

OPIS ZAŁOŻEŃ PROJEKTU INFORMATYCZNEGO

Tytuł projektu	Repozytorium CRIS Uniwersytetu Papieskiego Jana Pawła II w Krakowie		
Wnioskodawca	Minister Nauki i Szkolnictwa Wyższego		
Beneficjent	Uniwersytet Papieski Jana Pawła II w Krakowie		
Partnerzy			
Źródło finansowania	Środki UE – Działanie FERC.02.03 Cyfrowa dostępność i ponowne wykorzystanie informacji Budżet państwa – Część budżetowa: 28 – Szkolnictwo wyższe i nauka		
Całkowity koszt projektu	6 371 565,38 zł		
Planowany okres realizacji projektu	01-2027 do 12-2029		
Osoba kontaktowa	Sebastian Wojnowski	sebastian.wojnowski@upjp2.edu.pl	606330292

1. POWODY PODJĘCIA PROJEKTU

1.1. Identyfikacja problemu i potrzeb

Archiwum Arcybiskupa Eugeniusza Baziaka w Krakowie, działające od 2012 r. w strukturze UPJPII, przechowuje unikatowy zasób akt dawnej archidiecezji lwowskiej (XVI–XX w.), w tym księgi metrykalne o wysokiej wartości dla badań historycznych, genealogicznych i demograficznych. Mimo dużego zainteresowania (ok. 150 kwerend rocznie i 30 zapytań tygodniowo), interesariusze – naukowcy, genealodzy, studenci i instytucje kultury – napotykały istotne bariery: ograniczoną skalę digitalizacji, brak zaawansowanych narzędzi wyszukiwawczych, niską interoperacyjność danych oraz utrudniony zdalny dostęp do źródeł. Utrudnia to prowadzenie badań porównawczych, analiz ilościowych oraz ponowne wykorzystanie danych. Aktualnie wykorzystywana infrastruktura sprzętowa i programowa umożliwia podstawową obsługę procesów archiwalnych i repozytoryjnych, jednak nie zapewnia narzędzi do masowego przetwarzania danych, zaawansowanej analizy treści ani automatyzacji procesów, w tym z wykorzystaniem rozwiązań opartych na sztucznej inteligencji. Ograniczony poziom digitalizacji i interoperacyjności systemów istotnie zmniejsza dostępność zasobów oraz możliwości ich ponownego wykorzystania na potrzeby badań naukowych, edukacji i działalności społecznej. Projekt odpowiada na te potrzeby poprzez konserwację i digitalizację zasobu, jego udostępnienie w formie cyfrowej oraz wdrożenie nowoczesnych narzędzi analitycznych, w tym rozwiązań AI, co znacząco zwiększy dostępność, użyteczność i potencjał badawczy zbiorów. Kluczowe jest także zapewnienie interoperacyjności poprzez integrację z systemami zewnętrznymi. Integracja z KRONiK@ umożliwi włączenie danych do krajowej infrastruktury informacji naukowej, natomiast integracja z dLibra zapewni szeroką publikację zasobów w ekosystemie bibliotek cyfrowych oraz ich widoczność w agregatorach krajowych i międzynarodowych. Dzięki temu zasób stanie się częścią otwartej infrastruktury danych naukowych, zgodnej z zasadami FAIR i celami polityk krajowych oraz europejskich.

Interesariusz	Zidentyfikowany problem	Szacowana wielkość grupy
---------------	-------------------------	--------------------------

Interesariusz	Zidentyfikowany problem	Szacowana wielkość grupy
<p>Naukowcy i badacze (historycy, demografowie, badacze nauk humanistycznych i społecznych, nauczyciele historii, w tym wykładowcy akademicy)</p>	<p>(1) Ograniczony dostęp do unikatowych źródeł archiwalnych – naukowcy i badacze mają utrudniony dostęp do materiałów ze względu na konieczność korzystania z nich wyłącznie na miejscu w czytelni, która funkcjonuje w ograniczonych godzinach i dniach, co znacząco wydłuża proces badawczy. (2) Brak narzędzi umożliwiających zdalne zaawansowane przeszukiwanie i analizę danych zawartych w księgach metrykalnych – utrudnia to prowadzenie badań porównawczych, analiz ilościowych oraz ponowne wykorzystanie danych w projektach naukowych. (3) Niewystarczająca efektywność procesu kwerend archiwalnych – przygotowanie materiałów na potrzeby publikacji naukowych wymaga wielu godzin pracy z dokumentem źródłowym; badacze często spędzają w archiwum wiele tygodni, co znacząco wydłuża czas realizacji badań. (4) Ograniczona możliwość równoczesnego korzystania ze źródeł archiwalnych – brak digitalizacji powoduje, że z jednej jednostki archiwalnej może korzystać jednocześnie tylko jedna osoba, co ogranicza dostępność zasobu dla większej liczby badaczy.</p>	<p>10000</p>
<p>Potomkowie mieszkańców dawnej archidiecezji lwowskiej</p>	<p>(1) Ograniczony dostęp do archiwaliów ze względu na ich zły stan zachowania – znaczna część dokumentów jest w tak złym stanie, że nie może być udostępniana użytkownikom, co uniemożliwia zapoznanie się z ich treścią oraz prowadzenie badań genealogicznych. (2) Brak możliwości realizacji kwerend dla zniszczonych materiałów archiwalnych – w przypadku najbardziej zdegradowanych jednostek archiwalnych nie jest możliwe ani ich udostępnienie, ani wykonanie kwerendy przez pracowników archiwum, co całkowicie wyklucza te zasoby z wykorzystania. (3) Ograniczona dostępność archiwum wynikająca z godzin otwarcia – archiwum funkcjonuje w określonych dniach i godzinach, a liczba użytkowników, którzy mogą zostać obsłużeni jednocześnie, jest ograniczona, co wydłuża czas oczekiwania i utrudnia dostęp do zasobu. (4) Ograniczony dostęp do zasobów dla osób spoza Krakowa i zagranicy – potomkowie mieszkańców dawnej archidiecezji lwowskiej często mieszkają poza</p>	<p>150000</p>

Interesariusz	Zidentyfikowany problem	Szacowana wielkość grupy
	<p>regionem, a nawet poza granicami Polski (m.in. Francja, Wielka Brytania, Ukraina, Rosja, RPA, kraje Ameryki Południowej), co powoduje, że osobista wizyta w archiwum jest dla nich dużym obciążeniem logistycznym i finansowym lub jest całkowicie niemożliwa.</p> <p>(5) Brak możliwości równoczesnego korzystania z tych samych materiałów archiwalnych przez wielu użytkowników – dostęp do pojedynczych jednostek archiwalnych jest ograniczony do jednej osoby w danym czasie, co zmniejsza dostępność zasobu dla szerszego grona odbiorców.</p>	
<p>Genealogowie (zawodowi i amatorzy, firmy genealogiczne – oszacowano na podstawie dotychczasowego wykorzystania zasobu (liczba kwerend i zapytań) z uwzględnieniem efektu skali wynikającego z ich cyfrowego udostępnienia. Przyjęto, że zniesienie barier geograficznych i czasowych oraz możliwość równoczesnego dostępu do zasobów spowoduje kilkukrotny wzrost liczby użytkowników, w tym także odbiorców zagranicznych zainteresowanych historią rodzin z obszaru dawnej archidiecezji lwowskiej)</p>	<p>(1) Brak powszechnego, zdalnego dostępu do ksiąg metrykalnych z obszaru dawnej archidiecezji lwowskiej – genealogowie (zarówno zawodowi, jak i amatorzy oraz firmy genealogiczne) zmuszeni są do osobistego korzystania z zasobów lub rezygnacji z badań, co ogranicza dostępność materiałów. (2) Brak narzędzi umożliwiających efektywne wyszukiwanie danych w dużych zbiorach historycznych – przeszukiwanie ksiąg metrykalnych odbywa się w sposób manualny, co znacząco wydłuża proces identyfikacji konkretnych osób i zdarzeń. (3) Niewystarczająca efektywność prowadzenia kwerend genealogicznych – konieczność ręcznej analizy dokumentów powoduje, że odtworzenie historii rodzin jest bardzo czasochłonne i kosztowne, szczególnie w przypadku zleceń komercyjnych. (4) Ograniczone możliwości rekonstrukcji historii rodzin – brak szerokiego dostępu do danych oraz narzędzi analitycznych utrudnia łączenie informacji z różnych źródeł i budowanie pełnych drzew genealogicznych.</p>	1000
<p>Osoby prywatne poszukujące informacji o przodkach</p>	<p>(1) Ograniczony dostęp do archiwaliów ze względu na zły stan zachowania akt – znaczna część dokumentów jest w stanie, który uniemożliwia ich udostępnianie, co wyklucza możliwość zapoznania się z ich treścią. (2) Brak możliwości korzystania z najbardziej zniszczonych materiałów</p>	15000

