

Międzynarodowy Tydzień Archiwów Seminarium internetowe, 4-7 lipca 2024

Międzynarodowy Tydzień Archiwów (International Archives Week, IAW) to siedmiodniowe święto globalnej społeczności archiwalnej, zorganizowane po raz pierwszy w 2019 r. w związku z jubileuszem 60-lecia utworzenia Międzynarodowej Rady Archiwów (International Council on Archives, ICA, dalej: MRA). Jego punkt kulminacyjny przypada 9 czerwca w Międzynarodowy Dzień Archiwów (International Archives Day, IAD).

Tematem przewodnim tegorocznych obchodów Międzynarodowego Tygodnia Archiwów, wybranym przez członków organizacji pod koniec 2023 r., stały się #CyberArchiwa (#CyberArchives). Specjaliści z zakresu archiwistyki i zarządzania dokumentacją zaprosili wszystkich zainteresowanych do wirtualnej rozmowy o wpływie postępu technologicznego na dziedzinę archiwalną – dbającą o przeszłość, osadzoną w teraźniejszości, ale przede wszystkim nastawioną na przyszłość. Podkreślili, że w szybko ewoluującym krajobrazie cyfrowym archiwa mają szansę stać się podmiotami wytwarzającymi i skutecznie wdrażającymi nowe rozwiązania oparte na sztucznej inteligencji, uczeniu maszynowym i innych zaawansowanych technologiach, a archiwiści jako kustosze pamięci powinni odgrywać wieloaspektową rolę w promowaniu demokracji, ochronie praw obywateli i poprawie jakości życia. Istotę przedsięwzięcia, pobudzającego sektor publiczny i prywatny do podjęcia niezbędnych działań na rzecz przechowywania materiałów archiwalnych w sposób bezpieczny i trwały, prezydent MRA Josée Kirps określiła następująco: „To coś więcej niż temat. To wezwanie do działania, zapraszające nas do zagłębienia się w transformacyjną moc technologii w naszym zawodzie. Wdrażając nowe technologie, nie tylko dotrzymujemy kroku zmianom, ale też kształtujemy przyszłość. Umożliwiamy społeczeństwu podejmowanie świadomych decy-

zji, wspieramy przejrzystość i zachowujemy nasze bogate dziedzictwo kulturowe dla przyszłych pokoleń”¹.

Wirtualną rozmowę o charakterze interdyscyplinarnym napędzały otwarte i bezpłatne seminaria internetowe. Rozbudowany i różnorodny program webinariów składał się z siedmiu głównych sesji tematycznych, w tym pięciu anglojęzycznych, które zaprezentowano poniżej.

Podczas zorganizowanego 4 czerwca „FAN Virtual Colloquiums on Archival Matters 2024” Carlos Zapata Cárdenas – obecnie główny doradca ds. zarządzania informacją w kolumbijskim Banco de la República i członek grup eksperckich Iberoamerykańskiej Sieci Uniwersyteckiej Edukacji Archiwalnej Latinoamerykańskiego Stowarzyszenia Archiwów (Red Iberoamericana de Enseñanza Archivística Universitaria de la Asociación Latinoamericana de Archivos) – skupił się na poszukiwaniu odpowiedzi na najbardziej nurtujące pytania współczesnej archiwistyki – o możliwości, ograniczenia i zagrożenia związane z zastosowaniem sztucznej inteligencji w działalności archiwalnej. Czy archiwiści są gotowi na odpowiedzialne i profesjonalne wykorzystanie sztucznej inteligencji? Czy będą w stanie zastosować ją w takim samym tempie, w jakim się rozwija, bez narażania na ryzyko pamięci dokumentacyjnej i samego zawodu? Jak chronić autentyczność, integralność i wiarygodność elektronicznych dokumentów archiwalnych przy jednoczesnej ochronie zaufania publicznego? To jedynie niektóre z poruszanych przez prelegenta kwestii, stanowiących wzywanie dla społeczności archiwalnej. Przyznać należy, że pomimo optyki właściwej Ameryce Łacińskiej, kwestie podniesione na forum miały charakter uniwersalny. W dyskusji podjętej przez uczestników webinarium wyraźnie wybrzmiewał apel o konieczność wprowadzenia polityki standaryzacji szczegółowych informacji identyfikujących dokumenty elektroniczne (metadanych) dla wszystkich gałęzi sektora publicznego oraz zabezpieczenia interoperacyjności systemów teleinformatycznych, w których tworzone są i przechowywane materiały archiwalne. Zapata Cárdenas poruszył także zagadnienie użycia sztucznej inteligencji w działalności archiwalnej w erze transformacji cyfrowej, zwanej przez niego rewolucją, i masowego powstawania dokumentacji elektronicznej. Wskazanie przez niego obszarów szczególnie trudnych – ograniczeń o podłożu ekonomicznym i technologicznym oraz braku nowoczesnego programu kształcenia uniwersyteckiego przyszłych kadr – zmusiło

