

Archiwum społeczne w internecie

Marcin Wilkowski

Nie da się dziś myśleć o archiwistyce społecznej bez perspektywy internetu. Nie oznacza to oczywiście, że wszystkie zbiory gromadzone w ramach oddolnych i często nieformalnych inicjatyw muszą być udostępniane online. Nie zawsze jest to przecież możliwe choćby ze względów prawnych (prawo autorskie, ochrona wizerunku), etycznych, finansowych czy organizacyjnych. Budując lub prowadząc już archiwum społeczne można skutecznie wykorzystać internet do wsparcia swojej archiwistycznej działalności – stworzenie archiwum cyfrowego udostępniającego skany jest optymalnym, ale na pewno nie koniecznym jej rozwinięciem.

W idei archiwum społecznego zapisana jest jego otwartość. Zbiory gromadzone są w nim nie tylko ze względu na ich immanentną wartość, ale także ze względu na ich społeczne znaczenie. A jeśli tak, archiwum społeczne musi być publiczne, działać jawnie i oficjalnie oraz być odpowiednio widoczne. Dziś nie da się tego osiągnąć bez internetu.

1. FORMY OBECNOŚCI ARCHIWUM SPOŁECZNEGO W INTERNECIE

Zaprezentowane poniżej propozycje form obecności archiwum społecznego w internecie czytać można jako kolejne etapy ewolucji jego otwartości – od podstawowej strony informacyjnej aż po serwis udostępniający skany gromadzonych w archiwum zbiorów.

1.1 Informacyjna strona internetowa archiwum społecznego

Niewielkim nakładem pracy stworzyć można prostą stronę internetową, na której znajdują się podstawowe informacje o archiwum. Zwiększy ona zasięg

oddziaływania archiwum i ułatwi dotarcie do niego osobom zainteresowanym.

Na stronie znaleźć się powinny m.in.:

- » dobrze widoczne dane kontaktowe (adres, numer telefonu, adres email);
- » charakterystyka gromadzonych zbiorów;
- » zasady udostępniania zbiorów (od kwestii związanych z możliwością wykonywania kopii po informacje o godzinach otwarcia archiwum);
- » elementy uwiarygadniające archiwum (może to być informacja o afiliacji – instytucji czy grupie inicjatywnej, przy której budowane jest archiwum społeczne, o partnerach merytorycznych czy finansowaniu z publicznych grantów).

Dobłą praktyką jest przygotowanie zbioru najczęściej zadawanych pytań (*Frequently Asked Questions*, FAQ), w których wprost przekazać można najważniejsze informacje dotyczące archiwum i zbiorów, np.: Czy mogę przekazać archiwum swoje zbiory? Czy do korzystania ze zbiorów potrzebne są jakieś dokumenty? Jak są zasady kopiowania dokumentów?

Przemyślana i pełna treść strony informacyjnej serwisu internetowego archiwum ma kluczowe znaczenie, choć i strony wizualnej nie należy zaniedbywać. Nagromadzenie grafiki czy umieszczenie na stronie długo wczytujących się animacji utrudniać będzie korzystanie z informacji na niej publikowanych. Organizacja W3C, definiująca podstawowe standardy internetowe, wypracowała *Wytyczne dotyczące dostępności treści internetowych (Web Content Accessibility Guidelines 2.0)*, wśród których znalazły się m.in. zalecenia dotyczące możliwości odczytania publikowanych na stronie informacji przez osoby niepełnosprawne, przejrzystości i intuicyjności nawigacji czy zgodności strony z przeglądarkami. Specyfika archiwów społecznych powoduje, że korzysta z nich wiele osób starszych, które mogą mieć problemy z przeglądaniem stron wypełnionych grafiką i animacjami czy czytaniem tekstu o zbyt małych czcionkach i słabym kontraście. Zasady te obowiązywać powinny na każdym etapie rozwoju serwisu internetowego archiwum – od prostej strony informacyjnej aż po repozytorium cyfrowe udostępniające skany.

W przypadku archiwów społecznych tworzonych przy dużych organizacjach pozarządowych kwestie techniczne tworzenia strony informacyjnej mogą być powierzone odpowiedniemu pracownikowi czy nawet firmie tworzącej strony WWW, jednak małe inicjatywy archiwalne, budowane oddolnie przez wolontariuszy, muszą radzić sobie same. Na szczęście dziś stworzenie nawet rozbudowanego serwisu internetowego nie jest problemem także dla tych, którzy nie potrafią programować. Skorzystać z gotowych, darmowych i otwartych rozwiązań można nie tylko w celu zbudowania prostej, statycznej strony, ale nawet archiwum cyfrowego, które publikować będzie cyfrowe wersje gromadzonych zbiorów.