Interesariusz	Zidentyfikowany problem	Szacowana wielkość grupy
	<p>archiwalnych – dokumenty wymagające konserwacji nie są udostępniane, co powoduje trwałe wyłączenie części zasobu z wykorzystania przez użytkowników. (3) Niewystarczający poziom dostępności zasobu archiwalnego – brak digitalizacji powoduje, że dostęp do materiałów jest ograniczony wyłącznie do formy fizycznej i obecności w archiwum. (4) Brak kompetencji użytkowników w zakresie korzystania z ksiąg metrykalnych – osoby prywatne często nie posiadają wiedzy i doświadczenia w prowadzeniu kwerend archiwalnych, co znacząco utrudnia samodzielne poszukiwania genealogiczne. (5) Brak narzędzi wspierających samodzielne wyszukiwanie informacji – brak intuicyjnych, cyfrowych rozwiązań ogranicza możliwość efektywnego przeszukiwania zasobu oraz budowania drzew genealogicznych przez użytkowników.</p>	
<p>Studenci i doktoranci historii</p>	<p>(1) Ograniczony dostęp do źródeł archiwalnych niezbędnych w procesie dydaktycznym i badawczym – studenci i doktoranci mają utrudniony dostęp do materiałów źródłowych ze względu na konieczność pracy na miejscu oraz ograniczoną dostępność archiwum. (2) Brak zdalnego dostępu do materiałów źródłowych – niemożność korzystania z archiwaliów poza siedzibą archiwum ogranicza możliwości prowadzenia badań, przygotowywania prac dyplomowych i rozpraw doktorskich. (3) Brak narzędzi do efektywnego przeszukiwania i analizy źródeł historycznych – ręczne przeszukiwanie ksiąg metrykalnych utrudnia identyfikację potrzebnych informacji i wydłuża czas realizacji prac naukowych. (4) Niewystarczające możliwości wykorzystania danych archiwalnych w badaniach naukowych – brak zdigitalizowanych i ustrukturyzowanych danych ogranicza możliwość prowadzenia analiz porównawczych, statystycznych i interdyscyplinarnych. (5) Ograniczone możliwości nauki pracy ze źródłami historycznymi – brak cyfrowych narzędzi i szerokiego dostępu do materiałów ogranicza rozwój kompetencji badawczych, szczególnie w zakresie nowoczesnych metod analizy danych historycznych.</p>	<p>5000</p>

Interesariusz	Zidentyfikowany problem	Szacowana wielkość grupy
Regionaliści, nauczyciele, pasjonaci historii lokalnej	(1) Ograniczony dostęp do lokalnych źródeł historycznych – regionaliści, nauczyciele i pasjonaci historii mają utrudniony dostęp do archiwaliów dokumentujących dzieje społeczności lokalnych, co ogranicza możliwość prowadzenia badań i popularyzacji wiedzy. (2) Brak narzędzi umożliwiających łatwe wyszukiwanie informacji o historii lokalnej – przeszukiwanie materiałów archiwalnych jest czasochłonne i wymaga specjalistycznej wiedzy, co stanowi barierę dla osób niebędących profesjonalnymi badaczami. (3) Niewystarczające możliwości wykorzystania materiałów archiwalnych w edukacji i popularyzacji historii – brak dostępu do zdigitalizowanych zasobów ogranicza wykorzystanie źródeł w dydaktyce szkolnej, działalności regionalnej oraz inicjatywach społecznych.	10000
Instytucje naukowe	(1) Brak możliwości automatycznego pobierania danych archiwalnych – brak interfejsów API oraz ustandaryzowanych mechanizmów dostępu ogranicza możliwość integracji zasobów z systemami badawczymi oraz narzędziami analitycznymi wykorzystywanymi przez instytucje naukowe. (2) Niewystarczający poziom interoperacyjności i ustrukturyzowania danych – dane archiwalne nie są dostępne w formatach umożliwiających ich masowe przetwarzanie, co utrudnia ich wykorzystanie w badaniach naukowych, w tym analizach ilościowych i projektach opartych na dużych zbiorach danych.	100
Ministerstwo Nauki i Szkolnictwa Wyższego	(1) Niewystarczające wykorzystanie potencjału naukowego zasobów – brak pełnej digitalizacji i dostępności online ogranicza możliwość prowadzenia badań na skalę krajową i międzynarodową, co jest sprzeczne z celami MNiSW dotyczącymi rozwoju nauki i jej umiędzynarodowienia. (2) Brak spójności z polityką otwartej nauki i zarządzania danymi – ograniczona wyszukiwalność, brak standardów interoperacyjności i udostępniania danych utrudniają realizację strategii Ministerstwa w zakresie FAIR data oraz budowy zintegrowanej infrastruktury informacji naukowej. (3) Niedostateczna ochrona i cyfrowe zabezpieczenie dziedzictwa naukowego – brak pełnej digitalizacji	1

Interesariusz	Zidentyfikowany problem	Szacowana wielkość grupy
	zwiększa ryzyko utraty cennych materiałów, co stanowi problem systemowy dla MNiSW odpowiedzialnego za trwałość i dostępność narodowych zasobów wiedzy.	
Pracownicy Uniwersytetu Papieskiego Jana Pawła II w Krakowie	(1) Ograniczone wykorzystanie zasobu w pracy naukowej pracowników (ok. 300–500 osób dydaktycznych i badawczych) – brak pełnej digitalizacji utrudnia szybki dostęp do materiałów, co spowalnia prowadzenie badań, publikacje i projekty grantowe. (2) Niska efektywność pracy badawczej i dydaktycznej – brak zaawansowanych narzędzi wyszukiwania i analizy danych powoduje, że pracownicy muszą prowadzić czasochłonne kwerendy manualne zamiast korzystać z nowoczesnych metod (np. analizy cyfrowej). (3) Wysokie obciążenie i ograniczona efektywność pracy bibliotekarzy i archiwistów – konieczność manualnej obsługi zapytań, brak zautomatyzowanych narzędzi wyszukiwawczych i digitalizacji zasobu zwiększa nakład pracy, wydłuża czas realizacji kwerend i ogranicza możliwości rozwoju usług archiwalnych.	300

1.2. Opis stanu obecnego

Archiwum Arcybiskupa Eugeniusza Baziaka, będące jednostką organizacyjną Uniwersytetu Papieskiego Jana Pawła II w Krakowie, realizuje swoje procesy w oparciu o tradycyjny model archiwalny, wspierany wybranymi rozwiązaniami cyfrowymi. Działalność obejmuje gromadzenie, opracowywanie, przechowywanie oraz udostępnianie materiałów archiwalnych, w szczególności ksiąg metrykalnych parafii rzymskokatolickich z terenów dawnej archidiecezji lwowskiej z XVI–XX w. oraz innych akt o charakterze historycznym. Procesy udostępniania realizowane są głównie w formie kwerend prowadzonych na miejscu w czytelni Archiwum oraz poprzez obsługę zapytań indywidualnych. Opracowanie materiałów archiwalnych odbywa się z wykorzystaniem ewidencji tradycyjnej oraz podstawowych narzędzi informatycznych. Przetwarzane dane obejmują dane osobowe o charakterze historycznym (np. dane z ksiąg metrykalnych), dane opisowe dotyczące jednostek archiwalnych oraz dane bibliograficzne. Zasadnicza część zasobu archiwalnego nie została dotychczas zdigitalizowana i nie jest dostępna w formie cyfrowej ani udostępniona w repozytorium instytucjonalnym. Infrastruktura techniczna obejmuje podstawowe zasoby sprzętowe (serwery, stacje robocze) oraz repozytorium instytucjonalne oparte na platformie DSpace. System ten wspiera udostępnianie zasobów cyfrowych, jednak wykorzystywany jest w ograniczonym zakresie funkcjonalnym i nie obejmuje większości materiałów archiwalnych.

2. EFEKTY PROJEKTU

2.1. Cele i korzyści wynikające z projektu

Cel - 1	Konserwacja i digitalizacja zasobu archiwalnego
Cel strategiczny	(1) Strategia na rzecz Odpowiedzialnego Rozwoju do roku 2020 (z perspektywą do 2030), cel szczegółowy III: Skuteczne państwo i instytucje służące wzrostowi oraz włączeniu społecznemu i gospodarczemu, Obszar: E-Państwo. (2) Europejska strategia w zakresie danych (COM/2020/66) – w zakresie wspierania Europejskiej Chmury Otwartej Nauki poprzez udostępnianie i interoperacyjność danych. (3) Program otwierania danych na lata 2021–2027, cel czwarty: Stymulowanie rynku ponownego wykorzystywania zasobów kultury i danych naukowych
Korzyść:	(1) Zabezpieczenie unikatowych źródeł archiwalnych przed degradacją i utratą, co zapewnia ich długoterminową dostępność dla badań naukowych. (2) Ułatwienie dostępu badaczom do materiałów źródłowych niezależnie od miejsca i czasu, co sprzyja rozwojowi badań interdyscyplinarnych. (3) Zwiększenie możliwości analizy porównawczej i ilościowej dzięki wykorzystaniu zdigitalizowanych danych w narzędziach cyfrowych. (4) Upowszechnienie wyników badań naukowych oraz wsparcie nowych projektów badawczych opartych na wiarygodnych źródłach archiwalnych
KPI:	KPI 1: Liczba zdigitalizowanych dokumentów zawierających informacje sektora publicznego lub dane prywatne KPI 2: Rozmiar zdigitalizowanych informacji sektora publicznego lub danych prywatnych
Wartość aktualna i docelowa KPI:	Wartość aktualna KPI 1: 0 Wartość aktualna KPI 2: 0 Wartość docelowa KPI 1: 2700 Wartość docelowa KPI 2: 1 TB
Metoda pomiaru KPI	KPI 1 Metoda pomiaru – zliczenie liczby dokumentów poddanych procesowi digitalizacji w ramach projektu (np. skanowanie, OCR, opracowanie metadanych) Częstotliwość – na rozpoczęcie projektu oraz na zakończenie projektu Źródło – dokumentacja projektowa, raporty z procesu digitalizacji, raporty systemowe KPI 2 Metoda pomiaru – określeniu łącznego rozmiaru danych cyfrowych (w GB lub TB) powstałych w wyniku procesu digitalizacji, obejmujących pliki źródłowe, przetworzone oraz metadane zapisane w systemach repozytoryjnych Częstotliwość – na rozpoczęcie projektu oraz na zakończenie projektu. Źródło – raporty systemowe
Cel - 2	Zwiększenie dostępności i wykorzystania zasobów naukowych
Cel strategiczny	(1) Strategia na rzecz Odpowiedzialnego Rozwoju do roku 2020 (z perspektywą do 2030), cel szczegółowy III: Skuteczne państwo i instytucje służące wzrostowi oraz włączeniu społecznemu i gospodarczemu, Obszar: E-Państwo. (2) Europejska strategia w zakresie danych (COM/2020/66) – w zakresie wspierania Europejskiej Chmury Otwartej Nauki poprzez udostępnianie i interoperacyjność danych. (3) Program otwierania danych na lata 2021–2027, cel czwarty: Stymulowanie rynku ponownego wykorzystywania zasobów kultury i danych naukowych
Korzyść:	(1) Umożliwienie szerokiemu gronu odbiorców, w tym naukowcom i studentom, zdalnego dostępu do zasobów archiwalnych bez ograniczeń geograficznych. (2) Zwiększenie intensywności wykorzystania zasobów

