

OPIS ZAŁOŻEŃ PROJEKTU INFORMATYCZNEGO

Tytuł projektu	WUM Data Hub - Centrum Doskonałości Danych WUM – cyfrowe udostępnienie i ponowne wykorzystanie zasobów naukowych Warszawskiego Uniwersytetu Medycznego		
Wnioskodawca	Minister Zdrowia		
Beneficjent	Warszawski Uniwersytet Medyczny		
Partnerzy	brak		
Źródło finansowania	Źródło finansowania Środki UE – Działanie FERC.02.03 Cyfrowa dostępność i ponowne wykorzystanie informacji. Budżet państwa - część budżetowa 27		
Całkowity koszt projektu	32 000 000,00 zł		
Planowany okres realizacji projektu	12-2026 do 05-2029		
Osoba kontaktowa	Tadeusz Turzyński	tadeusz.turzynski@wum.edu.pl	602640003

1. POWODY PODJĘCIA PROJEKTU

1.1. Identyfikacja problemu i potrzeb

Warszawski Uniwersytet Medyczny dysponuje znacznymi zasobami danych naukowych, klinicznych, publikacyjnych, dydaktycznych i obrazowych. Zasoby te pozostają jednak rozproszone pomiędzy wieloma systemami, nie są w pełni interoperacyjne i nie funkcjonują w ramach jednolitego modelu katalogowania, anonimizacji, standaryzacji oraz udostępniania. Zidentyfikowany problem dotyczy braku jednolitego środowiska, które pozwalałoby na pozyskiwanie danych z wcześniej rozpoznanych źródeł, ich bezpieczne przyjmowanie, porządkowanie, anonimizację, standaryzację i dalsze przetwarzanie zgodnie z wymaganiami interoperacyjności, bezpieczeństwa i ponownego wykorzystania informacji. Obecny stan utrudnia zarówno budowę zasobu danych do celów badawczych i dydaktycznych, jak i wydzielanie danych odpowiednio przygotowanych do bardziej zaawansowanych zastosowań analitycznych i klinicznych. Utrudnia również przygotowanie danych do udostępniania w ujednoliconym modelu jakościowym i interoperacyjnym, odpowiadającym wymaganiom otwartych danych i ponownego wykorzystania informacji.

Z tego względu zidentyfikowana potrzeba obejmuje utworzenie spójnego środowiska przetwarzania danych, składającego się z warstwy gromadzenia danych z różnych źródeł (Data Lake), warstwy ich harmonizacji i analitycznego uporządkowania (hurtownia danych) oraz wydzielonej bazy danych zweryfikowanych jakościowo, przeznaczonej dla zastosowań wymagających podwyższonej wiarygodności, kontroli dostępu i możliwości dalszego wykorzystania klinicznego. W modelu tym warstwa Data Lake będzie pełniła funkcję przyjęcia i przechowywania danych o zróżnicowanym stopniu uporządkowania, kompletności i gotowości do udostępniania, natomiast dane przetworzone, zharmonizowane i udostępniane z hurtowni danych będą przygotowywane z założeniem osiągnięcia standardu interoperacyjności odpowiadającego modelowi 5-Star Open Data, tam gdzie pozwala na to charakter zasobu, model prawny i wymagania bezpieczeństwa.

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

Interesariusz	Zidentyfikowany problem	Szacowana wielkość grupy
Naukowcy i badacze WUM	Brak dostępu do uporządkowanych i możliwych do ponownego wykorzystania danych naukowych i medycznych pochodzących z obszaru działalności WUM, w tym możliwości ich zestawiania z własnymi danymi badawczymi oraz wykorzystania w analizach naukowych i projektach rozwojowych.	ok. 7 000 osób
Studenci i doktoranci WUM	Brak uporządkowanego dostępu do zasobów naukowych, dydaktycznych i zanonimizowanych danych, które mogłyby być wykorzystywane w kształceniu, projektach studenckich, pracach dyplomowych i doktoratach.	ok. 17 000 osób
Ministerstwo Zdrowia	Ograniczony dostęp do zagregowanych, zanonimizowanych i uporządkowanych informacji dotyczących danych medycznych i badawczych z obszaru działalności WUM, które mogłyby być zestawiane z danymi pochodzącymi z innych instytucji i systemów.	1
Warszawski Uniwersytet Medyczny	Brak spójnego, kontrolowanego i interoperacyjnego dostępu do danych oraz metadanych możliwych do wykorzystania przez narzędzia analityczne, systemy AI/ML (Artificial Intelligence / Machine Learning) i odbiorców maszynowych w procesach badawczych i rozwojowych.	1
Komisja Europejska – DG SANTE	Brak interoperacyjnego dostępu do uporządkowanych metadanych i uzgodnionych zasobów danych zdrowotnych oraz naukowych z obszaru działalności WUM, które mogłyby wspierać realizację założeń EHDS (European Health Data Space – Europejska Przestrzeń Danych Zdrowotnych) w zakresie wtórnego wykorzystania danych, w szczególności poprzez możliwość wyszukiwania, oceny przydatności i ponownego wykorzystania zasobów w analizach, politykach publicznych, badaniach naukowych oraz europejskich inicjatywach dotyczących danych zdrowotnych.	1
ECRIN-ERIC (European Clinical Research Infrastructure Network – European Research Infrastructure Consortium) –	Brak interoperacyjnego dostępu do uporządkowanych metadanych i uzgodnionych zasobów badawczych oraz klinicznych z obszaru działalności WUM, które mogłyby być wykorzystywane w międzynarodowej współpracy naukowej, w projektach badań klinicznych, badaniach translacyjnych oraz w europejskich sieciach	1

Interesariusz	Zidentyfikowany problem	Szacowana wielkość grupy
europejska niekomercyjna organizacja zrzeszająca krajowe sieci infrastruktury badań klinicznych państw członkowskich UE oraz krajów stowarzyszonych, wspierająca prowadzenie wieloośrodkowych, międzynarodowych badań klinicznych w Europie	infrastruktury badawczej, z zachowaniem wymagań dotyczących jakości danych, standaryzacji, identyfikowalności, dostępności i ponownego wykorzystania.	
Lekarze i klinicyści	Brak dostępu do uporządkowanych, zweryfikowanych i zanonimizowanych danych pochodzących z działalności WUM, które mogłyby wspierać analizę kliniczną, dydaktykę medyczną oraz praktykę opartą na danych i dowodach.	ok. 2 500 osób
Instytucje publiczne – w szczególności GUS, NFZ, AOTMiT.	Ograniczony dostęp do zagregowanych, zanonimizowanych i uporządkowanych informacji dotyczących danych medycznych i badawczych z obszaru działalności WUM, które mogłyby być zestawiane z danymi pochodzącymi z innych instytucji i systemów	ok. 10 podmiotów

1.2. Opis stanu obecnego

Warszawski Uniwersytet Medyczny (WUM), z ponad 200-letnią tradycją kształcenia medycznego, jest jednym z wiodących ośrodków naukowo-badawczych w Polsce w obszarze nauk medycznych, nauk o zdrowiu i nauk farmaceutycznych. Uczelnia prowadzi liczne projekty badawcze związane ze zwiększaniem wykorzystania danych zdrowotnych w działalności naukowej oraz z zastosowaniem rozwiązań bioinformatycznych i analitycznych w ochronie zdrowia. Działania te wspierają rozwój nowoczesnych metod diagnostycznych, terapeutycznych i organizacyjnych, ukierunkowanych na poprawę jakości opieki zdrowotnej.

WUM dysponuje istotnym potencjałem wyjściowym do realizacji projektu, obejmującym zasoby danych naukowych, klinicznych, publikacyjnych i dydaktycznych, a także doświadczenia w obszarze digitalizacji, anonimizacji, wyszukiwania informacji, udostępniania danych oraz budowy interfejsów integracyjnych. Punkt wyjścia projektu stanowią również wcześniej zidentyfikowane systemy źródłowe i repozytoryjne, które mogą zostać objęte wspólnym modelem integracji, porządkowania i dalszego wykorzystania danych.

W sektorze ochrony zdrowia funkcjonują rozwiązania zbliżone zakresem do wybranych elementów planowanego systemu, jednak mają one najczęściej charakter ograniczony do pojedynczych instytucji, wybranych zasobów lub określonych procesów. Zazwyczaj nie obejmują one jednocześnie szerokiej integracji źródeł danych, zaawansowanej anonimizacji, standaryzacji, agregacji, wielowarstwowego udostępniania oraz wykorzystania narzędzi analitycznych,

Business Intelligence i AI w jednym spójnym środowisku.

System CDD WUM nie jest projektowany jako rozwiązanie skalowalne do wdrożenia w innych uczelniach medycznych. Jest dedykowanym środowiskiem dla zasobów naukowych, klinicznych, publikacyjnych i dydaktycznych Warszawskiego Uniwersytetu Medycznego, ściśle zintegrowanym z konkretnymi systemami źródłowymi WUM oraz dostosowanym do specyfiki procesów badawczych i organizacyjnych uczelni.