Idea wolnego i otwartego oprogramowania polega (w dużym skrócie) na udostępnianiu skryptów (np. systemu zarządzania treścią strony) na takich zasadach, aby użytkownik mógł je swobodnie stosować, kopiować, rozposzechnić (a nawet sprzedawać). Większość oprogramowania tego typu udostępniana jest za darmo, a legalność korzystania z niego gwarantują odpowiednie publiczne licencje. Dzięki temu takie rozwiązania w bezpieczny sposób stosować można w instytucjach (np. organizacjach pozarządowych) czy nieformalnych grupach inicjatywnych. Do stworzenia strony archiwum wykorzystać można darmowe skrypty systemów zarządzania treścią (*Content Management System*, CMS) umożliwiające łatwe publikowanie artykułów, aktualności, galerii z materiałami wizualnymi itp. Na pewno polecić można dwa najpopularniejsze rozwiązania: Wordpress i Drupal – są one w pełni spolszczone, posiadają odpowiednio obszerną dokumentację i wsparcie techniczne ze strony społeczności koderów oraz oferują możliwość skorzystania z wielu gotowych szablonów wyglądu strony.

Instalacja i dostosowanie ich do własnych celów może być jednak sporym wyzwaniem dla osób nie mających doświadczenia. Oprócz tego pojawia się tu ważny problem bezpieczeństwa instalacji i odpowiedniej ochrony przed spamem. Na szczęście dostępne są gotowe rozwiązania. Serwisy takie jak wordpress.com czy blogger.com umożliwiają szybkie założenia strony internetowej i wybór jednego z wielu gotowych szablonów graficznych. Po zarejestrowaniu się i podaniu podstawowych parametrów strony możemy już dodawać do niej informacje.

Wiele społecznych i oddolnych inicjatyw historycznych wykorzystuje Facebooka jako podstawowe narzędzie publikowania informacji, a nawet gromadzenia czy opisywania zbiorów. Facebook ze względu na swoją popularność i łatwość w uruchomieniu podstawowej strony projektu (tzw. fan page) wydaje się dobrym rozwiązaniem problemu udostępniania informacji o archiwum społecznym oraz sposobem na promowanie jego zbiorów. Należy jednak mieć świadomość istniejących tam ograniczeń i ryzyka związanego z oparciem komunikacji wyłącznie o ten serwis. Po pierwsze, Facebook jako projekt komercyjny nastawiony jest na zysk, który – w obecnym modelu biznesowym – generowany jest ze sprzedaży reklam oraz gwarantowania lepszej widoczności informacji publikowanych przez użytkowników. Tylko niewielka część wszystkich wpisów publikowanych na facebookowej stronie naszego archiwum społecznego będzie widziana przez jej subskrybentów. Jeśli będziemy chcieli zwiększyć efektywność komunikacji, trzeba będzie zapłacić. Drugim ważnym wyzwaniem jest zupełnie niepartnerska relacja między administracją Facebooka a użytkownikami. Nie mają oni żadnej możliwości bezpośredniego skontaktowania się z przedstawicielem Facebooka w przypadku np. zablokowania wybranych materiałów czy nawet całej strony.

1.2 Katalog online archiwum społecznego

Jeśli archiwum społeczne ma już opracowany katalog czy inwentarz zbiorów (co jednak nie jest regułą), warto postarać się o jego udostępnienie online. Nie musi oznaczać to wielkich inwestycji w rozbudowę strony – najprostszym rozwiązaniem będzie opublikowanie spisu w przeszukiwalnym dokumencie tekstowym, arkuszu Excela czy pliku PDF. Dzięki temu użytkownik będzie miał możliwość zorientowania się w gromadzonych przez archiwum zasobach.

Rekomendacje. W zaplanowaniu bardziej rozbudowanego katalogu, który można by przeglądać bezpośrednio z poziomu strony WWW pomóc mogą rekomendacje wikipedystów dla współpracujących z nimi instytucji z sektora GLAM – galerii, bibliotek, archiwów i muzeów (*Galleries, Libraries, Archives, Museums*, GLAM). Dostępne na stronie http://meta.wikimedia.org/wiki/GLAM-WIKI_Recommendations#Requests_to_GLAM_2 zalecenia zwiększają użyteczność katalogu i ułatwiają korzystanie z niego: przeszukiwanie czy tworzenie odwołań (linków) do konkretnych pozycji. Przydaje się to bardzo np. przy gromadzeniu bibliografii w internecie. Rekomendacje wikipedystów dotyczą m.in.: publikowanie poszczególnych pozycji z katalogu na osobnych stronach o czytelnym adresie (jedna pozycja posiada jeden osobny adres URL, który w razie potrzeby umieścić można np. w zestawieniu bibliograficznym w tekście pracy naukowej). Oczywiście tego typu katalog nie powinien być tworzony ręcznie – w przypadku kilku tysięcy pozycji byłoby to zupełnie nieracjonalne. Na szczęście można to zrobić prawie automatycznie, korzystając z popularnych systemów CMS. Wikipedyści zalecają także czytelne udostępnianie wszystkich metadanych opisujących dokumenty przechowywane w archiwum.

