

Krzysztof Narojczyk  
(Uniwersytet Warmińsko-Mazurski)

## W kierunku historii cyfrowej. Nowe możliwości – nowe wyzwania

*Towards Digital History. New Possibilities – New Challenges*

### STRESZCZENIE

Wraz z rozwojem technologii informatycznych i telekomunikacyjnych wykształca się w wielu obszarach historii nowa interdyscyplinarna forma uprawiania badań, określana zbiorczo jako „historia cyfrowa”. Przedmiotem jej zainteresowania są problemy związane z cyfrowym przetwarzaniem informacji „z przeszłości” i „o przeszłości”. Stanowi ona zatem swoisty łącznik pomiędzy specjalistycznym zakresem badań historycznych a współczesną teorią i praktyką operowania informacją. Z jednej strony, historia cyfrowa praktycznie wspiera poszczególne etapy postępowania badawczego historyka, z drugiej zaś, z uwagi na możliwości integrowania różnych środowisk i pól badawczych, oddziałuje na organizację badań oraz teorię, metody i narzędzia poszczególnych dyscyplin „historycznych”. Pełni więc dziś zarówno funkcje praktyczne, jak i teoretyczne. Z jej pomocą można badać i rozwiązywać pragmatyczne oraz/lub koncepcyjne problemy związane z wykorzystaniem nauk o informacji i technologiach informatycznych w nauczaniu i badaniach historycznych oraz społecznym transferowaniu wiedzy o przeszłości. W artykule, oprócz omówienia kwestii terminologicznych i problemów związanych z przedmiotem badań historii cyfrowej, zasygnalizowano także jej najbardziej obiecujące metody badawcze oraz główne obszary zastosowań we współczesnej nauce historycznej. Dają one szansę na stosowanie wydajniejszych, bardziej wyrafinowanych i innowacyjnych praktyk badawczych. Wymagają jednakże zarazem zrozumienia istoty przemian, jakie dokonują się we współczesnym cyfrowym świecie.

**Słowa kluczowe:** historia cyfrowa, informatyka humanistyczna, nauki pomocnicze historii, informatologia, bazy danych, edycja źródłowa

Współcześnie, w dobie gwałtownego rozwoju technologii komputerowych, w tym zwłaszcza internetu, trwa stopniowy proces włączania osiągnięć informatyki do warsztatu badawczego historyka. W efekcie, w wielu

obszarach historii wykształciła się nowa, interdyscyplinarna forma uprawiania badań, określana zbiorczo jako „historia cyfrowa”. Początkowo była ona traktowana w kategoriach wyłącznie warsztatowych, jednak z czasem jej status zaczął się zmieniać i dzisiaj jest już czymś znacznie więcej niż tylko zaawansowanym narzędziem badawczym. Aspiruje coraz wyraźniej do miana samodzielnej subdyscypliny teoretycznej i praktycznej w historii.

### GENEZA HISTORII CYFROWEJ

W miarę postępu w rozwoju technologii komputerowej, a przede wszystkim dzięki jej upowszechnianiu w środowisku akademickim, w nauce zaczęło funkcjonować pojęcie „badanie wspomagane komputerowo” czy „analiza wspomagana komputerowo”. Dotyczyło to także dyscyplin humanistycznych. Specyfika badawcza nauk społecznych sprawiła, że stopniowo wyodrębniła się nowa subdyscyplina, określana początkowo jako „informatyka humanistyczna” (*e-humanities* lub *humanities computing*), zajmująca się, najogólniej rzecz ujmując, badaniem relacji zachodzących pomiędzy ludzką wiedzą i kulturą a technologią informatyczną. Wywodzi się ona zarówno z tradycji humanistycznej powiązanej z logiką, lingwistyką i filozofią, jak i z tradycji klasycznych nauk komputerowych. Owa interdyscyplinarność implikuje obalenie tradycyjnych paradygmatów humanistyki, jeżeli nie są one już w stanie w dobie komputerów poprawnie objaśniać duchowego świata człowieka<sup>1</sup>. Ważnym impulsem w rozwoju informatyki humanistycznej było pojawienie się na rynku w 1966 r. znakomitego, bardzo wysoko ocenianego przez profesjonalistów czasopisma naukowego „The Journal Computers and the Humanities” (od marca 2005 r. przemianowane na „Language Resources and Evaluation”). Kamieniem milowym okazał się także fakt powstania i szybkiego upowszechnienia globalnej sieci. Rozwijający się od połowy lat dziewięćdziesiątych XX w. internet stopniowo stał się centrum publicznej przestrzeni komunikacyjnej, integrując w swym obrębie większość, dotychczas analogowych, form wymiany informacji w postaci specjalizowanych aplikacji sieciowych (listy dyskusyjne, chat, e-mail itp.). Równocześnie pojawiły się specyficzne dla internetu formy wspierane przez technologię WWW (komunikatory wideo, blogi, wikis – czyli projekty kolektywnego gromadzenia i dystrybucji informacji wzorowane na modelu Wikipedii itp.). Wszystko to sprawia, iż zakres wzajemnego przenikania się

<sup>1</sup> I. Lytje, A. Bygholm, *Artificial Intelligence and Humanistic Informatics*, „Lecture Notes in Computer Science” 1990, 451, s. 114.

i interakcji świata człowieka oraz świata technologii informatycznych i telekomunikacyjnych ciągle wzrasta. Realne okazało się nawiązanie do ideałów humanistycznych w zakresie otwartości i zróżnicowania kierunków badań, dzielenia się wiedzą, współpracy czy interdyscyplinarności. W ślad za tym w przestrzeni publicznej pojawiło się szersze, pojemniejsze i bardziej elastyczne pojęcie na określenie tych procesów – „cyfrowa humanistyka” (*digital humanities*). W zamyśle Susan Schreibman, Johna Unswortha i Raya Siemensa<sup>2</sup>, którym przypisuje się stworzenie tego terminu, miał on wskazywać na praktykę stosowania humanistycznego podejścia w studiowaniu obiektów cyfrowych oraz cyfrowych metod i narzędzi w studiowaniu treści humanistycznych. Jako dyscyplina naukowa, lub jak chcą jej krytycy – jako nurt w nauce, cyfrowa humanistyka jest dziś uprawiana w wielu ośrodkach akademickich. Ma też już własną literaturę i periodyki naukowe.

Wspomaganie komputerowe badań dotyczyło także historii. Najwcześniej, bo już w latach sześćdziesiątych, zaczęto je stosować w Stanach Zjednoczonych. Ich motorem napędowym był głównie entuzjazm dla historii społecznej i gospodarczej, jaki ogarnął znaczną część środowiska historyków amerykańskich. Badania z tego zakresu opierały się na tradycyjnych metodach i technikach nauki historycznej, wzbogacanych o elementy teoretyczne i metodologiczne zaczerpnięte z socjologii. Koncentrowały się one głównie na kwantytatywnych aspektach dziejów społecznych i ekonomicznych. Technologia komputerowa, która właśnie w skali powszechnie dostępnej dla świata nauki zaczęła pojawiać się na rynku, wydawała się cudownym, uniwersalnym narzędziem do tego typu badań i obliczeń. Z jej pomocą historycy podejmowali i pomyślnie rozwiązywali problemy badawcze, które dotychczas wydawały się niemożliwe do zrealizowania. Pojawiały się wręcz głosy o rychłym schyłku tradycyjnej historii narracyjnej i jej abdykacji na rzecz kliometrii. Wkrótce okazało się, że ten entuzjazm był przedwczesny. Punktem zwrotnym było opublikowanie w 1974 r. głośnej pracy Richarda Williama Fogela i Stanleya Engermana *Time on the Cross: The Economics of American Negro Slavery*<sup>3</sup> oraz towarzysząca tej publikacji ożywiona dyskusja. Sztandarowe, wydawałoby się, osiągnięcie kliometrii spotkało się z ostrą krytyką wielu „tradycyjnych” historyków<sup>4</sup>. Autorom zarzucano przede wszystkim, że koncentrując się

<sup>2</sup> *A Companion to Digital Humanities*, eds. S. Schreibman, R. Siemens, J. Unsworth, Oxford 2004.

