



Fundusze Europejskie
na Rozwój Cyfrowy



Rzeczpospolita
Polska

Dofinansowane przez
Unię Europejską



200 lat
BIBLIOTEKI
KÓRNICKIEJ

1. WCAG 2.2 – technologie, narzędzia, doświadczenia

Pytanie:

Jakie technologie i rozwiązania rekomendują Państwo do zapewnienia zgodności interfejsu użytkownika z WCAG 2.2 AA w systemach bibliotecznych? Czy istnieją narzędzia usprawniające wdrożenie i utrzymanie dostępności na tym poziomie?

Odpowiedź:

Rekomendujemy podejście oparte na całkowitym oddzieleniu warstwy prezentacyjnej (frontend) od logiki biznesowej (backend) oraz wykorzystanie otwartych standardów webowych. Zgodność z WCAG 2.2 AA wymaga nie tylko technologii, ale i odpowiedniego podejścia już na etapie projektowania systemu (UX/UI). Projekty powinny być w pełni responsywne. Należy też pamiętać, że utrzymanie dostępności i zgodności z WCAG jest procesem ciągłym. Rekomendujemy wdrożenie hybrydowej strategii testowania z wykorzystaniem testów automatycznych (CI/CD), wtyczek w przeglądarkach (np. WAVE), czytników ekranowych oraz walidatorów kodu. Ponadto w skład naszego zespołu wchodzi osoby przeszkolone z zakresu weryfikacji zgodności serwisów internetowych z WCAG.

2. Integracja z Węzłem Krajowym (login.gov.pl / mObywatel)

Pytanie:

Jakie rozwiązania umożliwiają bezpieczną i efektywną integrację systemu bibliotecznego z Węzłem Krajowym przy zachowaniu obecnego SSO opartego na Keycloak? Na co zwrócić uwagę przy takiej integracji (np. wydajność, bezpieczeństwo, doświadczenia użytkownika)?

Odpowiedź:

Biorąc pod uwagę, że Państwa środowisko już wykorzystuje rozwiązanie Keycloak jako dostawcę tożsamości, najbezpieczniejszym i najefektywniejszym modelem integracji będzie wykorzystanie mechanizmu Identity Brokering. W tym modelu system biblioteczny nie łączy się bezpośrednio z Węzłem Krajowym. Zamiast tego, Keycloak odpowiada za komunikację między systemem bibliotecznym a Węzłem Krajowym.

W kontekście bezpieczeństwa wymagane jest skonfigurowanie dwustronnego zaufania. Keycloak musi podpisywać żądania wysyłane do Węzła Krajowego odpowiednim certyfikatem, a odpowiedzi z Węzła powinny być weryfikowane pod kątem podpisu cyfrowego. Należy zadbać o bezpieczne mapowanie unikalnego identyfikatora z Węzła Krajowego na konto użytkownika w systemie bibliotecznym. Rekomendujemy, aby Keycloak nie przechowywał trwale nadmiarowych danych



Zamek w Kórniku
ul. Zamkowa 5, 62-035 Kórnik
tel. (+48) 61 817 00 81
tel. (+48) 531 998 263
sekretariat.zamek@bk.pan.pl

Pałac Działyńskich
Stary Rynek 78/79, 61-772 Poznań
tel. (+48) 61 852 48 44
tel. (+48) 531 785 081
sekretariat.palac@bk.pan.pl | www.bkpan.poznan.pl | NIP 525-15-75-083



Fundusze Europejskie
na Rozwój Cyfrowy



Rzeczpospolita
Polska

Dofinansowane przez
Unię Europejską



200 lat
BIBLIOTEKI
KÓRNICKIEJ

wrażliwych. Po wykonaniu implementacji warto uwzględnić ją podczas testów bezpieczeństwa systemu. PCSS dysponuje zespołem specjalizującym się w wykonywaniu takich testów. Odnośnie doświadczenia użytkowników, należy zdecydować, czy użytkownik po wejściu na stronę logowania ma widzieć przycisk "Zaloguj przez Profil Zaufany", czy też domyślny formularz logowania. Krytycznym elementem UX jest obsługa sytuacji, gdy użytkownik posiada już konto w bibliotece, a loguje się pierwszy raz przez Węzeł Krajowy. System powinien pozwolić na jednorazowe "spięcie" konta bibliotecznego z tożsamością z Węzła Krajowego, aby uniknąć duplikowania użytkowników.

