



Koncepcja architektury informacji i koncepcja e-usług

Opracowanie zrealizowane w ramach projektu
„Workflow do opracowania dokumentów cyfrowych na uczelni”
[https://uni.wroc.pl/projekty-uwr/projekty-infrastrukturalne/
workflow-do-opracowania-publicacji-cyfrowych-na-uczelni/](https://uni.wroc.pl/projekty-uwr/projekty-infrastrukturalne/workflow-do-opracowania-publicacji-cyfrowych-na-uczelni/)

Opracowane przez:

Eunika Gielniewska
Tomasz Kalota
Piotr Malak
Rafał Olszak
Anna Osowska
Rafał Raczyński
Marcin Szala
Aleksandra Zawadzka
Marcin Zubel

Data opracowania:
17.04.2021 r.

Publikacja udostępniona na licencji
Creative Commons: Uznanie autorstwa,
Na tych samych warunkach 4.0,
Międzynarodowe ([CC BY-SA 4.0](https://creativecommons.org/licenses/by-sa/4.0/))

Kierownik projektu
pn. "Workflow do opracowania
publikacji cyfrowych na uczelni"
nr POPC.02.04.00-00-0018/20

Tomasz Kalota



Fundusze
Europejskie
Polska Cyfrowa



Rzeczpospolita
Polska

Unia Europejska
Europejski Fundusz
Rozwoju Regionalnego





Spis treści

1	Cel projektu	3
2	Terminy i skróty stosowane w dokumencie.....	3
3	Podstawowe założenia projektu	6
4	Opis zakresu prac	8
4.1	Zadanie nr 1. Modernizacja bazy wiedzy UWr pod kątem pełnotekstowego przeszukiwania zasobów.....	8
4.1.1	Import metadanych i uruchomienie modułu sterowni	8
4.1.2	Uprozczone mapowanie pomiędzy BPPiD, ADN i dLibrą (RUWr/BCUWr).....	8
4.1.3	Funkcje „sterowni” dla Wirtualnego Laboratorium Transkrypcji (WLT)	10
4.1.3.1	Transkrypcja rękopisów w TEI.....	10
4.1.3.2	Obsługa transkrypcji muzycznych.....	10
4.1.3.3	Adnotacje na obiektach graficznych	12
4.1.3.4	Obsługa OCR w formacie ALTO	12
4.2	Zadanie nr 2: Rozbudowa Wirtualnego Laboratorium Transkrypcji	13
4.2.1	Transkrypcja rękopisów w TEI	15
4.2.2	Obsługa transkrypcji muzycznych	17
4.2.3	Adnotacje na obiektach graficznych	18
4.2.4	Obsługa OCR w formacie ALTO	18
4.3	Zadanie nr 3. Rozbudowa Repozytorium UWr.....	18
5	Wymagania pozafunkcjonalne	22
5.1	Spełnienie standardów dostępności WCAG.....	22
5.2	Spełnienie standardów otwartości.....	23
5.3	Wymagania dotyczące dokumentacji API.	24
5.4	Spełnienie wymagań interoperacyjności	24
6	Platforma sprzętowa	25
7	Wymagania dotyczące realizacji projektu.....	25





1 Cel projektu

Celem głównym Projektu „WORKFLOW DO OPRACOWANIA PUBLIKACJI CYFROWYCH NA UCZELNI” jest modernizacja i rozbudowa posiadanych przez Uniwersytet Wrocławski narzędzi i usług wspierających zarządzanie, przetwarzanie i publikowanie archiwalnego dorobku naukowego Uniwersytetu Wrocławskiego. Przedsięwzięcie realizowane jest w ramach Programu Operacyjnego Polska Cyfrowa na lata 2014-2020. Nr. POPC.02.04.00-00-0018/20-00. Oś Priorytetowa nr 2 „E-administracja i otwarty rząd”, Działanie nr 2.4 „Tworzenie usług i aplikacji wykorzystujących e-usługi publiczne i informacje sektora publicznego”.

W ramach projektu realizowane będą trzy zadania:

Zadanie nr 1: Modernizacja bazy wiedzy UWr pod kątem pełnotekstowego przeszukiwania zasobów

Zadanie nr 2: Rozbudowa Wirtualnego Laboratorium Transkrypcji

Zadanie nr 3: Rozbudowa Repozytorium UWr

2 Terminy i skróty stosowane w dokumencie

ACUWr – Archiwum Cyfrowe Uniwersytetu Wrocławskiego, usługa cyfrowa dostępna pod adresem <https://archiwum.uni.wroc.pl/>.

ADN (Agregator Dorobku Naukowego) - system umożliwiający agregowanie metadanych z Bazy publikacji pracowników i doktorantów Uniwersytetu Wrocławskiego dostępnej pod adresem <https://luna.bu.uni.wroc.pl/publikacje/> oraz innych baz danych umożliwiających eksport metadanych przy pomocy protokołu OAI-PMH.

Agregacja danych - gromadzenie danych z różnych baz danych w celu przygotowania połączonych zestawów danych do ich wspólnego przetwarzania w jednej bazie (kompilacja różnych źródeł).

BCUWr – Biblioteka Cyfrowa Uniwersytetu Wrocławskiego, usługa cyfrowa dostępna pod adresem <https://www.bibliotekacyfrowa.pl/>.

DC (Dublin Core) – standard metadanych, stosowany do opisu zasobów cyfrowych i wymiany informacji o nich w środowisku Internetu, opracowany przez Dublin Core Metadata Initiative (DCMI). Powszechnie stosowany w polskich bibliotekach cyfrowych.

dLibra (ang. Digital Library Framework) – oprogramowanie służące do budowy bibliotek cyfrowych, rozwijane przez Poznańskie Centrum Superkomputerowo-Sieciowe. Umożliwia tworzenie repozytoriów dokumentów cyfrowych oraz udostępnianie ich użytkownikom





fizycznym i systemom w Internecie. Obsługiwane protokoły wymiany danych: RSS, RDF, MARC, DublinCore czy OAI-PMH. Wersja w posiadaniu UWr.: 6.2.7.

DOI (ang. Digital Object Identifier) – cyfrowy identyfikator dokumentu elektronicznego, który, w odróżnieniu od identyfikatorów URL, nie zależy od fizycznej lokalizacji dokumentu, lecz jest na stałe przypisany do dowolnego przedmiotu własności intelektualnej; jego zadaniem jest stałe identyfikowanie dowolnych obiektów własności intelektualnej w sieciach cyfrowych w powiązaniu z aktualnymi danymi na ich temat. Więcej informacji pod adresem: <http://www.doi.org/>.

ePUB/MOBI – formaty plików do publikacji elektronicznych.

Europeana – biblioteka cyfrowa, wirtualne muzeum i archiwum, mające za cel udostępnienie dziedzictwa kulturowego i naukowego Europy w Internecie, otwarta w ramach inicjatywy UE: i2010 na rzecz bibliotek cyfrowych. Jest to wspólny punkt dostępu do zbiorów bibliotek, archiwów i muzeów w całej Europie, który umożliwia użytkownikom znalezienie digitalizatów dzieł kultury, udostępnionych przez instytucje kulturalne w całej Unii Europejskiej.

FRBR (ang. Functional Requirements for Bibliographic Records) – koncepcyjny model danych bibliograficznych, służący do opisu dokumentów.

IIIF (ang. International Image Interoperability Framework) – specyfikacja obejmująca zestaw API (interfejsów programowania aplikacji) umożliwiająca uruchamianie interoperacyjnych funkcjonalności w obrębie cyfrowych repozytoriów obrazu. Strona główna inicjatywy: <http://iiif.io>.

Indeks (lista indeksowa) – lista słów (terminów) generowana automatycznie z opisów obiektów w systemie komputerowym.

IZZ (Inwentarz Zbiorów Zdigitalizowanych) – oprogramowanie własne Biblioteki UWr służące do obsługi procesu digitalizacji, z wykorzystaniem Sencha. Strona Platform. Strona WWW projektu: <https://www.sencha.com/>.

JPEG2000 – standard kompresji obrazu opracowany jako uzupełnienie istniejących, bardzo popularnych technik kompresji JPEG.