¹ International Council on Archives. International Archives Week. #IAW 2024, <https://www.ica.org/international-archives-week/international-archives-week-iaw2024/> [dostęp: 01.07.2024].

uczestników do głębszej refleksji nad przyszłością profesji archiwisty nie tylko w Iberoameryce, ale także w krajach, które dotychczas nie znalazły własnych sposobów na poradzenie sobie z wyzwaniami teraźniejszości.

Podbudowa teoretyczna przedstawiona przez Cárdenasa była wstępem do praktycznych rozważań o synergii archiwów i sztucznej inteligencji. Te podjęły w szczególności przedstawicielki archiwów Indonezji i Norwegii.

Wiwi Diana Sari, reprezentująca Narodowe Archiwum Republiki Indonezji (Arsip Nasional Republik Indonesia, ANRI), przybliżyła systemy wykorzystujące sztuczną inteligencję i technologie cyfrowe do usprawnienia procesów archiwizacji, umożliwiające wydajną digitalizację dokumentacji i ekstrakcję metadanych oraz zapewniające dostęp do narodowego dziedzictwa kulturowego. Działania podjęte przez Narodowe Archiwum Republiki Indonezji wpisują się w skodyfikowaną w siedmiu normatywach politykę rozwoju technologii cyfrowych w tym kraju. Dyrektywy prezydenta Joko Widodo z 2020 r. zobowiązały służbę archiwalną do przyspieszenia rozbudowy i poprawy infrastruktury cyfrowej. Długofalowy program transformacji technologicznej rozpisano na pięć etapów, a jego zakończenie zaplanowano na 2045 r. Faza obecna realizacji przedsięwzięcia obejmuje implementację Elektronicznego Systemu Rządowego (SPBE) z funkcjonalnościami aplikacji Zintegrowanego Dynamicznego Systemu Informacji Archiwalnej (SRIKANDI) oraz rejestrami programów Krajowy System Informacji Archiwalnej (SIKN) i Krajowa Sieć Informacji Archiwalnej (JIKN). Wdrożono ponadto projekt szkolenia sztucznej inteligencji do tworzenia opisów dokumentacji fotograficznej celem zarządzania rozproszonymi i niepełnymi danymi. Mówczyni przedstawiła przykład archiwum cyfrowego Sejarah Nusantara, dedykowanego badaczom historii Archipelagu Indonezyjskiego. Na zakończenie omówiła założenia innowacyjnego projektu ANRIpedia – łączącego powyższe systemy kompleksowego archiwum cyfrowego, stanowiącego transformacyjny krok w kierunku rozszerzenia dostępu do zasobów archiwalnych i ułatwiającego głębsze zrozumienie historii i kultury tego regionu geograficznego.

Kolejna prelegentka, Kristin Jacobsen, zaprezentowała myślenie projektowe w obszarze sztucznej inteligencji towarzyszące misji Archiwum Narodowego Norwegii (Arkivverket) od 2019 r. Dążenie do opracowania kreatywnych rozwiązań problemów archiwów cyfrowych skupiało się początkowo – jak podkreślała – na rozwijaniu procesu digitalizacji. W tym zakresie osiągnięto znaczne sukcesy. Szacuje się, że 90% wytworzonej na nośniku papierowym dokumentacji przeniesiono do formatu binarnego. Priorytetem norweskiej służby archiwalnej