3. MARC21, Kronik@, integrację z repozytoriami

Pytanie:

Jakie są najlepsze praktyki implementacji pełnej obsługi formatu MARC21 oraz integracji z repozytorium Kronik@ i podobnymi systemami? Czy rekomendują Państwo konkretne narzędzia lub biblioteki do konwersji i mapowania metadanych (MARC21 ⇔ RDF/JSON-LD itp.)?

Odpowiedź:

W kontekście implementacji formatu MARC21, system powinien być interoperacyjny. Metadane obiektów cyfrowych zapisane w formacie MARC21 powinny być łatwe do wyeksportowania do innych, wybranych w toku analizy formatów danych. Również import danych do systemu z formatu MARC21 lub innych wspieranych przez system formatów, powinien być łatwy do przeprowadzenia bez konieczności angażowania programistów. Rekomendujemy, aby system posiadał konfigurowalny silnik mapowań, który pozwoli administratorom biblioteki na samodzielną modyfikację reguł konwersji z MARC21 do formatów wyjściowych (np. Kronik@) bez konieczności angażowania developerów przy każdej zmianie standardu metadanych.

4. OCR – model lokalny czy chmurowy?

Pytanie:

Jaki model wdrożenia OCR jest korzystniejszy z Państwa perspektywy – jako usługa chmurowa czy lokalny moduł? Jak zapewnić zgodność takiego rozwiązania z wymogami RODO i standardem PDF/A? Prosimy o ewentualne przykłady rozwiązań OCR dobrze sprawdzających się w polskich realiach (np. pod kątem języka polskiego i materiałów historycznych).



Zamek w Kórniku
ul. Zamkowa 5, 62-035 Kórnik
tel. (+48) 61 817 00 81
tel. (+48) 531 998 263
sekretariat.zamek@bk.pan.pl

Pałac Działyńskich
Stary Rynek 78/79, 61-772 Poznań
tel. (+48) 61 852 48 44
tel. (+48) 531 785 081
sekretariat.palac@bk.pan.pl | www.bkpan.poznan.pl | NIP 525-15-75-083



Odpowiedź:

Wybór między usługą chmurową a lokalnym modulem zależy ściśle od priorytetów projektu, głównie w zakresie bezpieczeństwa oraz skali i kosztów. Usługa chmurowa cechuje się niskim kosztem początkowym (brak konieczności zakupu własnego sprzętu). Koszty eksploatacyjne są zmienne i zależne od użytych zasobów. Zapewnienie należytego poziomu bezpieczeństwa wymaga przeprowadzania audytów bezpieczeństwa i umów na przetwarzanie danych. Skalowalność jest wysoka, szybko można zwiększyć moce obliczeniowe. Lokalny moduł OCR wiąże się z wysokim kosztem początkowym – dedykowany serwer. Koszty eksploatacyjne są stałe a bezpieczeństwo łatwiejsze do kontrolowania. Skalowalność jest ograniczona i zależna od posiadanego sprzętu. Wdrożenie rozwiązania w modelu chmurowym może okazać się korzystniejsze z uwagi na możliwość zautomatyzowania procesu konwersji obiektów cyfrowych oraz przeniesienie ciężaru zarządzania infrastrukturą na dostawcę usługi. Konwersja do PDF/A wymaga, aby wynikowy plik PDF zawierał osadzoną, przeszukiwalną warstwę tekstową. Zaimplementowane rozwiązanie OCR (czy to chmurowe, czy lokalne) musi posiadać funkcjonalność eksportu do PDF/A, generując jeden spójny plik, w którym warstwa graficzna (obraz) i warstwa tekstowa (z OCR) są poprawnie zsynchronizowane. Wiele narzędzi OCR oferuje natywną obsługę PDF/A. W kontekście języka polskiego, a zwłaszcza materiałów historycznych, kluczowa jest jakość silnika OCR. W tym celu wykorzystane mogą zostać rozwiązania komercyjne np. ABBYY FineReader lub OpenSource np. Tesseract. W PCSS posiadamy doświadczenie z oboma rozwiązaniami.

