

EWA WIŚNIEWSKA
(ŁÓDŹ)

Pożytki ze znajomości osiągnięć digitalizacji w Stanach Zjednoczonych

1. Wprowadzenie; 2. Zarys historii ochrony praw autorskich: 2.1. Stany Zjednoczone; 2.2. Polska; 3. Zasoby Biblioteki Kongresu; 4. Zasoby Projektu Gutenberg; 5. Proces digitalizacji w Polsce; 6. Podsumowanie.

1

Digitalizacja to proces nadawania postaci cyfrowej danym pisany i drukowanym, zawartym na nośnikach magnetycznych lub innych. W praktyce oznacza to wprowadzanie danych do pamięci komputera za pomocą czytnika krokowego (skanera), odczytującego i rejestrującego kolejne warstwy materiału. Techniki cyfrowe służą do rozpoznawania załączonego tekstu (*Optical Character Recognition*, OCR)¹ lub obrazu, a nawet odczytywania pisma odręcznego, co znacząco przyspiesza i ułatwia pracę badacza. O ile skanowanie oraz rozpowszechnianie plików graficznych (na przykład w formatach JPEG, PNG lub TIFF) oraz plików dźwiękowych (w formatach MP3, WAV) nie jest procesem skomplikowanym, to dystrybucja tekstu może nastroczać trudności. Materiały tekstowe w postaci cyfrowej są stosunkowo rzadko udostępniane w niesformatowanych plikach, ze względu na brak możliwości wyszukiwania informacji. Popularnym formatem jest PDF (*Portable Document Format*), łatwo dostępny i odczytywany przez użytkowników na całym świecie. Dwa inne typy plików, HTML (*Hypertext Markup Language*)² i RTF (*Rich Text Format*)³, choć dają możliwość szybkiej prezentacji pliku w Internecie, nie wydają się odpowiednie ze względu na transformację wyglądu udostępnianego tekstu. Dokumenty są również często upowszechniane w postaci plików graficz-

¹ *What is OCR and OCR Technology*, ABBYY, <https://www.abbyy.com/en-ee/finereader/what-is-ocr> [dostęp: 5.07.2017].

² *HTML (Hypertext Markup Language)*, The Server Side, <http://searchmicroservices.techtarget.com/definition/HTML-Hypertext-Markup-Language> [dostęp: 5.07.2017].

³ *Rich Text Format (RTF)*, Search Exchange, <http://searchexchange.techtarget.com/definition/Rich-Text-Format> [dostęp: 5.07.2017].

nych, jak JPEG, PNG i inne, co umożliwia pozyskiwanie ich przez Internet w wysokiej rozdzielczości, a co za tym idzie – w dobrej jakości.

2

2.1. Ochrona praw autorskich w Stanach Zjednoczonych Ameryki Północnej przez wiele lat funkcjonowała w bardzo ograniczonym zakresie. Sytuacja uległa znaczącej zmianie dopiero w 1989 r. Co prawda pierwszy akt regulujący prawo własności intelektualnej, Copyright Act, został podpisany w 1790 r., ale stanowił, że autor ma prawo wyłączności do stworzonego przez siebie dzieła (książki, mapy lub wykresu⁴) jedynie przez okres 14 lat, a po jego upływie przez kolejne 14, o ile żył⁵. Wynika z tego, że prawa autorskie w tym okresie nie były chronione w przypadku śmierci uprawnionego.

Ustawa z 1831 r. przedłużyła okres ochrony własności intelektualnej z 14 do 28 lat, z możliwością przedłużenia o kolejne 14⁶. Do listy dzieł chronionych włączyła utwory muzyczne, choć jedynie w wersji zapisu nutowego⁷. Kwestia ich publicznego wykonywania nie była jeszcze w tym czasie regulowana.

Kolejne zmiany prawa autorskiego w XIX i XX w. wiązały się z rozwojem technologiczno-informatycznym. W 1865 r. ochrona prawna objęła również fotografie, w 1912 r. – filmy, w 1971 r. – nagrania głosowe⁸. W celu potwierdzenia prawa do dzieła, od 1909 r. konieczna była rejestracja materiału, nabycie koncesji i oznaczenie go znakiem © lub słowem *copyright*.

Tymczasem już w 1886 r. została podpisana Konwencja berneńska o ochronie dzieł literackich i artystycznych (dalej: Konwencja berneńska), czyli pierwsza wielostronna umowa międzynarodowa regulująca prawo własności intelektualnej⁹. Stany Zjednoczone aż do 1989 r. nie ratyfikowały tej konwencji i „należały jedynie do Konwencji powszechnej, której oddziaływanie w skali międzynarodowej jest

⁴ P. Goldstein, *International Copyright: Principles, Law, and Practice*, New York 2001, s. 6, Questia, <http://www.questia.com/read/105112909/international-copyright-principles-law-and-practice> [dostęp: 24.07.2017].

⁵ Copyright Act of 1790, U.S. Copyright Office, <https://copyright.gov/about/1790-copyright-act.html> [dostęp: 16.06.2017].

⁶ An Act to amend the several acts respecting copy rights z 3 lutego 1831 r., <http://memory.loc.gov/cgi-bin/ampage?collId=llsl&fileName=004/llsl004.db&recNum=483> [dostęp: 15.07.2017].

⁷ *Ibidem*.