2. EFEKTY PROJEKTU

2.1. Cele i korzyści wynikające z projektu

Cel - 1	Utworzenie w Warszawskim Uniwersytecie Medycznym spójnego systemu IT dla Centrum Doskonałości Danych, umożliwiającego pozyskiwanie, porządkowanie, anonimizację, standaryzację, archiwizację i udostępnianie danych naukowych, klinicznych, publikacyjnych oraz dydaktycznych w sposób wspierający ich bezpieczne i ponowne wykorzystanie przez odbiorców wewnętrznych i zewnętrznych.
Cel strategiczny	1. Program otwierania danych na lata 2021–2027 – w szczególności w zakresie celu pierwszego (zwiększenie dostępności danych w portalu Dane.gov.pl, w tym danych badawczych), udostępniania danych przez API oraz podnoszenia jakości danych. Projekt realizuje te kierunki w odniesieniu do danych naukowych, klinicznych, publikacyjnych i dydaktycznych WUM. 2. „Zdrowa Przyszłość. Ramy strategiczne rozwoju systemu ochrony zdrowia na lata 2021–2027 z perspektywą do 2030 r.” – w szczególności w zakresie cyfryzacji ochrony zdrowia, wykorzystania danych do wsparcia decyzji w nauce i klinice oraz interoperacyjności w środowisku zgodnym z założeniami FAIR (Findable, Accessible, Interoperable, Reusable) i Europejskiej Przestrzeni Danych Zdrowotnych (EHDS). 3. Polityka dla rozwoju Sztucznej Inteligencji w Polsce od roku 2020 – w szczególności w obszarach „AI w nauce” oraz wykorzystania AI w sektorze publicznym i ochronie zdrowia. Projekt buduje warstwę danych umożliwiającą zastosowania AI/BI w środowisku zgodnym z wymogami prywatności i bezpieczeństwa. 4. Strategia Rozwoju Warszawskiego Uniwersytetu Medycznego oraz Statut WUM – projekt wpisuje się w cele uczelni dotyczące doskonałości naukowej, jakości kształcenia i rozwoju infrastruktury badawczej.
Korzyść:	Najważniejsze korzyści wynikające z realizacji projektu obejmują: 1. Uporządkowanie i zintegrowanie procesów przetwarzania danych w obszarze nauki, dydaktyki i ochrony zdrowia, tak aby możliwe było ich przekształcanie w informacje i wiedzę użyteczną dla badań, analiz i decyzji. 2. Usystematyzowanie metod technicznych oraz zasad dostępu do zasobów informacyjnych WUM w ramach utworzonego Centrum Doskonałości Danych, obejmującego dane naukowe, kliniczne, publikacyjne i dydaktyczne. 3. Zwiększenie możliwości ponownego wykorzystania danych przez badaczy, studentów, doktorantów, klinicystów oraz partnerów instytucjonalnych dzięki ich standaryzacji, anonimizacji i udostępnianiu w uporządkowanym modelu. 4. Podniesienie kompetencji pracowników Beneficjenta w zakresie zaawansowanego przetwarzania, analizy i bezpiecznego wykorzystania danych zgodnie z wymaganiami prawnymi i organizacyjnymi.

	5. Wzmocnienie interoperacyjności i gotowości do współpracy krajowej oraz międzynarodowej, w tym w zakresie wymiany metadanych i dalszego rozwoju środowiska zgodnego z kierunkami FAIR, DCAT-AP PL (Data Catalogue Application Profile for Poland) i EHDS.
KPI:	KPI 1. Liczba podmiotów wspartych w zakresie rozwoju usług, produktów i procesów cyfrowych KPI 2. Liczba podmiotów, które udostępniły informacje sektora publicznego / dane prywatne on-line KPI 3. Użytkownicy nowych i zmodernizowanych publicznych usług, produktów i procesów cyfrowych KPI 4. Liczba udostępnionych on-line dokumentów zawierających informacje sektora publicznego / dane prywatne KPI 5. Liczba zdigitalizowanych dokumentów zawierających informacje sektora publicznego / dane prywatne KPI 6. Rozmiar udostępnionych on-line informacji sektora publicznego / danych prywatnych KPI 7. Rozmiar zdigitalizowanych informacji sektora publicznego / danych prywatnych KPI 8. Liczba utworzonych API KPI 9. Liczba pracowników IT objętych wsparciem szkoleniowym KPI 10. Liczba pracowników nie będących pracownikami IT, objętych wsparciem szkoleniowym KPI 11. Liczba odsłon zasobów Systemu CDD w okresie referencyjnym (wskaźnik rezultatu – monitorowanie popularności) KPI 12. Liczba nowych i zmodernizowanych systemów teleinformatycznych (wskaźniki zgodnie z FERC 2.3) KPI 1: Wartość aktualna – 0; wartość docelowa – 1. KPI 2: Wartość aktualna – 0; wartość docelowa – 1. KPI 3: Wartość aktualna – 0; wartość docelowa – 1 500. KPI 4: Wartość aktualna – 0; wartość docelowa – 80 000. KPI 5: Wartość aktualna – 0; wartość docelowa – 10 000. KPI 6: Wartość aktualna – 0; wartość docelowa – 5 TB. KPI 7: Wartość aktualna – 0; wartość docelowa – 30 TB. KPI 8: Wartość aktualna – 0; wartość docelowa – 3. KPI 9: Wartość aktualna – 0; wartość docelowa – 10 KPI 10: Wartość aktualna – 0; wartość docelowa – 80 KPI 11: Wartość aktualna – 0; wartość docelowa – 50 000 odsłon rocznie KPI 12: Wartość aktualna – 0; wartość docelowa – 2 (System CDD – nowy; openCARDIO – zmodernizowany w zakresie interfejsów wymiany danych)
Wartość aktualna i docelowa KPI:	0 1
Metoda pomiaru KPI	KPI 1. Liczba podmiotów wspartych w zakresie rozwoju usług, produktów i procesów cyfrowych Metoda oraz sposób pomiaru: pomiar na podstawie dokumentacji projektowej potwierdzającej objęcie wsparciem podmiotu (Beneficjenta) w zakresie rozwoju cyfrowych usług, produktów i procesów uruchomionych w ramach Systemu CDD WUM. Źródło: dokumentacja projektowa (umowa o dofinansowanie, protokoły odbioru, sprawozdania końcowe). Częstotliwość: pomiar jednorazowy na zakończenie rzeczowej realizacji projektu. KPI 2. Liczba podmiotów, które udostępniły informacje sektora publicznego /

dane prywatne on-line Metoda oraz sposób pomiaru: pomiar na podstawie potwierdzenia uruchomienia produkcyjnego komponentów Systemu CDD WUM odpowiedzialnych za on-line'owe udostępnianie informacji sektora publicznego oraz danych prywatnych objętych projektem. Źródło: dokumentacja techniczna i odbiorowa Systemu CDD WUM, rejestr usług udostępniania danych. Częstotliwość: pomiar jednorazowy na zakończenie rzeczowej realizacji projektu.

KPI 3. Użytkownicy nowych i zmodernizowanych publicznych usług, produktów i procesów cyfrowych Metoda oraz sposób pomiaru: pomiar liczby unikalnych użytkowników korzystających z nowych lub zmodernizowanych usług, produktów i procesów cyfrowych uruchomionych w ramach Systemu CDD WUM, w okresie referencyjnym. Źródło: system analityczny portalu CDD WUM, rejestr kont i sesji użytkowników, logi systemowe. Częstotliwość: pomiar jednorazowy po roku od zakończenia rzeczowej realizacji projektu.

KPI 4. Liczba udostępnionych on-line dokumentów zawierających informacje sektora publicznego / dane prywatne Metoda oraz sposób pomiaru: pomiar liczby rekordów, dokumentów i obiektów cyfrowych udostępnionych on-line w repozytoriach i zbiorach Systemu CDD WUM, kwalifikujących się jako informacje sektora publicznego lub dane prywatne objęte zakresem projektu. Źródło: rejestr repozytoriów, katalog metadanych i rejestr publikowanych zbiorów Systemu CDD WUM. Częstotliwość: pomiar na zakończenie rzeczowej realizacji projektu.

KPI 5. Liczba zdigitalizowanych dokumentów zawierających informacje sektora publicznego / dane prywatne Metoda oraz sposób pomiaru: pomiar liczby dokumentów i obiektów przetworzonych do postaci cyfrowej w ramach projektu, w tym poddanych skanowaniu, OCR (Optical Character Recognition), ekstrakcji metadanych i indeksacji. Wskaźnik obejmuje wszystkie zdigitalizowane dokumenty, niezależnie od tego, czy w okresie realizacji projektu zostały udostępnione on-line. Źródło: rejestr digitalizacji oraz rejestr zasobów Systemu CDD WUM. Częstotliwość: pomiar na zakończenie rzeczowej realizacji projektu.

KPI 6. Rozmiar udostępnionych on-line informacji sektora publicznego / danych prywatnych Metoda oraz sposób pomiaru: pomiar łącznego wolumenu (w TB) informacji sektora publicznego i danych prywatnych udostępnionych on-line za pośrednictwem Systemu CDD WUM. Źródło: rejestr repozytoriów, statystyki magazynów danych Systemu CDD WUM. Częstotliwość: pomiar na zakończenie rzeczowej realizacji projektu.

KPI 7. Rozmiar zdigitalizowanych informacji sektora publicznego / danych prywatnych Metoda oraz sposób pomiaru: pomiar łącznego wolumenu (w TB) informacji i zasobów zdigitalizowanych w ramach projektu, niezależnie od ich późniejszego trybu udostępnienia. Źródło: rejestr digitalizacji, statystyki repozytorium źródłowego Systemu CDD WUM. Częstotliwość: pomiar na zakończenie rzeczowej realizacji projektu.

KPI 8. Liczba utworzonych API Metoda oraz sposób pomiaru: pomiar liczby produkcyjnie uruchomionych interfejsów API wdrożonych w ramach Systemu CDD WUM, służących udostępnianiu danych i metadanych odbiorcom wewnętrznym i zewnętrznym. Źródło: rejestr usług API Systemu CDD WUM, dokumentacja interfejsów (np. specyfikacje OpenAPI). Częstotliwość: pomiar na zakończenie rzeczowej realizacji projektu.