5. Zastosowanie AI – chatbot, rekomendacje, automatyzacja

Pytanie:

Jakie narzędzia AI (np. silniki chatbotów, systemy rekomendacyjne) są obecnie dostępne i warte rozważenia dla biblioteki naukowej? Czy mają Państwo doświadczenia we wdrażaniu wirtualnych asystentów lub podobnych rozwiązań w sektorze kultury i nauki – jakie były wnioski, korzyści i ewentualne problemy?

Odpowiedź:

W kontekście modernizacji systemów bibliotecznych i podnoszenia jakości obsługi, wdrożenie sztucznej inteligencji może przybierać np. poniższe formy:

- Chatboty i Wirtualni Asystenci



Zamek w Kórniku
ul. Zamkowa 5, 62-035 Kórnik
tel. (+48) 61 817 00 81
tel. (+48) 531 998 263
sekretariat.zamek@bk.pan.pl

Pałac Działyńskich
Stary Rynek 78/79, 61-772 Poznań
tel. (+48) 61 852 48 44
tel. (+48) 531 785 081
sekretariat.palac@bk.pan.pl | www.bkpan.poznan.pl | NIP 525-15-75-083



- Chatboty oparte na AI mogą przejąć obsługę rutynowych zapytań, odciążając personel biblioteczny i oferując wsparcie 24/7
- Systemy Rekomendacyjne
 - Systemy zwiększają widoczność zbiorów i poprawiają doświadczenie użytkownika
- Narzędzia do Wzbogacania i Przetwarzania Danych
 - Wykorzystanie AI do poprawy jakości rozpoznawania tekstu, szczególnie przy materiałach historycznych i rękopisach
 - Automatyczne przypisywanie słów kluczowych
 - Tłumaczenie opisów bibliograficznych lub streszczeń w celu zwiększenia dostępności zbiorów obcojęzycznych dla polskiego użytkownika

W PCSS uruchomiono usługę dostępu do modeli sztucznej inteligencji (AI) poprzez interfejs API. Usługa ta została przygotowana z myślą o wsparciu badań naukowych, projektów akademickich oraz działalności dydaktycznej, oferując polskim naukowcom i studentom zaawansowane narzędzia oparte o bezpieczną infrastrukturę HPC PCSS. Obecnie udostępniamy szeroki wachlarz zaawansowanych modeli, w tym: DeepSeek-V3.1-vLLM, codellama:70b, Qwen3-Embedding-0.6B, llama3.3:70b, qwen3-coder:30b, bielik_11b, bielik_4.5b, Nanonets-OCR-s oraz gpt-oss_120b

Przykładowe realizacje usług z sektora kultury, zawierające mechanizmy sztucznej inteligencji:

- <https://pp.dlibra.psnc.pl/>
- <https://wlt.pcss.pl/>
- <https://tene.fbc.pionier.net.pl/>

6. Koszty i ryzyka

Pytanie:

Jak oceniają Państwo potencjalne koszty licencji, implementacji oraz utrzymania zaproponowanych rozwiązań (OCR, AI, nowe moduły)? Jakie główne ryzyka dostrzegają Państwo w realizacji tak szerokiej modernizacji istniejącego systemu i jak można im przeciwdziałać na etapie planowania oraz wdrożenia? Jakiego rodzaju testy bezpieczeństwa rekomendujecie Państwo?

Odpowiedź:

Największe nakłady finansowe i czasowe należy przewidzieć na implementację funkcjonalności AI (szczególnie na etapie przygotowania danych i pierwszego treningu modelu) oraz na zakup licencji profesjonalnego silnika OCR.