pozostaje zwiększanie dostępności zdigitalizowanych materiałów archiwalnych. Służyć temu celowi ma prowadzone na szeroką skalę testowanie technologii przetwarzania języka naturalnego, uczenia maszynowego oraz rozpoznawania tekstu odręcznego. Jak zauważyła Jacobsen, rozpoczęto już prace pilotażowe nad rozpoznawaniem mowy i budowaniem modeli językowych. Wprowadzone na ich podstawie narzędzia wyszukiwawcze pozwalają na szybką ekstrakcję treści dokumentów i precyzyjne odnalezienie szukanej informacji. Prelegentka, odwołując się do zapisów zawartych w rejestrach prawnych nieruchomości, będących przedmiotem częstego wykorzystania, przywołała przykład automatyzacji procesu ochrony danych stanowiących informację niejawną. Omówiła rezultaty ekstrakcji treści z 30 milionów zdigitalizowanych kart spisowych ludności i stworzenia pełnego wykazu osób w nich ujętych. Archiwum Narodowe Norwegii podjęło ponadto prace nad stworzeniem cyfrowego archiwum fotograficznego, umożliwiające – dzięki wspomaganium sztucznej inteligencji – wyszukiwanie obrazu za pomocą haseł, co przynosi lepsze rezultaty niż wyszukiwanie poprzez metadane.

Seminarium internetowe „FAN Virtual Colloquiums on Archival Matters 2024” zamknęło wystąpienie prof. Basmy Makhoul Shabou, wykładowczyni szwajcarskiej Szkoły Biznesu i Administracji w Genewie (Geneva School of Business Administration). Kierowniczką projektu „Maturity Assessment for Appraisal in the AI Age”, realizowanego przez InterPARES (International Research on Permanent Authentic Records in Electronic Systems), przybliżyła główne założenia badań teoretycznych prowadzonych na styku archiwistyki i nowoczesnych technologii. Celem przedsięwzięcia, opracowywanego na poziomie akademickim, jest identyfikacja archiwalnych, technicznych, technologicznych, kulturowych i strategicznych barier stosowania sztucznej inteligencji w archiwach oraz czynników ułatwiających skuteczne stosowanie jej narzędzi w procesach oceny wartości dokumentacji. W ramach badania przygotowany zostanie „model oceny gotowości” ukazujący spójność i powtarzalność decyzji podejmowanych przez archiwistów w tym obszarze działalności, wymagania dotyczące integracji sztucznej inteligencji z daną praktyką wartościowania oraz dojrzałość procesów archiwizacyjnych do stosowania „inteligentnej” oceny dokumentacji.

Podczas kolejnego, odbywającego się 5 czerwca, webinarium pt. „Recent ICA Projects: A Showcase of Products, Projects, and Ideas funded by the ICA Programme Commission” miała miejsce godzinna prezentacja projektów finansowanych przez Komisję Programową MRA w ramach corocznego składania wniosków. Seria dynamicznych pokazów innowacyjnych pomysłów archiwalnych

(„International Survey – Understanding the landscape of Trauma and Archives”, „Online Four-Week (4) Foundational Course in Archives and Records Management for the Southeast Asian Archival Community”, „ICA SAHR Reference Library Zotero”, „Caribbean Cultural Emergency Response Hub Advocacy Meeting”, „What Does Diversity mean to you? An international and Online discussion”, „Archival Summer School for Ukraine”) pozwoliła na ogólne zapoznanie się z ich tematyką i założeniami. Bardziej szczegółowe poznanie idei projektowych było możliwe dzięki sesjom grupowym opartym na bezpośrednim kontakcie z liderami.