MARC21 (ang. Machine-Readable Cataloging) – standard katalogowania danych bibliograficznych w formie elektronicznej, od wielu lat powszechnie stosowany w bibliotekach, w tym w bibliotece UWr.

MCUWr – Muzeum Cyfrowe Uniwersytetu Wrocławskiego, usługa cyfrowa dostępna pod adresem <https://muzeumcyfrowe.pl/>.

METS (ang. Metadata Encoding and Transmission Standard) – standard kodowania metadanych opisowych, strukturalnych i administracyjnych obiektów w bibliotekach cyfrowych.

OAI lub OAI-PMH (ang. Open Archives Initiative Protocol for Metadata Harvesting) – ustandaryzowany i rozpowszechniony protokół stosowany w sieci komputerowej (zwłaszcza w Internecie) przeznaczony do udostępniania i pobierania metadanych.





OCR (ang. Optical Character Recognition) – optyczne rozpoznawanie znaków. Proces tworzenia na podstawie obrazu rastrowego tekstu (np. zdjęcia lub skanu strony druku) tekstu, pozwalającego na edycję, przeszukiwanie treści itp.

Projekt – całość przedsięwzięcia pod nazwą: „WORKFLOW DO OPRACOWANIA PUBLIKACJI CYFROWYCH NA UCZELNI”

Publikacja wieloformatowa – typ publikacji w systemie dLibra (od wersji 6.0) charakteryzujący się możliwością jednoczesnego użycia w danej publikacji wielu formatów prezentacyjnych (np. pdf oraz djvu).

RUWr – Repozytorium Uniwersytetu Wrocławskiego, usługa cyfrowa dostępna pod adresem <https://www.repozytorium.uni.wroc.pl/>.

RWD (ang. Responsive Web Design) – technika projektowania stron www, zgodnie z zasadą responsywności, ideą responsywności jest tworzenie stron internetowych, których graficzna postać dostosowuje się do rozmiarów okna przeglądarki internetowej, przyjmując optymalny układ treści w zależności od rozdzielczości ekranu urządzenia. Rozwiązanie znajduje zastosowanie w szczególności w urządzeniach mobilnych (smartfon, tablet).

SSL (ang. Secure Socket Layer) – protokół typu klient-serwer pozwalający na nawiązanie bezpiecznego połączenia z użyciem certyfikatów x509. Umożliwia bezpieczną transmisję zaszyfrowanego strumienia danych.

TEI (ang. Text Encoding Initiative) – standard elektronicznej reprezentacji tekstu wraz z informacją o jego treści, pierwotnie opracowany w SGML, następnie także w XML. Więcej informacji pod adresem: <http://www.tei-c.org/>.

TIFF piramidalny – (ang. Tagged Image File Format) – prezentacyjny format plików graficznych, będący rozszerzeniem formatu TIFF.

Unicode – uniwersalny standard kodowania znaków przypisujący unikalny numer każdemu znakowi, niezależny od używanej platformy, programu czy języka.

UWr – Uniwersytet Wrocławski.

W3C (ang. World Wide Web Consortium (W3C)) – organizacja, która zajmuje się ustanawianiem standardów stron WWW.

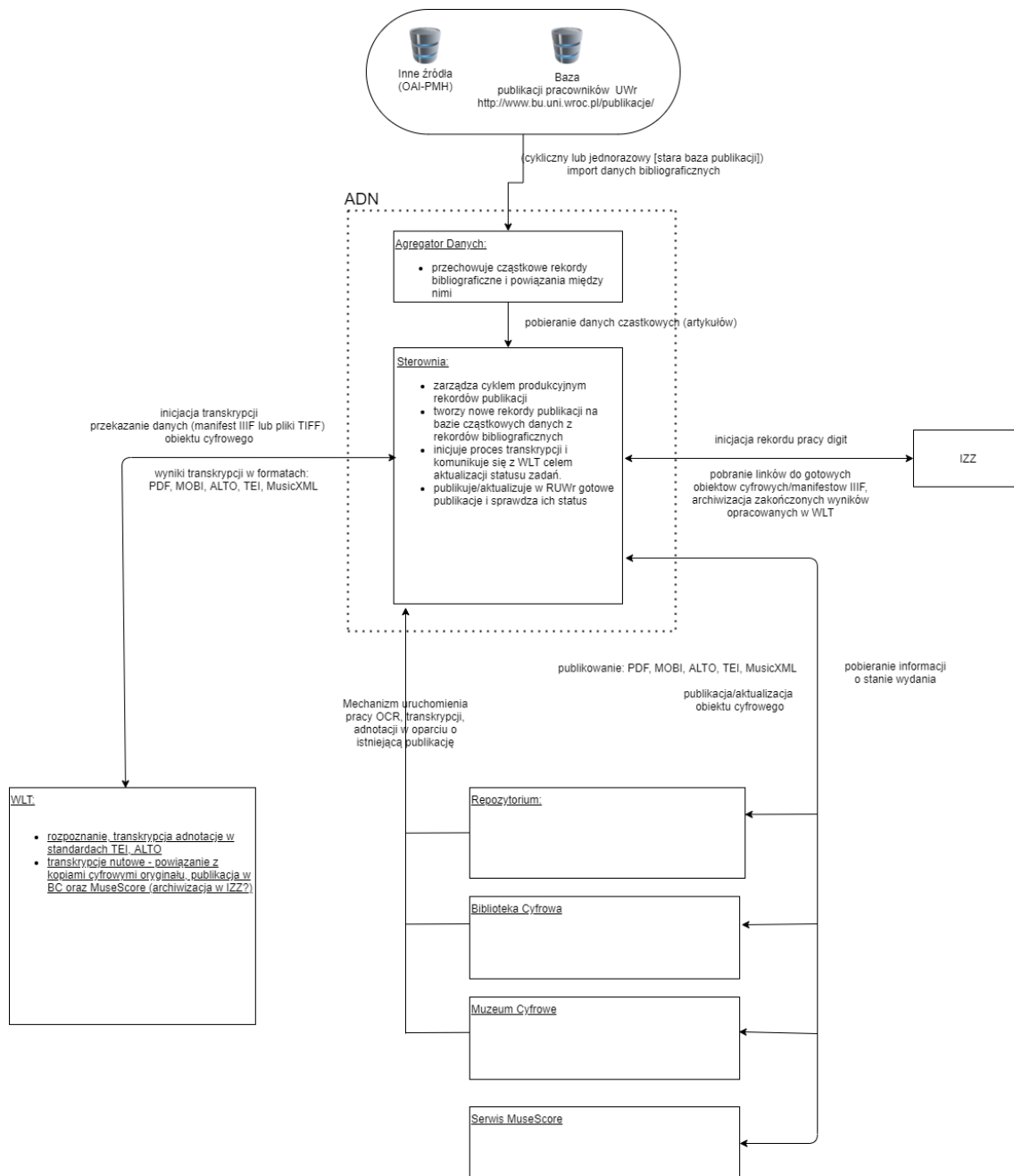
WCAG (ang. Web Content Accessibility Guidelines – Wytyczne Dotyczące Dostępności Treści Internetowych) – zbiór dokumentów opublikowany przez WAI (ang.: Web Accessibility Initiative, pl.: Inicjatywa dostępności do sieci); inicjatywa W3C mająca na celu zwiększenie szeroko rozumianej dostępności stron www, zawierająca zalecenia dotyczące tworzenia serwisów internetowych.

WLT (Wirtualne Laboratorium Transkrypcji) – autonomiczny moduł transkrypcji, konwertujący dane wejściowe w postaci obrazu (np. plik piramidalny, pdf, jpg, djvu) na postać tekstową, wykorzystujący narzędzia typu OCR oraz narzędzia umożliwiające tworzenie i przechowywanie adnotacji odnoszących się do zawartości obrazu cyfrowego.