Realizacja tak kompleksowej modernizacji systemu obarczona jest kilkoma kluczowymi ryzykami:

- ryzyko jakości danych
- ryzyko bezpieczeństwa
- ryzyko integracyjne
- ryzyko dostępności (WCAG)

Testy bezpieczeństwa powinny być realizowane zarówno na etapie wytwarzania oprogramowania jak i po jego implementacji. Uwzględnić należy testy penetracyjne, przegląd konfiguracji systemu, testy autoryzacji i testy kontroli dostępu. Rekomendujemy zlecenie testów bezpieczeństwa niezależnemu od wykonawcy, zewnętrznemu podmiotowi, posiadającemu doświadczenie w audytowaniu systemów IdP (Identity Provider) i integracji SAML.

PCSS może zaoferować przeprowadzenie audytu systemu wykorzystującego AI, mającego na celu wykrycie błędów, nieprawidłowości oraz podatności systemów korzystających z modeli AI, zanim potencjalnie wyrządzą one szkody.

7. Harmonogram wdrożenia

Pytanie:

Prosimy o przedstawienie orientacyjnego harmonogramu realizacji projektu (z podziałem na etapy, np. analiza przedwdrożeniowa, projektowanie, implementacja, integracje, testy, szkolenia, wdrożenie).

Odpowiedź:

Przygotowanie precyzyjnego harmonogramu wymaga dokładnej analizy wolumenu danych i złożoności systemów źródłowych, dlatego przedstawiamy orientacyjny harmonogram realizacji projektu:

- Analiza i projektowanie
 - analiza przedwdrożeniowa (3M)
 - projektowanie techniczne oraz UX (3M)
- Implementacja modułów podstawowych i dostępności
 - Moduł zarządzania metadanymi (4M)
 - Moduł konserwatora (4M)
 - Wirtualny Asystent Czytelnika (10M)





- Moduł OCR (4M)
- Moduł analityczny (4M)
- Implementacja WCAG (4M)
- Integracja i trening danych
 - Mechanizmy logowania i uwierzytelniania (4M)
 - Implementacja i integracja OCR (5M)
 - Przygotowanie danych dla AI (3M)
- Prace implementacyjne
 - Implementacja modułów AI (5M)
 - Testy i audyty końcowe (3M)
 - Szkolenia i wdrożenie produkcyjne (2M)

8. Wymagana infrastruktura

Pytanie:

Prosimy o przedstawienie orientacyjnych wymagań dla infrastruktury, niezbędnej do utrzymania systemu (serwery, macierze, GPU, backup, archiwizacja długoterminowa, administracja systemami operacyjnymi, monitorowanie usług).

Odpowiedź:

Poniżej przedstawiamy orientacyjne wymagania infrastrukturalne, z podziałem na kluczowe komponenty:

- Serwery aplikacyjne i bazy danych
 - Dwa serwery o parametrach:
 - 16vCPU
 - 64GB RAM
 - 500GB przestrzeń dyskowa
- Infrastruktura AI
 - Dostęp do GPU i szybki cache
 - Sparametryzowanie usługi wymaga wskazania wolumenu danych jaki ma być przetwarzany i ram czasowych w jakich mają być otrzymane wyniki. Prosimy o doprecyzowanie. Istnieje możliwość uruchomienia usługi rozliczanej w oparciu o tokeny lub abonament za zasoby.
- Infrastruktura OCR
 - serwer o parametrach:





- 16vCPU
- 64GB RAM
- 500GB przestrzeń dyskowa
- Macierze i archiwizacja długoterminowa
 - SSD (wolumen zależny od potrzeb, prosimy o doprecyzowanie ilości przetwarzanych danych lub wskazanie orientacyjnych wartości)
 - HDD (wolumen zależny od potrzeb, prosimy o doprecyzowanie ilości przetwarzanych danych lub wskazanie orientacyjnych wartości)
 - Przestrzeń dla długoterminowej archiwizacji (wolumen zależny od potrzeb, prosimy o doprecyzowanie ilości przetwarzanych danych lub wskazanie orientacyjnych wartości)
 - Przestrzeń obiektowa (wolumen zależny od potrzeb, prosimy o doprecyzowanie ilości przetwarzanych danych lub wskazanie orientacyjnych wartości)
- Administracja systemami i monitorowanie usług
 - Backup (wolumen zależny od potrzeb, prosimy o doprecyzowanie ilości przetwarzanych danych lub wskazanie orientacyjnych wartości)
 - Monitorowanie (Czy zakładane są konkretne SLA?)
 - Zarządzanie maszynami

