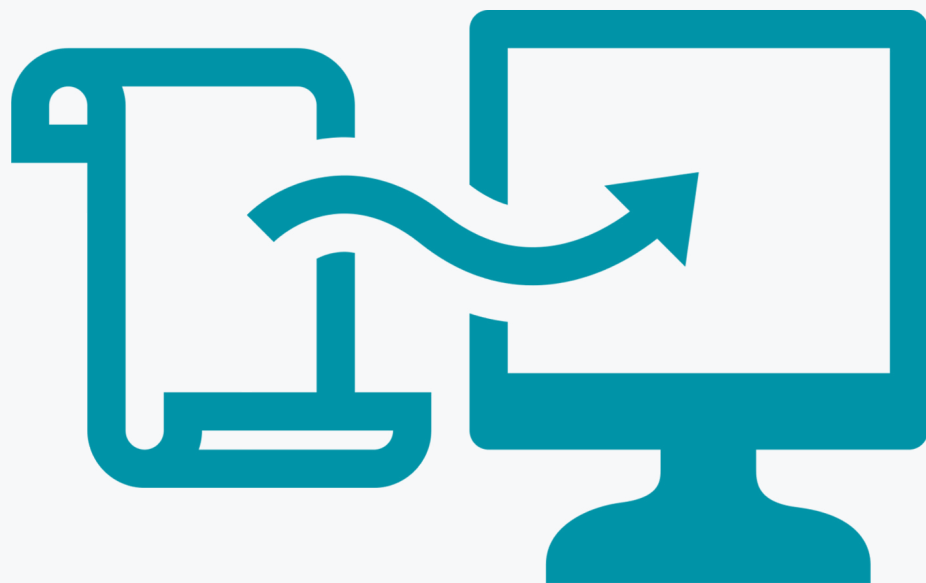


MARCIN WILKOWSKI



# WPROWADZENIE DO HISTORII CYFROWEJ

Instytut Kultury Miejskiej  
Gdańsk

# Wprowadzenie do historii cyfrowej

Marcin Wilkowski

Instytut Kultury Miejskiej  
Gdańsk 2013

## Wprowadzenie do historii cyfrowej

Marcin Wilkowski

Wydanie drugie uzupełnione - wersja elektroniczna

Gdańsk 2013

ISBN 978-83-64610-00-4

Wydawca:

Instytut Kultury Miejskiej

ul. Długi Targ 39/40,

80-830 Gdańsk

**Instytut  
Kultury  
Miejskiej**



Redakcja i korekta: Joanna Tyka (Filologos.pl), Paweł Koziół

Skład, projekt okładki: Tomasz Kielkowski

Ikona na okładce powstała w ramach Iconathonu (iconathon.org)

Informacja o prawach autorskich:

© Instytut Kultury Miejskiej 2013

Publikacja udostępniana jest na licencji Creative Commons Uznanie autorstwa. Na tych samych warunkach 3.0 Unported (CC BY-SA). Jej treść może być swobodnie kopiowana, rozpowszechniana i przetwarzana (także w celach komercyjnych) bez dodatkowych zezwoleń i umów, pod warunkiem udostępnienia informacji o autorze oraz zastosowania tej samej licencji w przypadku rozpowszechniania utworów zależnych. Więcej informacji na ten temat znaleźć można na stronie <http://creativecommons.org/licenses/by-sa/3.0/deed.pl>, gdzie dostępne jest także przystępne podsumowanie powyższych zasad.

*Dziękuję Radosławowi Bombie, Emanuelowi Kulczyckiemu, Marcinowi Werli, Rafałowi Wójcikowi oraz Katarzynie Ziętał za konsultacje, które pomogły w opracowaniu niniejszej publikacji.*

Wersja cyfrowa książki dostępna jest na stronie <http://historiacyfrowa.ikm.gda.pl/>

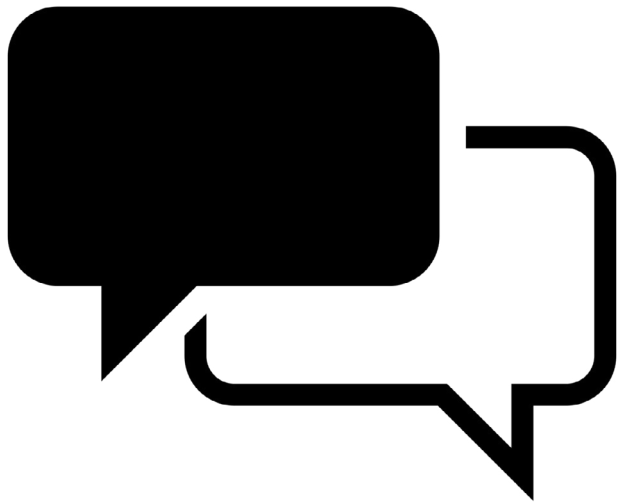
Treść książki jest otwarta, zachęcam więc do jej uzupełniania i rozbudowywania – Marcin Wilkowski (marcin@wilkowski.org).

# Spis treści

Definiowanie cyfrowej historii	6
Współczesny Internet nie tylko z perspektywy historyka	15
Digitalizacja dziedzictwa i cechy formatu cyfrowego	26
Wykorzystywanie repozytoriów cyfrowych i archiwalnych baz danych	31
Archiwa społeczne, oddolna digitalizacja i cyfrowa archiwistyka osobista	39
Wyszukiwanie pełnotekstowe w archiwum cyfrowym i <i>text mining</i>	48
Cyfrowe dziedzictwo	60

Instytucje dziedzictwa w kulturze cyfrowej	72
Historiografia poza drukiem	84
Otwarty dostęp do publikacji naukowych	92
Praca grupowa i nauka obywatelska	102
Nowe internetowe formy komunikacji naukowej	109
Internet i społeczna rola nauki historycznej	115

mojemu Ojcu



# Definiowanie cyfrowej historii

Zainteresowanie historyków komputerami nie jest niczym nowym. Już pod koniec lat 40. zakonnik o. Roberto Busa SJ pracował w firmie IBM nad konkordancją dzieł św. Tomasza (tzw. *Corpus Tomisticum*), zwieńczeniem czego były wydane w latach 90. edycje cyfrowe na płytach CD-ROM (Prinke, 2000, s. 7). Od końca lat 50. rozwijał się też nurt historii kwantytatywnej, w ramach którego historycy posługiwali się dostępnymi im współcześnie elektronicznymi narzędziami obliczeniowymi, aby zestawiać ze sobą dane historyczne, pozwalające opisywać złożone i wielowątkowe procesy społeczne z przeszłości (Grabski, 2006, s. 773 i nast.). Ich wiara w potencjał badań ilościowych była silna: 8 maja 1968 roku we francuskim tygodniku „Le Nouvel Observateur” Emmanuel Le Roy Ladurie, przedstawiciel trzeciej generacji szkoły Annales, stwierdził nawet, że odtąd albo historyk będzie programistą, albo nie będzie go wcale. Nawet dziś, mimo wciąż zwiększającej się wśród historyków popularności narzędzi cyfrowych, taka teza brzmi rewolucyjnie. Nic więc dziwnego, że jej krytyka miała swoje źródło także w środowisku samego Le Roy Ladurie’a. Fernand Braudel sprzeciwiał się myśleniu kategoriami narzędzi – to, co wartościowe w posługiwaniu się komputerami, miało pojawić się na poziomie samej strategii badawczej:

Obawiam się i będę się spierał z Emmanuelem Le Roy Ladurie o to, że istnieje jakieś złudzenie albo wytłumaczenie dla stwierdzenia, mówiąc z perspektywy historii statystycznej, że historyk w przyszłości będzie musiał być programistą, albo nie będzie historykiem. To, co mnie interesuje, to program



programisty (ang. *programmer's program*). Teraz powinien on dążyć do unifikacji wszystkich nauk społecznych (za Dosse, 1994, s. 116).

Co ciekawe, sam Le Roy Ladurie może być przez polskiego czytelnika kojarzony głównie za sprawą swojej pracy, mikrohistorii pod tytułem „Montaillou. Wioska heretyków 1294–1324”, wydanej w 1975 roku (wyd. polskie z 1988 roku), w której zastosował on metodę zdecydowanie odbiegającą od modelu badania ilościowego.

Rozwój historii cyfrowej nie byłby możliwy bez ewolucji komputerów, które stawały się z czasem coraz bardziej dostępne i użyteczne w codziennej pracy badawczej (np. edytory tekstu) czy dydaktycznej (płyty CD z multimedialnymi materiałami edukacyjnymi). Za ważną cezurę uznać można rok 1976. Utworzono wówczas Oxford Text Archive (OTA) – istniejące do dziś, najstarsze cyfrowe repozytorium dawnych tekstów literackich i korpusów wykorzystywanych w badaniach językowych. Powstanie i rozwój Internetu nie tylko ułatwiły komunikację między badaczami (np. za pomocą poczty elektronicznej), ale także pozwoliły na udostępnianie tekstów źródłowych czy opracowań naukowych online (za pomocą stron WWW). W 1994 roku, kiedy popularność zaczęły zdobywać pierwsze przeglądarki graficzne, takie jak Mosaic i Netscape Navigator, amerykański historyk Roy Rosenzweig założył Center for History and New Media (CHNM). Sukcesem tej instytucji i samego Rosenzweiga, określanego czasem mianem pierwszego cyfrowego historyka (*digital historian*), było uwiarygodnienie cyfrowej historii jako przedmiotu zainteresowań i praktyki akademickiej (zob. Zyziak, 2010, s. 154-155).

Warto w tym miejscu wspomnieć jeszcze, że w tym samym czasie, czyli w roku 1995, Wydawnictwo Uniwersytetu Mikołaja Kopernika opublikowało materiały z I Sympozjum Metod Komputerowych, organizowanego w Poznaniu w listopadzie 1994 roku przez, nieaktywny już dziś, polski oddział Association for History and Computing (AHC) i Komisję Metod Komputerowych Polskiego Towarzystwa Historycznego. W wydanym kilka lat później tomie dokumentującym II Sympozjum z grudnia 1995 roku Jerzy Wisłocki proponował utworzenie polskiego Centrum Elektronicznych Tekstów Historycznych, a Rafał T. Prinke opisywał wyzwania dotyczące komunikacji pomiędzy historykiem a informatykiem podczas wspólnej pracy nad projektami badawczymi (Ryszewski, 1997).

Pojęcie historii cyfrowej (ang. *digital history*) pojawiło się po raz pierwszy w projektach dydaktycznych realizowanych w Virginia Center for Digital

History w latach 1997–1998. Jeden z prowadzących te zajęcia, Edward L. Ayers, opublikował w 1999 roku esej *The Pasts and Futures of Digital History* (Ayers, 1999). Początkowo historii cyfrowej upatrywał on w hipertekstowych projektach edukacyjnych z lat 80., wykorzystujących system HyperCard, oraz w pierwszych edukacyjnych CD-ROM-ach z połowy lat 90., np. podręczniku *Who Built America? From the Centennial Celebration of 1876 to the Great War of 1914*, wydanym w 1995 roku. Ayers, pisząc o historii cyfrowej, nie koncentrował się wyłącznie na edukacji – pokazywał, że format cyfrowy pozwala na zbudowanie nowych interaktywnych modeli publikowania naukowego. Do tego wątku wrócę w jednym z ostatnich rozdziałów.

Od końca lat 90. wraz z ewolucją i upowszechnianiem się Internetu oraz technik cyfrowych ewoluowały najważniejsze narzędzia i kierunki historii cyfrowej. W pierwszej dekadzie XXI wieku historycy przyglądali się zarówno popularności Wikipedii, jak i samego modelu *wiki* (jako systemu pozwalającego na kolektywne budowanie wiedzy), a także serwisom kojarzonym z ideą Web 2.0, postępom projektów masowej digitalizacji, portalom społecznościowym itp. Idee historii cyfrowej kształtowane były zawsze w relacji z rozwijającym się Internetem i technologiami cyfrowymi, i tak jest do dziś.

W jaki sposób wobec tego należy skonstruować definicję cyfrowej historii i czy w ogóle jest sens wyodrębniania jej jako osobnej dziedziny? Czy da się znaleźć jakieś jej uniwersalne elementy, tak aby definicja utrzymała swoją moc mimo nieustannego rozwoju technologii? W opublikowanym w 2008 roku w „*The Journal of American History*” zapisie dyskusji, który opatrzone wiele mówiącym tytułem *The promise of Digital History*, kilku amerykańskich historyków próbowało określić ją na własny sposób. Zgodnie z ich propozycjami, historia cyfrowa miała być próbą badania i prezentacji przeszłości, wykorzystującą nowe technologie komunikacyjne i komputery, otwartą przestrzenią współpracy i komunikacji między historykami oraz zestawem narzędzi, pojęć i metod (ang. *framework*), pozwalającym w nowy sposób prowadzić dyskusje o przeszłości. Innym wyróżnikiem definicyjnym zaproponowanym przez uczestników dyskusji było umożliwienie historykom badania obszernych zestawów źródeł i danych, powstających dzięki programom masowej digitalizacji, oraz nowych źródeł istniejących już wyłącznie w formie cyfrowej. Zwrócono też uwagę na elastyczność pojęcia – historia cyfrowa miała być dziedziną ewoluującą od prostego wykorzystywania komputera i Internetu w pracy historyka (edytory tekstu, poczta elektroniczna) aż po rekonstrukcję przestrzeni historycznych w środowisku 3D

i nowe metody edukacji historycznej o bardzo demokratycznym, konstruktywistycznym charakterze. W dyskusji pojawił się także wątek zmian w formule archiwum i w sposobach samej pracy na źródłach oraz redukcji dystansu między historią akademicką i popularną. Z grona kilku dyskutantów tylko jeden krytycznie odniósł się do idei formułowania nowego kierunku w ramach dyscypliny historycznej i zwrócił uwagę na potencjalnie szybkie wyczerpanie się pojęcia cyfrowej historii z racji powszechnego wykorzystywania przez historyków narzędzi i źródeł dostępnych w formie cyfrowej. Skoro bowiem wszyscy historycy korzystają z komputerów pisząc swoje teksty naukowe, wszyscy są historykami cyfrowymi (Cohen i in., 2008).

CHNM na swojej oficjalnej stronie internetowej (<http://chnm.gmu.edu/>) proponuje własną definicję cyfrowej historii:

Historia cyfrowa to podejście do badania i reprezentowania przeszłości uwzględniające nowe technologie komunikacyjne, takie jak komputery i Sieć. Opiera się na istotnych właściwościach cyfrowej rzeczywistości, takich jak bazy danych, hipertekstualizacja i sieci, aby tworzyć i udostępnić wiedzę historyczną. Historia cyfrowa dopełnia inne formy historii – w istocie podkreśla siłę i metodologiczny rygor tego wiekowego sposobu ludzkiego rozumowania, używając najnowszych technologii.

Historia cyfrowa nie jest tu więc radykalnym zerwaniem z dotychczasowymi praktykami, propozycją nowego paradygmatu czy zwrotu, ale ewolucyjnym rozwinięciem starych metod.

Listę wyzwań stojących przed naukami historycznymi – w założeniu niepełną i otwartą – proponuje Toni Weller w redagowanym przez siebie tomie *History In the Digital Age* (Weller, 2012, s. 6–7). Są to problemy związane, m.in. z digitalizacją, zabezpieczaniem i zarządzaniem zasobami historycznymi, nieposiadającymi oryginalnie postaci cyfrowej (*born digital*), kwestią starych formatów i nośników danych, autentyczności źródeł cyfrowych i nowych metod ich badania, relacji między profesjonalnymi archiwami cyfrowymi a oddolną działalnością archiwalną, zaangażowaniem osób spoza historycznego środowiska akademickiego w działania o charakterze naukowym czy w końcu kwestii związanych z prawem autorskim. Jak podkreśla Weller, część z tych wyzwań dotyczy podstaw warsztatu badacza przeszłości i może być ważna dla wszystkich historyków, część jednak dotyczyć będzie jedynie specjalistów.

Obecnie ogromną popularność zdobywa pojęcie cyfrowej humanistyki (ang. *digital humanities*), które okazało się bardziej pojemne i elastyczne, a zatem także bardziej użyteczne, nie tylko w publicystyce naukowej, ale także w języku ofert grantowych czy nazwach instytucji badawczych. Podobnie jak w przypadku historii cyfrowej, mamy tu do czynienia z pojęciem obejmującym potencjalnie wiele zjawisk i bardzo trudnym do jednoznacznego zdefiniowania, które na dodatek stało się modne (ang. *buzzword*) i z pewnością bywa nadużywane.

O cyfrowej humanistyce piszą dziś nie tylko media branżowe, takie jak „Chronicle of Higher Education”, ale też popularne, jak choćby „New York Times”. Google w swojej ofercie grantowej dla naukowców posługuje się tym terminem. Pojawia się on także w ogłoszeniach o pracę, publikowanych przez uczelnie i instytuty badawcze, w nazwach instytucji oraz organizacji naukowych (Göttingen Centre for Digital Humanities, The Digital Humanities Centre – The University of Nottingham, UCL Centre for Digital Humanities), w tytułach konferencji i wykładów oraz oczywiście w literaturze naukowej. Uznaje się, że wydarzeniem, które w istotny sposób przyczyniło się do rozpowszechnienia pojęcia *digital humanities*, była odbywająca się w 2009 roku konferencja Modern Language Association of America (MLA). Krytycy, sceptycznie oceniając popularność tego pojęcia w międzynarodowym środowisku akademickim, wskazują na dominację Amerykanów. Jak zauważa Andrew Prescott z londyńskiego King’s College, to głównie oni są członkami międzynarodowych organizacji *digital humanities*. Według niego amerykańskim problemom, nie tylko tym dotyczącym materii badawczej, ale i systemowym, związanym choćby z modelem kariery akademickiej, niesprawiedliwie nadawane jest globalne znaczenie, co prowadzi do marginalizacji europejskich badaczy i ich bogatego, wieloletniego doświadczenia w badaniach komputerowych nad kulturą, językiem czy historią. W jego opinii cyfrowa humanistyka jako pewien nurt w nauce uznawany może być za przejaw globalizacji (Prescott, 2013).

Cyfrowa humanistyka definiowana jest tak samo niejednoznacznie, jak cyfrowa historia. Oczywiście nie ma sensu przytaczać choćby najważniejszych wątków w sporze na temat tego, czym jest, a czym nie jest ta nowa propozycja w naukach humanistycznych (i społecznych). Chciałbym jednak odwołać się do propozycji, która w tej dyskusji podoba mi się najbardziej i jest, według mnie, najatrakcyjniejsza. Jej autorką jest Lisa Spiro, jedna z czołowych postaci środowiska cyfrowej humanistyki, blogerka, dyrektor zarządzający pracami badawczymi w National Institute for Technology in Liberal Education (NITLE). Według niej

cyfrową humanistykę można definiować także przez takie atrybuty, jak: otwartość badań i komunikacji naukowej, współpraca między badaczami i umiejętność wyjścia poza środowisko akademickie, kolegalność, zróżnicowanie kierunków badawczych i samego środowiska czy eksperymentowanie z nowymi metodami prowadzenia badań i przekazywania wiedzy, wykorzystującymi technologie cyfrowe (Spiro, 2012, s. 24–30). Spiro pisze też, że wartości będące podstawą *digital humanities* wynikają wprost z ideałów tradycyjnej humanistyki.

O tym, że koncepcja cyfrowej humanistyki jest w pewnym sensie powrotem do renesansowych, humanistycznych koncepcji badacza i pracy naukowej, przekonuje też Piotr Celiński w artykule zamieszczonym w wydanym w maju 2013 roku zbiorczym tomie *Zwrot cyfrowy w humanistyce*. Wspólne mianowniki dwóch tak odległych czasowo kierunków to m.in. matematyczne widzenie świata (umiejętność opisywania rzeczywistości społecznej i kulturowej za pomocą liczb), interdyscyplinarność i łączenie refleksji teoretycznej z praktycznym warsztatem (także z tworzeniem własnych narzędzi), zaangażowanie badaczy w upowszechnianie wyników własnej pracy naukowej przez sprawne wykorzystywanie dostępnych mediów, przy jednoczesnym krytycznym rozpoznaniu ich kulturowego i komunikacyjnego znaczenia (Celiński, 2013).

Nie można oczywiście zapominać o kontekście współczesnych dyskusji na temat cyfrowej historii czy – szerzej – cyfrowej humanistyki. Jest nim wyraźny kryzys wizerunkowy dziedzin humanistycznych oraz liczne przekazy medialne wskazujące na ich komercyjną nieefektywność i społeczną nieprzydatność. Kiedy, jak pisze Michał Paweł Markowski, w *neoliberalnym społeczeństwie nastawionym na pomnażanie zysku uniwersytet jest fabryką produkującą korzyści*, istnieje pokusa, aby nowe, cyfrowe zaangażowanie humanistyki traktować jako próbę legitymizacji humanistyki w ogóle. Dziać się to może choćby przez wyszukiwanie w tym nowym kierunku elementów z natury nauk ścisłych i technicznych (*science*). Miałyby one neutralizować te cechy humanistyki, które są dziś najsilniej odrzucane przez społeczeństwo i decydentów wpływających na budżet nauki czy programy szkolnej edukacji: bezinteresowność i nieprzekładalność (przynajmniej w sposób bezpośredni) na rozwój ekonomiczny czy rynek pracy:

Porozumienie nie jest tu w ogóle możliwe, co widać jak na dłoni w bezskutecznych pohukiwaniach humanistów, zdziwionych, że podatnik nie chce finansować ich badań nad średniowieczną składnią zaginionych utworów lub osiemnastowieczną odą czy inną elegią. Neoliberalne społeczeństwo

nie życzy sobie, aby jakiegokolwiek publiczne pieniądze szły na bezużyteczne rzeczy, i ma rację, tyle tylko, że ta racja (całkowicie spójna z zasadami neoliberalnej gospodarki) kłóci się z racją humanistyki, opartej na całkiem przeciwnych zasadach, które – z definicji – znajdują się na przegranej pozycji (Markowski, 2013, s. 31-32).

Społeczna odpowiedzialność i użyteczność humanistyki to temat na zupełnie inne opracowanie. Jestem jednak przekonany, że chociaż jej nowy cyfrowy charakter nie może być przedstawiany jako argument za jej legitymizacją, to narzędzia cyfrowe, wolny dostęp do publikacji i jej nowe formy czy otwieranie się badaczy na naukę obywatelską może niwelować społeczny dystans do humanistyki, włączając w jej zasięg także osoby spoza środowiska akademickiego.

Na koniec chciałbym zasygnalizować pewne dylematy, jakie pojawiają się przy pisaniu o cyfrowej historii. Pierwszym z nich jest problem aktualności treści – treść podręcznika będzie z konieczności szkicowym opisem pewnego etapu rozwoju tej dynamicznej dziedziny, który już za jakiś czas, przynajmniej w niektórych punktach, może stać się nieaktualny. Kolejnym problemem jest sygnalizowany powyżej niewielki zasięg cyfrowej historii jako specyficznej niszy badawczej, w polskich realiach jeszcze pogłębiany ograniczonym dostępem do najnowszej zagranicznej literatury, brakiem profilowanych grantów badawczych czy instytucji, która mogłaby realizować podobną misję jak Center for History and New Media. Istotnym wyzwaniem dla autora jest także język, ponieważ wiele pojęć przytaczanych w tym opracowaniu nie ma polskich odpowiedników o ustalonym już w literaturze brzmieniu, a nawet sam termin „historia cyfrowa” nie musi być jedyną propozycją tłumaczenia *digital history*. Z uwagi na brak edycji polskojęzycznych podjąłem się też samodzielnego tłumaczenia zamieszczonych w tekście cytatów z prac anglojęzycznych.

W niniejszym skrypcie próbuję opisać historię cyfrową nie tylko przez pytania badawcze, metody czy wykorzystywane do pracy naukowej narzędzia, ale także przez wartości. Stąd nieodłączne są odwołania do idei otwartości badań, dostępności zasobów historycznych czy przekraczania granic dyscyplin, tradycyjnych instytucji dziedzictwa (archiwum, muzeum, biblioteka), jak i do modeli publikowania naukowego, dających większą wolność czytelnikom. Najważniejszym odniesieniem jest Internet, jako przestrzeń komunikacji i infrastruktura dla narzędzi cyfrowych i zasobów, stąd niewiele miejsca poświęcam samemu wykorzystaniu komputerów w badaniach historycznych. To osobny

wątek historii cyfrowej, który należałoby opisać oddzielnie, obejmujący takie zagadnienia, jak: tworzenie elektronicznych edycji tekstów źródłowych czy historycznych (np. językowych), baz danych, metody automatycznej analizy treści (*text mining*), wizualizacje, narzędzia GIS (ang. *Geographic Information System*) itp. Jeśli poruszam te tematy w swoim opracowaniu, to zawsze w perspektywie Internetu i bez zagłębiania się w ich specyfikę. Czytelnik zainteresowany tymi tematami bez problemu dotrze do odpowiedniej literatury, także w języku polskim.

Wszystkie odnośniki do stron i dokumentów internetowych były dostępne w dniu 10 grudnia 2013 roku.

## BIBLIOGRAFIA

- Ayers, E. L. (1999). *The Pasts and Futures of Digital History*. Pobrano z: <http://www.vcdh.virginia.edu/PastsFutures.html>.
- Celiński, P. (2013). Renesansowe korzenie cyfrowego zwrotu. W: Radomski, A., Bomba, R. (red.), *Zwrot cyfrowy w humanistyce. Internet / Nowe Media / Kultura 2.0* (s. 13-36). Lublin: E-naukowiec. Pobrano z: [http://e-naukowiec.eu/wp-content/uploads/2013/05/Zwrot\\_cyfrowy\\_w\\_humanistyce.pdf](http://e-naukowiec.eu/wp-content/uploads/2013/05/Zwrot_cyfrowy_w_humanistyce.pdf)
- Cohen, D. J., Frisch, M., Gallagher, P., Mintz, S., Sword, K., Taylor, A. M., & Turkel, W. J. (2008). Interchange: the promise of digital history. *Journal of American History*, 95(2), 452-491. Pobrano z: <http://www.journalofamericanhistory.org/issues/952/interchange/index.html>
- Dosse, F. (1994). *New History in France: The Triumph of the Annales*, University of Illinois.
- Grabski, A. (2006). *Dzieje historiografii*, Poznań: Wydawnictwo Poznańskie.
- Markowski, P. M. (2013). *Polityka wrażliwości. Wprowadzenie do humanistyki*, Kraków: Towarzystwo Autorów i Wydawców Prac Naukowych UNIVERSITAS.
- Prinke, T. (2000). *Fontes ex machina: komputerowa analiza źródeł historycznych*, Poznań: Polska Akademia Nauk.
- Prescott, A. (2013). Small Worlds and Big Tents. W: *Digital Riffs – Andrew Prescott*, Pobrano z: <http://digitalriffs.blogspot.co.uk/2013/05/small-worlds-and-big-tents.html>

- Ryszewski, B.** (1997). *Metody komputerowe w badaniach i nauczaniu historii: materiały II Sympozjum Polskiego Oddziału Association for History and Computing, Komisji Metod Komputerowych Polskiego Towarzystwa Historycznego, Poznań 1 i 2 grudnia 1995 r.* (red), Toruń, Wydawnictwo Uniwersytetu Mikołaja Kopernika.
- Spiro, L.** (2012). This Is Why We Fight: Defining the Values of the Digital Humanities. W: Gold M.K. (red.), *Debates in the Digital Humanities* (s.165). Minneapolis: University Of Minnesota Press. Pobrano z: <http://dhdebates.gc.cuny.edu/debates/text/13>.
- Weller, T.** (2012). Introduction: history in the digital age. W: T. Wellerv (red.), *History in the digital age* (s. 1-20). London, New York: Routledge.
- Zyziak, A.** (2010). Historia cyfrowa – nowe ramy badań naukowych? Przypadek Centrum Historii i Nowych Mediów Roya Rosenzweiga, *Przegląd Humanistyczny*, 5/6, s. 151-162.





# Współczesny Internet nie tylko z perspektywy historyka

Korzystanie z komputera i Internetu – także w nauce i edukacji – należy dziś do podstawowych kompetencji komunikacyjnych. Historycy powszechnie korzystają z edytorów tekstu, komunikują się za pomocą poczty elektronicznej czy przeglądają katalogi biblioteczne online. Nie zawsze jednak potrafią wyjść poza ten schemat i dostrzec w Internecie użyteczne narzędzie, mogące wspierać ich badania czy prowadzoną przez nich pracę dydaktyczną oraz popularyzatorską.

Na przeszkodzie często stają brak wiedzy i odpowiednich umiejętności. Rafał T. Prinke, wymieniając, za pracą D. I. Greensteina z 1994 roku, cztery zasadnicze mity związane z kwestią upowszechnienia stosowania przez historyków metod komputerowych, mówi o pozornej potrzebie nieustannego doskonalenia wiedzy technicznej, co ma odciągać badacza od podstawowej pracy naukowej (Prinke, 2000, s. 11). Argument ten jest nieaktualny, przynajmniej od czasu pojawienia się pierwszych arkuszy kalkulacyjnych uruchamianych na komputerach klasy PC, które pozwoliły na łatwe zarządzanie danymi i ich podstawowe wizualizowanie za pomocą prostych wykresów. Wiele dostępnych dziś narzędzi komputerowych i internetowych projektowanych jest w taki sposób, aby można było skorzystać z nich bez konieczności posiadania szczególnych kompetencji informatycznych. Organizacja pracy grupowej nad tekstem naukowym, wsparcie online dla transkrypcji dokumentów historycznych czy gromadzenie bibliografii z baz online nie stanowią problemu, jeśli można skorzystać z ogólnodostępnych i darmowych narzędzi (wiele z nich posiada nawet spolszczone interfejsy).

Korzystanie z komputera i Internetu – także w nauce i edukacji – należy dziś do podstawowych kompetencji komunikacyjnych. Historycy powszechnie korzystają z edytorów tekstu, komunikują się za pomocą poczty elektronicznej czy przeglądają katalogi biblioteczne online. Nie zawsze jednak potrafią wyjść poza ten schemat i dostrzec w Internecie użyteczne narzędzie, mogące wspierać ich badania czy prowadzoną przez nich pracę dydaktyczną oraz popularyzatorską.

Na przeszkodzie często stają brak wiedzy i odpowiednich umiejętności. Rafał T. Prinke, wymieniając, za pracą D. I. Greensteina z 1994 roku, cztery zasadnicze mity związane z kwestią upowszechnienia stosowania przez historyków metod komputerowych, mówi o pozornej potrzebie nieustannego doskonalenia wiedzy technicznej, co ma odciągać badacza od podstawowej pracy naukowej (Prinke, 2000, s. 11). Argument ten jest nieaktualny, przynajmniej od czasu pojawienia się pierwszych arkuszy kalkulacyjnych uruchamianych na komputerach klasy PC, które pozwoliły na łatwe zarządzanie danymi i ich podstawowe wizualizowanie za pomocą prostych wykresów. Wiele dostępnych dziś narzędzi komputerowych i internetowych projektowanych jest w taki sposób, aby można było skorzystać z nich bez konieczności posiadania szczególnych kompetencji informatycznych. Organizacja pracy grupowej nad tekstem naukowym, wsparcie online dla transkrypcji dokumentów historycznych czy gromadzenie bibliografii z baz online nie stanowią problemu, jeśli można skorzystać z ogólnodostępnych i darmowych narzędzi (wiele z nich posiada nawet spolszczone interfejsy). Przygotowanie własnych wyszukiwarek, katalogów czy repozytoriów online, prezentacji multimedialnych, jak i tematycznych stron WWW wymaga od historyka jedynie odpowiedniej otwartości i chęci uczenia się z ogólnodostępnych zasobów edukacyjnych, dotyczących programowania, edycji multimediiów, budowania baz danych itp. Naturalnie, takie narzędzia nie muszą powstawać przy każdym projekcie naukowym, zawsze też warto zachować odpowiednie proporcje między czysto technicznym zaangażowaniem a pracą merytoryczną.

Dan Cohen w eseju *History and the Second Decade of the Web* (Cohen, 2004) wskazuje na rodzaje uprzedzeń, jakie mogą mieć historycy wobec Internetu. Jednym z nich jest przywiązanie do tradycyjnej pracy z książką i dokumentami (historyk powinien przecież przede wszystkim czytać). Fizyczność i stabilność papierowych nośników informacji i uwiarygodnione setkami lat miejsce książki w modelu pracy naukowej konkurują tu z płynnością i niestabilnością zasobów Internetu. Historyk chce posługiwać się tym, co stabilne, autentyczne, trwałe i czytelne (ang. *stable, authentic, persistent, legible*), Internet natomiast zdaje się oferować mu treści, których autentyczności nie zawsze może być pewien i które, ze względu na ilość, nie mogą być w prosty sposób odczytywane.

Dodajmy do tego jeszcze jedno wyzwanie dla naukowca: Internet to przestrzeń, w której historyk funkcjonuje na równych prawach z osobami spoza świata akademickiego, nierzadko wypowiadającymi się na bliskie mu naukowo tematy poza jakimkolwiek systemem ocen, recenzji czy uwiarygodniającego wsparcia instytucji naukowej. Idealnym przykładem może być tu Wikipedia, ale także artykuły publikowane na portalach, blogowe notki czy wpisy na forach dyskusyjnych. Próba wejścia w taką przestrzeń z profesjonalną wiedzą historyczną wydaje się karkołomna i bezcelowa – merytoryczne treści pozornie i tak zginą w morzu populizmu i bylejakości, a publikowanie ich online dodatkowo naraża autora na to, że jego praca zostanie splagiatowana.

Internet jest wyzwaniem nie tylko dla pojedynczych badaczy, ale także dla całego systemu akademickiego. Źródłem nieufności akademików wobec tego nowego (i oczywiście niepozbawionego wad) środka komunikacji i przestrzeni tworzenia i dystrybucji wiedzy mogą być cechy kultury technicznej, w ramach której Internet powstał i wciąż się rozwija. Wskazał na to w jednym ze swoich artykułów amerykański historyk Edward L. Ayers (2004, s. 54), zestawiając ze sobą charakterystyczne cechy tradycyjnego świata akademickiego i cechy kultury komputerowej:

<b>Kultura IT/Internetu</b>	<b>Kultura świata akademickiego</b>
anonimowość	silna identyfikacja
krótka tradycja	tradycja antyczna
niestabilność	konserwatyzm
inkluzyjność	hierarchiczność
komercyjność	niekomercyjność
szybkie starzenie się rozwiązań	trwałość
szybki efekt	cierpliwość w oczekiwaniu na pozytywne rezultaty
transparentność	nieprzejrzystość

Naukowcom funkcjonującym w konserwatywnym systemie akademickim o silnej hierarchiczności i dość ekskluzywnym charakterze nie zawsze jest łatwo odnaleźć się w przestrzeni, w której dominuje anonimowość i niestabilność, a miejsce rozbudowanej, przygotowanej odpowiednio długo wypowiedzi zajmuje krótki komentarz. Zdobycie nowych kompetencji technicznych i często też pokonanie własnych uprzedzeń to dopiero początek procesu wdrażania się w tę nową przestrzeń – procesu, który tak naprawdę nigdy się nie kończy, ponieważ Internet nie jest dany raz na zawsze i wymaga nieustannego uczenia się i podnoszenia własnych kompetencji. W porównaniu ze światem akademickim jest sferą radykalnie niestabilną i cecha ta nie dotyczy wyłącznie zmieniającego się oprogramowania, ale też ewoluujących nieustannie trendów i praktyk społecznych z nim związanych.

Warto przytoczyć w tym miejscu przykład *Second Life*, wieloosobowej gry sieciowej i środowiska 3D (stworzonego przez firmę Linden Lab), w ramach którego użytkownicy tworzyć mogą własne awatary i konstrukcje przestrzenne oraz komunikować się także za pomocą głosu i obrazu. Oprogramowanie tego typu określa się jako MUVE (ang. *multi-user virtual environment*) lub MMO (ang. *massively multiplayer online game*). W czerwcu 2013 roku system ten obchodził 10 rocznicę powstania, przez ten czas założono tam około 36 milionów kont. W okresie największej popularności *Second Life* obok wirtualnych rekonstrukcji prywatnych domów i siedzib firm w grze pojawiały się muzea, biblioteki i uczelnie, na których w wirtualnym świecie gry organizowano zdalne wykłady i seminaria. Dziś środowisko to pod kątem wygody korzystania i potencjału komunikacyjnego, a co za tym idzie także popularności, przegrywa zdecydowanie z Twitterem, Facebookiem czy Google+.

Gdyby zebrać argumenty, jakimi swego czasu posługiwali się twórcy wirtualnych bibliotek, archiwów i sal wykładowych w *Second Life*, na pewno znaleźlibyśmy odwołania do popularności tej gry: instytucje kultury i naukowcy poprzez swoją obecność w tej przestrzeni chcieli skutecznie dotrzeć do odbiorców. Nie bez znaczenia, szczególnie dla instytucji, były na pewno możliwości tworzenia atrakcyjnych obiektów, w których prezentowano zdigitalizowane zbiory lub organizowano wykłady. *Second Life* jako system komunikacji umożliwiał też transmisję głosu i obrazu, co pozwalało organizować wirtualne spotkania, prezentacje czy dyskusje dla publiczności z całego świata. Steven Warburton w swojej analizie z 2009 roku, a więc napisanej w najlepszym okresie rozwoju tej gry, wymienia aż osiem kategorii cech sprawiających, że może być ona dobrym

narzędziem w edukacji akademickiej (Warburton, 2009, s. 421). W ówczesnej perspektywie organizacyjne i finansowe zaangażowanie w Second Life wydawało się racjonalne.

Nic więc dziwnego, że przez pewien czas instytucje kultury oraz uczelnie inwestowały w tę przestrzeń, zamawiając realizacje swoich wirtualnych siedzib, opłacając miejsce w przestrzeni gry, a także osoby, które je tam promowały i nimi zarządzały. Jedno z pierwszych badań dotyczących Second Life jako przestrzeni edukacyjnej – brytyjski projekt *The schome-NAGTY Teen Second Life Pilot* – zostało zrealizowane w 2007 roku, a więc trzy lata po tym, jak Linden Lab upublicznił przygotowany przez siebie system. Już wtedy zwracano uwagę na problem wysokiej bariery wejścia i wynikające z założeń komercyjnych ograniczenia w swobodnym wykorzystywaniu możliwości świata gry (Cheal, 2007, s. 209).

Dziś wiemy już, że Second Life rozwijało się w kierunku innym niż powszechne praktyki korzystania z Internetu. Zmieniały się one pod wpływem nowych technologii sieci WWW oraz rozwoju urządzeń mobilnych. Produkt Linden Lab przestał być wyjątkowy, kiedy efekty 3D, które były jedną z jego głównych wartości, stały się osiągalne na smartfonach. Co więcej, można je było odtąd zastosować poza przestrzenią strony WWW: budowanie w wirtualnym świecie makiet zabytków przestało być racjonalne, skoro dzięki odpowiednim aplikacjom na urządzenia mobilne (tzw. rzeczywistości rozszerzonej, ang. *augmented reality*) wyświetlały się one w rzeczywistej przestrzeni. Kiedy zmieniły się praktyki komunikacyjne i Facebook przez swoje aplikacje mobilne umożliwił użytkownikom bycie razem online w dowolnym miejscu, dyskusje między awatarami w Second Life przestały być atrakcyjne. Użytkownicy wybrali inne interfejsy i usługi, w których teraz od nowa szansa na dotarcie ze swoją ofertą szukają instytucje i edukatorzy.

Przykład Second Life wskazuje na problem zmienności Internetu jako infrastruktury technicznej i przestrzeni społecznej komunikacji. Projektując jakiegokolwiek działania edukacyjne, popularyzatorskie czy naukowe online warto bardzo krytycznie odnosić się do aktualnie dominujących trendów. Jeszcze niedawno jednym z takich trendów było *Web 2.0*. Pojęcie to opisywało nowy etap w ewolucji Internetu, w ramach którego globalna Sieć miała stawać się bardziej partycypacyjna. Miejsce jednokierunkowej komunikacji reprezentowanej przez wielkie portale sprzed kryzysu *dot-com bubble* z 2001 roku zajęły projekty otwarte na treści tworzone przez użytkowników (Wikipedia, Flickr, blogi). Centralne redakcje zastąpić miała zbiorowa inteligencja, a nowe oprogramowanie stawało



Wystawa w wirtualnej siedzibie Biblioteki Kongresu w Second Life, sierpień 2006 roku.  
Fot. John Lester (CC BY), <http://www.flickr.com/photos/pathfinderlinden/227332249/>

się infrastrukturą nowych, partnerskich i aktywnych relacji między wydawcami a użytkownikami i konsumentami, określanymi teraz mianem prosumentów. Takie myślenie o Sieci przyjęte zostało także w świecie naukowym, gdzie stawało się (często bezkrytycznym) źródłem argumentów wspierających rozwój nauki obywatelskiej czy ruchu na rzecz wolnego dostępu do publikacji naukowych.