⁸ P. Goldstein, *op. cit.*

⁹ Konwencja berneńska o ochronie dzieł literackich i artystycznych, z dnia 9 września 1886 r., przejrzana w Berlinie dnia 13 listopada 1908 r. i w Rzymie dnia 2 czerwca 1928 r., Kramarz.pl, http://www.kramarz.pl/KONWENCJA_BERNENSKA.pdf [dostęp: 2.07.2017].

w sumie mniejsze ze względu na niższy standard ochrony oraz zależność administracyjną od UNESCO¹⁰. Wszelkie umowy międzynarodowe, regulujące prawo autorskie w Stanach Zjednoczonych, miały charakter dwustronny i wymagały osobnych porozumień¹¹.

W art. 1 Konwencji berneńskiej zostały wyliczone typy dzieł podlegające ochronie: Wyrazy „dzieła literackie” obejmują wszelkie utwory literackie, naukowe i artystyczne, bez względu na sposób lub formę ich wyrażenia, jako to: książki, broszury i inne pisma; odczyty, przemówienia, kazania i inne dzieła tego samego rodzaju; dzieła dramatyczne lub dramatyczno-muzyczne ze słowami lub bez słów; utwory rysunkowe, malarskie, architektoniczne, rzeźby, utwory rytownicze i litograficzne; ilustracje, mapy geograficzne; plany, szkice i wyroby plastyczne [...] ¹².

Konwencja berneńska stanowiła kamień milowy w dziejach prawa autorskiego – gwarantowała twórcom prawo własności intelektualnej bez uprzedniej rejestracji dzieła lub postępowania sądowego. Każdy utwór lub przedmiot ochrony wymieniony w przytoczonym art. 1, zapisany lub w inny sposób zachowany na nośniku, stanowił własność autora. Wcześniej funkcjonowała zasada, zgodnie z którą prawo autorskie obowiązywało jedynie w kraju powstania dzieła. W praktyce oznaczało to brak jakiejkolwiek ochrony utworu poza granicami kraju, w którym powstało, jego swobodne kopiowanie, a nawet przywłaszczanie. Istotny był również art. 7 Konwencji, który stanowił, że: „Czas trwania ochrony [...] obejmuje życie autora i pięćdziesiąt lat po jego śmierci”¹³. Przewidziano zatem długi okres ochrony. Wraz z rozwojem technologiczno-informatycznym pojawiały się kolejne regulacje oraz uzupełnienia do Konwencji berneńskiej¹⁴. Akt był sygnowany przez 10 państw, m.in. Szwajcarię, Francję, Niemcy i Hiszpanię, jednak Stany Zjednoczone, obecne w czasie prac nad Konwencją, podpisały dokument dopiero 1 III 1989 r.¹⁵. Polska ratyfikowała Konwencję berneńską już w 1934 r.

¹⁰ M. Kępiński, *Przystąpienie Stanów Zjednoczonych do Konwencji Berneńskiej*, „Ruch Prawniczy, Ekonomiczny i Socjologiczny” 1990, R. LII, z. 3–4, s. 53, <https://repozytorium.amu.edu.pl/bitstream/10593/16234/1/006%20MARIAN%20K%C4%98PI%C5%83SKI.pdf> [dostęp: 20.06.2017].

¹¹ P. Goldstein, *op. cit.*

¹² Konwencja berneńska, art. 1.

¹³ *Ibidem*, art. 7.

¹⁴ *World Intellectual Property Organization. Policy*, WIPO, <http://www.wipo.int/policy/en/> [dostęp: 1.07.2017].

¹⁵ M. Kępiński, *op. cit.*

W 1967 r. powstała Światowa Organizacja Własności Intelektualnej¹⁶ (*World Intellectual Property Organization*), której celem była popularyzacja oraz ujednolicenie światowego prawa autorskiego. Ze względu na szybko postępujący rozwój wymiany danych, istotne było powołanie organu niezależnego, lecz działającego pod egidą Organizacji Narodów Zjednoczonych, skupiającego kraje całego świata wokół kwestii własności intelektualnej. Działalność tej instytucji, zarówno prawna, edukacyjna, jak i promocyjna, przyczyniła się do stworzenia oraz ratyfikacji w 1996 r. Traktatu WIPO o prawie autorskim (WIPO Copyright Treaty). Art. 5 tego aktu prawnego stanowi, że „zbiory danych lub innych materiałów, które ze względu na dobór lub układ treści stanowią wytwory intelektu, podlegają jako takie ochronie”¹⁷. Przez „zbiory innych materiałów” można rozumieć także zdigitalizowane zasoby archiwalne i biblioteczne, a więc zgodnie z tym przepisem podlegają one ochronie jako przedmiot własności intelektualnej.

Dokument ten stanowi uzupełnienie wcześniejszych aktów prawnych regulujących prawo autorskie. Rozwiązuje problem upowszechniania materiałów drogą elektroniczną, za pośrednictwem zdobywającego coraz większą popularność Internetu. Należy również wspomnieć, że art. 4 Traktatu WIPO omawia kwestię własności intelektualnej twórców programów komputerowych: „Programy komputerowe są chronione jak utwory literackie w rozumieniu artykułu 2 Konwencji berneńskiej. Ochrona ta odnosi się do programów komputerowych niezależnie od sposobu lub formy wyrażania”¹⁸. Traktat WIPO, choć sankcjonuje nietykalność wytworów pracy intelektualnej, nie wskazuje ani nie proponuje właściwej ochrony, sugeruje tylko konieczność jej implementacji przez twórcę i odsyła do przepisów Konwencji berneńskiej.