Mające miejsce 6 czerwca seminarium internetowe „Using the Universal Declaration on Archives as an Advocacy Tool” poprowadzili Margaret Crockett i Claude Roberto, reprezentanci grupy ekspertów MRA ds. Rzecznictwa (Expert Group on Advocacy, AEG). Wzajemnie uzupełniające się prezentacje uwypukliły znaczenie Powszechnej Deklaracji Archiwów (The Universal Declaration on Archives, UDA) jako dokumentu – jak wskazała Crockett – stanowiącego kamień węgielny działalności archiwów i archiwistów na całym świecie. Deklaracja, przyjęta przez MRA 17 września 2010 r., a następnie zaaprobowana 10 listopada 2011 r. przez UNESCO, w części wprowadzającej podkreśla, że archiwa, chroniąc pamięć indywidualną i zbiorową, odgrywają podstawową rolę w rozwoju społeczeństw. Wokół tego stwierdzenia zbudowano szerszą narrację, ukazującą naturę i wartość archiwaliów. Stają się one gwarantem odpowiedzialnych i przejrzystych działań, będących autentycznym świadectwem pracy administracyjnej, kulturalnej i intelektualnej, wspierających wiedzę obecną i przyszłą, a wreszcie – dzięki otwartemu dostępowi – promujących demokrację, chroniących prawa obywateli i podnoszących jakość życia. W wyróżnionych w manifeście punktach szczegółowo opisano cechy i znaczenie materiałów archiwalnych, ich różnorodność, formaty, wskazano na rolę archiwistów oraz zidentyfikowano interesariuszy ponoszących odpowiedzialność za zarządzanie dokumentacją. Kluczowym zapisem deklaracji jest, według Roberto, działanie rozumiane w kategoriach proaktywnego tworzenia polityki i praw dotyczących archiwaliów, promowania zasad zarządzania i przechowywania ich w sposób, który zapewnia im autentyczność, integralność i jak największą swobodę w zakresie wykorzystania. Eksperci apelowali o podjęcie zabiegów służących podnoszeniu świadomości i mobilizacji społeczeństw na rzecz ochrony dziedzictwa dokumentacyjnego, wskazując konkretne przykłady aktywności instytucji rządowych (podpisanie Powszechnej Deklaracji Archiwów przez autonomiczny parlament Katalonii, instytucje europej-

skie i amerykańskie), pozarządowych (lobbowane przez związki nauczycielskie rozszerzenie uniwersyteckich programów nauczania) i osób prywatnych (upowszechnianie przepisów wspomnianej deklaracji w mediach społecznościowych).

Sesje zaplanowane na 7 lipca zapoczątkowało webinarium „Digital Preservation 2049: The Past, Present and Futures of Cyber-Archives”. Dyrektor wykonawczy Koalicji Ochrony Cyfrowej (Digital Preservation Coalition, DPC) William Kilbride podzielił się na nim z uczestnikami spotkania refleksjami na temat ochrony zasobów cyfrowych wyłaniającymi się z jego 25-letniego doświadczenia zawodowego. Rozważania mówcy objęły także przyszłość – wskazania, rady i ostrzeżenia na kolejne ćwierćwiecze. Koalicję Ochrony Cyfrowej tworzy międzynarodowa społeczność, obecnie składająca się ze 159 organizacji z całego świata, która podjęła współpracę na rzecz trwałego, zrównoważonego i długoterminowego dostępu do treści i usług cyfrowych oraz zapewnienia bezpieczeństwa danych cyfrowych w zmieniających się warunkach (przestarzałość techniczna). Tytułowa ochrona informacji cyfrowych jest, zdaniem prelegenta, ciągłym wyzwaniem rzucanym środowisku archiwistów i zarządców dokumentacji, które – w zależności od kontekstu technologicznego, społecznego i organizacyjnego – za każdym razem na nowo definiuje środki zabezpieczające trwałą dostępność i czytelność zapisów cyfrowych. Jednym z nich jest, znajdująca licznych zwolenników w latach 90. XX w., migracja danych polegająca na ich sukcesywnym przenoszeniu z formatów starszych, wychodzących z użycia, do formatów nowej generacji. Konwersja danych może jednak powodować powstanie przekłamań i odstępstw od oryginalnej wersji, w związku z czym na początku obecnego stulecia postawiono na emulowanie ich pierwotnego otoczenia programowego, a więc na przeniesienie programów odtwarzających starsze platformy programowo-sprzętowe na te współcześnie wykorzystywane. Chronić długoterminowo można wreszcie, jak zauważył Kilbride, poprzez aktualizację dostępu do danych i procesów fizycznego środowiska obliczeniowego, w tym sprzętu i urządzeń peryferyjnych. Każde z rozwiązań będących odpowiedzią na zagrożenia świata cyfrowego pokazuje, że wyzwania związane z ochroną zasobów cyfrowych rodzą się zazwyczaj poza archiwum. W konsekwencji organizacje w coraz większym stopniu polegają na archiwistach, którzy na nowo odpowiadają na te same pytania. Jak zauważył prelegent, pomimo znacznego w ciągu ostatnich 25 lat postępu, kolejne 25 lat będzie nadal wymagało znaczącej i ciągłej transformacji. Tylko jedna teza pozostaje prawdziwa wtedy, teraz i w przyszłości: ochrona zasobów cyfrowych jest wspólnym wyzwaniem, które było, jest i będzie najlepiej realizowane poprzez współpracę.

