

# Digitalizacja materiałów archiwalnych

Joanna Łuba

Digitalizacja, inaczej cyfryzacja, polega na tworzeniu cyfrowych reprodukcji, a jej istotą jest przenoszenie informacji z obiektów analogowych (dokumenty, fotografie, mapy etc) na cyfrowe. Archiwiści poświęcają digitalizacji coraz więcej czasu, dostrzegając w niej szansę na zwiększenie zainteresowania zbiorami i szerzej – historią. Na szybę skanera trafiają archiwalia wszystkich typów. Skanują na coraz większą skalę również pasjonaci historii, osoby zainteresowane historią lokalną lub losami swoich rodzin. Dostępność do taniego sprzętu, bezpłatnego oprogramowania i pasja sprawiają, że w internecie obok zbiorów archiwów państwowych oraz społecznych, pojawiają się fotografie prywatne. W ten sposób rozproszone po całym kraju i za granicą archiwalia dotyczące makro- i mikrohistorii Polski tworzą bogate, otwarte i stale powiększające się archiwum cyfrowe.

## 1. KORZYŚCI Z DIGITALIZACJI

Skan zawsze jest gorszy od oryginału – już choćby dlatego, że dokument czy fotografia działa na co najmniej dwa zmysły, a plik cyfrowy tylko na wzrok. Dlaczego więc na tak dużą skalę trwa proces digitalizacji w archiwach państwowych i społecznych? Korzyści są ewidentne.

- » **Ochrona archiwaliów.** Po pierwsze na wypadek ich fizycznego zniszczenia (w szczególności dotyczy to materiałów archiwalnych, których nośnikiem jest kwaśny papier oraz negatywów nitrocelulozowych, które mogą ulec samodestrukcji) lub utracenia (np. na skutek klęsk żywiołowych), po drugie – ograniczamy ich eksploatację przez użytkowników archiwum.

- » **Poszerzenie dostępu do materiałów archiwalnych.** Dzięki dostępności archiwaliów w przestrzeni internetowej znikają ograniczenia terytorialne (przykładem może być projekt „Elektroniczne łączenie spuścizn”, którego efektem ma być wspólna, internetowa prezentacja zbiorów Instytutu Piłsudskiego w Ameryce i Archiwum Akt Nowych dotyczących Ignacego Paderewskiego, Władysława Sikorskiego i Ignacego Mościckiego). Poszerza się również grono użytkowników – dzięki internetowi do archiwaliów sięgają osoby, które nigdy nie przekroczyłyby progu archiwum. Idąc dalej, digitalizacja ułatwia również badania naukowe i wspiera działania edukacyjne (pliki cyfrowe są łatwiejsze do dydaktycznego wykorzystania).
- » **Integracja społeczności.** To cel digitalizacji szczególnie ważny dla archiwów społecznych. Przykładem może być wystawa *Ludzie od Schmitzlera: mieszkańcy przedwojennego Śniatynia*, eksponowana w 2012 w Domu Spotkań z Historią w Warszawie. Materiały do niej były gromadzone przede wszystkim dzięki wykorzystywaniu internetowych portali społecznościowych i genealogicznych. Wydaje się, że dzięki korzystaniu z takich form komunikacji i gromadzenia zbiorów cyfrowych można odbudować wiele spalonych lub wywiezionych archiwów polskich miast.
- » **Ułatwienie pracy archiwisty i badacza.** Zeskanowane w dużej rozdzielczości zdjęcie można powiększać i dzięki temu łatwiej je zidentyfikować.

Dziś nie trzeba już nikogo przekonywać do celowości digitalizacji. Nikt już nie pyta, czy digitalizować. Pojawiają się natomiast pytania, jak robić to profesjonalnie. Tu z odsieczą przychodzą archiwiści państwowi i pracownicy instytucji kultury specjalizujący się w digitalizacji. Najważniejszym źródłem informacji o cyfryzacji jest *Katalog Dobrych Praktyk Digitalizacji Materiałów Archiwalnych* (<http://www.nina.gov.pl/docs/instytut/katalog-dobrych-praktyk-digitalizacji-materia%C5%82%C3%B3w-archiwalnych.pdf>). Rzetelnym kompendium wiedzy jest również dostępny on-line (<http://www.bn.org.pl/download/document/1342175805.pdf>) pierwszy polski podręcznik digitalizacji obiektów piśmienniczych *Digitalizacja piśmiennictwa* pod redakcją Dariusza Parandowskiego (Warszawa 2010).