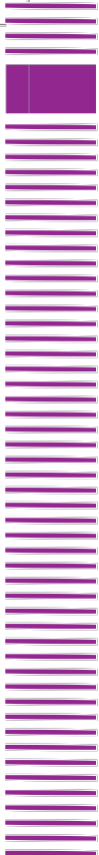
Indeksowanie przez wyszukiwarki. Ważne jest, aby katalog mógł być prawidłowo i efektywnie indeksowany przez wyszukiwarki takie jak np. Google. To właśnie z nich strona archiwum będzie miała potencjalnie najwięcej wejść. Należy zadbać, aby w kodzie każdej pojedynczej strony opisującej pojedynczy obiekt archiwalny (dokument, zdjęcie, nagranie relacji świadka historii itp.) pojawiły się odpowiednie metadane (zwłaszcza pole title i description, definiujące jakie informacje pojawiać się będą w wynikach wyszukiwania) oraz zapis zezwalający robotom wyszukiwarek na indeksowanie stron. Informację o tym, jak to zrobić, znaleźć można na wielu dostępnych w internecie kursach tworzenia stron WWW.

1.3 Internetowe repozytorium archiwalne

Archiwum społeczne może wykorzystać internet nie tylko do podstawowego informowania o swojej działalności i swoich zbiorach (strona informacyjna, katalog), ale także udostępniać publicznie online swoje zbiory w formie odpowiednio przygotowanych skanów.

Jakość skanów. Wspomniane wcześniej zalecenia Wikipedystów dla instytucji GLAM dotyczą także jakości zdigitalizowanych wersji archiwaliów: powinna być ona na tyle wysoka, aby można je było wykorzystać do pracy naukowej czy zastosowań edukacyjnych (dla fotografii to m.in. 1000px dłuższej krawędzi przy rozdzielczości 72 DPI). W bibliotekach, archiwach czy repozytoriach cyfrowych stosuje się jednak rozmaite strategie co do formy technicznej udostępnianych skanów. W repozytorium Wikipedii (Wikimedia Commons) znaleźć można cyfrowe reprodukcje w plikach o wysokiej jakości obok materiałów o jakości wyłącznie poglądowej. W przygotowanym przez Narodowe Archiwum Cyfrowe serwisie szukajwarchiwach.pl użytkownik ma dostęp do skanów o profesjonalnej jakości, dodatkowo skorzystać może z narzędzi ułatwiających przeglądanie ich treści. W polskich regionalnych bibliotekach cyfrowych zrzeszonych w Federacji Bibliotek Cyfrowych skany publikowane są głównie w formacie djvu, który gwarantuje niewielką objętość pliku przy zachowaniu jakości pozwalającej na podstawowe odczytanie treści – ze względu na wysoki stopień kompresji nie nadają się one już jednak np. do druku.

Wybór jakości skanów publikowanych online jest wypadkową wielu czynników, począwszy od wielkości kolekcji przez możliwości hostingowe, wydajność skryptu zarządzającego wyświetlaniem zbiorów aż po politykę



archiwum, która może ograniczać dostępność do dobrej jakości plików ze względu np. na prawa autorskie czy chęć zachowania kontroli nad rozpowszechnianiem zbiorów. Trudno tutaj mówić o najlepszych rozwiązaniach – zawsze zależą one od sytuacji archiwum czy charakteru zbiorów. Na pewno jednak społecznie korzystne jest jak najszersze udostępnianie zbiorów w jak najlepszej jakości skanach, szczególnie, jeśli przeszkodą nie są prawa autorskie (czyli w przypadku dokumentów z domeny publicznej). Takie podejście cechuje instytucje archiwalne i biblioteki współpracujące w ramach GLAM-Wiki.

Rozwiązaniem najbardziej optymalnym z perspektywy użytkowników archiwów, jednak równocześnie najbardziej pracochłonnym i skomplikowanym z perspektywy samego archiwum, jest udostępnianie oprócz skanów dokumentów także ich transkrypcji. Takie rozwiązanie pozwala łatwiej przeszukiwać zbiory archiwum pod względem wybranych treści. Użytkownicy mogą wykorzystać nie tylko standardowy opis zbudowany za pomocą metadanych (informujących np. o dacie wytworzenia dokumentu czy charakteryzujących jego treść), ale także odpowiednio przepisane do formatu tekstowego treści dokumentów czy nawet relacji historii mówionej. Przykładem archiwum udostępniającego zbiory w ten sposób jest brytyjski projekt London Lives (<http://www.londonlives.org/>). Użytkownicy tego archiwum cyfrowego, gromadzącego źródła do historii kryminalnej dawnego Londynu, mają do dyspozycji nie tylko skany, ale także przeszukiwalną pełnotekstową bazę danych z pełną transkrypcją. W tym przypadku nie polega ona na prostym sczytaniu i przepisaniu treści dokumentu. Jest ona opisana za pomocą odpowiednich tagów semantycznych, pozwalających m.in. na oznaczenie przekreśleń, błędów czy udostępnienie informacji o wzmiankowanych w dokumentach postaciach. W przypadku tworzonych oddolnie archiwów społecznych prostą wersję tekstową przygotować można za pomocą standardowego oprogramowania OCR, które jednak najbardziej efektywnie odczytuje pismo maszynowe. Niezbędna jest także korekta uzyskanego w ten sposób zapisu.