KPI 9. Liczba pracowników IT objętych wsparciem szkoleniowym Metoda oraz sposób pomiaru: pomiar liczby unikalnych pracowników WUM zatrudnionych na stanowiskach IT (administratorzy systemów, programiści, inżynierowie danych, specjaliści ds. bezpieczeństwa informacji, architekci rozwiązań),

	którzy ukończyli przynajmniej jedno szkolenie zrealizowane w ramach projektu. Źródło: dokumentacja szkoleniowa projektu – listy obecności, rejestr uczestników, wydane zaświadczenia / certyfikaty, dane HR pozwalające na klasyfikację stanowiska. Częstotliwość: pomiar na zakończenie rzeczowej realizacji projektu. KPI 10. Liczba pracowników nie będących pracownikami IT, objętych wsparciem szkoleniowym Metoda oraz sposób pomiaru: pomiar liczby unikalnych pracowników WUM nieposiadających stanowisk IT (m.in. badacze, klinicyści, dydaktycy, bibliotekarze, data stewardzi, pracownicy administracji), którzy ukończyli przynajmniej jedno szkolenie zrealizowane w ramach projektu (np. z zakresu pracy z danymi, anonimizacji, FAIR, korzystania z funkcji Systemu CDD WUM). Źródło: dokumentacja szkoleniowa projektu – listy obecności, rejestr uczestników, wydane zaświadczenia / certyfikaty, dane HR pozwalające na klasyfikację stanowiska. Częstotliwość: pomiar na zakończenie rzeczowej realizacji projektu. KPI 11. Liczba odsłon zasobów Systemu CDD w okresie referencyjnym Metoda oraz sposób pomiaru: pomiar łącznej liczby odsłon stron katalogu zasobów, kart datasetów, kart publikacji i innych zasobów udostępnianych przez Portal CDD WUM, rejestrowanych przez wbudowany system analityczny portalu. Wskaźnik rezultatu pozwalający na monitorowanie popularności platformy udostępniania. Źródło: system analityczny Portalu CDD WUM, logi serwera WWW, dashboardy monitoringu. Częstotliwość: pomiar w okresach 12-miesięcznych po zakończeniu rzeczowej realizacji projektu, do końca okresu trwałości. KPI 12. Liczba nowych i zmodernizowanych systemów teleinformatycznych Metoda oraz sposób pomiaru: pomiar liczby produkcyjnie uruchomionych nowych systemów teleinformatycznych oraz systemów zmodernizowanych w ramach projektu. Wartość docelowa obejmuje System CDD (nowy) oraz System openCARDIO (zmodernizowany w zakresie interfejsów wymiany danych z Systemem CDD). Źródło: protokoły odbioru komponentów, dokumentacja techniczna i odbiorowa Systemu CDD oraz Systemu openCARDIO. Częstotliwość: pomiar jednorazowy na zakończenie rzeczowej realizacji 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)

Rodzaj informacji/zasobów	Planowana data udostępnienia	Szacowana liczba obiektów objętych digitalizacją (udostępnianiem informacji)
Zdigitalizowane i udostępnione zasoby publikacyjne – prace naukowe, artykuły, monografie, rozdziały (pozyskane z PPM WUM) oraz prace doktorskie i habilitacyjne (część zasobu wymaga skanowania i OCR)	31-05-2029	75 785 rekordów (70 785 rekordów w PPM WUM + ok. 5 000 prac doktorskich i habilitacyjnych)
Udostępnione datasety FAIR – dane z projektów NCN / NCBR / ABM / EU Horizon, dane projektu openCardio (zanonimizowane zasoby kardiologiczne), dane eksperymentalne z laboratoriów WUM	31-05-2029	co najmniej 250 datasetów (w tym 213 datasetów już zidentyfikowanych w PPM WUM oraz nowe zbiory przygotowane i opisane w ramach projektu)
Zdigitalizowane i udostępnione zasoby obrazowe (naukowe, non-PHI) – obrazy histologiczne, mikroskopowe, mikro-CT po pełnej anonimizacji (DICOM non-PHI, TIFF, JPEG2000, NIfTI) oraz obrazy kliniczne po anonimizacji	31-05-2029	ok. 3 000 obiektów / serii obrazowych
Zdigitalizowane i udostępnione zasoby dydaktyczne i edukacyjne – wykłady, atlasy anatomiczne, materiały szkoleniowe (PDF, HTML5, SCORM), w tym wybrane materiały konferencyjne; część zasobu wymaga skanowania, część dostępna już w postaci cyfrowej w istniejącym Repozytorium materiałów dydaktycznych WUM. Do digitalizacji zasobów w ramach projektu CDD WUM nie będzie wykorzystywana infrastruktura ani systemy projektu openCARDIO – projekt CDD przewiduje wykorzystanie własnych komponentów skanowania,	31-05-2029	ok. 1 500 obiektów

Rodzaj informacji/zasobów	Planowana data udostępnienia	Szacowana liczba obiektów objętych digitalizacją (udostępnianiem informacji)
OCR i przetwarzania uruchamianych w środowisku Systemu CDD.		
Udostępnione metadane ISP (DCAT-AP PL) – katalog CDD WUM: opisy zbiorów danych do harvestingu przez PPM / dane.gov.pl / OpenAIRE / EOSC	31-05-2029	ok. 81 000 rekordów metadanych

Czy wszystkie zdigitalizowane zasoby objęte projektem będą udostępniane bezpłatnie?
TAK/NIE

2.4. Produkty końcowe projektu

Nazwa produktu	Planowana data wdrożenia
Raport z inicjalnego testu prywatności (DPIA (Data Protection Impact Assessment – ocena skutków dla ochrony danych))	12-2026
Uzupełniona infrastruktura techniczna Systemu CDD – serwery aplikacyjne, serwer bazodanowy PostgreSQL HA, klaster Elasticsearch, storage obiektowy S3 MinIO, NAS archiwalny, sieć 10 Gb/s oraz zasilanie awaryjne	06-2028
Interfejs API OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting) / DCAT-AP PL – harvesting metadanych do PPM / dane.gov.pl / OpenAIRE / EOSC	07-2028
Modyfikacja Systemu OpenCardio – uporządkowanie i konsolidacja zasobów danych, dostosowanie interfejsów wymiany danych z Systemem CDD oraz przygotowanie środowiska eksploatacyjnego do współpracy z platformą CDD	12-2028
Materiały szkoleniowe	12-2028
Raport z testów badań UX	02-2029
Raport z audytu dostępności WCAG (Web Content Accessibility Guidelines) 2.2 AA	03-2029
Raport z testów bezpieczeństwa (testy penetracyjne)	03-2029
Raport z testów wydajności	03-2029
Materiały informacyjno-promocyjne	03-2029
Raport z audytu wewnętrznego potwierdzającego gotowość Systemu CDD do ubiegania się o certyfikację ECRIN Data Centre Certification	03-2029
System CDD	05-2029
Interfejs API REST Systemu CDD (OpenAPI 3.1)	05-2029
Interfejs API FHIR R5 – warstwa kliniczna i interoperacyjność z E-Zdrowie P1	05-2029

3. KAMIENIE MIŁOWE

Kamienie milowe	Planowany termin osiągnięcia
Zatwierdzony Raport z inicjalnego testu prywatności (DPIA, w tym pogłębiona analiza w zakresie zastosowania sztucznej inteligencji zgodnie z opinią Europejskiej Rady Ochrony Danych 28/2024 z 17 grudnia 2024 r.)	2026-12-31
Opracowane szczegółowe założenia funkcjonalno-techniczne Systemu CDD	2027-03-31
Rozstrzygnięte główne postępowania przetargowe na komponenty Systemu CDD	2027-06-30
Zatwierdzona architektura Systemu CDD	2027-09-30
Uruchomione integracje z systemami źródłowymi	2027-11-30
Zatwierdzony Raport z weryfikacyjnego testu prywatności (DPIA)	2027-12-31
Uruchomione produkcyjnie repozytoria danych Systemu CDD	2028-03-31
Uzupełniona infrastruktura techniczna Systemu CDD	2028-06-30
Uruchomiony pipeline anonimizacji i pseudonimizacji danych	2028-06-30
Uruchomiony interfejs API OAI-PMH / DCAT-AP PL (harvesting metadanych)	2028-07-31
Rozpoczęte procesy testowania oprogramowania	2028-10-31
Zrealizowany I etap digitalizacji	2028-12-15
Zrealizowana Modyfikacja Systemu OpenCardio	2028-12-31
Zrealizowany II etap digitalizacji	2028-12-31
Odebrane Materiały szkoleniowe	2028-12-31
Pozytywny wynik testów badań UX	2029-02-28
Pozytywne wyniki testów oprogramowania – bezpieczeństwo, wydajność, dostępność WCAG 2.2 AA	2029-03-31
Odebrane Materiały informacyjno-promocyjne	2029-03-31
Potwierdzona gotowość Systemu CDD do ubiegania się o certyfikację ECRIN Data Centre Certification, na podstawie zatwierdzonego raportu z audytu wewnętrznego	2029-03-31
Odbiór końcowy Systemu CDD wraz z interfejsami API REST i FHIR R5 (zakończenie rzeczowej realizacji projektu)	2029-05-31

4. KOSZTY

4.1. Koszty ogólne projektu wraz ze sposobem finansowania

Całkowity koszt projektu (netto oraz brutto), w tym	Netto 26 000 000,00 zł Brutto 32 000 000,00 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 5 200 000,00 zł Brutto 6 240 000,00 zł
	2028	Netto 12 350 000,00 zł Brutto 15 620 000,00 zł
	2029	Netto 8 450 000,00 zł Brutto 10 140 000,00 zł