Web 2.0 nie jest jednak kierunkiem neutralnym i niepodlegającym krytyce. Można odrzucać poszczególne techniczne derywaty tego pojęcia, wskazując na niewystarczającą jakość treści tworzonych przez użytkowników czy udział ich powszechnego uczestnictwa w jej budowaniu. Można pokazywać problemy Wikipedii z wciąż spadającą liczbą aktywnych edytorów albo ujawniać, że narzędzia takie jak RSS czy blogi były skutecznie wykorzystywane zdecydowanie przez *założycielskim* dla pojęcia Web 2.0 artykułem Tima O'Reilly'ego z 2005 roku (O'Reilly, 2005). Można jednak spojrzeć na Web 2.0 jak na pewną ideologię, która przenoszona jest ze sfery biznesu do sfery nauki i edukacji: dość powiedzieć, że tekst O'Reilly'ego cytowany jest w ponad 5.5 tys. publikacji naukowych indeksowanych przez Google Scholar. Inspiracją co do kierunków krytyki może być artykuł Trebora Scholza w „Market Ideology and Myths of Web 2.0”, opublikowany w znanym internetowym czasopiśmie First Monday (Scholz, 2008). Według niego Web 2.0 jako ideologia kształtuje sposób, w jaki patrzymy i mówimy o współczesnym Internecie, neutralizując zupełnie takie problemy jak darmowa praca online (także praca w ramach zabawy, tzw. *playbor*)

lub cyfrowe wykluczenie i inne fizyczne ograniczenia i koszty dostępności treści i usług online. Nie bez znaczenia jest także kwestia prywatnej kontroli nad powszechnie wykorzystywanymi online narzędziami komunikacji i współpracy.

Web 2.0 przestało być atrakcyjnym, marketingowym terminem. Trudno też oszacować dziś, jaki będzie kierunek rozwoju Facebooka i innych usług mediów społecznościowych (ang. *social media*), nie wiemy, czy za kilka lat wśród jego uczestników dominować będzie ta sama grupa społeczna, co dziś. Projektując jakiegokolwiek działania online lub odwołując się w swoich publikacjach do opisujących Internet popularnych pojęć proponowanych przez biznes, warto zachować odpowiednio krytyczny dystans oraz mieć świadomość ewolucji rozwiązań technicznych i interfejsów, które nie są dane raz na zawsze.

Automatyzm w podejściu do Sieci przejawia się także często w myśleniu utopijnym i życzeniowym, w pewnej ideologizacji tej przestrzeni. Współczesna refleksja nad Internetem jako sferą komunikacji i środowiskiem społecznym odrzuca romantyczne wizje, w których Sieć stawała się narzędziem w radykalny sposób redukującym bariery społeczne w dostępie do wiedzy i kultury, zapewniającym każdemu egalitarne prawo do wypowiedzi i bycia zauważonym. Podejmując więc jakąkolwiek publiczną aktywność internetową, czy to w ramach indywidualnej działalności naukowej i popularyzatorskiej, czy też w ramach realizowania misji instytucji dziedzictwa (muzeum, archiwum), warto mieć świadomość ograniczeń i nieneutralności tego medium oraz zmienności tła społecznego, w którym się ono rozwija. Krytyczne spojrzenie na Internet to m.in. dostrzeganie problemu cyfrowego wykluczenia, dotyczącego w Polsce około 38 proc. osób powyżej 15 roku życia (badanie „World Internet Project: Polska 2013”). Myślenie o popularyzacji nauki w Internecie czy angażowaniu użytkowników cyfrowych repozytoriów archiwalnych i muzealnych nie może ignorować problemu kapitału kulturowego wśród odbiorców i musi zakładać, że mimo pozornie powszechnego i egalitarnego zasięgu Sieci komunikujemy się jedynie z wybraną grupą, która w pewien sposób jest już pozainternetowo zdeterminowana do korzystania z naszej oferty (np. poprzez edukację formalną).

Historyk czy instytucja naukowa lub instytucja kultury angażując się w projekty internetowe musi mieć też świadomość ograniczeń wynikających z infrastruktury technicznej, będącej tłem podejmowanych działań. Dostępnych jest wiele darmowych i otwartych rozwiązań związanych z gromadzeniem i udostępnianiem treści czy komunikowaniem się z użytkownikami – małym kosztem, bez specjalistycznej wiedzy informatycznej można przygotować własną stronę,



udostępniać własne publikacje czy nawet tworzyć repozytorium. Główne kanały komunikacji, uczestnictwo w których staje się niezbędne ze względu na możliwość dotarcia do odbiorców, pozostają jednak poza naszą kontrolą. Dobrym przykładem jest Facebook. Wiele instytucji naukowych, archiwów, bibliotek czy muzeów zakłada w tym serwisie reprezentujące ich strony a naukowcy budują swoje profile, chcąc skuteczniej dotrzeć do potencjalnych czytelników swoich prac. Facebook jako centralnie zarządzane przedsiębiorstwo nastawione na zysk nie stanowi jednak neutralnej infrastruktury dla inicjatyw naukowych czy edukacyjnych. W każdej chwili, bez żadnego ostrzeżenia i żadnych możliwości odwołania się od decyzji prowadzona od miesięcy a nawet lat strona lub profil mogą zostać zablokowane lub usunięte.

Na efektywność działań edukacyjnych i naukowych w Internecie wpływa także oprogramowanie. Przykładowo, działający na Facebooku algorytm EdgeRank sprawia, że tylko część z informacji publikowanych w poszczególnych stronach czy profilach jest widoczne dla odbiorców; dziś dla tzw. stron fanowskich to około 12 proc. wszystkich odbiorców (dane serwisu Edgerankchecker.com). Podobnie jak w przypadku wyszukiwarki Google, sama obecność w systemie (posiadanie strony, indeksacja zasobów) nie oznacza jeszcze bycia widocznym dla użytkowników. Gwarantowanie tej widoczności jest już jednak usługą, jaką w ramach swojego modelu biznesowego sprzedają te obie globalne firmy. Dopiero kupno reklamy pozwala na skuteczne dotarcie do odbiorców. Widać tu wyraźnie, że wyzwaniem jest sama ekonomia aktywności online. Przygotowywana pracownice notka (wpis) na blogu naukowym promowana na Facebooku jest tam widoczna jedynie przez kilka chwil, a potem ginie w zalewie kolejnych informacji publikowanych nieustannie w streamie. W takich sytuacjach racjonalne jest zastanowienie się, czy energia i czas przeznaczany na pracę online, z reguły przecież wydzielany już z czasu wolnego po pracy naukowej czy dydaktycznej, nie jest jednak marnowany.

Interfejsy, za pomocą których korzystamy z Internetu i oprogramowanie, które dekoduje zapisane w postaci liczbowej treści i dane także wpływają na to, jaka forma komunikacji staje się najbardziej efektywna. Obecny w języku Internetu akronim TL;DR (*too long; didn't read*) wskazuje właśnie na jedną z takich prawidłowości: teksty długie, pozbawione akapitów i czytelnego podziału są trudne w odbiorze na ekranie komputera, a zwłaszcza na ekranach urządzeń mobilnych (tabletów czy smartfonów). Efektywność tradycyjnego publikowania naukowego czy nawet działalności popularyzatorskiej i publicystycznej np.

w prasie nie jest aż tak bardzo uzależniona od objętości przekazu, jednak publikowanie online – właśnie ze względu na interfejsy, przez które korzystamy z treści – nakłada już duże ograniczenia. Oczywiście nie oznacza to, że jedyną formą publikowania powinny być abstrakty i streszczenia albo nagrania multimedialne – wiele czasopism naukowych z powodzeniem funkcjonuje w modelu OpenAccess. Projektując działania popularyzatorskie należy jednak wziąć pod uwagę ten problem.

Declan Butler w swoim komentarzu na łamach *Nature* poświęconym aktywności naukowców online zwraca uwagę na to, że Internet w swoich początkach (1989 rok) był projektem mającym usprawnić komunikację naukową w CERN (Europejska Organizacja Badań Jądrowych CERN). W toku swojej dynamicznej ewolucji stał się w pewnym momencie globalną biblioteką, aby w połowie pierwszej dekady XXI wieku m.in. wraz z upowszechnieniem blogów ponownie pełnić funkcję narzędzia komunikacji naukowej (Butler, 2005). Przez ten cały okres zasięg społeczny i natura Sieci zmieniła się radykalnie i środowisko akademickie musi dziś nauczyć się w niej funkcjonować bez żadnej taryfy ulgowej. Butler wskazuje na blogi jak na dodatkowy element systemu komunikacji naukowej, ułatwiający nie tylko porządkowanie nadmiaru informacji, ale też wymianę idei i odkrywanie nowego kontekstu własnych zainteresowań badawczych (np. przez lekturę blogów z zupełnie innych dziedzin wiedzy). Pokazuje jednak, że wielu autorów blogów naukowych publikuje je anonimowo, obawiając się, że tego typu aktywność byłaby źle odebrana w ich macierzystych uczelniach. Z jednej strony mamy więc nowe rozwiązania komunikacyjne zwiększające efektywność wymiany informacji o badaniach, konferencjach i publikacjach naukowych, z drugiej konserwatyzm środowiska akademickiego i wspomnianą już przez Ayersa *nieprzejrzystość*. Publiczne i bezpośrednie informowanie o aktywności badawczej czy otwarte konsultowanie własnych koncepcji uznawane jest wciąż za działanie nieracjonalne, zmniejszające szanse w rywalizacji na rynku akademickim poprzez łatwe *oddawanie* innym swojej pracy.

Internet może stanowić także infrastrukturę dla projektów naukowych wychodzących poza standardowy model pracy historyka, polegający na tworzeniu krytycznych narracji o przeszłości. Efektami pracy naukowej mogą być baza danych, wizualizacja danych historycznych czy nawet program komputerowy, np. przeliczający dawne jednostki miary, a nie tylko klasyczny artykuł w recenzowanym czasopiśmie czy monografia. Pozostaje oczywiście pytanie, jak taka forma aktywności naukowej zauważana i oceniana byłaby przez system akademicki.

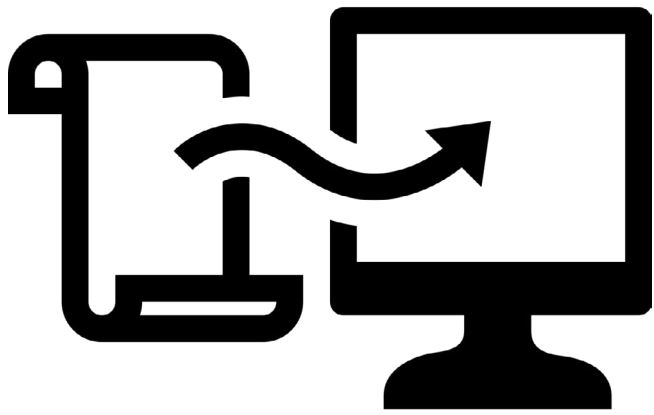
Próbując pisać o potencjale Sieci wobec nauki i edukacji historycznej, warto przyjąć postawę Daniela J. Cohena i Roy'a Rosenzweiga, którzy we wstępie swojego podręcznika do historii cyfrowej zachowują dystans wobec bezkrytycznych entuzjastów Internetu, nie bojąc się jednak wprost świadczyć o jego zaletach dla uprawianej przez siebie dziedziny. Kiedy ceną za różnorodność źródeł internetowych czy bezpośredni dostęp do informacji i zasobów historycznych online staje się obawa o ich autentyczność i jakość, historyk nie powinien dezertować i zamykać się we względnie bezpiecznej przestrzeni akademickiej. Jak piszą, naukowcy muszą skonfrontować się z problemem jakości, trwałości, czytelności czy dostępności zasobów i wiedzy historycznej online, a nie pozostawiać to wyzwanie korporacjom medialnym i firmom IT, ustawodawcom czy nawet instytucjom takim jak archiwa czy biblioteki. Dla jednych historyków odpowiedzią na to wyzwanie będzie publikowanie swoich prac w otwartym dostępie, dla innych – aktywny udział w oddolnym projekcie nauki obywatelskiej. Kluczem jest tu jednak aktywna, otwarta, ale też krytyczna postawa (Cohen i Rosenzweig, 2006, s. 13).

Na koniec warto zwrócić jeszcze uwagę na fakt, że Internet stanowi nie tylko nową przestrzeń komunikacji naukowej czy upowszechniania badań, ale także sam w sobie może być przedmiotem badań historycznych jako sfera społecznych interakcji i kulturowej aktywności, którą można analizować z perspektywy czasu. Wątek ten rozwijam w jednym z kolejnych rozdziałów, podejmując temat dziedzictwa cyfrowego.

## BIBLIOGRAFIA

- Ayers, E.L.** (2004). *The academic culture and the IT culture: Their effect on teaching and scholarship*, *Educause Review*, November/December 2004, s. 54.  
Pobrano z: <http://www.educause.edu/pub/er/erm04/erm0462.asp>
- Butler, D.** (2005). Science in the web age: Joint efforts. *Nature*, 438(7068), 548-549.
- Cheal, C.** (2007), Second Life: hype or hyperlearning? *On the Horizon*, 15(4), 204-210.
- Cohen, D. J.** (2004). History and the Second Decade of the Web, *Rethinking History*, (8)2, 293-301.

- Cohen, D.J., Rosenzweig, R.** (2006). *Digital history: a guide to gathering, preserving, and presenting the past on the Web*. University of Pennsylvania Press, Pobrano z: <http://chnm.gmu.edu/digitalhistory/introduction/>
- Scholz, T.** (2008). Market ideology and the myths of Web 2.0. *First Monday*, 13(3). Pobrano z <http://firstmonday.org/ojs/index.php/fm/article/viewArticle/2138/1945>
- O'Reilly, T.** (2005). What is web 2.0. Pobrano z: <http://oreilly.com/web2/archive/what-is-web-20.html>
- Prinke, T.** (2000). *Fontes ex machina: komputerowa analiza źródeł historycznych*, Poznań: Polska Akademia Nauk.
- Warburton, S.** (2009). Second Life in higher education: Assessing the potential for and the barriers to deploying virtual worlds in learning and teaching, *British Journal of Educational Technology*, 40(3), 414-426.



# Digitalizacja dziedzictwa i cechy formatu cyfrowego

Jednym z najważniejszych dla historyka pojęć i procesów związanych ze współczesnym Internetem jest digitalizacja (inaczej: cyfryzacja), czyli konwertowanie treści z nośników analogowych do postaci cyfrowej. Chociaż nie wszystkie zasoby źródłowe są dostępne online, programy masowej digitalizacji realizowane przez biblioteki i archiwa w radykalny sposób ułatwiają korzystanie ze zbiorów historycznych, tradycyjnie często trudno dostępnych lub niedostępnych wcale. Przed prezentacją przykładowych repozytoriów i nowych możliwości badawczych z nimi związanych warto zatrzymać się jeszcze na poziomie technicznych podstaw formatu cyfrowego. Pozwoli to zauważyć nie tylko pozytywny potencjał wynikający z digitalizacji i korzystania w pracy naukowej ze źródeł w wersji cyfrowej, ale też pewne wyzwania z tym związane. Kluczem do charakterystyki postaci cyfrowej mogą być wybrane pojęcia z zestawu cech nowych mediów, zaproponowanego przez Lva Manovicha w książce *Język nowych mediów* (Manovich, 2006, s. 91-118).

Podstawową cechą obiektów nowych mediów, według Manovicha, jest reprezentacja numeryczna. Skany dokumentów historycznych publikowane w Internecie są liczbami zapisanymi w postaci cyfrowej, które, aby mogły być zrozumiałe dla człowieka, muszą zostać odczytane za pomocą odpowiedniego oprogramowania. Dzięki liczbowej postaci są one w nieograniczony sposób kopiowane bez utraty jakości i w swobodny sposób rozpowszechniane za pomocą różnych internetowych kanałów komunikacji. Co więcej, mogą być także automatycznie przetwarzane – przykładowo, w obszernym zestawie fotografii

historycznych w formie cyfrowej można za pomocą odpowiedniego algorytmu poprawić kontrast czy nawet wyszukać ludzkie twarze. Takim rozwiązaniem jest oprogramowanie Open Computer Vision, zbudowane przez Gado Images. Pozwala ono przeszukiwać zasoby zdigitalizowanych fotografii historycznych w poszukiwaniu twarzy podobnych do tej, jaką ustawia się za wzór dla algorytmu. Przyspiesza to analizowanie zbiorów zdjęć, np. w ramach badań historii rodzinnej czy przy poszukiwaniu nieznanymi wizerunków znanych osób.

Kolejną ważną cechą nowych mediów jest modularność. Manovich opisuje tę cechę, odwołując się do konstrukcji zwykłej strony WWW, która składa się z kilku elementów: tekstu (treści merytorycznej), znaczników HTML (metatagów), odnośników, linków do plików graficznych czy kaskadowych arkuszy stylów (ang. *Cascading Style Sheets*, CSS), definiujących takie cechy dokumentu, jak rodzaj czcionki, kolor tła itp. Obiekty nowych mediów są modularne i składają się z wielu elementów mogących funkcjonować niezależnie od siebie. Dzięki temu pojawia się chociażby możliwość stworzenia wizualizacji, zbudowanej za sprawą nałożenia na siebie w programie graficznym różnych zdigitalizowanych map czy planów tego samego terenu, w tej samej skali, jednak powstałych w różnym czasie i ukazujących zmiany, np. w przestrzeni urbanistycznej.

Modularny charakter nowych mediów sprawia, że ich obiekty bardzo łatwo tracą oryginalny kontekst. W tradycyjnej drukowanej publikacji historycznej fotografia archiwalna przytaczana jest zawsze w czytelnym i stałym kontekście – autor umieszcza ją w odpowiedniej, merytorycznie uzasadnionej relacji do tekstu. W przypadku skanu kontekst publikacji może radykalnie się zmieniać, w konsekwencji czego pojawia się problem niemożności odczytania jej oryginalnego znaczenia. Łączy się on z inną cechą nowych mediów – wariacyjnością. Obiekt nowych mediów może istnieć w wielu różnych wersjach, w odróżnieniu od obiektów analogowych, utrwalanych fizycznie w materiale z zachowaną na stałe strukturą. Dobrym przykładem wariacyjności mogą być rozmaite cyfrowe formy narracyjnych źródeł historycznych – ten sam tekst kroniki może być dostępny dla historyka w formie skanów różnej jakości, transkrybowanego, przeszukiwalnego dokumentu tekstowego, dokumentu XML (Rozszerzalny Język Znaczników, ang. *Extensible Markup Language*) z odpowiednią strukturą semantyczną, czyli tagami informującymi o znaczeniu poszczególnych fragmentów tekstu, czy też bazy danych. Ta ostatnia zawiera odpowiednio ułożone poszczególne słowa, których wzajemnymi interakcjami zarządzać mogą rozmaite algorytmy i które wyświetlane być mogą za pomocą różnych interfejsów.

Charakterystykę nowych mediów zakończyć można na cesze, jaką jest transkodowanie. Według Manovicha nowe media składają się z dwóch różnych warstw: komputerowej i kulturowej. Historyk, korzystający z cyfrowych edycji źródeł średniowiecznych dostępnych online, rozpoznawać będzie ich gatunki, analizować charakterystyczne pojęcia czy badać kompozycję – jednak pod tą warstwą wiedzy kulturowej istnieć będzie porządek komputerowych mechanizmów: algorytmów pozwalających na wyświetlanie kolejnych stron, zarządzających zapytaniami do bazy danych, kodu odpowiadającego za wyświetlanie treści czy też standardowych funkcji serwera, pozwalających na publikowanie treści w Internecie. Istotny jest tutaj problem świadomości naukowca co do istnienia tego poziomu organizacji danych, szczególnie w przypadku zespołów historyków i informatyków pracujących nad cyfrowymi projektami naukowymi. Założenia merytoryczne projektowanej bazy danych czy interaktywnej wizualizacji muszą być wówczas „przetłumaczone” na język techniczny. Rafał T. Prinke (1997, s. 50) podaje tutaj bardzo czytelny przykład:

[...] zmienna typu „data” dla programisty składa się z trzech elementów: dnia, miesiąca i roku, podczas gdy dla historyka może mieć praktycznie nieograniczony zakres innych formatów, począwszy od niepełnych formatów standardowych („marzec 1523”), poprzez różne określenia niepewności („zapewne około 1623”), wieloczłonowe zakresy („między 23 lipca a 6 sierpnia przed 1418”), aż po określenia całkowicie relatywne („piątek po Wielkanocy między 1312 a 1323”).

Autor zwraca uwagę, że platformą komunikacji między historykiem a programistą jest umiejętność konstruowania algorytmów – logicznych schematów dochodzenia do odpowiedzi na postawione pytanie badawcze.

David Thomas i Valerie Johnson, doceniając potencjał digitalizacji dla rozwoju nauki historycznej, zwracają jednak uwagę na negatywne aspekty tego zjawiska (Thomas i Johnson, 2012, s. 180). Projekty masowej digitalizacji i udostępniania online zasobów archiwów i bibliotek nie są zazwyczaj wzajemnie skoordynowane, a wybór konkretnych zasobów zależy w dużej mierze od ich rodzaju czy statusu prawnoautorskiego. W rezultacie historyk otrzymuje eklektyczny zbiór kolekcji cyfrowych i nie może być pewien, czy oczekiwane przez niego zasoby zostaną wkrótce udostępnione w Internecie. Autorzy zastanawiają się też, czy brak dostępu do cyfrowych wersji źródeł i literatury może wpłynąć



negatywnie na poziom ich wykorzystania, ponieważ to, co nie jest osiągalne online, może zostać w łatwy sposób pominięte. Przy ocenie efektów digitalizacji warto również zwrócić uwagę na problem wierności odwzorowania oryginału. O ile w przypadku dokumentów tekstowych można przygotować dobrej jakości płaski skan, o tyle przygotowanie dobrych skanów artefaktów, takich jak medale, pieczęcie, rzeźby itp., jest zdecydowanie trudniejsze.

Dyskusja o naturze zbiorów cyfrowych z pozoru tylko wydaje się zajęciem przeznaczonym głównie dla wąskiego grona teoretyków. Odpowiednie rozpoznanie potencjału, jak i ograniczeń formatu cyfrowego staje się jednak kluczową sprawą dla historyków, archiwistów, bibliotekarzy, którzy w swojej pracy coraz częściej polegać będą na tego typu zasobach.

Na ważny problem związany z digitalizacją materiałów historycznych zwrócił uwagę Marek Kaźmierczak, pisząc o dostępności w Internecie treści związanych z tematem Zagłady. Według niego stają się one informacyjnym „surowcem”, przetwarzanym niemal automatycznie w ramach dostępnego oprogramowania, np. Google czy Facebooka (Kaźmierczak, 2012, s. 241). W konsekwencji mogą być one w łatwy sposób dekontekstualizowane – oderwane od oryginalnych nośników tracą pierwotne znaczenia, stają się symbolami. Jak pisze Kaźmierczak, *dla użytkownika Internetu obraz przedstawiający obóz zagłady może okazać się ekwiwalentem tego miejsca i tym, co zostanie zapamiętane lub utożsamione z Holocaustem* (s. 242). Konkretnie historyczne zdjęcie konkretnego



Digitalizacja książek w siedzibie fundacji Internet Archive.

Fot. liz west (CC BY), <http://www.flickr.com/photos/calliope/6883457517/>

obożu, po digitalizacji i udostępnieniu go online, może stać się ogólną ilustracją Zagłady. Przestaje być ważna data wykonania, jej autor, kontekst, w jakim powstała. Doskonale widać ten proces na przykładzie słynnego zdjęcia z raportu Jürgena Stroopa, pokazującego żydowskiego chłopca, który z podniesionymi rękoma wychodzi z budynku, otoczony niemieckimi żołnierzami pacyfikującymi tereny warszawskiego getta. Fotografia ta stała się ogólnym symbolem Zagłady i jest reprodukowana na niezliczonych stronach internetowych jako ilustracja tego tematu, zupełnie już oderwana od oryginalnego, szczegółowego kontekstu (powstania w getcie warszawskim, albumu Stroopa i oczywiście niezidentyfikowanego chłopca). Dużym zadaniem edukacji historycznej wobec szerokiej dostępności w Internecie wizualnych materiałów historycznych jest uczenie powrotu do faktu, przypominanie oryginalnego kontekstu reprodukowanych masowo obrazów oraz wspieranie krytycznego namysłu nad ich ponownym użyciem. W przypadku fotografii przemocy ma to szczególne znaczenie ze względów etycznych, o czym przypomina Janina Struk, a co można z powodzeniem odnieść także do Internetu:

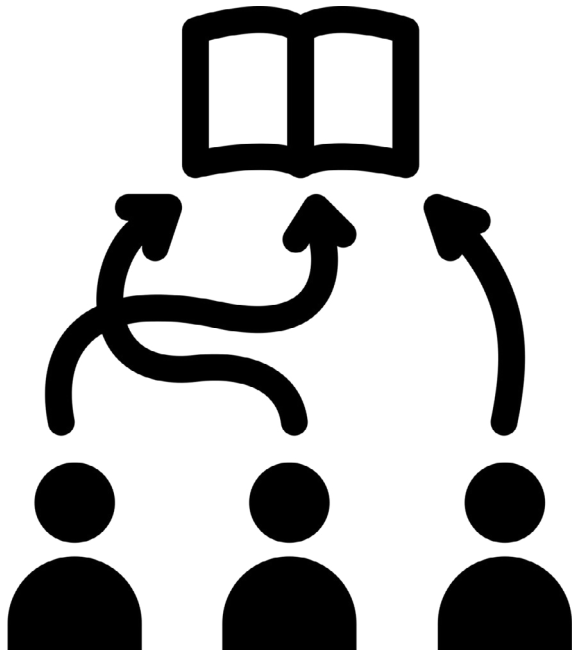
Większość fotografii prezentowanych na wystawach jako dowód zbrodni Holokaustu została zrobiona przez nazistów jako dowód ich potęgi. Można powiedzieć, że nazistowska obsesja na punkcie dokumentacji wizualnej przyczyniła się po części do upamiętnienia Holokaustu. [...] Naziści robili fotografie swoich ofiar, aby je upokorzyć i poniżyć. Czy nie jesteśmy z nimi w zmowie, kiedy sami wystawiamy je w muzeach? Czy mamy prawo pokazywać ludzi w ostatnich chwilach przed śmiercią, by osiągnąć cel propagandowy, wszystko jedno jaki? [...] Czy nie powinniśmy przywrócić tym fotografiom i nieszczęśliwym ludziom, którzy się na nich znaleźli, rangi dokumentów historycznych, zamiast wystawiać je na pokaz? (Struk, 2007, s. 282-283)

## BIBLIOGRAFIA

- Manovich, L.** (2006). *Język nowych mediów*. Warszawa: Wydawnictwa Akademickie i Profesjonalne.
- Każmierczak, M.** (2012). *Auschwitz w Internecie. Przedstawienia Holokaustu w kulturze popularnej*. Poznań: Wydawnictwo Naukowe Uniwersytetu im. Adama Mickiewicza.

- Prinke, T.** (1997). Techniki, metody i standardy komputerowe w naukach humanistycznych. W: Ryszewski B. (red.), *Metody komputerowe w badaniach i nauczaniu historii: materiały II Symposium Polskiego Oddziału Association for History and Computing, Komisji Metod Komputerowych Polskiego Towarzystwa Historycznego, Poznań 1 i 2 grudnia 1995 r.* Toruń: Wydawnictwo Uniwersytetu Mikołaja Kopernika.
- Struk, J.** (2007). *Holokaust w fotografiach. Interpretacje dowodów.* Warszawa: Prószyński i S-ka.
- Thomas, D., Johnson, V.** (2012). New universes or black holes? Does Digital change anything? W: T. Weller (red.), *History in the Digital Age* (s. 174-193). London, New York: Routledge.





# Wykorzystywanie repozytoriów cyfrowych i archiwalnych baz danych

Historia cyfrowa nie istnieje bez źródeł w formie cyfrowej. W poprzednim rozdziale, powołując się na zaprezentowany przez Lwa Manovicha zestaw pojęć charakteryzujących nowe media i cechy formatu cyfrowego, starałem się pokazać, jak dzięki digitalizacji zmienia się postać informacji zapisywanych oryginalnie na analogowych nośnikach i jak w związku z tym pojawiają się nowe możliwości ich przetwarzania i wykorzystania. W ślad za tymi zjawiskami idą także zmiany w funkcjonowaniu instytucji, takich jak archiwum czy biblioteka, chociaż nie można twierdzić, że mają one charakter rewolucyjny i podważają tradycyjne formy pracy archiwalnej czy bibliotecznej. Instytucje te raczej starają się płynnie dostosowywać się do nowych realiów cyfrowej kultury i ostrożnie wykorzystywać nowe możliwości.

Trudno zredukować wartość korzystania z zasobów źródłowych udostępnianych w repozytoriach internetowych wyłącznie do kwestii łatwiejszej i szybszej dostępności czy skuteczniejszego wyszukiwania. Archiwa i biblioteki cyfrowe udostępniają bowiem często narzędzia umożliwiające zupełnie nowe podejście do pracy badawczej (przynajmniej na poziomie gromadzenia bazy źródłowej). Oczywiście znów nie ma tu mowy o rewolucji, kwestionującej podstawy warsztatu historyka. To, że samo choćby przeszukiwanie zbiorów odbywać się może bez bezpośredniego wsparcia archiwisty czy bibliotekarza, stanowi już jednak wyraźną zmianę w stosunku do starego modelu pracy badawczej. Można tu mówić o perspektywie kryzysu władzy archiwum nad definiowaniem porządku

korzystania ze zbiorów czy nawet o przejawie końca pewnego etosu pracy w archiwum, opartego na doświadczeniu bezpośredniego kontaktu z historycznymi dokumentami, teraz dostępnymi online (Hitchcock, 2008, s. 81–83). Zmienia się środowisko pracy historyka:

Jesteśmy w centrum procesu przekształcania się samej natury manuskryptów i źródeł pierwotnych, na których polegamy. Digitalizacja, nowe ułatwienia w wyszukiwaniu, nowe sposoby prezentowania i łączenia informacji fundamentalnie zmieniają naturę archiwum – to, czym jest i jak jest wykorzystywane, i jak my, jako historycy, go doświadczamy (Hitchcock, 2008, s. 83).

Skoro społeczna legitymizacja pracy historyków opiera się na odwołaniu do archiwum i zawartych w nim źródeł, to wobec zmian wywołanych digitalizacją powinniśmy – zdaniem Tima Hitchcocka – przemyśleć ją na nowo. Oczywiście tego typu interpretację można krytykować, z uwagi na fakt, że mimo postępów digitalizacji wciąż stosunkowo niewielka część zbiorów archiwalnych dostępna jest w Internecie i trudno uznać, że tradycyjny model pracy w archiwum odchodzi już do lamusa. Jednak nawet jeśli nie zgodzimy się z twierdzeniami głoszącymi fundamentalne przemiany instytucji archiwum, trudno zaprzeczyć temu, że jego nowa, cyfrowa forma pozwala historykowi na stosowanie w pracy badawczej nowych narzędzi.

Jedną z najważniejszych i najbardziej użytecznych nowości są programy pozwalające na sprawne zarządzanie budowanym dla celów badawczych zbiorem źródeł i literatury. Warto wymienić tu Zotero – program rozwijany przez Center for History and New Media (CHNM), który można zainstalować nawet bezpośrednio w przeglądarce. Jego użytkownicy w trakcie przeglądania repozytoriów archiwalnych czy bibliotecznych mogą budować własną, roboczą bazę danych: bibliograficzną czy źródłową. Program pobiera nie tylko adresy URL z odpowiednimi danymi bibliograficznymi (tytuł zasobu, opis, słowa kluczowe itp.), ale także same pliki (dostępne online teksty naukowe czy pliki ze skanami). Gromadzone dane przechowywane być mogą na dysku twardym komputera lub na serwerach udostępnianych przez CHNM.

Archiwalne i biblioteczne repozytoria cyfrowe publikujące dane (skany) i metadane (opisy) mogą komunikować się między sobą, dzięki czemu możliwe jest tworzenie rozbudowanych serwisów, agregujących informacje z różnych źródeł. Serwisy te udostępniają użytkownikom jeden interfejs służący do

przeszukiwania wielu archiwów czy bibliotek internetowych. Przykładem tego typu agregatorów jest katalog i wyszukiwarka Federacji Bibliotek Cyfrowych, gdzie w jednym oknie przeszukiwać możemy zasoby kilkudziesięciu działających niezależnie od siebie polskich repozytoriów. Podobnie – tylko na poziomie europejskim – funkcjonuje Europeana, europejska biblioteka cyfrowa, do której trafiają informacje o zasobach udostępnianych w bibliotekach i archiwach cyfrowych z europejskich krajów, w tym także z Polski. W grudniu 2013 roku FBC agregowała ponad 1.6 mln obiektów, a Europeana – ponad 30 mln. Warto wiedzieć, że zawarte są tam informacje o obiektach, które nie zawsze są odpowiednio widoczne w popularnych wyszukiwarkach, takich jak Google. Do przesyłania informacji pomiędzy repozytoriami wykorzystywany jest protokół OAI-PMH (ang. *Protocol for Metadata Harvest*), będący otwartym standardem komunikacji między bibliotekami i archiwami cyfrowymi.

Opisy udostępnianych materiałów archiwalnych mogą zatem podlegać automatycznemu przetwarzaniu, w efekcie czego otwierają się nowe możliwości twórczego korzystania z zasobów bibliotek i archiwów cyfrowych. Od 2011 roku w całej Europie organizowane są warsztaty dla programistów, w ramach których tworzą oni aplikacje, korzystające z metadanych i danych agregowanych przez Europeanę. Dzięki publicznemu API, czyli – w uproszczeniu – systemowi eksportowania danych z serwisu internetowego, powstają gry, wyszukiwarki i aplikacje mobilne na telefony komórkowe oraz na inne urządzenia przenośne. Dobrym przykładem tego jest testowa polska aplikacja Art4Europe, która powstała w ramach konkursu Hack4Europe, organizowanego przez Komisję Europejską i Europeanę w 2011 roku. Aplikacja zainstalowana na telefonie komórkowym pozwala zidentyfikować fotografowany w muzeum obraz na podstawie zdjęcia zrobionego przez użytkownika oraz informacji przechowywanych w Europeanie. Chociaż narzędzia te zazwyczaj nie mają zastosowań ściśle naukowych, mogą stać się inspiracją dla badaczy, którzy chcieliby – wykorzystując API – w jakiś sposób automatyzować czy ułatwić sobie pracę nad analizą zbiorów udostępnianych przez archiwalne czy biblioteczne repozytoria. Obok technicznych kwestii związanych z publikowaniem metadanych dla swobodnego przetwarzania i budowania aplikacji ważny jest także ich status prawnoautorski – największe instytucje pamięci promują tu standard otwartych danych (*open data*), zakładający brak ograniczeń w rozpowszechnianiu, przetwarzaniu i komercyjnym wykorzystywaniu udostępnianych przez nie informacji o zasobach cyfrowych.



Problemem związanym z korzystaniem ze źródeł dostępnych online jest jednak nie tylko ich autentyczność czy poprawność opisów. W przypadku specjalistycznych repozytoriów, takich jak francuska Gallica czy dMGH (*Monumenta Germaniae Historica*), można mieć zresztą zaufanie co do jakości informacji opisujących zeskanowane zbiory. Wyzwaniem dla badacza natomiast jest zróżnicowanie schematów metadanych. Widać to wyraźnie choćby w kontekście polskich repozytoriów: biblioteki regionalne, akademickie i biblioteki organizacji pozarządowych publikują metadane w schematach wywodzących się ze schematu Dublin Core (DC), z którego korzystają także Federacja Bibliotek Cyfrowych i Europeana. Ten oryginalnie biblioteczny standard wykorzystywany jest jednak także do opisu dokumentów archiwalnych. Często zdarza się nawet, że w ramach projektów jednej instytucji wykorzystywane są różne modele metadanych, zależne od charakteru udostępnianych treści. Przykładowo, archiwalia publikowane przez Narodowe Archiwum Cyfrowe w serwisie Szukajwarchiwach.pl opisywane są zupełnie inaczej, niż zbiory w serwisie fotograficznym tejże instytucji.

Wybór systemu metadanych zależy też może od charakteru repozytorium. Od 2008 roku archiwa i biblioteki biorące udział w projekcie Flickr Commons publikują swoje zdigitalizowane zbiory, liczące ponad 250 tys. fotografii, w popularnym serwisie fotograficznym Flickr.com, a informacje o nich udostępniają za pomocą rozmaitych schematów. Co więcej, użytkownicy Flickr.com sami opisują udostępniane zdjęcia historyczne za pomocą wybranych przez siebie słów kluczowych. Popularyzatorski charakter tej inicjatywy oraz fakt, że jej infrastrukturę stanowi popularny serwis fotograficzny, sprawiły, że nie zdecydowano się na wykorzystanie żadnego spośród standardów metadanych używanego w profesjonalnych archiwach. Głównym wymogiem wobec instytucji uczestniczących w projekcie (dziś jest ich 78) było udostępnianie zbiorów, co do których wygasły prawa autorskie lub co do których zezwalają na dowolne wykorzystanie bez żadnych warunków. W czasie końcowych prac nad niniejszą książką zasoby Flickr Commons zostały wzbogacone przez British Library milionem skanów pochodzących z XVII, XVIII i XIX wieku. Co więcej, w 2014 roku biblioteka ma udostępnić aplikację ułatwiającą społecznościowe opisywanie umieszczonych na Flickr.com zbiorów.

Ważnym czynnikiem, który wpływa na pracę historyka z wykorzystaniem Internetu, jest także jakość udostępnianych skanów i ich dostępność. Ponieważ nie istnieje żaden powszechnie stosowany standard, związany bądź

z rozdzielczością, bądź z formatem pliku, za pomocą którego publikuje się treść skanu, poszczególne repozytoria definiują własne strategie, zależne wprost od polityki udostępniania, poziomu otwartości instytucji, możliwości technicznych (wydajności serwerów i łącz) czy wymogów instytucji finansujących digitalizację. W rezultacie historyk znaleźć może w Internecie doskonale przygotowane skany w wysokiej rozdzielczości, pozwalającej na wykorzystanie ich np. w druku, ale też pliki o słabej jakości, o funkcji jedynie poglądowej. Wiele instytucji dodaje do skanów znaki wodne, które w niektórych przypadkach mogą być przeszkodą w odczytywaniu treści dokumentów czy też zasłaniać ważne szczegóły fotografii.

Warto w tym miejscu zadać pytanie o naukowy potencjał repozytoriów, takich jak Federacja Bibliotek Cyfrowych, Flickr Commons czy Europeana. Udostępniane tam materiały źródłowe mają formę skanu z odpowiednim zestawem metadanych. O ile w przypadku fotografii historycznych taki model udostępniania wydaje się optymalny, o tyle skorzystanie z dokumentów źródłowych rozpowszechnianych w ten sposób może być dla historyka dużym problemem. Reprezentacja graficzna źródła (czyli skan) nie pozwala na przeszukiwanie treści czy umieszczanie odpowiednich adnotacji do tekstu, informujących o jego istotnych cechach i ułatwiających jego analizę.

W dyskusjach na temat historii cyfrowej problem ten podejmowany jest także w odniesieniu do samej natury pracy historyka i miejsca, jakie zajmują w jego badaniach źródła historyczne. Pojawiają się opinie, że profesja ta nie polega na analizie źródeł, ale na wytwarzaniu krytycznej wiedzy na temat dawnych społeczeństw. Zgodnie z takim podejściem, źródła – także te udostępniane online – pełnią jedynie pomocniczą rolę, niektóre repozytoria natomiast mogą w istotny sposób przydać się w pracy badawczej, ponieważ np. pozwalają na skorzystanie z materiałów ikonograficznych przy opracowywaniu historii danej miejscowości. Z drugiej strony, są jednak i tacy historycy, dla których praca ze źródłem jest podstawą procesu badawczego – rzeczywistość pozaźródłowa interesuje ich wyłącznie wtedy, jeśli pozwala na odczytanie znaczenia dokumentu. Nie ma tu mowy o budowaniu ogólnej narracji, jest za to praca nad poprawnym odczytaniem zawartego w źródle przekazu i ewentualnie odniesieniem go do wiedzy pozaźródłowej (Prescott, 2008, s. 12–13).