## 2. CZYM DIGITALIZOWAĆ?

Wybór sprzętu najczęściej zależy od dwóch czynników: rodzaju skanowanego materiału i – częściej – możliwości finansowych. Najtańsze i najpowszechniejsze są skanery płaskie, w których umieszczany na szybie obiekt jest przykrywany pokrywą. Są wśród nich takie, które skanują również materiały transparentne czyli negatywy i slajdy. Znacznie droższe są skanery planetarne, w których jednostka skanująca położona jest około metra nad specjalistycznym stołem, na którym układany jest digitalizowany obiekt. Używa się ich zwłaszcza do skanowania zbiorów bibliotecznych oraz archiwaliów o dużych formatach. Na rynku bywają również dostępne skanery przelotowe (pojedynczy arkusz jest przesuwany przez szczelinę we wnętrzu

urządzenia, w której znajduje się listwa matrycy światłoczułej), bębnowe (mające postać szerokiej rury wykonanej z przezroczystego tworzywa, która obraca się wokół własnej osi; bęben wiruje wraz z obiektem przytwierdzonym przezroczystym materiałem). W procesie digitalizacji używane są również cyfrowe aparaty fotograficzne.

### 3. CO DIGITALIZOWAĆ?

Wiele archiwów utożsamiając digitalizację z zabezpieczaniem zbiorów przed fizycznym zniszczeniem, skanuje wszystkie swoje materiały (digitalizacja masowa). Jeśli jednak digitalizacja jest selektywna, trzeba sobie jasno wyznaczyć kryteria dokonywania takiej selekcji. Można wartościować archiwalia poprzez ich znaczenie dla kultury i historii, stan zachowania, unikatowość, wartość edukacyjną oraz wartość naukową. Z digitalizacji powinny być wyłączone dublety.

W przypadku digitalizacji masowej, gdy archiwum nie ma odpowiedniej infrastruktury, łatwiej zlecić skanowanie zewnętrznej firmie (tzw. outsourcing). Czasem jest to rozwiązanie tańsze niż zakup skanera i przeszkolenie pracowników.

### 4. W JAKI SPOSÓB DIGITALIZOWAĆ?

- » **Przygotowanie sprzętu.** Przed skanowaniem należy zadbać o sprzęt – skaner powinien być ustawiony w miejscu nienarażonym na wstrząsy, w pomieszczeniu o stałej wilgotności i temperaturze bez obecności pyłów. Szybę skanera należy przecierać niepyłącym materiałem.
- » **Sprawdzenie obiektów.** Materiały przeznaczone do skanowania powinny być sprawdzone pod względem ich kompletności oraz stanu zachowania. Należy wyjąć z nich materiały mogące porysować szybę skanera. Podczas skanowania – tak jak zawsze w kontakcie z archiwaliami – należy używać rękawiczek. Przed zeskanowaniem obiekt musi być opatrzony opisem umożliwiającym jego identyfikację – konieczna jest paginacja i zasygnowanie obiektów.
- » **Parametry skanera.** Uzyskanie pożądanego efektu zależne jest od odpowiedniego ustawienia:
  - rozdzielczości fizycznej czyli maksymalnej liczby punktów na jednostkę długości obiektu, jaką może zarejestrować urządzenie. Im większa rozdzielczość, tym lepsze odwzorowanie obiektu analogowego. Jest wyrażana ilością punktów na cal – dpi (ang. dots per inch). Jej wybór zależy od typu i formatu skanowanego obiektu (definicja za *Standardy w procesie digitalizacji obiektów dziedzictwa kulturowego*, pod redakcją Grzegorza Płoszajskiego, Warszawa 2008);
  - formatu czyli pola skanowania. Powinien być większy niż obiekt, tak żeby można było zobaczyć kontury skanowanego materiału. Jego właściwe ustawienie jest możliwe dzięki skanowaniu wstępnemu (pre-scan). Dzięki tej funkcji, która tworzy szybki podgląd skanowanego

obrazu, można wyznaczyć pole skanowania i ocenić, czy obiekt jest dobrze ułożony na szybie skanera;

- **głębi bitowej.** Kolor obrazu jest rejestrowany w postaci składowych najczęściej światła czerwonego, zielonego i niebieskiego – w skrócie RGB (od ang. *red, green, blue*), które pozwalają odtworzyć większość widzianych kolorów. Wierność odwzorowania zależy od zdolności detektora do rozróżniania poziomów dla każdej z tych składowych oraz od dokładności, z jaką te wartości zostaną zapisane. Tę dokładność wyraża się w ilości bitów, w jakich są zapisane wartości, stąd mówi się o głębi kolorów np. 24-bitowej, 32-bitowej. Przy skanowaniu całkowicie czarno-białych dokumentów oraz czarno-białych negatywów zalecane jest stosowanie 8-bitowej skali szarości, w innych przypadkach należy używać 24-bitowej RGB. (definicja za *Standardy w procesie digitalizacji...* j.w.)
- » **Efekty skanowania powinny być kontrolowane.** Ważne jest również prawidłowe nazywanie plików – powinny zawierać numer zespołu i sygnaturę jednostki archiwalnej. Umożliwia to szybkie znalezienie oryginału.