4.2. Wykaz poszczególnych pozycji kosztowych

Nazwa pozycji kosztowej		Przewidywany koszt brutto	Uzasadnienie pozycji kosztowej (przeznaczenie)
Oprogramowanie	Wytworzenie dedykowanego oprogramowania przez pracowników WUM, zakup gotowego oprogramowania systemowego, narzędziowego i specjalistycznego, wszystkie rodzaje testów całego wytworzonego oprogramowania, opracowanie i wdrożenie komponentów pobierania, przetwarzania i publikowania wyników, walidacja danych, anonimizacja danych, katalogowanie danych.	25 540 000,00 zł	Koszty osobowe 25 pracowników etatowych WUM jako FTE ((Full-Time Equivalent - ekwiwalent pełnego czasu pracy) związane z: - opracowanie szczegółowych założeń technicznych i funkcjonalnych, - wytworzeniem całej części merytorycznej oraz aplikacyjnej systemu CDD, - przeprowadzenie wszystkich rodzajów testów, - zasilenie systemu w dane początkowe, - uruchomienie e-usług, - opracowanie dokumentacji technicznej oraz dla użytkowników końcowych, - walidacja danych źródłowych do CDD, - anonimizacja danych źródłowych do CDD - katalogowanie danych źródłowych do CDD
Infrastruktura	Zakup serwerów aplikacyjnych x2	2 500 000,00 zł	Zakup uzupełniające w zakresie infrastruktury:

Nazwa pozycji kosztowej		Przewidywany koszt brutto	Uzasadnienie pozycji kosztowej (przeznaczenie)
	(8-16 vCPU, 64-128 GB RAM, NVMe), serwer DB PostgreSQL HA (8 vCPU, 64 GB RAM), klaster Elasticsearch 3 węzły, Storage S3 MinIO 30 TB, NAS archiwalny 20 TB, networking 10 Gb/s, UPS, switche		- serwerowej, - dyskowej, - archiwizacji, - urządzeń sieciowych, - koszt utrzymania i wsparcia zakupionej infrastruktury w okresie realizacji projektu
Koszty UX i grafiki	Badania UX oraz opracowanie publikowanych danych merytorycznych i graficzne systemów CDD.	0,00 zł	Koszty pracowników merytorycznych WUM zaangażowanych w prace obejmujące badania potrzeb użytkowników oraz UX oraz prace związane z opracowaniem merytorycznym i graficznym systemów ujęte zostały w kosztach oprogramowania.
Bezpieczeństwo	Opracowanie przez pracowników WUM elementów CDD w zakresie bezpieczeństwa	700 000,00 zł	Keycloak 24+ wdrożenie, WAF ModSecurity, SIEM+ELK Stack, testy penetracyjne (zewn. firma), audyt WCAG 2.2 AA (zewn.), ISO 27001 SZBI wdrożenie + certyfikacja, konfiguracja BCP/DR, testy restore.
Wydajność rozwiązań	Ujęte w kosztach oprogramowania	0,00 zł	Ujęte w kosztach oprogramowania
Szkolenia	Szkolenia zespołu specjalistów AI/Data Science oraz szkolenia specjalistyczne zespołu IT	300 000,00 zł	Przewidziane są szkolenia: - podnoszenie kwalifikacje pracowników WUM analizy danych, - specjalistów AI/Data Science, - szkolenia specjalistyczne zespołu IT
Działania informacyjno-promocyjne	Działania informacyjne, promocja projektu (w tym przygotowanie materiałów informacyjno-promocyjnych)	120 000,00 zł	Wydatki pozwolą na prowadzenie działań informacyjnych na temat Projektu, a także mających na celu dotarcie do większej liczby użytkowników i rozszerzenie grupy odbiorców produktów projektu. Działania będą przede wszystkim promować efekty Projektu zarówno w mediach elektronicznych, jaki i tradycyjnych.
Koszty zarządzania	Wynagrodzenia	2 840 000,00 zł	Koszty personelu wchodzącego

Nazwa pozycji kosztowej		Przewidywany koszt brutto	Uzasadnienie pozycji kosztowej (przeznaczenie)
i wsparcia (w tym wynagrodzenia personelu wspomagającego)	członków zespołu projektowego w ramach kosztów pośrednich (7% kosztów kwalifikowanych)		w skład biura projektu zaangażowanego w realizację projektu w ramach kosztów pośrednich: - zarządzanie, - rozliczanie, - monitorowanie projektu, - prowadzenie działań administracyjnych, - obsługa kadrowa, finansowa, księgowa, prawna, zamówień publicznych, prawna.

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

Całkowity koszt utrzymania trwałości projektu (brutto)	13 614 000,00 zł		Źródło finansowania
Podział całkowitego kosztu utrzymania trwałości projektu na poszczególne lata (netto oraz brutto)	2029	1 100 000,00 zł (brutto) (920 000,00 zł netto)	krajowe środki publiczne - budżet państwa
	2030	2 400 000,00 zł (brutto) (2 000 000,00 zł netto)	krajowe środki publiczne - budżet państwa
	2031	2 600 000,00 zł (brutto) (2 170 000,00 zł netto)	krajowe środki publiczne - budżet państwa
	2032	2 880 000,00 zł (brutto) (2 400 000,00 zł netto)	krajowe środki publiczne - budżet państwa
	2033	3 050 000,00 zł (brutto) (2 541 700,00 zł netto)	krajowe środki publiczne - budżet państwa
	2034	1 584 000,00 zł (brutto) (1 320 000,00 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
Zatrudnienie etatowe w WUM wymaganych specjalistów	Duża	Średnie	Rozpoczynanie procedury zatrudnienia z wyprzedzeniem. Skuteczne przeprowadzenie procedury zatrudnienia wymaganych specjalistów. Zapewnienie atrakcyjnych warunków finansowych, stworzenie długookresowej perspektywy rozwoju zawodowego.
Przedłużające się procedury przetargowe	Średnia	Średnie	Rozpoczynanie procedur przetargowych z wyprzedzeniem. Właściwie opisanie przedmiotu zamówienia, dobrze opracowanie kryteriów i sposobów oceny ofert oraz warunków udziału w postępowaniu. Zapewnienie współpracy z komórkami organizacyjnymi WUM w celu zapewnienia prawidłowego przeprowadzenia postępowań.
Zmiany rynkowe związane ze zmianami cen usług oraz wynagrodzeń podczas realizacji projektu	Średnia	Niskie	Dogłębna analiza związana z wszelkimi aspektami budowy i wdrożenia systemu. Ciągłe monitorowanie trendów rynkowych. Zabezpieczenie odpowiednich rezerw finansowych.
Zmiany w otoczeniu zewnętrznym (w obszarze prawnym i obszarze IT oraz AI)	Średnia	Wysokie	Stały monitoring zmian w otoczeniu prawnym i technologicznym, a także elastyczne podejście do zarządzania projektem, które umożliwi dostosowanie się do nowych okoliczności. Rozpoznanie i monitorowanie czy planowane są zmiany w otoczeniu prawnym i otoczeniu IT oraz AI i bieżące reagowanie na zachodzące zmiany.
Brak możliwości zapewnienia na każdym etapie realizacji projektu składu Zespołu Projektowego o	Mała	Średnie	Zarządzanie tymi ryzykami wymaga kompetentnego i doświadczonego kierownictwa projektu, efektywnej komunikacji oraz ciągłego monitorowania i szkolenia zespołu. Zapewnienie możliwości przesunięcie

Nazwa ryzyka	Siła oddziaływania	Prawdopodobieństwo wystąpienia ryzyka	Sposób zarządzania ryzykiem
wystarczających: 1. umiejętnościach technicznych, 2. doświadczeniu z projektami tej skali i specyfiki			zasobów ludzkich aby zapewnić odpowiednią ilość doświadczonego personelu w kluczowych momentach projektu. Skuteczne zastosowanie metodyk zarządzania projektami, które mają na celu minimalizowanie tego typu ryzyk. Monitorowanie doświadczenia członków zespołu projektowego, regularne szkolenia i podnoszenie kwalifikacji członków zespołu projektowego. Odpowiednie prowadzenie procesów rekrutacyjnych.
Nieprecyzyjne określenie wymagań biznesowych i funkcjonalnych w stosunku do wymagań prawnych i potrzeb użytkowników.	Średnia	Niskie	Precyzyjne określenie potrzeb i wymagań dla poszczególnych elementów systemu. Zatwierdzania przez kierownictwo projektu oraz władze WUM koncepcji biznesowej, funkcjonalnej i sposobu realizacji poszczególnych elementów systemu.
Wyciek PHI (Protected Health Information – chronione informacje zdrowotne) przez błąd anonimizacji	Duża	Niskie	Niepełna anonimizacja PHI przed załadowaniem do repozytorium publicznego (ryzyko RODO Art. 9). Automatyczne testy anonimizacji przed każdym importem danych; DPIA; zewnętrzny audyt RODO (M16); logi KMS/HSM; procedura rollback i powiadomienia UODO
Odrzucenie harvestingu przez PPM (Polską Platformę Medyczną)	Średnia	Niskie	Błędy w schemacie DCAT-AP PL 1.2 uniemożliwiające pobieranie metadanych do dane.gov.pl Automatyczna walidacja JSON-LD przy każdym imporcie; szablony metadanych zatwierdzone przez OPI PIB.
Niezgodność z EHDS, FHIR lub Beacon v2	Średnia	Niskie	Zmiana specyfikacji Beacon v2 lub FHIR R5 w trakcie realizacji projektu Stały monitoring zmian w standardach EHDS, GA4GH, Beacon v2 i HL7 FHIR (Health Level 7 Fast Healthcare Interoperability Resources), udział w pracach grup roboczych oraz bieżąca aktualizacja założeń interoperacyjności projektu.
Przekroczenie	Średnia	Niskie	Stosowanie metodyk zarządzania