Oferta cyfrowych repozytoriów zwiększa się z roku na rok coraz bardziej. Warto jednak zapytać o to, czy historycy rzeczywiście chcą i potrafią z niej korzystać, a także czy jest ona dla nich naprawdę użyteczna. Jak pokazują Valerie Harris i Peter Hepburn, zwiększony dostęp do zasobów ikonograficznych nie przekłada się wprost na ich przytaczanie w historycznych artykułach naukowych

(Harris i Hepburn, 2013). Autorzy po przeanalizowaniu 1366 artykułów z lat 2000–2009, pochodzących z kilku głównych anglojęzycznych recenzowanych czasopism historycznych (wydawanych w tradycyjnym modelu), oraz z kilku dostępnych wyłącznie online wykazali, że pomimo rozwoju bibliotek i archiwów cyfrowych nie wzrósł poziom wykorzystania materiałów wizualnych w artykułach naukowych. Okazuje się, że oferta projektów digitalizacyjnych nie zawsze musi odpowiadać zapotrzebowaniom historyków, zainteresowanych wciąż głównie źródłami tekstowymi. Problemem może być też brak wsparcia ze strony bibliotekarzy i archiwistów cyfrowych w zakresie wyszukiwania i wykorzystywania zasobów wizualnych w historycznych pracach naukowych. Wobec ogromnej liczby materiałów dostępnych dziś online kwestia ich odpowiedniej promocji i prezentacji – skierowanej także do środowiska naukowego – staje się kluczowa.

Edukowanie historyków w zakresie efektywnego korzystania z zasobów cyfrowych opierać się powinno na zrozumieniu sposobów, w jaki wyszukują oni online potrzebne w swoich badaniach informacje i materiały. Próbą przygotowania takiej diagnozy był *Primarily History Project*, w ramach którego w latach 2001–2002 badano kompetencje historyków w zakresie gromadzenia źródeł oraz dydaktyki akademickiej w tym zakresie (Tibbo, 2003). Odpowiednie ankiety rozesłano m.in. do 700 amerykańskich historyków, uzyskując odpowiedzi od nieco ponad jednej trzeciej adresatów. Badanie pokazało m.in. nieufność historyków do wyszukiwarek w repozytoriach i katalogach. Preferowali oni bardziej bezpośredni kontakt z archiwistami w celu uzyskania potrzebnych informacji czy materiałów. Przy kwestii korzystania z zasobów internetowych ujawnił się też podział wiekowy, a także zależność od specyfiki tematu badawczego. Trudno dziś określić, jak sposoby wyszukiwania źródeł w badaniach historycznych zmieniły się przez 10 lat od publikacji wyników *Primarily History Project*. Na pewno jednak przesłanie tego badania, ukazującego konieczność edukacji historyków w zakresie korzystania z narzędzi internetowych czy dbania o użyteczność repozytoriów cyfrowych, jest wciąż aktualne. Dziś w niektórych specjalnościach historycznych – np. mediewistyce – korzystanie z archiwów cyfrowych staje się standardem i obowiązkowym elementem dydaktyki akademickiej. Dzięki digitalizacji badacze mogą analizować wybrane rękopisy w szerszym niż do tej pory kontekście, a nawet współpracować ze sobą przy opracowywaniu poszczególnych tekstów (jak np. w projekcie T-PEN – *Transcription for Paleographical and Editorial Notation* czy w projekcie MOM-CA – *Monasterium Collaborative Archive* istniejącego w ramach serwisu *Monasterium.net*).

Należy pamiętać także, że obok rozwijających się wciąż repozytoriów cyfrowych publikujących skany od dawna dostępne są online także rozmaite pomoce archiwalne oraz bazy biograficzne. Zwrócić uwagę warto na Zintegrowany System Informacji Archiwalnej ZoSIA, który pozwala archiwistom na prowadzenie ewidencji i inwentaryzacji materiałów archiwalnych, a użytkownikom na łatwy dostęp do informacji (i często także do skanów) za pomocą serwisu [szukajwarchiwach.pl](http://szukajwarchiwach.pl). ZoSIA używana jest dziś w 15 polskich instytucjach. Obok tego w badaniach wykorzystać można inne polskie bazy, których ograniczony wybór przedstawia poniższa tabela. Szerszy i obejmujący instytucje z całej Europy katalog dostępnych online archiwalnych baz danych i repozytoriów znaleźć można w serwisie [Archivesportaleurope.net](http://Archivesportaleurope.net).

Nazwa bazy	URL	Zawartość
Szukaj w Archiwach	<a href="http://www.szukajwarchiwach.pl/">http://www.szukajwarchiwach.pl/</a>	Zawiera opisy materiałów archiwalnych oraz skany z kilkunastu instytucji, głównie archiwów państwowych. Baza z założenia ma dawać dostęp do zbiorów cyfrowych, dlatego większość instytucji zamieszcza w niej informacje o zasobie, który jest lub ma być w najbliższym czasie zdigitalizowany. Wyszukiwarka ma zaawansowane kryteria wyszukiwania.
Wyszukiwarka zespołów narodowego zasobu archiwalnego	<a href="http://baza.archiwa.gov.pl/sezam/sezam.php">http://baza.archiwa.gov.pl/sezam/sezam.php</a>	Wyszukiwarka zawiera informacje o zasobie wszystkich archiwów państwowych oraz kilkunastu instytucji polskich i polonijnych spoza państwowej sieci archiwalnej. Pozwala zobaczyć spis zespołów danego archiwum oraz wyszukiwać według sygnatury, klasyfikacji dziedziny twórcy zespołu, dziejów twórcy itp.

Wyszukiwarka ogólna narodowego zasobu archiwalnego	<a href="http://baza.archiwa.gov.pl/sezam/index.php">http://baza.archiwa.gov.pl/sezam/index.php</a>	Zawiera opisy jednostek archiwalnych pochodzące z różnych baz danych prowadzonych przez archiwa i instytucje współpracujące. Prezentowanych jest ponad 19 700 inwentarzy zespołów, co stanowi ok. 22 proc. wszystkich zespołów przechowywanych w archiwach państwowych (dane Naczelnej Dyrekcji Archiwów Państwowych).
Program Rejestracji Akt Metrykalnych i Stanu Cywilnego PRADZIAD	<a href="http://mak.bn.org.pl/cgi-bin/makwww.exe?BM=23">http://mak.bn.org.pl/cgi-bin/makwww.exe?BM=23</a>	Baza zawiera informacje o księgach metrykalnych i stanu cywilnego w archiwach państwowych i w instytucjach współpracujących. Nie znajdziemy tam jednak spisu nazwisk występujących w dokumentach bądź informacji o konkretnych osobach.
Program Straty osobowe i ofiary represji pod okupacją niemiecką	<a href="http://www.straty.pl/">http://www.straty.pl/</a>	Baza udostępnia dane o blisko trzech i pół milionach ofiar i osobach represjonowanych. Podając imię, nazwisko i ew. datę urodzenia – można odnaleźć informacje o konkretnych osobach i ich losach w czasie II wojny światowej.
Indeks Represjonowanych przez organa władzy sowieckiej w latach 1939–56	<a href="http://www.indeks.karta.org.pl/pl/">http://www.indeks.karta.org.pl/pl/</a>	Ponad 316 tys. biogramów, w tym osoby z dwóch oryginalnych zestawień – tzw. listy ukraińskiej (więźniowie rozstrzelani na podstawie decyzji władz sowieckich z 5 marca 1940) oraz wykazu spraw prowadzonych przez organa NKWD Zachodniej Ukrainy i Białorusi, będącego wiarygodnym spisem osób, które aresztowano i przeciw którym wszczęto śledztwo podczas okupacji sowieckiej 1939-41).
Archiwum Historii Mówionej	<a href="http://www.audiohistoria.pl/web/">http://www.audiohistoria.pl/web/</a>	Archiwum Historii Mówionej to największy w Polsce zbiór relacji biograficznych oraz innych archiwalnych świadectw, obejmujących niemal cały XX wiek. Zbiór obejmuje około 4 tysięcy nagrań, z których część dostępna jest online.

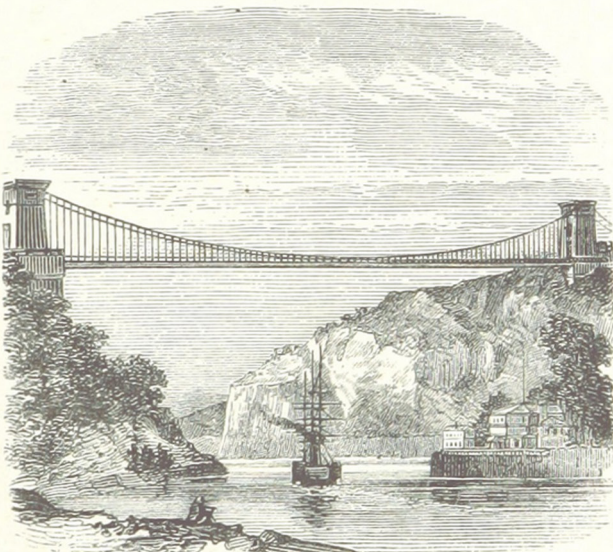
Wiedza o dostępnych archiwalnych i biograficznych bazach danych jest kluczowa. Repozytoria publikujące skany pozwalają na dotarcie jedynie do części zasobów – we wspomnianym wyżej serwisie [szukajwarchiwach.pl](http://szukajwarchiwach.pl) przeglądać dziś można około 7 mln zdigitalizowanych dokumentów, a to wciąż niewiele w porównaniu z 70 mln skanów przygotowanych przez polskich archiwistów, ale z różnych przyczyn nieopublikowanych. To także niewiele w porównaniu z liczbą niezdigitalizowanych dokumentów przechowywanych w zasobach polskich archiwów.

Dominacja Google w zakresie wyszukiwania informacji w Internecie nie wiąże się jedynie z pojedynczym narzędziem (wyszukiwarką tej firmy), ale także z samym modelem wyszukiwania. Google sprawia, że przyzwyczajamy się do wyszukiwania pełnotekstowego, o nieograniczonym zasięgu, prostego i jak najbardziej intuicyjnego, które od razu gwarantuje nam dostęp do poszukiwanej informacji. Tymczasem wyszukiwarki baz archiwalnych czy innych narzędzi naukowych bardzo rzadko pozwalają na taki sposób odnajdywania informacji, wymagają też większej wiedzy na temat poszukiwanych zasobów oraz umiejętności szczegółowego określania zapytania. Pamiętajmy jednak, że Google indeksuje jedynie publicznie dostępne zasoby WWW, a to tylko część informacji, do których możemy dotrzeć dzięki Internetowi.

Standardowy mechanizm wyszukiwania wymaga od użytkownika wiedzy o tym, czego dokładnie chce wyszukać. W przypadku dostępności milionów skanów nie jest to rozwiązanie optymalne: system wyświetlając wyniki wyłącznie na bazie słów kluczowych może pomijać zasoby, których treść pasuje do zapytania, ale których opis nie jest z nim zbieżny. Eksperyment, który ma pozwolić na dotarcie do dotąd nieopisanych treści skanów rozwijają Ben O'Steen i James Baker w ramach British Library Labs. Stworzone przez nich oprogramowanie (Mechanical Curator) wycięło automatycznie ponad milion grafik, rycin i innych obiektów wizualnych z 46 tys. XVII, XVIII i XIX-wiecznych zdigitalizowanych książek przechowywanych przez British Library. Grafiki te są co godzinę automatycznie i losowo publikowane na Twitterze i na blogu (<http://mechanicalcurator.tumblr.com/>) wraz z opisem bibliograficznym źródłowej publikacji. Udostępniono je także, o czym pisałem wyżej, w popularnym serwisie fotograficznym Flickr.com w ramach projektu Flickr Commons. British Library chce wykorzystać społeczność użytkowników Flickr'a do opisywania treści opublikowanych grafik – wypracowane w taki sposób metadane pozwolą skuteczniej przeszukiwać zbiory i staną się podstawą bazy danych, z której korzystał będzie Mechanical Curator przy kolejnych automatycznych kwerendach

the  
and  
thin  
dicts  
all  
arer  
nich  
to  
the  
stle  
the  
ide.  
ipal  
ck-

Fig. 66.—CLIFTON SUSPENSION BRIDGE.



ll-slope overlooking the valley of the Wye

Image from 'The Earth and its Inhabitants. The European section of the Universal Geography by E. Reclus. Edited by E. G. Ravenstein. Illustrated by ... engravings and maps', [003055004](#)


- **Author:** RECLUS, Élisée.
- **Volume:** 04
- **Page:** 155
- **Year:** 1878
- **Place:** London
- **Publisher:**

Following the link above will take you to the British Library's integrated catalogue. You will be able to download a PDF of the book this image is taken from, as well as view the pages up close with the 'itemViewer'. Click on the 'related items' to search for the electronic version of this work.

- Open the page in the [British Library's itemViewer \(page. 000155\)](#)
- [Download the PDF for this book](#)

#ldigital #bl\_labs #britishlibrary #1878 #similar\_to\_70375062170\_published\_date  
#similar\_to\_70375062170\_slantyness #similar\_to\_70375062170\_bubblyness\_y  
#upperbody\_detected\_410left\_550top\_512right\_634bottom

[Tweet](#)

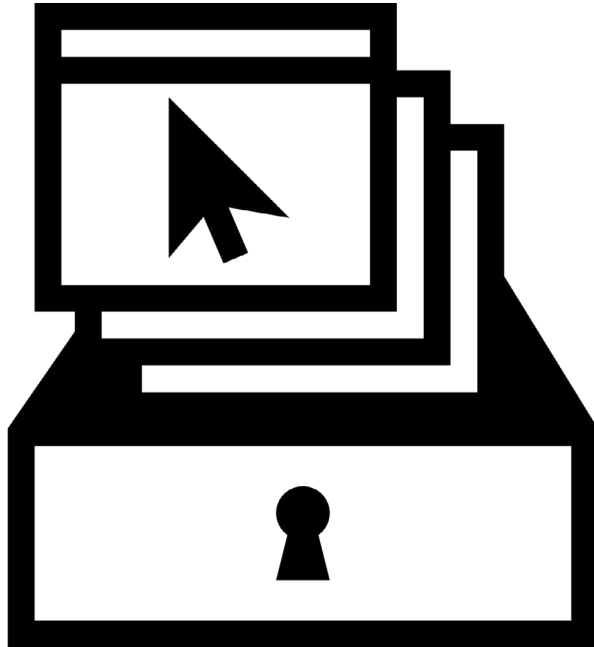


Efekt działania programu Mechanical Curator: grafika wycięta automatycznie z treści skanu zdigitalizowanej przez *British Library* książki *The Earth and its Inhabitants. The European section of the Universal Geography* (London 1878). Oprogramowanie publikuje przy tym podstawowe informacje bibliograficzne oraz odnośnik do skanu książki dostępnego w repozytorium cyfrowym.

## BIBLIOGRAFIA

- Harris, V., Hepburn, P.** (2013). Trends in Image Use by Historians and the Implications for Librarians and Archivists. *College & Research Libraries*, 74(3), 272-287. Pobrano z: <http://crl.acrl.org/content/74/3/272.short>
- Hitchcock, T.** (2008). Digital Searching and the Reformulation of Historical Knowledge. W: Greengrass, M., Hughes L. (red.), *The Virtual Representation of the Past*. Burlington: Ashgate Publishing.
- Prescott, A.** (2008). The imaging of historical documents, W: Greengrass, M., Hughes, L. (red.), *The Virtual Representation of the Past*. Burlington: Ashgate Publishing.
- Tibbo, H. R.** (2003). Primarily History in America: How U.S. Historians Search for Primary Materials at the Dawn of the Digital Age. *American Archivist*, 6(1), 9-50. Pobrano z: <http://archivists.metapress.com/content/b120370l1g718n74/>





# Archiwa społeczne, oddolna digitalizacja i cyfrowa archiwistyka osobista

Historyk w swojej pracy badawczej korzystać może nie tylko z repozytoriów powstających w ramach instytucji publicznych czy ośrodków akademickich. Również wiele organizacji pozarządowych udostępnia w Internecie własne zbiory. Jakość tych zbiorów, merytoryczność opisu czy też czas istnienia tego typu inicjatyw są różne, jednak – zwłaszcza dla historyków badających lokalną przeszłość – mogą być one potencjalnie cennymi zasobami i źródłami informacji.

Pojęcie archiwów społecznych pojawiło się w diagnozie Ministerstwa Kultury i Dziedzictwa Narodowego określonej jako „Strategia Rozwoju Kapitału Społecznego”. W wydanym przez Ośrodek KARTA podręczniku archiwistyki społecznej zdefiniowano je w następujący sposób:

Archiwum społeczne prowadzone jest przez: organizację pozarządową (w rozumieniu art. 3 ust. 2 *Ustawy o działalności pożytku publicznego i o wolontariacie*) lub powstaje przy jednostce samorządu terytorialnego, w efekcie celowej aktywności obywatelskiej. Archiwum gromadzi, przechowuje, opracowuje (w celu udostępniania) materiały dokumentowe, ikonograficzne lub audiowizualne. Dokumentacja ta tworzy niepaństwowy zasób archiwalny (Ziętał, 2012, s. 9).

Spółeczny, oddolny charakter tej misji podkreślał w swojej analizie Tomasz Czarnota, zestawiając ją z nie zawsze skuteczną polityką archiwów państwowych wobec materiałów wytworzonych poza nadzorem państwowej służby archiwalnej. Jego zdaniem

[...] archiwa społeczne są nie tyle konkurencją dla AP [Archiwów Państwowych], ile alternatywą dla tych podmiotów prawnych i osób fizycznych, które nie są z mocy prawa zobowiązane do przekazywania AP swojej dokumentacji o wartości historycznej. Dla tych, które mając prawo wyboru, wolą ją powierzać nie państwowym, ale społecznym ośrodkom. Z możliwości tej nie omieszkano skorzystać już całkiem niewąskie grono. Z faktu tego dla AP, nie rezygnujących z misji opieki nad całym narodowym zasobem archiwalnym, powinien chyba wynikać imperatyw nawiązania z nimi stałej partnerskiej (z naciskiem na ów przymiotnik) współpracy (Czarnota, 2011, s. 24).

Chociaż dyskusja na temat pojęcia, jakim jest *archiwum społeczne*, szczególnie w odniesieniu do teorii i prawa archiwalnego, dopiero się w Polsce zaczyna, zbiory archiwów powstałych przy organizacjach pozarządowych i samorządach już teraz obecne są w Internecie i można wykorzystywać je w badaniach. W ramach Federacji Bibliotek Cyfrowych archiwalia udostępniają m.in. Ośrodek KARTA, Teatr NN Brama Grodzka, Biblioteka Cyfrowa Małopolskiego Towarzystwa Genealogicznego i cyfrowe archiwum KK NSZZ Solidarność. Ośrodek KARTA prowadzi także program Cyfrowe Archiwa Tradycji Lokalnej (CATL), w którym ponad sto bibliotek gminnych z niewielkich miejscowości animuje proces gromadzenia, opisywania i udostępniania lokalnych zbiorów. Na stronie projektu dostępnej pod adresem [archiwa.org](http://archiwa.org) umieszczono dotąd prawie 14 tys. skanów.

Na przełomie 2012 i 2013 roku odbyły się w Polsce dwie konferencje naukowe dotyczące tematyki archiwów społecznych. Pierwsza została zorganizowana w grudniu 2012 roku na Uniwersytecie Marii Curie-Skłodowskiej, a druga w marcu, w Warszawie, w siedzibie Naczelnej Dyrekcji Archiwów Państwowych. Historycy zainteresowani historią lokalną powinni śledzić rozwój dyskusji na temat oddolnie tworzonych archiwów.

Dość interesującym przykładem archiwum społecznego jest katalog archiwum dworu Saari (<http://saarenkartano.muuritutkimus.fi/>), archiwum

prowadzonego przez fińską organizację pozarządową Kone Foundation w mieście Hietamäki. Specyfika internetowego repozytorium Saari Manor polega m.in. na wykorzystaniu modelu wiki, czyli specjalnego rodzaju serwisu internetowego, którego treść edytować może każdy. Fińskie archiwum ma zatem charakter otwarty i partycypacyjny, a sama konstrukcja repozytorium cyfrowego konsultowana była z badaczami, którzy realizowali swoje badania w Saari Manor. Badacze ci uzupełniają także opisy dokumentów, które zostały udostępnione online przez Fundację (Huvila, 2008).

Pozostając przy temacie archiwów społecznych, warto zwrócić uwagę na inicjatywy i formy aktywności, ujawniające się już poza instytucjami, potencjalnie mogące mieć znaczenie dla badaczy. Przykładowo: forum Dawny Gdańsk ([www.forum.dawnnygdansk.pl](http://www.forum.dawnnygdansk.pl)) to jedna z najpopularniejszych przestrzeni dyskusji na temat historii Gdańska (także Trójmiasta i całego Pomorza) w Internecie. To także doskonały przykład oddolnej digitalizacji – nieformalnego rozpowszechniania zasobów historycznych poza zasięgiem instytucji państwowych czy archiwów społecznych. Użytkownicy forum publikują tam setki fotografii historycznych, które są następnie analizowane przez członków społeczności. Źródłem zdjęć są ich prywatne kolekcje, profesjonalne repozytoria, a nawet... aukcje internetowe, z których kopiowane są reprodukcje. Nie można mówić tutaj oczywiście o jakimkolwiek archiwum, zakładającym w swojej istocie określony porządek i władzę katalogowania – zamiast tego jednak historyk prowadzący badania nad historią lokalną znajdzie tam wsparcie i pomoc w rozwiązaniu problemów z interpretacją fotografii czy w poszukiwaniach nowych materiałów, mogących dokumentować podejmowany przez niego temat. Materiały źródłowe, zazwyczaj fotografie, publikowane w ramach tego typu działań niearchiwalnych – ale o archiwalnym charakterze – udostępniane są oczywiście z pominięciem jakichkolwiek standardów co do opisu, jakości skanu czy kwestii praw autorskich.

Dość interesującą perspektywę oddolnej digitalizacji opisuje Melissa Terras. Jej zdaniem amatorskie skanowanie i udostępnianie materiałów historycznych przez użytkowników Internetu nie powinno być ignorowane przez profesjonalne instytucje. Organizowane oddolnie „archiwa” mogą być dla instytucji inspiracją do poprawy użyteczności repozytoriów cyfrowych czy do usprawnienia komunikacji z jej użytkownikami. Poza tym zbiory gromadzone oddolnie mogą też uzupełniać ofertę profesjonalnych zbiorów:

Entuzjastyczna amatorska digitalizacja, zjawisko dotąd ignorowane przez profesjonalistów, tworzy bogate internetowe źródło zasobów dziedzictwa kulturowego, zawierające nierzadko zbiory, którymi nie są zainteresowane tradycyjne instytucje. Faktycznie, efemerydy i zasoby kultury popularnej są często lepiej udostępniane przez amatorskie społeczności niż przez instytucje pamięci. Energia i zapał amatorskich digitalizatorów jest warty dalszego zainteresowania, jako że amatorskie kolekcje często uzupełniają istniejące już zbiory, tworząc alternatywną przestrzeń dyskusji między entuzjastami. Społeczność amatorów w integrowaniu użytkowników przewyższa instytucje pamięci, a statystyki generowane przez użytkowników takich zbiorów wskazują, że częste aktualizacje i interakcje są kluczem do budowania grupy stałych odbiorców. W rezultacie instytucje pamięci powinny być świadome, że nie jest już możliwe realizowanie polityki „zeskanuj i porzuć” (*scan and dump*) wobec swoich digitalizowanych zbiorów i biernie oczekiwanie na ich wykorzystywanie (Terras, 2010, s. 436).

W literaturze przedmiotu znaleźć możemy wiele przykładów nieformalnych i oddolnych inicjatyw digitalizacyjnych i archiwalnych, które zyskały nawet pewną przewagę nad archiwami profesjonalnymi i instytucjonalnymi. Alan McKnee z Queensland University of Technology w artykule *YouTube versus the National Film and Sound Archive: Which Is the More Useful Resource for Historians of Australian Television?* porównuje potencjał australijskiego profesjonalnego archiwum telewizyjnego National Film and Sound Archive (NFSA) z YouTube, najpopularniejszym dziś serwisem udostępniającym video. Okazuje się, że dla historyka telewizji, zainteresowanego badaniem tego medium z perspektywy popkultury, serwis YouTube jest nie tylko zdecydowanie bardziej dostępny (ang. *accessibility*), ale także pozwala dotrzeć do nagrań, które z różnych powodów nie są zamieszczone w repozytorium NFSA (McKee, 2011). Podobną sytuację obserwować możemy w Polsce: instytucjonalny dostęp do archiwalnych zbiorów telewizji publicznej jest bardzo utrudniony, jednak wiele nagrań dostępnych jest niezależnie (i niestety nie zawsze zgodnie z zapisami prawa autorskiego) na YouTube. Dla polskiego historyka, zainteresowanego wykorzystywaniem materiałów telewizyjnych, YouTube z pewnością nie będzie wystarczająco dobrą alternatywą, ale może być świetnym uzupełnieniem oferty archiwum TVP. Już teraz w wielu pracach historycznych czy kulturoznawczych przypisy do archiwalnych audycji telewizyjnych kierują nie do archiwów instytucjonalnych, ale właśnie do YouTube.

Ostatnim zjawiskiem, które warto przedstawić w niniejszym rozdziale, są cyfrowe banki pamięci (ang. *digital memory banks*). Pod nazwą tą kryją się internetowe archiwa budowane przez organizacje pozarządowe czy instytucje akademickie, których celem jest gromadzenie dokumentacji na temat ważnych aktualnych wydarzeń czy problemów. Przykładem takiej inicjatywy jest stworzone m.in. przez Center For History and New Media (CHNM) archiwum relacji, dokumentów i wspomnień związanych z tragicznym przejściem przez południowe stany USA huraganów Rita i Katrina w 2005 roku. Hurricane Digital Memory Bank ([Hurricanearchive.org](http://Hurricanearchive.org)), podobnie jak inne projekty archiwalne tego typu, powstał w reakcji na wydarzenie o wyjątkowym znaczeniu. Jego celem jest zgromadzenie bazy źródłowej dla przyszłych badań (Brennan, Kelly, Mills, 2010). Inne archiwa tego typu to September 11 Digital Archive (gromadzące źródła cyfrowe dotyczące zamachów z 11 września 2001) czy Canterbury Earthquake Digital Archive (poświęcone trzęsieniom ziemi w nowozelandzkiej prowincji Canterbury). Projekty te mają charakter otwarty, a zatem każdy może dodać do repozytorium swoją relację czy posiadane przez siebie dokumenty i fotografie. Gromadzą one zbiory w aktywny sposób, przykładowo inspirując do pisania relacji oraz do udostępniania fotografii. Cyfrowe banki pamięci są zatem archiwami aktywnie poszukującymi zbiorów. Zdaniem Paula Millara, jednego z twórców wspomnianego wyżej projektu z Canterbury:

[...] archiwiści muszą być znacznie bardziej zręczni i aktywni niż w przeszłości. To nie jest dla nich łatwa rola, ponieważ archiwa są często postrzegane jako bierne repozytoria, a nie aktywni twórcy wiedzy (za: Wilkowski, 2012).

Dla badacza historii każda tego typu inicjatywa archiwalna może być inspiracją do wypracowania nowych metod gromadzenia przez Internet źródeł (relacji, dokumentów). Jak przekonują jednak Mills T. Kelly i Sheila A. Brennan, Internet może służyć tu jedynie wsparciem, bo na pewno nie może być jedyną przestrzenią pracy. Zbieranie materiałów źródłowych jedynie online ogranicza zasięg projektu archiwalnego, jego organizatorzy nie mogą też ignorować problemów związanych z kompetencjami potrzebnymi do udostępniania zasobów czy ze spamem (Brennan, Kelly, Mills, 2010).

Opisane wyżej trendy i projekty odnieść można do popularnego na Zachodzie pojęcia *historii publicznej* (ang. *public history*). Nie jest ono jednoznacznie definiowane i bywa wykorzystywane w wielu kontekstach. Pojęcie

*historii publicznej* jest stosowane np. w analizach społecznej recepcji filmów historycznych, opisach obecności w przestrzeni publicznej miejsc pamięci (jak choćby pomników), czy w omawianiu wszelkich aktywności o charakterze wiedzotwórczym, odnoszących się do historii, podejmowanych poza albo na granicach systemu akademickiego i profesjonalnych instytucji pamięci (muzeów, archiwów). W tym przypadku idea historii publicznej akcentuje podmiotowość tych, którzy dotąd byli raczej przedmiotem badania historycznego (de Groot, 2008, s. 60). Najbardziej rozpowszechnioną formą *public history* są więc badania genealogiczne i historia lokalna, uprawiana przez osoby nieposiadające formalnego wykształcenia historycznego, ale – nierzadko bardzo skutecznie – posługujące się metodami naukowymi. Szczególnie w anglosaskim systemie edukacji historyków duży nacisk jest kładziony na tego typu wyjście poza środowisko akademickie czy mury instytucji, stąd wiele amerykańskich, a także brytyjskich uniwersytetów oferuje studentom historii specjalizacje w kierunku *public history*. Na gruncie tej dziedziny rozwijanych jest też wiele wątków teoretycznych, związanych np. z wytyczaniem granicy między historią (historiografią) a przeszłością, definiowaniem dziedzictwa, etyką badań historycznych czy ich politycznymi aspektami, zwłaszcza w perspektywie obiektów tych badań, np. wspólnot lokalnych, etnicznych itp. (zob. np. Ashton i Kean, 2009).

Opis oddolnie tworzonych archiwów byłby niepełny, gdyby zabrakło w nim odniesienia do cyfrowych archiwów osobistych, które są już obiektem zainteresowania teoretyków archiwów oraz archiwistów i bibliotekarzy. Cyfryzacja życia codziennego polega na upowszechnieniu się narzędzi pozwalających na nieustanną rejestrację (telefony komórkowe z aparatami, archiwa maili i dokumentów w komputerach, aplikacje rejestrujące pozycje geograficzne dzięki GPS itp.), które sprawiają, że ich użytkownicy gromadzą coraz więcej danych, mogących mieć znaczenie dla historyków. Historyk piszący np. biografię Jacka Kuronia dociera i analizuje dostępne na jego temat dokumenty oraz zasoby z jego prywatnego archiwum, czyli listy, notatki i zdjęcia. Tymczasem badanie biografii osoby współcześnie żyjącej oznaczać będzie konieczność skorzystania ze źródeł cyfrowych, których może być zdecydowanie więcej, ale które mogą być przy tym trudniej dostępne, oczywiście, o ile w ogóle zostaną zachowane. W 2006 roku Salman Rushdie, autor m.in. powieści *Szatańskie wersety*, przekazał swoje archiwum osobiste bibliotece Uniwersytetu Emory. Oprócz archiwum papierowego biblioteka otrzymała komputer osobisty, trzy laptopy, zewnętrzny dysk twardy i smartfona. Znajdujące się na tych nośnikach dokumenty, czyli

maile, SMS-y, zdjęcia itp., mogą zostać wykorzystane w pracach naukowych dotyczących biografii pisarza czy jego twórczości literackiej.

Archiwizacja codzienności wynika nie tylko z możliwości technologicznych i lepszej dostępności narzędzi cyfrowych, ale także ze zmian kulturowych. Jak przekonuje Mateusz Halawa, archiwum jest narzędziem tożsamościowym, bo pozwala na budowanie własnego *ja* (Halawa, 2011). Staje się też elementem procesów komunikacyjnych, co doskonale widać na przykładzie facebookowego publikowania na bieżąco informacji o sobie (ang. *lifestreaming*). Już dziś tego typu zasoby to z jednej strony przedmiot zainteresowania badaczy społecznych, z drugiej – akcji edukacyjnych, promujących wiedzę o metodach zabezpieczania osobistych zbiorów cyfrowych (ang. *personal digital archiving*). Akcje te prowadzi np. Biblioteka Kongresu. W 2013 roku opublikowano raport podsumowujący podejmowane w ostatnich latach aktywności Biblioteki w tym zakresie i podkreślający, że osobiste zbiory cyfrowe – także te *born digital*, nieposiadające pierwotnej postaci analogowej, stanowią ważny element narodowego dziedzictwa:

Wyobraź sobie kogoś, kto za 50 lat zechce badać życie w początkach XXI wieku. Jakiękolwiek rozsądne zrozumienie tego czasu wymagać będzie korzystania ze źródeł *born digital*: stron internetowych, wiadomości poczty elektronicznej, fotografii cyfrowych i szeregu innych materiałów, które istnieją jedynie w postaci cyfrowej. Oznacza to konieczność zabezpieczania wielu cyfrowych informacji do długotrwałego przechowywania. Jest to przy tym coś zupełnie innego niż zabezpieczanie książek, dokumentów i innych tradycyjnych form, za pomocą których przekazujemy informacje (LeFurgy, 2013, s. 3).

Działania podejmowane na rzecz ochrony takich zbiorów są uzupełnieniem pracy profesjonalnych archiwistów i bibliotekarzy, którzy zabezpieczają cyfrowe zbiory tworzone przez instytucje państwowe czy organizacje społeczne. Warto podkreślić, że zbiory stające się elementem osobistego archiwum cyfrowego to nie tylko fotografie cyfrowe czy wiadomości email, ale także np. profile z serwisów społecznościowych. Duże zróżnicowanie typu zasobów oraz ich cyfrowa postać generuje wiele problemów związanych np. z kwestiami praw autorskich, ochroną prywatności czy też z kwestiami dziedziczenia.



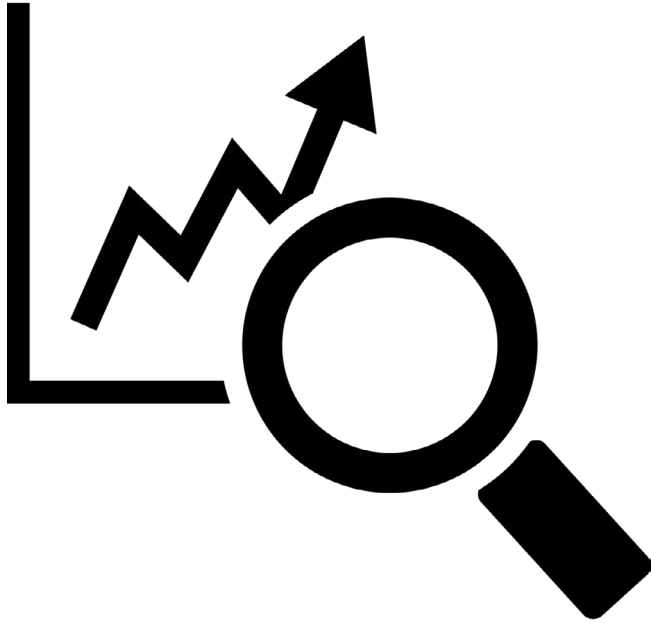
Problem jest o tyle skomplikowany, że duża część prywatnych zbiorów cyfrowych gromadzona jest dziś w tzw. chmurze, czyli na serwerach dostawców usług pocztowych czy wirtualnych kont pozwalających na przechowywanie plików. W przypadku śmierci właściciela tych zasobów osoby bliskie lub badacze mogą chcieć uzyskać do nich dostęp, na zasadnie podobnej do tej, w ramach której przejmowane są tradycyjne prywatne archiwa papierowe. Często jednak nie posiadają oni danych pozwalających na zalogowanie się do konkretnych usług poczty elektronicznej czy dysku internetowego. Pomiędzy rodziną czy badaczami a archiwum zmarłego pojawia się pośrednik w postaci – jak w przypadku Gmaila – dużej globalnej firmy, z którą bezpośredni kontakt jest bardzo utrudniony. Google co prawda proponuje pewne rozwiązanie – tzw. Menedżer Nieaktywnego Konta, który pozwala złożyć deklarację co do dostępności prywatnych zasobów w usługach tej firmy dla innych osób po śmierci właściciela. Bez tej deklaracji uzyskanie możliwości wglądu do zgromadzonych treści jest bardzo utrudnione i wymaga korespondencji z administracją Google'a w Mountain View oraz wykazania się odpowiednimi dokumentami.

Więcej na temat dziedzictwa cyfrowego przeczytać można w jednym z kolejnych rozdziałów.

## BIBLIOGRAFIA

- Ashton, P., Kean, H.** (2009). *People and their Pasts: public history today*, Basingstoke, New York: Palgrave Macmillan.
- Brennan, S., Kelly, A., Mills, T.** (2010). *Why Collecting History Online is Web 1.5*, Center for History and New Media, Case Study. Pobrano z: <http://chnm.gmu.edu/essays-on-history-new-media/essays/?essayid=47>.
- Czarnota, P.** (2011). Komu są potrzebne społeczne archiwa? *Archiwista Polski*, 4(64),15-33.
- de Groot, J.** (2008). *Consuming history: Historians and heritage in contemporary popular culture*, London ; New York : Routledge.
- LeFurgy, B. i in.** (2013), *Perspectives on Personal Digital Archiving*. *National Digital Information Infrastructure and Preservation Program*, Library of Congress, [http://www.digitalpreservation.gov/documents/ebookpdf\\_march18.pdf](http://www.digitalpreservation.gov/documents/ebookpdf_march18.pdf)

- Halawa, M.** (2011), Nowe media i archiwizacja życia codziennego, *Kultura Współczesna*, 4(70), 27-41.
- Huvila, I.** (2008). Participatory archive: towards decentralised curation, radical user orientation, and broader contextualisation of records management, *Archival Science*, 8(1), 15-36.
- Terras, M.** (2010). Digital curiosities: resource creation via amateur digitization, *Linguist Computing*, 25(4), 425-438.
- McKee, A.** (2011). YouTube versus the National Film and Sound Archive: Which Is the More Useful Resource for Historians of Australian Television? *Television New Media*, 12(2), 154-173.
- Wilkowski, M.** (2012). *Canterbury Earthquake Digital Archive: archiwum cyfrowe to archiwum aktywne*. Pobrano z: <http://historiaimedia.org/2012/07/09/canterbury-earthquake-digital-archive-archiwum-cyfrowe-to-archiwum-aktywne/>.
- Ziętał, K.** (2012). Wstęp. W: Ziętał, K. (red.), *Archiwistyka społeczna* (7-13). Warszawa: Ośrodek KARTA.



# Wyszukiwanie pełnotekstowe w archiwum cyfrowym *i text mining*

Tradycyjne archiwum udostępniające oryginalne dokumenty lub ich analogowe reprodukcje, np. w postaci mikrofilmów, przerzuca na badacza całą pracę związaną z analizowaniem ich treści. Archiwiści przygotowują co prawda podstawowy opis na poziomie zestawu metadanych, jednak odczytanie treści dokumentu w poszukiwaniu konkretnych informacji jest już zadaniem osoby korzystającej ze zbioru. Jak odbywa się to w sytuacji, kiedy archiwalia udostępniane są w Internecie w formie pełnotekstowej, tzn. pozwalającej na automatyczne przeszukiwanie ich treści pod kątem wybranych słów? Co zmienia się, kiedy badacz ma dostęp nie tylko do pełnego tekstu, ale także odpowiednich adnotacji w nim zapisanych, wskazujących na określone cechy treści?

Rafał T. Prinke zwraca uwagę, że tekst pozbawiony adnotacji praktycznie nie istnieje:

Sposoby wyróżniania granic wyrazów, zdań czy akapitów, oznaczanie cytatów, odsyłaczy przypisów i ich treści, układ dialogów albo tytułów poszczególnych części dzieła – to wszystko są sposoby adnotowania, które ewoluowały i zmieniały się wraz z rozwojem pisma, sztuki drukarskiej i analitycznego rozumienia natury tekstu (Prinke, 2000, s. 53).

Znaki takie jak kropka czy przecinek, odstępy między wyrazami, akapity itp. to pozbawione treści podstawowe elementy tekstu, wskazujące na jego strukturę czy prawidłowy sposób odczytywania. Transkrypcje tekstów historycznych wzbogacone adnotacjami to rozwinięcie metastruktury tekstu, możliwe dzięki istnieniu odpowiednich szablonów zapisów (np. plików XML – *Extensible Markup Language*, rozszerzalny język znaczników) i standardów ich przygotowywania.