Implementację Traktatu WIPO w prawie amerykańskim zapewnia Digital Millennium Copyright Act, ustawa Kongresu Stanów Zjednoczonych z 28 października 1998 r., regulująca prawo własności intelektualnej. Ustawa ta, szeroko krytykowana ze względu na brak jednoznacznej interpretacji, doczekała się wielu poprawek, jak choćby The Unlocking Technology Act z 2013 i Unlocking Consumer Choice and Wireless Competition Act z 2014 r. Szczególną wartością stanowi zawarta w niej idea zwolnienia z odpowiedzialności prawnej dostawców internetowych, którzy nie mają wpływu na przesyłane treści i nie są w stanie zapobiec

¹⁶ *World Intellectual Property Organization. Policy.*

¹⁷ Traktat WIPO o prawie autorskim (WCT) Genewa (1996 r.), art. 5, Dziennik Urzędowy Wspólnot Europejskich z dnia 11 kwietnia 2000 r., [http://eur-lex.europa.eu/legal-content/PL/TXT/PDF/?uri=CELEX:22000A0411\(01\)&from=PL](http://eur-lex.europa.eu/legal-content/PL/TXT/PDF/?uri=CELEX:22000A0411(01)&from=PL) [dostęp: 1.07.2017].

¹⁸ *Ibidem*, art. 4.

naruszeniu prawa autorskiego. Jednakże akt zobowiązuje ich do niezwłocznego usunięcia spornej treści¹⁹.

Ustawa reguluje kwestię rozpowszechniania dzieł w Internecie, gwarantując bibliotekom, archiwom oraz instytucjom edukacyjnym możliwość nieodpłatnego udostępniania materiałów, pod warunkiem, że działają w dobrej wierze²⁰. Za naruszenie prawa uznawane jest czerpanie korzyści finansowych lub osiągnięcie innych zysków komercyjnych²¹.

Instytucja dysponująca danymi jest zobligowana do ich upublicznienia (w ramach tzw. otwartego dostępu) zwłaszcza badaczom zajmującym się konkretną dziedziną. Za sprzeczne z prawem uznaje się udostępnienie danych jedynie badaczom związanym z instytucją nimi dysponującą²².

Ustawa Digital Millenium Copyright Act uregulowała zasady dotyczące powoływania oraz kompetencje nowego organu – Bibliotekarza Biblioteki Kongresu. Pełni on dwojaką funkcję: jest kierownikiem Biblioteki Kongresu i przewodniczącym United States Copyright Office, urzędu utworzonego w 1897 r., którego głównym celem jest ochrona prawa autorskiego²³. Rozstrzyga spory dotyczące prawa autorskiego – według ustawy Bibliotekarz ma prawo arbitrażu w przypadku braku odpowiednich umów licencyjnych²⁴. Jest również uprawniony do wyrażania zgody na wykonywanie utworów muzycznych przechowywanych w Bibliotece Kongresu (jeśli brak oznaczenia autora) i może wskazać warunki ich upowszechniania²⁵.

2.2. Ustawa o prawie autorskim została uchwalona 29 marca 1926 r. W art. 1 stanowiła: „Przedmiotem prawa autorskiego jest od chwili ustalenia w jakiegobądź postaci (słowem żywym, pismem, drukiem, rysunkiem, barwą, bryłą, dźwiękiem, mimiką, rytmiką) każdy przejaw działalności duchowej, noszący cechę osobistej twórczości”²⁶. Znowelizowana w 1935 r. i obowiązująca aż do 1952 r. ustawa

¹⁹ Digital Millennium Copyright Act z 28 października 1998 r., title II: Online Copyright Infringement Liability Limitation, <https://www.gpo.gov/fdsys/pkg/PLAW-105publ304/pdf/PLAW-105publ304.pdf> [dostęp: 17.07.2017].

²⁰ *Ibidem*.

²¹ *Ibidem*.

²² *Ibidem*.

²³ *Overview of the Copyright Office*, <https://www.copyright.gov/about/> [dostęp: 15.07.2017].

²⁴ Digital Millennium Copyright Act, sec. 404: Exemption for Libraries and Archives, <https://www.gpo.gov/fdsys/pkg/PLAW-105publ304/pdf/PLAW-105publ304.pdf> [dostęp: 17.07.2017].

²⁵ *Ibidem*.

²⁶ Ustawa z dnia 26 marca 1926 r. o prawie autorskim, Dz.U. z 1926 r., nr 48, poz. 286; <http://isap.sejm.gov.pl/DetailsServlet?id=WDU19260480286> [dostęp: 5.09.2017].

przyznawała prawo własności twórcy dzieła²⁷ na okres życia i pięćdziesięciu lat po śmierci²⁸. Art. 55 tego aktu gwarantował również ochronę majątkową: „Twórca (lub jego prawny następca) może żądać od wkraczającego bezprawnie w jego prawa, by zaniechał naruszenia, wydał to, czym się z bogacił, a w razie winy wynagrodził wszelką szkodę”²⁹. Zaproponowane rozwiązanie było nowatorskie i jak podają źródła: „System zastosowany w polskiej ustawie o prawie autorskim nie miał odpowiednika w ówczesnych ustawodawstwach światowych. Nie definiował on co prawda prawa autorskiego jako własności, może dlatego, by uniknąć terminologicznego sporu, ale przeprowadzał uregulowanie go jako specyficznej własności”³⁰.

Ustawa z 10 lipca 1952 r. o prawie autorskim skupiała się na zapobieganiu oraz penalizacji plagiatu. Skracala czas obowiązywania ochrony dzieła do 20 lat:

od śmierci twórcy, przy dziełach zaś wspólnych i łącznych – od śmierci twórcy, który innych przeżył, 2) od opublikowania utworu anonimowo lub pod pseudonimem, chyba że twórca wcześniej ujawnił publicznie swoje autorstwo, 3) od opublikowania utworu, jeżeli autorskie prawa majątkowe służą osobie prawnej³¹.