Praca nad transkrypcją dokumentów z archiwum społecznego może być częścią projektu edukacyjnego i odbywać się online. Wirtualne laboratorium transkrypcji (<http://wlt.synat.pcss.pl/wlt-web/index.xhtml>) to stworzone przez Poznańskie Centrum Superkomputerowo-Sieciowe darmowe narzędzie, które pozwala tworzyć przeszukiwalny, cyfrowy tekst na podstawie zeskanowanych dokumentów historycznych. Wykorzystać je można do wspólnej pracy nad przepisywaniem treści skanów rękopisów czy archiwalnych wydań książek i prasy lokalnej.

Dostępność. Sposobem ograniczenia dostępności publikowanych online skanów są znaki wodne. Powszechnie stosują je biblioteki czy archiwa, które w taki sposób chcą potwierdzić swoje prawo do zarządzania udostępnianymi zbiorami. Znaki wodne utrudniają korzystanie ze skanów (często dodawane są automatycznie i mogą zasłaniać duży fragment treści), a ich stosowanie nie zawsze jest uprawnione (jak w przypadku dokumentów z domeny publicznej).

Należy wystrzegać się wszystkich rozwiązań pozwalających pozornie zabezpieczyć publikowany online skan przed pobraniem na dysk twardy komputera. Blokowanie menu kontekstowego myszy czy wyświetlanie dokumentów za pomocą animacji flashowych utrudnia przeglądanie zbiorów, a na pewno nie jest skutecznym zabezpieczeniem.

Informacja o zasadach korzystania ze skanów. Ważnym elementem re-

pozytorium powinna być informacja o zasadach, na jakich można korzystać z udostępnianych zbiorów. Pojawić tu się powinien zapis o prawach autorskich (albo ich braku w przypadku zasobów z domeny publicznej) oraz dane kontaktowe, które pozwolą użytkownikowi zgłosić konieczność uzupełnienia lub poprawienia opisu dokumentu. Informując o prawach autorskich należy wystrzegać się zapisów, które mogą wprowadzać w błąd użytkowników. Jeśli udostępniamy zbiory z domeny publicznej, są one własnością publiczną i informowanie o tym, że mamy do nich wyłączne prawa jest nieporozumieniem,

chyba że jako pierwsi publikujemy utwór (patrz artykuł: *Udostępnianie materiałów archiwalnych*). Informacje o statusie prawnoautorskim udostępnianego dokumentu archiwalnego i zasadach jego rozpowszechniania i wykorzystania zapisać można w odpowiednich polach metadanych.

Gotowe rozwiązania. Budując repozytorium cyfrowe dla archiwum społecznego skorzystać można z darmowego oprogramowania takiego jak CollectiveAccess (<http://www.collectiveaccess.org/>) czy Omeka (Omeka.org). CollectiveAccess składa się z dwóch modułów – pierwszy odpowiada za wewnętrzne katalogowanie zbiorów, drugi – opcjonalny, pozwala publikować katalog i skany online. Omeka i CollectiveAccess wspierają najpopularniejsze standardy metadanych (w tym popularny w polskich bibliotekach cyfrowych Dublin Core), Omeka dodatkowo pozwala na eksport metadanych opisujących zbiory za pomocą protokołu OAI-PMH (*Open Archives Initiative Protocol for Metadata Harvesting*) – twórcy CollectiveAccess zapowiadają wprowadzenie tej funkcji już wkrótce. Pozwala ona podłączyć katalog archiwum społecznego do ogólnopolskiej wyszukiwarki Federacji Bibliotek Cyfrowych i przez nią także do Europeany – biblioteki europejskiej. Integracja z tymi serwisami przebiega jedynie na poziomie metadanych i nie ma wpływu na działanie repozytorium archiwum.

przykład: *Wdrożenie darmowego narzędzia.* Katalog archiwum dworu Saari (http://saarenkartano.muurututkimus.fi/index.php/Main_Page), prowadzonego przez organizację pozarządową Kone Foundation z Finlandii, zbudowany jest w modelu wiki. Każdy dokument posiada tam własną podstronę, która może być edytowana przez zarejestrowanych użytkowników. Dodawanie czy uzupełnianie informacji o obiekcie archiwalnym jest bardzo proste, dostępna jest także wyszukiwarka i system kategorii porządkujących katalog. To dobry przykład na skuteczne archiwalne wdrożenie darmowego narzędzia (MediaWiki), które obsługiwać można bez konieczności znajomości programowania.

przykład: *DSpace to popularne narzędzie do tworzenia i zarządzania repozytoriami cyfrowymi, rozwijane w ramach współpracy między firmą Hewlett Packard a MIT (Massachusetts Institute of Technology).* DSpace wykorzystywane być może także do budowy archiwum cyfrowego, czego doskonałym przykładem jest serwis Travellers in the Middle East Archive (TIMEA), publikowany przez amerykański Rice University (<http://timea.rice.edu/>). Oprogramowanie wspiera Dublin Core jako standard metadanych oraz pozwala na rozszerzone przeszukiwanie zasobów czy automatyczne importowanie dużych kolekcji dokumentów. Ważną zaletą DSpace jest też wspieranie dbałości o integralność udostępnianych plików za pomocą systemu tzw. sum kontrolnych.