Przykładem archiwum udostępniającego transkrybowane i adnotowane teksty źródłowe jest brytyjski projekt London Lives (<http://www.londonlives.org/>), udostępniający materiały archiwalne dokumentujące historię kryminalną dawnego Londynu. Każdy dokument publikowany jest w dwóch wersjach: standardowego skanu (pliku graficznego) i w formie tekstowej, opartej na odpowiednim schemacie XML. Nie jest to więc zwykła transkrypcja treści, polegająca na udostępnieniu jednego ciągu tekstu, ale ustandaryzowany zestaw pól semantycznych. Każdy element treści, np. podpis sędziego, zapis informacyjny o dacie rozpatrywania sprawy karnej, imię i nazwisko osoby sądzonej, skreślenie itp., opisany jest odpowiednim znacznikiem (tagiem), pozwalającym systemowi rozpoznać znaczenie treści przetwarzanych dokumentów:

```
<rs type="persName" id="OA16901024_n12-1">Peter
Vallard</rs>
<interp inst="OA16901024_n12-1" type="given"
value="PETER"></interp>
<interp inst="OA16901024_n12-1" type="surname"
value="VALLARD"></interp> <interp
inst="OA16901024_n12-1" type="gender" value="male">
</interp> [...] He was a working <rs type="occupa-
tion" id="OA16901024_occ2">Silver-Smith</rs>
```

Powyższy fragment pochodzi z wydanej drukiem w 1690 roku publikacji opracowanej przez więźnia w Newgate, w której zamieszczano krótkie biografie osób skazanych na śmierć oraz przytaczano ich ostatnie słowa przed wykonaniem egzekucji. Informacja o Peterze Vallardzie, Francuzie skazanym za zdradę stanu, została uzupełniona wpisem o jego zawodzie (był złotnikiem, ang. *silver-smith*). Treść tego fragmentu opisana jest za pomocą odpowiednich tagów, pozwalających tworzyć specjalistyczne indeksy czy przeszukiwać treść

dokumentów udostępnianych w serwisie pod kątem płci, imion i nazwisk skazanych czy zawodów przez nich wykonywanych. XML określa sposób, w jaki zapisywane są dane wraz z ich strukturą, pozwalając oprogramowaniu rozpoznawać znaczenie zgromadzonych informacji. Najprostszym przykładem może być taki fragment kodu:

```
<cytat>to jest treść cytatu</cytat>
```

Spróbujmy zanalizować jeszcze jeden przykład, tym razem fikcyjnej bazy biogramów ofiar cywilnych powstania warszawskiego. Fragment przygotowanego przez historyka rekordu w postaci zwykłego tekstu:

5 sierpnia 1944 Jan i jego matka zostali wraz z innymi mieszkańcami Woli wypędzeni z domów w okolice torów kolejowych przy ulicy Górczewskiej.

po uzupełnieniu odpowiednio zaprojektowanymi znacznikami (tagami) XML uzyskać może następującą postać:

```
<data>5 sierpnia 1944</data> <osoba id="Jan Adamczewski"><imie>Jan</imie></osoba> i jego <osoba id="Teresa Adamczewska">matka</osoba> zostali wraz z innymi mieszkańcami <miejsce typ="dzielnica">Woli</miejsce> wypędzeni z domów w okolice torów kolejowych przy ulicy <miejsce typ="ulica">Górczewskiej</miejsce>.
```

Dzięki takiemu zapisowi użytkownik bazy biogramów mógłby przeszukiwać ją pod kątem imion i nazwisk wspomnianych w nich osób czy robić kwerendę szukając konkretnych nazw dzielnic lub ulic czy dat. W przypadku wyszukiwania pełnotekstowego w pozbawionej znaczników XML treści, wyszukiwarka nie mogłaby rozpoznać, że wspomniany w relacji *Jan* to *Jan Adamczewski*. Praca polegająca na semantycznym opisanu narracji pozwoli użytkownikowi skuteczniej badać udostępniane w bazie materiały. Przygotowanie takiej formy udostępniania treści historycznych jest bardzo pracochłonne, ponieważ poza standardową transkrypcją należy opisać treść dokumentu za pomocą semantycznych znaczników (adnotacji) i trzymać się przy tym przyjętego standardu opisu.

Jednym z najważniejszych źródeł standardów dotyczących cyfrowej reprezentacji tekstów historycznych i przestrzenią prac nad ich rozwijaniem jest konsorcjum Text Encoding Initiative (TEI). TEI wyewoluowało z grupy badaczy współpracujących ze sobą pod koniec lat 80. i rozwinęło się po przystąpieniu do inicjatywy stowarzyszeń naukowych związanych z nurtem cyfrowej humanistyki, przede wszystkim Association for Computers and the Humanities (ACH), do którego dołączyło Association for Computational Linguistics (ACL) oraz Association for Literary and Linguistic Computing (ALLC). Od 1994 roku standard TEI (*Guidelines for Electronic Text Encoding and Interchange*) wykorzystywany jest przez biblioteki i archiwa udostępniające swoje zbiory w formie cyfrowej oraz przez wydawców i badaczy pracujących nad naukowymi cyfrowymi edycjami tekstów historycznych. Jest to obszerny zestaw zaleceń dotyczących konstruowania plików XML będących semantycznym opisem transkrybowanych materiałów – ich struktury i treści, pozwalającym na automatyczne przetwarzanie ich i analizowanie przez komputer. W ramach TEI dostępne jest oprogramowanie ułatwiające zgodne ze standardem kodowanie tekstów historycznych, organizowane są także konferencje i szkolenia, publikowane są czasopisma.

Zdaniem Tima Hitchcocka, twórcy archiwum London Lives i kilku podobnych projektów, dostępność przeszukiwanych pełnotekstowo wersji dokumentów historycznych prowadzi do redukcji roli archiwum i znaczenia struktury, w ramach której udostępniane są dokumenty (zespoły i jednostki archiwalne). Do informacji źródłowych interesujących badacza dotrzeć można przecież bezpośrednio, po wpisaniu odpowiednich fraz w wyszukiwarke, przeszukującą wszystkie treści ponad strukturą zbiorów. Dzięki digitalizacji uwalniamy się zatem w pewnym sensie spod władzy organizacji archiwum (Hitchcock, 2008, s. 89). Efektem ubocznym pełnotekstowego przeszukiwania zdigitalizowanych zbiorów może być jednak utrata kontekstu, nadającego sens wyszukiwanym informacjom:

[...] przeszukiwanie po słowach kluczowych (*keyword searching*) tekstów drukowanych i – podobnie – kolekcji wizualnych dramatycznie wpłynęło na to, jak prowadzimy badania i czego szukamy. [...] Wybieramy frazę wyszukiwania lub zestaw takich fraz i przeszukujemy bezkrytycznie literaturę, dowody zbrodni, elektroniczne katalogi i gazety. W tym procesie często zanika wymóg (albo choćby możliwość) zrozumienia kontekstu, w ramach którego ujawnia się nam wybrana informacja (Hitchcock 2008, s. 85).

Hitchcock zwraca także uwagę na to, że dzięki digitalizacji i udostępnianiu miliardów słów z dawnych publikacji historycy coraz chętniej obudowują przygotowywane przez siebie przekazy o faktach źródłami narracyjnymi.

Pójdźmy jednak krok dalej. Skoro archiwum cyfrowe może udostępniać dokumenty historyczne w formie ustandaryzowanych plików XML z semantycznym opisem poszczególnych fragmentów ich treści, do przeszukiwania tych dokumentów można wykorzystać odpowiedni, automatycznie działający algorytm. W archiwum cyfrowym Old Bailey Online (<http://www.oldbaileyonline.org>), podobnym do projektu London Lives, można analizować automatycznie 127 milionów słów, zapisanych w dokumentach ze śledztw i procesów z lat 1674–1913, z wykorzystaniem API. Dzięki zastosowaniu odpowiednich skryptów, korzystających z danych udostępnianych przez Old Bailey Online, praca historyka ulega usprawnieniu, ponieważ w krótkim czasie może on zanalizować nawet obszerne zbiory. Oczywiście w tym miejscu pojawia się pytanie o wartości poznawcze takiego automatycznego badania, w porównaniu z tradycyjnym, manualnym przeglądaniem dokumentów, jeden po drugim. Przygotowanie algorytmu ściągającego odpowiednie informacje z bazy danych archiwum wymaga odpowiednich kompetencji programistycznych, co skłania z kolei do postawienia pytania o edukację informatyczną historyków. Ci ostatni mogą przecież nie tylko korzystać z tego typu archiwów, ale też je tworzyć, np. udostępniając za ich pomocą zbiory wykorzystywane do własnych badań. Problem kompetencji wiąże się tutaj także z problemem współpracy między historykiem, archiwistą i informatykiem: czy studia historyczne przygotowują do pracy grupowej, uczą zarządzania projektami badawczymi, w których bierze udział nie jedna, a kilka czy nawet kilkanaście osób? Model kolektywnej pracy jest obecnie standardem w naukach ścisłych i biologicznych – w humanistyce i naukach społecznych to wciąż nowość.

Duże zestawy transkrybowanych źródeł wymagają odpowiednich metod analizy. Tu z pomocą historykom przychodzą metody *text mining*. W dużym skrócie *text mining* to koncepcje i metody automatycznej eksploracji (czyli analizy) tekstów, zazwyczaj dużych korpusów dokumentów, pozwalające na wydobycie z nich określonych, nieujawniających się na pierwszy rzut oka informacji i prawidłowości. Pojęcie to warto odróżnić od pojęcia *data mining*, które opisuje metody analizy zestawów danych, posiadających już jednak – w odróżnieniu od tekstów – ściśle określoną strukturę. Ciekawym przykładem zastosowania w Polsce metod komputerowej analizy tekstów historycznych są badania nad



tożsamością i pisarstwem Galla Anonima, prowadzone przez prof. Tomasza Jasińskiego (Biblioteka Kórnicka PAN). Porównywał on za pomocą odpowiednich statystycznych narzędzi komputerowych styl *Kroniki Polskiej* oraz utworów, które przypisywał temu samemu autorowi, m.in. *Historii o translacji św. Mikołaja Wielkiego* (Jasiński, 2011).

Dla polskich badaczy opracowywane są narzędzia pozwalające na automatyczne analizowanie dużych zestawów tekstów. Dzieje się to za sprawą konsorcjum tworzonego przez Politechnikę Wrocławską, Instytut Podstaw Informatyki Polskiej Akademii Nauk, Polsko-Japońską Wyższą Szkołę Technik Komputerowych, Instytut Sławistyki Polskiej Akademii Nauk oraz Uniwersytet Łódzki i Uniwersytet Wrocławski w ramach europejskiej sieci CLARIN (Common Language Resources & Technology Infrastructure). Prace dotyczą tworzenia narzędzi umożliwiających automatyczne ustalanie autorstwa tekstów anonimowych, tworzenie indeksów, analizowanie relacji między dokumentami czy ich zaawansowane przeszukiwanie. Wymaga to naturalnie posiadania odpowiedniego korpusu transkrybowanych tekstów źródłowych. Jedną z dostępnych już usług oferowanych przez CLARIN-PL jest SłowoSieć, leksykalna baza wiedzy zawierająca słownik wyrazów bliskoznacznych i możliwość opisywania relacji semantycznych między poszczególnymi wyrazami.

Internet jest bogatym źródłem korpusów tekstów i zestawów danych, które można wykorzystywać w badaniach czy projektach naukowych (bazach danych, wizualizacjach, mashupach, prezentacjach). Wiele instytucji, takich jak np. europejska biblioteka cyfrowa Europeana, udostępnia swoje dane online na zasadach pozwalających każdemu eksplorować je do własnych celów – mówimy tu o tzw. *open data*. Czasem dostęp do korpusów tekstów bywa płatny – przykładem może być angielski serwis Electronic Enlightenment, zawierający w pełni transkrybowaną i uzupełnioną hipertekstualnymi odnośnikami korespondencję uczestników oświeceniowej republiki uczonych i poetów określaną jako *Respublica literaria*. Mediewiści skorzystać mogą z baz, takich jak choćby *Patrologia Latina Database*, zestawu publikowanych przez firmę ProQuest cyfrowych wydań tekstów Ojców Kościoła, przygotowanego na podstawie pracy XIX-wiecznego wydawcy Jacques'a Paula Migne'a. Baza ta pozwala na przeglądanie skanów, ale umożliwia także przeszukiwanie pełnotekstowe (proste i zaawansowane), korzystanie z aparatu krytycznego (przypisy), dodawanie własnych notatek itp. Wszystkie te możliwości oparte są na w pełni transkrybowanym tekście, uzupełnionym o odpowiednie semantyczne znaczniki. Jest to

wymagająca żmudnej pracy, ale z pewnością efektywna forma publikowania tekstów historycznych online, o wiele lepsza niż udostępnianie jedynie skanów.

Nie wszystkie historyczne bazy danych powstają wyłącznie jako efekty pracy zespołów specjalistów i w ramach dużych budżetów, często także przy udziale komercyjnych podmiotów. Najważniejszym celem takiej pracy nad bazą danych jest przekształcenie ciągu narracyjnego, tj. tekstu, zdań, akapitów, nieregularnych zestawów wyrazów, w katalog odpowiednio uszeregowanych danych, co pozwala później na ich szczegółowe przeszukiwanie. W realizowanym przez amerykańskie archiwa narodowe The U.S. National Archives and Records Administration (NARA) projekcie pod nazwą 1940 Census Community Indexing Project, każdy użytkownik Internetu mógł włączyć się w prace nad przenoszeniem informacji ze skanów kart ze spisu ludności z 1940 roku, udostępnionych online, do bazy danych. Dzięki pomocy ponad 150 tys. wolontariuszy udało się opracować przeszukiwalny indeks liczący 132 mln nazwisk. Dziś ta wypracowana oddolnie baza danych może być swobodnie wykorzystywana przez historyków, genealogów i wszystkich zainteresowanych historią rodzinną. Więcej o tego typu inicjatywach piszę w jednym z ostatnich rozdziałów podręcznika.

Korzystanie z internetowych historycznych baz danych niesie ze sobą jednak wyzwania dla naukowca, związane z nowymi metodami pracy. Efektywne korzystanie z bazy pełnotekstowej wymaga odpowiednich umiejętności, związanych z formułowaniem zapytań do jej wyszukiwarek i projektowaniem kwerend. Wątpliwości wzbudza także kwestia poprawnego informowania o wykorzystanych bazach danych w przypisach do pracy naukowej: jak pisze Krzysztof Narojczyk, poprawnie sformułowane odwołanie do tego typu źródła to coś więcej niż standardowe podanie adresu URL (jak w przypadku zwykłych stron WWW). Wyniki wyszukiwania w bazie danych mają charakter dynamiczny, bo zależą przecież od sformułowanych przez użytkownika zapytań (Narojczyk, 2005, s. 38–39).

Nowoczesne teorie wydań cyfrowych źródeł historycznych akcentują konieczność wyjścia poza klasyczny, narracyjny model wykorzystywania tych źródeł, w ramach którego to badacz samodzielnie analizuje udostępniane treści. Jak pisze Dino Buzzetti (2012), mówiąc o edycjach cyfrowych wciąż myślimy o człowieku, chociaż powinniśmy myśleć raczej o tym, czy i jak komputer może przetwarzać do celów badawczych zawarte w nich informacje. Buzzetti zgadza się ze stwierdzeniem, że naczelną ideą edycji cyfrowych powinno być przeniesienie kompetencji czytelnika-badacza na oprogramowanie (Buzzetti, 2012, s. 45).

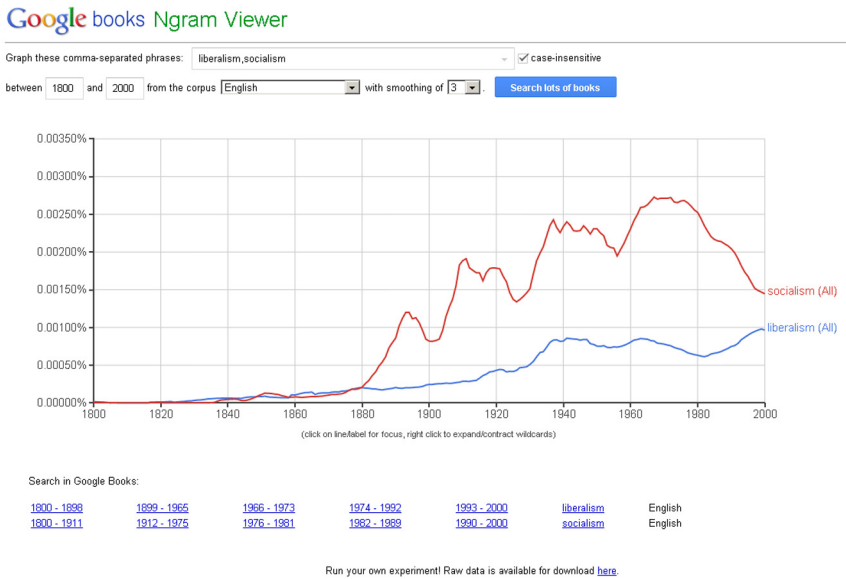
Rosnące zainteresowanie historyków automatycznym przetwarzaniem i analizowaniem dużych korpusów tekstów, obrazów czy zestawów danych wywołuje wiele opinii głoszących moment przejścia historii akademickiej w okres post-teoretyczny. Dyskusje teoretyczne związane z kolejnymi zwrotami w humanistyce przestają być atrakcyjne, z uwagi na szerokie możliwości badawcze, oferowane przez narzędzia cyfrowe i modele współpracy online. Praktykowanie historii cyfrowej – podobnie jak to było w przypadku historyków działających pod koniec XIX wieku – to koncentracja na metodach, narzędziach, na organizacji i przetwarzaniu zbiorów informacji, gdzie zdecydowanie mniejszą wagę poświęca się ideologiom i wielkim narracjom (Scheinfeldt, 2012, s. 124).

Konieczność wyjścia poza badawcze ideologie i popularne dyskursy opisujące przeszłość akcentuje także Lev Manovich i Jeremy Douglas (Manovich, Douglas, 2010), nawołując wprost do zejścia na najniższy poziom opisu: do danych (*we have to turn „culture” into „data”*). Jak pisze, interpretacje w badaniach zawsze motywowane są określonymi przekonaniem, ideologiami czy modami, w praktyce mamy jednak do czynienia zawsze z materialnymi – i dziś dostępnymi w zdigitalizowanej formie – obiektami kultury, przechowywanymi w muzeach, archiwach, galeriach i bibliotekach. Tworząc narracje o historii kultur i społeczeństw posługujemy się abstrakcyjnymi pojęciami. Według wspomnianych autorów historie te można opisywać równie dobrze za pomocą odpowiednich wizualizacji, opartych na analizowaniu ewolucji podstawowych cech obiektów kultury, tj. kształtów, kolorów, struktury wizualnej obrazów, rzeźb, okładek czasopism itp.

Jedną z oponentek takiego programu jest Johanna Drucker. Jej zdaniem konieczny jest ciągły krytyczny namysł nad nowym, cyfrowym kierunkiem badań humanistycznych. Wizualizacje danych historycznych czy automatyczne analizy tekstu mogą być przejawem pozytywistycznego, mechanistycznego, dosłownego i ahistorycznego traktowania zjawisk społecznych i kulturowych. W tego typu badaniach faktycznie odrzuca się podstawowe zasady humanistycznej refleksji, które każą podejrzliwie patrzeć na wszelkie pozornie oczywiste fakty i szukać determinujących je czynników, takich jak: historia, polityka, kultura, procesy społeczne, gospodarcze czy choćby gatunkowość, retoryka itp. Johanna Drucker zwraca też uwagę na niebezpieczeństwo przenoszenia na metody badań humanistycznych ograniczeń związanych z wykorzystywaną w nich technologią. Jej zdaniem, o ile idea książki została historycznie wykształcona w środowisku humanistów, którzy dostosowali to medium do charakteru własnych badań, o tyle

narzędzia cyfrowe, oprogramowanie czy interfejsy są już produktem zupełnie innej kultury wiedzy (Drucker, 2012, s. 85–86).

Wróćmy jeszcze na chwilę do potencjału automatycznej eksploracji tekstu w badaniach historycznych. Skoro można automatycznie analizować zasoby pełnotekstowych archiwów, takich jak London Lives czy Old Bailey Online, dlaczego nie zastosować takiego podejścia do zdecydowanie większych korpusów tekstów? W opublikowanym w 2011 roku w „Science” artykule *Quantitative Analysis of Culture Using Millions of Digitized Books* (Michel i in., 2011) autorzy zaproponowali interesujący model ilościowych badań kultury, który określono pojęciem *kulturomii* (ang. *culturomics*). Przygotowane przez Google narzędzie Ngram Viewer pozwala na wizualizowanie na osi czasu częstości występowania wybranych fraz w korpusach treści książek zdigitalizowanych i transkrybowanych w ramach programu Google Books. Szacuje się, że w ramach projektu Google Books zeskanowano do tej pory i czytano za pomocą OCR około 12 proc. wszystkich książek, opublikowanych drukiem od połowy XV wieku (daje to w sumie około 15 milionów tytułów). Na ich bazie Google zbudowało korpus składający się z ponad 500 miliardów wyrazów, w którym – co nie powinno dziwić – dominuje głównie język angielski (361 miliardów wyrazów). Są w nim także słowa z języków: francuskiego (45 miliardów), hiszpańskiego (45 miliardów), niemieckiego (37 miliardów), chińskiego (13 miliardów), rosyjskiego (35 miliardów) i hebrajskiego



Wykres dla fraz *liberalizm* i *socjalizm* wygenerowany w bazie Ngram Viewer na podstawie treści książek wydanych w latach 1800-2000 i indeksowanych w ramach usługi Google Books.

(2 miliardy). Chociaż w bazach Google Books można znaleźć książki w języku polskim, narzędzie Ngram Viewer nie pozwala na przeszukiwanie korpusu języka polskiego. Wszystkie dostępne tam korpusy mają zresztą ograniczoną objętość – twórcy systemu z zasobów 15 milionów dostępnych książek wybrali treść 5 milionów. Kryterium wyboru była m.in. jakość OCR i metadanych. Wykluczono także – co ważne! – czasopisma.

Zasada działania Ngram Viewer jest dość prosta – system wizualizuje w czasie częstość występowania w określonych korpusach wybranych przez użytkownika słów lub fraz, w odniesieniu do ogólnej liczby wyrazów zindeksowanych w zebranych publikacjach z danego roku. Przykładowo, można dzięki temu sprawdzić, jak w anglojęzycznej literaturze funkcjonowało na skali popularności pojęcie *niewolnictwo* (ang. *slavery*). Analiza wykresu pozwala zaznaczyć wartości maksymalne, pokrywające się z ważnymi wydarzeniami historycznymi i społecznymi (np. dla abolicjonizmu są nimi zniesienie niewolnictwa w koloniach Wielkiej Brytanii w 1833 roku czy amerykańska Wojna Domowa).

Czy takie narzędzie może być pomocne w badaniach historycznych? Ngram Viewer wizualizuje trendy kulturowe jedynie z perspektywy języka, dzięki czemu można przyglądać się ewolucji popularności określonych pojęć czy porównywać ze sobą ich rozmaite zestawy. Twórcy systemu przekonują, że odpowiednio przemyślane użycie tego narzędzia pozwala np. badać historię cenzury, związaną z występowaniem albo niewystępowaniem określonych pojęć (np. nieobecność frazy *Marc Chagall* w literaturze niemieckiej z lat 30. spowodowana była uznaniem kubizmu za sztukę zdegenerowaną). Umożliwia też badanie popularności określonych postaci historycznych, historii adaptacji wynalazków, trendów kulturowych, ideologii, np. socjalizmu czy feminizmu.

Brzmi to bardzo atrakcyjnie, jednak warto uświadomić sobie ograniczenia tej metody. Po pierwsze, z perspektywy badaczy zajmujących się historią Polski poważnym mankamentem tego narzędzia jest brak możliwości analizy korpusu książek w języku polskim. Co prawda, wykorzystać można korpus książek rosyjskich czy niemieckich. Poważniejszy problem wiąże się z faktem, że korpusy analizowane przez Ngram Viewer zbudowane są na bazie jedynie części publikowanych tytułów, z wykluczeniem czasopism czy druków ulotnych. Wadą systemu, bardzo często podkreślaną przez jego krytyków, są podstawowe błędy w automatycznym rozpoznawaniu wyrazów ze skanów, mogące poważnie zaburzać jakość wizualizacji (Hayes, 2011, s. 193).

Interesującą krytykę kulturomii jako metody badań nad kulturą zaproponował Tim Hitchcock, twórca wspomnianych już cyfrowych projektów

archiwalnych. W opublikowanej na swoim blogu naukowym notce odwołuje się do popularności klometrii, przypadającej na lata 70. i 80. XX wieku. Zaproponowany przez twórców Ngram Viewer model badania kultury – w odróżnieniu od tego stosowanego przez wcześniejszych praktyków historycznych badań kwantytatywnych – działa bez odniesienia do historycznej wiedzy o dawnych społeczeństwach, która nadaje odpowiedni kontekst analizom danych liczbowych (Hitchcock, 2011). Nie da się kultury czy procesu historycznego zredukować jedynie do ciągu danych. Chociaż użycie metod statystycznych w naukach historycznych ma długą tradycję, nie mogą być one przecież uznawane za wystarczające. Z drugiej strony – dostępność danych dotyczących kultury pozwala jednak stawiać nowe pytania badawcze, tworzyć wizualizacje, infografiki, kształtować nowe, wizualne formy mówienia o przeszłości.

Ważnym problemem, pojawiającym się w kontekście pracy badawczej, jest dostępność danych i ich otwartość. Ngram Viewer to narzędzie zbudowane przez Google, a korpus tekstów, z którego to narzędzie korzysta, jest także własnością Google, chociaż firma ta nie ma praw do poszczególnych treści wchodzących w skład korpusu. Stąd też system wykorzystuje w swoim działaniu nie standardowe słowa kluczowe, ale n-elementowe zestawy wyrazów (ang. *n-gram*) – których bazy zresztą można pobrać ze strony projektu na jednej z licencji Creative Commons – a sam zestaw n-gramów dla korpusu języka angielskiego zajmuje w skompresowanej formie 340 gigabajtów.

Warto zatem podkreślić to, że pojęcie *open data* związane jest z zagadnieniem statusu prawnoprawa autorskiego danych (badawczych) i nakłada na naukowca konieczność udostępniania ich na zasadach pozwalających wykorzystywać je swobodnie. Dobrym przykładem takiego podejścia jest Europeana, która agregując metadane z bibliotek i archiwów europejskich publikuje własne API, pozwalające na ich przeszukiwanie i wykorzystywanie (także do celów komercyjnych). Na podobnych zasadach udostępnia swoje zbiory i metadane także amerykańskie Walters Art Museum (<http://www.thedigitalwalters.org>).

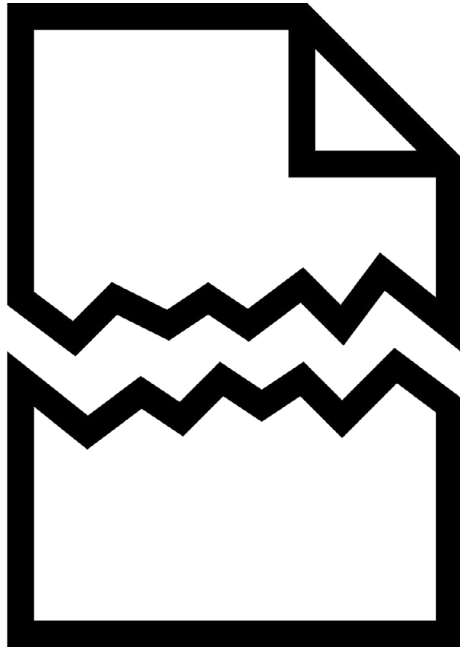
## BIBLIOGRAFIA

- Buzzetti, D.** (2012). Digital editions and text processing. W: Deegan M., Sutherland, K. (red.), *Text Editing Print and the Digital World* (s. 45-61). Burlington: Ashgate Publishing.

- Drucker, J.** (2012). Humanistic theory and digital scholarship. W: Gold M. K. (red.), *Debates in the Digital Humanities* (s.85-95). Minneapolis: University of Minnesota Press. Pobrano z: <http://dhdebates.gc.cuny.edu/debates/text/34>
- Hayes, B.** (2011). Bit Lit: With digitized text from five million books, one is never at a loss for words, *American Scientist*, 99(3), 190-194. Pobrano z: <http://www.americanscientist.org/libraries/documents/2011412130118486-2011-05CompSciHayes.pdf>
- Hitchcock, T.** (2008). Digital Searching and the Reformulation of Historical Knowledge. W: Greengrass, M., Hughes, L. (red.), *The Virtual Representation of the Past*. Burlington: Ashgate Publishing.
- Hitchcock, T.** (2011). *Culturomics, Big Data, Code Breakers and the Casaubon Delusion, Historyonics*. Pobrano z: <http://historyonics.blogspot.com/2011/06/culturomics-big-data-code-breakers-and.html>.
- Jasiński, T.** (2011). *Kronika Polska Galla Anonima w świetle unikatowej analizy komputerowej nowej generacji. Wykłady Inauguracyjne Instytutu Historii Uniwersytetu im. Adama Mickiewicza ; semestr letni 2010/2011, 6*, Instytut Historii UAM: Poznań.
- Manovich, L., Douglass, J.** (2010). *Mapping Time: Visualization of Temporal Patterns in Media and Art*. Pobrano z: [http://softwarestudies.com/cultural\\_analytics/visualizing\\_temporal\\_patterns.pdf](http://softwarestudies.com/cultural_analytics/visualizing_temporal_patterns.pdf)
- Michel, J. B., Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., Pickett, J. P., Hoiberg, D., Aiden, E. L.** (2011). Quantitative Analysis of Culture Using Millions of Digitized Books, *Science*, 331(6014), 176-182.
- Narojczyk, K.** (2005). Internetowe bazy danych w badaniach historycznych. W: Narojczyk, K., Ryszewski, B. (red.), *Metody komputerowe w badaniach i nauczaniu historii*. Olsztyn: Wydawnictwo Uniwersytetu Warmińsko-Mazurskiego.
- Scheinfeldt, T.** (2012). Sunset for Ideology, Sunrise for Methodology? W: Matthew K. Gold (red.), *Debates in the Digital Humanities* (s.124-126). Minneapolis: University of Minnesota Press. Pobrano z: <http://dhdebates.gc.cuny.edu/debates/text/39>.
- Prinke, R. T.** (2000). *Fontes ex machina: Komputerowa analiza źródeł historycznych*. Poznań: Centrum Elektronicznych Tekstów Humanistycznych PAN.







# Cyfrowe dziedzictwo

W październiku 2003 roku Organizacja Narodów Zjednoczonych do Spraw Oświaty, Nauki i Kultury (UNESCO) wydała *Kartę w sprawie Ochrony Dziedzictwa Cyfrowego (Charter on the Preservation of Digital Heritage)*, w której zdefiniowano cyfrowe dziedzictwo jako *źródła informacji i kreatywnej ekspresji wytwarzane, udostępniane i zarządzane w formie cyfrowej*. W zakres definicji UNESCO włączono zasoby cyfrowe powstałe w ramach digitalizacji (posiadające analogowe oryginały) i zasoby nieposiadające oryginalnie formy fizycznej (*born digital*):

Dziedzictwo cyfrowe tworzą unikatowe zasoby ludzkiej wiedzy i formy wyrazu. Obejmuje ono zasoby o charakterze kulturowym, edukacyjnym, naukowym i administracyjnym, jak również informacje o charakterze technicznym, prawnym, medycznym i inne rodzaje informacji tworzone w postaci cyfrowej lub na tę postać konwertowane z dostępnych źródeł analogowych. W przypadku zasobów wytworzonych cyfrowo nie występują one w postaci innej niż obiekt cyfrowy.

Materiały cyfrowe obejmują teksty, bazy danych, fotografie i materiały filmowe oraz dźwiękowe, grafikę, oprogramowanie i witryny internetowe występujące w szerokim i coraz większym zakresie formatów. Mają one często charakter efemeryczny i w związku z tym wymagają odpowiedniego wytwarzania, utrzymania i zarządzania, by mogły przetrwać.

Wiele zasobów cyfrowych to materiały o trwałej wartości i znaczeniu, i dlatego zaliczane są do dziedzictwa zasługującego na ochronę i zachowanie dla współczesnych i przyszłych pokoleń. To ciągle rosnące dziedzictwo może występować w każdym języku, w każdej części świata i w każdej dziedzinie wiedzy lub działalności ludzkiej (UNESCO, 2013).

W niniejszym rozdziale temat dziedzictwa cyfrowego podejmuję więc z perspektywy obu typów zasobów, które dla historyka mogą być wartościowe. W socjologii *dziedzictwo* rozumiane jest z kolei przez pryzmat wartości niematerialnych, np. tradycji czy języka. Tego wątku jednak nie będę podejmował w swoim opracowaniu (o różnicach między historycznym a socjologicznym ujęciem problemu tradycji zob: Szpociński, 2008, s. 29–30).

Coraz więcej źródeł historycznych jest digitalizowanych i udostępnianych online w bibliotekach i archiwach cyfrowych albo w ramach rozmaitych oddolnych i nieformalnych inicjatyw. Mimo że digitalizacja ma charakter masowy, wciąż nie obejmuje większości zgromadzonych zbiorów. Według danych z 2012 roku udostępnianych przez ENUMERATE w raporcie *Survey Report on Digitisation in European Cultural Heritage Institutions*, w europejskich instytucjach kultury zdigitalizowano dotąd około 20 proc. zasobów przeznaczonych do digitalizacji, a w ciągu najbliższych dwóch lat ta wielkość ma się podwoić. Wciąż jednak 57 proc. europejskich zbiorów czeka na cyfryzację. Dane z 2007 roku dotyczące amerykańskich archiwów National Archives and Records Administration (NARA) pokazywały jeszcze większą dysproporcję (Hafner, 2007). W przypadku dokumentów tekstowych zdigitalizowano i udostępniło 8 milionów spośród liczby około 9 miliardów dokumentów archiwalnych (1:1800). W przypadku map było jeszcze gorzej (1:20000). W polskiej Federacji Bibliotek Cyfrowych udostępniono do kwietnia 2013 ponad 1.2 mln publikacji, Narodowe Archiwum Cyfrowe (NAC) przechowuje 15 mln fotografii, 30 tys. nagrań i 2 tys. filmów, co roku wykonując około 10 mln skanów. Dziś na stronie Szukajwarchiwach.pl dostępne są online wersje cyfrowe 7.8 mln dokumentów archiwalnych. Spośród przechowywanych przez NAC fotografii zdigitalizowanych i opublikowanych online jest jedynie nieco ponad 1 proc.

Zakres i rozwój projektów digitalizacyjnych ograniczane są przez brak odpowiednich informacji o zasobach, ograniczone możliwości finansowe instytucji oraz przepisy prawa autorskiego, zwłaszcza w przypadku utworów o niejasnym statusie (tzw. dzieł osieroconych). Dużym wyzwaniem dla badacza są też różne

standardy udostępniania zbiorów i ich metadanych. Nie można więc uznać, że korzystanie z archiwów i bibliotek cyfrowych staje się popularną alternatywą wobec korzystania ze zbiorów na miejscu, tym bardziej że historycy korzystają nie tylko z dokumentów źródłowych, ale też z literatury naukowej. Ta ostatnia z kolei jest słabo reprezentowana w bibliotekach cyfrowych, ze względu na obowiązujące wciąż autorskie prawa majątkowe, blokujące swobodne rozpowszechnianie treści. Sytuację mogłoby poprawić rozprzestrzenienie się modelu otwartego dostępu do literatury naukowej. Szerzej na ten temat piszę w osobnym rozdziale.

Do zasobów dziedzictwa cyfrowego włączyć należy także zbiory audiowizualne. O ich ogromnym znaczeniu kulturowym świadczy chociażby sam fakt ustanowienia przez UNESCO Międzynarodowego Dnia Dziedzictwa Audiowizualnego (ang. *World Day for Audio Visual Heritage*), który przypada 27 października. Zbiory fotografii, filmy, archiwa telewizyjne i radiowe czy amatorskie kolekcje nagrań przechowywane na kasetach VHS stanowić mogą podstawę historycznych analiz. Dostępność tego typu zasobów ograniczana bywa już na etapie korzystania z tradycyjnego archiwum. Przykładowo, archiwa polskiej telewizji publicznej są źle skatalogowane, a pozyskanie nagrań, nawet do celów edukacyjnych, łączy się z opłatami licencyjnymi (Sosnowska, 2012). Historyk ponosi także koszt opłaty związanej z dostępem do wybranego materiału archiwalnego i z jego ewentualnym kopiowaniem – kopia taka oczywiście udostępniana jest bez prawa do rozpowszechniania. Nie są to warunki sprzyjające efektywnemu gromadzeniu materiałów źródłowych do pracy badawczej. Funkcjonowanie repozytorium EUscreen.eu, w którym znajdują się także nagrania TVP, nie rozwiązuje tego problemu ze względu na niewielką liczbę zasobów. Prace nad udostępnieniem archiwów Telewizji Polskiej podjął Narodowy Instytut Audiowizualny – na razie jednak badacze zmuszeni są do osobistych wizyt w archiwum. Warto w tym miejscu zaznaczyć, że ta sytuacja nie musi wynikać bezpośrednio z winy TVP. Zarządzanie archiwum telewizyjnym jest niezwykle trudnym wyzwaniem, choćby ze względu na to, że stan prawny dużej części zasobów nie jest jednoznacznie określony, a prawo autorskie skutecznie blokuje możliwość udostępniania zbiorów archiwum w Internecie i tym samym utrudnia pracę badawczą.

W dniu 8 grudnia 1941 roku, czyli dzień po japońskim ataku na bazę marynarki wojennej USA, Alan Lomax, pracownik archiwum Biblioteki Kongresu, zaczął wysyłać telegramy do zaprzyjaźnionych badaczy w całym kraju z prośbą o zbieranie informacji na temat postaw wobec tego wydarzenia wśród zwykłych

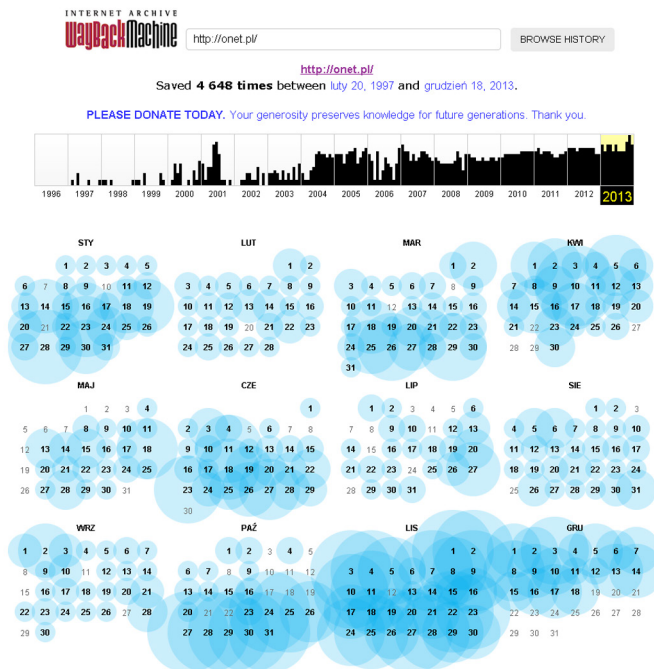
Amerykanów. Wykorzystując ówczesnie dostępne narzędzia do nagrywania głosu, udało mu się w trzy dni uzyskać łącznie prawie pięć godzin wypowiedzi. Była to pierwsza z całej serii podobnych inicjatyw archiwistycznych i badawczych. Sześć dekad później, w samym tylko archiwum stworzonym w ramach projektu Biblioteki Kongresu dokumentującego wydarzenia z 11 września 2001 roku, znalazło się ponad 300 godzin wywiadów oraz 150 tys. dokumentów cyfrowych przekazanych przez zwykłych Amerykanów do otwartego archiwum 911digitalarchive.org. Biblioteka Kongresu we współpracy z organizacją Internet Archive zarchiwizowała także 30 tys. stron internetowych, które na bieżąco opisywały zamachy na USA i dokumentowały reakcje społeczne na to wydarzenie (Cohen, 2005, s. 6–7). Z tego względu 11 września 2001 roku można uznać za wydarzenie, które symbolicznie otwiera nową perspektywę badań historycznych. Jak przekonuje Daniel Cohen, po 11 września wielu historyków, archiwistów, badaczy społecznej musiało na nowo określić, czym może być źródło historyczne i w pewnym stopniu zmodyfikować swoje tradycyjne metody badawcze przystosowane do źródeł analogowych.

Strony WWW, dokumenty elektroniczne czy fotografie cyfrowe to dziś źródła pozwalające w pełniejszy sposób opisać wydarzenie historyczne. Transfer od tradycyjnej do internetowej komunikacji sprawia, że do przeszłości odchodzą dawne rodzaje źródeł historycznych, zastępowane obecnie przez nowe formy publikowania i dzielenia się informacją. Próba napisania analizy historycznej wydarzeń takich jak zamachy na USA z 11 września 2001 roku czy „arabska wiosna” (od 2010 roku) pozostanie niepełna bez – nawet ograniczonego – odwołania się do źródeł internetowych: wpisów na blogach i Twitterze, fotografii umieszczanych w mediach społecznościowych, amatorskich nagrań dostępnych na YouTube itp. Oczywiście tego typu zasoby zainteresują wyłącznie historyków dziejów najnowszych, trudno też oczekiwać, że uda się je odpowiednio wykorzystać i interpretować bez odniesienia do źródeł tradycyjnych i literatury naukowej. Na tworzone współcześnie zasoby Internetu warto spojrzeć z perspektywy badań historycznych. Przykładem takiego podejścia jest zaproponowany przez Jerzego Topolskiego model dynamicznego rozumienia źródła historycznego. Upraszczając, notka na blogu czy fotografia cyfrowa opublikowana online stają się źródłami historycznymi, kiedy tylko historyk jest w stanie, poprzez stawianie właściwych pytań, uzyskać na ich podstawie odpowiednie informacje na temat badanej przeszłej rzeczywistości (zob. Topolski, 1983, s. 257).