	naukowych w badaniach, dydaktyce oraz projektach interdyscyplinarnych. (3) Skrócenie czasu dotarcia do materiałów źródłowych i usprawnienie procesu prowadzenia badań naukowych. (4) Upowszechnienie wiedzy oraz wyników badań opartych na zasobach archiwalnych poprzez ich łatwiejszą dostępność w formie cyfrowej
KPI:	<p>KPI 1: Liczba rozwiązań wykorzystujących informacje sektora publicznego lub dane prywatne</p> <p>KPI 2: Liczba baz danych udostępnionych online poprzez API</p> <p>KPI 3: Liczba udostępnionych online dokumentów zawierających informacje sektora publicznego lub dane prywatne</p> <p>KPI 4: Rozmiar udostępnionych online informacji sektora publicznego lub danych prywatnych</p> <p>KPI 5: Użytkownicy nowych i zmodernizowanych publicznych usług, produktów i procesów cyfrowych</p> <p>KPI 6: Liczba pracowników niebędących pracownikami IT, objętych wsparciem szkoleniowym</p>
Wartość aktualna i docelowa KPI:	<p>Wartość aktualna KPI 1: 0</p> <p>Wartość aktualna KPI 2: 0</p> <p>Wartość aktualna KPI 3: 0</p> <p>Wartość aktualna KPI 4: 0</p> <p>Wartość aktualna KPI 5: 0</p> <p>Wartość aktualna KPI 6: 0</p> <p>Wartość docelowa KPI 1: 1</p> <p>Wartość docelowa KPI 2: 1</p> <p>Wartość docelowa KPI 3: 2700</p> <p>Wartość docelowa KPI 4: 1 TB</p> <p>Wartość docelowa KPI 5: 10000</p> <p>Wartość docelowa KPI 6: 5</p>
Metoda pomiaru KPI	<p>KPI 1</p> <p>Metoda pomiaru – osiągnięcie wartości docelowej (1) zostanie potwierdzone uruchomieniem co najmniej jednego rozwiązania</p> <p>Częstotliwość – na rozpoczęcie i zakończenie projektu</p> <p>Źródło – dokumentacja techniczna, protokoły odbioru</p> <p>KPI 2</p> <p>Metoda pomiaru – zliczenie liczby baz danych, które zostały udostępnione online za pośrednictwem API w ramach projektu</p> <p>Częstotliwość – na rozpoczęcie i zakończenie projektu</p> <p>Źródło – dokumentacja projektowa, dokumentacja techniczna API, protokoły odbioru</p> <p>KPI 3</p> <p>Metoda pomiaru – zliczenie liczby dokumentów udostępnionych online na podstawie raportów systemowych</p> <p>Częstotliwość – na rozpoczęcie projektu oraz na zakończenie projektu</p> <p>Źródło – dokumentacja projektowa, raporty systemowe</p> <p>KPI 4</p> <p>Metoda pomiaru – pomiar łącznego rozmiaru (w GB lub TB) informacji archiwalnych udostępnionych online</p> <p>Częstotliwość – na rozpoczęcie projektu oraz na zakończenie projektu</p> <p>Źródło – dokumentacja projektowa, raporty systemowe</p> <p>KPI 5</p> <p>Metoda pomiaru – analiza wykorzystania systemu na podstawie logów systemowych oraz narzędzi analitycznych, potwierdzająca korzystanie z</p>

	<p>udostępnionej usługi cyfrowej w danym okresie referencyjnym</p> <p>Częstotliwość – na rozpoczęcie projektu oraz na zakończenie projektu</p> <p>Źródło – raporty systemowe (logi), narzędzia analityczne, raporty z monitorowania wykorzystania systemu.</p> <p>KPI 6</p> <p>Metoda pomiaru – zliczenie liczby pracowników nie – IT (np. bibliotekarzy, redaktorów, pracowników merytorycznych), którzy uczestniczyli w szkoleniach w ramach projektu</p> <p>Częstotliwość – na rozpoczęcie i zakończenie projektu</p> <p>Źródło – dokumentacja projektowa, listy obecności, certyfikaty ukończenia</p>
Cel - 3	Rozwój infrastruktury informacji naukowej i interoperacyjności
Cel strategiczny	(1) Strategia na rzecz Odpowiedzialnego Rozwoju do roku 2020 (z perspektywą do 2030), cel szczegółowy III: Skuteczne państwo i instytucje służące wzrostowi oraz włączeniu społecznemu i gospodarczemu, Obszar: E-Państwo. (2) Europejska strategia w zakresie danych (COM/2020/66) – w zakresie wspierania Europejskiej Chmury Otwartej Nauki poprzez udostępnianie i interoperacyjność danych. (3) Program otwierania danych na lata 2021–2027, cel czwarty: Stymulowanie rynku ponownego wykorzystywania zasobów kultury i danych naukowych.
Korzyść:	(1) Zapewnienie integracji zasobów archiwalnych z innymi systemami informacji naukowej, co umożliwia ich łączne wyszukiwanie i analizę. (2) Zwiększenie interoperacyjności danych poprzez stosowanie wspólnych standardów i formatów, ułatwiających wymianę informacji między instytucjami. (3) Umożliwienie ponownego wykorzystania danych archiwalnych w różnych środowiskach badawczych i narzędziach cyfrowych. (4) Wzmocnienie infrastruktury wspierającej nowoczesne metody badań naukowych, w tym analizy masowych zbiorów danych i badań interdyscyplinarnych.
KPI:	<p>KPI 1: Liczba podmiotów, które udostępniły informacje sektora publicznego lub dane prywatne online</p> <p>KPI 2: Liczba podmiotów wspartych w zakresie rozwoju usług, produktów i procesów cyfrowych</p> <p>KPI 3: Liczba pracowników IT objętych wsparciem szkoleniowym</p> <p>KPI 4: Liczba utworzonych API</p>
Wartość aktualna i docelowa KPI:	<p>Wartość aktualna KPI 1: 0</p> <p>Wartość aktualna KPI 2: 0</p> <p>Wartość aktualna KPI 3: 0</p> <p>Wartość aktualna KPI 4: 0</p> <p>Wartość docelowa KPI 1: 1</p> <p>Wartość docelowa KPI 2: 1</p> <p>Wartość docelowa KPI 3: 0</p> <p>Wartość docelowa KPI 4: 1</p>
Metoda pomiaru KPI	<p>KPI 1</p> <p>Metoda pomiaru – osiągnięcie wartości docelowej (1) zostanie potwierdzone wdrożeniem rozwiązania umożliwiającego publikację danych przez co najmniej jeden podmiot</p> <p>Częstotliwość – na rozpoczęcie i zakończenie projektu</p> <p>Źródło – dokumentacja projektowa, protokoły odbioru</p> <p>KPI 2</p> <p>Metoda pomiaru – osiągnięcie wartości docelowej (1) potwierdzone udzieleniem wsparcia co najmniej jednemu podmiotowi</p>