<sup>3</sup> R.W. Fogel, S. Engerman, *Time on the Cross: The Economics of American Negro Slavery*, London 1974.

<sup>4</sup> Zob. np. H.G. Gutman, *Slavery and the Numbers Game: A Critique of Time on the Cross*, Urbana 1975.

na czysto ekonomicznych aspektach tego systemu organizacji produkcji, takich jak zyskowność czy względna produktywność gospodarstw, pominieli zupełnie społeczny i polityczny kontekst niewolnictwa. Obliczyli np., że przeciętnie niewolnik był chłostany 0,7 raza rocznie, co jednakowoż nie przybliżało zrozumienia roli publicznie wymierzanych kar cielesnych jako instrumentu utrzymywania dyscypliny wśród niewolników. Okazało się zatem, że wielu najważniejszych z punktu widzenia nauki historycznej kwestii nie da się rozwiązać wyłącznie metodami kwantytatywnymi. Bardzo wysoki był także koszt stosowania nowych metod, bowiem oparte na nich projekty wymagały najczęściej dużych zespołów pracowników do opracowywania wstępnego i wprowadzania danych. Także sama technologia informatyczna była początkowo stosunkowo kosztowna i nie wszystkie ośrodki badawcze mogły z niej korzystać. Na szczęście ta przeszkoda została szybko zniwelowana przez radykalne obniżanie cen produktów komputerowych oferowanych na rynku indywidualnemu odbiorcy. Komputery trafiły wprawdzie w skali masowej „pod strzechy”, jednak absorpcja metod komputerowych przez historyków była procesem powolnym, wymagającym zarówno „oswojenia się” z nową technologią, jak i przełamania wielu nawyków, uprzedzeń i stereotypów.

W Europie metody obliczeniowe nie znalazły początkowo uznania. Komputery zaczęto natomiast wykorzystywać do edytorstwa źródłowego. W tym kontekście krokiem milowym, ukazującym potencjał technologii komputerowej, była elektroniczna edycja konkordancji dzieł wszystkich św. Tomasza z Akwinu<sup>5</sup>.

Nauka historyczna, jako istotna część składowa humanistyki, także podlegała podobnym procesom, starając się absorbować i dostosowywać technologie komputerowe do potrzeb własnego warsztatu badawczego. Wykształcająca się w toku tych usiłowań subdyscyplina – informatyka historyczna – była rozmaicie sytuowana w obrębie warsztatu naukowego historyka. W USA, Wielkiej Brytanii i Skandynawii dominował pogląd, że stanowi ona część nowej, wspomnianej wyżej, dynamicznie rozwijającej się dyscypliny naukowej – informatyki humanistycznej (*Humanities Computing*) i jako taka miała stanowić naukę posiłkową historii, zaś np. w Niemczech traktowano ją raczej jako naukę pomocniczą historii. Roz-

---

<sup>5</sup> Jezuita Roberto Busca w 1947 r. rozpoczął projekt wydania konkordancji dzieł wszystkich św. Tomasza z Akwinu. Jednak skala przedsięwzięcia (ponad 70 tys. stron, 1,5 mln wierszy, 9 mln słów) przerosła możliwości dokonania tego metodami tradycyjnymi. Dopiero wykorzystanie komputerów firmy IBM pozwoliło opublikować do 1970 r. łącznie 56 encyklopedycznych tomów *Index Thomisticus*. W 1989 r. ukazała się wersja na CD-ROM zaś w 2005 r. udostępniono indeks *on-line* <<http://www.corpusthomaticum.org/it/index.age>>.

różnienie to jest być może pochodną odmiennego usytuowania samej nauki historycznej w anglosaskiej tradycji dyscyplin akademickich (wyodrębnienie humanistyki i sztuki – *Arts and Humanities* z zakresu nauki – *Science*, w tym także z zakresu nauk społecznych – *Social Sciences*).

Także nomenklatura stosowana na świecie nie jest, niestety, jednoznaczna. Używane określenia zależą od treści (idei), jaką wiążą z tymi terminami ich twórcy oraz po części od właściwości słowotwórczych konkretnych języków.

### WOKÓŁ POJĘCIA – PROBLEMY TERMINOLOGICZNE

Termin „informatyka” pojawił się po raz pierwszy w 1957 r. w tytule pracy niemieckiego cybernetyka Karla Steinbucha, poświęconej automatycznemu przetwarzaniu informacji<sup>6</sup>. Łączył on w sobie dwa pojęcia: „informacja” i „automatyka”, znakomicie oddając istotę nowych propozycji w operowaniu informacją. Termin ten, a raczej technologia jego tworzenia szybko została zaimplementowana w wielu językach. W 1962 r. pojawiło się francuskie *informatique* (*Société d’Informatique Appliquée*), rok później rosyjskie *информатика*. Termin ten w formie *informatica* używany jest także np. w języku holenderskim, rumuńskim, włoskim, hiszpańskim, portugalskim i wielu innych. Współcześnie pojęcie „informatyka” występuje w większości języków świata, odnosząc się wszakże zasadniczo do komputerowego przetwarzania informacji. W języku angielskim *informatics* jest wręcz tożsame z określeniem *computer science*.

Wraz z coraz szerszym, a nie incydentalnym, wykorzystywaniem komputerów do badań historycznych pojawił się problem wypracowania jednoznacznego i adekwatnego terminu dla tego rodzaju aktywności naukowej. W 1977 r. historycy z kręgu *Annales* wydali pracę<sup>7</sup>, w której połączyli w jedną zbitkę pojęciową informatykę i historię – *l’histoire et l’informatique* (rok później na fali obrony języka francuskiego termin *l’informatique* zastąpiono rdzennie francuskim określeniem *les ordinateurs*<sup>8</sup>, co nie zmienia wszakże logiki konstrukcji tego związku frazeologicznego). Angielska wersja terminu – *history and computing* szybko rozprzestrzeniła się w świecie, po części za sprawą powołanej w 1987 r. w Londynie międzynarodowej organizacji *Association for History and Computing*

<sup>6</sup> K. Steinbuch, *Informatik: Automatische Informationsverarbeitung*, „SEG-Nachrichten” 1957, 4.

<sup>7</sup> L. Fossier, A. Vauchez, V. Cinzio, *Informatique et histoire médiévale*, Rome 1977.

<sup>8</sup> Por. *L’Histoire médiévale et les ordinateurs: Medieval history and computers: rapports d’une table ronde internationale*, ed. K. Ferdinand Werner, Paris 1978.

i powstawania jej licznych oddziałów narodowych. Samo określenie, z punktu widzenia precyzji językowej, jest niezbyt fortunate, a na pewno nieostre. Równorzędnie, niejako neutralnie, traktuje obydwie części wchodzące w jego skład. Wskazuje to wprawdzie, że przedmiotem zainteresowania nowej dyscypliny jest styk obu pól badawczych, tj. historii i przetwarzania komputerowego, nie określa jednak jednoznacznie, które z tych pól jest służebne, a które wiodące. Brzmieniowo termin jest bardzo podobny do powszechnie używanego określenia *history of computing*, co wśród laików może prowadzić do pomyłek. *History and computing* nie był jedynym terminem stosowanym dla określenia komputerowego wspomaganie badań historycznych. Równolegle w użyciu pozostawały terminy: *historical computing* propagowany np. przez portal Virtual Library<sup>9</sup>, *historical information processing* czy *Historical Informatics*. Tego ostatniego określenia – *историческая информатика* – konsekwentnie od połowy lat osiemdziesiątych używali np. Rosjanie. Początkowo odnoszono je bardziej do technicznych aspektów stosowania technologii komputerowej, z czasem przenosząc jednak akcent na zagadnienia metodologiczne. Z kolei najbardziej precyzyjni w terminologii okazali się Niemcy, którzy wykorzystali właściwość swojego języka pozwalającą, od ręki niejako, stworzyć nowe rzeczowniki i powołali do życia pojęcie *Historische Fachinformatik*. Manfred Thaller – guru pierwszego pokolenia historyków informatyków, twórca legendarnego programu Kleo – zwrócił uwagę na fakt, iż termin ten odnosi się przede wszystkim do komputerowo wspomaganym prac nad źródłami historycznymi, a więc akcentuje samą istotę postępowania badawczego historyka, które wspierane jest przez technologię komputerową. Rozwiązanie niemieckie powielili także praktyczni Holendrzy, którzy obok terminu *geschiedenis en informatica* stosują określenie *Historische Informatiekunde*.