3 Podstawowe założenia projektu



Rys.1 Blokowy schemat powiązań pomiędzy systemami





Przebieg procesu ucyfrowienia archiwalnej publikacji naukowej:

1. Użytkownik loguje się do systemu zarządzania zadaniami - ADN (Agregator Dorobku Naukowego) i tworzy nowe zadanie.
2. Do nowego zadania stara się wyszukać w bazie ADN rekord bibliograficzny artykułu oraz powiązany z nim (poprzez atrybut źródło) wpis formy wydawniczej (zeszytu w tym przypadku) określającej fizyczny materiał do zeskanowania.
3. Na podstawie wyszukanych danych bibliograficznych użytkownik określa stan praw autorskich i możliwość udostępniania w Internecie za pośrednictwem Repozytorium Uniwersytetu Wrocławskiego.
4. Użytkownik, mając komplet danych, oznacza status zadania na „gotowy do przetwarzania/wymagana digitalizacja”. Jest to pierwszy (startowy) poziom gotowości digitalizacyjnej.
5. ADN tworzy automatycznie rekord w IZZ (Inwentarz Zbiorów Zdigitalizowanych) i napełnia go podstawowymi danymi pozyskanymi z rekordu zadania.
6. Użytkownik przekazuje fizyczny egzemplarz czasopisma do BUWr i czeka na wykonanie digitalizacji lub dostarcza skany wykonane we własnym zakresie.
7. PRiD (Pracownia Reprografii i Digitalizacji) wykonuje digitalizację, jednocześnie ADN cyklicznie sprawdza w IZZ status realizacji zadań. W momencie osiągnięcia końca procesu digitalizacji (IZZ kończy tworzyć wersję IIP/IIIF) ADN automatycznie tworzy rekord zadania w WLT (Wirtualne Laboratorium Transkrypcji) i przekazuje do niego niezbędne informacje inicjujące (link do manifestu IIIF, nazwę zadania, identyfikator użytkownika ADN/WLT (SSO) itp.).
8. WLT wykonuje automatyczną transkrypcję, a ADN cyklicznie sprawdza, czy pliki są już przetworzone. W momencie osiągnięcia końca procesu wstępnego przetwarzania (transkrypcji) ADN powiadamia użytkownika (zmiana statusu zadania/wysyłka maila itd.) o konieczności jego ręcznej interwencji/decyzji.
9. Użytkownik pracuje z dokumentem w WLT (link do danego zadania w WLT powinien być dostępny z poziomu odpowiedniego zadania w ADN i nie powinno być wymagane dzięki SSO kolejne logowanie). Opcjonalnie użytkownik powinien mieć opcję ustawienia znacznika pracy w ADN „potrzebuję pomocy przy transkrypcji”, tak aby zasygnalizować, iż temat transkrypcji w WLT jest mu obcy i liczy na to, iż inna osoba (redaktor/moderator z WLT) wykona za niego korektę transkrypcji w WLT.
10. Po zakończeniu pracy w WLT (przez użytkownika bądź inną osobę realizującą daną transkrypcję) ADN wykrywa jej ukończenie i inicjuje publikację rekordu w Repozytorium UWr.
11. ADN pobiera z RUWr link do gotowej publikacji oraz ustawia status zadania na „zrobione”, kończąc proces.





4 Opis zakresu prac

4.1 Zadanie nr 1. Modernizacja bazy wiedzy UWr pod kątem pełnotekstowego przeszukiwania zasobów

4.1.1 Import metadanych i uruchomienie modułu sterowni

W ramach tego zadania zostanie wykonany import metadanych z obecnie istniejącej bazy publikacji pracowników naukowych dostępnej pod adresem - <https://luna.bu.uni.wroc.pl/publikacje/index.php> do Agregatora Dorobku Naukowego (ADN). Agregator Dorobku Naukowego jest usługą cyfrową umożliwiającą gromadzenie informacji na temat obiektów wymagających opracowania cyfrowego oraz zarządzanie i sterowanie procesem cyfryzacji realizowanym przy pomocy zintegrowanych usług, czyli Wirtualnego Laboratorium Transkrypcji oraz Repozytorium UWr. ADN zostanie wyposażony w funkcjonalności umożliwiające autorowi publikacji zainicjowanie procesu przygotowania własnej publikacji w wersji cyfrowej i zamieszczenie jej w Repozytorium Uniwersytetu Wrocławskiego (moduł „Sterownia” na Rys. 1). W tym celu zostanie wykonana integracja ADN z Wirtualnym Laboratorium Transkrypcji (WLT), w którym będzie można przygotować plik zgodnie ze standardami zalecanymi przez Google Scholar. Po wyprodukowaniu cyfrowej wersji publikacji zostanie ona udostępniona w Repozytorium Uniwersytetu Wrocławskiego (RUWr). Do zrealizowania tego zadania konieczne jest:

- import metadanych przechowywanych w bazie publikacji pracowników i doktorantów UWr - <https://luna.bu.uni.wroc.pl/publikacje/index.php> do Agregatora Dorobku Naukowego (ADN);
- wyposażenie Agregatora Dorobku Naukowego (ADN) w zestaw statusów określających poziomy gotowości digitalizacyjnej, umożliwiających komunikację pomiędzy użytkownikami obsługującymi proces przygotowania publikacji cyfrowej;
- integracja ADN z Wirtualnym Laboratorium Transkrypcji (WLT) umożliwiająca zainicjowanie produkcji wersji cyfrowej na podstawie danych przechowywanych w ADN;
- integracja Agregatora Dorobku Naukowego (ADN) z Repozytorium Uniwersytetu Wrocławskiego (RUWr), umożliwiająca synchronizację metadanych, inicjowanie procesu przesyłania publikacji do RUWr oraz dostęp do wersji cyfrowych publikacji w RUWr.

4.1.2 Uprozczone mapowanie pomiędzy BPPiD, ADN i dLibrą (RUWr/BCUWr)

Tabela pokazuje powiązania między atrybutami, które w zależności od typu obiektu muszą być rozpisane na szczegółowe przypadki (łączenie wartości z BPPiD w jedną wartość do ADN lub rozbicie zawartości z jednego pola z BPPiD na więcej niż jeden atrybut w ADN, opcjonalnie użycie tylko części pola).





pobierane z BPPiD - MARC21	atrybuty w ADN	eksport do dLibry (rekord inicjalny RUWr/BCUWr do uzupełnienia) - RDF
020a	ISBN	Identifier
022a	ISSN	Identifier
024a	DOI	DOI
041a	język publikacji	Language
041b	język streszczenia	LanguageOfAbstract
090a	data wydania	Issued
100a	autor	Creator
245abnp	tytuł	Title
246ab	tytuł odmienny	AlternativeTitle
250	wydanie	Description
260a	miejsce wydania	PlaceOfPublishing
260b	wydawca	Publisher
260c	rok wydania	Issued
300	opis fizyczny	PhysicalDescription
392	tom	Title
440	seria/tytuł tomu	GroupTitle
490	seria/tytuł tomu	GroupTitle
500	opis	Description
504	opis	Description
520	opis	Description
546	opis	Description
592a	opis	Description
700a	autor	Creator
700ae	współtwórca	Contributor
773t	seria/tytuł tomu	GroupTitle
773d	miejsce wydania, wydawca, rok wydania	PlaceOfPublishing, Publisher, Issued
773k	seria/tytuł tomu	GroupTitle
773g	opis fizyczny	PhysicalDescription
773tcdkgz	źródło	BibliographicCitation
787tg	źródło	BibliographicCitation
830	seria/tytuł tomu	GroupTitle
856u	wersja cyfrowa	HasVersion
856f	licencja	License
LDR d	typ obiektu	DetailedType
dodatkowe atrybuty do ADN:	prawa dostępu	AccessRight
	właściciel praw	RightsHolder
	data zastrzeżenia praw autorskich	DateCopyrighted
	katalog BUWr	CatalogBUWr
	sygnatura	Source





4.1.3 Funkcje „sterowni” dla Wirtualnego Laboratorium Transkrypcji (WLT)

4.1.3.1 *Transkrypcja rękopisów w TEI*

Warunki startowe:

- użytkownik posiada wiedzę w zakresie standardu TEI (Text Encoding Initiative), służącego do opracowania i reprezentacji tekstów w formie cyfrowej;
- dostępna jest cyfrowa kopia obiektu oryginalnego hostowanego na infrastrukturze UWr lub innej zewnętrznej instytucji kultury w standardzie IIIF Presentation API;
- użytkownik posiada konto użytkownika w systemie Workflow.