	<p>Częstotliwość – na rozpoczęcie i zakończenie projektu Źródło – dokumentacja projektowa, wykaz podmiotów objętych wsparciem KPI 3 Metoda pomiaru – zliczenie liczby pracowników IT, którzy uczestniczyli w szkoleniach finansowanych w ramach projektu Częstotliwość – na rozpoczęcie i zakończenie projektu Źródło – dokumentacja projektowa, listy obecności, certyfikaty ukończenia KPI 4 Metoda pomiaru – zliczenie liczby API opracowanych, wdrożonych i uruchomionych w ramach projektu Częstotliwość – na rozpoczęcie i zakończenie projektu Źródło – dokumentacja projektowa, dokumentacja techniczna API, protokoły odbioru</p>
Cel - 4	Wsparcie dostępności i wykorzystania zasobów przez wdrożenie narzędzi AI
Cel strategiczny	(1) Strategia na rzecz Odpowiedzialnego Rozwoju do roku 2020 (z perspektywą do 2030), cel szczegółowy III: Skuteczne państwo i instytucje służące wzrostowi oraz włączeniu społecznemu i gospodarczemu, Obszar: E-Państwo. (2) Europejska strategia w zakresie danych (COM/2020/66) – w zakresie wspierania Europejskiej Chmury Otwartej Nauki poprzez udostępnianie i interoperacyjność danych. (3) Program otwierania danych na lata 2021–2027, cel czwarty: Stymulowanie rynku ponownego wykorzystywania zasobów kultury i danych naukowych.
Korzyść:	(1) Ułatwienie wyszukiwania i analizy zasobów archiwalnych dzięki inteligentnym mechanizmom rozpoznawania treści i kontekstu. (2) Zwiększenie efektywności pracy naukowej poprzez automatyzację procesów analizy i opracowania materiałów źródłowych. (3) Poprawa jakości dostępu do zasobów dla różnych grup użytkowników dzięki spersonalizowanym i intuicyjnym narzędziom AI. (4) Stworzenie warunków do rozwoju innowacyjnych metod badawczych opartych na analizie dużych zbiorów danych archiwalnych.
KPI:	<p>KPI 1: Użytkownicy nowych i zmodernizowanych publicznych usług, produktów i procesów cyfrowych KPI 2: Liczba pracowników niebędących pracownikami IT, objętych wsparciem szkoleniowym KPI 3: Liczba zapytań skierowanych do systemu z wykorzystaniem modułu zaawansowanego wyszukiwania zasobów cyfrowych opartego na sztucznej inteligencji KPI 4: Liczba dokumentów udostępnionych z wykorzystaniem modułu automatycznej transformacji dokumentów do postaci dostępnej opartej na sztucznej inteligencji</p>
Wartość aktualna i docelowa KPI:	<p>Wartość aktualna KPI 1: 0 Wartość aktualna KPI 2: 0 Wartość aktualna KPI 3: 0 Wartość aktualna KPI 4: 0 Wartość docelowa KPI 1: 10000 Wartość docelowa KPI 2: 5 Wartość docelowa KPI 3: 20000 Wartość docelowa KPI 4: 1000</p>
Metoda pomiaru KPI	<p>KPI 1 Metoda pomiaru – analiza wykorzystania systemu na podstawie logów systemowych oraz narzędzi analitycznych, potwierdzająca korzystanie z</p>

udostępnionej usługi cyfrowej w danym okresie referencyjnym
Częstotliwość – na rozpoczęcie projektu oraz na zakończenie projektu
Źródło – raporty systemowe (logi), narzędzia analityczne, raporty z monitorowania wykorzystania systemu.

KPI 2

Metoda pomiaru – zliczenie liczby pracowników nie-IT (np. bibliotekarzy, redaktorów, pracowników merytorycznych), którzy uczestniczyli w szkoleniach w ramach projektu

Częstotliwość – na rozpoczęcie i zakończenie projektu

Źródło – dokumentacja projektowa, listy obecności, certyfikaty ukończenia

KPI 3

Metoda pomiaru – zapytań skierowanych do systemu z wykorzystaniem modułu zaawansowanego wyszukiwania zasobów cyfrowych opartego na sztucznej inteligencji

Częstotliwość – na rozpoczęcie i zakończenie projektu

Źródło – narzędzia analityczne, raporty z monitorowania wykorzystania systemu.

KPI 4

Metoda pomiaru – zliczenie liczby dokumentów udostępniionych z wykorzystaniem modułu automatycznej transformacji dokumentów do postaci dostępnej oparty na sztucznej inteligencji

Częstotliwość – na rozpoczęcie projektu oraz na zakończenie projektu

Źródło – raporty systemowe

Należy zaznaczyć, że w projekcie świadomie odstąpiono od definiowania sztywnych mierników jakościowych AI (np. określonego procentu skuteczności OCR, poziomu precyzji wyszukiwania semantycznego) na poziomie globalnych wskaźników zadeklarowanych we wniosku o dofinansowanie. Wynika to z faktu, iż skuteczność algorytmów zależy będzie od zmiennego stanu zachowania fizycznego ksiąg oraz zróżnicowanych stylów pisma odręcznego (z lat od XVI do XX wieku). Mierniki jakości oraz dopuszczalny poziom błędów transformacji dokumentów zostaną natomiast rygorystycznie zdefiniowane jako parametry techniczne i warunki odbioru oprogramowania (SLA/kryteria jakościowe) na etapie tworzenia Opisu Przedmiotu Zamówienia (OPZ). Ponadto, zgodnie z założeniami mitygacji ryzyk, ostateczne wartości progowe dla skuteczności algorytmów sztucznej inteligencji zostaną dostrojone na podstawie wyników pilotażu digitalizacyjnego, przewidzianego w początkowej fazie realizacji projektu.

W ramach wskaźników (KPI) celowo nie zdefiniowano mierników opartych na tradycyjnej bibliometrii (np. liczba nowych publikacji naukowych, liczba cytowań). Wynika to wprost z wytycznych naboru FERC 2.3, w ramach którego wskaźniki programowe i własne zostały dobrane z naciskiem na budowę infrastruktury cyfrowej oraz otwarty wolumen i wykorzystanie udostępniionych danych (np. liczba użytkowników usług cyfrowych, rozmiar udostępniionych informacji, wygenerowane klucze API). Dodatkowo, wdrożenie w projekcie najwyższego standardu otwartości zasobów (5 Star Open Data) oraz licencji bezwarunkowych (CC0 Public Domain) poprzez zautomatyzowane interfejsy API, pozbawia Uczelnię możliwości technicznego śledzenia i wymuszania na

globalnych badaczach raportowania powstałych na tej bazie publikacji. Definiowanie tego typu wskaźników naukowych wiązałoby się z brakiem obiektywnej możliwości ich zmierzenia i udokumentowania w okresie trwałości projektu.

2.2. Udostępnione e-usługi

Lp.	Nazwa e-usługi	Typ	Zakres oddziaływania	Poziom dojrzałości e-usługi

2.3. Udostępnione informacje sektora publicznego i zdigitalizowane zasoby

Rodzaj informacji/zasobów	Planowana data udostępnienia	Szacowana liczba obiektów objętych digitalizacją (udostępnianiem informacji)
Zdigitalizowane i udostępnione księgi metrykalne parafii rzymskokatolickich (urodzeń, małżeństw i zgonów) z terenu dawnej archidiecezji lwowskiej	31-10-2029	2700
Zakonserwowany zasób archiwalny	30-04-2029	2700
Cyfrowe kopie materiałów archiwalnych	31-10-2029	2700

Czy wszystkie zdigitalizowane zasoby objęte projektem będą udostępniane bezpłatnie?

TAK/NIE

Liczba zasobów jaka nie zostanie udostępniona bezpłatnie: 34 %

2.4. Produkty końcowe projektu

Nazwa produktu	Planowana data wdrożenia
Raport z inicjalnego testu prywatności	01-2027
Infrastruktura obejmująca serwery aplikacyjne	05-2027
Digitization Process Management (DPM)	06-2027
Raport z testu prywatności	12-2027

Nazwa produktu	Planowana data wdrożenia
Raport z testów bezpieczeństwa	02-2028
Raport z testów wydajności	02-2028
Raport z testów badań UX	02-2028
Raport z audytu dostępności	02-2028
Repozytorium CRIS Uniwersytetu Papieskiego Jana Pawła II w Krakowie w zakresie wdrożenia modułu zaawansowanego wyszukiwania zasobów cyfrowych i modułu automatycznej transformacji dokumentów do postaci dostępnej opartych na sztucznej inteligencji oraz rozbudowa do wersji CRIS	03-2028
Interfejs API	08-2028
Materiały szkoleniowe	08-2028
Materiały informacyjno-promocyjne	02-2029
Zakonserwowany zasób archiwalny	04-2029
Cyfrowe kopie materiałów archiwalnych	10-2029
Konferencja na zakończenie projektu	12-2029

3. KAMIENIE MIŁOWE

Kamienie milowe	Planowany termin osiągnięcia
Przeprowadzony inicjalny test prywatności	2027-01-31
Powołanie Zespołu Zarządzania Projektem (ZZP), uruchomienie procedur zamówieniowych	2027-02-28
Rozstrzygnięte postępowania przetargowe na główne produkty projektu	2027-04-30
Zakończenie pilotażu digitalizacji ksiąg metrykalnych oraz ocena jakości przetwarzania	2027-04-30
Zakupiona infrastruktura obejmująca serwery aplikacyjne	2027-05-31
Wdrożony system Digitization Process Management (DPM)	2027-06-30
Wdrożony testowo system DSpace CRIS i rozpoczęte procesy testowania oprogramowania	2027-06-30
Wdrożony na środowisku testowym moduł zaawansowanego wyszukiwania zasobów cyfrowych opartego na mechanizmach inteligentnych	2027-08-31
Wdrożony na środowisku testowym moduł automatycznej transformacji skanów do postaci dostępnej	2027-09-30
Zakonserwowana i zdigitalizowana I część ksiąg metrykalnych (700 ksiąg)	2027-12-31
Zintegrowany DSpace CRIS z DataCite	2028-02-29
Uzyskany pozytywny wynik testów bezpieczeństwa	2028-02-29