Korzystanie ze źródeł *born digital* niesie ze sobą kilka ważnych problemów związanych z dostępem, gromadzeniem, analizowaniem i przechowywaniem tych materiałów. Tylko w ograniczonym stopniu historyk może korzystać przy tym z profesjonalnych archiwów i repozytoriów zasobów tego typu – jak zobaczymy w dalszej partii mojego tekstu, nie jest ich zbyt wiele, gromadzą one też jedynie niewielki wycinek potencjalnie interesujących treści, a system udostępniania często utrudnia korzystanie z nich w celach naukowych. Dlatego historyk musi nauczyć się w odpowiedni sposób gromadzić zasoby internetowe, tworząc na potrzeby badawcze własne archiwum. Potrzebne są tu wiedza i umiejętności dotyczące przeszukiwania i pozyskiwania treści z sieci WWW, ich umiejętnego katalogowania, opisywania i analizowania – także automatycznego (*text mining*).

Przykładem badawczego archiwum Internetu jest Digital Archive for Chinese Studies (DACHS), budowane przez biblioteki instytutów sinologicznych z uniwersytetów w Heidelbergu i Lejdzie. Celem archiwum jest gromadzenie zasobów źródłowych, m.in. do badań nad recepcją oficjalnej polityki władz w chińskim społeczeństwie. Gromadzi ono takie zasoby, jak dokumenty cyfrowe, archiwa z chińskich forów dyskusyjnych, materiały video, internetowe czasopisma i newslettery oraz strony WWW. W 2005 roku było to ponad 36 GB danych, w 2006 – już ponad 60 GB. Ze względu na ograniczenia wynikające z prawa autorskiego, dostęp do archiwum możliwy jest wyłącznie w celach naukowych, po uzyskaniu odpowiedniej zgody i przedstawieniu projektu badań. Hanno E. Lecher w przygotowanym przez siebie studium przypadku opisującym tworzenie DACHS wskazuje na problemy, na które również napotka historyk dziejów najnowszych, chcący stworzyć takie archiwum na własny naukowy użytek (Lecher, 2006).

Wspomniana wyżej fundacja Internet Archive odpowiada za powstanie najbardziej znanej inicjatywy, związanej z archiwizowaniem zasobów WWW. Dzięki niej mamy dostęp do historycznych wersji także polskich stron internetowych, jednak zasięg archiwizacji jest ograniczony, zarówno pod względem liczby zabezpieczonych stron, jak i głębokości ich sczytania (strona główna, podstrony dalszego rzędu, dalsze linki itp.). Co istotne, Internet Archive publicznie udostępnia zgromadzone zasoby. Zasada ta nie zawsze jednak obowiązuje w przypadku tego typu repozytoriów. Przykładowo, archiwum WWW domeny krajowej (.at), tworzone przez Austriacką Bibliotekę Narodową, ze względu na ograniczenia prawnoautorskie przeglądać można tylko na terminalach na miejscu i w wybranych bibliotekach krajowych. Archiwum Internetu tworzone



#### Note

This calendar view maps the number of times <http://onet.pl/> was crawled by the Wayback Machine, *not* how many times the site was actually updated. More info in the [FAQ](#).



The Wayback Machine is an initiative of the [Internet Archive](#), a 501(c)(3) non-profit, building a digital library of Internet sites and other cultural artifacts in digital form. Other projects include [Open Library](#) & [archive.org](#).

Your use of the Wayback Machine is subject to the Internet Archive's [Terms of Use](#).

Archiwalne wersje strony głównej portalu onet.pl z lat 1997-2013, zabezpieczone i udostępniane przez fundację Internet Archive za pomocą narzędzia Wayback Machine.

przez polskie Narodowe Archiwum Cyfrowe dostępne jest publicznie online – znajdują się w nim jednak wyłącznie strony internetowe z domeny rządowej (gov.pl). Dużą barierą w archiwizacji internetu są prawa autorskie. Biblioteka Narodowa, jako polski członek International Internet Preservation Consortium (IIPC), bierze obecnie aktywny udział w pracach nad projektem nowelizacji ustawy o obowiązkowych egzemplarzach bibliotecznych, której zapisy pozwolić by mogły na samodzielne gromadzenie przez Bibliotekę zawartości stron internetowych z domeny .pl, bez konieczności uzyskiwania zgód od wydawców i twórców. Nowelizacje takie wprowadzono w ostatnich latach w wielu krajach Unii Europejskiej (Austria, Dania, Finlandia, Estonia, Łotwa, Francja, Litwa, Niemcy, Słowenia).

Szczególnym przypadkiem, z uwagi na problemy z udostępnianiem archiwum zasobów *born digital*, może być archiwum Twittera. Twitter to amerykański serwis pozwalający na komunikowanie się między użytkownikami za pomocą krótkich, 140-znakowych wpisów, które wysyłać i odczytywać można nie tylko przez standardową przeglądarkę na komputerze, ale także na telefonach komórkowych czy tabletach. Jego mobilny charakter oraz łatwość użycia sprawiły, że stał się ważnym kanałem komunikacji w trakcie wydarzeń „arabskiej wiosny” oraz przestrzeni ich dokumentowania. Znaczenie Twittera jako narzędzia wspierającego bliskowschodnie rewolucje doczekało się ujęcia naukowego (zob. np. Khondker, 2011; Howard, Hussain, 2011; Howard, Parks, 2012 czy bardzo krytycznie Morozov, 2013).

W kwietniu 2010 roku Biblioteka Kongresu ogłosiła rozpoczęcie prac nad archiwizacją zasobów Twittera. W ciągu jednego dnia na Twitterze publikowanych jest około 140 milionów wiadomości (i związanych z nimi metadanych) – to pokazuje skalę problemu, jaki stoi przed archiwistami i badaczami, chcącymi wykorzystać zasoby Twittera w swoich analizach. Pomimo około 400 zgłoszeń od badaczy, do dziś Biblioteka Kongresu nie pozwala na eksplorację ogromnej bazy 170 miliardów tweetów z lat 2006–2010. Samo przeszukanie archiwum za pomocą tylko jednego słowa kluczowego zajmuje 24 godziny (*Update on the Twitter Archive At the Library of Congress 2013*). Problemy z wykorzystaniem naukowym korpusu wiadomości z Twittera mają nie tylko charakter techniczny. Mamy tu bowiem do czynienia z sytuacją, w której prywatna firma staje się dyspozytariuszem jedyne w swoim rodzaju zasobu o dużym znaczeniu społecznym i może dowolnie kreować politykę jego udostępniania do celów naukowych. Badacze próbujący obecnie analizować treści Twittera zmuszeni są do korzystania wyłącznie z tych narzędzi, na jakie pozwala oficjalne API.

Przykład archiwum Twittera daje dobry pogląd na przyszłe wyzwania dla nauki historycznej. O ile coraz większa sfera komunikacji społecznej zapośredniczona będzie przez Internet w infrastrukturze zdominowanej przez komercyjne podmioty, o tyle problem regulacji dostępu do danych historycznych będzie się potęgował. W pracy historyka dużą rolę zaczną odgrywać kompetencje związane z programowaniem i automatycznym analizowaniem dużych zestawów danych. Z drugiej strony, dostęp do tego typu zasobów być może pozwoli spojrzeć na przeszłość z perspektywy źródeł mniej oczywistych i mniej centralnych, za to bardziej oddolnych i zróżnicowanych, oddających punkt widzenia marginalizowanych do tej pory środowisk (Cohen, 2005, s. 10).



Można w tym miejscu zadać także pytanie o to, czy pokolenia przyszłych historyków będą w ogóle mogły skorzystać z generowanych współcześnie źródeł cyfrowych. Dyrektorka Biblioteki Narodowej Nowej Zelandii Penny Carnaby, odpowiedzialna m.in. za program archiwizacji zasobów nowozelandzkiej domeny krajowej, w swoim wystąpieniu podczas międzynarodowej konferencji biblioteczej w 2009 roku wspominała o straconej generacji (ang. *delete generation*), bezpowrotnie tracącej cyfrową część własnej historii:

Wszyscy jesteśmy częścią *generacji delete*. W każdej sekundzie każdej minuty każdego dnia ludzie na całym świecie wymazują własną historię, swoje myśli i opinie, które w dzisiejszych czasach są zawsze prezentowane w środowisku cyfrowym. Nasza wiedza na temat znaczenia tego rodzaju strat nie jest jeszcze dojrzała. Cała generacja może przeminąć, zanim uświadomimy sobie, co to oznacza dla komunikacji wiedzy i idei. Czy rozumiemy, co tracimy i czy jest to dla nas ważne? Jakie są ekonomiczne, społeczne i kulturowe konsekwencje tej straty? (Carnaby, 2009, s. 2)

Być może dziś nie uświadamiamy sobie tego problemu, ale zetkną się z nim z pewnością przyszłe pokolenia historyków. Zadanie związane z archiwizacją zasobów WWW jest o tyle bardziej skomplikowane, że nie jest jedynie zagadnieniem technicznym, ale też problemem politycznym (Illien, 2011). Bardzo trudno jednoznacznie określić granice odpowiedzialności poszczególnych narodowych instytucji pracujących nad zabezpieczeniem treści WWW dla przyszłych pokoleń. Próbą podjęcia współpracy w tym zakresie jest powstanie i działalność IIPC (International Internet Preservation Consortium, Międzynarodowe Konsorcjum Archiwizacji Internetu). Ważną barierą blokującą rozwój inicjatyw wieczystej archiwizacji dokumentów elektronicznych są koszty: przykładowo, stworzony przez amerykańskie archiwa narodowe (NARA) system Electronic Records Archive kosztował 567 mln dolarów i obecnie gromadzi rocznie około 40 terabajtów danych. W przygotowaniu jest jednak m.in. import dokumentów ze spisu ludności USA z 2010 roku o objętości około 500 terabajtów (Thomas i Johnson, 2012, s. 174).

Problem leży jednak nie tylko w zachowaniu zasobów cyfrowego dziedzictwa. Zasoby te wymagają przecież także narzędzi pozwalających je odczytać. Dokumenty zapisane w starych formatach nie zawsze mogą być otwarte za pomocą dostępnych dziś programów. Odczytanie niektórych treści wymaga nie

tylko odpowiedniego programu, ale także dostępu do starego komputera, obsługującego nieużywane dziś nośniki danych, np. dyskietki 8- czy 3,5-calowe. Wobec dynamicznego rozwoju oprogramowania i sprzętu komputerowego coraz ważniejszą – w tym również badawczą – rolę pełnią muzea komputerów i oprogramowania, które gromadzą, zabezpieczają i mogą udostępniać historykom archiwalne wersje programów i sprzętu.

Warto dodać, że pierwsza polska instytucja tego typu została otwarta w grudniu 2012 roku. Muzeum Historii Komputerów i Informatyki mieści się dziś w tymczasowej siedzibie, udostępnionej w Wyższej Szkole Mechatroniki w Katowicach-Szopienicach. Wiedzę na temat archiwalnych wersji oprogramowania i komputerów, także tych produkowanych w Polsce, uzyskać można w Stowarzyszeniu Miłośników Zabytków Informatyki, prowadzącym m.in. Polski Rejestr Zabytków Informatycznych, lub w Sekcji Historycznej Polskiego Towarzystwa Informatycznego.

Ważnym elementem cyfrowego dziedzictwa historycznego, ze względu na swoją społeczną rolę i zasięg oddziaływania, są gry komputerowe. Amerykański historyk Jeremy K. Saucier przekonuje, że gry komputerowe podlegać mogą badaniom historycznym (Saucier, 2012), ponieważ, podobnie jak filmy, powstają w określonych warunkach społeczno-politycznych, a w ich treści odbijają się dominujące w społeczeństwie idee, stereotypy, mity, wspomnienia wydarzeń (np. II wojna światowa, katastrofa w Czarnobylu). Saucier zwraca uwagę także na to, że gry często stanowią narzędzie oddziaływania rozmaitych instytucji – zresztą sam opisuje to zjawisko na przykładzie internetowych gier strategicznych tworzonych przez armię amerykańską. Według niego pytaniem nie jest tylko to, w jaki sposób historycy powinni badać takie teksty kultury, ale też to, czy będą mieli do nich dostęp. To ostatnie jest szczególnie ważne w przypadku gier starych i produkowanych przez niewielkie, niszowe firmy o niewielkim zasięgu. Czasem jedynym źródłem dostępu stają się nielegalne repozytoria internetowe tworzone z pominięciem ograniczeń prawnoautorskich. Czy możemy mówić – jak proponuje historyk gier komputerowych Benj Edwards – o zabezpieczającym efekcie piractwa gier komputerowych (dosłownie *piracy's preserving effect*)? Edwards przekonuje, że zabezpieczanie software'owego dziedzictwa nie może być domeną jednej czy nawet kilku instytucji, ze względu na dostępność i bezpieczeństwo takich pojedynczych zbiorów. Nieautoryzowane kopiowanie neutralizuje te problemy. Oczywiście dziś oznacza to złamanie prawa, przynajmniej z punktu widzenia zamkniętego, komercyjnego oprogramowania (Edwards, 2012).

Gry rozumiane jako narzędzie rozrywki, a także – co pokazuje Saucier – jako forma transmisji rozpowszechnionych w społeczeństwie idei, stereotypów czy postaw, nie są dziś przypisane jedynie do pojedynczych komputerów, na których instalują je gracze. Grać można także przez Internet, co sprawia, że gra staje się przestrzenią bezpośredniej komunikacji, w ramach której obserwować można określone społeczne zjawiska. Niektóre gry, takie jak Sims Online czy World of Warcraft, doczekały się już wielu naukowych opracowań. Czy tymi wirtualnymi światami również powinni zainteresować się historycy? Henry Lowood, historyk pracujący w bibliotece Uniwersytetu Stanforda, w swoim artykule *Memento Mundi: Are Virtual Worlds History?* przekonuje, że mogą być one przedmiotem badań historycznych przynajmniej z dwóch względów. Po pierwsze, są historyczne, ponieważ wiele z nich dziś już nie istnieje. Przykładowo, w 2008 roku zamknięto utrzymywany przez firmę Electronic Arts wirtualny świat The Sims Online. Można więc badać je z perspektywy czasu pod kątem ich ewolucji technologicznej, profilu gatunkowego czy stopnia popularności wśród graczy (Lowood, 2009).

Gry MMOG (ang. *Massively Multiplayer Online Game*) wyznaczają z kolei przestrzeń interakcji między ich użytkownikami. Lowood zwraca uwagę na to, że najważniejszą rzeczą jest tam właśnie aktywność graczy, określona przez ich działania, postawy i wzajemne interakcje w środowisku gry. Drugi argument za uwzględnieniem tego zjawiska w perspektywie badań historycznych jest następujący: przyszłych historyków interesować mogą np. motywacje graczy organizujących wspólnie protest wobec zmiany zasad funkcjonowania gry, jak to było w przypadku protestów w Second Life w 2003 roku, czy reakcje na zamachy terrorystyczne na USA z 11 września 2001 roku, obserwowane w grach takich, jak Everquest czy Asheron's Call. Dodać do tego można perspektywę badania biografii tych osób, w przypadku których spora część aktywności realizowana jest wyłącznie w grach online.

W 2000 roku na uniwersytecie Stanforda uruchomiono projekt The How They Got Game, którego celem było rozpoczęcie badań nad możliwością analizy historycznej gier MMOG oraz wynikającymi z tego problemami rejestracji i archiwizacji obserwowanych w nich zjawisk społecznych. Osiem lat później inicjatywa ta rozwinęła się, dzięki wsparciu Biblioteki Kongresu, w międzynarodowy projekt badawczo-archiwalny Preserving Virtual Worlds. Część zbiorów projektu, wśród nich choćby rejestracja ostatnich chwil istnienia środowiska The Sims Online, udostępnianych jest w repozytorium Internet Archive.

Na koniec chciałbym zaproponować jeszcze jedną perspektywę spojrzenia na dziedzictwo cyfrowe. Programy komputerowe, gry, język stron WWW, protokoły Internetu podlegać mogą analizie nie tylko jako nośniki treści, ale także jako konstrukcje informatyczne reprezentujące określone cechy, które mogą mieć istotne znaczenie społeczne – a to już jest zagadnienie, które może zainteresować historyka:

Fundament studiów nad oprogramowaniem stanowi założenie, że tradycyjne podejście do badania mediów, obok właściwości środków przekazu skoncentrowane na przekazywanych w nich tekstach i tym, co robią z tymi tekstami odbiorcy, nie wystarcza – jest opisem efektów, a nie przyczyn. Pomija bowiem kod komputerowy, ukrytą warstwę warunkującą funkcjonowanie środków przekazu (Filiciak, 2008, s. 15–16).

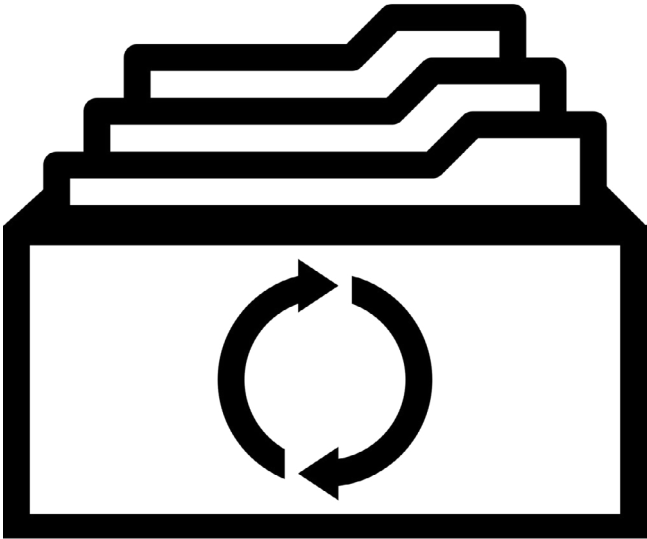
Historyk analizujący media w swoich badaniach może zejść poniżej poziomu dostępnej z zewnątrz treści, aby dostrzec i krytycznie ocenić budujące je interfejsy i algorytmy.

## BIBLIOGRAFIA

- Carnaby, P.** (2009). *Citizen-created content, digital equity and the preservation of community memory*, World Library and Information Congress: 75th IFLA. Pobrano z: <http://conference.ifla.org/past/ifla75/146-carnaby-en.pdf>.
- Cohen, D.J.** (2005). The Future of Preserving the Past, *CRM: The Journal of Heritage Stewardship*, 2(2), 6-19.
- Edwards, B.** (2012). *Why History Needs Software Piracy. How copy protection and app stores could deny future generations their cultural legacy*, Technologizer. A smarter Take on Tech. Pobrano z: <http://technologizer.com/2012/01/23/why-history-needs-software-piracy/>
- Filiciak, M.** (2008). Zajrzeć pod powierzchnię ekranu. Wprowadzenie do studiów nad oprogramowaniem, *Kultura Popularna*, 4(22), 15-18.
- Hafner, K.** (2007). History, Digitized (and Abridged), *New York Times*. Pobrano z: <http://www.nytimes.com/2007/03/10/business/yourmoney/11archive.html>

- Howard, P.N., Hussain, M.M.** (2011). The Role of Digital Media, *Journal of Democracy*, 22(3), 35-48.
- Howard, P.N., Parks, M.R.** (2012). Social Media and Political Change: Capacity, Constraint, and Consequence, *Journal of Communication*, 62(2), 359-362. Pobrano wersję zaktualizowaną z 2013 roku: [http://www.loc.gov/today/pr/2013/files/twitter\\_report\\_2013jan.pdf](http://www.loc.gov/today/pr/2013/files/twitter_report_2013jan.pdf)
- Illien, G.** (2011). Une histoire politique de l'archivage du web: le Consortium International pour la Preservation de l'Internet, *Bulletin des Bibliothèques de France*, 56(2), 60-68.
- Khondker, H.H.** (2011). Role of the New Media in the Arab Spring, *Globalizations*, 8(5), 675-679.
- Lecher, H.E.** (2006). Small Scale Academic Web Archiving: DACHS. W: Masanès, J. (red.), *Web Archiving*, (s. 213–225). Berlin, Heidelberg, New York: Springer.
- Lowood, H.** (2009). *Memento Mundi: Are Virtual Worlds History?*, iPres Sixth International Conference on Preservation of Digital Objects Joint Annual Meeting of the Society of American Archivists and the Council of State Archivists, San Francisco. Pobrano z: <http://www.escholarship.org/uc/item/2gs3p6jx>
- Morozov, E.** (2013). *To save everything, click here: Technology, solutionism, and the urge to fix problems that don't exist*. London: Allen Lane.
- Saucier, J.K.** (2012). Playing the Past: The Video Game Simulation as Recent American History. W: Potter C.B. i Romano R.C. (red.), *Doing recent history: On privacy, copyright, video games, institutional review boards, activist scholarship, and history that talks back* (s. 201-224), Athens: University of Georgia Press.
- Sosnowska, A.** (2012). Zamknięte archiwa publiczne TVP. Relacja z researchu, *Mała Kultura Współczesna*, 17.03.2012. Pobrano z: <http://malakulturawspolczesna.org/2012/03/17/agnieszka-sosnowska-zamkniete-archiwa-publiczne-tvp-relacja-z-researchu/>
- Szpociński, A.** (2008). Nowe formy tradycji. W: Krzemień-Ojak S. (red.), *Przyszłość tradycji* (s. 29–36). Białystok : Libra Wydawnictwo i Drukarnia.
- Topolski J.** (1983). *Teoria wiedzy historycznej*, Poznań: Wydawnictwo Poznańskie.
- Thomas, D., Johnson, V.** (2012). New universes or black holes? Does digital change anything? W: T. Weller (red.), *History in the Digital Age* (s. 174-193). London, New York: Routledge.

**UNESCO (2013)**, Karta w sprawie ochrony dziedzictwa cyfrowego, Konferencja Generalna UNESCO, 32 sesja: 2003, rezolucja nr 42 z dnia 15 października 2003, Pobrano z: <http://www.archiwa.gov.pl/pl/dokument-elektroniczny/961-karta-unesco.html>



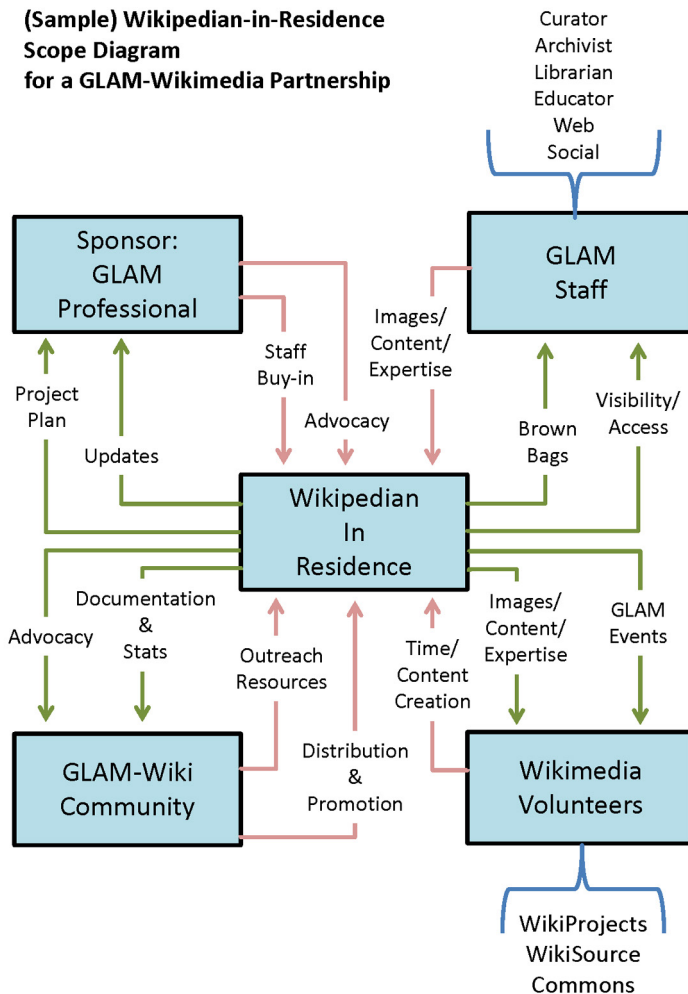
# Instytucje dziedzictwa w kulturze cyfrowej

Galerie, biblioteki, archiwa i muzea określa się czasami wspólnym terminem GLAM (ang. *Galleries, Libraries, Archives, Museums*) lub LAM (ang. *Libraries, Archives, Museums*). Pojęcia te funkcjonują dziś głównie w odniesieniu do projektów współpracy między Wikipedią a tymi instytucjami (GLAM-Wiki), jednak oczywiście można ich użyć poza tym kontekstem. Nie są to też pojęcia oznaczające zamkniętą grupę instytucji, a w dyskusjach czy opracowaniach wykorzystywane są głównie jako zwięzłe określenie instytucji kultury i dziedzictwa.

Z punktu widzenia historyka wiedza o kondycji instytucji kultury zarządzających materialnymi zasobami dziedzictwa historycznego czy udostępniających literaturę naukową jest niezwykle ważna. W kulturze cyfrowej archiwa, muzea, galerie sztuki i biblioteki podlegają ogromnym zmianom. Z instytucji zamkniętych, niedostępnych, posiadających niemal sakralny status stają się przestrzeniami otwartymi, udostępniającymi swoje zasoby, intensyfikującymi komunikację z odbiorcami i animującymi ich aktywność. Wspomniane już wyżej inicjatywy GLAM-Wiki to projekty współpracy między instytucjami kultury a środowiskiem Wikipedii. Dziś taką formę wspólnych działań na rzecz zwiększenia dostępności i popularyzacji zbiorów podejmują 64 archiwa, biblioteki, muzea i galerie, także z krajów europejskich: Niemiec, Francji, Norwegii, Wielkiej Brytanii i Holandii. Z wikipedystami współpracują m.in. Biblioteka Kongresu, amerykańskie archiwa narodowe (National Archives and Records Administration, NARA) oraz brytyjskie archiwa narodowe (The National Archives), Szwajcarskie Archiwa Federalne (Schweizerischen Bundesarchiv),



**(Sample) Wikipedian-in-Residence  
Scope Diagram  
for a GLAM-Wikimedia Partnership**



Schemat współpracy między społecznością Wikipedii a pracownikami instytucji GLAM w ramach modelu wikipedian in residence. Opracowała Lori Byrd Phillips (CC BY-SA)

biblioteki narodowe Szkocji i Holandii, liczne muzea (w tym Muzeum Picassa w Barcelonie, British Museum, Muzeum Archeologiczne w Hamburgu czy nowojorskie Museum of Modern Art). W Polsce ze Stowarzyszeniem Wikimedia Polska współpracują Narodowy Instytut Dziedzictwa, Archiwum Główne Akt Dawnych oraz Kancelaria Senatu, rozpoczęcie współpracy deklaruje także Narodowe Archiwum Cyfrowe, którego zdigitalizowane zbiory z domeny publicznej już teraz w nieformalny sposób pojawiają się na Wikimedia Commons.

Współdziałanie w ramach GLAM-Wiki przybierać może np. postać kilkumiesięcznego zatrudnienia wikipedysty w instytucji w celu edukacji zespołu

i wsparcia udostępniania zbiorów w repozytorium Wikipedii oraz tworzenia nowych haseł (to tzw. *wikipedian in residence*). Współpraca instytucji kultury z Wikipedią możliwa jest także w przypadku takich inicjatyw jak organizowanie wycieczek wikipedystów do muzeów i galerii w celu fotografowania (często niedostępnych dla zwykłych zwiedzających) zbiorów i udostępniania ich w Wikipedii (*Wikipedia Loves Art*) czy akcji wspólnego skanowania zbiorów (*scan-a-thon* organizowany wielokrotnie w waszyngtońskiej siedzibie amerykańskich archiwów narodowych – NARA). W ramach GLAM-Wiki w latach 2010-2011 pozyskano 194 tys. fotografii, którymi zilustrowano ponad 2 tys. artykułów.

Chociaż wciąż nie jest to trend powszechny, czołowe instytucje GLAM, takie jak: Biblioteka Kongresu, amerykańskie archiwa narodowe National Archives and Records Administration (NARA), British Library i brytyjskie archiwa narodowe The National Archives wypracowują nowe standardy otwartości i nowe modele działania, które kopiowane są przez mniejsze instytucje na całym świecie. Profesjonalne instytucje GLAM nie działają w próżni – otaczają je niezliczone oddolne, „amatorskie” projekty i inicjatywy archiwalne, muzealne czy biblioteczne, rozwijane przez organizacje pozarządowe, społeczności lokalne, nieformalne grupy inicjatywne czy indywidualne osoby, korzystające z możliwości, jakie daje Internet. Jedną z takich inicjatyw jest zestaw zaleceń OpenGLAM Principles (*Zasady otwartych instytucji kultury*), którego wersja 1.0 została niedawno wypracowana w społeczności animowanej przez Open Knowledge Foundation (<http://openglam.org/principles/>). OpenGLAM Principles wprowadza konkretne standardy otwartości instytucji kultury i dziedzictwa, podkreślając konieczność gwarantowania pełnych praw dostępu i ponownego wykorzystania zbiorów z domeny publicznej oraz opisujących je metadanych.

Współpraca instytucji GLAM z Wikipedią wymusza taką otwartość, Wikipedia bowiem akceptuje i korzysta jedynie z treści dostępnych na wolnych licencjach lub w domenie publicznej. Efektem tego współdziałania jest nie tylko większa dostępność zbiorów, ale też lepsza promocja samej instytucji. Amerykańskie Walter Art Museum przekazało do repozytorium Wikimedia Commons 18 tys. skanów dokumentów i fotografii obiektów wraz z opisami. Obejmują one okres do końca XIX wieku. Po udostępnieniu zbiorów wykorzystywano je do uzupełnienia 1357 haseł w Wikipedii, publikowanych w 40 językach i wyświetlonych ponad 10 mln razy. Jeszcze przed zainicjowaniem projektu serwis Walters Art Museum generował około 2 tys. wizyt dziennie. Po

skopiowaniu części skanów do Wikimedia Commons popularność repozytorium na stronach Walters Art Museum wzrosła o ponad 500 proc. w porównaniu ze stanem przed udostępnieniem zbiorów do Wikipedii.

Digitalizacja niesie ze sobą spore wyzwania. Programy masowej digitalizacji zasobów dziedzictwa historii i kultury bardzo często pozbawione są elementów edukacyjnych, pozwalających przygotować odbiorców do pełnego, aktywnego i kreatywnego korzystania z udostępnianych online materiałów, co w tej sytuacji staje się ważnym zadaniem dla edukacji historycznej. Postępująca digitalizacja i popularność cyfrowych repozytoriów wywołują też zjawisko analogowego wykluczenia. Skoro wygodny dostęp online staje się preferowany, to zasoby, do których dotrzeć można wciąż jedynie w tradycyjny sposób, to jest za pośrednictwem bibliotek czy archiwów, mogą być niezauważane i pomijane.

Nic więc dziwnego, że pojawiają się głosy mówiące o potrzebie powrotu do fizycznych przestrzeni i fizycznych obiektów. W wydanym kilka lat temu raporcie American Association of Museums (AAM) i Center for the Future of Museums (CFM) przewiduje się, że wobec postępującej cyfryzacji kultury muzea stać się powinny rezerwuarami autentyczności (AAM i CFM 2008, s. 15), udostępniającymi fizyczne obiekty i pozwalającymi na bezpośredni kontakt z nimi. Dużym problemem staje się też niska aktywność użytkowników GLAM-ów, ponieważ w wielu przypadkach uniemożliwia ona realizację pewnych działań, takich jak wspólne opisywanie zasobów czy kreatywne korzystanie z danych udostępnianych przez instytucję (Werla, 2012). Koncepcja Web 2.0, tak popularna jeszcze kilka lat temu i wykorzystywana do projektowania nowej komunikacji między instytucjami GLAM a ich odbiorcami, nawet jeśli dziś nie wydaje się być w odwrocie, to jest traktowana z dużą dozą rezerwy.

Z punktu widzenia badacza historia otwartość instytucji GLAM może mieć podstawowe znaczenie przy planowaniu i prowadzeniu badań. Jej pierwszym elementem jest na pewno dobrze przygotowana, czytelna, użyteczna i aktualna strona internetowa, pozwalająca szybko zorientować się w zasadach korzystania z biblioteki czy archiwum i w charakterze zbiorów. Uzupełnieniem oferty informacyjnej mogą być blogi prowadzone przez indywidualnych pracowników instytucji lub reprezentujące pracę całych wydziałów, jak to odbywa się w przypadku Biblioteki Kongresu czy NARA. Narzędzia takie jak newsletter czy kanał RSS pozwalający w łatwy sposób otrzymywać najnowsze informacje ze strony to standardowe rozwiązania usprawniające komunikację. Ale komunikacja ta nie musi odbywać się wyłącznie między użytkownikiem (badaczem) a instytucją.

Our Archives ([www.ourarchives.wikispaces.net](http://www.ourarchives.wikispaces.net)) to platforma internetowa oparta o model wiki, pozwalający na łatwe dodawanie i edytowanie treści stron WWW. Amerykańskie Archiwa Narodowe (NARA) przygotowały ten serwis z myślą o komunikacji między badaczami korzystającymi z udostępnianych przez nich zbiorów. Dzięki temu osoby te wymieniać się mogą informacjami na temat dostępu do określonych dokumentów, tworzyć własne katalogi czy współpracować z innymi prowadzącymi badania na podobny temat. Na nieco podobnej zasadzie funkcjonuje katalog archiwum Saari Manor, wspomniany już w jednym z poprzednich rozdziałów, przy czym historycy korzystający z jego zbiorów sami zaprojektowali sposób ich katalogowania i uzupełniają przygotowane przez pracowników podstawowe informacje o poszczególnych dokumentach.

Oczywiście dla historyka najbardziej użyteczną formą otwartości instytucji GLAM będzie swobodny dostęp do źródeł, czyli np. dokumentów i obiektów muzealnych oraz literatury naukowej. Nie wszystko może być jednak swobodnie publikowane online – na drodze postępu digitalizacji i udostępniania zasobów stoją także kwestie finansowe, zasady ochrony danych osobowych czy stan opisu zbiorów.

Ze względu na ograniczenia prawa autorskiego biblioteki umieszczają w swoich internetowych repozytoriach najczęściej książki z domeny publicznej, które nie zawsze są użyteczne z punktu widzenia współczesnych badań. Korzystanie z trudno dostępnej książki czy czasopisma nadal wiąże się często z koniecznością wyjazdu do innego miasta czy z długim czasem oczekiwania na dostęp do publikacji. Na szczęście pojawiają się takie usługi biblioteczne, jak serwis zdalnego zamawiania kopii z materiałów bibliotecznych, uruchomiony przez Bibliotekę Uniwersytecką w Poznaniu. W serwisie Libsmart Copy zamówić można skan dowolnej publikacji z katalogu biblioteki. Usługa jest płatna, ale pozwala bardzo szybko skorzystać z treści publikacji – po uiszczeniu opłaty skan można ściągnąć ze strony biblioteki. Jest on zabezpieczony znakiem wodnym, zawierającym imię i nazwisko oraz adres email zamawiającego, i może być wykorzystywany jedynie w ramach dozwolonego użytku. Podobne usługi dotyczyć mogą także archiwaliów. W podobnym modelu funkcjonuje program eBooks on Demand (EOD), prowadzony przez 30 europejskich bibliotek z 12 krajów. Usługa ta umożliwia zdalne zamawianie przygotowania i udostępnienia cyfrowej wersji dowolnej książki z domeny publicznej, o ile ta oczywiście znajduje się w katalogu współpracujących bibliotek.

Wiele instytucji GLAM współpracuje w celu udostępniania wspólnego interfejsu wyszukiwania w swoich zasobach. Przykładem takiego rozwiązania jest wyszukiwarka polskiej Federacji Bibliotek Cyfrowych, agregująca ponad milion obiektów z ponad 100 bibliotek, archiwów i repozytoriów. Dane z FBC eksportowane są do wyszukiwarki Europejskiej Biblioteki Cyfrowej Europeana, gdzie przeszukiwać można zasoby instytucji z całego kontynentu. Uruchomiony przez Narodowe Archiwum Cyfrowe portal Szukajwarchiwach.pl umożliwia przeglądanie części zbiorów z 13 archiwów, m.in. Archiwum Akt Nowych, archiwów państwowych w Lublinie i Poznaniu, Archiwum m. st. Warszawy, Archiwum Muzeum Zamoyskich w Kozłowie czy Archiwum Hoovera. Swoje zbiory fotograficzne i dźwiękowe publikuje online Narodowe Archiwum Cyfrowe. Naturalnie oprócz zdigitalizowanych zbiorów archiwa i muzea udostępniają także liczne katalogi internetowe, bazy danych itp.

Otwierając dostęp do swoich zasobów instytucje GLAM podejmuje czasem współpracę z Wikipedią, którą, ze względu na zasięg tego serwisu, uznać można za jedno z najważniejszych repozytoriów zasobów o charakterze historycznym – pomimo jego oczywistych wad związanych z poziomem opisów czy brakiem jednolitego standardu metadanych. Przykładem takiej współpracy są archiwa niemieckie (Bundesarchiv), które w 2008 roku umieściły w repozytorium Wikimedia Commons prawie 100 tys. fotografii archiwalnych. Są one teraz dostępne swobodnie na jednej z wolnych licencji Creative Commons, umożliwiającej m.in. ich komercyjne wykorzystywanie. Opisy fotografii również przejęto bezpośrednio z Bundesarchiv, co gwarantuje ich merytoryczną jakość. Poza Wikipedią zdjęcia te dostępne są także w oficjalnym serwisie <http://www.bild.bundesarchiv.de>, gdzie pobrać je można w plikach bardzo dobrej jakości. Ponieważ Wikimedia Commons już dziś gromadzi wybrane zbiory z wielu instytucji GLAM, stanowi ona rodzaj wspólnego globalnego repozytorium, które jednak – co trzeba podkreślić – nie udostępnia jedynie zbiorów o charakterze historycznym. Jest to jednak na pewno dobre źródło materiałów historycznych, do wykorzystania choćby w szkolnym nauczaniu historii.

Otwartość instytucji GLAM polegać może nie tylko na udostępnianiu samych zbiorów w formie skanów, ale także innego rodzaju zasobów. Przykładowo, biblioteki zrzeszone w FBC pozwalają skorzystać z gotowych zestawów metadanych, dostępnych za pomocą protokołu

OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting). Zestawy te wykorzystane mogą być choćby do skutecznego przeszukiwania zbiorów lub przygotowywania wizualizacji. Instytucje mogą publikować także takie dane, jak: zestawienia ekonomiczne, geograficzne, statystyczne itp. (przykładowe rodzaje zestawów danych z holenderskich instytucji GLAM dostępne są na stronie <http://www.opencultuurdata.nl/datasets/>). Istotną cechą danych eksportowanych z instytucji GLAM powinna być ich otwartość – muszą być one rozpowszechnione na takich zasadach prawnych, które pozwalają na ich swobodne wykorzystywanie i republikowanie, także w projektach komercyjnych.

Trudno na kilku stronach opisać problem ewolucji instytucji GLAM, gdyż składa się na niego wiele wątków technicznych, organizacyjnych, finansowych czy politycznych. Z pewnością w ofercie studiów historycznych powinny pojawić się kursy pozwalające studentom zdobyć odpowiednią wiedzę i umiejętności związane z korzystaniem z, nowej internetowej oferty bibliotek, archiwów i muzeów. Oferta ta, jak pokazuje to wspomniany już prof. Tomasz Jasiński z Biblioteki Kórnickiej PAN, niekiedy w radykalny sposób usprawnia prowadzenie badań i pozwala zwiększyć ich skuteczność, także przez to, że źródła dostępne w formie cyfrowej mogą być automatycznie przetwarzane i analizowane:

Jeden z polskich mediewistów zapytał mnie, czy przytoczone zdanie z pewnej kroniki można w taki a nie inny sposób przetłumaczyć. Kilka chwil poszukiwań po Internecie i dysponowałem tekstem cyfrowym tej kroniki; następnie zestawilem wszystkie zdania z tej kroniki, w których występował dwuznaczny wyraz, który utrudniał zrozumienie przysłanego mi zdania. W chwilę później tekst został poprawnie przetłumaczony. Następnie w wyszukiwarce znalazłem sto siedemdziesiąt kilka pozycji napisanych przez uczonych polskich, niemieckich, francuskich, węgierskich, rosyjskich etc., którzy analizowali interesujące zdanie. Oczywiście, w każdej pozycji bibliograficznej podałem dokładną stronę i wysłałem odpowiedź mojemu koledze. Najbardziej zdumiała go owa lista stu siedemdziesięciu pozycji bibliograficznych (Jasiński, 2010, s. 12).

Przywołany tutaj przykład wskazuje również na to, że historyk prowadzący badania z użyciem Internetu nie powinien ograniczać się do korzystania wyłącznie ze znanych już sobie repozytoriów i narzędzi, ale nieustannie się uczyć

i poznawać nowe. Jednak to nie wszystko: Internet wytworzył przecież także nowe formy komunikacji i współpracy między badaczami – piszę o tym w trzeciej części podręcznika.