Problematyka wykorzystania technologii komputerowej w badaniach historycznych nieobca była także środowisku uczonych polskich. Po fazie wstępnych dyskusji i ustaleń na spotkaniu zainteresowanych historyków w Krakowie w 1993 r. powołano Polski Oddział Association for History and Computing. Wkrótce po tym przy Zarządzie Głównym PTH powołano Komisję Metod Komputerowych. Pod auspicjami tych instytucji zaczęto organizować ogólnopolskie sympozja: „Metody komputerowe w badaniach i nauczaniu historii”, które stanowiły podstawową platformę wymiany poglądów i dyskusji środowiska polskich historyków stosujących w pracy badawczej metody komputerowe. Rozpoczęto wydawanie,

<sup>9</sup> WWW Virtual Library, <<http://www.vl-ghw.uni-muenchen.de/hwEnglish.html#sect31>> [dostęp: 15 XII 2014].

redagowanej przez Bohdana Ryszewskiego, serii „Historia i Komputery”. Doceniając tę aktywność Zarząd Główny Stowarzyszenia powierzył Polsce organizację w 2001 r. XV Międzynarodowej Konferencji AHC. Odbyła się ona w dniach od 28 do 31 sierpnia w Poznaniu pod hasłem *New Methodologies for the New Millennium*.

Od początku istniał poważny problem, jak w języku polskim adekwatnie określać nowo powstałą sferę aktywności badawczej. Termin *history and computing* okazał się nieprzetłumaczalny z uwagi na brak w naszym języku odpowiednika angielskiego czasownika *computing*. W tym wypadku nie nastąpiło obserwowane niekiedy wykształcenie nowego czasownika odrzeczownikowego, jak np. *kamerowanie* od rzeczownika *kamera* (termin niechętnie widziany przez purystów językowych, ale pozostający w szerokim obiegu). Stosowanie rzeczownika *komputer* lub w liczbie mnogiej *komputery* (zob. wspomniana wyżej seria wydawnicza) sugerowało wprawdzie jakiś związek historii z techniką komputerową, nie oddawało jednak zupełnie istoty tego związku. Prawdopodobnie podobne problemy językowe mieli Rosjanie i dlatego zaimplementowali termin „informatyka historyczna”. Podobnymi ścieżkami podażyły także procesy językowe w Polsce. Jakiś czas w użyciu pozostawało opisowe określenie – „metody komputerowe” (np. Komisja Metod Komputerowych ZG PTH czy Zakłady Metod Komputerowych w składzie Instytutów Historii). Wąski, techniczno-utilitytarny zakres tego pojęcia rychło przestał wystarczać. Dlatego powoli, nie bez oporów, zaczęło się przebić stosowanie terminu „informatyka historyczna”. Oponenci podnosili wprawdzie problem niejednoznaczności tego związku frazeologicznego, bowiem może on oznaczać zarówno wykorzystanie informatyki w badaniach historycznych, jak i informatykę ogólną, ujmowaną w jej rozwoju dziejowym. Jednak w środowisku historyków, dylematów, jak rozumieć to pojęcie raczej nie było i funkcjonowało ono w obiegu wzorem wielu dyscyplin, które w nazwie mają przymiotnik *historyczna* (geografia, statystyka, demografia).

W latach dziewięćdziesiątych ub. wieku postępował proces ewolucji znaczenia terminu *Historical Informatics* w kierunku poszerzania jego zakresu pojęciowego. Obok tradycyjnego, kwantytatywnego przetwarzania danych historycznych za pomocą komputera włączono doń także różnorodne metody ich analizy i oceny jakościowej. Rozwój dyscypliny zmierzał w kierunku wypracowywania podstaw teoretycznych dla całościowego, komputerowego wspomagania procesu rekonstrukcji dziejów człowieka i społeczeństwa. Od mniej więcej początku obecnego wieku na świecie stopniowo weszło do obiegu, zwłaszcza w USA, pojęcie *Historical Information Science*. Termin ten sugeruje, że mamy już do czynienia z pew-

nym działem nauki o informacji (informatologii), ukierunkowanym na informacje przydatne w badaniach nad przeszłością (historyczne). Zaproponowany został przez Lawrence'a McCranka w 1995 r. na łamach „Annual Review of Information Science and Technology”<sup>10</sup>. Uważał on, że lepiej oddaje istotę nowej formy aktywności naukowej, akcentując zwłaszcza wpływ zastosowania technologii informacyjnej na wszystkie elementy bezpośrednio lub pośrednio wpływające na tok postępowania badawczego historyka i jego otoczenie, do których zaliczył m.in.: proces dziejowy, pamięć, strukturę organizacyjną instytucji i obieg informacji w jej obrębie, kwerendy archiwalne i biblioteczne, badania bibliometryczne literatury, nauki archiwalne ze szczególnym uwzględnieniem archiwów elektronicznych, socjoekonomiczne metody badawcze, bezpośrednie wykorzystanie komputerów w badaniach, rozwój kooperacji naukowej i sprawne zarządzanie projektami o dużej skali złożoności, integrację tekstu, obrazu i dźwięku, systemy bazodanowe, wykorzystanie interaktywnych mediów w symulacjach i wizualizacjach, problemy z zapewnieniem trwałości danych elektronicznych, edukację historyczną społeczeństwa i wykształcenie warsztatowe historyków<sup>11</sup>. Takie widzenie zakresu pojęciowego *Historical Information Science* sytuuje je jako coś więcej niż tylko zbiór tradycyjnych intelektualnych, interdyscyplinarnych zapożyczeń z zakresu różnych dziedzin nauki i praktyki. Jest to raczej próba przełamania barier pomiędzy poszczególnymi dyscyplinami w kierunku nadania *Historical Information Science* charakteru multidyscyplinarnego.

Podobnie jak w przypadku informatyki humanistycznej, także dla informatyki historycznej ważnym momentem zwrotnym okazał się rozwój internetu. Technologia WWW potrafiła zintegrować większość dotychczasowych mediów przekazu treści. Internet sam w sobie jest niezwykle cennym źródłem historycznym, tym bardziej że wiele materiałów, zwłaszcza niemających odpowiedników analogowych (*born-digital*), dla historyków dostępnych jest obecnie, i zapewne będzie także w przyszłości, wyłącznie za pośrednictwem sieci. To rodzi pytania o to, co charakteryzuje materiał dostępny w Web i jak ta charakterystyka wpływa na jego wykorzystywanie przez historyka oraz jakie są możliwości i ograniczenia stosowania cyfrowych narzędzi analitycznych w stosunku do tego typu materiału.

Internet jest w ciągłym dynamicznym rozwoju. Badacz, który będzie chciał w przyszłości rozumieć nasze czasy, będzie musiał rozumieć in-

---

<sup>10</sup> L.J. McCrank, *History, Archives, and Information Science*, „Annual Review of Information Science and Technology” (ARIST) 1995, 30, s. 281–382.