Kamienie milowe	Planowany termin osiągnięcia
Uzyskany pozytywny wynik testów wydajności	2028-02-29
Uzyskany pozytywny wynik testów badań UX	2028-02-29
Uzyskany pozytywny wynik audytu dostępności WCAG 2.1 AA	2028-02-29
Zintegrowany system DSpace CRIS z systemami zewnętrznymi KRONiK@, dLibra	2028-06-30
Pracownicy przeszkoleni w zakresie wykorzystywania systemu i w tematyce dostępności cyfrowej	2028-08-31
Zakonserwowana i zdigitalizowana II część ksiąg metrykalnych (1000 ksiąg)	2028-12-31
Udostępniona I i II część ksiąg metrykalnych (łącznie 1700 ksiąg)	2028-12-31
Opracowane i opublikowane materiały informacyjno-promocyjne	2029-02-28
Zakonserwowana III część ksiąg metrykalnych (1000 ksiąg – łącznie całość zasobu: 2700 ksiąg)	2029-04-30
Zdigitalizowana i udostępniona III część ksiąg metrykalnych (1000 ksiąg – łącznie całość zasobu: 2700 ksiąg)	2029-10-31
Zorganizowana konferencja kończąca projekt	2029-12-31

4. KOSZTY

4.1. Koszty ogólne projektu wraz ze sposobem finansowania

Całkowity koszt projektu (netto oraz brutto), w tym	Netto 5 561 005,36 zł Brutto 6 371 565,38 zł	
Procent dofinansowania ze środków UE (brutto)	79,71%	
Procent środków z budżetu państwa (brutto)	20,29%	
Podział całkowitego kosztu projektu na poszczególne lata (netto oraz brutto)	2027	Netto 3 000 883,50 zł Brutto 3 520 532,37 zł
	2028	Netto 1 267 157,43 zł Brutto 1 409 807,14 zł
	2029	Netto 1 292 964,43 zł Brutto 1 441 225,87 zł

4.2. Wykaz poszczególnych pozycji kosztowych

Nazwa pozycji kosztowej		Przewidywany koszt brutto	Uzasadnienie pozycji kosztowej (przeznaczenie)
Oprogramowanie	<p>Koszt obejmuje wdrożenie oprogramowania DSpace-CRIS oraz systemu Digitization Process Management (DPM), w tym przygotowanie dokumentacji analitycznej, konfigurację i dostosowanie systemów do specyfiki zasobów archiwalnych oraz rozwój dodatkowych funkcjonalności. W ramach prac przewidziano realizację testów (deweloperskich, funkcjonalnych i integracyjnych), wsparcie ekspertów programistycznych oraz przygotowanie prototypów rozwiązań. Koszt obejmuje również działania związane z integracją systemów, procesem digitalizacji oraz udostępnianiem i publikacją danych i zasobów archiwalnych. Pozycja obejmuje także wynagrodzenia dwóch archiwistów jest niezbędnych do merytorycznego inwentaryzowania unikatowych metrykaliów oraz</p>	3 180 244,60 zł	<p>Zakup i wdrożenie systemów DSpace-CRIS oraz Digitization Process Management są niezbędne do zapewnienia profesjonalnego przechowywania, przetwarzania i udostępniania unikatowych zasobów archiwalnych, których obsługa nie jest możliwa w obecnym środowisku. Wdrożenie tych rozwiązań umożliwi efektywne zarządzanie procesem digitalizacji, automatyzację opracowania danych oraz zapewnienie ich jakości dzięki zaangażowaniu archiwistów i prac konserwatorskich. Inwestycja zapewni zwiększenie dostępności danych oraz ich interoperacyjność w ramach systemów informacji naukowej.</p>

Nazwa pozycji kosztowej		Przewidywany koszt brutto	Uzasadnienie pozycji kosztowej (przeznaczenie)
	zapewnienia poprawności danych wprowadzanych do systemów cyfrowych oraz nakłady na konserwację zachowawczą będącą kluczowym warunkiem digitalizacji.		
Infrastruktura	Infrastruktura teleinformatyczna, obejmująca zasoby sprzętowe, serwerowe oraz wartości niematerialne i prawne (w szczególności oprogramowanie systemowe i narzędziowe wraz z licencjami), niezbędna do uruchomienia, utrzymania i zarządzania systemem DSpace CRIS z nowymi funkcjonalnościami i zasobami.	300 000,00 zł	<p>Rozbudowa zasobów serwerowych o dodatkowe jednostki vCPU oraz pamięć RAM jest niezbędna, gdyż obecna infrastruktura nie pozwala na masowe przetwarzanie danych ani zaawansowaną analizę treści z użyciem AI. Nowe serwery umożliwią stabilne działanie systemów DSpace CRIS oraz Digitization Process Management, zapewniając wydajność niezbędną do obsługi baz danych i silników wyszukiwania.</p> <p>Koszty uwzględniają wyposażenie sprzętowe w dedykowane akceleratory graficzne (karty GPU), które są technicznie niezbędne do obsługi zaplanowanych w projekcie komponentów sztucznej inteligencji (algorytmów uczenia maszynowego HTR/OCR oraz NLP). W projekcie nie planuje się uruchamiania silników generatywnych tworzących nowe treści.</p>
Koszty UX i grafiki	Koszt obejmuje opracowanie projektu UX oraz warstwy graficznej systemów, z uwzględnieniem standardów dostępności. W ramach prac	312 122,30 zł	Opracowanie UX i grafiki jest niezbędne do stworzenia intuicyjnej i czytelnej platformy, która umożliwi szerokiemu gronu odbiorców, w tym genealogom i badaczom, sprawne przeszukiwanie oraz ponowne wykorzystanie unikatowych zasobów archiwalnych.

Nazwa pozycji kosztowej		Przewidywany koszt brutto	Uzasadnienie pozycji kosztowej (przeznaczenie)
	<p>przewidziano przeprowadzenie badań użytkowników, przygotowanie koncepcji funkcjonalnej i wizualnej, a także testowanie rozwiązań z udziałem użytkowników docelowych. Zakres obejmuje również wprowadzanie poprawek wynikających z testów oraz przeprowadzenie audytu dostępności i dostosowanie interfejsu do wymagań dostępności dla różnych grup odbiorców.</p>		<p>Uwzględnienie standardów dostępności (WCAG) stanowi kluczowy wymóg programu FERC i gwarantuje pełną inkluzywność zdigitalizowanych materiałów dla osób o zróżnicowanych potrzebach oraz z niepełnosprawnościami. Profesjonalne zaprojektowanie interfejsu znacząco podniesie użyteczność e-usług, wspierając realizację założeń otwartej nauki i ułatwiając dostęp do cyfrowego dziedzictwa międzynarodowej społeczności badaczy.</p>
Bezpieczeństwo	<p>Koszt obejmuje wdrożenie mechanizmów szyfrowania połączeń opartych na protokole HTTPS (SSL/TLS) oraz realizację audytów bezpieczeństwa i testów podatności systemu.</p>	191 414,90 zł	<p>Wydatki na bezpieczeństwo są niezbędne do zapewnienia poufności, integralności oraz stałej dostępności udostępnianych zasobów cyfrowych, zgodnie z ogólnymi zasadami bezpieczeństwa informacji. Środki te pozwolą na wdrożenie mechanizmów szyfrowania połączeń opartych na protokole HTTPS (SSL/TLS) oraz procedur technicznych chroniących dane przed utratą lub nieuprawnioną modyfikacją. Zastosowanie rozwiązań zgodnych z normami infrastruktury IT uczelni umożliwi skuteczne zarządzanie ryzykiem związanym z długoterminowym przechowywaniem unikatowego dziedzictwa archiwalnego.</p>
Wydajność rozwiązań	<p>Zapewnienie odpowiedniej mocy</p>	669 952,20 zł	<p>Zapewnienie odpowiedniej mocy obliczeniowej jest niezbędne dla</p>

Nazwa pozycji kosztowej		Przewidywany koszt brutto	Uzasadnienie pozycji kosztowej (przeznaczenie)
	obliczeniowej, w tym testów wydajnościowych		prawidłowego działania systemów przetwarzających i udostępniających duże zbiory danych archiwalnych, w tym obsługi procesów indeksowania, wyszukiwania oraz analizy treści. Przeprowadzenie testów wydajnościowych pozwoli na weryfikację stabilności i skalowalności systemu, identyfikację potencjalnych wąskich gardeł oraz dostosowanie konfiguracji infrastruktury do zakładanego obciążenia i liczby użytkowników.
Szkolenia	Szkolenia z dostępności cyfrowej i obsługi kluczowych systemów	195 000,00 zł	Szkolenia z zakresu dostępności (WCAG) są niezbędne dla zapewnienia pełnej inkluzywności cyfrowej udostępnianych zasobów oraz spełnienia wymogów formalnych programu FERC. Profesjonalne przygotowanie personelu do obsługi wdrożonych systemów (DSpaceCRIS, DPM) oraz narzędzi AI umożliwi sprawne zarządzanie procesem digitalizacji i poprawne publikowanie unikatowych materiałów archiwalnych. Inwestycja ta stanowi kluczowy element zarządzania ryzykiem, gwarantując transfer wiedzy oraz ciągłość kompetencji zespołu niezbędnych do długofalowego utrzymania efektów projektu.
Działania informacyjno-promocyjne	Koszty organizacji konferencji podsumowującej projekt oraz i przygotowania i wytwarzania materiałów promujących projekt	170 000,00 zł	Zorganizowanie konferencji naukowej oraz działań promocyjnych jest niezbędne do zwiększenia widoczności unikatowych zasobów archiwalnych dawnej archidiecezji lwowskiej w międzynarodowym obiegu naukowym. Działania te pozwolą skutecznie dotrzeć do szerokiego grona interesariuszy, informując ich o nowych możliwościach zdalnego dostępu oraz zaawansowanych narzędziach