Nazwa ryzyka	Siła oddziaływania	Prawdopodobieństwo wystąpienia ryzyka	Sposób zarządzania ryzykiem
harmonogramu realizacji projektu			projektami z bieżącym monitoringiem postępu prac na poziomie pakietów roboczych i kamieni milowych. Cykliczne (miesięczne) raporty postępu prowadzone przez biuro projektu z wczesnym sygnalizowaniem odchyłeń od ścieżki krytycznej. Zaplanowane bufory czasowe przy kluczowych kamieniach milowych – analizie, postępowaniach przetargowych, odbiorach testowych i końcowym. Wyprzedzające uruchamianie procedur przetargowych i rekrutacyjnych z marginesem na ewentualne unieważnienia lub przedłużenia. Eskalacja zagrożeń harmonogramowych do Komitetu Sterującego projektu z możliwością przesunięć zasobów, korekty zakresu w ramach uzgodnionego budżetu lub realokacji prac między pakietami roboczymi.
Nieosiągnięcie wskaźników produktu oraz celu projektu	Duża	Niskie	Konserwatywne wartości docelowe wskaźników oparte na zweryfikowanych szacunkach zespołów merytorycznych oraz na analizie istniejących zasobów źródłowych (PPM WUM, OpenCardio, Repozytorium materiałów dydaktycznych WUM, REDCap). Bieżący monitoring postępu osiągania wskaźników w cyklach kwartalnych, z mapowaniem każdego KPI na konkretne produkty i komponenty Systemu CDD oraz na właściwe pakiety prac. Przy odchyleniach od trajektorii docelowej – uruchamianie działań korygujących, w tym rozszerzenia zakresu udostępniania w ramach istniejących źródeł danych, intensyfikacji digitalizacji w komponentach z nadwyżką zasobów oraz wzmocnienia działań informacyjno-promocyjnych zwiększających liczbę użytkowników nowych usług cyfrowych. Identyfikacja alternatywnych źródeł zasobów możliwych do uwzględnienia we wskaźnikach, jeżeli pierwotnie planowany zasób nie zostanie udostępniony zgodnie z harmonogramem. Zatwierdzanie

Nazwa ryzyka	Siła oddziaływania	Prawdopodobieństwo wystąpienia ryzyka	Sposób zarządzania ryzykiem
			istotnych zmian w wartościach wskaźników przez kierownictwo projektu i władze WUM po analizie wpływu na cele projektu.
Brak wystarczających środków na realizację projektu	Duża	Niskie	Bieżący monitoring wydatków w odniesieniu do planowanego harmonogramu kosztowego (forecast vs budget) prowadzony przez biuro projektu w cyklach miesięcznych, z analizą wpływu zmian rynkowych na pozycje kosztowe. Zabezpieczenie odpowiednich rezerw finansowych w pozycjach kosztowych wrażliwych na zmiany cen rynkowych (infrastruktura, wynagrodzenia specjalistów). Ścisła współpraca z dysponentem części budżetowej 27 w zakresie zapewnienia płynności finansowania kolejnych transz, w tym wczesne sygnalizowanie potrzeb płatniczych i terminowe składanie wniosków o płatność. W razie zagrożenia płynnością – przesunięcia między pozycjami kosztowymi w granicach dopuszczonych umową o dofinansowanie. Optymalizacja procedur zamówień publicznych pod kątem najlepszego stosunku jakości do ceny oraz zabezpieczanie warunków cenowych w umowach wieloletnich, ograniczające ekspozycję na inflację. W ostateczności podejmowanie działań w celu pozyskania finansowania pomostowego do czasu rozliczenia transz, w uzgodnieniu z władzami WUM i dysponentem.

5.2. Ryzyka wpływające na utrzymanie efektów

Nazwa ryzyka	Siła oddziaływania	Prawdopodobieństwo wystąpienia ryzyka	Sposób zarządzania ryzykiem
Incydent cyberbezpieczeństwa (np. ransomware lub	Duża	Niskie	WAF+ModSecurity CRS 3.3+; SIEM+ELK; Backup 3-2-1 (RPO<4h); testy penetracyjne (1x/rok); ISO 27001 SZBI; procedury BCP/DR

Nazwa ryzyka	Siła oddziaływania	Prawdopodobieństwo wystąpienia ryzyka	Sposób zarządzania ryzykiem
DDoS na infrastrukturę platformy CDD WUM)			
Niższe niż planowane zainteresowanie wykorzystaniem produktów projektu przez różne grupy Interesariuszy	Średnia	Średnie	Doskonalenie funkcjonalności systemu stosownie do zgłaszanych propozycji usprawnień oraz uwag Interesariuszy. Prowadzenie stałego procesu promocji projektu. Rozszerzanie zakresu danych oraz gotowych scenariuszy do przetwarzania danych własnych Interesariuszy
Brak zastępowalności i niedostępność specjalistów	Średnia	Wysokie	Racjonalna przemyślana polityka kadrowa oraz system motywacyjny. Racjonalne nabywanie usług zewnętrznych.
Brak wsparcia technicznego dla technologii wspierającej produkty projektu	Mała	Niskie	Zapewnienie odpowiednich kompetencji poprzez zawieranie umów serwisowych z dostawcami technologii, które gwarantują odpowiednie wsparcie techniczne. Dokładne badanie różnych dostępnych technologii i kierowanie się nie tylko ich obecną popularnością czy łatwością użycia, ale również ich potencjałem do długoterminowego wsparcia i rozwoju. Utrzymywanie regularnych kontaktów z dostawcami technologii oraz monitorowanie ich polityk wsparcia. Inwestowanie w szkolenia dla zespołu utrzymaniowego i rozwijającego system CDD aby zwiększyć ich kompetencje techniczne.
Znaczące zmiany w przepisach prawa i w zakresie standardów obowiązujących całą administrację publiczną, które wpływają na system wytworzony w ramach projektu	Mała	Niskie	Modyfikowanie systemów w sposób elastyczny i modularny, co umożliwi łatwe wprowadzanie zmian. Utrzymywanie ścisłej współpracy z organami regulacyjnymi i prawnymi, aby na bieżąco monitorować wszelkie zmiany w przepisach i standardach, które mogą wpłynąć na system
Niewystarczające środki w budżecie dysponenta na	Średnia	Średnie	Odpowiednio wczesne planowanie środków na zapewnienie usług utrzymania i rozwoju.

Nazwa ryzyka	Siła oddziaływania	Prawdopodobieństwo wystąpienia ryzyka	Sposób zarządzania ryzykiem
zapewnienie efektów projektu			Przedstawienie problemu kierownictwu wyższego szczebla i władzom WUM. Wygospodarowanie środków z innych zadań realizowanych u Beneficjenta lub podejmowanie działań mających na celu pozyskanie z budżetu środków na utrzymanie efektów Projektu po jego zakończeniu, a także w pozostałym, całym okresie trwałości.
Brak zapewnienia odpowiednich zasobów personalnych odpowiedzialnych za rozwój i utrzymanie systemów	Średnia	Średnie	Konieczność zabezpieczenia kadry merytorycznej i informatycznej do utrzymania systemów oraz ich dalszego rozwoju
Nieosiągnięcie wszystkich zaplanowanych korzyści	Średnia	Niskie	Bieżący monitoring stopnia wykorzystania Systemu CDD i osiąganych wskaźników korzyści w okresie trwałości, prowadzony przez wyznaczonego koordynatora ds. CDD przy WUM. Cykliczne (półroczne) przeglądy realizacji korzyści (benefits realization review) z udziałem przedstawicieli kluczowych grup interesariuszy – naukowców, doktorantów, klinicystów oraz jednostek dydaktycznych. Stałe rozszerzanie zakresu udostępnianych zasobów oraz dostosowywanie funkcjonalności Systemu CDD do potrzeb zgłaszanych przez użytkowników. Kontynuowanie działań informacyjno-promocyjnych oraz programu szkoleń po zakończeniu projektu, w celu utrzymania i rozszerzania bazy aktywnych użytkowników. Wzmacnianie współpracy strategicznej z partnerami krajowymi (PPM, dane.gov.pl, Ministerstwo Zdrowia) oraz międzynarodowymi (HealthData@EU, OpenAIRE, EOSC, ECRIN-ERIC), zwiększające widoczność zasobów CDD i tworzące zewnętrzne mechanizmy walidacji wartości projektu.