<sup>11</sup> Idem, *Historical Information Science: History in Information Science; Information Science in History*, w: *History under Debate. International Reflection on the Discipline*, eds. C. Baros, L.J. McCrank, New York 2004, s. 177–198.

ternet. Zatem niezależnie od nastawienia do nowych technologii historycy nie mogą już ignorować faktu, iż coraz częściej źródła, od których zależy postęp w ich pracy, mają postać cyfrową i dostępne są zdalnie. Lawinowy wzrost liczby dostępnych źródeł cyfrowych o rozmaitej proweniencji – od monografii naukowych do wymiany poglądów w sieciach społecznościowych – jest jednym z istotnych argumentów przemawiających za implementacją metod cyfrowych w badaniach historycznych. Zdecydowana większość powstających obecnie informacji ma postać cyfrową i wobec mnogości rozmaitych ścieżek dotarcia do nich łatwo jest istotne naukowo dane przeoczyć. Z drugiej strony, brak umiejętności selekcji przydatnych w badaniach informacji szybko prowadzi do szumu informacyjnego i przekroczenia możliwości percepcyjnych człowieka<sup>12</sup>. Dodatkowo, dzięki digitalizacji unikatowych materiałów analogowych (*reborn-digital*), dziś dostęp do nich ma praktycznie każdy zainteresowany, a nie jak dotąd nieliczni badacze. Już w 2003 r. Roy Rosenzweig określił, iż następuje w historii swoista zmiana paradygmatu i przejście od kultury niedoboru do kultury obfitości źródeł<sup>13</sup>. Na tym gruncie, poczynając od przełomu XX i XXI w., zaczęto powszechnie używać terminu „cyfrowa historia” (*digital history*), odnosząc się do praktyk wykorzystywania technologii komputerowej do kreowania, wzbogacania i prezentowania wiedzy o przeszłości społecznej człowieka (historycznej) zarówno w ujęciu *stricte* akademickim, jak i popularnonaukowym czy edukacyjnym. W poszczególnych krajach termin ten zastąpił określenia używane dotychczas. Także w Polsce został powszechnie zaakceptowany, czego dowodem jest opublikowanie w sieci podręcznika *Wprowadzenie do historii cyfrowej* autorstwa Marcina Wilkowskiego<sup>14</sup>. Jedynie Niemcy pozostali przy swoim *Historische Fachinformatik*, nadając jednak tej konstrukcji frazeologicznej nowe, poszerzone znaczenie.

## PRZEDMIOT BADAŃ

Informatyka jako samodzielna dyscyplina naukowa datuje się od lat sześćdziesiątych ub. wieku. Zajmuje się całokształtem zagadnień związanych z przechowywaniem, przesyłaniem, przetwarzaniem, klasyfikowaniem i interpretowaniem informacji. Czerpie z dorobku m.in. elektroni-

<sup>12</sup> W.J. Turkel, Shezan Muhammedi, Mary Beth Start, *Grounding Digital History in the History of Computing*, „IEEE Annals of the History of Computing” 2014, 36, 2, s. 72–75.

<sup>13</sup> R. Rosenzweig, *Scarcity or Abundance? Preserving the Past in a Digital Era*, „American Historical Review” 2003, 108, 3, s. 735–762.

<sup>14</sup> M. Wilkowski, *Wprowadzenie do historii cyfrowej*, Instytut Kultury Miejskiej, Gdańsk 2013, <<http://www.historiacyfrowa.ikm.gda.pl>> [dostęp: 12 XII 2014].

ki, nauki o informacji, bibliotekoznawstwa itd. Wiele dyscyplin z różnych dziedzin wiedzy włączyło elementy informatyki do kanonu swych nauk pomocniczych lub nauk posiłkujących. Niezależnie od stopnia absorpcji metod komputerowych, w dzisiejszym świecie niewiele już pozostało obszarów naukowych, które mogłyby ignorować informatykę bez szkody dla swego rozwoju. W każdym jednak wypadku jej rola i miejsce w całości kształcie postępowania badawczego mogą być różne. Różnie też bywa definiowany obszar badawczy.

Trzymając się sztywno pierwotnego kanonu Nauk Pomocniczych Historii [dalej: NPH], określającego, iż są to nauki „dające poznawać źródła”, a więc związane z etapem krytyki źródłowej i ustalaniem wartości informacyjnej źródeł, trzeba by stwierdzić, iż historia cyfrowa znakomicie się w ów kanon wpisuje. Rozwój technologii komputerowej spowodował, że pojawił się nowy, zupełnie nieznan wcześniej, sposób kodowego zapisu, przechowywania i dystrybuowania informacji w postaci cyfrowej. W ślad za tym w obiegu, na coraz szerszą skalę, znalazł się nowy typ dokumentu – dokument elektroniczny. Unikalność jego cech fizycznych i często funkcjonalnych<sup>15</sup> wymaga stosowania całkowicie nowych metod krytyki zewnętrznej. Także dotychczasowy katalog narzędzi krytyki wewnętrznej w wielu przypadkach okazuje się niewystarczający. Z tego punktu widzenia sprawa wydaje się prosta – pojawił się nowy, odmienny od dotychczasowych, rodzaj źródła – konieczne jest wyodrębnienie na gruncie NPH nowej metody badawczej. Bez znajomości jej podstaw historyk nie będzie w stanie skutecznie weryfikować źródeł elektronicznych i efektywnie odczytywać treści, które za sobą niosą. Jednak zawężanie obszaru zainteresowań historii cyfrowej do krytyki źródłowej nie oddaje w pełni jej potencjału i znaczenia w procesie poznawania minionej rzeczywistości. Historia jako nauka zajmuje się transferem informacji z przeszłości do terażniejszości. Istota poznania historycznego polega na wpisaniu informacji uzyskanych o przeszłości z rozmaitych źródeł i różnymi metodami we współczesną logiczną rekonstrukcję minionych wydarzeń i procesów. Historia jest więc tak naprawdę nauką o informacji. W swej istocie informuje terażniejszość o przeszłości<sup>16</sup>.

Informacją jako taką zajmują się dwie odrębne dyscypliny naukowe. Informatyka (*computer science*) to generalnie nauka o przetwarzaniu infor-

<sup>15</sup> Zob. szerzej na temat nowych cech formalnych i funkcjonalnych dokumentów elektronicznych w kontekście badań naukowych: K. Narojczyk, *Dokument elektroniczny w badaniach naukowych. Szanse i zagrożenia*, w: *Uniwersum piśmiennictwa wobec komunikacji elektronicznej*, red. K. Migoń, M. Skalska-Zlat, Wrocław 2009, s. 199–213.

<sup>16</sup> L.J. McCrank, *Historical Information Science: History in Information Science; Information Science in History*, w: *History under debate*, s. 179.

macji, w tym zwłaszcza w sposób zautomatyzowany za pomocą technologii komputerowej. W centrum jej dociekań są matematyczne i logiczne podstawy przetwarzania danych, uwarunkowania techniczne: hardware i software oraz praktyczne aspekty pracy z technologią komputerową, w tym programowanie, projektowanie systemów, praca w sieciach itd. Nie zajmuje się ona natomiast w jakikolwiek sposób analizowaniem treści i znaczenia informacji, a jedynie problemami związanymi z jej cyfrowym przekształcaniem i transferem danych. Inaczej jest w przypadku informatologii<sup>17</sup> (*information science*). Bada ona problemy informacyjne poszczególnych obszarów ludzkiej aktywności, takich jak: nauka, kultura, gospodarka, administracja itd., itd. Wypracowując podstawy teoretyczne funkcjonowania informacji w społeczeństwie, zmuszona jest informację nie tylko kwantyfikować i opisać formalnie – jak informatyka, ale także ocenić oraz przeanalizować jej treść i kontekst funkcjonalny. Przedmiot jej zainteresowań dotyczy więc w istocie wiedzy, jaka wynika z operowania informacją. Informatologia nie kwantyfikuje informacji jako takiej, lecz ustala prawidłowości jej obiegu społecznego od chwili powstania, poprzez procesy modyfikowania treści, na dystrybuowaniu różnymi kanałami medialnymi kończąc. Analizuje także potrzeby informacyjne odbiorcy oraz reguły zarządzania informacją w celu optymalizowania jej przepływów i budowy efektywnych systemów informacyjnych.