## 5. JAK ARCHIWIZOWAĆ PLIKI CYFROWE?

- » **Format pliku.** Pliki powinny być zapisywane w bezstratnym formacie tiff. Po zeskanowaniu są zazwyczaj poddawane przekształceniom bezstratnym, czyli kadrowaniu, korekcji kolorów, obracaniu oraz przekształceniom stratnym (nie zalecane na plikach, które będą zarchiwizowane), czyli kompresji np. formatu tiff do jpg, zmianie skali, modyfikacjom (np. kontrastowaniu, usuwaniu zabrudzeń, które były na oryginałach). Można je wykonywać przy pomocy bezpłatnych programów, np. IrfanView i GIMP.
- » **Nośniki do przechowywania plików cyfrowych.** Ich wybór zależy od możliwości finansowych. Im droższy sprzęt, tym lepsze zabezpieczenie plików cyfrowych. Najpopularniejszym sposobem przechowywania plików jest zapisywanie ich na dyskach twardych podłączonych bezpośrednio do komputera (on-line) i dysków zewnętrznych (off-line). Do tanich, ale nietrwałych nośników, należą: pamięci flash (np. pendrive i karty pamięci) oraz dyski optyczne (np. płyty CD i DVD). Zalecanym sprzętem są macierze dyskowe, nośniki taśmowe, biblioteki taśmowe oraz repozytorium.

## 6. NAJCZĘSTSZE BŁĘDY POPEŁNIANE PRZY ŚKANOWANIU

**przekład** : *Odwrotne zeskanowanie negatywu*



Archiwum  
Fotografii Ośrodka  
KARTA

**przekład** : *Zagięcie dokumentu*



Archiwum  
Fotografii Ośrodka  
KARTA

**przykład:** Umieszczenie obiektu poza obszarem skanowania



Archiwum  
Fotografii Ośrodka  
KARTA

**przykład** : Nadanie zbyt niskiej rozdzielczości

Archiwum Fotografii Ośrodka KARTA



**przykład** : Nieprawidłowe przyciśnięcie zdjęcia do szyby skanera

Archiwum Fotografii Ośrodka KARTA



# Załącznik nr 1 STANDARDY W PROCESIE DIGITALIZACJI OBIEKTÓW DZIEDZICTWA KULTUROWEGO

## wymagania minimalne

Grupa materiałów (patrz tabela)	Grupa A	Grupa B	Grupa C	Grupa D	Grupa E	Grupa F	Grupa G
<b>Materiały przykładowe</b>	teksty drukowane	teksty drukowane z ilustracjami	rysunki i grafiki monochromatyczne, rękopisy, odbitki foto. cz/b	materiały fotograficzne: negatywy i przezroczka	mikrofilmy	malarstwo, odbitki foto. barwne, małe obiekty muzealne	plakaty, duże mapy, duże obiekty muzealne
<b>Format</b>	TIFF 6.0 z kompresją CCITT Group4*	TIFF 6.0 dopuszcza się kompresję bezstratną LZW**	TIFF 6.0 dopuszcza się kompresję bezstratną LZW**	TIFF 6.0 dopuszcza się kompresję bezstratną LZW**	TIFF 6.0 dopuszcza się kompresję bezstratną LZW**	TIFF 6.0 dopuszcza się kompresję bezstratną LZW**	TIFF 6.0 dopuszcza się kompresję bezstratną LZW**
<b>Rozdzielczość</b>	400 ppi	300 ppi	300 ppi, lecz nie mniej niż 3000 pikseli na dłuższym wymiarze	300 ppi, lecz nie mniej niż 3000 pikseli na dłuższym wymiarze	jak dla mikrofilmowego oryginału w granicach przenoszenia jego cech przez mikrofilm	300 ppi, lecz nie mniej niż 3000 pikseli na dłuższym wymiarze	300ppi
<b>Bitów na piksel</b>	1	8-bitowa skala szarości	8-bitowa skala szarości	8 bitów na kolor* 24-bit RGB albo 8-bitowa skala szarości	8 bitów na kolor* 24-bit RGB albo 8-bitowa skala szarości	8 bitów na kolor 24-bit RGB	8 bitów na kolor 24-bit RGB
<b>Wzorce szarości/koloru</b>	nie dotyczy	Gray Gamma 2.2	Gray Gamma 2.2	Adobe RGB 1998	nie dotyczy	Adobe RGB 1998	Adobe RGB 1998