Cel:

- wykonanie w WLT struktury rękopisu zgodnie ze schematem TEI;
- publikacja opracowanej struktury rękopisu w BCUWr, MCUWr jako część IIIF Presentation API w części „structures” prezentowane podobnie jak w niniejszym przykładzie: Pan Tadeusz T2 - https://glam.ossolineum.pl/browser/uv2/app.html?manifestUri=https://iip2.bu.uni.wroc.pl/ms/manifest/ZNIO/6932_2_9.json - zakładka Index.

Przebieg procesu:

- a) Użytkownik inicjuje nową transkrypcję rękopisu w WLT za pośrednictwem systemu ADN.
- b) ADN automatycznie odszukuje rekord w Inwentarzu Zbiorów Zdigitalizowanych IZZ, aby stworzyć komunikację na potrzeby późniejszej archiwizacji, uruchamia nową transkrypcję w WLT na podstawie IIIF manifest.
- c) Po zakończeniu prac transkrypcyjnych i edytorskich w WLT (100% OCR) właściciel transkrypcji oznacza status np. „Publikacja”, który jest wyrażeniem zgody na rozpoczęcie procedury archiwizacji i publikacji.
- d) ADN rozpoczyna procedurę archiwizacji nowej zawartości treściowej w systemie IZZ oraz publikacji wyników w RUWr, BCUWr lub MCUWr za pomocą IIIF Presentation API w części „structures”.

4.1.3.2 *Obsługa transkrypcji muzycznych*

Warunki startowe:

- użytkownik posiada wykonaną za pomocą zewnętrznego oprogramowania do pisania nut (Muscore, Capella, Sibelius, Finale) transkrypcję muzyczną, której źródło w postaci kopii cyfrowej obiektu oryginalnego jest hostowane na infrastrukturze UWr;



- dostępna jest cyfrowa kopia obiektu oryginalnego hostowanego na infrastrukturze UWr lub innej zewnętrznej instytucji kultury w standardzie IIIF Presentation API;
- użytkownik posiada konto użytkownika w systemie Workflow (ADN/WLT - konto zintegrowane).

Cel:

- import pliku transkrypcji nutowej, opatrzenie jej podstawowymi wartościami opisu oraz powiązanie jej z cyfrową kopią obiektu oryginalnego hostowanego na infrastrukturze UWr;
- publikacja transkrypcji nutowej w RUWr, BCUWr lub MCUWr jako obiekt wieloformatowy, gdzie pierwszym formatem będzie manifest IIIF (widok oryginału), który łączy wprowadzone do projektu w WLT linki, zaś drugim sama transkrypcja w postaci pliku MusicXML;
- efektem prezentacji w systemach RUWr, BCUWr lub MCUWr będzie prezentacja wiążąca materiał źródłowy (reprodukcję cyfrową) i wykonaną na tej podstawie transkrypcję;
- zalecaną formą prezentacji pliku w formacie MusicXML jest zastosowanie dedykowanej przeglądarki umożliwiającej wyświetlenie oraz odtworzenie w jakości MID zawartości muzycznej z animacją przebiegu odtwarzanych nut;
- przekazanie zakończonej w WLT transkrypcji do archiwizacji w infrastrukturze macierzowej z uwzględnieniem systemu IZZ.

Przebieg procesu:

- a) Użytkownik inicjuje nową transkrypcję rękopisu w WLT za pośrednictwem systemu ADN.
- b) ADN automatycznie odszukuje rekord w IZZ, aby stworzyć komunikację na potrzeby późniejszej archiwizacji, uruchamia nową transkrypcję w WLT na podstawie IIIF manifest.
- c) Po zakończeniu prac transkrypcyjnych i edytorskich w WLT (100% OCR) właściciel transkrypcji oznacza status np. „Publikacja”, który jest wyrażeniem zgody na rozpoczęcie procedury archiwizacji i publikacji.
- d) ADN rozpoczyna procedurę archiwizacji nowej zawartości treściowej w systemie IZZ oraz publikacji wyników do RUWr, BCUWr lub MCUWr. Zostanie dodany obiekt wieloformatowy, gdzie pierwszym formatem będzie manifest IIIF (widok oryginału), który łączy wprowadzone do projektu w WLT linki, zaś drugim sama transkrypcja w postaci pliku MusicXML. Efektem prezentacji w RUWr, BCUWr lub MCUWr będzie prezentacja wiążąca materiał źródłowy (reprodukcję cyfrową) i wykonaną na tej podstawie transkrypcję.





4.1.3.3 Adnotacje na obiektach graficznych

Warunki startowe:

- dostępna jest cyfrowa kopia obiektu oryginalnego hostowanego na infrastrukturze UWr lub innej zewnętrznej instytucji kultury w standardzie IIIF Presentation API;
- użytkownik posiada konto użytkownika w systemie Workflow.

Cel:

- publikacja opracowanych adnotacji obiektu w RUWr, BCUWr lub MCUWr jako część IIIF Presentation API w części adnotacji zgodne z <https://iiif.io/api/presentation/3.0/#55-annotation-page>;
- przekazanie zakończonej w WLT adnotacji do archiwizacji w infrastrukturze macierzowej z uwzględnieniem systemu IZZ.

Przebieg procesu:

- a) Użytkownik inicjuje nową transkrypcję rękopisu w WLT za pośrednictwem systemu ADN.
- b) ADN automatycznie odszukuje rekord w IZZ, aby stworzyć komunikację na potrzeby późniejszej archiwizacji, uruchamia nową transkrypcję w WLT na podstawie IIIF manifest.
- c) Po zakończeniu prac transkrypcyjnych i edytorskich w WLT (100% OCR) właściciel transkrypcji oznacza status np. „Publikacja”, który jest wyrażeniem zgody na rozpoczęcie procedury archiwizacji i publikacji.
- d) ADN rozpoczyna procedurę archiwizacji nowej zawartości treściowej w systemie IZZ w formacie TEI oraz publikacji wyników adnotacji w RUWr, BCUWr lub MCUWr za pomocą IIIF Presentation API w części „canvas”.

4.1.3.4 Obsługa OCR w formacie ALTO

Warunki startowe:

- Dostępna jest cyfrowa kopia obiektu oryginalnego hostowanego na infrastrukturze UWr lub innej zewnętrznej instytucji kultury w standardzie IIIF Presentation API;
- użytkownik posiada konto użytkownika w systemie Workflow.

Cel:

- wykonanie w WLT automatycznego OCR i zapisanie w formacie ALTO;
- możliwość opracowania strukturalnego rozpoznanego tekstu, którego efektem będzie możliwość udostępnienia obrazów i tekstu poszczególnych artykułów jak w





przytoczonym przykładzie Biblioteki Narodowej Walii

<https://newspapers.library.wales/view/4342439/4342445/53/>;

- udostępnienie warstwy tekstowej w RUWr, BCUWr lub MCUWr jako część IIIF Presentation API w części Canvas, <https://iiif.io/api/presentation/3.0/#53-canvas> prezentowane podobnie jak w niniejszym przykładzie - <http://universalviewer.io/examples/#?c=&m=&s=&cv=12&manifest=http%3A%2F%2Fwellcome.library.org%2Fiiif%2Fb18035723%2Fmanifest&xywh=1805%2C2274%2C944%2C709>

Przebieg procesu:

- a) Użytkownik inicjuje nową transkrypcję rękopisu w WLT za pośrednictwem systemu ADN.
- b) ADN automatycznie odszukuje rekord w IZZ, aby stworzyć komunikację na potrzeby późniejszej archiwizacji, uruchamia nową transkrypcję w WLT na podstawie IIIF manifest.
- c) Po zakończeniu prac transkrypcyjnych i edytorskich w WLT (100% OCR) właściciel transkrypcji oznacza status np. „Publikacja”, który jest wyrażeniem zgody na rozpoczęcie procedury archiwizacji i publikacji.
- d) ADN rozpoczyna procedurę archiwizacji nowej zawartości treściowej w systemie IZZ oraz publikacji wyników w RUWr, BCUWr lub MCUWr za pomocą IIIF Presentation API w części „structures”.