przykład: *AtoM to oprogramowanie pozwalające na stworzenie profesjonalnego katalogu zbiorów archiwalnych oraz repozytorium udostępniającego skany. Program w opublikowanej w listopadzie 2013 roku wersji 2.0 umożliwia prowadzenie dokumentacji zbiorów z wykorzystaniem najważniejszych archiwalnych standardów metadanych – projekt wspiera merytorycznie Międzynarodowa Rada Archiwalna. AtoM może być uruchamiany w dwóch trybach: na lokalnym komputerze jako wewnętrzne narzędzie do zarządzania archiwaliami lub jako serwis WWW, pozwalający każdemu na przeszukiwanie informacji o dostępnych w archiwum zbiorach. Warto dodać, że opisy archiwaliów mogą być w AtoMie w prosty sposób importowane i eksportowane, a administrator może także regulować dostęp do określonych zasobów. Przy wprowadzaniu danych do opisu archiwalnego system generuje automatyczne podpowiedzi na bazie najczęściej powtarzających się fraz. AtoM wykorzystywać można do budowania katalogów i repozytoriów archiwalnych współtworzonych przez kilka instytucji.*

Narodowe Archiwum Cyfrowe zbudowało system ZoSIA (Zintegrowany System Informacji Archiwalnej), który również potencjalnie wykorzystywać można do udostępniania zbioru archiwum społecznego. Chociaż system ten przeznaczony jest przede wszystkim dla archiwów państwowych, korzystają z niego także – na zasadach specjalnych umów – organizacje pozarządowe (Ośrodek KARTA, Fundacja Dziedzictwa i Kultury Ormian Polskich). Skorzystanie z systemu ZoSIA wymaga przejścia odpowiednich szkoleń przez przedstawicieli archiwum społecznego, jakość skanów i ich opis muszą zostać zatwierdzone przez administratorów z NAC.

Metadane. Przy projektowaniu systemu archiwum należy wybrać także odpowiedni standard opisu metadanych. Nie jest to sprawa prosta, bo do wyboru jest przynajmniej kilka rozmaitych schematów wykorzystywanych już w wielu instytucjach gromadzących zbiory o charakterze archiwalnym. Twórcy archiwum społecznego do dyspozycji mają standardy archiwalne (np. ISAD(G)) i biblioteczne (np. Dublin Core, powszechnie wykorzystywany do opisów zbiorów w polskich bibliotekach cyfrowych). Często już wybór systemu udostępniania zbiorów determinuje wybór schematu opisu – np. w przypadku wymienionej wyżej Omeki publikowane materiały opisywać możemy jedynie według schematu Dublin Core. Warto pamiętać też, że przygotowując metadane należy respektować zasady konstruowania ich treści: korzystać z odpowiednich skrótów, dbać o prawidłową formę zapisu daty itp. Dokumentacja standardów metadanych znaleźć można bez trudu w internecie.

2. BEZPIECZEŃSTWO ZBIORÓW CYFROWYCH

W dwóch rozdziałach tego podręcznika (patrz: *Przechowywanie dokumentacji aktywnej* i *Przechowywanie fotografii*) poruszony został problem odpowiedniego przechowywania zbiorów archiwalnych. W części poświęconej archiwum społecznemu w internecie również pojawić się powinien ten wątek, jednak już w nieco innym kontekście. Obok dbałości o stan i bezpieczeństwo zbiorów archiwalnych w ich fizycznej postaci należy również odpowiednio dbać o zabezpieczenie zbiorów cyfrowych – skanów, które w ramach działań digitalizacyjnych powstają w archiwum.

W praktyce działań archiwalnych prowadzonych w ramach Cyfrowych Archiwów Tradycji Lokalnej bywa tak, że archiwum społeczne skorzystać może z

dokumentu historycznego w jego pierwotnej fizycznej postaci jedynie przez krótki czas wymagany na przygotowanie skanu – następnie materiał ten jest zwracany właścicielowi. Archiwum pozostaje zatem wyłącznie z cyfrową wersją dokumentu nie mając gwarancji tego, że w przyszłości możliwy będzie jeszcze dostęp do analogowego oryginału. Jeśli tak, to profesjonalne wykonanie cyfrowej kopii oraz jej późniejsze zabezpieczenie staje się kluczowym wyzwaniem.

Problem zabezpieczania zasobów cyfrowych (digital preservation) jest złożony i wynika z całego katalogu zagrożeń.

Nietrwałość nośników. Integralność zbiorów cyfrowych i możliwość ich odczytania ogranicza przede wszystkim nietrwałość nośników. Płyty CD/DVD, wciąż popularne jako forma przechowywania danych, podatne są na zarysowania i inne mechaniczne zniszczenia. Jakość materiałów, z których obecnie są wytwarzane sprawia, że trudno uznać je za bezpieczną formę długoterminowego przechowywania zbiorów. Popularne dziś nośniki takie jak pendrive czy dysk zewnętrzny podłączany do portu USB komputera także narażone są na zniszczenie np. w przypadku zalania czy upadku. W tego typu nośnikach nawet niewielkie uszkodzenie powodować może, że zapisane dane przestaną być odczytywalne.