Zmianę, jaka zaszła w formule pracy historyka w związku z postęпами digitalizacji i możliwością skorzystania z narzędzi cyfrowych, w bardzo czytelny sposób przedstawiają William J. Turkel, Kevin Kee i Spencer Roberts, podsumowując swoje doświadczenia z zajęć dydaktycznych z metodologii historii od lat 80. XX w. aż do dziś. Stary schemat badań streszczają oni w następujących punktach (Turkel i in., s. 61):

1. Przeprowadź systematyczny przegląd literatury przedmiotu;
2. Sformułuj pytania badawcze;
3. Przeprowadź odpowiednie analizy w archiwum, przepytaj świadków historii itp., w zależności od specyfiki przedmiotu badania;
4. Opanuj dogłębnie badane zagadnienie;
5. Opisz swoje wyniki w ramach publikacji pokonferencyjnej, artykułu w drukowanym czasopiśmie lub w monografii;
6. Zanalizuj najnowszą literaturę przedmiotu w poszukiwaniu informacji istotnych dla wyników twojego badania;
7. Opublikuj ostateczne wyniki pracy.

Kiedy dostępne stają się źródła i literatura przedmiotu w formie cyfrowej, wtedy można skorzystać z wyszukiwarek, narzędzi do automatycznej analizy treści czy wizualizacji. Pojawia się też możliwość zdalnej współpracy z innymi badaczami. W takim przypadku podstawy i warunki pracy historyka można opisać następująco (zob. Turkel i in., 2012, s. 64–65):

1. Zdigitalizuj wszystko;
2. Przechowuj swoje zasoby w chmurze (tj. online na serwerach, tak aby mieć do nich dostęp z każdego komputera);
3. Zarządzaj gromadzonymi cytatami z literatury w ramach odpowiedniej bazy danych (np. menadżera bibliografii);
4. Pamiętaj, że interesujące cię informacje wciąż mogą do ciebie docierać (np. za pomocą subskrybowanych kanałów RSS czy newsletterów naukowych);

5. Pamiętaj, że uwaga to najrzadszy zasób;
6. Współpracuj z innymi;
7. Dziel się wynikami swojej pracy.

Ważnym kontekstem działalności współczesnych instytucji dziedzictwa jest prawo autorskie. Kwestie tę przywołuje w niniejszym podręczniku wielokrotnie, akcentując głównie ograniczenia, jakie wynikają z zapisów prawa autorskiego dla pracy naukowej czy udostępniania treści archiwalnych online. Nic w tym dziwnego, jeśli uświadomimy sobie, że głównym celem tych przepisów jest ochrona praw twórcy, a nie odbiorców jego dzieł, nawet jeśli on sam i jego spadkobiercy od lat już nie żyją. Ochrona ta trwa w Polsce z reguły 70 lat. Dostrzec można przy tym starania polityczne, których celem jest dalsze wydłużenie tego okresu. Efektem tego jest sytuacja, w której wiele potencjalnie wartościowych badawczo zasobów, np. rękopisów, fotografii, filmów, nie może być swobodnie udostępnianych online przez biblioteki czy archiwa.

Problemem wpływu systemu praw autorskich na praktykę naukową historyka i współczesną kulturę zajmuje się m.in. amerykańska badaczka Gail Drakes. W jednym ze swoich artykułów opisuje ona przypadek spuścizny Martina Luthera Kinga: jego spadkobiercy działający przez King Center w Atlancie restrykcyjnie kontrolują możliwość wykorzystywania jej w badaniach naukowych. W 1987 roku rodzina Kinga wytoczyła proces Uniwersytetowi Bostońskiemu, któremu King osobiście zapisał część swojego archiwum osobistego. W 2009 roku na skutek porozumienia rodziny z koncernem EMI, jednym z czterech czołowych potentatów na rynku muzycznym, ograniczono możliwość nieautoryzowanego wykorzystywania archiwalnych nagrań Kinga w Internecie. Umowa obejmowała także kontrolę nad przytaczaniem słynnego zwrotu z 1963 roku (*I had a dream...*), który formalnie w USA pozostaje pod ochroną autorskich praw majątkowych do 2038 roku (Drakes, 2012, s. 101). Mamy więc do czynienia z sytuacją, w której zasoby dokumentujące nawet przełomowe momenty czy procesy historyczne pozostają pod kontrolą osób prywatnych, instytucji publicznych czy nawet nastawionych na zysk firm działających na rynku komercyjnym.

Na szczęście biblioteki i archiwa mogą swobodnie rozpowszechniać w Internecie utwory z tzw. domeny publicznej. Pojęcie to, pochodzące jeszcze z Anglii czasów nowożytnych, oznaczało pierwotnie ziemie należące bezpośrednio do państwa, dostępne wszystkim (ang. *common law*). W znaczeniu prawnym termin ten pojawił się także w angielskim Statucie Anny z 1710 roku. Według



niego twórcy tracili prawo do dysponowania własnym utworem po upływie 14 lat, po czym stawał się on własnością publiczną. Domena publiczna (ang. *public domain*) to w sensie prawnym zbiór tych utworów, do których wygasły już autorskie prawa majątkowe lub które nigdy nie były nimi objęte. Dzięki jej istnieniu kultura nieustannie rozwija się i ewoluuje. Jak czytamy w *Manifeście Domeny Publicznej*, wydanym przez sieć Communia:

domena publiczna leży u podstaw naszej świadomości, wyrażonej przez wspólnotę wiedzy i kultury. Jest surowcem, dzięki któremu tworzymy nową wiedzę i dzieła kultury (...) jest niezbędna dla zapewnienia społecznego i ekonomicznego dobrobytu naszych społeczeństw (*The Public Domain Manifesto*, 2010).

Historyk prowadząc badania korzysta aktywnie z zasobów domeny publicznej. Nawet jeśli wykorzystuje źródła, które wciąż są objęte ochroną prawną autorską, np. nagrania telewizyjne w badaniach historii najnowszej, podstawy jego pracy w jakimś stopniu oparte są na ideach, dziełach czy faktach naukowych wypracowanych dużo wcześniej, które – tak jak język czy przestrzeń publiczna – są własnością wszystkich i przez nikogo nie powinny być zawłaszczane.

Jednak nawet pobieżna analiza zbiorów udostępnianych online przez najważniejsze polskie i zagraniczne instytucje GLAM pokazuje, że dzieje się inaczej (Bednarek-Michalska, Werla, 2011). Powszechnie na przykład jest oznaczanie znakami wodnymi dokumentów, książek czy obrazów pochodzących sprzed wielu stuleci. W wielu muzeach zakazane jest fotografowanie wystawianym tam obiektom, ogranicza się też możliwość wykorzystania udostępnianych w cyfrowych repozytoriach zbiorów z domeny publicznej. Niektóre instytucje GLAM czy komercyjne podmioty, takie jak Google Books, skanują wiele materiałów z domeny publicznej, po czym nie udostępniają ich swobodnie online, powołując się nieprawomocnie na wciąż aktywne autorskie prawa majątkowe do tych zasobów.

W materiałach przygotowanych w ramach projektu badawczego *Reclaiming New Zealand's digitised heritage* możemy znaleźć informacje dotyczące tego, w jaki sposób Google czy instytucje zawłaszczają domenę publiczną, nakładając dodatkowe ograniczenia na wykorzystywanie dostępnych w niej zasobów. Takimi metodami blokowania wolności korzystania z domeny publicznej są dodatkowe licencje, opłaty za dostęp do pliku, nakładanie ograniczeń

komercyjnego wykorzystania utworów, umożliwianie dostępu jedynie w sieci akademickiej czy różnicowanie możliwości pełnego korzystania z zasobów w zależności od kraju (Clark, 2012).

Do domeny publicznej należą dzieła klasyków literatury i sztuki, dawne traktaty naukowe, ale także niezliczona liczba dokumentów, rękopisów, fotografii i innych materiałów o niezwykle zróżnicowanym przeznaczeniu i reprezentujących różny poziom intelektualny, literacki czy artystyczny. Jednym z podstawowych błędów w rozumieniu idei domeny publicznej jest właśnie utożsamianie jej zasobów z dziełami o wielkiej wartości artystycznej. Tymczasem prawo pozwala nam korzystać swobodnie i rozpowszechniać w Internecie nie tylko cyfrowe reprodukcje dzieł Kopernika czy Mickiewicza, ale także np. archiwalne amatorskie zdjęcia historyczne, dokumentujące historie lokalne, które interesować będą jedynie wybranych badaczy. Powinni oni mieć świadomość swojego prawa do wykorzystywania i rozpowszechniania tych zasobów także poza zakresem podstawowych prac badawczych (np. w publikacjach książkowych).

Lawrence Lessig zwraca uwagę, że efektywna domena publiczna jest naturalną dla kultury przestrzenią zasobów i idei pozostających poza zasięgiem wpływu prawników i korzystającą z efektywności ograniczeń prawa autorskiego (Lessig, 2006, s. 59). Jego zdaniem mamy dziś do czynienia z ograniczaniem tej przestrzeni: mimo że dzięki masowej digitalizacji zbiorów bibliotek, archiwów czy muzeów, coraz więcej treści dostępnych jest online, ich status prawnoautorski nie zawsze jest wystarczająco czytelny. Nie istnieją odpowiednie regulacje prawne w jasny i przejrzysty sposób pozwalające sprawdzić, czy dany zasób jest wciąż objęty autorskim prawem majątkowym, czy też można go w dowolny sposób wykorzystywać bez zgody kogokolwiek. Lessig podkreśla też, że w odróżnieniu od korzystania z archiwów i bibliotek na miejscu, obecność treści archiwalnych w Internecie generuje wiele wyzwań związanych z prawem autorskim. Podaje przykład usługi Google Books i porównuje ją z klasycznym katalogiem książek w bibliotece: odpowiednikiem bibliograficznej usługi w przestrzeni analogowej jest katalog kartkowy, który, podobnie jak Google Books, indeksuje dostępne książki; podobnie jego celem jest ułatwienie korzystania z tych prac. Trudno jednak sobie wyobrazić sytuację, w której wydawcy albo właściciele praw autorskich mogą mieć prawo do kontrolowania, w jaki sposób ten katalog jest przygotowywany. Dzieje się tak, ponieważ przestrzeń analogowa, zdaniem Lessiga, nie wywołuje tak wielu problemów z prawem autorskim (Lessig 2006, s. 69–70)

– trudniej tu choćby o błyskawiczne i niekontrolowane kopiowanie czy o masowe rozpowszechnianie treści, których status nie jest jasny.

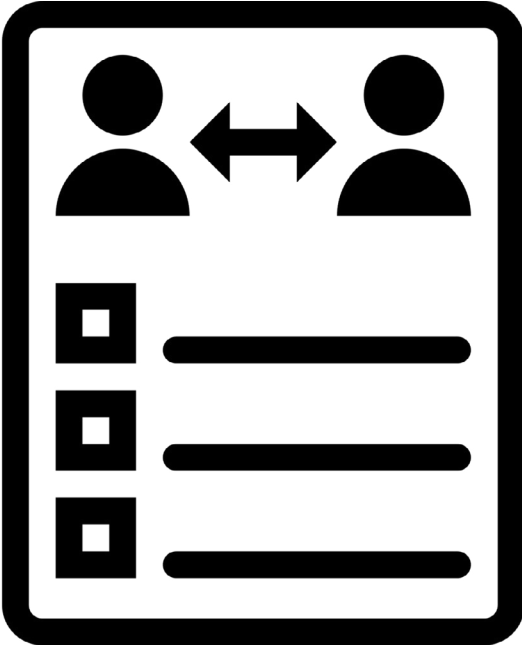
Postęp digitalizacji blokowany jest także przez tzw. utwory osierocone. To utwory, co do których nie można jednoznacznie określić, czy znajdują się już w domenie publicznej, czy nie, co wynikać może choćby z braku informacji na temat autorów i daty ich śmierci, kluczowej przy wyliczeniu okresu ochrony prawnoautorskiej. Nie wszystkie biblioteki czy archiwa skłonne są umieszczać reprodukcje cyfrowe tego typu zbiorów we własnych repozytoriach. Te, które się na to decydują, przerzucają konieczność zbadania statusu udostępnianego materiału na użytkownika – osobę lub instytucję.

Historyk pracujący z utworami w domenie publicznej powinien orientować się także w zasadach prawa do pierwszego wydania niechronionego i nigdy niepublikowanego utworu (może to być rękopis, list, szkic, fotografia itp.) oraz w zasadach prawa chroniącego wydania naukowe i krytyczne tych utworów – prawa dotyczącego w znacznym stopniu praktyki badawczej historyków i ich relacji z pracodawcą (uczelnia czy instytutem badawczym). Struktura niniejszego podręcznika nie pozwala jednak na rozwinięcie tego wątku. Wiedzę z zakresu prawa autorskiego i domeny publicznej w kontekście pracy historyków zdobyć można czytając odpowiednie podręczniki i opracowania prawnicze. Podstawowe informacje na temat prawa autorskiego, dozwolonego użytku edukacyjnego i naukowego czy domeny publicznej uzyskać można także w przewodniku opublikowanym w ramach prowadzonego przez Fundację Nowoczesna Polska portalu Prawokultury.pl.

## BIBLIOGRAFIA

- American Association of Museums, Center for the Future of Museums (2008). *Museums & society 2034: Trends and potential futures*, American Association of Museums, Washington. Pobrano z: <http://www.aam-us.org/docs/center-for-the-future-of-museums/museumssociety2034.pdf?sfvrsn=0>
- Jasiński, T. (2010). Biblioteka cyfrowa i informatyka w warsztacie mediewisty. W: Mazurek, C., Stroiński, M., Węglarz, J., (red.), *Polskie biblioteki cyfrowe 2010* (s. 1–15). Poznań: Ośrodek Wydawnictw Naukowych (IChB PAN). Pobrano z: <http://lib.psnc.pl/dlibra/docmetadata?id=293&from=publication>.

- Turkel, W.J., Kee, K., Roberts, S.** (2012). A method for navigating the infinite archive. W: T. Weller (red.), *History in the Digital Age* (s. 61-75). London, New York: Routledge.
- Bednarek-Michalska B., Werla M.** (2011). Status prawny obiektów cyfrowych w polskich bibliotekach cyfrowych — analiza techniczna stanu w roku 2011, *Biuletyn EBIB*, 7(125), 1-8. Pobrano z: [http://www.nowyebib.info/images/stories/numery/125/125\\_werla\\_michalska.pdf](http://www.nowyebib.info/images/stories/numery/125/125_werla_michalska.pdf)
- Clark, A.** (2012). *Privatising the Public Domain? Exploring access and usage regimes for public domain works in a digital environment* [draft]. Pobrano z: [http://www.kiwialex.com/uploads/1/1/8/5/11856774/\\_privatising\\_the\\_public\\_domain\\_-\\_mar\\_2013\\_draft.docx](http://www.kiwialex.com/uploads/1/1/8/5/11856774/_privatising_the_public_domain_-_mar_2013_draft.docx)
- Drakes, G.** (2012). Who Owns Your Archive? Historians and the Challenge of Intellectual Property Law. W: Potter, C.B. i Romano, R.C. (red.), *Doing recent history: On privacy, copyright, video games, institutional review boards, activist scholarship, and history that talks back* (s. 83–111). Athens: University of Georgia Press.
- Lessig, L. (2006). Re-crafting a public domain, *Yale Journal of Law & the Humanities*, 18, 56–83.
- The Public Domain Manifesto* (2010). Pobrano z: <http://www.publicdomainmanifesto.org/manifesto>.
- Werla, M.** (2012). Web 2.0 i (polskie) biblioteki cyfrowe, *Biuletyn EBIB*, 2(129). Pobrano z: <http://ebib.info/biuletyn/plan-numerow/986-ebib-22012-1294>.



# Historiografia poza drukiem

Nauka historyczna począwszy od XIX wieku posługuje się drukiem. Co więcej – jest z tym medium zintegrowana. Tradycyjną narrację pisaną uznaje się za najbardziej naukową formę pisarstwa historycznego, ponieważ, jak się powszechnie uważa, pozwala ona na najbardziej bezpośrednie przekazanie prawdy o przeszłości.

Teoretycy historiografii przynajmniej od lat 60. XX wieku podważają ten wyjątkowy status druku i linearnej narracji w historiografii, starając się przekonać co do potencjału innych jej form, dostosowanych do nowej wizualnej i cyfrowej kultury. Wyrazem tego są m.in. tezy Roberta Rosenstone'a czy Haydena White'a o filmie jako równorzędnym wobec tekstów pisanych nośniku wiedzy historycznej. Krytyka druku wynika u nich z odpowiednio krytycznego podejścia epistemologicznego i językowego: Rosenstone i White pokazują, że film jako medium historiograficzne ma tej samej natury wady i ograniczenia, co tradycyjnie pisana monografia, a jednym z argumentów jest to, że w podobny sposób nie pozwalają one na pełną rekonstrukcję przeszłej rzeczywistości (Rosenstone, 2008; White, 2008). Problemy filmu jako nośnika naukowej wiedzy historycznej od lat zajmują polskich teoretyków historiografii. W polskich bibliotekach dostępna jest też bogata literatura na ten temat (oryginalne prace w języku polskim oraz tłumaczenia).

Kiedy przedmiotem badań historycznych stają się społeczeństwa i jednostki funkcjonujące w kulturze cyfrowej, trzeba zastanowić się nad tym, czy druk wciąż pozostaje optymalną formą historiografii i czy – jako jednowymiarowe,

statyczne medium – pozwala na efektywne opisanie rzeczywistości. Obok klasycznej krytyki teoretycznej w dyskusji tej pojawiają się wątki bardzo praktyczne: nie da się w treści książki zamieścić filmu z YouTube ani oddać w piśmie dynamiki dyskusji na Twitterze, co w przypadku tematów takich jak np. „arabska wiosna” staje się akurat bardzo istotne.

Pod koniec lat 90. Robert Darnton, ówczesny przewodniczący Amerykańskiego Towarzystwa Historycznego, zaproponował model wielopoziomowej, analogowo-cyfrowej monografii historycznej. W relacji Edwarda L. Ayersa skonstruowany miał być on następująco:

Najwyższa warstwa powinna być zwięzłym podsumowaniem tematu, być może dostępnym także w druku. Kolejna warstwa może zawierać rozbudowane omówienia różnorodnych aspektów wykładu, ułożone nieliniarnie w ramach narracji, ale dostępne jako samodzielne jednostki uzupełniające główną narrację. Trzecia warstwa mogłaby zawierać dokumentację źródłową... Czwarta warstwa mogłaby mieć charakter historiograficzny... Piąta warstwa mogłaby pełnić rolę dydaktyczną... Szósta warstwa zawierać mogłaby komentarze od czytelników, korespondencję między autorem i wydawcą oraz listy od czytelników... (Ayers, 1999).

W modelu Darntona mamy do czynienia z monografią spełniającą nie tylko swoją podstawową funkcję historiograficzną, ale także będącą rodzajem repozytorium źródeł, forum dyskusyjnym i broszurą edukacyjną. Pełna treść monografii dostępna jest w formie cyfrowej, niektóre jej elementy mogą być także publikowane w druku.

W poszukiwaniu nowych form monografii historycy zwrócili się w stronę eksperymentów z hipertekstem. Pojęcie to, tak dobrze znane wszystkim użytkownikom Internetu, w teorii literatury ma znaczenie wychodzące poza kontekst sieci WWW – przykładowo według niektórych badaczy *Ulisses* Jamesa Joyce'a jest hipertekstem *Odysei* Homera (zob. Vandendorpe, 2008, s. 101). O hipertekstowej historiografii pisać więc można także bez odniesień do systemów informatycznych – robi tak np. Andrzej Radomski, opisując wydaną w latach 70. książkę A. Kerstena *Na tropach Napierskiego* (Radomski, 2002). Można też szukać hipertekstowych struktur w filmie i utworach multimedialnych, traktowanych jako alternatywna forma pisanej narracji historycznej – recepcja i badanie takich dzieł niewiele ma wspólnego z tradycyjnym linearnym odczytywaniem następujących po sobie fragmentów tekstu.

Pojęcie hipertekstu zaproponował w połowie lat 60. amerykański socjolog i filozof Ted Nelson, który wykorzystał je do zaprojektowania i udostępnienia w 1974 roku programu komputerowego Xanadu, pozwalającego na tworzenie połączeń między poszczególnymi dokumentami poza strukturą uporządkowanej listy. taki model uporządkowania treści sprawiał, że jej linearne odczytywanie stawało się niemożliwe i wymagało od czytelnika nawigowania przez wybrane połączenia. Na bazie teorii Nelsona, na początku lat 90., brytyjski naukowiec Tim Berners-Lee wypracował system World Wide Web (WWW), który dziś jest najpopularniejszą usługą internetową, często w ogóle kojarzoną błędnie z całym Internetem.

Gdyby uprzytomnić sobie, w jaki sposób korzystamy ze stron WWW, podstawowe cechy hipertekstu staną się dla nas czytelne. Mamy więc do



Interaktywna wizualizacja komunikacji listowej Woltera w okresie od 1750 do 1770 roku, przygotowana w ramach projektu Mapping the Republic of Letters realizowanego na Uniwersytecie Stanforda (<http://republicofletters.stanford.edu/>).



czynienia z różnymi dokumentami, pojedynczymi stronami o indywidualnych adresach URL, połączonymi ze sobą hiperłączami (linkami). O ile można stworzyć indeksy takich linków, jak czyni to choćby Google, trudno uszeregować je w ramach jednej linearnej kolejności – zupełnie inaczej niż w tradycyjnej książce, której struktura zazwyczaj nakłada na czytelnika obowiązek podążania w jednym kierunku, strona za stroną, zgodnie z zamysłem autora. Hipertekst daje użytkownikowi więcej wolności w wyborze własnej drogi i eksploracji tekstu, co oczywiście nie oznacza, że wolności tej nie ma w ogóle w przypadku klasycznego czytania drukowanego egzemplarza.

Hipertekst pozwala na eksplorację treści na wielu poziomach. O ile w przypadku druku mamy do czynienia wyłącznie z jedną płaszczyzną prezentacji, czyli zadrukowaną powierzchnią kartki, o tyle w systemach hipertekstowych czytelnik i/lub użytkownik może poruszać się w głąb eksplorowanej treści. Podobnie jak w propozycji Darntona, pod poziomem klasycznej narracji naukowej zamieszczone mogą być cyfrowe reprodukcje źródeł wykorzystanych do jej sformułowania, wizualizacje, grafiki itp. Ciekawą propozycję hipertekstowego systemu historiograficznego przedstawiają Sara Roegiers i Frederik Truyen. Według nich hipertekst jest formą zdecydowanie lepiej przystosowaną do uprawiania historiografii, z uwagi na złożoność materii historycznej, niż tradycyjny jednowymiarowy tekst pisany, i lepiej niż ten drugi pozwala na zaprezentowanie aparatu krytycznego w treści pracy historycznej. Autorzy ci proponują trójwymiarowy model struktury reprezentacji przeszłej rzeczywistości, składający się z takich elementów jak: czas, przestrzeń i kontekst społeczny (ang. *community*). Hipertekstowa monografia historyczna ma umożliwiać takie przedstawienie opisu przeszłości, którego elementy można dynamicznie przeglądać w perspektywie zdefiniowanego przez autora okresu czasu, zasięgu geograficznego i społecznego (Roegiers, Truyen, 2008, s. 71 i nast.) Spróbujmy wyobrazić sobie wdrożenie takiego modelu w postaci przykładowej monografii poświęconej powojennej odbudowie historycznej zabudowy centrum Gdańska:

Po uruchomieniu programu lub wejściu na stronę WWW czytelnik miałby do dyspozycji nie tylko klasyczny tekst z pełnym, naukowym przedstawieniem zagadnienia, ale także rozbudowane indeksy poszczególnych budynków, ulic czy nazwisk pojawiających się w publikacji. Wybrane słowa w narracji byłyby aktywnymi linkami, umożliwiającymi rezygnację z linearnej lektury i przejście do kolejnych wątków lepiej pasujących do

potrzeb czytelnika. Dostępna w ramach publikacji baza danych z informacjami o budynkach prezentowałaby je w dynamicznie zmieniającym się kontekście, pokazując np. postępy odbudowy w poszczególnych latach czy zmiany w najbliższym otoczeniu. Każdy cytat z literatury naukowej byłby jednocześnie linkiem do podstrony prezentującej cytowaną publikację wraz z odpowiednim komentarzem autora.

Warto zwrócić uwagę na fakt, że tego typu prezentacja zachowywałaby naukowy charakter, ponieważ wciąż dostępny byłby cały aparat krytyczny (przypisy, bibliografia). Zmieniłby się jednak model czytania – z jednowymiarowej i jednokierunkowej narracji na wielowymiarowy, hipertekstowy system powiązanych ze sobą elementów. Jak pisze Piotr Witek:

Historyk mający do dyspozycji nowe audiowizualne technologie nie musi, choć może, pisać historii. Badacz przeszłości może nie być pisarzem, może też nie naśladować konwencji literackich. Historyk nie musi myśleć o przeszłości w kategoriach dobrze skonstruowanej opowieści. Refleksja historyczna nie musi opierać się na linearnych schematach przyczynowo-skutkowych. Oznacza to, że w dobie symulacji elektronicznej badacz przeszłości nie musi dokonywać wiązania faktów i zdarzeń w spójne opowieści z początkiem, rozwinięciem i zakończeniem (Witek, 2005, s. 411).

Oczywiście hipertekst i hipertekstowe formy mają swoje wady: czy rzeczywiście czytelnik potrzebuje aż takiej swobody w recepcji przekazu naukowego? Czy tak wielkie możliwości wyboru własnej ścieżki lektury nie prowadzą do chaosu, w którym trudno zorientować się, jakie elementy treści są najważniejsze, a jakie poboczne? Z punktu widzenia autora historyka problemem są także kompetencje informatyczne, pozwalające na przygotowanie tak złożonej prezentacji oraz umiejętność odpowiedniego zaplanowania badań. Nie bez znaczenia pozostaje także problem kosztu produkcji i użyteczności takich multimedialnych, historycznych projektów: czy można z nich korzystać bez odpowiedniego oprogramowania? Czy – tak jak w przypadku książki – czytelne są zasady korzystania z jej treści, dotyczące kopiowania i kwestii praw autorskich do poszczególnych elementów?

Hipertekstowe, multimedialne historyczne publikacje naukowe pozostaną pewnie jedynie niszową, awangardową formą historiografii. Z drugiej strony,

potencjał hipertekstu pozwala myśleć w nowy sposób o naukowej monografii historycznej.

Wspomniane wyżej propozycje odejścia od tradycyjnej pisanej narracji historycznej mają charakter awangardowy, a multimedialna forma publikacji nie musi być wcale bardziej użyteczna i wygodniejsza w odbiorze od standardowej, drukowanej książki czy artykułu. Dodajmy do tego konieczne kompetencje techniczne i koszty przygotowania tego typu prezentacji. Z drugiej strony sama idea wyjścia poza tekst i szukania nowych form naukowego przekazu historycznego, zwłaszcza w odniesieniu do wadliwej pisanej narracji, jest bardzo atrakcyjna. Z tego też powodu warto zainteresować się wizualizacjami i infografikami, czyli formą prezentacji informacji i danych historycznych pośrednią między tekstem a obrazem.

Wizualizacje i infografiki wykorzystywane są w publikacjach historycznych od dawna jako uzupełnienie (ilustracja) tekstu pisanego. Są to rozmaite tabele, wykazy, wykresy, zestawienia i schematy. W odróżnieniu od skomplikowanych multimedialnych prezentacji, ich przygotowanie nie jest trudne, a forma często intuicyjna, np. oś czasu lub wykres liniowy. Niektórzy teoretycy historiografii zwracają jednak uwagę na możliwość budowania wizualizacji będących czymś więcej niż jedynie dodatkiem do głównego tekstu, prezentujących w skrótovej formie zawarte w nim najważniejsze informacje.

David J. Staley w swojej książce *Computers, Visualization and History* wskazuje na dominację myślenia tekstem nad myśleniem obrazem w szkolnej i akademickiej edukacji historycznej. Studentów uczy się głównie analizy tekstów pisanych, ale nie wizualnych, które dziś kulturowo dominują i – co pokazuje telewizja czy Internet – są równorzędną formą komunikacji, posiadającą własną gramatykę i rozpoznawalne cechy gatunkowe (memy, teledyski hiphopowe itp). Teksty wizualne kojarzone są wciąż z gorszą, trywialną formą prezentacji informacji i wiedzy, z telewizją albo z grami komputerowymi, którym odmawia się jakiegokolwiek edukacyjnego znaczenia. Zakłada się także błędnie, że analiza tekstów wizualnych jest raczej formą subiektywnego doświadczenia, porównywalnego z tym, wywoływanym kontaktem z dziełami sztuki, i z tego powodu nie ma większego poznawczego i obiektywnego potencjału (Staley, 2003, s. 39).

Dla Staleya wizualizacja czy infografika to coś więcej niż zwykła ilustracja tekstu: to sposób organizacji znaczącej (ang. *meaningful*) informacji w formie przestrzennej, przeznaczony do dalszej systematycznej analizy (Staley, 2003, s. 42). Wizualizacja nie pełni więc roli dekoracji czy prostej ilustracji

uatrakcyjniającej tekst, a jest osobnym przekazem, który oczywiście może – a nawet powinien – posiadać odpowiednie cechy estetyczne. Te ostatnie nie są jednak najważniejsze. Jak zresztą podkreśla Staley, tak jak krytykować można styl tekstu pisanego, tak krytykować można estetykę wizualizacji. W obu jednak przypadkach mamy do czynienia z pewnymi elementami badania, które w swojej istocie koncentrować powinno się głównie na warstwie merytorycznej przekazu.

Jeśli przypomnimy sobie oglądane w zwykłych naukowych książkach historycznych elementy graficzne włączone w tekst, łatwo zauważymy, że wizualizacje mogą mieć dwojaką naturę. Niektóre mają charakter reprezentacji, tak jak fragmenty map, reprodukcje budynków, szkice posągów itp. Wizualizacja korzysta tutaj z informacji, które i tak oryginalnie dostępne są jedynie w wizualnej formie. Innym – i według Staleya – bardziej interesującym rodzajem wizualizacji są te oparte o obiekty abstrakcyjne, czyli rozmaite diagramy, osie czasu, schematy, w których oryginalnie niewizualna informacja uzyskuje wizualną formę. Prosta oś czasu jest tutaj doskonałym przykładem: chronologiczny zestaw dat zostaje uporządkowany wzdłuż jednej linii, co pozwala pokazać niewidoczne na pierwszy rzut oka relacje między poszczególnymi datami.

Jak zacząć pracę nad wizualizacją? Staley proponuje przyjąć kilka podstawowych zasad, bazując na tradycyjnych zaleceniach dla autorów tekstów pisanych:

- Unikaj bałaganu: zadbaj o to, aby wizualizacja była odpowiednio prosta i czytelna, ale na tyle złożona, aby przekazać twoją ideę;
- Wizualna prezentacja informacji musi bezpośrednio zakomunikować ideę samą w sobie, inaczej staje się zwykłą ilustracją;
- Weź pod uwagę odbiorców swojej wizualizacji;
- Nigdy nie wykorzystuj wizualizacji tylko jako formy dekoracji czy wypełniacz miejsca;
- Pamiętaj o tym, że wizualizacja zniekształca dane, co jest nieuniknione (tak samo zresztą jak w przypadku tekstu pisanego);
- Używaj koloru wyłącznie, kiedy jest to niezbędne; nie wykorzystuj kolorów, które mogą rozpraszać uwagę lub być słabo rozróżnialne.

(na podstawie: Staley, 2003, s. 41)

Oczywiście analiza wizualizacji wymaga innych narzędzi poznawczych niż w przypadku tekstu pisanego, a przede wszystkim zrozumienia fundamentalnych różnic między pisaną a wizualną historiografią: po pierwsze, tekst pisany porządkuje logika, wizualizację – analogia; po drugie, badanie tekstu opiera się na krytycznej analizie jego elementów, podczas gdy badanie wizualizacji musi być z konieczności syntezą całości; po trzecie, poszczególne elementy narracji łączą się w liniowe ciągi (zdanie po zdaniu), a w przypadku wizualizacji mamy do czynienia raczej z siecią (ang. *network*); po czwarte, w przypadku narracji pisanej liczy się przede wszystkim fakt, a w wizualizacjach najważniejsza jest struktura (Staley, 2003, s. 55). Świadomość tych fundamentalnych różnic jest konieczna do krytycznej analizy wizualizacji.

Narzędzia cyfrowe pozwalają w prosty sposób tworzyć wizualizacje historyczne. Nie mniej istotna niż kompetencje techniczne jest wyobraźnia, która pozwala myśleć o niewizualnych danych w formie wizualnej prezentacji.

## BIBLIOGRAFIA

- Ayers, Edard L.** (1999). *The Pasts and Futures of Digital History*. Pobrano z: <http://www.vcdh.virginia.edu/PastsFutures.html>.
- Radomski, A.** (2002). Historiografia jako hipertekst, *Kultura i Historia*, 3/2002. Pobrano z: <http://www.kulturaihistoria.umcs.lublin.pl/wp-content/uploads/2008/01/Kultura-i-Historia-nr-3.pdf>.
- Roegiers S., Truyen F.** (2008). History is 3D: Presenting a framework for meaningful historical representations in digital media. W: Kalay, Y. E., Kvan, T., Affleck J., (red.), *New Heritage, New media and cultural heritage* (s. 67-77). London, New York: Routledge.
- Rosenstone, R.** (2008). Historia w obrazach/historia w słowach: rozważania nad możliwością przedstawienia historii na taśmie filmowej. W: Kurz, I. (red.), *Film i historia. Antologia* (s. 93-116). Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.
- Staley, D. J.** (2003). *Computers, visualization, and history: how new technology will transform our understanding of the past*, Armonk, New York: M.E. Sharpe.
- White, H.** (2008). Historiografia i historiofotia. W: Kurz, I. (red.) *Film i historia. Antologia* (s. 117-130). Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.

- Witek, P.** (2005). Multimedia jako jeden z wariantów kulturowej gry w historię. Metodologiczne problemy przedstawiania przeszłości w epoce ekranów. W: Dytman-Stasienko, A., Stasienko, J. (red.), *Język a multimedia* (s. 367-390). Wrocław: Wydawnictwo Naukowe Dolnośląskiej Szkoły Wyższej Edukacji TWP.
- Vandendorpe, Ch.** (2008). *Od papirusu do hipertekstu. Esej o przemianach tekstu i lektury*, Warszawa: Wydawnictwo Uniwersytetu Warszawskiego.



# Otwarty dostęp do publikacji naukowych

Internet jest dziś niezwykle efektywnym kanałem dystrybucji publikacji naukowych, potrzeba jednak odpowiedniej wiedzy i kompetencji, aby w pełni wykorzystać jego możliwości w tym zakresie. Na najbardziej podstawowym poziomie kompetencje te związane są z umiejętnością sprawnego wyszukiwania informacji o publikacjach z interesującej nas dziedziny i tematu.

Przygotowanie zestawu lektur do opracowania przed napisaniem własnego tekstu to jedno z podstawowych zadań autora. Zdobycie informacji o niezbędnych tytułach możliwe jest przede wszystkim dzięki wyszukiwarkom bibliotecznym działającym w modelu OPAC (ang. Online Public Access Catalog). To często rozbudowane o wiele użytecznych narzędzi wyszukiwania katalogi informacji bibliograficznych. Obok katalogów bibliotek uczelnianych skorzystać można z katalogów o charakterze centralnym, gromadzących największą liczbę informacji o dostępnej literaturze przedmiotu. W Polsce są to katalogi Biblioteki Narodowej oraz Narodowy Uniwersalny Katalog Centralny NUKAT. Przeszukiwanie zasobów bibliograficznych Biblioteki Narodowej jest o tyle zasadne, że jako jedna z kilkunastu bibliotek (wyszczególnionych pod adresem [http://pl.wikipedia.org/wiki/Egzemplarz\\_obowi%C4%85zkowy](http://pl.wikipedia.org/wiki/Egzemplarz_obowi%C4%85zkowy)) ma ona prawa do tzw. egzemplarza obowiązkowego, czyli egzemplarza każdej publikacji wydawanej w Polsce. Biblioteka Narodowa rozwija także zbiór baz bibliograficznych składających się na Polską Bibliografię Narodową oraz bibliografie specjalne: Przewodnik Bibliograficzny, Czasopisma Polskie, Bibliografia Zawartości Czasopism, Polonica Zagraniczne. Wspomniani wcześniej NUKAT



pozwala tymczasem na łączne przeszukiwanie polskich bibliotek naukowych i akademickich, podobnie jak projekt KARO (Katalog rozproszony Bibliotek Polskich).

Obok katalogów bibliotecznych w poszukiwaniu literatury do swojej pracy badawczej skorzystać można z bibliograficznych baz danych. W badaniach humanistycznych i społecznych źródłem informacji o artykułach z polskich czasopism może być baza BazHum, rozwijana przez Muzeum Historii Polski, indeksująca obecnie 338 czasopism i mająca wkrótce udostępniać pełne wersje publikowanych w nich artykułów. Dostępna jest także opracowywana przez Uniwersytet Śląski w Katowicach baza ARTON, rejestrująca informacje bibliograficzne i cytowania zawarte w polskiej literaturze humanistycznej. Dane z kolejnych roczników Bibliografii Historii Polskiej wydawanych przez Instytut Historii PAN przeszukiwać można online w ramach wyszukiwarki przygotowanej przez Instytut Pamięci Narodowej. Poniżej krótkie zestawienie wybranych polskich baz:

Nazwa bazy	URL	Zawartość
BazHum	<a href="http://bazhum.pl/">http://bazhum.pl/</a>	BazHum to baza bibliograficzna polskich naukowych czasopism humanistycznych i społecznych. Baza gromadzi metadane bibliograficzne: rejestruje całą zawartość czasopism — od pierwszych numerów do bieżących. Twórcą kolekcji jest Muzeum Historii Polski w Warszawie.
Bibliografia Historii Polskiej	<a href="http://www.bibliografia.ipn.gov.pl/">http://www.bibliografia.ipn.gov.pl/</a>	Wersja elektroniczna Bibliografii Historii Polskiej, opracowana we współpracy z Instytutem Pamięci Narodowej, zawiera te same dane, co kolejne roczniki wydawane w formie książkowej przez Instytut Historii PAN. Baza obejmuje publikacje wydawane w latach 1980-2007.
ARTON: Polska literatura humanistyczna	<a href="http://www1.bg.us.edu.pl/arton_inf/arton.htm">http://www1.bg.us.edu.pl/arton_inf/arton.htm</a>	Informacje bibliograficzne (ponad 12 tys. artykułów) i cytowania (ponad 183 tys.) zawarte w polskiej literaturze humanistycznej.

Przewodnik Bibliograficzny	<a href="http://mak.bn.org.pl/cgi-bin/makwww.exe?BM=23">http://mak.bn.org.pl/cgi-bin/makwww.exe?BM=23</a>	Przewodnik Bibliograficzny rejestruje publikacje, których wydawca lub współwydawca ma siedzibę na terenie Polski oraz wydawnictwa ogłoszone poza terenem państwa, jeżeli ich wydawcą lub współwydawcą są instytucje krajowe lub ich agendy zagraniczne.
Bibliografia Zawartości Czasopism	<a href="http://mak.bn.org.pl/w14.htm">http://mak.bn.org.pl/w14.htm</a>	Jej zadaniem jest rejestrowanie artykułów opublikowanych w polskich czasopismach, bez względu na język publikacji, wpływających jako egzemplarz obowiązkowy do Biblioteki Narodowej.
Bibliografia Wydawnictw Ciągłych	<a href="http://mak.bn.org.pl/w12.htm">http://mak.bn.org.pl/w12.htm</a>	W bibliografii wydawnictw ciągłych uwzględnia się wydawnictwa ciągłe publikowane w Polsce, niezależnie od języka tekstu, o częstotliwości od dziennika do rocznika, czasopisma publikowane nieregularnie, wydawnictwa zbiorowe, kalendarze, sprawozdania.

Poważnym ograniczeniem wymienionych wyżej katalogów jest stosunkowo niewielka dostępność informacji o publikacjach wydawanych za granicą. W tym przypadku proste, najbardziej intuicyjne skorzystanie z wyszukiwarki Google to zdecydowanie za mało. Google oprócz globalnej wyszukiwarki przeszukującej zasoby World Wide Web udostępnia także dwie usługi kluczowe dla każdego, kto poszukuje informacji o wydawanych na całym świecie publikacjach naukowych. Pierwsza z nich to Google Scholar, będąca połączeniem dynamicznie rozwijanego katalogu tekstów naukowych z systemem ich oceniania i prezentowania. Druga usługa to Google Books (Google Książki). W Google Books znajdziemy informacje, a bardzo często również skany książek naukowych i czasopism. Wiele z nich ze względu na ograniczenia prawa autorskiego dostępnych jest jedynie w ograniczonej postaci – możemy przeglądać jedynie fragmenty zeskanowanych treści. System pozwala jednak na przeszukiwanie całości (wszystkich słów zawartych w tekście książki), co ma fundamentalne znaczenie. Dzięki temu jesteśmy w stanie sprawdzić, czy interesujące nas pojęcie, osoba lub fakt zostało w jakikolwiek sposób opisane w przeglądanej publikacji, a więc czy warto jest się nią zainteresować i do niej dotrzeć.