4.2 Zadanie nr 2: Rozbudowa Wirtualnego Laboratorium Transkrypcji

Wirtualne Laboratorium Transkrypcji (WLT) <https://translab.uni.wroc.pl> jest obecnie funkcjonującym narzędziem będącym częścią zaplecza usług cyfrowych Uniwersytetu Wrocławskiego. Obecnie realizowane funkcjonalności przedstawione są pod adresem <https://translab.uni.wroc.pl/how-to-transcribe>. W tym zadaniu planowana jest rozbudowa niniejszego narzędzia o nowe funkcjonalności umożliwiające opracowanie treści zdigitalizowanego obiektu cyfrowego w następujących obszarach:

- Transkrypcja rękopisów w standardzie TEI (Text Encoding Initiative),
- Obsługa transkrypcji muzycznych,
- Adnotacje na obiektach graficznych,
- Obsługa OCR w formacie ALTO (Analyzed Layout and Text Object),

Ponadto system WLT zostanie zmodernizowany o następujące funkcjonalności:

1. Szczegółowość rozpoznania tekstu OCR obecnie realizowana jest na poziomie „wiersza”, tzn. koordynaty opisujące pozycję rozpoznanego najmniejszego fragmentu tekstu opisują wiersz w kolumnie. Oczekiwany efektem modernizacji jest rozpoznanie najmniejszego fragmentu tekstu na poziomie „słowa”.
2. Wdrożenie modułu rozpoznawania pisma ręcznego.
3. Rozbudowa o narzędzia umożliwiające trenowanie systemu rozpoznającego na podstawie wybranych transkrypcji. Oczekiwany efektem modernizacji jest





- dedykowany panel, w którym można wskazać dokumenty, na podstawie których będzie trenowany silnik OCR dla różnych typów zasobów oraz rozpoznawania pisma ręcznego. Alternatywnym rozwiązaniem może być uruchomienie kilku silników zorientowanych na określone typy zasobów lub rozpoznawanie pisma ręcznego.
4. Rozbudowa WLT o możliwość crowdsourcingu przez użytkowników zupełnie zewnętrznych dla projektu, którzy zdalnie będą mogli opracowywać materiały źródłowe podczas organizowanych cyklicznie Transcribathonów wzorowanych na rozwiązaniu - <https://transcribathon.eu/>. W ramach rozbudowy powinny zostać uruchomione następujące funkcjonalności umożliwiające wsparcie dla organizacji rozgrywek w transkrypcji materiałów źródłowych:
 - możliwość tworzenia zespołów/drużyn startujących w rozgrywkach;
 - statystyki rozpoznanych dokumentów oraz znaków z podziałem na drużyny oraz użytkowników;
 - moduł instruktarzowy określający zasady organizacji rozgrywek w transkrypcji materiałów źródłowych.
 5. Rozbudowa WLT o API, które pozwoli tworzyć gry lub inne aplikacje typu „recaptcha” (np. dostarczanie przez WLT linii tekstu i odpowiadającej linii obrazka lub dostarczanie zdjęć oznaczonych w dokumentach).
 6. Zmiany funkcjonalne w interfejsie użytkownika:
 - a) Dodanie informacji o trwającej transkrypcji również w widoku „moje transkrypcje” (oprócz obecnej informacji w szczegółach transkrypcji).
 - b) Zmiana trybu działania importu publikacji z dotychczasowego synchronicznego (przeglądarka oczekuje na zwrotny komunikat z serwera o zakończeniu importu) na asynchroniczny (proces importu powinien odbywać się w tle, zaś przeglądarka od razu po jego inicjacji powinna przekierować użytkownika do podstrony transkrypcji prezentującej stan zaawansowania trwającego importu).
 - c) Zmiana algorytmu dodawania plików do nowej transkrypcji w trybie WLT:
 - użytkownik jest zalogowany, transkrypcja założona;
 - użytkownik chce dodać pliki do istniejącej transkrypcji i przejść do okna szczegółów transkrypcji, aby rozpocząć transkrybowanie.

Obecne kroki:

- Wejdź na <https://translab.uni.wroc.pl>
- Kliknij „transkrypcje”
- Kliknij „moje transkrypcje”
- Kliknij „zobacz więcej” przy wybranej transkrypcji
- Kliknij „dodaj nowe strony”
- Kliknij „wybierz plik”
- Wybierz plik do przesłania
- Kliknij „prześlij plik”
- Kliknij „anuluj” (sic!) **lub** kliknij „wstecz” w przeglądarce **lub** kliknij nazwę transkrypcji w okruszkach **lub** kliknij „transkrypcja”-> „moje transkrypcje”-> „zobacz więcej”





Proponowane kroki:

- Wejść na <https://translab.uni.wroc.pl>
- Kliknij „transkrypcje”
- Kliknij „moje transkrypcje”
- Kliknij „zobacz więcej” przy wybranej transkrypcji
- Kliknij „dodaj nowe strony”
- Kliknij „wybierz plik”
- Wybierz plik do przesłania
- Kliknij „prześlij plik”
- Kliknij „zakończ”
- Przycisk „zakończ” powinien przekierowywać do Home/Transkrypcje/Transkrypcja_do_której_dodawano_pliki

- d) Poprawa nawigacji i przemieszczania obszaru w przeglądarce pliku graficznego w szczególności przy dużych zbliżeniach. Narzędzie tyłu „łapka”.
- e) Utworzenie opcji cofnięcia zmian w transkrypcji, w ramach edycji pojedynczej strony. Po zamknięciu edycji danej strony historia zmian może zostać usunięta.
- f) Oznaczanie bloku tekstu na stronie zawierającej przypisy - oddzielenie tych rodzajów treści co powinno skutkować również podziałem tych treści w pobranym/eksportowanym tekście w postaci pliku bez współrzędnych, ale z zachowaniem struktury.
- g) Rozbudowa panelu użytkownika o widoki „moich transkrypcji” oraz „transkrypcje udostępnione dla mnie” (obecnie obie grupy wyświetlane są wspólnie pod nazwą „moje transkrypcje”).
- h) Automatyczne powiadomienia mailowe dla użytkowników, którym nadano uprawnienia w transkrypcji (np. informacja „Użytkownik x przydzielił ci rolę y w transkrypcji z” z załączonym linkiem do transkrypcji).

Nowa transkrypcja może być zainicjowana poprzez:

- System ADN - jest to preferowana metoda dla obiektów zdigitalizowanych i prezentowanych w RUWr, BCUWr, MCUWr lub ACUWr oraz zarchiwizowanych w infrastrukturze UWr.
- WLT - dla obiektów prywatnych, których wyniki transkrypcji nie będą przedmiotem archiwizacji i prezentacji w systemach UWr.

4.2.1 Transkrypcja rękopisów w TEI

Użytkownik w WLT posiada narzędzia do transkrypcji w TEI, które umożliwiają tworzenie następujących znaczników:

Znaczniki określające strukturę danego tekstu:

- **<div>** – podział tekstu na różne części; występuje z atrybutem **@type**, któremu można nadawać następujące wartości:
- **prologue** – prolog, wstęp;





- **part** – część, księga itp.;
- **chapter** – rozdział;
- **section** – podział wewnątrz rozdziału; nie jest to aż tak często potrzebne, ale może lepiej od razu dodać;
- **<head>** – umieszcza się w nim tytuły wszelkich elementów strukturalnych z możliwością ręcznego dodania numeru rozdziału właśnie do tytułu, np. `<head>Caput primum</head>`;
- **<p>** – wskazuje poszczególne akapity;
- **<pb>** – nowa strona; w tym znaczniku warto wskazywać numer karty strony/zgodnie z rękopisem, np. `<pb n="1r"/>`; to również ułatwiłoby nawigację.

Przykłady użycia:

```
<text>
<body>
<pb n="1r"/>
<div type="prologue">
<head>tytuł</head>
<p>tekst</p>
</div>
<pb n="1v"/>
<div type="part">
<head>tytuł</head>
<div type="chapter">
<head>tytuł</head>
<p>tekst</p>
<p>tekst</p>
</div>
<pb n="2r"/>
<div type="chapter">
<head>tytuł</head>
  <div type="section">
    <head>tytuł</head>
    <p>tekst</p>
  </div>
  <div type="section">
    <head>tytuł</head>
    <p>tekst</p>
  </div>
</div>
</div>
</body>
</text>
```

Oraz znaczniki dodatkowe:

- **** – jeśli skryba lub późniejszy czytelnik chciał usunąć jakąś partię tekstu np. przez skreślenie lub wypunktowanie np. *In principio creavit Diabolus Deus caelum et terram.*
- **<add>** – jeśli jakiś tekst został dopisany; można jeszcze wskazać miejsce dopisania





dodając atrybut **@place** z wartością:

- **above** – dopisek nad linią tekstu
- **margin** – dopisek na marginesie np. *In principio creavit* <add place="margin">Deus</add> *caelum et terram.*
- **<subst>** – jeśli jedna fraza została poprawiona na inną, np.