Rozwój sprzętu komputerowego i oprogramowania. Komputery i nośniki danych ewoluują. Dobrym tego przykładem są dyskietki, z których nikt już dziś nie korzysta, a dostępne na rynku komputery od lat produkuje się bez napędów pozwalających je odczytać. Nie mamy pewności, czy za kilka lat nie pojawi się nowa technologia zapisywania, przechowywania i odczytywania danych, która wyprze obecne dziś rozwiązania.

Podobnie jak sprzęt komputerowy i nośniki danych, ewoluuje także oprogramowanie. Zapisywane na nośnikach dane muszą być odczytywane w odpowiedni sposób – jeśli niedostępne są już programy, za pomocą których je zapisywano, mogą pojawić się problemy w korzystaniu z treści przechowywanych plików.

W jaki zatem sposób można efektywnie i bezpiecznie przechowywać zbiory cyfrowe? Po pierwsze: należy zadbać o dywersyfikację nośników i przygotować kopie swojego zasobu np. na dwóch dyskach zewnętrznych, przechowywanych w różnych miejscach oraz udostępnić je do przechowywania w tzw. chmurze, np. w usłudze Dropbox, na serwerach wirtualnych (np. w ramach odpowiednio zabezpieczonych kont FTP). Wiele z usług zdalnego przechowywania danych jest bezpłatnych, a zarządzanie plikami nie sprawia kłopotu. W przypadku Dropboxa po założeniu konta i instalacji aplikacji, na dysku twardym komputera pojawia się specjalny katalog. Wszystkie umieszczane w nim pliki automatycznie (w tle) wysyłane są na internetowe konto konkretnego użytkownika, na podobnej zasadzie aktualizowane są wszystkie zmiany robione w tych plikach. Po archiwizacji należy regularnie sprawdzać stan nośników, integralność zgromadzonych na nich danych, w razie potrzeby aktualizować formaty do aktualnie obowiązujących jako standardowe.

Skomplikowana struktura zbiorów cyfrowych. Zabezpieczanie zbiorów cyfrowych to także problem związany z ich organizacją i zarządzaniem. Cyfrowe zasoby archiwum przechowywane mogą być w różnych formatach, na różnych nośnikach i odczytywane za pomocą rozmaitych programów. Dobra organizacja archiwum cyfrowego oznacza także wdrożenie odpowiednich zasad porządkujących organizację zbiorów. Warto ustalić odpowiednie standardy co

do wyboru nośników, sposobu archiwizacji i zabezpieczania oraz zasad udostępniania kopii.

Dbanie o integralność i bezpieczeństwo danych konieczne jest na poziomie ich gromadzenia, zarządzania (przetwarzania, kopiowania, katalogowania), przechowywania oraz udostępniania.

Aby kontrolować integralność zgromadzonych danych, tzn. sprawdzać, czy z powodu jakiegoś błędu nośnika ich treść nie została zmodyfikowana i nie może przez to być poprawnie odczytana, wykorzystać można pliki tzw. sum kontrolnych. Generowane są one na bazie treści zabezpieczanych plików - kiedy przy każdorazowym sprawdzeniu generowana jest nowa suma kontrolna, porównywana jest ona z pierwotną, co pozwala sprawdzić, czy plik na pewno nie został zmodyfikowany. Do generowania i sprawdzania sum kontrolnych wykorzystać można darmowy program Rapid CRC.

Stosowanie sum kontrolnych powinno być uzupełnieniem dla inwentarza zbiorów cyfrowych, określającego, jakie dane, na jakich nośnikach i w jaki sposób zostały zarchiwizowane i są przechowywane.

Zróznicowanie formatów danych. Liczba formatów plików interesujących nas z punktu widzenia społecznego archiwum cyfrowego wciąż się zwiększa. Tylko do plików dźwiękowych upowszechniły się przynajmniej cztery podstawowe formaty (wav, mp3, ogg/vorbis, wma), różniące się stopniem kompresji i dostępną jakością dźwięku – w przypadku plików graficznych czy video dostępnych formatów jest jeszcze więcej. Do tego dodać można także wiele różnych standardów kompresji plików w archiwa (np. rar, zip, gzip, 7z, tar) Projektując archiwizację cyfrowych zasobów archiwum społecznego należy mieć świadomość, że przechowywanie danych w postaci niepopularnych, nieposiadających wsparcia producentów oprogramowania czy społeczności użytkowników formatów może skutkować problemami w ich odczytaniu w przyszłości. Zaleca się, aby w projektach archiwizacyjnych wykorzystywać pliki o otwartym formacie – będą one odczytywalne bez względu na system operacyjny komputera. Ich pełna dokumentacja jest powszechnie dostępna, nie są też obwarowane żadnymi patentami ani licencjami. Przykładowo, dla tekstowych dokumentów biurowych takim standardem jest .odt, zalecany przez Międzynarodową Organizację Normalizacyjną (International Organization for Standardization, ISO). Użytecznym otwartym formatem archiwizacyjnym jest szczególnie format plików PDF - PDF/A, przystosowany specjalnie do długoterminowego przechowywania danych elektronicznych (także fotografii) i również zalecany przez ISO. Na koniec warto dodać jeszcze rzecz oczywistą: treść każdego skanu przeznaczzonego do archiwizacji i funkcjonowania jako tzw. master file powinna być zapisana w maksymalnej dostępnej jakości i w formacie, który ją gwarantuje. Z tego też powodu stosowanie plików jpg czy mp3 nie jest prawidłowe (pliki te powstają w wyniku kompresji danych).