# bibliograficzna baza czasopism humanistycznych i społecznych

bazhum.pl

 260  
czasopism

 300 tysięcy  
artykułów



otwarte dane - możliwość  
ponownego wykorzystania  
materiałów

GNU

oprogramowanie dostępne  
na otwartej licencji



opracowanie całego  
zakresu czasopisma:  
od pierwszych  
do bieżących numerów



metadane dla każdego  
artykułu

## Cechy:



metadane wysokiej jakości  
tworzone zgodnie ze  
standardami opracowania  
bibliograficznego



pełne zakresy czasopism:  
od pierwszych numerów  
do najnowszych (baza  
aktualizowana na bieżąco)



najważniejsze polskie  
czasopisma z dziedzin  
 nauk humanistycznych  
i społecznych



stałe rozbudowywana  
o nowe tytuły



bez konieczności  
logowania i bezpłatna

## Funkcjonalność:



dostęp do informacji  
- szybki, łatwy



wygodne przeglądanie  
i wyszukiwanie



eksport metadanych do  
wielu formatów



czytelny na telefonie  
i w tablecie

## Zastosowania:



tworzenie bibliografii  
do badań i projektów



poszukiwanie informacji  
o ludziach, miejscach  
i wydarzeniach



tworzenie przypisów



kwerenda naukowa



uzupełnianie katalogów  
bibliotecznych



Plakat dostępny jest na licencji Creative  
Commons Uznanie Autorstwa 3.0 Polska

W usłudze Google Books dostęp do informacji o książce połączony jest często z dostępem do jej treści. W kwietniu 2013 roku w bazie tej usługi znalazło się ponad 30 mln tytułów. Co więcej, w listopadzie amerykański sąd oddalił pozew wydawców, zagrażający dalszemu rozwojowi tego projektu, możemy spodziewać się więc, że liczba indeksowanych publikacji będzie przyrastać.

Korzystając z Google Scholar i Google Books należy mieć jednak na uwadze, że nie gromadzą one informacji o bezwzględnie wszystkich publikacjach i szukając książek, szczególnie w mniej popularnych językach, należy wybrać także inne katalogi i narzędzia.

Użytecznym narzędziem może być serwis i wyszukiwarka WorldCat. Zawiera on informacje o 2 miliardach obiektów (książek – także wydawnictw w formatach czytnikowych, artykułów naukowych, wydawnictw DVD i CD). Rozwijany przez organizację Online Computer Library Center projekt wykorzystać można do budowania bibliografii publikacji zagranicznych, które zamierzamy wykorzystać przy pisaniu własnej pracy. Skorzystać warto także z wyszukiwarki założonego w 2004 roku czasopisma Central European Journal Of Social Sciences And Humanities (CEJSH), publikującego anglojęzyczne abstrakty artykułów i rozpraw z czasopism z zakresu nauk społecznych i humanistycznych, wydawanych w Europie Środkowej.

Wymienione wyżej projekty i usługi zapewniają dostęp do informacji o publikacjach naukowych, nie gwarantują jednak możliwości ich przeczytania. Jest to duży problem, szczególnie wobec wysokich cen książek zagranicznych i ograniczonego dostępu do nich w polskich bibliotekach. Ponieważ jest to problem globalny, powstają nieformalne jego rozwiązania również o globalnym zasięgu – dziś funkcjonują przynajmniej dwie inicjatywy internetowe, w ramach których publikowane są odnośniki do umieszczanych w serwisach hostujących pliki cyfrowych wersji książek naukowych wydawanych w ostatnich latach. Są wśród nich także naukowe publikacje historyczne. Jeszcze do niedawna jednym z najbardziej popularnych projektów tego typu był Library.nu, dostępny wcześniej pod domenami ebooksclub.org i gigapedia.com. Tuż przed zamknięciem udostępniał on odnośniki do 400 tys. tytułów, umieszczanych na zewnętrznych serwerach zazwyczaj ze złamaniem autorskich praw majątkowych wydawców. W lutym 2012 roku, po sprawie wytoczonej przez grupę 17 z nich, serwis Library.nu został zamknięty decyzją sądu w Monachium.

Na przełomie 2010 i 2011 roku Aaron Swartz, aktywista internetowy i programista podłączył się do sieci komputerowej w bibliotece Massachusetts Institute of Technology. Uzyskał nieautoryzowany dostęp do bogatego repozytorium czasopism naukowych JSTOR i pobrał z niej automatycznie prawie 5 milionów artykułów naukowych, które następnie nielegalnie republikaował w sieciach P2P. Odtąd korzystać mógł z nich każdy. Kilka lat wcześniej sam opublikował „Manifest Partyzanckiego Open Access” (*Guerilla Open Access*

*Manifesto*), w którym wskazywał na problem słabości ruchu na rzecz otwartego dostępu, niesprawiedliwą politykę wydawców czy monopol Google w digitalizacji tego typu zasobów. Wzywał do oddolnego i masowego republikowania w ogólnodostępnym Internecie artykułów pobieranych z płatnych i zamkniętych repozytoriów czasopism naukowych:

Do zapłacenia jest zbyt wysoka cena. Zmuszanie akademików do płacenia za czytanie prac ich kolegów? Skanowanie całych bibliotek, ale pozwalanie jedynie pracownikom Google na czytanie tych zdigitalizowanych materiałów? Dostarczanie artykułów naukowych elicie uniwersytetów w Pierwszym Świecie, ale nie odbiorcom z Globalnego Południa? To jest oburzające i nie do przyjęcia. [...] Powinniśmy pozyskiwać zasoby, gdziekolwiek są przechowywane, robić własne kopie i dzielić się nimi ze światem. Powinniśmy odnajdować materiały, do których prawa autorskie wygasły i i dodawać je do archiwum. Powinniśmy kupować dostęp do zamkniętych baz i udostępniać je w Sieci. Powinniśmy ściągać artykuły naukowe i umieszczać je w sieciach wymiany plików... (Swartz, 2008)

W styczniu 2011, tuż po akcji na MIT, Swartz został aresztowany. W lipcu postawiono mu zarzuty włamania do systemu komputerowego z zamiarem popełnienia przestępstwa. Po wpłaceniu kaucji został zwolniony, ale sprawa karna toczyła się nadal przed sądem federalnym. Pod koniec 2012 roku prokuratura ujawniła, że grozi mu kara do 50 lat więzienia oraz 1 mln dolarów grzywny. Sprawę zamknięto wkrótce po tym, jak 11 stycznia 2013 roku Swartz popełnił samobójstwo w swoim mieszkaniu na Brooklynie.

Działania Swartza czy popularność Library.nu i następców tego serwisu wskazują na głębokie systemowe problemy związane z dostępem do literatury naukowej. Rosnące wciąż koszty subskrypcji czasopism naukowych sprawiają, że – szczególnie w czasach kryzysu ekonomicznego – biblioteki często zmuszone są rezygnować z zakupu dostępu do części tytułów. Wiele uczelni i badaczy z krajów rozwijających się nie może pozwolić sobie na kupno drogich zagranicznych książek, niezbędnych w ich pracy naukowej. Co więcej, zyski z komercyjnej dystrybucji trafiają głównie do pośredników, a nie do autorów wydawanych prac. Chociaż oferta darmowego dostępu online do książek i czasopism naukowych jest coraz bogatsza, wciąż jednak – przynajmniej w przypadku prac historycznych – nie jest powszechnym standardem.

Warto wiedzieć, że bezpłatny dostęp do najważniejszych globalnych baz czasopism naukowych gwarantowany jest polskim uczelniom i innym instytucjom naukowym w ramach programem Wirtualna Biblioteka Nauki, finansowanego przez Ministerstwo Nauki i Szkolnictwa Wyższego, a prowadzonego przez Centrum Otwartej Nauki Interdyscyplinarnego Centrum Modelowania Matematycznego i Komputerowego Uniwersytetu Warszawskiego (ICM UW). Objęte programem są takie bazy jak SCOPUS, Science, EBSCO, Springer (ponad 1800 tytułów, roczniki od 1996), Web of Knowledge, Nature, Wiley-Blackwell i Elsevier (ponad 1700 tytułów, roczniki od 1995). Alternatywną i nieformalną formą dostępu (także do rosnącej liczby polskich artykułów naukowych z zamkniętych czasopism) są serwisy takie jak Academia.edu czy Research Gate, w ramach których autorzy na własną rękę republikują swoje prace.

Barierą dla otwartego publikowania online są przepisy prawa autorskiego i nieczytelny status prawnoautorski artykułów: bardzo często ich autorzy nie podpisują z wydawcą czasopisma żadnych umów; bywa też tak, że nawet jeśli istnieje jakaś umowa, to nie obejmuje ona możliwości rozpowszechniania tekstu w Internecie. Nie zawsze też wydawca ma możliwości organizacyjne i techniczne pozwalające na uruchomienie internetowej wersji czasopisma – z drugiej strony bez problemu znaleźć można wsparcie w tym zakresie, a koszty uruchomienia funkcjonalnej strony internetowej czasopisma, nawet w pełni wspierającej proces gromadzenia tekstów, ich recenzowania i publikacji, nie są wysokie. Skorzystać tu można na przykład z gotowego rozwiązania: darmowego, otwartego oprogramowania Open Journal Systems, posiadającego również polską wersję językową.

Zalety internetowej, publicznej dostępności tekstów naukowych są oczywiste. Dostępność ta wpływa pozytywnie na zasięg i poziom cytowania pracy naukowej, a dokumenty w formie cyfrowej pozwalają na błyskawiczne przeszukiwanie treści, mogą być przy tym łatwo rozpowszechniane i kopiowane. Wciąż jednak nie wszyscy naukowcy pozytywnie oceniają te możliwości, widząc w nich niebezpieczeństwo nieautoryzowanego rozpowszechniania, plagiatowania czy wykorzystania komercyjnego swoich artykułów. Podejście takie wynikać może z uprzedzeń wobec Internetu i nieracjonalnej nostalgii za tradycyjnym modelem obiegu naukowego opartego na druku: w perspektywie możliwości Internetu wydawanie czasopism naukowych jedynie w kilkuset trudno dostępnych egzemplarzach wydaje się czymś zdecydowanie nieracjonalnym i nieefektywnym. Tezy o niebezpieczeństwie komercyjnego wykorzystania tekstów łatwo zbić przypominając, że ochrona prawnoautorska tekstów naukowych jest aktywna bez

względu na medium, za pomocą którego są one rozpowszechniane. Zazwyczaj także niszowy, specjalistyczny charakter ich treści faktycznie nie pozwala myśleć o jakimkolwiek komercyjnym użyciu. Obawa przed nieautoryzowanym rozpowszechnianiem wydaje się nieracjonalna, skoro w dobie powszechnego dostępu do skanerów czy ksero nie można tego zjawiska zablokować nawet w przypadku tradycyjnych papierowych wydań.

Czasopisma i prace naukowe finansowane ze środków publicznych powinny być swobodnie dostępne online. W tym kontekście wybór tradycyjnej papierowej dystrybucji zamiast dystrybucji online uznany być może za formę ograniczania społecznego dostępu do wiedzy wytwarzanej przez naukowców. Oczywiście trudno oczekiwać, że publikowane online specjalistyczne teksty naukowe uzyskają dużą popularność: specyfika materii historycznej pozwala jednak sądzić, że możliwość swobodnego z nich skorzystania może okazać się wartościowa także dla osób czy instytucji funkcjonujących poza systemem akademickim: historyków obywatelskich prowadzących lokalne badania, organizacji pozarządowych oraz szkół. Warto też zapytać, czy blokowanie publikowania zawartości czasopism nie jest czasem próbą ukrycia niewysokiej jakości artykułów naukowych.

Jeśli chodzi o wsparcie organizacyjne, prawne i techniczne w zakresie przenoszenia zawartości czasopism naukowych do Internetu bądź organizowania ich od początku w modelu mieszanym (druk-Internet) lub wyłącznie w formie internetowej, można je uzyskać w Polsce ze strony szeregu instytucji i organizacji. Centrum Cyfrowe: Projekt Polska prowadzi projekt Biblioteka Otwartej Nauki (BON), który koncentruje się głównie na pozyskiwaniu i udostępnianiu online monografii z nauk humanistycznych i społecznych. Do projektu zgłosić się mogą indywidualni autorzy i wydawnictwa naukowe. W ramach BON autorzy i wydawcy mogą bezpłatnie zdigitalizować swoje publikacje oraz uzyskać wsparcie w zakresie ustalenia ich statusu prawnoautorskiego. Pomoc uzyskać można także od ekspertów Koalicji Otwartej Edukacji (KOED). Na początku 2013 roku trwała dyskusja o nowych inicjatywach legislacyjnych zmierzających do wprowadzenia ogólnej zasady pełnej dostępności publikacji naukowych w Internecie: zainteresować się tym problemem powinni nie tylko wydawcy i redakcje czasopism naukowych, ale także autorzy, ponieważ wkrótce okazać się może, że publikowanie online stanie się obligatoryjne.

Wolny, powszechny, trwały i natychmiastowy dostęp do cyfrowych wersji tekstów naukowych określa się mianem otwartego dostępu (Open Access).

Publikowanie w tym modelu daje użytkownikowi pewien podstawowy zasób praw, pozwalający przynajmniej na zapoznanie się z treścią, skopiowanie czy wydrukowanie tekstu z czasopisma. Niektórzy wydawcy wychodzą jednak poza to minimum, publikując swoje tytuły na zasadach pełnej otwartości (z wykorzystaniem wolnych licencji, umożliwiających nawet komercyjne ich wykorzystanie i neutralizujących większość ograniczeń związanych z rozpowszechnianiem). W modelu tym – podobnie jak w żadnej innej formie Open Access – nie dochodzi do rezygnacji z praw autorskich (zrzeczenia się autorskich praw majątkowych).

Oprócz rozmaitych poziomów otwartości otwartego dostępu mamy do czynienia też z dwiema strategiami publikowania w tym modelu. Tzw. złota droga oznacza publikowanie artykułów naukowych wyłącznie w recenzowanych czasopismach naukowych, których treść dostępna jest swobodnie w Internecie. Ponieważ czasopisma te nie mogą sprzedawać płatnej subskrypcji (opłaty za dostęp), koszty procesu recenzentkiego przerzucane mogą być (ale nie muszą) na autora. Polscy badacze skorzystać mogą z finansowania tych opłat, udostępnianego w ramach Narodowego programu publikacji naukowych w Otwartym Dostępie, prowadzonego przez Ministerstwo Nauki i Szkolnictwa Wyższego. Istnieje także tzw. zielona droga Open Access – strategia ta polega z kolei na publikowaniu tekstów naukowych w zamkniętych czasopismach i po okresie embarga umieszczaniu ich w ogólnodostępnych uczelnianych repozytoriach (oczywiście po uzyskaniu zgody wydawcy). Przykładowo, na początku stycznia 2013 roku rektor Uniwersytetu Mikołaja Kopernika w Toruniu wydał zarządzenie, zgodnie z którym UMK powołuje cyfrową Platformę Czasopism UMK, na której obligatoryjnie umieszczane będą artykuły z czasopism wydawanych przez uczelnię. Teksty te dostępne będą na licencji Creative Commons Uznanie Autorstwa-Bez Utworów Zależnych (CC BY-ND 3.0), pozwalającej na swobodne ich rozpowszechnianie (także na zasadach komercyjnych), jednak bez prawa do tworzenia adaptacji (np. tłumaczeń) i ingerowania w oryginalną treść.

Dobłą okazją do zdobycia wiedzy na temat licznych aspektów idei otwartego dostępu jest obchodzony co roku w październiku Tydzień Open Access, w ramach którego organizuje się – także w Polsce – wiele spotkań, konferencji, wykładów czy warsztatów. Fachową pomocą dla wydawców jest także publikacja dra Krzysztofa Siewicza *Otwarty dostęp do publikacji naukowych. Kwestie prawne*, dostępna online i wydana przez Centrum Otwartej Nauki (CEON) działające w ramach Interdyscyplinarnego Centrum Modelowania Matematycznego i Komputerowego Uniwersytetu Warszawskiego. Międzynarodowy katalog



czasopism naukowych, które czytać można w formule otwartego dostępu, rozwijany jest w ramach projektu DOAJ (ang. Directory of Open Access Journals).

Warto też spojrzeć na problem Open Access w Polsce przez pryzmat pewnych faktów. Otóż już dziś studenci i wykładowcy rozpowszechniają w nieautoryzowany przez autorów i wydawnictwa sposób publikacje naukowe, umieszczając je w popularnych repozytoriach plików (takich jak np. Chomikuj.pl). Nieformalny, oddolny dostęp jest więc faktem. Wprowadzenie obligatoryjnej dla wszystkich badaczy polityki otwartego dostępu pozwoliłoby udostępniać zasoby czasopism i prace naukowe w bardziej efektywny i użyteczny sposób.

Internet wpływa nie tylko na dostępność literatury naukowej, ale także ułatwia jej gromadzenie w ramach prowadzonych badań. Dobrym przykładem takiego usprawnienia jest system DOI (ang. Digital Object Identifier) i nowoczesne narzędzia do zarządzania bibliografią, takie jak Zotero czy Mendeley. DOI to nic innego jak cyfrowy identyfikator dokumentu elektronicznego, przypisany bezpośrednio do pliku i zawierający podstawowe informacje o jego cechach. Jeśli artykuły lub monografie naukowe publikowane są profesjonalnie w formie cyfrowej, zawierają identyfikator DOI z takimi informacjami, jak tytuł, imię i nazwisko autora, data wydania itp. DOI pozwala uniezależnić informacje o dokumencie od miejsca jego przechowywania – nawet jeżeli kopie dokumentu rozpowszechniane są niezależnie na wielu stronach internetowych, to zapisane w pliku metadane jednoznacznie go identyfikują. Takie rozwiązanie nie byłoby specjalnie atrakcyjne dla badaczy, o ile metadane zapisane w certyfikacie DOI nie byłyby rozpoznawane przez menadżery bibliografii.

Zotero czy Mendeley to programy pozwalające stworzyć na własnym komputerze bibliograficzną bazę danych. Posiadają one odpowiednie funkcje dotyczące gromadzenia, kategoryzowania i zarządzania informacjami o publikacjach naukowych, które wykorzystujemy w badaniu. Oprócz danych bibliograficznych pozwalają katalogować także pliki publikacji, wyświetlać je, dodawać notatki i komentarze czy generować przypisy w określonym stylu cytowania. Oba programy dostępne są za darmo i posiadają rozszerzenia internetowe, dzięki czemu ich użytkownicy mogą wymieniać się online informacjami o literaturze naukowej oraz tworzyć kopie bezpieczeństwa zbudowanych na komputerach bibliograficznych baz danych.

Rozbudowywanie bazy bibliograficznej nawet z wykorzystaniem menadżera bibliografii może być zajęciem dość czasochłonnym, gdyż każdą pojedynczą pozycję (artykuł w czasopiśmie, artykuł w zbiorze, monografię itp.)

należy osobno opisać, podając przynajmniej podstawowe dane, pomocne potem przy wygenerowaniu odpowiedniego przypisu. Dzięki zapisanemu w plikach publikacji identyfikatorowi DOI menedżer bibliografii błyskawicznie czyta metadane i skopiuje je do bazy.

Zotero czy Mendeley to oczywiście jedynie dwa przykładowe menadżery bibliografii. Programów tego typu jest zdecydowanie więcej, brakuje tu jednak miejsca, aby nawet skrótowo opisać te najważniejsze. Kończąc ten wątek chciałbym polecić szczególnie Zotero, które można zainstalować bezpośrednio w przeglądarce Firefox (co ułatwia prowadzenie badania internetowego) i które posiada interfejs i dokumentację w języku polskim. Zotero jest też też wolnym i otwartym oprogramowaniem.

Temat dostępności w Internecie publikacji naukowych przedstawić można także z perspektywy modeli oceny pracy naukowej. Ocenia się ją dziś za pomocą różnorodnych wskaźników, czyli, mówiąc w bardzo uproszczony sposób, zestawień pokazujących liczbę nawiązań do prac wybranego autora w innych publikacjach. Przykładem takiego rozwiązania jest h-indeks (tzw. indeks Hirscha), przeznaczony także dla nauk społecznych i humanistycznych. Nie wchodząc w szczegóły, można powiedzieć, że im wyższy h-indeks, tym wyższa ocena znaczenia określonych publikacji naukowych oraz całego dorobku autora. Jaki to ma jednak związek z Internetem? Okazuje się, że rozpowszechnianie publikacji naukowych online w trybie otwartego dostępu wpływa pozytywnie na poziom cytowań (Eysenbach, 2006), i co za tym idzie, może przyczynić się do poprawienia oceny autora umieszczonej online pracy. Ma to znaczenie choćby we wnioskach o granty naukowe, gdzie poziom cytowań jest ważnym parametrem podlegającym ocenie. Oczywiście łatwy dostęp do publikacji naukowej jest tylko jednym z czynników wpływających na jej popularność i liczbę odniesień w tekstach innych autorów (Craig i in., 2007).

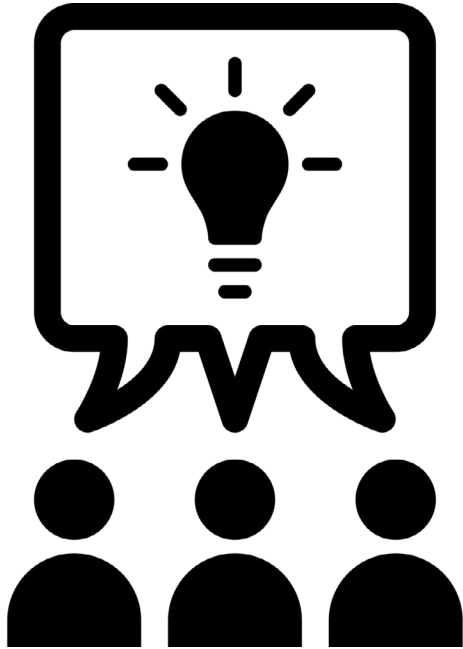
Liczba cytowań wpływa także na ocenę jakości czasopism naukowych. Uznaje się, że im wyższy poziom cytowań dla konkretnego periodyku, tym większe jest jego znaczenie naukowe. Na bazie poziomu cytowań czasopisma ustala się też poziom punktów, które naukowiec otrzymuje, publikując w nim swoje artykuły. Punkty zgromadzone przez naukowców sumuje się w ramach wybranej jednostki naukowej i na tej podstawie ocenia jej aktywność, co ma istotne znaczenie np. przy rozdziale funduszy publicznych i aplikowaniu o zespołowe granty badawcze.

Dobrą praktyką – zwłaszcza w przypadku młodych autorów, zaczynających dopiero swoją działalność naukową – jest udostępnianie w Internecie wczesnych wersji swoich publikacji jeszcze przed ich umieszczeniem w drukowanych, recenzowanych czasopismach (tzw. *preprint*). Biorąc pod uwagę, że proces recenzyjny w redakcji czasopisma może trwać bardzo długo, pozwala to na przyspieszenie recepcji pracy wśród innych badaczy zainteresowanych tematem. Co więcej, taki krok umożliwia zebranie dodatkowych ocen i sugestii, które mogą pomóc w wypracowaniu finalnej wersji publikacji przeznaczonej już do druku.

## BIBLIOGRAFIA

- Craig, I. D., Plume, A. M., Mcveigh, M. E., Pringle, J., Amin, M.** (2007). Do Open Access Articles Have Greater Citation Impact?, *Journal of Informetrics*, 1(3), 239-248. Pobrano z: <http://www.sciencedirect.com/science/article/pii/S1751157707000466>.
- Eysenbach, G.** (2006), Citation advantage of Open Access articles, *PLoS biology*, 4(5). Pobrano z: <http://www.plosbiology.org/article/info:doi/10.1371/journal.pbio.0040157>.
- Swartz, A.** (2008), Guerilla Open Access Manifesto, Pobrano z: <https://archive.org/details/GuerillaOpenAccessManifesto>.





# Praca grupowa i nauka obywatelska

Internet może być dla historyka nie tylko źródłem literatury naukowej, zasobów źródłowych czy informacji o konferencjach lub grantach, ale także przestrzenią współpracy. Przykładowo, wspomniane w poprzednim rozdziale programy do zarządzania bibliografią, takie jak Zotero czy Mendeley, ułatwiają wymianę danych bibliograficznych i informacji o zasobach internetowych i publikacjach, przydatnych z punktu widzenia badacza. Użytkownicy w ramach kont rejestrowanych w serwisach Zotero.org czy Mendeley.com tworzą własne, tematyczne katalogi bibliograficzne, do których uzupełniania mogą zaprosić innych.

Oczywiście wymiana informacji o literaturze naukowej to tylko jedna z możliwości współdziałania online. W polskim systemie edukacji akademickiej, przynajmniej na kierunkach humanistycznych, bardzo rzadko zwraca się uwagę na możliwość kolektywnej pracy badawczej – efektem tego jest np. niewielka popularność wieloautorskich artykułów w czasopismach historycznych czy ogólna niechęć do dzielenia się wynikami własnej pracy naukowej w publicznym Internecie, z reguły z obawy przed komercyjnym ich wykorzystaniem (na co, szczerze mówiąc, z reguły nie ma nawet szans). Nie chodzi o to, aby kolektywny model publikowania stał się dominujący – tak jak np. w naukach biologicznych – ale aby nie odrzucać od razu jego potencjału i wykorzystywać go, kiedy wydaje się najbardziej optymalny.

W jaki jednak sposób dwoje czy kilku nawet autorów może wspólnie sprawnie pracować nad jednym tekstem? Można do tego celu wykorzystać narzędzia udostępniane bezpłatnie przez Google w ramach usługi Google Dysk

(Google Drive). Po zarejestrowaniu konta mamy do dyspozycji w pełni funkcjonalny pakiet biurowy z edytorem tekstu (Google Docs), pozwalającym na grupową pracę nad dokumentami. Dzięki opcjom udostępniania treści innym użytkownikom, archiwizacji wersji, komentarzom czy wsparciu dla komunikacji między piszącymi, Google Docs może zostać wykorzystane np. do wspólnego redagowania sprawozdań z konferencji naukowych, współpracy między studentem a promotorem czy wspólnego gromadzenia danych i notatek w ramach grupowej pracy badawczej. Użyteczny interfejs i możliwość pracy na dokumentach bez względu na komputer, z którego się korzysta – to cechy, które do rozwiązań Google przekonać mogą także naukowców pracujących samodzielnie. Wykorzystując Google Drive w swojej pracy naukowej warto jednak mieć świadomość kontrowersyjnych zasad polityki Google wobec treści tworzonych w ramach swoich usług: wszystkie te materiały Google ma prawo niezależnie wykorzystywać, modyfikować i publikować. Nie oznacza to oczywiście, że treści pisane w ramach naszego konta Google będą udostępniane publicznie, jednak przez taki, a nie inny regulamin Google ich status prawnoautorski wydaje się niejasny. W takiej sytuacji warto pomyśleć o alternatywnych rozwiązaniach.

O tym, że dzięki Internetowi w pracy grupowej nad publikacją można wyjść poza ograniczenia edytora tekstu, świadczy projekt *Writing History in the Digital Age* (<http://writinghistory.trincoll.edu/>). Książka, kolektywnie napisana i zredagowana na przełomie 2011 i 2012 roku, została udostępniona na jednej z licencji Creative Commons w serii *Digital Culture Books* (2012), prowadzonej przez wydawnictwo Uniwersytetu Michigan i czeka na swoją ebookową wersję. Do przygotowania publikacji wykorzystano narzędzie Comment Press, wtyczkę do popularnego systemu zarządzania treścią, Wordpressa, które pozwala w prosty sposób komentować poszczególne paragrafy publikowanych treści. Do współpracy przy książce zaproszono wszystkich chętnych, każdy mógł zgłosić propozycję rozdziału, zaakceptowane rozdziały były także publicznie recenzowane (zebrano ponad 900 komentarzy), a nad całością publikacji czuwali oczywiście redaktorzy. W projekcie wzięło udział 73 autorów. *Writing History in the Digital Age* jest poniekąd eksperymentem na swój własny temat – traktując o nowych formach pisania o przeszłości, wdraża taką właśnie formę pracy.

Opisany wyżej projekt jest nowatorski przynajmniej z trzech powodów: po pierwsze, pozwala on na uruchomienie otwartego procesu konsultacyjnego

i recenzenckiego już na samym początku planowania publikacji (wybór tematów rozdziałów) i proces ten właściwie nigdy się nie kończy, ponieważ nawet po publikacji finalnej wersji książki można dodawać publiczne komentarze do poszczególnych akapitów jej treści. Po drugie, każdy zainteresowany i posiadający odpowiednią wiedzę może mieć swój udział w tworzeniu treści, co daje szansę na nowe, interesujące spojrzenie na poruszany problem. Po trzecie wreszcie, książka – jak dotąd – nie posiada swojej drukowanej wersji, bo w końcu liczy się nie forma dystrybucji, ale jakość treści. Zresztą wykorzystanie licencji Creative Commons pozwala na wydrukowanie książki i jej niezależną redystrybucję bez konieczności uzyskania osobnej zgody wydawcy.

W 2006 roku Roy Rosenzweig opublikował esej *Can History be Open Source? Wikipedia and the Future of the Past*. Opisał w nim zasady funkcjonowania Wikipedii i charakter publikowanych w niej haseł historycznych, przy czym zwrócił uwagę na to, że dla historyków może być ona przykładem nie tylko skuteczności otwartego modelu rozpowszechniania wiedzy, ale też skuteczności demokratycznego modelu jej produkcji (Rosenzweig, 2006). Idąc tym torem rozumowania, można wyobrazić sobie taką organizację pracy badawczej zespołu historyków, w której działają oni online, wykorzystując system wiki do przygotowania publikacji albo raportu z badań, co daje łatwą kontrolę nad kolejnymi wersjami tworzonych wspólnie tekstów i zarazem umożliwia łatwe przeszukiwanie treści czy tworzenie indeksów. Ostateczny efekt ich pracy zostaje udostępniony w Internecie na wolnej licencji – a przynajmniej w trybie otwartego dostępu – i może być wykorzystywany w innych projektach badawczych czy w edukacji. Pojawia się także tutaj – sygnalizowana także w eseju Rosenzweiga – możliwość udziału w pracach badawczych osób spoza systemu akademickiego, w ramach nauki obywatelskiej (ang. *citizen science*).

Dobrym przykładem realizacji idei pracy grupowej i nauki obywatelskiej w środowisku historyków jest istniejący w ramach portalu Monasterium.net projekt *Monasterium Collaborative Archive*. Monasterium.net to bogate repozytorium zawierające cyfrowe wersje (skany) ponad 250 tys. średniowiecznych i wczesnonowożytnych dokumentów udostępnianych przez prawie 60 archiwów i bibliotek 8 europejskich krajów. Skany te opisywane i transkrybowane są przez użytkowników serwisu, a nad jakością edycji czuwa 14 moderatorów, ekspertów w dziedzinie dyplomatyki. Zatwierdzone edycje są dostępne dla wszystkich użytkowników. Na stronie internetowej projektu możemy przeczytać, że nad skanami pracuje około 150 osób: profesjonalnych badaczy, ale także studentów i wolontariuszy niezwiązanych ze środowiskiem akademickim.



Projekty nauki obywatelskiej tym głównie różnią się od wielu rozmaitych okołonaukowych inicjatyw popularyzatorskich i edukacyjnych realizowanych w Internecie, że ich zakładanym efektem i celem jest uzyskanie konkretnego, merytorycznego wsparcia w realizacji jakiegoś zadania o charakterze badawczym (taki jest choćby efekt społecznościowej transkrypcji we wspomnianym wyżej portalu *Monasterium.net*). Oczywiście nie oznacza to, że nie można mówić o edukacyjnym czy popularyzatorskim potencjale takich działań – nie jest on jednak najważniejszy.

Udział w projektach nauki obywatelskiej biorą osoby nieposiadające odpowiedniej wiedzy naukowej, ale chcące się zaangażować np. w analizowanie dużych zestawów danych czy transkrypcję dokumentów. Przykładowo, w organizowanym przez archiwum amerykańskie *1940 Census Community Indexing Project* wolontariusze za pomocą specjalnie przygotowanego serwisu internetowego transkrybowali zapisy z ponad 3 milionów skanów kart statystycznych, za pomocą których gromadzono dane o amerykańskich obywatelach podczas spisu ludności w 1940 roku. Dzięki pracy tysięcy osób udało się stworzyć przeszukiwany indeks nazwisk, tak pomocny w badaniach genealogicznych. Na podobnej zasadzie zorganizowano brytyjski projekt *Old Weather*, polegający na transkrybowaniu kart pogodowych z archiwów marynarki wojennej w celu zgromadzenia danych dotyczących zmian klimatu. To, że uczestnicy takich naukowych projektów nie muszą mieć wiedzy na temat przedmiotu badań, które wspierają swoją pracą, pokazuje dobrze projekt *AncientLives* – podobnie jak *Old Weather* organizowany w ramach brytyjskiego programu *Zooniverse*, prowadzonego przez *Citizen Science Alliance (CSA)*. W *AncientLives* wolontariusze otrzymują do zanalizowania fragmenty spisanego na papirusie greckiego rękopisu Nowego Testamentu, datowanego na początek III wieku (*papyrus Oxyrhynchus*). Przygotowany dla nich system pozwala na transkrybowanie poszczególnych fragmentów rękopisu nawet bez znajomości greckiego alfabetu. W projekcie tym *Citizen Science Alliance* współpracuje bezpośrednio z badaczami z uniwersytetu oksfordzkiego.

Nieco inny model partycypacji nie-naukowców w działaniach o charakterze naukowym zastosowano w ramach projektu *Children of the Lodz Ghetto* (<http://online.ushmm.org/lodzchildren/>), prowadzonego przez *United States Holocaust Memorial Museum*. Tutaj, w oparciu o przygotowaną wcześniej alfabetyczną listę nazwisk i udostępniane przez muzeum bazy danych, wolontariusze tworzą profile biograficzne dzieci z łódzkiego getta. Praca uczestników projektu polega więc nie tyle na transkrypcji, ile na uzupełnianiu informacji o konkretnej

osobie, w oparciu o dostępne w systemie indeksy, bazy danych, repozytoria fotografii, historii mówionej itp. Jest to więc zadanie, któremu, ze względu choćby na zróżnicowanie formatów danych, nie podoła żaden automatycznie działający program komputerowy.

Podczas jednej z edycji międzynarodowej konferencji *Museums and the Web*, w 2011 roku, zaprezentowano studium przypadku wspomnianego przeze mnie wyżej projektu *Old Weather* (Romeo, Blaser, 2011). W opublikowanym na stronie konferencji artykule znalazły się bardzo ciekawe dane, dotyczące m.in. motywacji osób uczestniczących w działaniach w ramach nauki obywatelskiej. Okazało się, że oprócz zainteresowania konkretnym tematem dużą rolę przy podjęciu decyzji o wsparciu badań pełnią też zabawa i wzajemna, pozytywna rywalizacja uczestników. W przypadku *Old Weather* zorganizowano nawet grę, w ramach której wolontariusze, w zależności od postępów w transkrybowaniu kart pogodowych z konkretnego statku, uzyskiwali wirtualny tytuł kapitana. Postępy w zbieraniu danych wizualizowane były za pomocą mapy, po której poruszały się statki, uczestnicy natomiast starali się jak najszybciej docierać do docelowych portów – oczywiście jedynym na to sposobem była intensywniejsza praca.

Warto wspomnieć, że również w Polsce podejmowane są próby uruchomienia podobnych inicjatyw. Fundacja Centrum Cyfrowe: Projekt Polska prowadzi projekt *Otwarte zabytki* ([Otwartzabytki.pl](http://Otwartzabytki.pl)), w ramach którego siłami wolontariuszy rozbudowywany jest katalog polskich zabytków, oparty o podstawowe dane zebrane z krajowego Rejestru Zabytków, publikowanego przez Narodowy Instytut Dziedzictwa. Prosty indeks obiektów ma zostać przekształcony w obszerną bazę danych udostępniającą wiele informacji i zasobów związanych z konkretnym budynkiem (np. fotografie, dane geolokalizacyjne czy informacje o wydarzeniach historycznych z nim związanych).

Polscy badacze, którzy chcieliby przetestować potencjał nauki obywatelskiej, mogą skorzystać z gotowego narzędzia, jakim jest Wirtualne Laboratorium Transkrypcji (<http://wlt.synat.pcass.pl/>), przygotowane przez Poznańskie Centrum Superkomputerowo-Sieciowe (PCSS). Serwis ten pozwala na publikowanie skanów dokumentów historycznych i ich społecznościową transkrypcję, z zachowaniem odpowiedniej kontroli nad wprowadzanymi przez użytkowników zmianami. W prosty sposób można też transkrybować publikacje dostępne w Federacji Bibliotek Cyfrowych dzięki funkcji ich automatycznego importu do systemu.

Problemem przy budowaniu społecznościowego wsparcia dla projektów badawczych w ramach koncepcji nauki obywatelskiej są wysokie koszty organizacyjne: opisane wyżej inicjatywy wykorzystują specjalnie przygotowane narzędzia internetowe, a angażujący się w działania wolontariusze otrzymują pomoc i są nieustannie zachęceni do dalszej pracy. Z pewnością dużym wyzwaniem – jeszcze na poziomie planowania – jest odpowiednie określenie ich zadania: nie może być ono zbyt skomplikowane czy zbyt nudne i wymagające specjalistycznej wiedzy, a przy tym musi przynosić konkretne efekty. Mimo tego warto zastanowić się, czy nie jest to dobry model na realizację przynajmniej tych specyficznych celów badawczych, które ze względu na swoją skalę mogą być bardzo czasochłonne dla jednego czy niewielkiej grupy badaczy. W razie potrzeby odpowiedniego wsparcia technicznego szukać można we wspomnianym już wyżej Citizen Science Alliance, które udostępnić może zestaw narzędzi internetowych, np. dla projektu transkrypcyjnego.

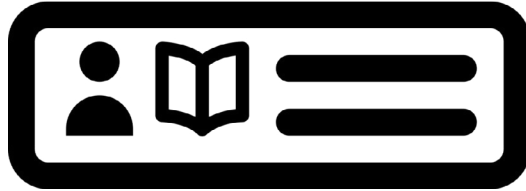
Przedstawione na poprzednich stronach rozmaite przykłady współdziałania online – od prostej wymiany informacji bibliograficznych w Zotero czy Mendeley, przez alternatywne formy redagowania publikacji, aż po naukę obywatelską – są przejawami tego samego zjawiska: otwierania się systemu naukowego. Wydaje mi się, że wobec tak wielkiego komunikacyjnego potencjału Internetu historyk pracujący na własną rękę i nieinformujący na zewnątrz o swoich działaniach, choćby przez publikowanie w modelu *open access*, działa nieracjonalnie oraz wbrew interesowi społeczeństwa, które ma prawo do wiedzy o rezultatach badań finansowanych z publicznych źródeł.

Wielu naukowców obawia się tej otwartości, uznając Internet za przestrzeń, w której oferowana przez nich wiedza zostanie niebezpiecznie zagłuszona przez wszechobecną ignorancję. Krytycy Wikipedii podkreślają na przykład, że w jej strukturach promowane są nieprofesjonalna wiedza i niesprawdzone informacje, na skalę uniemożliwiającą jakiegokolwiek naukowego zaangażowanie w ten projekt. Roy Rosenzweig w cytowanym wyżej eseju odnosi się do tego problemu i nie stara się bynajmniej wybielać wad Wikipedii: rzeczywiście, choćby ze względu na zasadę NOR (*no original research – zakaz opierania się na twórczości własnej*) aktywny udział naukowców w redagowaniu haseł na tematy, którymi profesjonalnie się zajmują, jest mocno utrudniony. Z drugiej jednak strony proponuje pozytywny program: zamiast biernie krytykować jakość wiedzy dostępnej online, lepiej aktywnie wzbogacić zasoby Internetu o dobre, merytoryczne treści naukowe. Można to zrobić otwierając system naukowy:

publikując w trybie otwartego dostępu czy wzmacniając obecność naukowców w Sieci i społeczną rozpoznawalność ich pracy. Skoro nauka tak słabo obecna jest w tradycyjnych, tabloidyżujących się mediach, powinna być jak najsilniej reprezentowana w Internecie. Tutaj nie istnieje problem *prime time* czy oczekiwań reklamodawców, sprawiających, że oferta telewizyjna czy prasowa staje się coraz bardziej infantylna i nauka funkcjonuje tam wyłącznie w kategorii ciekawostek. O tym, jak historycy mogą stać się aktywnie obecni w Internecie, piszę w kolejnym, ostatnim rozdziale tego podręcznika.