In principio creavit

<subst>

Diabolus

<add place="above">Deus</add>

</subst>

caelum et terram.

Oraz znacznik **<gap>** - stosowany, gdy słowo jest nieczytelne np. z powodu fizycznego uszkodzenia rękopisu, np.

In principio creavit <gap/> caelum et terram.

Rozbudowa WLT o eksport opracowanych obiektów w formacie tekstowym z zachowaniem struktury tekstu, lecz bez koordynatów. W poniższym linku znajdują się przykłady wyeksportowanych transkrypcji zgodnie z obecnie dostępnymi możliwościami oraz wersje poprawione, które są oczekiwane w wyniku rozbudowy narzędzia:

https://drive.google.com/drive/folders/1EDDS5wJjbeLWIBEY86LcMp-qdw_X8NI?usp=sharing

4.2.2 Obsługa transkrypcji muzycznych

Celem tej funkcjonalności jest stworzenie interfejsu umożliwiającego dołączenie do prezentacji i archiwizacji wykonanej za pomocą zewnętrznych narzędzi transkrypcji nutowej na podstawie zbiorów zawierających treści muzyczne, zdigitalizowanych i prezentowanych w systemach RUWr, BCUWr, MCUWr i ACUWr. WLT oprócz dodania pliku w formacie MusicXML umożliwia powiązanie transkrypcji z graficzną prezentacją obiektu oryginalnego prezentowanego w standardzie IIIF.

- Przy pomocy ADN użytkownik inicjuje w WLT „nową transkrypcję muzyczną” na podstawie linku IIIF manifest (np: https://iip.bu.uni.wroc.pl/iiif/OZM_60471_MUZ_VIOLINO_PRIMO_33598/manifest), jeśli transkrypcja odnosi się do zeskanowanego pojedynczego obiektu muzycznego. W przypadkach gdy transkrypcja odnosi się do kilku zeskanowanych obiektów muzycznych np. książeczki głosowe, WLT umożliwia dodanie dodatkowych linków IIIF manifest.
- Możliwość dodania do utworzonego projektu pliku transkrypcji w formacie MusicXML którą użytkownik będzie mógł powiązać z zeskanowanymi obiektami, korzystając z wewnętrznej przeglądarki w WLT. Przewidywane powiązania to:
 - plik transkrypcji do pierwszej strony z danym utworem obiektu nutowego,
 - plik transkrypcji do pierwszych stron z danym utworem każdej książeczki





- głosowej obiektu nutowego,
 - plik transkrypcji z fragmentem (oznaczenie ramką) strony np. gazety, w której znajduje się krótki utwór muzyczny,
 - do systemu WLT wgrany może być tylko jeden plik z transkrypcją lecz z możliwością połączenia go z różnymi obszarami strumieni IIF prezentujących kopię cyfrową oryginału.
- c) Możliwość podłączenia transkrypcji nutowej do standardowych dokumentów transkrybowanych w WLT. Funkcjonalność będzie umożliwiała oznaczenie fragmentu skanu jako zapisu nutowego z możliwością wgrania pliku transkrypcji.
- d) Po wykonaniu powyższych kroków (a. i b.) użytkownik będzie mógł wgrać te informacje w sposób automatyczny.
- e) Do RUWr, BCUWr lub MCUWr za pośrednictwem ADN zostanie dodany obiekt wieloformatowy, gdzie pierwszym formatem będzie manifest IIF (widok oryginału), który łączy wprowadzone do projektu w WLT linki, zaś drugim sama transkrypcja w postaci pliku MusicXML. Efektem prezentacji w systemach RUWr, BCUWr lub MCUWr będzie prezentacja wiążąca materiał źródłowy (reprodukcję cyfrową) i wykonaną na tej podstawie transkrypcję. Oczekowaną formą prezentacji pliku Music XML będzie zastosowanie w systemach RUWr, BCUWr oraz MCUWr dedykowanej przeglądarki umożliwiającej poprawne wyświetlanie zawartości muzycznej oraz jej odtworzenie w jakości MID z animacją przebiegu odtwarzanych nut.

4.2.3 Adnotacje na obiektach graficznych

WLT zostanie wyposażone w zestaw narzędzi do tworzenia adnotacji analogicznie do funkcjonalności przeglądarki Mirador Viewer 3, lub posiada zaimplementowaną przeglądarkę Mirador Viewer umożliwiające tworzenie adnotacji na nietekstowych plikach graficznych z użyciem znaczników takich jak punkt, użycie różnych rodzajów obrysów np. kształty: prostokąt, owal czy obrys ręczny.

4.2.4 Obsługa OCR w formacie ALTO

Możliwość opracowania strukturalnego rozpoznanego tekstu zgodnie ze schematem ALTO dostępnym na stronie: <https://www.loc.gov/standards/alto/news.html#4-2-released>. WLT zawiera zestaw narzędzi umożliwiających wskazywanie bloków tekstu stanowiących odrębne części np. artykuły gazet. Oczekiwane są przynajmniej dwie wartości wyróżniające tytuł artykułu oraz blok, bądź kolejne bloki jego tekstu. Rozbudowa WLT o możliwość eksportu wyników OCR zarówno przed jak i po ręcznej korekcie w formacie ALTO.

4.3 Zadanie nr 3. Rozbudowa Repozytorium UWr

W ramach tego zadania zostanie wykonane rozdzielnie do osobnej instancji systemu Repozytorium Uniwersytetu Wrocławskiego (RUWr), które aktualnie jest częścią Biblioteki Cyfrowej UWr (BCUWr). Rozdzielenie to powinno być przeprowadzone w taki sposób, aby





zachowane zostały istniejące identyfikatory i linki do obiektów. Nowa instancja RUWr powinna opierać się o osobny serwer systemu dLibra oraz osobną bazę danych i być powiązana z istniejącą już aplikacją czytelnika (www.repozytorium.uni.wroc.pl).

Zostaną rozbudowane funkcjonalności istniejącego RUWr:

W zakresie filtrowania:

- przyporządkowanie różnych wariantów nazwiska autora pod jeden rekord (załącznik „Nazwiska”-> Katarzyna Guz – Regner),
- zmiana formatu daty na datę roczną zamiast dziennej

Pozostałe:

- dodanie mechanizmu monitorującego publikacje planowane z nadanym numerem DOI. Mechanizm ten powinien umożliwiać tworzenie zestawień publikacji planowanych z nadanym DOI, do których w zadanym czasie (np. miesiąca) nie została dodana treść (stworzone wydanie);
- rozbudowa modułu statystycznego o funkcjonalności:
 - roczny licznik wejść do bazy (ogółem oraz z oddzielnym podziałem na pobrania oraz sesje),
 - ranking najbardziej poczytnych publikacji (co najmniej pierwsze 10 najczęściej czytanych z możliwością określenia liczby pozycji na liście);
- zaimplementowanie historii zmian rekordu (w formie osobnej zakładki w aplikacji redaktora);
- zaimplementowanie panelu zmian dla administratorów merytorycznych portalu;
- rozbudowa mechanizmu statusów obiektów, aby umożliwić wdrożenie procedur kontroli nad obiektami i ich zatwierdzania (do publikacji) bądź wycofywania (do poprawy).
- możliwość tworzenia manifestu IIIF Presentation API <https://iiif.io/api/presentation/3.0/> dla obiektów w formatach: pdf, audio, video i udostępniania linku do niego w aplikacji czytelnika
- dodanie w aplikacji redaktora panelu tłumaczeń dla wybranych atrybutów, umożliwiających wiązanie wartości z zakładek językowych
- dodanie nowego typu pola - „link” składającego się z podpola „etykieta” oraz podpola „link” mogącego zawierać odnośniki do zasobów
- stworzenie pola - „opiekun rekordu”, zawierającego wskazanie na konto (w dLibrze) osoby odpowiedzialnej za zawartość danego rekordu. Domyślnie opiekunem rekordu byłby jego twórca (z możliwością późniejszej zmiany przypisania)
- stworzenie możliwości tworzenia zasad walidacji w formie pliku reguł (np. rodzaj pseudojęzyka z instrukcjami warunkowymi typu: if (attrA_value_id == 'xy') then attrB_value_id IN (qb, bc, cd, ...)