Oprócz plików zawierających skany czy nagrania relacji historii mówionej archiwizować należy także metadane (opisy zbiorów) - można robić to archiwizując całą bazę danych, w oparciu o którą funkcjonuje nasz internetowy katalog czy repozytorium.

3. PROMOCJA

Chociaż internet to ważna przestrzeń informowania o archiwum społecznym i promowania jego zbiorów, nie można w planowanej komunikacji ograniczać się wyłącznie do niego. Jeżeli archiwum działa w środowisku lokalnym i gromadzi

zbiory do niego się odnoszące, skuteczniejsze od promocji w internecie może być przygotowanie plakatów, nawiązanie współpracy z lokalnymi mediami, parafią, władzami samorządowymi, biblioteką. Warto pamiętać, że nie wszyscy korzystają z internetu i opierając komunikację wyłącznie na tym medium w istotny sposób ograniczyć możemy jej zasięg.

Wikipedia. Ważnym elementem internetowej promocji powinna być Wikipedia, jako najpopularniejsze źródło wiedzy online. W przypadku Wikipedii – ze względu na encyklopedyczny charakter tego serwisu – promocja projektu archiwalnego polegać może wyłącznie na uzupełnianiu istniejących już haseł, tworzeniu nowych lub przekazywaniu skanów do repozytorium Wikimedia Commons. Tego typu współpraca podejmują już duże archiwa i biblioteki na całym świecie (m.in. amerykańskie archiwa narodowe, *National Archives and Records Administration* – NARA). Wikipedyści chętnie podejmują współpracę przy tego typu inicjatywach – budowana jest ona jednak zawsze w relacji między instytucją a krajową organizacją zarządzającą lokalną edycją Wikipedii. Polskie archiwa społeczne powinny kontaktować się ze Stowarzyszeniem Wikimedia Polska, mającym siedzibę w Łodzi. Jednym z warunków działania na Wikipedii – bez względu na to, czy działań podejmowanych nieformalnie (np. prostej edycji hasła na temat archiwum społecznego) czy oficjalnych relacji (umowa ze Stowarzyszeniem Wikimedia Polska) – jest udostępnianie wszelkich materiałów (informacji i skanów zbiorów) na wolnych licencjach albo wykorzystywanie jedynie tych, będących w domenie publicznej. Umieszczenie własnych zbiorów na Wikipedii pozwolić może na radykalny wzrost ich społecznego zasięgu, może być też rozwiązaniem podstawowych problemów związanych z infrastrukturą udostępniania skanów.

Przed przystąpieniem do jakichkolwiek działań na Wikipedii warto zapoznać się z jej podstawowymi celami i zasadami tworzenia, edytowania i udostępniania treści. Jedną z takich zasad jest np. encyklopedyczność (notability) – blokuje ona publikowanie haseł, na które nie powinno być miejsca w encyklopedii (przykładowo biogramów osób prywatnych). Tworząc hasło dotyczące naszego archiwum, powinniśmy go jak najbardziej uwiarygodnić – dobrym sposobem jest zamieszczanie odnośników do publikacji na jego temat (także pozainternetowych).

Media społecznościowe. Ważnym kierunkiem promocji archiwum w internecie mogą być media społecznościowe (np. Facebook, Twitter, Google Plus). Skuteczne działania w tej przestrzeni wymagają jednak dość dużego zaangażowania czasu i środków oraz odpowiednich strategii wobec nieustannie zmieniającego się środowiska tych serwisów (co jest szczególnie widoczne w przypadku Facebooka). Obecność w serwisach społecznościowych pozwala jednak dotrzeć do nowych odbiorców oferty archiwum społecznego i może mieć duży potencjał popularyzatorski. Nie zastąpi jednak na pewno działań promocyjnych na poziomie lokalnym.

Istnienie archiwum społecznego warto zasygnalizować instytucjom bliskim jego misji, np. regionalnemu archiwum państwowemu, bibliotekom, szkołom, lokalnej prasie czy historycznym serwisom i forom internetowym o zbliżonej tematyce. Warto udostępnić podstawowe informacje o archiwum do rozwijanej w Ośrodku KARTA ogólnopolskiej bazy danych tych instytucji

Na koniec zasygnalizować warto jeszcze jedną możliwość wykorzystania internetu w archiwum społecznym. W październiku 2013 roku Fundacja AriAri prowadząca projekty historii mówionej dotyczące okolic Włocławka zainicjowała internetową akcję zbierania funduszy na opracowanie i transkrypcję zgromadzonych nagrań. Tego typu działanie określa się mianem crowdfundingu – finansowania społecznościowego, w którym niewielkie wpłaty dużej liczby osób pozwalają w sumie na wsparcie wybranych inicjatyw. Realizowany w ramach portalu wspieram.to projekt Fundacji Ari Ari zakończył się co prawda niepowodzeniem (zebrano 570 z oczekiwanych 3 tys. złotych), jednak istnieją przykłady projektów archiwalnych czy digitalizacyjnych uzyskujących takie wsparcie (np. szkocki projekt uporządkowywania archiwum Hamisha Hendersona czy finansowanie digitalizacji książek z domeny publicznej w ramach portalu Wolnelektury.pl). Crowdfunding w Polsce czeka wciąż na odpowiednie regulacje ze strony Ministerstwa Administracji i Cyfryzacji (konieczna nowelizacja ustawy o zbiorach publicznych).