## wymagania zalecane

<b>Format</b>	TIFF 6.0 z kompresją CCITT Group4	TIFF 6.0 dopuszcza się kompresję bezstratną LZW**	TIFF 6.0 dopuszcza się kompresję bezstratną LZW**	TIFF 6.0 dopuszcza się kompresję bezstratną LZW**	TIFF 6.0 dopuszcza się kompresję bezstratną LZW**	TIFF 6.0 dopuszcza się kompresję bezstratną LZW**	TIFF 6.0 dopuszcza się kompresję bezstratną LZW**
<b>Rozdzielczość</b>	600 ppi	400 ppi	400 ppi, lecz nie mniej niż 5000 pikseli na dłuższym wymiarze	600 ppi, lecz nie mniej niż 5000 pikseli na dłuższym wymiarze	jak dla mikrofilmowego oryginału w granicach przenoszenia jego cech przez mikrofilm	400 ppi, lecz nie mniej niż 5000 pikseli na dłuższym wymiarze	300ppi
<b>Bitów na piksel</b>	1	16-bitowa skala szarości	16-bitowa skala szarości	16 bitów na kolor 48-bit RGB albo 16-bitowa skala szarości	8 bitów na kolor**** 24-bit RGB albo 8-bitowa skala szarości	16 bitów na kolor 48-bit RGB	16 bitów na kolor 48-bit RGB
<b>Wzorce szarości/koloru</b>	-	Gray Gamma 2.2	Gray Gamma 2.2	Gamma 2.2. albo Adobe RGB 1998 lub lepszy***	nie dotyczy	Adobe RGB 1998 lub lepszy***	Adobe RGB 1998 lub lepszy

\* kompresja opisana w zaleceniach ITU-T T6 (11/1988) dla faksów grupy 4 [P08];

źródło: „standardy w procesie digitalizacji obiektów dziedzictwa kulturowego” pod redakcją grzegorza ploszajskiego, warszawa 2008

Utwór udostępniony na licencji Creative Commons Uznanie autorstwa-Użycie niekomercyjne-  
-Na tych samych warunkach 2.5 Polska (CC BY-NC-SA 2.5 PL)



**Artykuł jest częścią publikacji pt. „Archiwistyka Społeczna”,  
którego pełna wersja dostępna jest pod adresem: [www.archiwa.org](http://www.archiwa.org)**

redakcja merytoryczna: Katarzyna Ziętał

opracowanie redakcyjne: Aleksandra Janiszewska

projekt graficzny i skład: Diana Gawronkiewicz / Podpunkt druk: TZG

Zapolex Sp. z o.o.

Wydanie I, Warszawa 2012

Ośrodek KARTA

Ul. Narbutta 29, 02-536 Warszawa

Tel. (48-22) 848-07-12, faks (48-22) 646-65-11

Email: [ok@karta.org.pl](mailto:ok@karta.org.pl), [www.karta.org.pl](http://www.karta.org.pl)

**Ośrodek**  
**Karta**

ISBN: 978-83-61283-87-4

Tekst publikacji (z wyłączeniem materiałów wizualnych) dostępny jest na licencji **Creative Commons Uznanie autorstwa – Na tych samych warunkach 3.0**. Zezwala się na dowolne wykorzystanie treści pod warunkiem wskazania autorów oraz Fundacji Ośrodka KARTA jako właściciela praw do tekstu oraz zachowania niniejszej informacji licencyjnej. Utwory zależne zbudowane w oparciu o treść niniejszej publikacji muszą być rozpowszechniane na tych samych warunkach. Tekst prawny licencji z przystępnym podsumowaniem dostępny jest na stronie <http://creativecommons.org/licenses/by-sa/3.0/deed.pl>

Dofinansowano ze środków Programu „Archiwistyka społeczna”.



**NARODOWY  
INSTYTUT  
AUDIOWIZUALNY**

**Ministerstwo  
Kultury  
i Dziedzictwa  
Narodowego**