Następne seminarium internetowe poświęcono tematowi „#CyberArchives and AI: Safeguarding and Enhancing Digital Information”. Archiwistka Eng Sengsavang zapoznała uczestników sesji z wynikami eksperymentów przeprowadzonych w archiwach UNESCO w obszarze wzbogacania metadanych i analizy dyplomatycznej wspomaganą przez sztuczną inteligencję, prowadzonych wspólnie z członkami grup badawczych InterPARES. Prelegentka skupiła się na innowacyjnych badaniach archiwalnych kolekcji nagrań dźwiękowych (audycje radiowe, wywiady, przemówienia, wydarzenia i muzyka) odzwierciedlających międzynarodową działalność UNESCO. W okresie 2017–2021 archiwa organizacji zdigitalizowały około 177 000 jednostek archiwalnych dokumentacji dźwiękowej wytworzonej w latach 50.–80. ubiegłego stulecia. Zbiór ten cechuje znaczne zróżnicowanie w zakresie języka (70 różnych języków państw członkowskich). Z uwagi na brak opisów zawartości umożliwiających zapoznanie się z ich treścią, podjęto się prób zautomatyzowania procesu inwentaryzacji. W tym celu, po pierwsze, zastosowano maszynowe modele rozpoznawania mowy oparte na transkrypcji i identyfikacji języka. W efekcie uzyskano bardzo dobre wyniki dla języków kongresowych (francuskiego i angielskiego) i przyjęto, że w odniesieniu do nich osiągnięto poziom akceptowalny pod względem dokładności i precyzji zapisów. Sztuczna inteligencja w znacznie mniejszym stopniu poradziła sobie z posiadającymi mniejsze zasoby modelami testowymi języków katalońskiego i walijskiego. Problemów nastręczyły także nagrania wielojęzyczne, częste w przypadku dokumentów wytwarzanych przez urzędników pracujących w organizacji międzynarodowej oraz rejestrujące wypowiedzi osób porozumiewających się w języku obcym z akcentem właściwym dla ich języka naturalnego. Po drugie, wykorzystano analizę dyplomatyczną – metodę badania cech wewnętrznych i zewnętrznych dokumentów. Bazując na niej, każde nagranie analizowane jest pod kątem struktury narracji i kategoryzowane gatunkowo. Przykładowo ustalono, że programy radiowe rozpoczynają się od wprowadzenia muzycznego, trwającego od 10 do 30 sekund, a wywiady otwiera prezentacja sylwetki gościa (imię, nazwisko, stanowisko) oraz wskazanie celu rozmowy. Elementy te uwzględniono w opisie archiwalnym.

Pilotażowy program wykorzystania skonfigurowanego chatbota w działalności archiwalnej opisał przewodniczący grupy roboczej UNESCO ds. sztucznej inteligencji Thomas Chanussot. Zwiększenie ilościowe i jakościowe rezultatów wyszukiwania zaprezentował w odniesieniu do konkretnych kolekcji dostępnych za pośrednictwem centralnego repozytorium dokumentów UNESDOC, groma-

dzącego około 180 000 plików zapisanych w formacie PDF. Jako próbę reprezentatywną tak dużego zbioru wybrano publikowane corocznie przez organizację raporty z monitoringu edukacji globalnej i dokonano ekstrakcji zawartych w nich danych. W konkluzji badań testowych stwierdzono, że strategia oparta na pracy z mniejszym materiałem przynosi lepsze efekty niż trenowanie bardzo dużych zbiorów informacji, których sztuczna inteligencja nie mogła zrozumieć, a zatem nie była w stanie udzielić wyczerpujących i dokładnych odpowiedzi na zadane pytania. Ponadto podkreślono przyszłość generatywnie wspomaganego systemu wyszukiwawczego jako narzędzia ułatwiającego użytkownikom bardziej zaawansowane badania i odkrywanie zasobów UNESCO.