Prawdziwym przełomem w prawie własności intelektualnej w Polsce była ustawa z 4 lutego 1994 r. o prawie autorskim i prawach pokrewnych, która wprowadzała nowe przedmioty ochrony. Były to przede wszystkim wymienione w art. 1 programy komputerowe:

W szczególności przedmiotem prawa autorskiego są utwory:

- 1) wyrażone słowem, symbolami matematycznymi, znakami graficznymi (literackie, publicystyczne, naukowe, kartograficzne oraz programy komputerowe),
- 2) plastyczne,
- 3) fotograficzne,
- 4) lutnicze,
- 5) wzornictwa przemysłowego,
- 6) architektoniczne, architektoniczno-urbanistyczne i urbanistyczne,
- 7) muzyczne i słowno-muzyczne,

²⁷ *Ibidem*, art. 6.

²⁸ *Ibidem*, art. 20.

²⁹ *Ibidem*, art. 55.

³⁰ L. Górnicki, *Rozwój idei praw autorskich: od starożytności do II wojny światowej*, Wrocław 2013, s. 238, http://www.bibliotekacyfrowa.pl/Content/42471/Rozwoj_idei_praw_autorskich.pdf [dostęp: 7.09.2017].

³¹ Ustawa z dnia 10 lipca 1952 r. o prawie autorskim, art. 26, Dz.U. z 1952, nr 34, poz. 234, <http://prawo.sejm.gov.pl/isap.nsf/download.xsp/WDU19520340234/O/D19520234.pdf> [dostęp: 7.09.2017].

- 8) sceniczne, sceniczno-muzyczne, choreograficzne i pantomimiczne,
- 9) audiowizualne (w tym wizualne i audialne)³².

Ustawa nadal obowiązuje, a jej ostatnia nowelizacja miała miejsce 5 kwietnia 2017 r.³³.

Warto również podkreślić, że w myśl wymienionej ustawy: „przedmiotem prawa autorskiego jest każdy przejaw działalności twórczej o indywidualnym charakterze, ustalony w jakiejkolwiek postaci, niezależnie od wartości, przeznaczenia i sposobu wyrażenia (utwór)”, bez konieczności jego bezpośredniego wskazania bądź nazwania³⁴.

Ustawa ponownie zmieniła czas obowiązywania ochrony autorskich praw majątkowych – wydłużyła go do 70 lat od śmierci autora lub od śmierci współtwórcy, który żył najdłużej³⁵. Wymienione po raz pierwszy programy komputerowe, w myśl art. 74 „podlegają ochronie jak utwory literackie”³⁶. Ustawa określa reguły dysponowania utworami, jak również ustanawia Komisję Prawa Autorskiego, czyli organ nadzorujący prawidłowe egzekwowanie praw autorskich.

3

Biblioteka Kongresu Stanów Zjednoczonych jest największą biblioteką świata. W jej zbiorach, zgromadzonych w trzech imponujących rozmiarów budynkach głównych (im. Thomasa Jeffersona, im. Johna Adamsa i im. Jamesa Madisona), znajdują się łącznie ponad 164 miliony pozycji³⁷.

Według oficjalnych informacji, zawartych na stronie internetowej, Biblioteka posiada 24 189 688 książek (w tym ponad dwie trzecie to książki nieanglojęzyczne); 14 660 079 materiałów drukowanych, wśród których są monografie i serie, gazety, broszury, raporty, inkunabula oraz zbiory specjalne; ponad 3,5 mln

³² Ustawa z dnia 4 lipca 1994 r. o prawie autorskim i prawach pokrewnych, art. 1, Dz.U. z 1994 r., nr 24, poz. 83, <http://prawo.sejm.gov.pl/isap.nsf/download.xsp/WDU19940240083/U/D19940083Lj.pdf> [dostęp: 7.09.2017].

³³ Obwieszczenie Marszałka Sejmu RP z dnia 5 kwietnia 2017 r. w sprawie ogłoszenia jednolitego tekstu ustawy o prawie autorskim i prawach pokrewnych, Dz.U. z 2017 r., poz. 880, <http://prawo.sejm.gov.pl/isap.nsf/download.xsp/WDU20170000880/U/D20170880Lj.pdf> [dostęp: 7.09.2017].

³⁴ Ustawa z dnia 4 lipca 1994 r. o prawie autorskim i prawach pokrewnych, art. 1.

³⁵ *Ibidem*, art. 36.

³⁶ *Ibidem*, art. 74.

³⁷ *General Information*, Library of Congress, <https://www.loc.gov/about/general-information/> [dostęp: 20.04.2017].

materiałów audio, takich jak dyski, taśmy, audiobooki; a także licznie mapy, plakaty, rysunki, fotografie i rękopisy³⁸.

Biblioteka powstała w 1800 r. z inicjatywy prezydenta Johna Adamsa i od początku istnienia była jednostką nie tylko przechowującą i upowszechniającą materiały, ale i badawczą. W 1989 r. wśród pracowników 101 instytucji należących do Association of Research Libraries oraz 50 bibliotek stanowych przeprowadzono ankietę badającą zainteresowanie rozpowszechnianiem zbiorów bibliotecznych za pomocą nowoczesnych technologii informatycznych. Wynik badania wskazywał na duży popyt na ofertę, zwłaszcza w jednostkach badawczych³⁹.