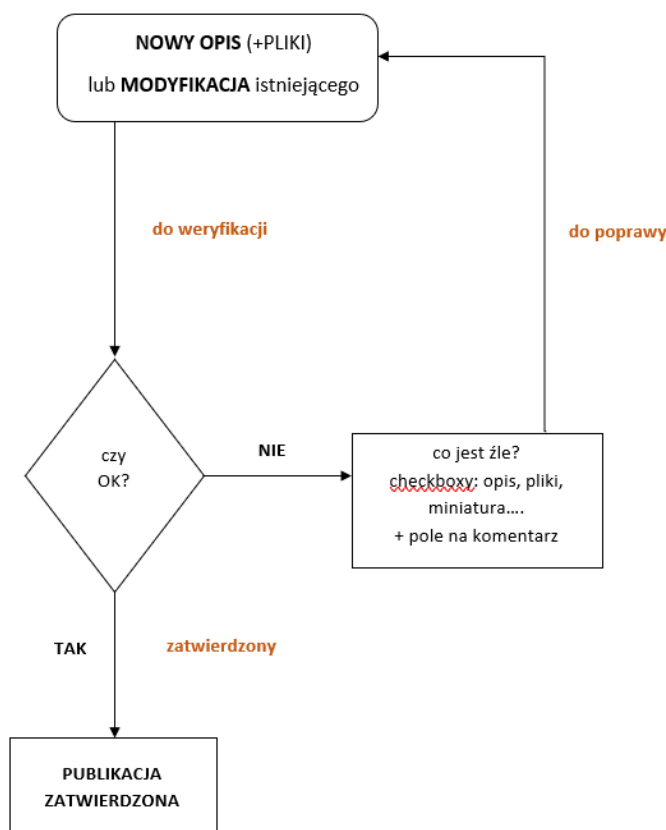




- rozbudowanie konfiguracji szablonów (masek) danych w aplikacji redaktora, tak by możliwe było zdefiniowanie atrybutów wymaganych bez względu na wybrany widok danych (maskę)
- rozszerzenie atrybutu (formularza) typu „Osoba”, tak aby zawierał pola: imię, nazwisko, data urodzenia, data śmierci, dopowiedzenie, ORCID, link oraz możliwe było przypisanie roli/rodzaju współpracy.
- dodanie możliwości przerwania dziedziczenia dla danego obiektu
- obsługa danych badawczych – rozbudowa aplikacji (głównie części czytelnika) o dodatkowe możliwości prezentujące dane badawcze, w szczególności:
 - dedykowany interfejs, uwzględniający prezentację struktury danych (np. umieszczenie danych bezpośrednio pod opisem lub zmiana dla obiektów tego typu zakładki „Struktura” na bardziej intuicyjną)
 - umożliwienie kontaktu z osobą odpowiedzialną w sprawie dostępu do danych chronionych (nieudostępnych publicznie), np. przycisk „Poproś o dostęp”

Wymagania dotyczące implementacji panelu zmian oraz rozbudowy mechanizmu statusów obiektów:





Rys. 2. Schemat zmiany statusów obiektu (publikacji)

Opis procesu przepływu obiektów:

- W Aplikacji zarządzania biblioteką powinien być dodatkowy panel/okienko, gdzie widać rekordy tak, jak w Aplikacji czytelnika (w postaci płaskiej listy - poglądowo poniżej) z tym, że można posortować i wyfiltrować po zadanych polach (min. data, kolekcja, status).
- Każdy nowy lub zmodyfikowany rekord otrzymuje status „do weryfikacji”.
- Redaktor wybiera rekord, sprawdza poprawność obiektu (opis/treść) - jeśli weryfikacja jest pozytywna nadaje status „zatwierdzony”; jeśli nie, wybiera status „do poprawy”.
- W przypadku wyboru statusu „do poprawy” powinno pojawić się okno zawierające listę klikalnych opcji (np. Przycisk typu radio/checkbox itd..) służących do określenia zagadnień wymagających poprawy (np. błędnie wypełnione pole, błędnie podpisany skan, brak miniatury itd.) oraz jedno wspólne dla wszystkich opcji pole tekstowe (textarea) służące do uzupełnienia opisu błędu.
- Po wypełnieniu okno opisu błędów i zatwierdzeniu zmian twórca/ostatni modyfikator





danego rekordu powinien otrzymać powiadomienie (np. email) z informacją o rekordzie do poprawy.

- Po wdrożeniu wymaganych przez korektora zmian edytor ponownie ustawia status „do weryfikacji” i dalej proces przebiega w ten sam sposób.

FILTRY:

wybierz status.. wybierz datę publikacji.. wybierz kolekcję..

zatwierdzony
 do poprawy
 do weryfikacji

El		Status	Identyfikator
			68
<input type="checkbox"/>	100 % Arabica : chiNOISEry	<input type="radio"/>	79961
<input type="checkbox"/>	1600 postaci literackich	<input type="radio"/>	41084
<input type="checkbox"/>	Analyzing quantitative data : an introduction for social researchers	<input checked="" type="radio"/>	81793
<input type="checkbox"/>	Anatomia człowieka	<input type="radio"/>	27718
<input type="checkbox"/>	Andrzej Zdanowicz : dystans : monografia subiektywna	<input checked="" type="radio"/>	92814
<input type="checkbox"/>	Antropologia kultury : zagadnienia i wybór tekstów	<input type="radio"/>	40812
<input type="checkbox"/>	Applied international economics	<input type="radio"/>	81794
<input type="checkbox"/>	Astronomia ogólna	<input checked="" type="radio"/>	27934
<input type="checkbox"/>	Astronomical photometry : past, present, and future	<input type="radio"/>	81883
<input type="checkbox"/>	Astrophysics. Pt. A,Optical and infrared	<input type="radio"/>	81541
<input type="checkbox"/>	Automating Linux and Unix System Administration	<input type="radio"/>	81594
<input checked="" type="checkbox"/>	Barchańska, Maria (1955-), 2016, Parafia św. Teresy Benedykty	<input type="radio"/>	84269
<input type="checkbox"/>	Basiu, zatańczmy... Barbara Bittnerówna, Witold Gruca	<input type="radio"/>	118829
<input checked="" type="checkbox"/>	Biochemia	<input type="radio"/>	28110
<input type="checkbox"/>	Biochemia Harpera ilustrowana	<input type="radio"/>	27939
<input type="checkbox"/>	Biotechnological approaches to barley improvement	<input type="radio"/>	81565
<input type="checkbox"/>	Boże igrzysko : historia Polski	<input type="radio"/>	40653
<input type="checkbox"/>	Być, albo nie być : (Mój głos w sprawie żydowskiej)	<input type="radio"/>	84632
<input type="checkbox"/>	Chemia analityczna	<input type="radio"/>	28101
<input type="checkbox"/>	Chemia nieorganiczna : podstawy	<input type="radio"/>	38614
<input type="checkbox"/>	Chemia organiczna. T. 1-2	<input type="radio"/>	38801
<input type="checkbox"/>	Chemistry of heterocycles : structure, reactions, synthesis and applications	<input type="radio"/>	82149
<input type="checkbox"/>	Chrestomatia staropolska : teksty do roku 1543	<input type="radio"/>	40658
<input type="checkbox"/>	Cisza : lektura krytyczna	<input type="radio"/>	82834
<input type="checkbox"/>	Classics in spectroscopy : isolation and structure elucidation of natural products	<input type="radio"/>	82202
<input type="checkbox"/>	Commercial plant-produced recombinant protein products : case studies	<input type="radio"/>	81787
<input type="checkbox"/>	Community of Europe : a history of European integration since 1945	<input type="radio"/>	81791
<input type="checkbox"/>	Contemporary innovative and information technologies of social development : educational and legal aspects	<input type="radio"/>	103963
<input type="checkbox"/>	Convergence of food security, energy security and sustainable agriculture	<input type="radio"/>	81788

Rys.3. Panel kontroli przepływu obiektów (publikacji)

5 Wymagania pozafunkcjonalne

5.1 Spełnienie standardów dostępności WCAG.





Graficzny interfejs (GUI) części prezentacyjnej będzie spełniać wymagania **WCAG 2.1** dla systemów teleinformatycznych w zakresie dostępności dla osób niepełnosprawnych zgodnie z załącznikiem nr 4 Rozporządzenia Rady Ministrów z dnia 12 kwietnia 2012 r. w sprawie Krajowych Ram Interoperacyjności, minimalnych wymagań dla rejestrów publicznych i wymiany informacji w postaci elektronicznej oraz minimalnych wymagań dla systemów teleinformatycznych.