Dyrektor Departamentu Archiwów w Bibliotece Narodowej i Archiwach (National Library and Archives, NLA) Zjednoczonych Emiratów Arabskich Hamad Al Mutairi podniósł kwestię ochrony informacji cyfrowych. Stanowiące temat przewodni Międzynarodowego Tygodnia Archiwów cyberarchiwa, zgodnie z definicją mówcy rozumiane jako dokumenty cyfrowe, szerzej – dane przechowywane elektronicznie lub online, wymuszają rewizję dotychczasowego podejścia do bezpieczeństwa informacji. W środowisku cyfrowym bezpieczeństwo jest budowane na trzech podstawowych filarach: poufności informacji, a więc pewności, że dane niejawnie są odpowiednio zabezpieczone, integralności (kompletności), wykluczającej wprowadzenie do nich zmian w sposób nieautoryzowany, oraz dostępności, która sprowadza do poziomu minimalnego lub (lepiej) całkowicie eliminuje prawdopodobieństwo nieuprawnionego działania przeciwko infrastrukturze komputerowej. Cyberataki, na których szczególnie skupił się prelegent, przynoszą niepowetowane szkody w zasobach informacyjnych, finansowych i wizerunkowych instytucji. W samych Zjednoczonych Emiratach Arabskich odnotowano 58 000 zamachów przeprowadzonych przez hakerów. W skali globalnej szacuje się, że dokonano ich prawie 8 milionów. Zaprezentowana przez Al Mutairiego mapa cyberataków w czasie rzeczywistym potwierdzała te dane. Szerokie tło funkcjonowania informacji cyfrowej w cyberprzestrzeni dało prelegentowi asumpt do przedstawienia polityki swojego kraju w dziedzinie cyberbezpieczeństwa. Strategia rządowa wdrożona i rozwijana w Zjednoczonych Emiratach Arabskich określa wizję zapewnienia bezpieczeństwa cybernetycznego poprzez: kodyfikowanie polityki i standardów w zakresie nowych technologii, wypracowywanie najlepszych praktyk obejmujących kontrolę szyfrowania danych, przeprowadzanie audytów wewnętrznych i zewnętrznych, rewidowanie programów tworzenia kopii zapasowych, dokonywanie testów penetracyjnych,

wdrażanie środków reagowania na incydenty, a nade wszystko szkolenie personelu własnego i podnoszenie świadomości zagrożeń związanych z cyberarchiwami.

Reasumując, seminaria internetowe towarzyszące Międzynarodowemu Tygodniowi Archiwów były ważną płaszczyzną wymiany poglądów środowiska na tematy istotne i aktualne, oscylujące wokół funkcjonowania archiwów w erze cyfrowej. Ich wartość, w moim odczuciu, podnosiło przywoływanie konkretnych przykładów wykorzystania najnowszych technologii, jak choćby sztucznej inteligencji, w działalności archiwalnej, a tym samym ukazanie interesujących perspektyw dla przyszłości zawodu. Zróżnicowane pod względem treści i reprezentacji prelegentów webinaria stały się przedmiotem ożywionych dyskusji, ukazujących sprzężenie zwrotne między MRA a globalną społecznością archiwistów. Badania nad cyberarchiwami zachęcały do stawiania dalszych hipotez, tworzenia nowych projektów, prowokowały do myślenia perspektywicznego o archiwach i ich misji. Z całą pewnością, zgodnie z życzeniem prezydent MRA Josée Kirps, stały się wyzwaniem, które należy podjąć dla kształtowania przestrzeni archiwalnej przy pomocy nowoczesnej technologii i wypełnienia jej bogatą i różnorodną treścią cyfrowego dziedzictwa kulturowego.

Kamila Pawełczyk-Dura

Naczelna Dyrekcja Archiwów Państwowych / The State Archives Head Office (Poland)
kpawelczyk@archiwa.gov.pl, ORCID 0000-0002-3615-8120