Nazwa pozycji kosztowej		Przewidywany koszt brutto	Uzasadnienie pozycji kosztowej (przeznaczenie)
			wyszukiwania opartych na sztucznej inteligencji. Promocja projektu bezpośrednio wspiera realizację założeń otwartej nauki (Open Science), upowszechniając wiedzę o zdigitalizowanym zasobie i zachęcając do jego ponownego wykorzystania w badaniach interdyscyplinarnych.
Koszty zarządzania i wsparcia (w tym wynagrodzenia personelu wspomagającego)	Koszty wynagrodzeń personelu wspomagającego i koszty pośrednie	1 352 831,38 zł	<p>Kwota 936 000,00 zł: koszty personelu wspomagającego, czyli wynagrodzenia osób odpowiedzialnych za zarządzanie projektem, obsługę administracyjną, finansową i prawną, niezbędnych do prawidłowej realizacji przedsięwzięcia, zapewnienia zgodności z wymogami formalnymi oraz sprawnego rozliczenia projektu.</p> <p>Kwota 416 831,38 zł: koszty pośrednie (wyliczone jako zryczałtowana stawka 7 proc. od kwalifikowalnych kosztów bezpośrednich, zgodnie z mechaniką systemu LSI i zasadami naboru FERC) obejmujące wydatki związane z funkcjonowaniem zaplecza organizacyjnego projektu, w tym koszty administracyjne, obsługę finansowo-księgową, utrzymanie infrastruktury organizacyjnej oraz inne koszty ogólne, których nie można bezpośrednio przypisać do konkretnych działań merytorycznych, a które są niezbędne do zapewnienia ciągłości i efektywności realizacji projektu.</p>

4.3. Koszty ogólne utrzymania wraz ze sposobem finansowania (okres 5 lat)

Całkowity koszt utrzymania	1 234 743,60 zł	Źródło finansowania
-----------------------------------	-----------------	----------------------------

trwałości projektu (brutto)			
Podział całkowitego kosztu utrzymania trwałości projektu na poszczególne lata (netto oraz brutto)	2030	234 906,09 zł (brutto) (190 980,56 zł netto)	krajowe środki publiczne - budżet państwa
	2031	240 778,75 zł (brutto) (195 755,08 zł netto)	krajowe środki publiczne - budżet państwa
	2032	246 798,21 zł (brutto) (200 648,95 zł netto)	krajowe środki publiczne - budżet państwa
	2033	252 968,17 zł (brutto) (205 665,18 zł netto)	krajowe środki publiczne - budżet państwa
	2034	259 292,38 zł (brutto) (210 806,81 zł netto)	krajowe środki publiczne - budżet państwa

4.4. Planowane koszty ogólne realizacji (w przypadku projektu współfinansowanego – wkład krajowy z budżetu państwa) oraz koszty utrzymania projektu:

- zostaną pokryte w ramach budżetów odpowiednich dysponentów części budżetowych bez konieczności występowania o dodatkowe środki z budżetu państwa
- ~~- będą powodować konieczność przyznania dodatkowych kwot~~

5. GŁÓWNE RYZYKA

5.1. Ryzyka wpływające na realizację projektu

Nazwa ryzyka	Siła oddziaływania	Prawdopodobieństwo wystąpienia ryzyka	Sposób zarządzania ryzykiem
Niezrealizowanie zamówień publicznych i zakupów w planowanym terminie	Duża	Średnie	Wczesne rozpoczęcie procedur zakupowych, wsparcie wyspecjalizowanej jednostki zakupowej, bufor czasowy w harmonogramie.
Zły stan fizyczny części ksiąg metrykalnych	Duża	Wysokie	Przeprowadzenie pilotażu konserwatorskiego, etapowanie prac, ścisła współpraca z konserwatorami, elastyczne planowanie digitalizacji
Niska jakość skanów lub	Średnia	Średnie	Etap pilotażu i testy jakości, definiowanie minimalnych parametrów jakości,

Nazwa ryzyka	Siła oddziaływania	Prawdopodobieństwo wystąpienia ryzyka	Sposób zarządzania ryzykiem
przetwarzania OCR/HTR			odbiory cząstkowe, korekty parametrów przetwarzania
Problemy z integracją systemów (DSpace CRIS, DataCite, KRONiK@, dLibra)	Duża	Średnie	Wykorzystanie standardowych protokołów (API, OAI PMH), testy integracyjne na środowiskach testowych, wsparcie producentów systemów
Niedoszacowanie nakładu pracy na opracowanie metadanych	Średnia	Średnie	Standaryzacja schematów metadanych, automatyzacja części procesów, monitorowanie postępu, korekta harmonogramu
Brak wystarczających środków na realizację projektu	Średnia	Średnie	Zastosowanie rezerwy budżetowej oraz bieżącego monitoringu kosztów i harmonogramu w celu szybkiego identyfikowania ryzyka niedoboru środków, a także pozyskiwanie dodatkowego finansowania lub ograniczanie zakresu projektu poprzez priorytetyzację kluczowych działań.
Brak wystarczających zasobów kadrowych o odpowiednich kompetencjach do realizacji projektu	Duża	Średnie	Zapewnienie odpowiedniej rezerwy kadrowej, planowanie zasobów oraz bieżące monitorowanie obciążenia zespołu, a także w razie potrzeby pozyskanie dodatkowych pracowników lub outsourcing. Szkolenia pracowników.
Nieosiągnięcie wskaźników celów projektu	Duża	Średnie	Bieżące monitorowanie postępu realizacji wskaźników, dostosowywanie działań i zasobów do założeń oraz wprowadzanie działań naprawczych w przypadku odchyleń od planu.
Niepełne zidentyfikowanie zależności z innymi projektami	Średnia	Średnie	Identyfikowanie i analiza powiązań z innymi projektami na etapie planowania, bieżąca koordynacja między zespołami oraz regularna aktualizacja harmonogramów i zależności.

5.2. Ryzyka wpływające na utrzymanie efektów

Nazwa ryzyka	Siła oddziaływania	Prawdopodobieństwo wystąpienia ryzyka	Sposób zarządzania ryzykiem
--------------	--------------------	---------------------------------------	-----------------------------

Nazwa ryzyka	Siła oddziaływania	Prawdopodobieństwo wystąpienia ryzyka	Sposób zarządzania ryzykiem
Brak środków na utrzymanie infrastruktury informatycznej po zakończeniu projektu	Duża	Średnie	Zapewnienie finansowania utrzymaniowego w budżecie uczelni, włączenie systemu do stałej infrastruktury IT UPJPII
Brak wystarczających zasobów kadrowych o odpowiednich kompetencjach do utrzymania efektów projektu	Duża	Średnie	Dokumentacja techniczna i użytkowa, przekazanie wiedzy, szkolenia administratorów, zapewnienie ciągłości kadrowej
Dezaktualizacja oprogramowania lub komponentów integracyjnych	Średnia	Średnie	Wykorzystanie rozwiązań open source (DSpace CRIS), plan aktualizacji, monitorowanie zmian standardów i protokołów
Problemy z długoterminowym przechowywaniem danych	Duża	Niskie	Stosowanie polityk backupu, replikacja danych, utrzymanie infrastruktury zgodnie z dobrymi praktykami archiwizacji cyfrowej
Zmiany regulacyjne w zakresie udostępniania danych archiwalnych	Średnia	Niskie	Oparcie integracji na standardach (API, OAI PMH), elastyczna architektura systemu
Nieosiągnięcie wszystkich zaplanowanych korzyści	Średnia	Średnie	Zapewnienie mechanizmów utrzymania efektów poprzez planowanie działań po zakończeniu projektu, zabezpieczenie zasobów na fazę operacyjną oraz bieżące monitorowanie trwałości osiągniętych rezultatów.

6. OTOCZENIE PRAWNE

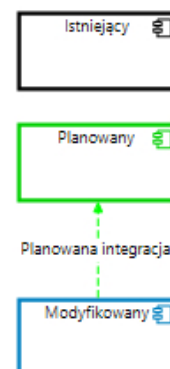
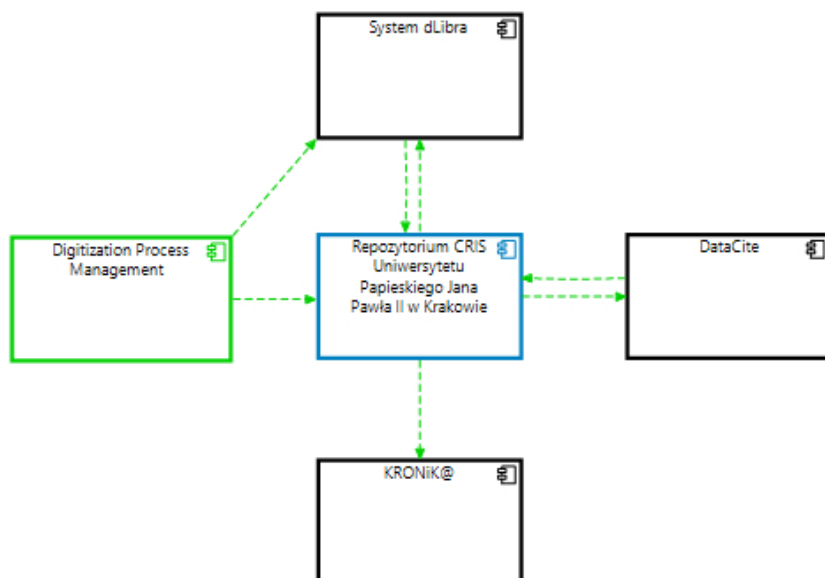
Lp.	Tytuł aktu prawnego	Czy wymaga zmian	Opis zmian (jeśli dotyczy)	Etap prac legislacyjnych (jeśli dotyczy)
1	Ustawa z dnia 20 lipca 2018 r. – Prawo o szkolnictwie wyższym i nauce (Dz.U. z 2024	TAK/NIE		