## BIBLIOGRAFIA

- Romeo, F., Blaser L.** (2011). Bringing Citizen Scientists and Historians Together. W: Trant, J., Bearman, D., (red.), *Museums and the Web 2011: Proceedings, Archives & Museum Informatics*, Toronto. Pobrano z: [http://conference.archimuse.com/mw2011/papers/bringing\\_citizen\\_scientists\\_historians\\_together](http://conference.archimuse.com/mw2011/papers/bringing_citizen_scientists_historians_together).
- Rosenzweig, R.** (2006). Can History Be Open Source? Wikipedia and the Future of the Past, *Journal of American History*, 93(1), 117-146.



# Nowe internetowe formy komunikacji naukowej

Recenzowane publikacje naukowe to podstawowa forma dystrybucji wiedzy, jednak z pewnością nie jedyna. Dlatego mówiąc o otwieraniu się systemu naukowego, nie wystarczy wspomnieć o otwartym dostępie do tekstów naukowych. Internet pozwolił na wykształcenie się innych form komunikacji w ramach systemu naukowego i poza nim. Jedną z najbardziej interesujących i mających chyba największy potencjał są blogi.

Blogi – obok Wikipedii – to jeden z najczęstszych argumentów wysuwanych przez krytyków otwierania się systemu nauki na Internet. W ich opinii symbolizują one największe wady tego medium: niskiej jakości treści, brak poszanowania dla praw autorskich czy chamstwo wypełniające komentarze. Takie podejście uniemożliwia jakąkolwiek refleksję nad możliwościami wykorzystania blogów w nauce i jest poważnym błędem. Uznanie, że blog jest formą internetowego pamiętnika i z tego powodu jego prowadzenie nie przystoi naukowcowi, to nieporozumienie: nie można traktować tej formy e-pisarstwa jako gatunku medialnego, gdyż jest to jedynie forma organizacji treści, oparta o schemat chronologicznie dodawanych wpisów (notek). Jedyne gatunkowe cechy bloga, o których można mówić, związane są z cechami wykorzystywanego w nim języka: z reguły nawet w profesjonalnych blogach naukowych nie jest on tak skostniały i formalny jak w literaturze naukowej. W formie bloga mieścić się może wiele rozmaitych gatunków pisarskich, w tym recenzje artykułów naukowych, czego świetnym przykładem jest serwis [Researchblogging.org](http://Researchblogging.org), agregujący właśnie notki blogowe o takiej treści.

W anglosaskiej kulturze naukowej posiadanie bloga nie jest może standardem, ale na pewno nie jest też ewenementem. Dość powiedzieć, że Amerykańskie Towarzystwo Historyczne już od 2005 roku patronuje dorocznej nagrodzie dla najlepszych historycznych blogów naukowych (Cliopatria Awards), w 2013 roku ufundowało też nagrodę dla najlepszego artykułu historycznego publikowanego w Internecie poza systemem czasopism naukowych. Notki z blogów bywają publikowane także w profesjonalnych wydawnictwach naukowych, jak choćby we wspomnianej na pierwszych stronach tego podręcznika książce *Debates In the Digital Humanities*. W czasopismach naukowych, np. *Journal of American History*, publikowane są zasady cytowania blogów.

Różne są teorie dotyczące pochodzenia blogów naukowych. Niektórzy zwracają uwagę na to, że są one naturalną kontynuacją istnienia stron domowych naukowców, którzy tworzyli niezależne od macierzystej uczelni strony WWW, informując na nich o swoich zainteresowaniach badawczych i publikując zestawienia bibliograficzne własnych prac. Można też zastanawiać się nad rolą usenetowych list dyskusyjnych – wcześniejszych internetowych form komunikacji między naukowcami a osobami zainteresowanymi nauką.

Inna Kouper wylicza cztery cele blogów naukowych w swoim artykule, poświęconym ich roli w promocji nauki: informowanie o aktualnościach naukowych (ang. *scientific news*), wyjaśnianie skomplikowanych zagadnień naukowych za pomocą odpowiednio dostosowanego języka, recenzowanie prac naukowych publikowanych przez innych badaczy oraz formułowanie z perspektywy naukowej stanowisk wobec ważnych społecznie kwestii (Kouper, 2010, s. 2). Warto zauważyć, że taki profil treści może być atrakcyjny nie tylko dla osób spoza środowiska naukowego, zainteresowanych nauką (o czym właśnie pisze Kouper), ale także dla osób w tym środowisku funkcjonujących. Blogowanie może wspierać komunikację naukową: ułatwiać rozpowszechnianie informacji o konferencjach naukowych, nowych publikacjach, dostępnych grantach badawczych itp.

Próby opisu stanu polskiej blogosfery naukowej podjął Emanuel Kulczycki w swoim artykule przygotowanym w ramach odbywającej się w dniach 13–14 kwietnia 2012 w Toruniu III Międzynarodowej Konferencji Open Access pt. „Otwarta nauka i edukacja” (Kulczycki, 2012a). Na podstawie przeprowadzonych wywiadów z polskimi blogerami-naukowcami zwraca on uwagę m.in. na to, że podjęcie decyzji o blogowaniu może być także wyrazem niezgody na przestarzały i niesprawny model komunikacji naukowej oraz próbą znalezienia

efektywniejszego sposobu nie tylko informowania o własnych badaniach, ale także promocji zainteresowania nauką wśród osób spoza systemu akademickiego.

Przykładem polskiego historycznego bloga naukowego jest Herstoria.blox.pl. Jest on prowadzony przez Annę Nowakowską-Wierzchoś, doktorantkę Instytutu Historii Polskiej Akademii Nauk, pracującą na stanowisku starszej archiwistki w Archiwum Akt Nowych w Warszawie. W wywiadzie przeprowadzonym przez Emanuela Kulczyckiego wspomina ona o wyzwaniach i problemach związanych z taką formą działalności (około) naukowej, jakimi są m.in. obawy o reakcję przełożonych na prowadzenie bloga pod własnym imieniem i nazwiskiem:

pisanie na stronach internetowych, a tym bardziej blogów, nadal nie jest traktowane poważnie, a już pisanie blogów naukowych odbierane jest przez środowisko naukowe jako strata czasu. Bo nikt za to nie płaci, udostępnia się powszechnie swój dorobek i odkrycia (Kulczycki, 2012b).

Blogi są jedynie pojedynczym elementem bardzo zróżnicowanego systemu komunikacji internetowej, w którym dominującą rolę odgrywają media społecznościowe (ang. *social media*), czyli serwisy takie jak Facebook, Twitter czy Google+. To również alternatywne kanały dystrybucji informacji naukowych i promocji nauki, w tym także nauki historycznej, archiwalnej czy muzealnictwa. Przykładowo w serwisie Twitter, umożliwiającym komunikację między użytkownikami za pomocą wpisów o objętości jedynie 140 znaków, istnieje wiele grup integrujących badaczy i pracowników instytucji, takich jak archiwa czy muzea z całego świata. Wpisy na Twitterze opisywane są odpowiednimi słowami kluczowymi – hashtagami (ang. *hashtags*). Jednym z najpopularniejszych i używanych przez historyków obecnych na Twitterze jest #twitterstorians, stworzony przez niemiecką badaczkę Katrinę Gulliver. We wpisach oznaczonych tym tagiem historycy znaleźć mogą informacje o planowanych konferencjach, nowych publikacjach, *call for papers* czy możliwościach grantowych. Z perspektywy polskich badaczy najbardziej użyteczne będą oczywiście informacje o nowych, wydawanych za granicą książkach z interesujących ich dziedzin – w tym sensie jest to zdecydowanie efektywniejsze źródło informacji niż recenzje w polskich czasopismach historycznych czy strony polskich bibliotek akademickich. W artykule *Twitter in der Wissenschaft: Ein Leitfaden für Historiker/innen* Mareike König (König, 2012) opisuje inne tagi, które warto obserwować w ramach



zarejestrowanego na Twitterze konta. Zwraca też uwagę na fakt, że, oprócz historyków, na Twitterze komunikuje się także wiele instytucji badawczych czy bibliotek, co może być również użyteczne. Dzięki temu narzędziu badacze z Polski w prosty sposób włączyć się mogą w globalny system komunikacji naukowej – jedyną barierą korzystania z niego jest znajomość języka.

Pisząc o znaczeniu mediów społecznościowych we współczesnej komunikacji naukowej nie można nie wspomnieć o Facebooku, który także w Polsce zdobył ogromną popularność. Serwis ten oferuje wiele narzędzi komunikacyjnych – czat, system wiadomości wysyłanych bezpośrednio między użytkownikami czy formułę strony fanowskiej (*fanpage*), w ramach której funkcjonuje już wiele polskich akademickich instytucji historycznych czy promowane są konferencje naukowe. Facebook stał się także infrastrukturą służącą komunikacji w ramach projektów naukowych – dobrym przykładem jest tutaj forum Memory at War. Jest to facebookowa grupa dyskusyjna poświęcona problematyce historycznej pamięci zbiorowej społeczeństw Polski, Ukrainy i Rosji (w szerszej perspektywie oczywiście także innych krajów Europy Środkowej i Wschodniej). Grupa została założona przez uczestników międzynarodowej inicjatywy badawczej, eksplorującej temat konfliktów pamięci w tym regionie, zainicjowanej przez uniwersytet Cambridge i prowadzonej we współpracy z uczelniami z Holandii, Estonii, Finlandii i Norwegii. Jeśli zdecydujemy się na zorganizowanie internetowej infrastruktury projektu naukowego czy stworzenie profilu czasopisma lub konferencji akurat na Facebooku, warto pamiętać o tym, że administratorzy tego serwisu często dość swobodnie interpretują zapisy obowiązującego wszystkich użytkowników regulaminu, dlatego też twórcy, w przypadku zablokowania lub usunięcia strony, nie mają zbyt wielu możliwości dochodzenia własnych praw. Z drugiej jednak strony Facebook umożliwia potencjalnie łatwe dotarcie do wielu odbiorców oraz pozytywnie wpływa na komunikację między instytucją (uczelnia, archiwum, muzeum) a odbiorcami informacji. To także użyteczna przestrzeń promocji nauki, o czym świadczą liczne prowadzone tam profile blogów naukowych czy profile samych naukowców.

Serwisem społecznościowym już o wyraźnie naukowym charakterze jest, zdobywająca użytkowników także wśród polskich badaczy, Academia.edu. Po założeniu konta można zbudować tam swój publiczny profil naukowy, w ramach którego udostępniane są: lista publikacji (można zezwolić na ich pobieranie), informacje o zainteresowaniach naukowych, o afiliacji, podstawowe dane składające się na akademickie CV itp. Posiadanie konta umożliwia łatwe kontaktowanie się z innymi badaczami oraz dotarcie do publikowanych przez nich materiałów.

Pokazane wyżej formy obecności naukowca w Internecie różnią się poziomem otwartości i zaangażowania. Nie można zapominać, że obecność ta jest jedynie dodatkiem do pracy naukowej i, choć może ją wspierać, np. dzięki umożliwieniu szybkiej komunikacji z innymi badaczami, nie powinna być rozwijana kosztem czasu i sił przeznaczonych na prace merytoryczne, tym bardziej że w obowiązującym dziś systemie ewaluacji aktywności naukowej polscy badacze nie otrzymują żadnych punktów za prowadzenie bloga naukowego ani za inne formy promocji nauki. Z drugiej jednak strony, aktywna obecność w Internecie jest dziś standardem. Już sama dbałość o aktualizowanie bibliografii własnych prac na witrynie uczelnianej przynosi pozytywny efekt, nie mówiąc już o publikowaniu własnych prac w otwartym dostępie, np. w uczelnianym repozytorium. W dobie medialnych dyskusji kwestionujących zasadność finansowania nauk humanistycznych i społecznych powinno nam zależeć na jak najlepszej widoczności nauk historycznych w Internecie, ponieważ jest to jedyne medium, które pozwala komunikować treści naukowe w obiegu społecznym w sposób nieograniczony wolą reklamodawców czy objętością szpalt czasopism drukowanych.

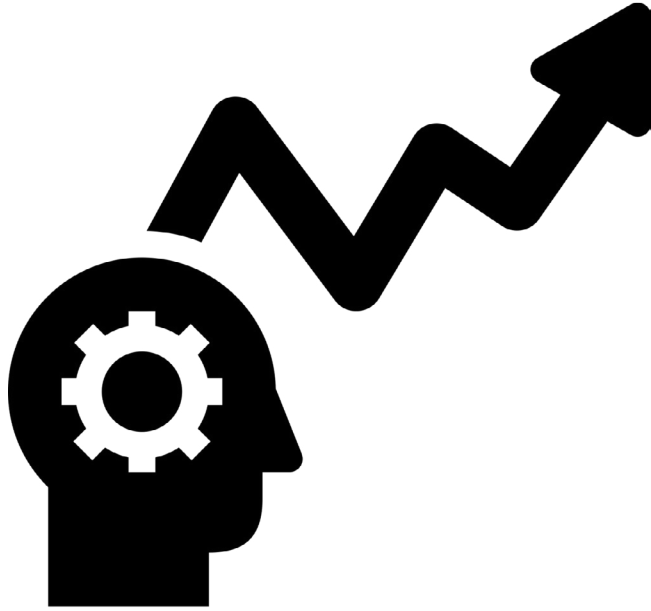
Promocja nauki nie musi odbywać się wyłącznie w Internecie. Można wspomnieć tu choćby o piknikach naukowych – organizowanych także w Polsce – czy o innych publicznych wydarzeniach tego typu. Inicjatywy te skierowane są jednak raczej do odbiorcy masowego, stąd ich poziom merytoryczny jest ograniczony. Formą promocji nauki i projektów naukowych oraz intensyfikowania obiegu koncepcji badawczych, skierowaną raczej do samego środowiska naukowego, wciąż jednak otwartą na uczestników spoza niego, są niekonferencje (*unconference*). Koncepcja niekonferencji pochodzi wprost ze środowiska informatycznego i oznacza otwarte, nieformalne, ale merytoryczne spotkania, podczas których wymienia się wiedzę i umiejętności oraz nawiązuje kontakty. W odróżnieniu od konferencji naukowych, nie prezentuje się tam klasycznych wystąpień i nie publikuje artykułów pokonferencyjnych – dominują krótkie prezentacje multimedialne, rejestrowane też zazwyczaj na video i udostępniane później w Internecie. Otwarty model niekonferencji oznacza, że każdy może zgłosić chęć aktywnego uczestnictwa. Z tego powodu często nawet nie przygotowuje się wstępnych list wystąpień i dopiero tuż przed rozpoczęciem prezentacji zamyka się listę występujących osób. Spotkania mają charakter nieformalny – rezygnuje się z akademickiej tytułomanii, skupiając uwagę na potencjale bezpośredniej rozmowy i swobodnej wymiany wiedzy, często między przedstawicielami różnych dziedzin.

Przykładem humanistycznej niekonferencji jest organizowany od 2008 roku THATCamp (ang. *The Humanities And Technology Camp*), gromadzący m.in. historyków, bibliotekarzy, pracowników muzeów i archiwów oraz informatyków. Organizatorem pierwszego THATCampu i koordynatorem organizowanych na całym świecie kolejnych jego odsłon jest wspomniane już w tym podręczniku amerykańskie Center For History and New Media (CHNM). Pierwsza polska edycja THATCampu została zorganizowana w Lublinie w październiku 2012 roku, kolejne odbywały się w 2013 roku w Poznaniu, Lublinie, Warszawie i Gdańsku. Więcej informacji o THATCampie znaleźć można na stronach THATCamp.org i THATCamp.pl.

## BIBLIOGRAFIA

- Kouper, I.** (2010). Science blogs and public engagement with science: Practices, challenges, and opportunities, *Journal of Science Communication* (JCOM), 9 (March), 1-10.
- König, M.** (2012). Twitter in der Wissenschaft: Ein Leitfaden für Historiker/innen. *Digital Humanities am DHIP*, 21.08.2012. Pobrano z: <http://dhdhi.hypotheses.org/1072>
- Kulczycki, E.** (2012a). Blogi i serwisy naukowe. Komunikacja naukowa w kulturze konwergencji. W: Sójkowska, I. (red.), *Materiały konferencyjne EBIB nr 22* (s. 1-24), Toruń: Stowarzyszenie EBIB.
- Kulczycki, E.** (2012b). Sylwetka blogerki naukowej: Anna Nowakowska-Wierzchoś, *Warsztat badacza*. Pobrano z: [http://ekulczycki.pl/teoria\\_komunikacji/sylwetka-blogerki-naukowej-anna-nowakowska-wierzchos/](http://ekulczycki.pl/teoria_komunikacji/sylwetka-blogerki-naukowej-anna-nowakowska-wierzchos/).





# Internet i społeczna rola nauki historycznej

Teoretycy historii czy socjologowie badający społeczne odniesienia do przeszłości zwracają dziś uwagę na występowanie oznak wyraźnego kryzysu społecznego znaczenia nauki historycznej. Jak pisze o tym Jan Pomorski, akademicka historiografia utraciła swój monopol na określanie tego, co należy do historii, i na potwierdzanie prawdy o faktach z przeszłości (Pomorski, 2011, s. 76–77). Utrata tej dominującej pozycji nie musi być oceniana wyłącznie negatywnie: do głosu dochodzą przecież historie alternatywne, lokalne, mniejszościowe, eksperymentalne, na które do tej pory brakowało miejsca w przestrzeni akademickiej. Nie bez znaczenia jest tutaj Internet jako przestrzeń udostępniania źródeł i opracowań oraz promocji materiałów edukacyjnych. Pojawiają się nowe media, za pomocą których relacjonowane i dokumentowane są aktualne wydarzenia, mogące być w przyszłości przedmiotem badań historyków: serwisy społecznościowe, blogi, Wikipedia itp.

Z drugiej jednak strony obserwować można postępującą inflację powszechnej wiedzy historycznej, świadczącą o wyraźnym kryzysie efektywności tradycyjnej szkolnej edukacji w tym zakresie oraz zdominowanie przez media społecznej refleksji nad historią. Nośnikami wiedzy historycznej – a może bardziej pamięci historycznej, łatwiejszej do przyjęcia ze względu na jej naturalny, niekrytyczny charakter – są dziś filmy, komiksy, książki, pokazy organizowane przez grupy odtwórstwa historycznego, a także Internet: Wikipedia, fora dyskusyjne, a nawet memy. Oczywiście treści te mieszają się ze sobą, co opisał Henry Jenkins, mówiąc o zjawisku konwergencji mediów i wskazując na nowe

kompetencje kulturowe, wykształcane przez użytkowników Internetu (Jenkins, 2007). Współczesna kultura historyczna, jak przekonuje z kolei Andrzej Szpociński, jest wizualna i afektywna: społeczne odwołania do przeszłości formatowane są przez potrzeby uczestnictwa: rekonstrukcje historyczne, koncerty patriotyczne czy rowerowe masy krytyczne z okazji rocznic wybuchu powstania warszawskiego, gry komputerowe. Do tego dochodzą potrzeby wpływające z emocji i poczucia estetyki (filmy, plakaty, piosenki itp.):

Kultura historyczna wieku XX była kulturą intelektualną – zamkniętą w przekazach, które należało interpretować, poddawać krytycznym rozbiorom. Dominacja przekazów pozawerbalnych – obrazów, happeningów, performansów – sprawia, że ta forma obcowania z przeszłością zanika. W kulturze wizualnej to nie intelekt, lecz zmysły odgrywają pierwszoplanową rolę w doświadczaniu przeszłości (Szpociński, 2010, s. 16).

Rozpowszechnianie krytycznej wiedzy historycznej jest w takich warunkach trudne, tym bardziej, że media nie są neutralnym przekazywaczem informacji, jako że programują ją w celu zwiększenia swojego zasięgu i zysków (tabloidyzacja). Przykładem tego zjawiska mogą być wypowiedzi historyków przytaczane w telewizyjnych programach informacyjnych – kilkudziesięciosekundowe, zdawkowe, nigdy nie przedstawiające podejmowanego problemu w szerszym i krytycznym kontekście. Można dodać do tego kwestię fatalnego poziomu czytelnictwa – według najnowszych badań Biblioteki Narodowej z 2013 roku, żadnej książki w ciągu roku nie przeczytała nawet jedna trzecia osób z wyższym wykształceniem. Czy w tak scharakteryzowanej społeczno-medialnej przestrzeni krytyczna wiedza historyczna może być efektywnie komunikowana za pomocą wydawanych w niskich nakładach czasopism i publikacji? Czy szkolna edukacja historyczna daje kompetencje do krytycznego odbioru narracji o przeszłości, których źródłem jest popkultura, media i rynek?

Trudno uznawać Internet za cudowne lekarstwo na wymienione wyżej problemy. Po okresie zachłyśnięcia się pozytywną narracją o Sieci, rozwijającą się wokół nieco fantastycznej koncepcji Web 2.0, przychodzi czas na nowe krytyczne otwarcie. Należy jeszcze raz spojrzeć na Internet z perspektywy nauki historycznej i dostrzec niepokojące zjawiska: komercyjną kontrolę nad przestrzeniami społecznej komunikacji online (Twitter, Facebook), które stają się przedmiotem badań historycznych i społecznych, nadmiar dostępnych

informacji i ograniczone kompetencje techniczne jej użytkowników czy patologie na rynku czasopism naukowych, dotykające także czasopism Open Access. Problemy te nie powinny jednak przesłaniać potencjału Internetu w zakresie integrowania środowiska naukowego, polepszania komunikacji naukowej, redukcji kosztów dostępu do wiedzy i do źródeł.

Jednak myślenie o Internecie wyłącznie jako o przestrzeni i narzędziu dla *wewnętrznego* rozwoju nauki historycznej byłoby błędne i ograniczające. Historia jako nauka, na co wskazywał cytowany wyżej Jan Pomorski, nie rozwija się w społecznej próżni. Hiszpański historyk Fernando Sánchez-Marcos proponuje wprost:

[...] winniśmy [jako historycy] zaoferować współobywatelom naszą umiejętność przerzucania mostów zrozumienia. Dzięki naszej zawodowej działalności nawykliśmy do budowania takich mostów, łączących grupy ludzkie zlokalizowane na różnych poziomach czasowości. Nawykliśmy do podróży w czasie, które zwykle odbywa się w obrębie tej samej kultury. Tak więc możemy skorzystać z tych nawyków podróżnika i wspomóc zadanie budowy mostów intelektualnego zrozumienia i porozumienia, tym razem jednak pomiędzy społecznościami żyjącymi współcześnie, zlokalizowanymi i zakorzenionymi w kulturach, w sensach antropologicznych, które są bardzo odmienne, a czasem wręcz sobie przeciwstawne. Historycy są znakomicie przygotowani do takiej podróży, do przekraczania granic odrębnej tożsamości i ten багаż doświadczeń umożliwi ich współobywatelom przystosowanie się do życia w wielokulturowej cywilizacji (Sánchez-Marcos, 2002, s. 38-39).

Tak ambitnie określony społeczny program nauki historycznej realizowany być może także za pomocą Internetu, który redukuje ograniczenia w dostępie do wiedzy, umożliwia jej skuteczne komunikowanie z pominięciem tabloidyżujących się mediów, a nawet – w ramach nauki obywatelskiej – włączenie w proces jej tworzenia osób spoza środowiska akademickiego. Historyk, jak pisze Sánchez-Marcos, może (a nawet powinien) być tłumaczem, przenoszącym sensy z jednej kultury do drugiej między przeszłością a teraźniejszością i przekładać dawne i inne doświadczenia na zrozumiały dziś język. Zwraca też uwagę, że dzięki nowym technologiom nasze codzienne doświadczanie inności zintensyfikowało się (s. 41). Bezpośredni dostęp online do materiałów historycznych pozwala łatwiej niż kiedykolwiek spojrzeć na przeszłość z nowego



punktu widzenia. Zwłaszcza materiały wizualne, skany starych kodeksów czy fotografie historyczne przechowywane w internetowych archiwach zapraszają do kontaktu z nieistniejącym już światem i zachęcają do jego zrozumienia. Czy rolą historyków i systemu edukacyjnego nie powinno być przeciwstawianie się wyłącznie estetycznej recepcji tych materiałów, nadawanie im odpowiedniego kontekstu, rozwijanie empatii wobec ludzkich doświadczeń przez nie dokumentowanych, często wbrew *historii brązowniczej* (L. Gonzales), wspieranej przez politykę historyczną czy populistyczne media? Jeśli tak, wówczas ich obecność oraz obecność ich badań w Internecie, także w formie otwartych zasobów edukacyjnych, staje się niezbędna, chociaż na pewno niewystarczająca.

## BIBLIOGRAFIA

- Jenkins, H.** (2007). *Kultura konwergencji: Zderzenie starych i nowych mediów*. Warszawa: Wydawnictwa Akademickie i Profesjonalne.
- Pomorski, J.** (2011). Wybieram czerwoną pigułkę – czyli historia i Matrix. W: Witek, P., Mazur, M., Solska, E. (red.), *Historia w kulturze współczesnej. Niekonwencjonalne podejścia do przeszłości* (s. 72-85), Lublin: Edytor.org.
- Sánchez-Marcos F.** (2002). Historyk jako tłumacz. W: Domańska, E. (red.). *Pamięć, etyka i historia: Anglo-amerykańska teoria historiografii lat dziewięćdziesiątych; antologia przekładów*. Poznań: Wydawnictwo Poznańskie.
- Szpociński, A.** (2010). Współczesna kultura historyczna, *Kultura Współczesna*, 1(63), 9-17.

## Słownik pojęć używanych w publikacji

**API** – w perspektywie repozytoriów cyfrowych zestaw protokołów i narzędzi programistycznych pozwalających na korzystanie z ich zasobów przez zewnętrzne aplikacje. Przykładowo, Europeana udostępnia API pozwalające na tworzenie aplikacji komórkowych wykorzystujących wizerunki obiektów i metadane publikowane w tej bibliotece.

**archiwum społeczne** – archiwum lub działalność archiwalna prowadzona przez organizację pozarządową, jednostkę samorządu terytorialnego lub grupę inicjatywną czy osobę prywatną. Zbiory gromadzone w ten sposób zasilają niepaństwowy zasób archiwalny.

**Creative Commons (CC)** – amerykańska organizacja pozarządowa oraz rozwijany przez nią system licencji prawnoautorskich, które odwracają standardową zasadę „wszystkie prawa zastrzeżone” na rzecz „pewnych praw zastrzeżonych”. Stosując licencje CC, autor może łatwo zdefiniować zasady wykorzystania swoich prac, a ich użytkownicy uzyskują czytelną informację na ten temat. Najbardziej liberalne wersje licencji CC przyjęte zostały do rozpowszechniania treści Wikipedii oraz jako podstawa tworzenia i rozpowszechniania Otwartych Zasobów Edukacyjnych.

**historia cyfrowa, historia digitalna** (*digital history*) – wykorzystywanie cyfrowych narzędzi i zasobów w praktyce badań historycznych oraz działalności instytucji dziedzictwa. Także: nowy kierunek lub osobna specjalizacja w nauce historycznej.

**humanistyka cyfrowa** (*digital humanities*) – szeroka kategoria problemów, narzędzi i praktyk badawczych w ramach badań humanistycznych, dotyczących lub wykorzystujących narzędzia i zasoby cyfrowe.

**dziedzictwo cyfrowe** – zasoby o charakterze kulturowym, edukacyjnym, naukowym i administracyjnym, jak również informacje o charakterze technicznym, prawnym, medycznym i inne rodzaje informacji tworzone w postaci cyfrowej lub na tę postać konwertowane z dostępnych źródeł analogowych. W przypadku zasobów wytworzonych cyfrowo nie występują one w postaci innej niż obiekty cyfrowe. (*Karta w sprawie zachowania dziedzictwa cyfrowego*, UNESCO)

**digitalizacja, cyfryzacja** – odwzorowanie obiektów analogowych w postaci cyfrowej zapisanej np. za pomocą skanowania lub fotografowania. Na zmianę analogowego oryginału na jego cyfrowy odpowiednik używa się

także pojęć takich jak *reformatting* czy *digital reformatting*. Z perspektywy technicznej digitalizacja oznacza konwersję sygnału ciągłego na dyskretny poprzez próbkowanie, np. w procesie skanowania fotografii podzielenie jej treści na kwadraty o określonej wartości koloru (piksele).

**DOI** (ang. *digital object identifier*) – przypisany do dowolnego dokumentu elektronicznego cyfrowy identyfikator niezależny od miejsca zapisania pliku. Znając numer DOI cyfrowej wersji publikacji naukowej, posługując się menadżerem bibliografii można automatycznie pobrać informacje na jego temat (metadane).

**domena publiczna** – zbiór utworów, do których wygasły autorskie prawa majątkowe lub które nigdy nie były objęte ochroną prawnoautorską. Utwory dostępne w domenie publicznej mogą być bez ograniczeń publikowane w bibliotekach i archiwach cyfrowych.

**EdgeRank** – algorytm zarządzający wyświetlaniem informacji w serwisie Facebook w zależności od preferencji użytkownika definiowanych automatycznie m.in. na bazie jego dotychczasowych interakcji z publikowanymi treściami.

**Europeana** – europejska biblioteka cyfrowa agregująca i publikująca informacje (metadane) o zbiorach cyfrowych z europejskich archiwów, bibliotek, muzeów i galerii. Obecnie Europeana gromadzi informacje o 30 mln obiektów cyfrowych z ponad 2 tys. instytucji, w tym także instytucji z Polski; <http://europeana.eu/>.

**Federacja Bibliotek Cyfrowych (FBC)** – wspólny interfejs przeszukiwania zbiorów polskich bibliotek, archiwów i repozytoriów cyfrowych, gwarantujący dziś dostęp do ponad 1.6 mln obiektów ze 109 instytucji; <http://fbc.pionier.net.pl/owoc>.

**GLAM** (Galleries, Libraries, Archives, Museums) – akronim oznaczający instytucje kultury i dziedzictwa – galerie, biblioteki, archiwa i muzea.

**GLAM-Wiki** – kierunek współpracy między instytucjami dziedzictwa a społecznością Wikipedii, w ramach którego realizowane są projekty zwiększające dostępność zbiorów instytucji GLAM w zasobach Wikipedii.

**Google Książki** (*Google Books*) – usługa Google pozwalająca na przeszukiwanie i w niektórych wypadkach pełny lub ograniczony dostęp do treści zdigitalizowanych książek (w kwietniu 2013 było to około 30 mln tytułów).

**Google Scholar** – specjalistyczna wyszukiwarka prac naukowych oraz system cytowań.

- historia publiczna** (*public history*) – praktyka historyczna podejmowana poza lub na granicach instytucji naukowych i instytucji dziedzictwa.
- historia ratownicza** – projekt praktyki historycznej skoncentrowanej na zabezpieczeniu i opisywaniu i upowszechnianiu najbardziej zagrożonych fragmentów przeszłości, leżących poza głównym nurtem badań akademickich czy zainteresowaniem mediów.
- licencja** – forma umowy prawnoautorskiej, w ramach której twórca lub wydawca (licencjodawca) przekazuje część uprawnień do korzystania ze swojego utworu innym osobom lub podmiotom. Przykładem licencji są rozwiązania z katalogu Creative Commons.
- media społecznościowe** (ang. *social media*) – bardzo ogólna kategoria serwisów i usług internetowych, których główną cechą jest umożliwianie natychmiastowej i bezpośredniej komunikacji oraz łatwej wymiany dowolnych treści między użytkownikami. Kategoria ta w zależności od wybranej definicji obejmować może blogi, serwisy takie jak Facebook, Twitter czy Google+, a nawet Wikipedię i często przyjmuje bardzo ideologiczny i marketingowy charakter. Media społecznościowe są szeroko wykorzystywane w komunikacji naukowej, promocji nauki oraz działalności instytucji dziedzictwa.
- metadane** – dane opisujące zasoby lub inne dane; w przypadku bibliotek i archiwów cyfrowych metadane to tworzone według określonych standardów opisy udostępnianych publikacji i dokumentów.
- menedżer bibliografii** – program komputerowy pozwalający na łatwe gromadzenie i zarządzanie bibliografią, wspierający proces cytowania w ramach pracy nad tekstem naukowym.
- mikrohistoria** – kierunek w historiografii koncentrujący się na badaniu i opisywaniu historii niewielkich środowisk czy grup społecznych w bardzo ograniczonym zakresie czasowym, zainteresowany przedstawianiem ich życia codziennego, obyczajów, mentalności itp., wspomagany metodami badań kulturoznawczych czy antropologicznych.
- Mendeley** – darmowe zamknięte oprogramowanie pełniące funkcję menadżera bibliografii, pozwalające na sprawne zarządzanie przechowywanymi w postaci cyfrowej publikacjami naukowymi (PDF), przeszukiwanie ich treści i generowanie cytatów w wybranych stylach. Internetowe rozszerzenie Mendeley Web pozwala na tworzenie badawczych grup tematycznych, wymienianie się informacją o czytanych publikacjach i dyskusję na ich temat.
- nauka obywatelska** (*citizen science*) – praktyka badań naukowych, do której

włączane są osoby spoza środowiska akademickiego, posiadająca dość długą tradycję (np. akcje liczenia ptaków organizowane już 1900 roku przez Audubon Society). Nauka obywatelska to także forma edukacji naukowej i promocji nauki.

**PageRank** – algorytm zarządzający wyświetlaniem wyników wyszukiwania w wyszukiwarce Google.

**P2P** (*Peer to Peer*) – model komunikacji w sieci komputerowej, którego wszystkie elementy mają takie same prawa; komunikacja między komputerami użytkowników Internetu z pominięciem lub ograniczeniem roli centralnych serwerów np. w ramach protokołu BitTorrent pozwalającego na efektywną wymianę w Internecie nawet dużych plików; model P2P stosuje się także do opisu rozproszonych form współpracy i produkcji bez jednego centralnego ośrodka zarządzania czy dystrybucji (*peer production*).

**OAI-PMH** (*Open Archives Initiative Protocol for Metadata Harvesting*) – protokół wymiany metadanych w standardzie Dublin Core między repozytoriami cyfrowymi, rozwijany przez społeczność Open Archive Initiative. Z OAI-PMH korzysta polska Federacja Bibliotek Cyfrowych oraz Europeana.

**OCR** (*Optical Character Recognition*) – metoda pozwalająca na automatyczne rozpoznawanie znaków odczytywanych ze skanów dokumentów w formie graficznej, wykorzystywana szeroko w procesie digitalizacji; także określenie kategorii oprogramowania wykorzystywanego do tych zadań.

**OPAC** (*Online Public Access Catalog*) - internetowy i dostępny publicznie katalog zbiorów biblioteki (bibliotek), udostępniający informacje bibliograficzne o dostępnych pozycjach, czasem także pozwalający na ich zamawianie czy recenzowanie.

**Open Access** (*Otwarty dostęp*) – model publikowania prac naukowych w Internecie, pozwalający na darmowy, powszechny, trwały i natychmiastowy dostęp do ich treści; także rozwijający się od lat 80. ruch na rzecz wolnego dostępu do tych zasobów (Open Access Movement).

**OpenGLAM** – wypracowany przez społeczność wokół Open Knowledge Foundation zestaw zaleceń (standard) na rzecz pełnej otwartości instytucji kultury i dziedzictwa (galerii, bibliotek, archiwów i muzeów).

**open source** (*otwarte źródło*) – oprogramowanie udostępniane wraz z kodem źródłowym i prawem do jego modyfikacji oraz rozpowszechniania zmieniowanych wersji. Także: ruch na rzecz jego rozwijania; często stosowane wymiennie z pojęciem wolne oprogramowanie (*free software*).

**rzeczywistość rozszerzona** (ang. *augmented reality*) – rozwiązanie techniczne pozwalające w czasie rzeczywistym nakładać wygenerowane przez oprogramowanie obiekty na obraz rzeczywistej przestrzeni. Umożliwia to np. wyświetlanie komputerowych rekonstrukcji 3D zabytku na obrazie z kamery obejmującym rzeczywisty, ale zniszczony już obiekt dostępny w przestrzeni miejskiej.

**eksploracja tekstu** (*text mining*) – metody automatycznej komputerowej analizy obszernych zestawów tekstów, pozwalające na wyodrębnienie z nich wybranych elementów oraz ułatwiające interpretację.

**THATCamp** – (*The Humanities and Technology Camp*) cykl otwartych, oddolnych, nieformalnych i niskobudżetowych spotkań i niekonferencji (*unconference*) poświęconych zagadnieniom cyfrowej humanistyki. W Polsce spotkania w ramach inicjatywy THATCamp organizowane są od 2012 roku; <http://thatcamp.pl/>.

**transkrypcja** – zapis dźwięków danego języka (wypowiedzi głosowej) za pomocą symboli graficznych lub – szczególnie w perspektywie digitalizacji dokumentów historycznych – przepisanie treści dokumentów pozostających w postaci graficznej do pliku tekstowego; zob. też OCR. Transkrypcja łączyć się może z anotacją semantyczną wykorzystującą format XML.

**wiki** – model serwisu internetowego; serwisy internetowe działające w modelu wiki pozwalają użytkownikom w prosty sposób edytować treści poszczególnych stron bezpośrednio z poziomu przeglądarki, bez konieczności znajomości kodu HTML oraz w trybie wieloosobowym.

**Wikipedia** – wolna internetowa encyklopedia, którą współtworzyć może każdy. Treść haseł Wikipedii ustalana jest na zasadzie konsensusu przez poszczególnych edytorów ze zróżnicowanymi prawami do wprowadzania czy akceptowania zmian i na podstawie zasad wypracowanych przez społeczność tego projektu. Infrastrukturą techniczną Wikipedii zarządza Wikimedia Foundation, nie kontroluje ona jednak publikowanych haseł. Treść haseł Wikipedii dostępna jest na wolnej licencji (obecnie Creative Commons Uznanie Autorstwa – Na Tych Samych Warunkach).

**wolne licencje** – sposób licencjonowania praw autorskich dający szerokie możliwości wykorzystania i przerabiania licencjonowanych utworów (oprogramowania, fotografii, tekstów itp.). Idea wolnych licencji została stworzona przez Richarda Stallmana i pierwotnie odnosiła się do zasad tworzenia i dystrybucji wolnego oprogramowania (*free software*) za

pomocą licencji takich jak GNU GPL. Dziś do rozpowszechniania publikacji naukowych czy materiałów edukacyjnych wykorzystać można dwie wolne licencje z katalogu Creative Commons: Uznanie Autorstwa (CC BY) oraz Uznanie Autorstwa – Na Tych Samych Warunkach (CC BY-SA).

**WWW** (*World Wide Web*) – oparty o protokół http hipertekstowy multimedialny system prezentacji informacji w Internecie, bardzo często błędnie utożsamiany z całym Internetem. Zasoby WWW podlegają archiwizacji w ramach programów realizowanych przez niektóre biblioteki narodowe oraz fundację Internet Archive.

**XML** (*Extensible Markup Language*) – rozszerzalny język znaczników; język służący do reprezentowania danych w ustrukturyzowany sposób.

**Zotero** – darmowy menadżer bibliografii i zarządzania cytatami dostępny w postaci wtyczki do przeglądarki Firefox oraz samodzielnej aplikacji; zainstalowany w przeglądarce pozwala na łatwe importowanie metadanych ze stron internetowych, repozytoriów czy baz danych dostępnych online.

Historia cyfrowa to nowa i jedna z najbardziej dynamicznie rozwijających się subdyscyplin badań historycznych. Książka Marcina Wilkowskiego to kompetentne, przystępne i syntetyczne wprowadzenie do podstawowych jej zagadnień: digitalizacji i prezentacji online źródeł historycznych, nowych form publikowania naukowego i dostępu do literatury, nowoczesnej komunikacji naukowej. Książka przybliży także najważniejsze problemy wynikające z historycznego myślenia o Internecie jako przestrzeni społecznej. Jest to podręcznikowa publikacja, którą powinien przeczytać każdy zainteresowany badaniami przeszłości podejmowanymi w sposób otwierający je na przyszłość.

**Ewa Domańska**

Uniwersytet im. Adama Mickiewicza/Stanford University

Wprowadzenie do historii cyfrowej służyć może wszystkim badaczom, szczególnie nauk humanistycznych, do refleksji nad ich osobistym stosunkiem do nowych technologii i stanowi wartościową pomoc w pracy ze studentami.

**Rafał Wójcik**

Biblioteka Uniwersytecka Uniwersytetu im. Adama Mickiewicza

**Marcin Wilkowski** ([marcin@wilkowski.org](mailto:marcin@wilkowski.org)) - absolwent Instytutu Historii Uniwersytetu Gdańskiego, uczestnik programu doktorskiego Collegium Civitas w Warszawie. Twórca portalu Historia i Media ([historiaimedia.org](http://historiaimedia.org)). Pracuje w Fundacji Nowoczesna Polska, koordynując działania Koalicji Otwartej Edukacji (KOED). Inicjator spotkań i warsztatów w ramach polskiej edycji THATCamp (The Humanities and Technology Camp). Współpracuje z Ośrodkiem Karta, ruchem społecznym Obywatele Nauki oraz Instytutem Kultury Miejskiej w Gdańsku.

**edu**kacja

ISBN 978-83-938408-6-1