W latach 1990–1994 prowadzono pilotażowy program, nazwany American Memory Project, który miał na celu udostępnienie zbiorów Biblioteki Kongresu. Wśród upublicznionych na płytach CD i kasetach video zasobów znalazły się nagrania, obrazy, materiały tekstowe, mapy oraz zapisy nutowe⁴⁰. Zdigitalizowana kolekcja cieszyła się ogromnym zainteresowaniem nie tylko badaczy i pracowników szkół, ale również uczniów, którzy docenili możliwość nauki bez konieczności fizycznego odwiedzania Biblioteki. Ten fakt zdopingował pracowników Biblioteki, z jej szefem, Jamesem H. Billingtonem na czele, do uruchomienia The Library of Congress National Digital Library Program (NDLP), zainicjowanego 1 maja 1995 r.⁴¹. Choć proces digitalizacji był czasochłonny i generował wysokie koszty, do 2000 r. udało się udostępnić online 5 mln dzieł, które zostały użyte do stworzenia zbioru zwanego American Memory Historical Collections for The National Digital Library (AMHC)⁴².

Strona internetowa Biblioteki Kongresu imponuje dostępnością materiałów oraz przejrzystością układu⁴³. W zdigitalizowanych zbiorach Biblioteki znajduje się 19 102 621 materiałów, z czego online dostępnych jest 2 213 921 obiektów⁴⁴.

W kolekcji znajduje się baza Chronicling America. Historic American Newspapers, powstała dzięki zainicjowanemu przez National Endowment for the Humanities (NEH) w 1982 r. U.S. Newspaper Program, mającemu na celu ochronę

³⁸ *Ibidem*.

³⁹ *National Digital Library Program*, Library of Congress, <https://memory.loc.gov/ammem/dli2/html/lcndlp.html> [dostęp: 21.04.2017].

⁴⁰ *American Memory. Mission and History*, The Library of Congress, <https://memory.loc.gov/ammem/about/index.html> [dostęp: 19.04.2017].

⁴¹ *Ibidem*.

⁴² *Ibidem*.

⁴³ Library of Congress, <https://www.loc.gov> [dostęp: 15.04.2017].

⁴⁴ *Ibidem*.

i przechowywanie prasy amerykańskiej⁴⁵. NEH to utworzona w 1965 r. agencja federalna, wspierająca rozwój humanistyki w Stanach Zjednoczonych. Jednym z działań agencji jest dotowanie instytucji takich jak muzea, uczelnie, archiwa i biblioteki oraz finansowanie programów wspierających i promujących omawianą dziedzinę nauki⁴⁶.

Rządowy program finansujący digitalizację prasy w ogóle (National Digital Newspaper Program, NDNP) zapewnia stały dostęp, również internetowy, do krajowych gazet i czasopism, przechowywanych w bazie Biblioteki Kongresu. Program jest finansowany w całości przez NEH, przy wsparciu technicznym Biblioteki Kongresu, pod której egidą działa⁴⁷.

Strona internetowa bazy Chronicling America oferuje nieodpłatny przegląd prasy ze zbiorów instytucji biorących udział w NDNP. Celem programu było odnalezienie i zachowanie przez katalogowanie oraz fotografowanie prasy – od najstarszych wydań z 1690 r. aż do współczesnych⁴⁸. Istotnym aspektem programu jest współpraca i wymiana informacji pomiędzy instytucjami, dzięki czemu biblioteki i archiwa przechowujące wydania prasy mają pewność, że nie duplikują istniejących już materiałów i dokonywana przez nie archiwizacja danych jest ważnym elementem procesu. Dokumenty w zasobach są zapisane w postaci plików TIFF, JPEG2000, PDF, plików tekstowych oraz obiektów METS XML⁴⁹. Poza zachowaniem egzemplarzy prasy dla przyszłych pokoleń, Chronicling America wskazuje cele edukacyjne i badawcze jako swoją podstawową misję.

Korzystanie z zasobów wchodzących w skład Chronicling America nie nastrocza problemów nawet początkującemu użytkownikowi Internetu: należy otworzyć stronę główną Biblioteki Kongresu (<https://www.loc.gov/>), która oferuje szereg modułów, takich jak Library Catalog (katalog główny, <https://catalog.loc.gov/>), funkcjonujący w ramach Biblioteki Copyright Office (urząd ds. praw autorskich, <https://copyright.gov/>) czy Digital Collections (katalog zasobów zdigitalizowanych, <https://www.loc.gov/collections/>). Ostatni z wymienionych odnośników przenosi użytkownika wprost do bazy Chronicling America, czyli do zdigitalizowanych artykułów prasowych będących w posiadaniu Biblioteki Kongresu.

⁴⁵ *U.S. Newspaper Program*, National Endowment for the Humanities, <https://www.neh.gov/us-newspaper-program> [dostęp: 29.07.2017].

⁴⁶ *About NEH*, National Endowment for the Humanities, <https://www.neh.gov/about> [dostęp: 30.07.2017].

⁴⁷ *Ibidem*.

⁴⁸ *About Chronicling America*, Chronicling America. Historic American Newspapers, <http://chroniclingamerica.loc.gov/about> [dostęp: 20.07.2017].

⁴⁹ *Ibidem*.

Zbiór zawiera pliki gazet codziennych z lat 1789–1924 pochodzących z 48 stanów. Dzięki łatwo dostępnej wyszukiwarce możliwe jest odnalezienie numeru gazety ze wskazanego przez użytkownika stanu, roku i przedziału czasowego. Dodatkowym udogodnieniem jest możliwość wpisania w pole wyszukiwarki słowa lub frazy i uzyskania artykułów prasowych je zawierających. Innym wartym wspomnienia narzędziem pomocniczym jest znajdująca się również na stronie Chronicling America wyszukiwarka tematyczna, która do często pojawiających się zapytań lub wydarzeń historycznych dopasowuje artykuły prasowe. Reading Room, czyli czytelnia jest dostępna pod adresem <http://www.loc.gov/rr/news/topics/topics.html> i stanowi ogromną wartość edukacyjną.