Lp.	Tytuł aktu prawnego	Czy wymaga zmian	Opis zmian (jeśli dotyczy)	Etap prac legislacyjnych (jeśli dotyczy)
	poz. 1571)			
2	Ustawa z dnia 11 sierpnia 2021 r. o otwartych danych i ponownym wykorzystywaniu informacji sektora publicznego (tekst jednolity: Dz.U. 2023 poz. 1524)	TAK/NIE		
3	Ustawa z dnia 17 lutego 2005 r. o informatyzacji działalności podmiotów realizujących zadania publiczne (tekst jednolity: Dz.U. z 2024 r. poz. 1557 i 1717)	TAK/NIE		
4	Rozporządzenie Rady Ministrów z dnia 21 maja 2024 r. w sprawie Krajowych Ram Interoperacyjności, minimalnych wymagań dla rejestrów publicznych i wymiany informacji w postaci elektronicznej oraz minimalnych wymagań dla systemów teleinformatycznych (Dz.U. 2024 poz. 773)	TAK/NIE		
5	Ustawa z dnia 5 lipca 2018 r. o krajowym systemie cyberbezpieczeństwa (tekst jednolity: Dz.U. 2024 poz. 1077 i 1222)	TAK/NIE		
6	Ustawa z dnia 27 lipca 2001 r. o ochronie baz danych (tekst jednolity: Dz.U. 2024 poz. 1769)	TAK/NIE		
7	Ustawa z dnia 11 sierpnia 2021 r. o otwartych danych i ponownym wykorzystywaniu informacji sektora publicznego (tekst jednolity: Dz.U. 2023 poz. 1524)	TAK/NIE		
8	Ustawa z dnia 4 kwietnia 2019 r. o dostępności cyfrowej stron internetowych i aplikacji mobilnych podmiotów publicznych (tekst jednolity: Dz.U. 2023 poz. 1440)	TAK/NIE		
9	Ustawa z dnia 7 maja 2010 r. o wspieraniu rozwoju usług i sieci telekomunikacyjnych (tekst jednolity: Dz.U. 2026 poz. 562)	TAK/NIE		
10	Ustawa z dnia 18 listopada 2020 r. o doręczeniach elektronicznych (tekst jednolity: Dz.U. 2026 poz. 3)	TAK/NIE		
11	Ustawa z dnia 5 września 2016 r. o usługach zaufania oraz identyfikacji elektronicznej (tekst jednolity: Dz.U. 2024 poz. 1725)	TAK/NIE		
12	ROZPORZĄDZENIE PARLAMENTU EUROPEJSKIEGO I RADY (UE) 2016/679 z dnia 27 kwietnia 2016 r. w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych i w sprawie swobodnego przepływu takich danych oraz uchylenia dyrektywy 95/46/WE (ogólne rozporządzenie	TAK/NIE		

Lp.	Tytuł aktu prawnego	Czy wymaga zmian	Opis zmian (jeśli dotyczy)	Etap prac legislacyjnych (jeśli dotyczy)
	o ochronie danych)			
13	Rozporządzenie Prezesa Rady Ministrów z dnia 20 lipca 2011 r. w sprawie podstawowych wymagań bezpieczeństwa teleinformatycznego (Dz.U. 2011 nr 159 poz. 948)	TAK/NIE		
14	Rozporządzenie Prezesa Rady Ministrów z dnia 14 września 2011 r. w sprawie sporządzania pism w formie dokumentów elektronicznych, doręczania dokumentów elektronicznych oraz udostępniania formularzy, wzorów i kopii dokumentów elektronicznych (tekst jednolity: Dz.U. 2018 poz. 180)	TAK/NIE		
15	Ustawa z dnia 14 lipca 1983 r. o narodowym zasobie archiwalnym i archiwach (tekst jednolity: Dz.U. 2020 poz. 164)	TAK/NIE		
16	Rozporządzenie Ministra Cyfryzacji z dnia 10 marca 2020 r. w sprawie szczegółowych warunków organizacyjnych i technicznych, które powinien spełniać system teleinformatyczny służący do uwierzytelniania użytkowników (Dz.U. 2020 poz. 399)	TAK/NIE		
17	Rozporządzenie Ministra Kultury i Dziedzictwa Narodowego z dnia 20 października 2015 r. w sprawie klasyfikowania i kwalifikowania dokumentacji, przekazywania materiałów archiwalnych do archiwów państwowych i brakowania dokumentacji niearchiwalnej (Dz.U. 2019 poz. 246)	TAK/NIE		
18	Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2024/1689 z dnia 13 czerwca 2024 r. ustanawiające zharmonizowane przepisy dotyczące sztucznej inteligencji (akt w sprawie sztucznej inteligencji)	TAK/NIE		
19	Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2022/868 z dnia 30 maja 2022 r. w sprawie europejskiego zarządzania danymi (Akt w sprawie zarządzania danymi)	TAK/NIE		
20	Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2023/2854 z dnia 13 grudnia 2023 r. w sprawie zharmonizowanych przepisów dotyczących sprawiedliwego dostępu do danych i ich wykorzystywania (Akt w sprawie danych)	TAK/NIE		

7. ARCHITEKTURA

7.1. Widok kooperacji aplikacji



Lista systemów wykorzystywanych w projekcie

Lp.	Nazwa systemu	Gestor systemu	Opis systemu	Status	Krótki opis ewentualnej zmiany
1	Repozytorium CRIS Uniwersytetu Papieskiego Jana Pawła II w Krakowie	Uniwersyte t Papieski Jana Pawła II w Krakowie	Repozytorium CRIS Uniwersytetu Papieskiego Jana Pawła II w Krakowie to system teleinformatyczny wspierający procesy gromadzenia, przechowywania,	Modyfikowany	Rozbudowa do klasy CRIS i dostosowanie do obsługi złożonych zbiorów archiwalnych. Wdrożenie

Lp.	Nazwa systemu	Gestor systemu	Opis systemu	Status	Krótki opis ewentualnej zmiany
			<p>opracowania i udostępniania zasobów archiwalnych wraz z metadanymi. System nie prowadzi rejestrów publicznych. Główne funkcjonalności obejmują archiwizację danych, wyszukiwanie pełnotekstowe oraz udostępnianie otwartych danych przez interfejsy API. System jest zintegrowany z systemami DataCite, KRONiK@ oraz System dLibra.</p>		<p>modułu automatycznej transformacji dokumentów do postaci dostępnej (funkcjonalności: mechanizmy HTR i OCR do transkrypcji pisma) oraz modułu zaawansowanego wyszukiwania zasobów cyfrowych (funkcjonalności: wyszukiwanie semantyczne NLP zintegrowane z silnikiem Apache Solr).</p>
2	DataCite	DataCite	<p>DataCite to system wspierający rejestrację i zarządzanie identyfikatorami DOI dla repozytoriów i zasobów badawczych. System nie prowadzi rejestrów publicznych. Główne funkcjonalności to nadawanie, walidacja i zarządzanie trwałymi identyfikatorami obiektów cyfrowych. System integruje się z platformami naukowymi na całym świecie, w tym z systemem Repozytorium CRIS Uniwersytetu Papieskiego Jana Pawła II w Krakowie.</p>	Istniejący	
3	KRONiK@	Ministerstwo Cyfryzacji	<p>KRONiK@ to system wspierający przechowywanie i udostępnianie zasobów z</p>	Istniejący	

Lp.	Nazwa systemu	Gestor systemu	Opis systemu	Status	Krótki opis ewentualnej zmiany
			zakresu nauki i kultury w jednym miejscu. System nie prowadzi rejestrów publicznych. Główne funkcjonalności to działanie jako multiwyszukiwarka integrująca rozproszone zasoby instytucji oraz pełnienie funkcji repozytorium zapasowego. Integruje się z systemami instytucji, w tym z systemem Repozytorium CRIS Uniwersytetu Papieskiego Jana Pawła II w Krakowie.		
4	System dLibra	Uniwersyte t Papieski Jana Pawła II w Krakowie	System dLibra to system wspierający gromadzenie, długoterminowe przechowywanie i udostępnianie cyfrowych zasobów naukowych oraz dziedzictwa kulturowego uczelni. System nie prowadzi rejestrów publicznych. Główne funkcjonalności to zarządzanie publikacjami, obsługa metadanych oraz prezentacja treści użytkownikom. System jest zintegrowany z wewnętrznymi systemami uczelni, w tym z Repozytorium CRIS oraz Digitization Process Management.	Istniejący	
5	Digitization Process Management	Uniwersyte t Papieski Jana Pawła II w Krakowie	Digitization Process Management to system wspierający zarządzanie procesem digitalizacji obiektów archiwalnych. System nie prowadzi rejestrów publicznych. Główne funkcjonalności obejmują rejestrowanie zadań, monitorowanie prac przy cyfryzacji oraz	Planowany	

