl/32267,najciemniejszy-zakatek-internetu-naprawde-istnieje-ukryta-siec-tor-lewe-papiery-pedofilia-przekrety-i-narkotyki> [accessed: June 18, 2015].
- Wawrzyn M., *Psychologia trolla. Czemu niektórzy internauci tak lubią wkurzać innych*, <http://gadzetomania.pl/3968,psychologia-trolla-czemu-niktorzy-internauci-tak-lubia-wkurzac-innych> [accessed: June 17, 2015].
- Wilkowski M., *Archiwum brytyjskiego WWW: UK Web Archive*, November 9, 2011, Historia i Media, <http://historiaimmedia.org/2011/11/09/archiwum-brytyjskiego-www-uk-web-archive/> [accessed: June 24, 2015].
- Wilkowski M., *Archiwum Twittera w Bibliotece Kongresu*, June 14, 2011, Historia i Media, <http://historiaimmedia.org/2011/06/14/archiwum-twittera-w-bibliotece-kongresu/> [accessed: March 19, 2014].

- Wilkowski M., *Biuletyn EBIB o archiwizacji Internetu*, February 25, 2012, Historia i Media, <http://historiaimmedia.org/2012/02/25/biuletyn-ebib-o-archiwizacji-internetu/> [accessed: March 19, 2014].
- Wilkowski M., *Dostępne dla ludu – Brewster Kahle o Internet Archive*, July 19, 2011, Historia i Media, <http://historiaimmedia.org/2011/07/19/dostepne-dla-ludu-brewster-kahle-o-internet-archive/> [accessed: March 19, 2014].
- Wilkowski M., *Internetowe archiwum 10 kwietnia*, April 20, 2010, Historia i Media, <http://historiaimmedia.org/2010/04/20/internetowe-archiwum-10-kwietnia/> [accessed: March 19, 2014].
- Wilkowski M., *Od osobistej archiwistyki cyfrowej do edukacji medialnej*, EBiB, no. 6(151), 2014, <http://open.ebib.pl/ojs/index.php/ebib/article/view/274/436> [accessed: June 21, 2015].
- Wilkowski M., *Trzy argumenty przeciwko archiwizowaniu Internetu*, October 4, 2011, Historia i Media, <http://historiaimmedia.org/2011/10/04/trzy-argumenty-przeciwko-archiwizowaniu-internetu/> [accessed: March 19, 2014].
- Wilkowski M., *Web Archives: The Future(s) – trzy scenariusze rozwoju archiwistyki internetowej*, August 10, 2011, Historia i Media, <http://historiaimmedia.org/2011/08/10/web-archives-the-futures-trzy-scenariusze-rozwoju-archiwistyki-internetowej/> [accessed: March 19, 2014].

Summary

Ślad człowieka w Internecie - zagadnienie archiwizacji Internetu od strony archiwistyki zantropologizowanej

The article is an attempt to answer the question about presence and products of a man on the Web from the point of view of archival science, which meets a challenge of contemporary humanities. The author studied actions taken by a man on the Internet in the context of possibilities offered. Some attention was paid to currently used methods of archiving the Web, what enabled the author to draw some general conclusions about factors, which mostly decide about archiving the Internet. The approach of anthropology-oriented archival science to archiving the Internet requires implementation of proper archiving strategies. A key in studying these strategies should be analysis of actions and products of people, paying attention to activities of a man on the Internet and his/her products – put on the Web and created on the Web.