Dziś trudno wyobrazić sobie dziedzinę nauki lub praktyki, która nie korzysta w mniejszym lub większym zakresie ze wsparcia informatycznego i informatologicznego. Na styku poszczególnych dyscyplin naukowych i wspierających je nauk o informacji wykształca się przestrzeń dla informatyk specjalistycznych, sprofilowanych pod kątem specyficznych potrzeb owych dyscyplin. Wykorzystują one zarówno zaadaptowane uniwersalne narzędzia, techniki i metody informatyczne, jak i rozwiązania *stricte* utylitarne, dostosowane do wąskich sprofilowanych potrzeb. Już w 1985 r. Heiner Müller-Merbach, parafrazując Kanta, sformułował tezę, że „Każda pojedyncza dyscyplina naukowa będzie mogła osiągnąć tylko taką sprawność, jaką posiada jej informatyka”<sup>18</sup>. Praktyka dowiodła, że takie wsparcie może okazać się bardzo konstruktywne także dla humanistycznego modelu przetwarzania informacji. W przypadku historii cyfrowej możemy mówić o przetwarzaniu informacji w naukach historycznych na wszystkich etapach postępowania badawczego oraz kształto-

<sup>17</sup> W Polsce można też spotkać określenia „nauka o informacji” lub „informacja naukowa”.

<sup>18</sup> H. Müller-Merbach, *Ansätze zu einer informationsorientierten Betriebswirtschaftslehre*, w: *Information und Wirtschaftlichkeit*, Hrsg. W. Ballwieser, K.-H. Berger, Wiesbaden 1985, s. 121.

wania społecznej roli historii poprzez udział w transferze wiedzy do społeczeństwa. Oznacza to, że zakresy tematyczne informatyki i informatologii zostały ściśle skojarzone z teorią, metodami i technikami nauk historycznych. Nowe technologie nie służą wyłącznie jako medium prezentacji ustaleń historyków, ale także jako efektywne narzędzia wspomagające procesy badawcze. Niejednokrotnie są też swoistym katalizatorem w poszukiwaniach nowych innowacyjnych metod badawczych.

Nasza znajomość niegdysiejszych cywilizacji zależy w dużej mierze od charakteru mediów, za pomocą których w czasach owych cywilizacji kreowano i dystrybuowano informację oraz od możliwości ich przetrwania, odnalezienia i zrozumienia. Każde z mediów informacyjnych (zabytki piśmiennictwa, zabytki kultury materialnej, rozmaite artefakty, kości ludzkie, ikonografia itd.) ma określone właściwości i wewnętrzne ograniczenia. Historyk odczytuje nie tylko treść medium, ale także jego ukonstytuowanie w rzeczywistości społecznej. Każde medium, jakiegokolwiek by ono było, samo w sobie stanowi dla historiografii komunikat. Dziejopisarstwo ma, zdaniem Johna Durhama Petersa, wiele wspólnego z ogólną teorią komunikacji. W obu przypadkach mamy do czynienia z operowaniem informacją przesuniętą w czasie i przestrzeni. W komunikacji nagranie informacji to nadanie jej jakiejś utrwalonej formy zapisu, transmisja to przesłanie tej formy na określoną odległość i/lub po upływie określonego czasu, interpretacja to jej odebranie i odczytanie (zdekodowanie). Badanie historyczne zawsze polega na odnalezieniu takich zapisów z przeszłości, przesłedzeniu ich transmisji w czasie (ewentualnie także w przestrzeni) oraz adekwatnej interpretacji treści przekazu. Badanie form ludzkiego komunikowania się w czasie i przestrzeni jest nie tylko uzupełnieniem dociekań historycznych, ale może także stanowić asumpt do nowego podejścia wobec badań historycznych traktowanych jako rozwiązywanie problemów w komunikacji teraźniejszości z przeszłością<sup>19</sup>.

Przedmiot zainteresowań historii cyfrowej stanowią zasadniczo formalne procedury badawcze nauki historycznej, funkcjonujące w dwóch wzajemnie się uzupełniających i wspierających zakresach praktycznych. Z jednej strony, wspiera ona cały proces badawczy historyka – od dokumentowania źródeł, zjawisk i procesów, poprzez ich analizowanie, do prezentowania uzyskiwanych wyników. Z drugiej zaś, z uwagi na jej praktyczną użyteczność i możliwość integrowania różnych pól badawczych, oddziałuje ona na teorię, metody i narzędzia poszczególnych dyscyplin

---

<sup>19</sup> J.D. Peters, *History as a Communication Problem*, <<https://drive.google.com/file/d/0B4YlqWJ4zziOMjU1MTIkOTctZWZiNi00NDIwLTgwMWUtNDRhZDI5MDA2YmZh/view?pli=1>> [dostęp: 12 XII 2014].

„historycznych”. Historia cyfrowa jest więc dziś dyscypliną zarówno naukową (teoretyczną), jak i praktyczną, za pomocą której można badać i rozwiązywać pragmatyczne oraz/lub koncepcyjne problemy, związane z wykorzystaniem nauk o informacji i technologii informatycznej w nauczaniu, badaniach oraz społecznym upowszechnianiu historii. Odgrywa ona także ogromną rolę w publicznym transferowaniu wiedzy o przeszłości i tym samym odpowiada w znacznej mierze za wypełnianie przez naukę historyczną jej funkcji społecznych.

Jako specjalistyczne pole badań historycznych, zajmujące się przetwarzaniem informacji z przeszłości i o przeszłości, historia cyfrowa stanowi swoisty interdyscyplinarny łącznik pomiędzy specjalistycznym zakresem badawczym historii a teorią i praktyką operowania informacją. W centrum jej zainteresowań teoretycznych leżą zagadnienia wzajemnych relacji pomiędzy rzeczywistością historyczną a poświadczającymi ją źródłami. Jedną z istotnych kwestii jest np. problem niejednoznaczności informacyjnej przekazów źródłowych i ich często kontekstowy charakter. Wymaga to stosowania specyficznych metod automatycznego przetwarzania takich informacji, w tym stosowania logiki rozmytej (*fuzzy logic*). Modele cybernetyczne zaimplementowane z informatologii pomagają opisywać procesy informacyjne towarzyszące zarówno samemu powstaniu źródeł, jak i ich późniejszemu społecznemu odczytywaniu oraz reinterpretacji. Pozwala to uchwycić diachroniczną perspektywę komunikacji społecznej i towarzyszące jej procesy poznawcze. Przedmiotem rozważań jest tu więc nie tylko sama treść informacji, ale także wszystkie okoliczności (ludzie i wydarzenia) związane z jej powstaniem, przekształcaniem, rozprzestrzenianiem czy wykorzystywaniem. Jak już wspomniano, przedmiotem badań są także zagadnienia związane ze źródłoznawstwem, czyli analizowaniem formalnych cech przekazów źródłowych, uwzględniające zarówno dotychczasowe typologie, jak i kryterium wewnętrznej struktury informacyjnej źródła. Historia cyfrowa wspiera także procesy badania wartości informacyjnej źródeł, a więc tradycyjną krytykę źródłową (w przypadku materiałów elektronicznych jej metody i narzędzia stanowią podstawę krytyki erudycyjnej).

Historia cyfrowa stawia sobie dwa zasadnicze zadania:

- rozwój systemu informacyjnego nauki historycznej, począwszy od analizy systemowej i planowania, na inżynierii wiedzy kończąc;
- rozwijanie adekwatnych do potrzeb rozwiązań aplikacyjnych – od tworzenia podstaw teoretycznych do praktycznej implementacji gotowego oprogramowania.