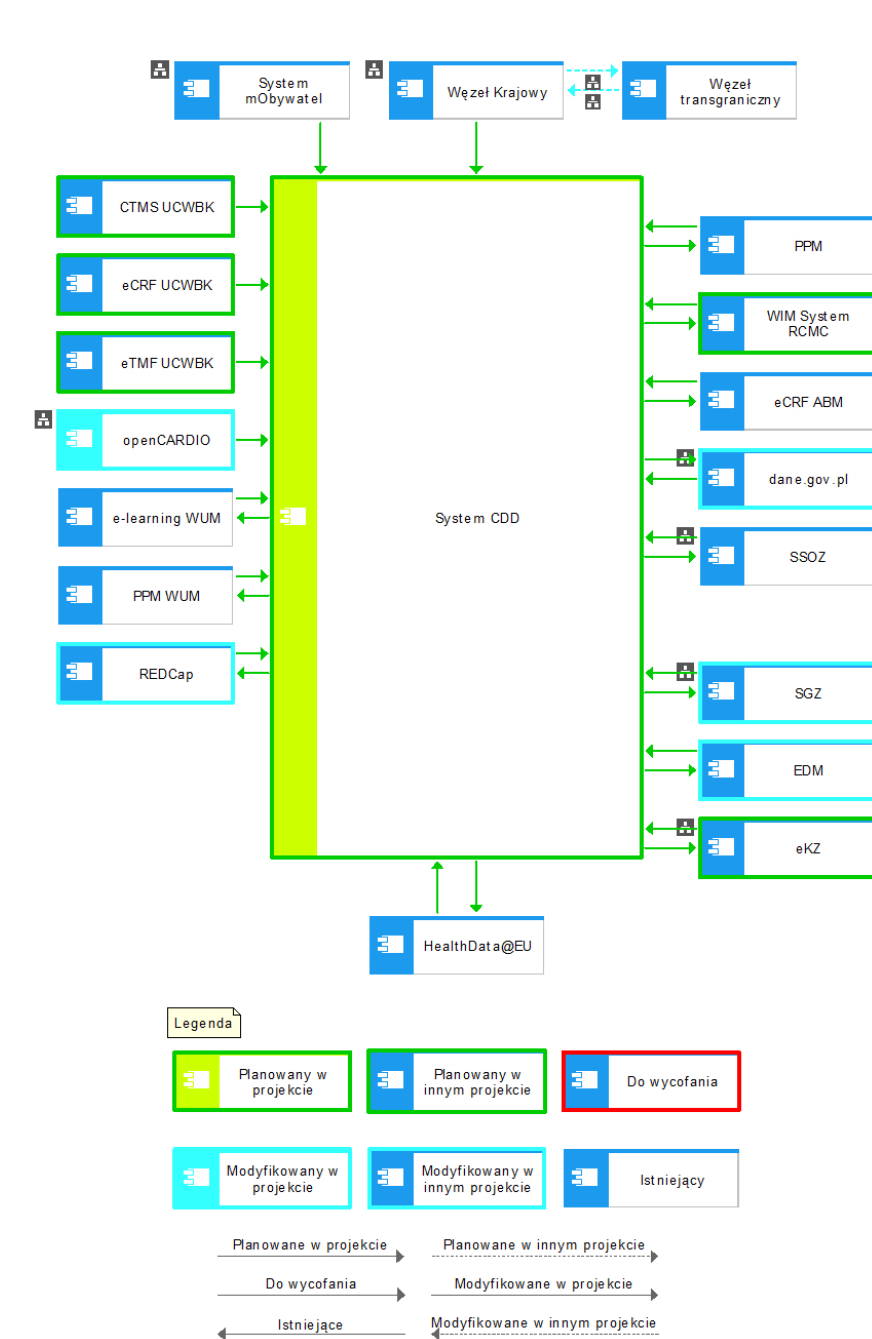
6. OTOCZENIE PRAWNE

Lp.	Tytuł aktu prawnego	Czy wymaga zmian	Opis zmian (jeśli dotyczy)	Etap prac legislacyjnych (jeśli dotyczy)
1	Prawo o szkolnictwie wyższym i nauce (t.j. Dz. U. z 2024 r. poz. 1571 z późn. zm.)	TAK /NIE		
2	Ustawa z dnia 28 kwietnia 2011 r. o systemie informacji w ochronie zdrowia (t.j. Dz.U. z 2021 r. poz. 657, z późn. zm.)	TAK /NIE		
3	Ustawa z dnia 17 lutego 2005 r. o informatyzacji działalności podmiotów realizujących zadania publiczne (t.j. Dz. U. z 2025 r. poz. 1703	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 (Dz.U. 2018 r. poz. 1560, z późn. zm.)	TAK /NIE		
6	Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2016/679 (RODO)	TAK /NIE		
7	Ustawa z dnia 28 kwietnia 2022 r. o zasadach realizacji zadań finansowanych ze środków europejskich w perspektywie finansowej 2021-2027	TAK /NIE		
8	Rozporządzenie Parlamentu Europejskiego i Rady (UE) nr 2021/1060 z dnia 24 czerwca 2021r	TAK /NIE		
9	Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2021/1058 z dnia 24 czerwca 2021r. ustanawiające wspólne przepisy dotyczące funduszy europejskich	TAK /NIE		
10	Ustawa z dnia 5 września 2016 r. o usługach zaufania oraz identyfikacji elektronicznej	TAK /NIE		
11	Kodeks postępowania administracyjnego (KPA) z dnia 14 czerwca 1960 r.	TAK /NIE		
12	Ustawa z dnia 11 sierpnia 2021 r. o otwartych danych i ponownym wykorzystywaniu informacji sektora publicznego	TAK /NIE		
13	Ustawa z dnia 27 lipca 2001 r. o ochronie baz danych (t.j. Dz. U. z 2024 r. poz. 1254)	TAK /NIE		

Lp.	Tytuł aktu prawnego	Czy wymaga zmian	Opis zmian (jeśli dotyczy)	Etap prac legislacyjnych (jeśli dotyczy)
14	Ustawa z dnia 4 kwietnia 2019 r. o dostępności cyfrowej stron internetowych i aplikacji mobilnych podmiotów publicznych	TAK /NIE		
15	Ustawa z dnia 18 listopada 2020 r. o doręczeniach elektronicznych (t.j. Dz. U. z 2026 r. poz. 3)	TAK /NIE		
16	Rozporządzenie Ministra Cyfryzacji z dnia 16 czerwca 2021 r. w sprawie profilu zaufanego i podpisu zaufanego	TAK /NIE		
17	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	TAK /NIE		
18	Rozporządzenie Prezesa Rady Ministrów z dnia 20 lipca 2011 r. w sprawie podstawowych wymagań bezpieczeństwa teleinformatycznego	TAK /NIE		
19	Dyrektywa 2022/2555 z dnia 14 grudnia 2022 w sprawie środków na rzecz wysokiego wspólnego poziomu cyberbezpieczeństwa w całej Unii (Dyrektywa NIS2 - Network and Information Security Directive 2)	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	CTMS UCWBK	Warszawski Uniwersytet Medyczny	CTMS UCWBK (Clinical Trial Management System Uniwersyteckiego Centrum Wsparcia Badań Klinicznych) to system wspierający zarządzanie badaniami klinicznymi prowadzonymi w	Planowany	

Lp.	Nazwa systemu	Gestor systemu	Opis systemu	Status	Krótki opis ewentualnej zmiany
			środowisku Uniwersyteckiego Centrum Wsparcia Badań Klinicznych Warszawskiego Uniwersytetu Medycznego. Celem systemu jest zapewnienie zorganizowanego środowiska planowania, monitorowania i nadzorowania procesu badań klinicznych. System nie prowadzi rejestru publicznego. Główne grupy funkcjonalności obejmują: • planowanie badań klinicznych i harmonogramów, • zarządzanie ośrodkami i wizytami badanych, • ewidencję statusów oraz etapów realizacji badań, • obsługę procesów administracyjnych i organizacyjnych związanych z badaniami klinicznymi, • gromadzenie metadanych dotyczących przebiegu badań. System integruje się z innymi systemami wykorzystywanymi w środowisku badań klinicznych Warszawskiego Uniwersytetu Medycznego.		
2	dane.gov.pl	Ministerstwo Cyfryzacji	Dane.gov.pl to system wspierający udostępnianie informacji sektora publicznego w celu ich ponownego wykorzystywania. System został utworzony na podstawie ustawy o otwartych danych i	Modyfikowany	Modyfikowany poza Projektem CDD

Lp.	Nazwa systemu	Gestor systemu	Opis systemu	Status	Krótki opis ewentualnej zmiany
			ponownym wykorzystywaniu informacji sektora publicznego oraz rozporządzenia Rady Ministrów w sprawie portalu danych. System nie prowadzi rejestrów publicznych, lecz umożliwia dostęp do danych pochodzących z takich rejestrów, udostępnianych przez inne podmioty administracji publicznej. Główne grupy funkcjonalności obejmują: • publikację zbiorów danych w różnych formatach, • wyszukiwarkę danych, • API do automatycznego pobierania danych, • mechanizmy oceny jakości danych, • możliwość zgłaszania zapotrzebowania na dane. System integruje się z innymi krajowymi systemami informatycznymi.		
3	eCRF ABM	Agencja Badań Medycznych	eCRF ABM (Electronic Case Report Form Agencji Badań Medycznych) to system wspierający prowadzenie badań niekomercyjnych finansowanych przez Agencję Badań Medycznych. Celem systemu jest zapewnienie standaryzowanego środowiska gromadzenia danych badawczych powstających w toku badań klinicznych prowadzonych w jednostkach realizujących projekty Agencji Badań	Istniejący	

Lp.	Nazwa systemu	Gestor systemu	Opis systemu	Status	Krótki opis ewentualnej zmiany
			Medycznych. System nie prowadzi rejestru publicznego. Główne grupy funkcjonalności obejmują: • zarządzanie danymi badawczymi, • obsługę formularzy obserwacji klinicznych, • realizację wybranych procesów raportowych, • kontrolę zmian danych badawczych, • obsługę identyfikatorów i statusów badań. System integruje się z systemami informatycznymi wykorzystywanymi w obszarze badań niekomercyjnych oraz z infrastrukturą informatyczną Agencji Badań Medycznych.		
4	eCRF UCWBK	Warszawski Uniwersytet Medyczny	eCRF UCWBK (Electronic Case Report Form Uniwersyteckiego Centrum Wsparcia Badań Klinicznych) to system wspierający gromadzenie ustrukturyzowanych danych badawczych pozyskiwanych w toku badań klinicznych prowadzonych w Uniwersyteckim Centrum Wsparcia Badań Klinicznych Warszawskiego Uniwersytetu Medycznego. Celem systemu jest zapewnienie standaryzowanego środowiska dokumentowania danych obserwacyjnych zgodnie z wymaganiami protokołów badań klinicznych. System nie prowadzi	Planowany	