9. Przewidywany podział kosztów

Pytanie:

Prosimy o przedstawienie orientacyjnego modelu podziału kosztów na składowe, np.:

- licencja (moduł konserwatora, moduł asystenta czytelnika, ...)
- prace analityczne i projektowe
- prace UX z uwzględnieniem UCD
- szkolenia
- warstwa interfejsu (frontend)
- backend, API, integracje
- Wirtualny Asystent Czytelnika





Fundusze Europejskie
na Rozwój Cyfrowy



Rzeczpospolita
Polska

Dofinansowane przez
Unię Europejską



200 lat
BIBLIOTEKI
KÓRNICKIEJ

- Moduł przetwarzania języka naturalnego (NLP)
- Moduł wyszukiwania i eksploracji zasobów BK PAN
- Moduł interfejsu użytkownika BK PAN
- Moduł generowania raportów i eksportu danych BK PAN
- Moduł integracji z systemami bibliotecznymi BK PAN
- Moduł administracji dla bibliotekarzy BK PAN
- Moduł automatycznego tworzenia rekordów MARC21
- Moduł OCR
- Moduł Konserwatora
- Moduł AI
- Trenowanie modułu AI
- Integracja z węzłem krajowym
- Obsługa SSO
- Moduł Analityczny
- Migracja i integracje MARC21
- Testy funkcjonalne systemu
- Testy interfejsu pod kątem WCAG
- Audyt bezpieczeństwa
- Utrzymanie roczne systemu (pomoc techniczna, obsługa błędów)
- Utrzymanie roczne infrastruktury



Zamek w Kórniku
ul. Zamkowa 5, 62-035 Kórnik
tel. (+48) 61 817 00 81
tel. (+48) 531 998 263
sekretariat.zamek@bk.pan.pl

Pałac Działyńskich
Stary Rynek 78/79, 61-772 Poznań
tel. (+48) 61 852 48 44
tel. (+48) 531 785 081
sekretariat.palac@bk.pan.pl | www.bkpan.poznan.pl | NIP 525-15-75-083



Fundusze Europejskie
na Rozwój Cyfrowy



Rzeczpospolita
Polska

Dofinansowane przez
Unię Europejską



200 lat
BIBLIOTEKI
KÓRNICKIEJ

Odpowiedź:

Orientacyjny podział kosztów może wyglądać następująco, jednak wymaga on jeszcze pogłębionej analizy.

Licencje (w tym na silniki OCR) ok. 300 000,00 zł netto

Prace analityczne ok. 150 000,00 zł netto

Projektowanie UX ok. 150 000,00 zł netto

Prace programistyczne backend i frontend ok. 2 700 000,00 zł netto

Szkolenia ok. 70 000,00 zł netto

Testy użyteczności i testy bezpieczeństwa ok. 200 000,00 zł netto

Koszt utrzymania infrastruktury jest zależny od dodatkowych odpowiedzi wskazanych wyżej.

III. Dodatkowe rekomendacje lub uwagi uczestnika (opcjonalne)

Potrzebne są dodatkowe informacje dot. planowanego wolumenu przetwarzanych danych.



Zamek w Kórniku

ul. Zamkowa 5, 62-035 Kórnik
tel. (+48) 61 817 00 81
tel. (+48) 531 998 263
sekretariat.zamek@bk.pan.pl

Pałac Działyńskich

Stary Rynek 78/79, 61-772 Poznań
tel. (+48) 61 852 48 44
tel. (+48) 531 785 081

sekretariat.palac@bk.pan.pl | www.bkpan.poznan.pl | NIP 525-15-75-083