Zadaniem przyszłościowym jest natomiast stworzenie swoistej historycznej stacji roboczej (*Historical Workstation*) opierającej się na architekту-

rze *client-server*, łączącej w sobie koncepcyjnie i praktycznie zarówno pakiety oprogramowania, jak i podstawy ogólnej wiedzy oraz metodologii historycznej.

Jak już wcześniej wspomniano, istotą historii cyfrowej jest przetwarzanie informacji z przeszłości lub o przeszłości. Przy całej złożoności tej problematyki można owe problemy podzielić na trzy zasadnicze grupy:

1. Związane z procesem ustalania tematu badawczego oraz kwerendy bibliograficzno-archiwalnej. Ten etap postępowania badawczego dla większości dyscyplin naukowych jest bardzo zbliżony i w jego trakcie stosuje się ogólne technologie pozyskiwania adekwatnej informacji. W naukach historycznych nieco odmiennym elementem na tym etapie badań jest zazwyczaj kwerenda archiwalna, która wiąże się zwykle z przejrzaniem dostępnych pomocy archiwalnych. Z uwagi na materię, o której pomoge te traktują, metodyka ich tworzenia jest nieco inna niż metodyki budowy bibliografii i pomocy bibliotecznych.

2. Problemy informacyjne związane bezpośrednio z samym postępowaniem badawczym historyka. Możliwe jest tu wyodrębnienie dwóch kategorii:

- α) Problemy informacyjne związane z krytyką źródłową i interpretacją materiału źródłowego, a więc te najbardziej zbliżone do klasycznego kanonu Nauk Pomocniczych Historii. Jednak historia cyfrowa odnosi się do wszystkich możliwych kategorii i rodzajów źródeł wykorzystywanych w badaniach. Niezależnie od całej swej złożoności i różnorodności form mają one określoną precyzję i dokładność. Często zawierają nieścisłości, informacje nieprawdziwe lub niejednoznaczne, bywają niekompletne lub niecałkowicie zachowane, mogą tworzyć skomplikowaną wielowarstwową strukturę wewnętrzną. Bywa też, że rozmaicie aranżując lub agregując dane źródłowe, możliwe jest uzyskanie różnego poziomu oglądu minionej rzeczywistości, np.: człowiek, rodzina, gospodarstwo domowe, społeczność lokalna itd. Historia cyfrowa potrafi efektywnie wspierać rozstrzygnięcie tych wszystkich kwestii. Jednak niejednokrotnie zdarza się, że historyk, przystępując do eksploracji źródeł, nie ma jeszcze w pełni skryształizowanych pytań, na które chciałby znaleźć odpowiedź. Kwestionariusz badawczy historyka powstaje wówczas w miarę postępów prac nad źródłami. Dlatego nie można stworzyć modelu systemu informacyjnego o charakterze uniwersalnym, to znaczy takiego, aby odpowiadał na pytania, których jeszcze nie sformułowano. Powstają wszakże efektywne systemy dla eksploracji rozpoznanych już pojedynczych katego-

rii źródeł, jednak jednostkowe dane wydobyte ze źródła są bezwartościowe bez znajomości ich szerokiego kontekstu dziejowego. Dlatego ów kontekst, pozwalający dane interpretować i zrozumieć, powinien być także dostępny w systemie informacyjnym (w praktyce chodzi zazwyczaj o system informatyczny) zbudowanym na podstawie owych danych. W trakcie badań źródłowych często okazuje się, że badany problem jest poświadczony poszlakowo lub nieostro w wielu różnorodnych przekazach. Powstaje wtedy pytanie, czy odnoszą się one do tożsamyh postaci, wydarzeń, procesów itd. W takim wypadku znaczenie informacyjne źródła zależy od kontekstu czasoprzestrzennego, który system informacyjny także powinien uwzględniać. Jednakże każdy zaproponowany przez wydawcę (edytora) model interpretacji danych źródłowych (choćby poprzez ich uporządkowanie lub strukturyzację) jest z definicji w pewien sposób subiektywny. Dlatego poprawnie zaprojektowany system powinien umożliwiać oddzielenie obu tych warstw, tj. treści źródła od metadanych dodanych przez wydawcę.

- β) Problemy informacyjne podczas analizy historycznej. Badanie historyczne musi uwzględniać zmienność w czasie i przestrzeni opisywanych zjawisk i procesów. Niezbędne są do tego narzędzia analityczne uwzględniające ową zmienność. Większość metod statystycznych wykorzystywanych przez historyków wywodzi się z nauk społecznych. Jeżeli uwzględniają one wpływ czasu, to bardziej jako element (parametr) doświadczalny (eksperymentalny). Także metody analiz przekrojowych nie zawsze dobrze sprawdzają się w badaniach diachronicznych. Zdarzają się zapożyczenia bardzo użyteczne, np. metody analizy jednostkowych przypadków (*event history*), ale ich szersze wykorzystanie wymaga chyba jeszcze badań. Podobne trudności występują także w przypadku konieczności interpretacji poszczególnych części i poziomów danych zagregowanych (zespolonych). Zazwyczaj przy próbach badania szczególnie złożonych procesów i wydarzeń historycznych dwuwymiarowy model okazuje się niewystarczający. Dzisiaj właściwie tylko systemy informatyczne potrafią poradzić sobie z tworzeniem i badaniem wielowymiarowych modeli minionej rzeczywistości.

3. Problemy informacyjne związane z prezentacją poszczególnych etapów i faz postępowania badawczego (prezentacja źródeł, hipotez roboczych, wykorzystanych metod badawczyh, przesłanek wnioskowania, przeprowadzonych dowodów, symulacji, eksperymentów czy w końcu

wyników badań – całościowych lub cząstkowych). Podstawowym problemem w tym względzie jest wypracowanie metod pozwalających oddać możliwie jasno i precyzyjnie zmienność czasu i przestrzeni historycznej oraz wielopoziomowość analizy historycznej (np. niedające się *explicite* zdefiniować prawidłowości postępowania człowieka, tzw. logika dziejów, analizy wariantowe i hipotetyczne, nieostrość pojęć itp.). Systemy informatyczne wykorzystywane do prezentowania danych historycznych z łatwością potrafią odzwierciedlać zmienność procesów w czasie za pomocą dobranego według odpowiedniego przelicznika czasu rzeczywistego.

### GŁÓWNE OBSZARY ZASTOSOWAŃ HISTORII CYFROWEJ

- Bazy danych i edycje źródłowe

Bez wątplenia podstawowym polem zastosowań historii cyfrowej jest dzisiaj zarządzanie dużymi zbiorami danych, realizowane najczęściej za pomocą relacyjnych baz danych. Jeśli chodzi o bazy typu źródłowego, rodzi się podstawowe pytanie, jaką metodologię należy stosować w trakcie ich projektowania. Czy powinny one możliwie wiernie odwzorowywać treść i strukturę materiału źródłowego, czy też raczej odzwierciedlać przyjęty model badawczy, a więc *de facto* pewną wizję dziejów skonstruowaną przez historyka. W pierwszym wypadku relacyjne bazy danych pozwalają na zaimplementowanie hierarchicznych modeli danych typowych dla tekstów o w miarę jednorodnej strukturze treści. Jednak wielu kategorii źródeł nie da się wtłoczyć, bez strat dla kompletności przekazu, w ramy relacyjnego modelu danych. Struktura danych w obrębie zbioru źródeł, a nawet w pojedynczym źródle, może się zmieniać, co sprawia, że budowanie na jej podstawie modeli bazodanowych bywa trudne, a czasami wręcz niemożliwe. Źródło to zazwyczaj złożona sieć informacji, zwykle dobrze rozumianych przez historyka, ale trudnych do odwzorowania w systemach informatycznych. Dlatego często źródłowe bazy danych zawierają jedynie część wystandaryzowanej i znormalizowanej według określonego schematu informacji źródłowej dostępnej w zgromadzonych materiałach. Schemat selekcji owych informacji może być związany z problemem badawczym, jaki ma być rozwiązany z pomocą bazy. Modelowanie struktury bazy przez pryzmat celu badawczego, a nie przez treść i układ źródła, jest zazwyczaj prostsze w praktycznej realizacji, jednak uzyskuje się tym sposobem narzędzie mniej uniwersalne.