Lp.	Nazwa systemu	Gestor systemu	Opis systemu	Status	Krótki opis ewentualnej zmiany
			rejestru publicznego. Główne grupy funkcjonalności obejmują: • obsługę elektronicznych formularzy obserwacji klinicznych, • gromadzenie wyników wizyt badanych, • ewidencję parametrów klinicznych, • kontrolę kompletności i jakości danych badawczych, • obsługę procesów eksportu danych do dalszego przetwarzania. System integruje się z innymi systemami wykorzystywanymi w środowisku badań klinicznych Warszawskiego Uniwersytetu Medycznego.		
5	EDM	Ministerstwo Zdrowia	System EDM (Elektronicznej Dokumentacji Medycznej) to system wspierający obsługę indeksów elektronicznej dokumentacji medycznej. Celem systemu jest umożliwienie wymiany informacji o istnieniu dokumentu medycznego i jego lokalizacji w zewnętrznym repozytorium dokumentów wykorzystywanym przez usługodawcę. System prowadzi rejestr indeksów elektronicznej dokumentacji medycznej. Główne grupy funkcjonalności obejmują: • obsługę indeksów elektronicznej dokumentacji medycznej,	Modyfikowany	Modyfikowany poza Projektem CDD

Lp.	Nazwa systemu	Gestor systemu	Opis systemu	Status	Krótki opis ewentualnej zmiany
			• ewidencję informacji o lokalizacji dokumentów w repozytoriach źródłowych, • udostępnianie informacji o dokumentacji medycznej uprawnionym podmiotom, • wymianę metadanych dokumentów pomiędzy uczestnikami systemu ochrony zdrowia. System jest komponentem Systemu P1 i współpracuje z innymi krajowymi systemami informatycznymi w obszarze ochrony zdrowia.		
6	eKZ	Ministerstwo Zdrowia	System generowania elektronicznych kart urodzeń i zgonów (eKZ) to system wspierający proces tworzenia, korekty oraz anulowania dokumentów związanych z urodzeniami i zgonami. Celem systemu jest zapewnienie elektronicznego obiegu tych dokumentów oraz ich udostępnienie uprawnionym podmiotom, takim jak Urzędy Stanu Cywilnego (USC), Główny Urząd Statystyczny (GUS) i Główny Inspektorat Sanitarny (GIS). System nie prowadzi odrębnego rejestru publicznego, lecz wspiera istniejące rejestry poprzez przekazywanie danych. Główne grupy funkcjonalności obejmują: • generowanie elektronicznych kart urodzeń i zgonów, • obsługę korekt i	Planowany	

Lp.	Nazwa systemu	Gestor systemu	Opis systemu	Status	Krótki opis ewentualnej zmiany
			anulowania dokumentów, • udostępnianie dokumentów uprawnionym instytucjom, • integrację z aplikacją gabinet.gov.pl oraz aplikacjami dostawców komercyjnych. System jest komponentem Systemu P1 i współpracuje z innymi krajowymi rozwiązaniami informatycznymi w obszarze ochrony zdrowia.		
7	e-learning WUM	Warszawski Uniwersytet Medyczny	System e-learning WUM (System E-Learningu Warszawskiego Uniwersytetu Medycznego) to system wspierający proces dydaktyczny prowadzony w Warszawskim Uniwersytecie Medycznym. Celem systemu jest zapewnienie elektronicznego środowiska kształcenia oraz udostępniania materiałów dydaktycznych dla studentów, doktorantów i pracowników uczelni. System nie prowadzi rejestru publicznego. Główne grupy funkcjonalności obejmują: • gromadzenie i organizację materiałów dydaktycznych, • udostępnianie wykładów, prezentacji oraz materiałów multimedialnych, • obsługę atlasów anatomicznych i specjalistycznych zasobów dydaktycznych,	Istniejący	

Lp.	Nazwa systemu	Gestor systemu	Opis systemu	Status	Krótki opis ewentualnej zmiany
			• zarządzanie zasobami szkoleniowymi, • wsparcie procesu kształcenia w trybie elektronicznym. System integruje się z innymi systemami informatycznymi wykorzystywanymi w środowisku Warszawskiego Uniwersytetu Medycznego.		
8	eTMF UCWBK	Warszawski Uniwersytet Medyczny	eTMF UCWBK (Electronic Trial Master File Uniwersyteckiego Centrum Wsparcia Badań Klinicznych) to system wspierający prowadzenie elektronicznej teczki badania klinicznego w Uniwersyteckim Centrum Wsparcia Badań Klinicznych Warszawskiego Uniwersytetu Medycznego. Celem systemu jest zapewnienie zorganizowanego środowiska przechowywania i obsługi dokumentacji formalnej, regulacyjnej i operacyjnej związanej z badaniami klinicznymi. System nie prowadzi rejestru publicznego. Główne grupy funkcjonalności obejmują: • prowadzenie elektronicznej teczki badania klinicznego, • ewidencję metadanych dokumentów badania, • obsługę statusów obiegu dokumentów, • monitorowanie kompletności dokumentacji,	Istniejący	

Lp.	Nazwa systemu	Gestor systemu	Opis systemu	Status	Krótki opis ewentualnej zmiany
			• obsługę identyfikatorów dokumentów oraz informacji formalnych dotyczących badań klinicznych. System integruje się z innymi systemami wykorzystywanymi w środowisku badań klinicznych Warszawskiego Uniwersytetu Medycznego.		
9	HealthData@EU	Komisja Europejska	HealthData@EU to system wspierający wtórne wykorzystanie danych zdrowotnych w ramach Europejskiej Przestrzeni Danych dotyczących Zdrowia (EHDS). Celem systemu jest umożliwienie wyszukiwania, katalogowania oraz wymiany informacji o zasobach danych zdrowotnych pomiędzy krajowymi punktami dostępu do danych w państwach członkowskich Unii Europejskiej. System nie prowadzi rejestru publicznego. Główne grupy funkcjonalności obejmują: • wyszukiwanie zasobów danych zdrowotnych, • katalogowanie zbiorów danych zdrowotnych, • obsługę dostępu do danych zdrowotnych do celów wtórnego wykorzystania, • wspieranie współpracy pomiędzy krajowymi punktami dostępu do danych, • udostępnianie	Istniejący	

Lp.	Nazwa systemu	Gestor systemu	Opis systemu	Status	Krótki opis ewentualnej zmiany
			metadanych i informacji o dostępności zasobów. System integruje się z krajowymi punktami dostępu do danych zdrowotnych państw członkowskich Unii Europejskiej w ramach infrastruktury Europejskiej Przestrzeni Danych dotyczących Zdrowia.		
10	openCARDIO	Warszawski Uniwersytet Medyczny	openCARDIO (Bank Otwartych Danych Naukowych w Zakresie Diagnostyki i Leczenia Chorób Układu Krążenia) to system wspierający gromadzenie i udostępnianie zbiorów danych naukowych dotyczących diagnostyki i leczenia chorób układu krążenia. Celem systemu jest umożliwienie dostępu do uporządkowanych zbiorów danych medycznych, obrazowych i publikacyjnych z obszaru kardiologii na potrzeby badań naukowych, dydaktyki i analiz. System nie prowadzi rejestru publicznego. Główne grupy funkcjonalności obejmują: • gromadzenie i udostępnianie danych klinicznych: historii chorób, czynników ryzyka, diagnostyki przedszpitalnej, badań podmiotowych i przedmiotowych, chorób towarzyszących, wyników badań laboratoryjnych i obrazowych, • przechowywanie danych obserwacyjnych dotyczących leczenia i	Modyfikowany	

Lp.	Nazwa systemu	Gestor systemu	Opis systemu	Status	Krótki opis ewentualnej zmiany
			zaleceń poszpitalnych oraz zapisów EKG, • gromadzenie wyników badań i obrazów echokardiograficznych, tomografii komputerowej klatki piersiowej oraz badań ultrasonograficznych żył kończyn dolnych i jamy brzusznej u chorych na żylną chorobę zakrzepowo-zatorową, • udostępnianie obrazów innych badań radiologicznych (tomografia jamy brzusznej, rezonans magnetyczny serca, naczyń płucnych, głowy, RTG klatki piersiowej), • gromadzenie publikacji, rozpraw doktorskich i habilitacyjnych, wystąpień zjazdowych oraz monografii pracowników Kliniki Chorób Wewnętrznych i Kardiologii. System integruje się z innymi systemami informatycznymi wykorzystywanymi w środowisku Warszawskiego Uniwersytetu Medycznego. Zmiany w ramach projektu CDD: uporządkowanie i konsolidacja zasobów danych, dostosowanie interfejsów wymiany danych z Systemem CDD oraz przygotowanie środowiska eksploatacyjnego do współpracy z platformą CDD.		