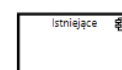
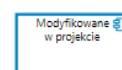
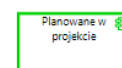
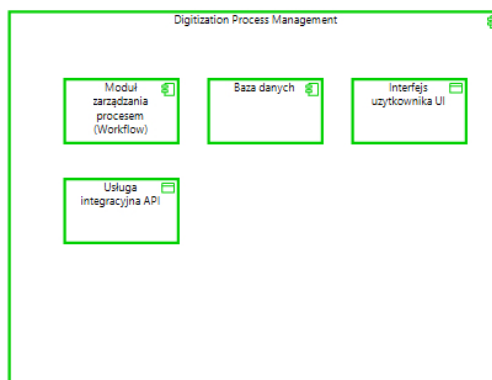
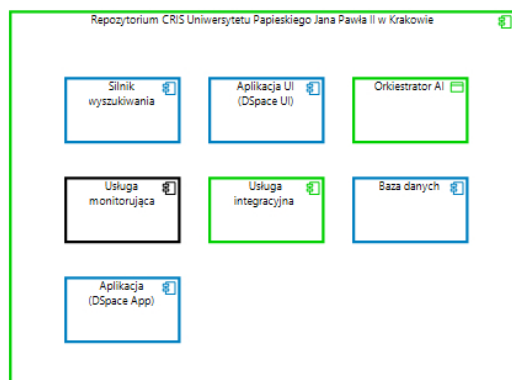
Lp.	Nazwa systemu	Gestor systemu	Opis systemu	Status	Krótki opis ewentualnej zmiany
			przygotowywanie obiektów cyfrowych do publikacji. System integruje się z Repozytorium CRIS oraz System dLibra		
6	ZoSIA (Zintegrowany System Informacji Archiwalnej)	Narodowe Archiwum Cyfrowe	ZoSIA jest systemem archiwalnym wykorzystywanym do ewidencji, opracowania oraz zarządzania materiałami archiwalnymi, który nie prowadzi rejestrów publicznych w rozumieniu ustawowym, lecz służy do tworzenia i utrzymywania uporządkowanej dokumentacji zasobu. System umożliwia wprowadzanie i zarządzanie szczegółowymi opisami archiwaliów wraz z ich strukturą, wspiera procesy klasyfikacji, porządkowania i opracowania materiałów, a także zapewnia funkcje wyszukiwania i przeglądania zasobu. Wspomaga również udostępnianie informacji o zbiorach użytkownikom. ZoSIA może współpracować z innymi systemami, w tym repozytoriami cyfrowymi i rozwiązaniami do prezentacji zasobów, poprzez mechanizmy eksportu danych i integracji, co umożliwia ich dalsze udostępnianie w krajowych i międzynarodowych środowiskach informacyjnych.	Istniejący	

Lista przepływów

Lp.	System źródłowy	System docelowy	Zakres wymienianych danych	Sposób wymiany danych	Typ modyfikacji	Typ interfejsu
1	Repozytorium CRIS Uniwersytetu Papieskiego Jana Pawła II w Krakowie	DataCite	Metadane obiektów cyfrowych (autorzy, tytuł, data publikacji, typ zasobu, identyfikatory lokalne) oraz adresy URL/ URI obiektów w repozytorium	tryb odwołań bezpośrednich	krytyczny dla sukcesu projektu	REST API
2	DataCite	Repozytorium CRIS Uniwersytetu Papieskiego Jana Pawła II w Krakowie	Metadane publikacji na podstawie numeru DOI	Kopiowanie danych	realizowalną inną metodą	REST API
3	Repozytorium CRIS Uniwersytetu Papieskiego Jana Pawła II w Krakowie	Krajowe Repozytorium Obiektów Nauki i Kultury (KRONIK@)	Metadane obiektów cyfrowych oraz informacje o zasobach (opisy, identyfikatory, linki do pełnych treści)	Kopiowanie danych	realizowalną inną metodą	REST API
4	Repozytorium CRIS Uniwersytetu Papieskiego Jana Pawła II w Krakowie	System dLibra	Pliki cyfrowe (skany, publikacje) oraz metadane opisowe i strukturalne	kopiowanie danych	realizowalną inną metodą	OAI-PMH
5	System dLibra	Repozytorium CRIS Uniwersytetu Papieskiego Jana Pawła II w Krakowie	Metadane zdigitalizowanych obiektów oraz informacje o plikach cyfrowych	kopiowanie danych	realizowalną inną metodą	OAI-PMH

Lp.	System źródłowy	System docelowy	Zakres wymienianych danych	Sposób wymiany danych	Typ modyfikacji	Typ interfejsu
6	Digitization Process Management	Repozytorium CRIS Uniwersytetu Papieskiego Jana Pawła II w Krakowie	Metadane obiektów, pliki cyfrowe przygotowane do publikacji	kopiowanie danych	krytyczny dla sukcesu projektu	API REST, import masowy (XML/JSON)
7	Digitization Process Management	System dLibra	Metadane obiektów, pliki cyfrowe przygotowane do publikacji	kopiowanie danych	krytyczny dla sukcesu projektu	API REST, import masowy (XML)
8	Repozytorium CRIS Uniwersytetu Papieskiego Jana Pawła II w Krakowie	ZoSIA (Zintegrowany System Informacji Archiwalnej)	Kopie cyfrowe materiałów archiwalnych oraz metadane	Kopiowanie danych, eksport	Planowany	Import masowy API REST

7.2. Kluczowe komponenty architektury rozwiązania



7.3. Przyjęte założenia technologiczne

Lp.	Obszar	Założenie technologiczne
1.	Infrastruktura	
2.	Sieć i bezpieczeństwo	Wdrożenie mechanizmów szyfrowania połączeń z repozytorium opartych na protokole HTTPS (SSL/TLS)
3.	Standardy wymiany danych	Realizacja komunikacji repozytorium ze środowiskiem zewnętrznym z użyciem standardu HTTPS
4.	Systemy operacyjne serwerowe	Oparcie środowiska systemowego na open source'owym systemie Linux z linii Debian
5.	Bazy danych	Oparcie warstwy bazodanowej na otwartoźródłowym systemie PostgreSQL
6.	Serwery aplikacji	Zastosowanie serwera aplikacyjnego Apache Tomcat w wersji open source do obsługi repozytorium; wykorzystanie systemu DSpace-CRIS jako platformy repozytorium cyfrowego
7.	Portale	
8.	Inne	Oparcie mechanizmów wyszukiwania na open source'owym rozwiązaniu Apache Solr oraz wykorzystanie Redis jako serwera cache

7.4. Opis zasobów danych przetwarzanych w planowanym rozwiązaniu

Czy nowy system będzie tworzył zasoby danych o charakterze rejestru publicznego?

TAK/NIE

Czy nowy system będzie przetwarzał (używał, zmieniał) zawartość innych rejestrów publicznych?

TAK/NIE

7.5. Bezpieczeństwo

Planowany poziom zapewnienia bezpieczeństwa (w rozumieniu przepisów §20 rozporządzenia Rady Ministrów z dnia 12 kwietnia 2012 r. w sprawie Krajowych Ram Interoperacyjności [...] (Dz. U. 2012, poz. 526 z późn. zm.) w zakresie dot. systemu zarządzania bezpieczeństwem informacji:

- system nie podlega rygorom KRI – należy wyjaśnić czy istnieją inne normy bezpieczeństwa, które będą spełnione przez system zgodnie z wymogami KRI

Planowany system repozytoryjny nie będzie przetwarzał danych wrażliwych ani pełnił funkcji rejestru publicznego, w związku z czym nie podlega rygorom Krajowych Ram Interoperacyjności w zakresie minimum dla systemów teleinformatycznych o podwyższonym poziomie bezpieczeństwa, o których mowa w § 20 rozporządzenia KRI. System będzie wykorzystywany do udostępniania zdigitalizowanych materiałów archiwalnych w celach naukowych i informacyjnych, z zachowaniem ograniczeń wynikających z przepisów prawa archiwalnego. Jednocześnie system zostanie zaprojektowany i utrzymywany zgodnie z ogólnymi zasadami bezpieczeństwa informacji określonymi w KRI, w tym w szczególności w zakresie zapewnienia poufności, integralności i dostępności informacji. Zastosowane zostaną procedury i mechanizmy organizacyjne oraz techniczne zgodne z dobrymi praktykami bezpieczeństwa, w tym elementy systemu zarządzania bezpieczeństwem informacji opartego na normach obowiązujących w infrastrukturze IT Uniwersytetu Papieskiego Jana Pawła II w Krakowie.

Należy podkreślić, że wdrażane w projekcie narzędzia sztucznej inteligencji (HTR, OCR, NLP) nie stanowią systemów wysokiego ryzyka w świetle przepisów AI Act, ponieważ ich jedynym przeznaczeniem jest transkrypcja i indeksowanie historycznych dóbr nauki. Ponadto dla wdrożenia systemu nie zachodzi konieczność przeprowadzania Oceny Skutków dla Ochrony Danych (DPIA). RODO nie ma zastosowania do informacji o osobach zmarłych, a w przypadku nowszych ksiąg objętych ograniczeniami udostępniania (ok. 34 proc. całego zasobu) ochrona realizowana jest poprzez bezwzględne zautomatyzowane embargo czasowe (100 lat dla urodzeń, 80 lat dla małżeństw i zgonów) w warstwie dostępu publicznego systemu DSpace-CRIS, co jednoznacznie i z góry wyklucza wystąpienie wysokiego ryzyka dla praw osób fizycznych.

~~-dodatkowe zabezpieczenia powyżej wymogów KRI: należy wskazać uzasadnienie~~