Trudno na kilku stronach choćby streścić wszystkie wątki funkcjonowania archiwum społecznego w internecie. Wiele zależy przecież od charakteru jego zbiorów, misji, jaką zdefiniowali dla siebie jego organizatorzy oraz możliwości finansowych i organizacyjnych. Inną perspektywę internetową będą miały duże archiwa społeczne powstające np. przy organizacjach pozarządowych, inną niewielkie oddolne i nieformalne inicjatywy archiwalne.

Wiedzę i inspiracje na temat prowadzenia archiwum społecznego w internecie znaleźć można na stronach portali takich jak: Digitalz.pl – dobra cyfrowe w teorii i praktyce, Archiwa.net – Naukowy Portal Archiwalny czy HistoriaiMedia.org – historia i dziedzictwo w kulturze cyfrowej.

Słownik wybranych terminów technicznych wykorzystanych w tekście

metadane - dane opisujące inne dane; w perspektywie archiwum społecznego metadanymi są merytoryczne opisy gromadzonych zbiorów oraz ich techniczne specyfikacje;

OAI PMH (Open Archives Initiative Protocol for Metadata Harvesting) - protokół pozwalający na wymianę metadanych między repozytoriami cyfrowymi, wykorzystywany m.in. w Federacji Bibliotek Cyfrowych oraz Europeanie.

repozytorium cyfrowe - w kontekście archiwum społecznego: internetowy system uporządkowanego przechowywania i udostępniania dokumentów oraz ich opisów (metadanych).

OCR (Optical Character Recognition) - automatyczne rozpoznawanie znaków, pozwalające na przeniesienie treści dokumentu ze skanu w postaci pliku graficznego do pliku tekstowego. OCR-owane skany udostępniać można za pomocą plików djvu.

digitalizacja, cyfryzacja - odwzorowywanie w postaci cyfrowej wizerunków (treści) dokumentów lub nagrań analogowych.

robots.txt - zapisywany na serwerze plik tekstowy definiujący, jakie wyszukiwarki w jaki sposób indeksować mogą zasoby publikowanego serwisu.

GLAM - skrót oznaczający branżę instytucji kultury i dziedzictwa: galerii, bibliotek, archiwów i muzeów (galleries, libraries, archives, museums).

GLAM-Wiki - projekty współpracy między instytucjami kultury a środowiskiem Wikipedystów, zmierzające do zwiększenia obecności zbiorów tych instytucji w Wikipedii i w repozytorium Wikimedia Commons.

wiki / mediawiki - model serwisu internetowego, w którym edycja treści poszczególnych stron możliwa jest w łatwy i szybki sposób z poziomu przeglądarki oraz przez wielu użytkowników. Przykładem wdrożenia tego modelu jest Wikipedia, której hasła może tworzyć i poprawiać każdy.

CMS (Content Management System) - system zarządzania treścią serwisu internetowego, pozwalający na łatwe dodawanie i modyfikowanie treści oraz tworzenie jej kategorii i spisów (taksonomii bez konieczności znajomości programowania).

dostępność (web accessibility) - cecha serwisu internetowego określająca możliwość korzystania z niego przez jak najszersze grono użytkowników, w tym przez osoby niepełnosprawne, używające np. specjalnego oprogramowania czytającego strony WWW. Zasady tworzenia dostępnych stron internetowych określa m.in. standard Web Content Accessibility Guidelines (WCAG) 2.0.

użyteczność (usability) - miara wydajności, efektywności i satysfakcji użytkownika z jaką dany serwis internetowy może być używany do osiągnięcia określonych przez danych użytkowników w danym kontekście (norma Norma ISO 9241); przystępność, łatwość i efektywność używania serwisu internetowego.

chmura - w kontekście archiwum społecznego np. usługa przechowywania kopii zbiorów cyfrowych na zdalnym serwerze o określonych parametrach, pozwalająca je dodatkowo zabezpieczyć bez konieczności instalowania dodatkowego rozbudowanego oprogramowania czy zwiększania parametrów komputera lokalnego. Przykładem usługi przechowywania danych w chmurze jest Dropbox lub Gmail.

serwer - w perspektywie serwisu internetowego archiwum społecznego: konieczna infrastruktura (oprogramowanie i komputer) pozwalająca na uruchomienie ogólnodostępnej strony internetowej, zazwyczaj kupowana w ramach usługi zewnętrznej (np. serwer wirtualny).

Artykuł jest częścią publikacji pt. „Archiwistyka Społeczna”, którego pełna wersja dostępna jest pod adresem: 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, 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.