- Pełnotekstowe bazy danych

Osobną kategorią źródłowych baz danych są tzw. bazy pełnotekstowe, to znaczy zawierające pełny, przytoczony *in extenso* tekst źródłowy.

Podstawą operowania przez system informatyczny na takim materiale jest wyodrębnienie i jednoznaczne oznaczenie wartości semantycznej<sup>20</sup> możliwie wielu fragmentów przechowywanych tekstów. Jest to proces bardzo pracochłonny i wymagający wysokich kwalifikacji merytorycznych, a przez to kosztowny. Jednak bardzo udana edycja elektroniczna *Herbarza polskiego* Adama Bonieckiego<sup>21</sup> z wykorzystaniem języka adnotowania XML dowodzi, że możliwe jest skonstruowanie półautomatycznego systemu tagowania, co stwarza nowe perspektywy dla rozwoju baz pełnotekstowych. Jest to o tyle istotne, iż wyodrębnianie struktur i łączenie ustrukturyzowanego tekstu z modelami semantycznymi stanowi kwintesencję przetwarzania informacji historycznej. Niezależnie od formy źródłowej bazy danych zawsze aktualny pozostaje postulat zachowania w takich edycjach także możliwości korzystania z tekstu pierwotnego (zeskanowany obraz, tekst w transkrypcji itp.).

- Analiza zawartości tekstów

Obszarem bezpośrednio łączącym się z bazami pełnotekstowymi są metody analizy zawartości i tematyki tekstu. Trzeba do nich zaliczyć w pierwszym rzędzie rozmaite stylometry (przypisywanie autorstwa, wyodrębnianie i porównywanie określonych słów, fraz, konstrukcji stylistycznych). Stylometryczna analiza dokumentów z określonego obszaru pomaga np. określać kierunki i szybkość rozprzestrzeniania się badanych idei, poglądów, prądów umysłowych itp. W przypadku źródeł masowych (np.: rejestry chrztu, małżeństw, zgonów, spisy podatkowe, spisy wyborców, poborowych, meldunkowe itp.) bardzo użyteczne mogą okazać się technologie, które badają i łączą ze sobą według określonego klucza wzmiankowane w różnych źródłach tekstowych postacie (*Nominal Record Linkage*).

- Statystyka historyczna i demografia dziejów

Tradycyjnym, mocno wspieranym przez historię cyfrową, obszarem zainteresowań jest statystyka historyczna i demografia dziejów. Narzędzia statystyczne wykorzystywane były w badaniach historycznych od dawna. Jednak ich funkcja w ostatnim czasie nieco się zmieniła. W latach siedemdziesiątych i osiemdziesiątych ub. wieku (czasy zafascynowania kliometrią) statystyka służyła głównie do weryfikacji hipotez, podobnie jak w naukach ekonomicznych i społecznych. Współcześnie traktowana jest bardziej jako narzędzie badania i opisu dużych zbiorów danych historycznych w kierunku wykrywania w nich prawidłowości (schematów, struktury, korelacji, cykli). W badaniach demograficznych z kolei, obok

<sup>20</sup> Proces ten określa się często jako adnotowanie tekstu lub tagowanie tekstu – od ang. *tag* – znacznik.

<sup>21</sup> A. Boniecki, *Herbarz polski*, CD-ROM v. 2.1, oprac. M. Minakowski, Kraków 2005.

rozmaitych baz danych, duże usługi oddają np. elektroniczne systemy rekonstrukcji rodzin.

- Digitalizacja i sieciowe udostępnianie źródeł

Jest to proces, który nabiera tempa i rozmachu. Można mieć nadzieję, że w przyszłości całkowicie zostanie pokonana bariera dostępu do źródeł związana z ich odległym miejscem przechowywania lub potrzebą ochrony rzadkich egzemplarzy. Digitalizacja nie pokona jednak, niestety, barier wynikających z praw komercyjnych lub przyczyn politycznych.

- Wykorzystanie potencjału informacyjnego danych wizualnych

Według amerykańskiego przysłowia „obraz wart jest tysiąca słów”. Łatwość generowania i niewielkie koszty reprodukcji obrazów za pomocą technologii cyfrowej sprawiły, iż we współczesnych badaniach coraz częściej wykorzystuje się informacje, jakich mogą dostarczyć odpowiednio zaprezentowane dane wizualne. Informatyka Historyczna z powodzeniem przyswaja różnorodne technologie wizualizacji danych naukowych. Znakomitym przykładem udanego importu jest graficzne prezentowanie danych historycznych w ujęciu geoprzestrzennym za pomocą technologii GIS (*Geographic Information System*).

## METODY BADAWCZE

Wśród wielu rozmaitych metod badawczych pozostających w sferze zainteresowań historii cyfrowej do najciekawszych należą<sup>22</sup>:

- Analiza jednostkowych przypadków (*event history*)

Analiza jednostkowych przypadków wywodzi się w prostej linii z nauk biomedycznych, gdzie od dawna stosowano tzw. analizę danych przeżycia. W ujęciu historycznym przedmiotem zainteresowania są okoliczności życiowe związane z jednostką ludzką, traktowane jako dynamiczny proces kolejnych wydarzeń, które mogą wystąpić lub nie w życiu konkretnego, objętego badaniem człowieka. Zmienna zależna w analizie wyraża prawdopodobieństwo lub przypadkowość zajścia określonego zdarzenia<sup>23</sup>, przy czym analiza nie koncentruje się na konkretnym momencie czasowym w przeszłości, w którym wydarzenie miało miejsce, ale

<sup>22</sup> Interesujący przegląd metod komputerowych wraz z przykładami zawiera raport: O. Boonstra, L. Breure, P. Doorn, *Past, Present and Future of Historical Information Science*, Amsterdam 2004, <<http://www.hist.msu.ru/Departments/Inf/Stud/Heuristics/Reader/PPFofHIS.pdf>> [dostęp: 20 XI 2014].

<sup>23</sup> Zob. szerzej: B. Francis, J. Pritchard, *Visualisation of Historical Events Using Lexis Pencils*, w: *Case Studies of Visualization in the Social Sciences*, eds. D. Unwin, P. Fisher, <[http://www.agocg.ac.uk/reports/visual/casestud/francis/conten\\_1.htm](http://www.agocg.ac.uk/reports/visual/casestud/francis/conten_1.htm)> [dostęp: 15 XI 2014].

na okresie przed wystąpieniem owego wydarzenia. Bada się, jak jedna lub więcej zmiennych niezależnych mogły wpłynąć na ryzyko wystąpienia zdarzenia. Ponieważ przedmiotem zainteresowania jest okres czasu, a nie określony moment w jego obrębie – do analizy można włączyć wyłącznie dane z owego przedziału czasowego. Są one przydatne w badaniu, nawet jeśli nie wiemy, co działo się z człowiekiem przed i po wskazanym okresie.

- Dedukcja środowiskowa

Z powodu niedostatku danych na poziomie indywidualnym historycy często przeprowadzają wnioskowanie na podstawie danych zbiorowych, np. próbują określić zachowanie jednostki na podstawie zachowań grupowych.

- Analiza szeregu czasowego

Modele szeregu czasowego mogą być stosowane do wszystkich źródeł, w których występują diachroniczne serie (źródła ekonomiczne, takie jak statystyka handlowa albo demograficzne, takie jak np. dane o populacji). Dlatego analiza szeregu czasowego może być stosowana w badaniach o różnorodnej problematyce.

- Badawcza analiza danych i *data mining* (drażenie danych)

Zestaw metod i technik umożliwiających eksplorację danych (struktura zbioru, najważniejsze zmienne modulujące, połączenia między danymi, odnajdywanie danych pojedynczych, niedających się powiązać z innymi).