Strona internetowa Chronicling America zawiera jedynie część zasobów prasowych Biblioteki Kongresu. Wynika to z braku zdigitalizowanych kopii pozostałych materiałów. By móc odnaleźć inne zachowane artykuły prasowe, zwłaszcza te sprzed 1789 r., warto udać się do sekcji U.S. Newspaper Directory, 1690–Present (<http://chroniclingamerica.loc.gov/search/titles/>), podającej miejsce przechowywania fizycznej kopii artykułu.

4

Projekt Gutenberg to inicjatywa stworzona przez Michaela Harta, mająca na celu rozpowszechnienie książek w wersji elektronicznej⁵⁰. W 1971 r. na Uniwersytecie w Illinois Hart ręcznie wpisał tekst Deklaracji niepodległości Stanów Zjednoczonych do pamięci komputera, tworząc w ten sposób pierwszą elektroniczną książkę na świecie⁵¹. Zainicjowana przez niego aktywność rozwinęła się przez dekady w największą na świecie bibliotekę cyfrową, obecnie mającą w swych zasobach 55 444 pozycje książkowe (nie pojedyncze pliki)⁵².

W początkowej fazie istnienia inicjatywy wolontariusze i osoby zainteresowane zbudowaniem pierwszej biblioteki cyfrowej ręcznie wpisywali teksty dzieł z tzw. *public domain*. Jest to pojęcie z prawa anglosaskiego, oznaczające twórczość nieobjętą prawem autorskim. Najczęściej oznacza to, że dany dokument nie podlegał prawu autorskiemu lub że takie prawo wygasło. W powszechnym rozumieniu przyjmuje się, że dzieło nigdy nieopublikowane, opublikowane pod pseudonimem lub anonimowo, lecz niezarejestrowane przed 1897 r., lub opublikowane i zareje-

⁵⁰ P. Hane, *Project Gutenberg Progresses*, "Information Today" 2004, vol. 21, nr 5, <http://www.questia.com/read/1G1-117037627/project-gutenberg-progresses> [dostęp: 3.09.2017].

⁵¹ *Ibidem*.

⁵² *Free ebooks – Project Gutenberg*, https://www.gutenberg.org/wiki/Main_Page [dostęp: 2.09.2017].

strowane w USA przed 1923 r. może przejść do *public domain*, czyli powszechnego i bezpłatnego dostępu⁵³.

Celem Projektu Gutenberg jest popularyzacja klasycznych utworów literackich oraz ich łatwa i bezpłatna dystrybucja. Według inicjatorów projektu każdy użytkownik sprzętu komputerowego z dostępem do Internetu może z łatwością przeczytać, jak również pobrać na swe urządzenie wybrany materiał, bez limitu danych. Wobec braku praw autorskich do dzieła proces jest w pełni legalny, a pobrane dokumenty (w postaci plików tekstowych lub skompresowanych) mogą być kopiowane i rozpowszechniane do celów edukacyjnych i badawczych.

Większość materiałów w bibliotece cyfrowej Projektu Gutenberg jest napisana w języku angielskim, ale znajdują się tam też dokumenty m.in. w języku chińskim, niemieckim, greckim, włoskim, portugalskim, szwedzkim i hiszpańskim (powyżej 50 pozycji) oraz arabskim, rosyjskim, rumuńskim, staroangielskim, słoweńskim i wielu innych (mniej niż 50 pozycji w danym języku). W języku polskim udostępniono dzieła 42 autorów. Najstarsze z nich to *Treny* Jana Kochanowskiego, w wersji tekstu opublikowanej w 1583 r.

5

Za początek cyfryzacji zasobów znajdujących się w polskich jednostkach badawczych oraz archiwach można uznać utworzenie Poznańskiego Centrum Superkomputerowo-Sieciowego (PCSS) w 1993 r. albo jego podłączenie do sieci naukowej TEN-155 w 1999 r., co umożliwiło wymianę informacji z innymi ośrodkami⁵⁴. Ale dla planu powszechnej digitalizacji dóbr kultury przełomowym był rok 2001, kiedy został wdrożony Program Powszechnej Informatyzacji „Nowoczesna Polska”⁵⁵. Celem Programu było wprowadzenie edukacji informatycznej na wszystkich szczeblach, w jego ramach „po raz pierwszy zaplanowano powstanie Polskiej Biblioteki Internetowej oraz, z czasem, organizacji pochodnych (Skarbnicy Wiedzy Narodowej i zorganizowanie Nowej Edukacji Multimedialnej)”⁵⁶.

Zagadnienie cyfryzacji, niezwykle popularne i przyciągające uwagę, wymagało olbrzymiego nakładu pracy oraz środków finansowych. W odpowiedzi na prośby

⁵³ Copyright Information Center, Cornell University Library, <https://copyright.cornell.edu/publicdomain> [dostęp: 5.09.2017].

⁵⁴ Kamienie milowe. Najważniejsze wydarzenia z historii PCSS, Poznańskie Centrum Superkomputerowo-Sieciowe, <http://www.man.poznan.pl/online/pl/strona/680> [dostęp: 8.09.2017].

⁵⁵ J. Komusińska, *Kto jest odpowiedzialny za digitalizację – organizacja i finansowanie cyfryzacji polskiej kultury*, s. 4, http://ekultura.org/wp-content/uploads/2014/09/ekspertyza_koordinacja-dzia%C5%82a%C5%84-digitalizacyjnych.pdf [dostęp: 8.09.2017].