Ponadto oprogramowanie będzie wyposażone w funkcjonalności dla osób niepełnosprawnych **wykraczające poza standard WCAG 2.1** zgodny z załącznikiem nr 4 Krajowych Ram Interoperacyjności. Poniżej zawarte zostały zalecenia dotyczące standardu AAA, które muszą zostać uwzględnione w procesie projektowania interfejsu użytkownika:

- a. Funkcjonalność interfejsów graficznych będzie ograniczona limitem czasowym o długości 24 h. Dzięki temu użytkownicy będą posiadać tyle czasu, ile potrzebują na wykonanie określonego zadania na stronie internetowej.
- b. Serwis nie będzie posiadał żadnych elementów, które migają z większą częstotliwością niż trzy razy na sekundę.
- c. Strona internetowa zostanie wyposażona w mechanizm wskazania miejsca aktualnej strony w strukturze serwisu za pomocą ścieżki okruszków (breadcrumbs) lub określając bieżący krok w sekwencji.

System będzie wyświetlać powiadomienie o używaniu i zapisywaniu w przeglądarce użytkownika plików „cookies” wraz ze zgodą na ich użycie. Zakres informacyjny tego powiadomienia musi być zgodny z obowiązującymi przepisami. Oprogramowanie zostanie wykonane z zastosowaniem najlepszych praktyk w dziedzinie budowania witryn WWW i w zgodności z najnowszymi standardami, wyznaczonymi przez W3C. Wszystkie strony części prezentacyjnej muszą być co najmniej zgodne ze standardem HTML 5 i CSS 3. Wymagana jest prawidłowa walidacja tworzonego kodu HTML i CSS za pomocą udostępnionego na stronach W3C walidatora (<http://validator.w3.org>).

Publicznie dostępna część będzie posiadać interfejs użytkownika w językach:

- a. polskim, wraz z obsługą polskich liter diakrytycznych i sortowania wg polskiego alfabetu
- b. angielskim

Elementy tekstowe muszą być kodowane zgodnie ze standardem Unicode UTF-8 wersja 3.0. Grafika i kolorystyka części prezentacyjnej będą zgodne z Systemem Identyfikacji Wizualnej wprowadzonym zarządzeniem JM rektora UW r nr 129/2015 z dnia 29 grudnia 2015 lub jego znowelizowana wersją.

5.2 Spełnienie standardów otwartości.

Zasoby nauki udostępniane cyfrowo poprzez oprogramowanie realizowane w projekcie muszą być dostępne zgodnie z otwartym standardem pozwalającym na ich maszynowe przetwarzanie, odpowiadającym poziomowi 5 gwiazdek w skali opisanej w standardzie „5 Star



**Open Data”:**

- a. dane muszą być opublikowane w ustrukturyzowanej postaci z wykorzystaniem otwartego formatu danych: **CSV (***)**,
- b. dane muszą być opublikowane wraz z odnośnikami do poszczególnych elementów zbioru (np. za pomocą metody opisu RDF), co pozwoli na bezpośrednie linkowanie do nich w sieci,
- c. dane muszą być opublikowane wraz z odnośnikami do innych zbiorów, stanowiących dla tych informacji kontekst (tzw. Linked Open Data). Wymienione powyżej funkcjonalności zapewnią spełnienie standardu „**5 Star Open Data**” na poziomie 5 gwiazdek.

5.3 Wymagania dotyczące dokumentacji API.

Zostanie przygotowana dokumentacja interfejsu dla programistów (API) oraz szczegółowa specyfikacja interfejsów API, w tym opisy struktur danych, dostępy do interfejsu, przypadki użycia, przykładowe pliki interfejsu i inne, niewymienione lecz konieczne elementy umożliwiające integrację udostępnianych usług z innymi systemami podmiotów zewnętrznych. Szczegółowość tego opracowania musi być wystarczająca do sprawnego i samodzielnego przyłączenia do interfejsu API systemu zewnętrznego.

5.4 Spełnienie wymagań interoperacyjności

Zaprojektowane i wdrożone oprogramowanie musi spełniać wymagania dotyczące interoperacyjności wynikające z Rozporządzenia Rady Ministrów z dnia 12 kwietnia 2012 r. w sprawie Krajowych Ram Interoperacyjności, minimalnych wymagań dla rejestrów publicznych i wymiany informacji w postaci elektronicznej oraz minimalnych wymagań dla systemów teleinformatycznych.

W celu spełnienia wymagań interoperacyjności na poziomie organizacyjnym przygotowane zostaną informacje, umożliwiające skuteczne zapoznanie innych podmiotów ze sposobem dostępu do serwisów i usług oraz zakresem użytkowym uruchomionego serwisu.

W celu osiągnięcia interoperacyjności na poziomie semantycznym w pracach programistycznych zostaną zastosowane struktury danych i znaczenia danych zawartych w tych strukturach zgodnie z Krajowymi Ramami Interoperacyjności oraz danych publikowanych w repozytorium interoperacyjności.

W celu osiągnięcia interoperacyjności na poziomie technologicznym stosowane będą minimalne wymagania dla systemów teleinformatycznych określonych w rozdziale IV





Rozporządzenia w sprawie Krajowych Ram Interoperacyjności, minimalnych wymagań dla rejestrów publicznych i wymiany informacji w postaci elektronicznej oraz minimalnych wymagań dla systemów teleinformatycznych.

6 Platforma sprzętowa

Uniwersytet Wrocławski posiada zakupioną wcześniej platformę sprzętowo-programową, na której zostaną uruchomione usługi cyfrowe realizowane w ramach projektu. Platforma sprzętowa składa się z siedmiu serwerów fizycznych (2 procesory, 512 GB RAM) oraz macierzy dyskowej (o pojemności 16 TB), które tworzą klaster zarządzany przez oprogramowanie VMWare. Archiwizacja odbywa się za pomocą pamięci taśmowej

7 Wymagania dotyczące realizacji projektu

Cały projekt począwszy od fazy badania i dokumentowania potrzeb biznesowych, poprzez budowę aż do wdrożenia będzie prowadzony w oparciu o metody zorientowane na użytkownika ze szczególnym uwzględnieniem następujących aspektów:

- a. dostępność i użyteczność funkcjonalności i graficznych interfejsów dla wszystkich grup użytkowników,
- b. możliwość korzystania przez użytkowników z zasobów nauki niezależnie od miejsca przebywania i stosowanej technologii,
- c. umożliwienie monitorowania pracy pod kątem: dostępności i użyteczności graficznych interfejsów dla wszystkich grup użytkowników, ciągłości działania, powszechności wykorzystania oraz satysfakcji użytkowników.

W celu zapewnienia wygodnego dostępu do zasobów niezależnie od używanej technologii oprogramowanie musi być przystosowane do pracy za pomocą popularnych przeglądarek Internetowych. Wszystkie ekrany, które służą udostępnianiu zasobów nauki zostaną dostosowane do standardu RWD. Zgodnie z zasadami RWD, wygląd graficzny tych ekranów powinien zmieniać swój stan po przekroczeniu pewnych punktów granicznych (breakpoints) w szerokości obszaru okna przeglądarki, w którym wyświetlana jest treść. Proponowane wartości dla punktów granicznych wynoszą:

- Desktop: 1024px (szerokość wyjściowa)
- Tablet: 768px
- Mobile: 320px