- Analiza grupowa i symulacje komputerowe

Analiza grupowa usiłuje zidentyfikować jednorodne podgrupy przypadków w populacji (znaleźć maksymalną cechę łączącą w obrębie grupy i maksymalną cechę różnicą od innych grup). Symulacje komputerowe z kolei są projektowane najczęściej dla oceny zachowań w granicach szczególnie zdefiniowanego systemu. Najpierw definiuje się oraz ustawia strukturę modelu i kryteria operacji na nim, następnie analizuje się i rozwija strategię działania, przewiduje wynik końcowy i jego prawdopodobieństwo, a także określa prawidłowości i relacje pomiędzy działaniami. Warunkiem dobrej analizy jest stworzenie poprawnego modelu wraz ze wszystkimi operandami, czyli ustawieniem zmiennych i połączeń między nimi. W trakcie kolejnych symulacji można zmieniać i modyfikować dane wejściowe lub zmieniać parametry modulujące. W pierwszym wypadku (tzw. technika Monte Carlo) uzyskujemy hipotetyczny zestaw danych, który może być poddany kolejnym analizom statystycznym, w drugim badamy siłę czynników modulujących.

Wydaje się, iż nie ulega wątpliwości, że rozwój i upowszechnianie się technologii komputerowej (zwłaszcza internetu) wpłynął na organizację pracy oraz metody pozyskiwania i dystrybuowania informacji historycznej. Z pewnością umożliwił podjęcie badań, których dotychczas nie można było realizować z uwagi na barierę efektywności lub brak odpowiednich narzędzi i metod. Zmiany, jakie się dokonują w tym względzie, są oczywiste, chociaż chyba jeszcze nie do końca zrozumiałe. Czy mamy do czynienia z radykalną zmianą, przynajmniej w odniesieniu do niektórych pól badawczych, czy też ze stopniową, adaptacyjną zmianą środowiska pracy historyków? Zapewne w niektórych wypadkach zmiany są wręcz rewolucyjne, np. możliwość operowania wielkimi bazami danych, dostęp *on-line* do elektronicznych katalogów i baz pełnotekstowych, możliwości nawiązywania na niespotykaną dotąd skalę kontaktów naukowych na całym świecie. Wszystko to sprawia, że historia cyfrowa zmienia jakość środowiska, w jakim przychodzi działać historykowi. Wiadomo, że komputer nigdy nie zastąpi człowieka i nie napisze automatycznie syntezy historycznej, może jednak umożliwić bardziej wyrafinowaną i innowacyjną analizę wybranych problemów badawczych. Wykorzystując potencjał technologii informatycznej, można udzielać lepszych, pełniejszych odpowiedzi na stare pytania lub stawiać nowe kwestie, które do tej pory ze względu na barierę niezbędnego nakładu sił i środków nie były podejmowane.

#### BIBLIOGRAFIA

##### Literatura

- A Companion to Digital Humanities*, eds. S. Schreibman, R. Siemens, J. Unsworth, Oxford 2004.
- Boniecki A., *Herbarz polski*, CD-ROM v. 2.1, oprac. M. Minakowski, Kraków 2005.
- Fogel R.W., Engerman S., *Time on the Cross: The Economics of American Negro Slavery*, London 1974.
- Fossier L., Vauchez A., Cinzio V., *Informatique et histoire médiévale*, Rome 1977.
- Gutman H.G., *Slavery and the Numbers Game: A Critique of Time on the Cross*, Urbana 1975.
- L'Histoire médiévale et les ordinateurs: Medieval history and computers: rapports d'une table ronde internationale*, ed. K.F. Werner, Paris 1978.
- Lytje I., Bygholm A., *Artificial Intelligence and Humanistic Informatics*, „Lecture Notes in Computer Science” 1990, 451.
- McCrank L.J., *Historical Information Science: History in Information Science; Information Science in History*, w: *History under Debate. International Reflection on the Discipline*, eds. C. Baros, L.J. McCrank, New York 2004.
- McCrank L.J., *History, Archives, and Information Science*, „Annual Review of Information Science and Technology” 1995, 30.

- Müller-Merbach H., *Ansätze zu einer informationsorientierten Betriebswirtschaftslehre*, w: *Information und Wirtschaftlichkeit*, Hrsg. W. Ballwieser, K.H. Berger, Wiesbaden 1985.
- Narajczyk K., *Dokument elektroniczny w badaniach naukowych. Szanse i zagrożenia*, w: *Uniwersum piśmiennictwa wobec komunikacji elektronicznej*, red. K. Migoń, M. Skalska-Zlat, Wrocław 2009.
- Rosenzweig R., *Scarcity or Abundance? Preserving the Past in a Digital Era*, „American Historical Review” 2003, 108, 3.
- Steinbuch K., *Informatik: Automatische Informationsverarbeitung*, „SEG-Nachrichten” 1957.
- Turkel W.J., Muhammedi S., Start M.B., *Grounding Digital History in the History of Computing*, „IEEE Annals of the History of Computing” 2014, 36, 2.

### Źródła internetowe

- Boonstra O., Breure L., Doorn P., *Past, Present and Future of Historical Information Science*, Amsterdam 2004, <http://ww.hist.msu.ru/Departments/Inf/Stud/Heuristics/Reader/PPFofHIS.pdf> [dostęp: 20 XI 2014].
- Francis B., Pritchard J., *Visualisation of Historical Events Using Lexis Pencils*, w: *Case Studies of Visualization in the Social Sciences*, eds. D. Unwin, P. Fisher, <[http://www.agocg.ac.uk/reports/visual/casestud/francis/conten\\_1.htm](http://www.agocg.ac.uk/reports/visual/casestud/francis/conten_1.htm)> [dostęp: 15 XI 2014].
- Peters J.D., *History as a Communication Problem*, <<https://drive.google.com/file/d/0B4YlqWJ4zziOMjU1MTkOTctZWZiNi00NDIwLTgwMWUtNDRhZDI5MDA2YmZh/view?pli=1>> [dostęp: 12 XII 2014].
- Wilkowski M., *Wprowadzenie do historii cyfrowej*, Instytut Kultury Miejskiej, Gdańsk 2013, <<http://www.historiacyfrowa.ikm.gda.pl>> [dostęp: 12 XII 2014].
- WWW Virtual Library, <<http://www.vl-ghw.uni-muenchen.de/hwEnglish.html#sect31>> [dostęp: 15 XII 2014].

### ABSTRACT

Following the information and communication technologies development, the new kind of interdisciplinary research has been formed and termed the *digital history*. It focuses on the issues regarding data processing “from and about the past”. It’s a unique link between a specialist range of historical research and a contemporary theory and practice of information use. On the one hand, the digital history helps with particular stages of the research carried out by the historian. On the other hand, owing to abilities of various communities’ integration, it can have an effect on the research organization, theory, tools and methods. Therefore, it fulfils practical and theoretical function. The digital history allows one to examine and solve pragmatic and/or conceptual problems connected with the use of information science and information technologies in teaching and historical research and social transferring of knowledge about the past. Apart from terminological questions and issues regarding the digital history, its edifying research methods and the main areas of its implementation have been discussed as well. They give a chance to implement much more productive, sophisticated and innovative practices. What is more, they require understanding the essence of transformation taking place today.

**Key words:** digital history, humanistic computer science, useful disciplines of history, informatology, data base, source edition

## NOTA O AUTORZE

Krzysztof Narojczyk – dr hab., prof. UWM, kierownik Zakładu Archiwistyki, Bibliotekoznawstwa i Informatyki Naukowej Uniwersytetu Warmińsko-Mazurskiego w Olsztynie, badacz problematyki zastosowania metod komputerowych w historii i archiwistyce. Jest autorem wielu publikacji oraz wdrożeń praktycznych z zakresu historii cyfrowej. Członek Association for History and Computing i International Society for Knowledge Organization.