⁵⁶ *Ibidem*.

płynące ze środowiska polskich bibliotekarzy, w 1999 r. Rada Ministrów poparła utworzenie międzysektorowego planu „Kwaśny papier. Ratowanie w skali masowej zagrożonych polskich woluminów, których papierowe strony ze względu na skład chemiczny ulegają degradacji w wyniku zakwaszenia”⁵⁷. W ramach programu, realizowanego w latach 2000–2008 (a w niektórych ośrodkach do 2012 r.), finansowano laboratoria badawcze, konserwatorskie oraz cyfryzacyjne, w których metodą mikrofilmowania digitalizowano zbiory⁵⁸. W 2007 r., w celu zapobieżenia chaosowi organizacyjnemu, zbiory podzielono na kategorie:

- archiwalne (ich cyfryzację nadzoruje Centrum Kompetencji – Narodowe Archiwum Cyfrowe),
- biblioteczne (Centrum Kompetencji – Biblioteka Narodowa),
- muzealne (Centrum Kompetencji – Narodowy Instytut Dziedzictwa),
- audiowizualne (Centrum Kompetencji – Narodowy Instytut Audiowizualny)⁵⁹.

Istotną pomoc w finansowaniu cyfryzacji zbiorów polskich ośrodków stanowił unijny Program Kultura, realizowany w Polsce w latach 2007–2013, a zwłaszcza Priorytet XI „Kultura i dziedzictwo kulturowe”⁶⁰. Według raportu końcowego, podsumowującego działania w zakresie cyfryzacji w latach 2008–2013, zrealizowano 991 projektów digitalizacyjnych o łącznej wartości ponad 363 mln zł⁶¹.

Wydarzeniem przełomowym w polityce cyfryzacji okazał się Kongres Kultury Polskiej, zorganizowany w 2009 r. Dyskusja na temat cyfryzacji dziedzictwa narodowego zaowocowała projektem działań mających na celu wsparcie nowoczesnych metod⁶². Również w 2009 r. został opublikowany Program digitalizacji dóbr kultury, oraz gromadzenia, przechowywania obiektów cyfrowych w Polsce 2009–2020, który nie tylko podsumował dotychczasowe osiągnięcia na tym polu, ale przede wszystkim ustanowił standardy oraz wskazał podmioty odpowiedzialne. Organem bezpośrednio odpowiedzialnym za proces digitalizacji jest Ministerstwo Kultury i Dziedzictwa Narodowego przy współpracy Ministerstwa Nauki i Szkolnictwa

⁵⁷ *Ibidem*, s. 9.

⁵⁸ *Ibidem*.

⁵⁹ *Ibidem*, s. 6.

⁶⁰ Program Kultura 2007–2013 w Polsce – analiza i podsumowanie, Warszawa 2013, http://kreatywna-europa.eu/wp-content/uploads/2016/01/Publikacja_PKK_www-1.pdf [dostęp: 8.09.2017].

⁶¹ Raport końcowy. Analiza i ocena działań w zakresie digitalizacji oraz tworzenia wirtualnych instytucji kultury w kontekście projektów zrealizowanych w ramach XI Priorytetu POLIŚ, https://www.pois.2007-2013.gov.pl/AnalizyRaportyPodsumowania/Documents/Digitalizacje_XI.pdf [dostęp: 8.09.2017].

⁶² J. Komusińska, *op. cit.*, s. 11.

Wyższego, zaś instytucjami współpracującymi są Biblioteka Narodowa oraz Naczelna Dyrekcja Archiwów Państwowych⁶³.

We wskazanym dokumencie zaplanowano następujące kwoty na potrzeby poszczególnych procesów:

- digitalizacja zasobów kultury niebędących materiałami audialnymi i audiowizualnymi – 420 mln zł (środki z budżetu 370 mln, UE – 50 mln),
- digitalizacja zasobów audialnych i audiowizualnych – 680 mln zł (środki z budżetu 520 mln, UE – 60 mln zł, inne – 100 mln zł),
- budowa cyfrowych repozytoriów i bezpiecznych banków zasobów cyfrowych – 230 mln zł (środki z budżetu – 180 mln zł, UE – 50 mln zł),
- budowa pracowni digitalizacyjnych – 238 mln zł (środki z budżetu – 138 mln zł, UE – 50 mln zł, inne – 50 mln zł),
- Krajowe Centra Kompetencji (badania, szkolenia, edukacja, promocja) – 155 mln zł (środki z budżetu – 150 mln zł, UE – 5 mln zł)⁶⁴.

Ogółem, według projektu, digitalizacja będzie kosztować 2883 mln zł, z czego większość nakładów (2338 mln zł) ma pokryć rząd polski z budżetu państwa⁶⁵.

Zadaniem Centrów Kompetencji jest nie tylko popularyzacja procesu cyfryzacji, ale przede wszystkim ochrona materiałów znajdujących się w ich zbiorach. Jedno z nich, Narodowy Instytut Audiowizualny, w ramach bazy NInA oferuje możliwość bezpłatnego dostępu do materiałów audio i wideo z lat 1927–2017⁶⁶. Całkowity koszt finansowania cyfryzacji w Narodowym Instytucie Audiowizualnym wyniósł w latach 2011–2015 20 mln zł (w ramach programu Kultura+)⁶⁷. Obecnie jest aktywnych ponad 180 projektów, których głównym celem jest digitalizacja zbiorów⁶⁸.

Drugie Centrum Kompetencji, Biblioteka Narodowa, prowadzi portal Polona, czyli bazę internetową Nowej Cyfrowej Biblioteki Narodowej⁶⁹. W zasobach

⁶³ Program digitalizacji dóbr kultury oraz gromadzenia, przechowywania i udostępniania obiektów cyfrowych w Polsce 2009–2020, Warszawa 2009, s. 35, <http://www.nac.gov.pl/wp-content/uploads/2015/05/Program-digitalizacji-2009-2020-1.pdf> [dostęp: 7.09.2017].

⁶⁴ *Ibidem*, s. 63.

⁶⁵ *Ibidem*.

⁶⁶ Ninateka, <http://ninateka.pl/> [dostęp: 8.09.2017].

⁶⁷ J. Komusińska, *op. cit.*, s. 12.

⁶⁸ *Digitalizacja w Polsce następuje najszybciej w obszarach o dużej konkurencji*, 1.08.2017, <https://resellernews.pl/digitalizacja-polsce-następuje-najszybciej-obszarach-o-duzej-konkurencji> [dostęp: 7.09.2017].

⁶⁹ Polona, <http://www.polona.pl/> [dostęp: 9.09.2017].

Polony znajduje się 1 932 260 obiektów⁷⁰ – należy zwrócić uwagę na niezwykle wartość historyczną niektórych z nich, jak Deklaracja cenzusowa z 161 r. n.e.⁷¹ czy Pokwitowania podatkowe z 208 r. n.e.⁷², których nie sposób udostępnić do badań metodą tradycyjną. Warto też zauważyć, że choć Polona działa w ramach i pod egidą Biblioteki Narodowej, prezentuje materiały również z innych ośrodków, między innymi Biblioteki Jagiellońskiej, Instytutu Badań Literackich PAN czy Archiwum Kraszewskich. Przepływ informacji między instytucjami oraz wzajemny dostęp do ewidencji danych pozwala na sprawny podział pracy nad digitalizacją oraz zapobiega dublowaniu plików.

6

Choć proces digitalizacji dziedzictwa narodowego jest w Polsce opóźniony w stosunku do dokonań Stanów Zjednoczonych, wzorce amerykańskie mogą posłużyć jako inspiracja. Biblioteka Kongresu jest nie tylko największą biblioteką świata, z milionami dzieł oraz rocznym budżetem przekraczającym 600 mln dolarów⁷³, ale jednym ze światowych pionierów szybko postępującej digitalizacji źródeł. Równocześnie jako urząd federalny (U.S. Copyright Office) stoi na straży prawa własności intelektualnej w Stanach Zjednoczonych: rejestruje dzieła i ich autorów, pobiera opłaty rejestracyjne oraz wysuwa sugestie zmian w obowiązującym prawie autorskim.

Biblioteka stale modernizuje posiadane zbiory, zmieniając nośniki danych, poszerzając zasoby fizyczne oraz elektroniczne. Serwisy na jej stronie internetowej umożliwiają wyszukanie tematyczne oraz wyszukiwanie według nazwiska autora. Nieodpłatny dostęp do zasobu źródeł umożliwia badaczowi pracę zdalną. Dodatkowym ułatwieniem jest funkcja przeszukiwania zbiorów specjalnych, takich jak artykuły prasowe. Dzięki niej badacz może łatwo i szybko dotrzeć do potrzebnego materiału, bez konieczności żmudnego czytania każdej ze stron w zbiorze. Szeroki wybór materiałów, ich dobra jakość oraz możliwość wyszukiwania tekstu na obrazie (dzięki metodzie OCR) pomagają w popularyzacji zbiorów i chronią je od zapomnienia.

⁷⁰ *Ibidem*.

⁷¹ Deklaracja cenzusowa (zeznanie majątkowe Hatresa syna Satabousa z 28 lipca 161 r.), Polona, <https://polona.pl/item/7036392/0/> [dostęp: 9.09.2017].

⁷² Pokwitowania podatkowe (dwa pokwitowania wpłat dokonanych przez Abousa syna Satabousa, 3 i 4 sierpnia 208 r.), Polona, <https://polona.pl/item/7036830/0> [dostęp: 9.09.2017].

⁷³ Library of Congress Fiscal 2017 Budget Justification, <https://www.loc.gov/portals/static/about/reports-and-budgets/documents/budgets/fy2017.pdf> [dostęp: 2.07.2017].

EWA WIŚNIEWSKA

(ŁÓDŹ)

*Pożytki ze znajomości osiągnięć digitalizacji
w Stanach Zjednoczonych*

Digitalizacja materiałów, czyli ich wprowadzanie do pamięci komputerów, jest obecnie najpopularniejszą metodą zachowywania i upowszechniania dziedzictwa kultury. Pionierami w tej dziedzinie byli Amerykanie, np. Michael Hart, który już w latach 70. XX w. dokonał pierwszej udanej cyfryzacji Deklaracji niepodległości. Powstałe tam bazy danych, oferujące dostęp do źródeł, imponują mnogością plików, a brak opłat i możliwość precyzyjnego przyjrzenia się i dokładnego zapoznania z materiałami ułatwiają pracę badawczą. Wzorce amerykańskie mogą posłużyć jako inspiracja do działań mających na celu ochronę i popularyzację dzieł zgromadzonych w polskich instytucjach.

Słowa kluczowe: digitalizacja, prasa amerykańska, biblioteka cyfrowa

EWA WIŚNIEWSKA

(ŁÓDŹ)

*The benefits from the knowledge of the digitalization achievements
in the United States*

The digitization of materials, which means their introduction into the memory of computers, is now the most popular method of preservation, as well as the popularization of cultural heritage. The pioneers in this field were Americans, including Michael Hart, who as early as in the 1970s made the first successful digitization of the Declaration of Independence. Emerging databases that offer access to resources are impressive with their multitude of files. The lack of fees and the ability to look closely at the materials and familiarize with them make research easier and faster. US standards may serve as inspiration for actions aimed at protecting and popularizing works collected in Polish institutions.

Key words: digitalization, American press, digital library