Lp.	Nazwa systemu	Gestor systemu	Opis systemu	Status	Krótki opis ewentualnej zmiany
11	PPM	Uniwersyte t Medyczny im. Piaśtów Śląskich we Wrocławiu	PPM (Polska Platforma Medyczna) to system wspierający zarządzanie wiedzą i potencjałem badawczym w obszarze nauk medycznych. Celem systemu jest umożliwienie gromadzenia i udostępniania publikacji, datasetów, metadanych oraz innych zasobów nauki tworzonych przez polskie ośrodki medyczne. System nie prowadzi rejestru publicznego. Główne grupy funkcjonalności obejmują: • ewidencjonowanie i katalogowanie publikacji naukowych, • udostępnianie datasetów i metadanych badawczych, • obsługę identyfikatorów obiektów naukowych, • synchronizację metadanych pomiędzy modułami uczelnianymi a platformą wspólną, • udostępnianie zasobów odbiorcom zewnętrznym. System integruje się z modułami uczelnianymi obsługującymi zasoby poszczególnych polskich uczelni medycznych (m.in. PPM WUM) oraz z innymi systemami informatycznymi obsługującymi zasoby naukowe.	Istniejący	
12	PPM WUM	Warszawsk i Uniwersyte t Medyczny	PPM WUM (Polska Platforma Medyczna – Moduł Warszawskiego Uniwersytetu Medycznego) to repozytorium lokalne	Istniejący	

Lp.	Nazwa systemu	Gestor systemu	Opis systemu	Status	Krótki opis ewentualnej zmiany
			systemu Polskiej Platformy Medycznej, wspierające zarządzanie dorobkiem naukowym i danymi badawczymi Warszawskiego Uniwersytetu Medycznego w ramach Polskiej Platformy Medycznej. Celem systemu jest umożliwienie ewidencjonowania, opisywania, katalogowania i udostępniania publikacji, datasetów oraz metadanych badawczych powstających w uczelni. System nie prowadzi rejestru publicznego. Główne grupy funkcjonalności obejmują: • ewidencjonowanie publikacji naukowych i datasetów, • katalogowanie metadanych badawczych, • udostępnianie publikacji i danych badawczych odbiorcom zewnętrznym, • obsługę identyfikatorów powiązanych z obiektami naukowymi, • wsparcie procesu identyfikacji dorobku naukowego uczelni. System integruje się z Polską Platformą Medyczną oraz z innymi systemami informatycznymi obsługującymi zasoby naukowe Warszawskiego Uniwersytetu Medycznego.		
13	REDCap	Warszawsk i Uniwersyte	REDCap (Research Electronic Data Capture) to system wspierający	Modyfikowany	

Lp.	Nazwa systemu	Gestor systemu	Opis systemu	Status	Krótki opis ewentualnej zmiany
		t Medyczny	prorowadzenie badań naukowych i klinicznych w obszarze gromadzenia ustrukturyzowanych danych badawczych. Celem systemu jest zapewnienie elastycznego środowiska tworzenia formularzy elektronicznych oraz obsługi procesów związanych z zarządzaniem danymi projektowymi. System nie prowadzi rejestru publicznego. Główne grupy funkcjonalności obejmują: • tworzenie elektronicznych formularzy badawczych, • gromadzenie ustrukturyzowanych danych badawczych, • obsługę ankiet i rejestrów badawczych, • zarządzanie projektami badawczymi, • obsługę identyfikatorów rekordów oraz informacji o statusie danych, • wsparcie procesów eksportu danych. System integruje się z innymi systemami informatycznymi wykorzystywanymi w środowisku badawczym Warszawskiego Uniwersytetu Medycznego.		
14	SGZ	Ministerstwo Zdrowia	SGZ (System Gromadzenia Zdarzeń Medycznych) to system wspierający rejestrowanie, gromadzenie i przetwarzanie danych o zdarzeniach medycznych. Celem systemu jest	Modyfikowany	Modyfikowany poza Projektem CDD

Lp.	Nazwa systemu	Gestor systemu	Opis systemu	Status	Krótki opis ewentualnej zmiany
			zapewnienie centralnej infrastruktury obsługi zdarzeń medycznych powstających w trakcie udzielania świadczeń zdrowotnych. System prowadzi rejestr zdarzeń medycznych. Główne grupy funkcjonalności obejmują: • rejestrowanie zdarzeń medycznych zgłaszanych przez usługodawców, • gromadzenie i przetwarzanie danych o zdarzeniach medycznych, • udostępnianie zdarzeń medycznych podmiotom uprawnionym, • wymianę danych o zdarzeniach z systemami lokalnymi usługodawców, • udostępnianie informacji o zdarzeniach medycznych użytkownikom końcowym za pośrednictwem aplikacji portalowych, w tym Internetowego Konta Pacjenta. System jest komponentem Systemu P1 i współpracuje z innymi krajowymi systemami informatycznymi w obszarze ochrony zdrowia.		
15	SSOZ	Ministerstwo Zdrowia	SSOZ (Statystyczny System Ochrony Zdrowia) to system wspierający ministra właściwego do spraw zdrowia oraz Prezesa Głównego Urzędu Statystycznego w gromadzeniu i opracowywaniu danych statystycznych związanych z sektorem	Istniejący	

Lp.	Nazwa systemu	Gestor systemu	Opis systemu	Status	Krótki opis ewentualnej zmiany
			ochrony zdrowia w Polsce. Celem systemu jest umożliwienie analizy sytuacji zdrowotnej oraz zjawisk zachodzących w ochronie zdrowia. System nie prowadzi rejestru publicznego. Główne grupy funkcjonalności obejmują: • gromadzenie danych ze sprawozdań statystycznych jednostek sprawozdających, • gromadzenie danych ze sprawozdań finansowych nieobjętych statystyką publiczną, • obsługę formularzy sprawozdawczych przypisanych do jednostek, • weryfikację danych przez przedstawicieli administracji wojewódzkiej oraz inne jednostki nadzorujące, • opracowywanie i udostępnianie danych statystycznych dotyczących ochrony zdrowia. System integruje się z innymi krajowymi systemami informatycznymi w obszarze ochrony zdrowia oraz statystyki publicznej.		
16	System CDD	Warszawski Uniwersytet Medyczny	System CDD (System Centrum Doskonałości Danych) to system wspierający gromadzenie, przetwarzanie i udostępnianie danych badawczych, medycznych i dydaktycznych w jednolitym środowisku danych WUM. Celem systemu jest zapewnienie	Planowany	

Lp.	Nazwa systemu	Gestor systemu	Opis systemu	Status	Krótki opis ewentualnej zmiany
			centralnej infrastruktury wtórnego wykorzystania danych naukowych oraz danych zdrowotnych pochodzących z wielu źródeł, z zachowaniem wymogów anonimizacji, jakości danych i interoperacyjności. System nie prowadzi rejestru publicznego. Udostępnia katalog metadanych zasobów oraz dane przygotowane do ponownego wykorzystania. System CDD składa się z następujących modułów: • Portal CDD – warstwa prezentacji i interakcji z użytkownikiem końcowym, udostępniająca katalog zasobów, wyszukiwarę metadanych, wnioskowanie o dostęp do zbiorów oraz funkcje raportowe; główny kanał interakcji dla badaczy, doktorantów, klinicystów i partnerów instytucjonalnych, • API Gateway CDD – centralny punkt ekspozycji interfejsów API udostępniających dane i metadane odbiorcom wewnętrznym i zewnętrznym (REST/ OpenAPI 3.1, HL7 FHIR R5, OAI-PMH/DCAT-AP PL); zapewnia zarządzanie ruchem, autoryzację, limitowanie i wersjonowanie, • DataLake CDD – pierwsza warstwa architektury danych; gromadzi dane źródłowe		

Lp.	Nazwa systemu	Gestor systemu	Opis systemu	Status	Krótki opis ewentualnej zmiany
			po wstępnej anonimizacji i pseudonimizacji o zróżnicowanym stopniu uporządkowania, • Hurtownia danych CDD – druga warstwa architektury; zawiera dane ustrukturyzowane, zwalidowane jakościowo i zharmonizowane do wspólnych modeli danych; wspiera analizy naukowe i raportowanie, • Repozytorium danych CDD – trzecia warstwa; przechowuje datasey FAIR oraz zbiory zweryfikowane. System integruje się z systemami źródłowymi środowiska WUM (CTMS UCWBK, eCRF UCWBK, eTMF UCWBK, openCARDIO, REDCap, PPM WUM, System e-learning WUM), z systemami inf ochrony zdrowia (EDM, eKZ, SGZ, SSOZ), dane.gov.pl, Węzeł Krajowy, System mObywatel, z systemami badawczymi innych jednostek (eCRF ABM, WIM System RCMC), z europejską infrastrukturą wtórnego wykorzystania danych zdrowotnych HealthData@EU oraz z PPM.		
17	System mObywatel	Ministerstwo Cyfryzacji	System mObywatel to system teleinformatyczny wspierający obywateli w pobieraniu i prezentowaniu dokumentów elektronicznych oraz w korzystaniu z usług cyfrowych za pomocą aplikacji mobilnej	Modyfikowany	Modyfikowany poza Projektem CDD

Lp.	Nazwa systemu	Gestor systemu	Opis systemu	Status	Krótki opis ewentualnej zmiany
			mObywatel. Celem systemu jest zapewnienie wygodnego dostępu do danych pochodzących z rejestrów publicznych oraz innych systemów teleinformatycznych podmiotów współpracujących, wspierając cyfrową obsługę spraw urzędowych. W systemie mObywatel prowadzone są rejestry i mechanizmy udostępniania danych obejmujące m.in.: • dokumenty elektroniczne zawierające dane osobowe użytkownika, • dokumenty potwierdzające sytuację prawną lub prawa użytkownika, • dokumenty umożliwiające identyfikację rzeczy powiązanych z użytkownikiem, • elektroniczne odpowiedniki dokumentów urzędowych wydawanych pierwotnie w postaci papierowej. Główne grupy funkcjonalności obejmują: • obsługę aplikacji mobilnej mObywatel, Portalu dla Szkół i Uczelni oraz Portalu Administracyjnego i back-endu, • zarządzanie dokumentami i usługami dostępnymi w aplikacji, • prezentację powiadomień PUSH, • obsługę środka identyfikacji		

Lp.	Nazwa systemu	Gestor systemu	Opis systemu	Status	Krótki opis ewentualnej zmiany
			elektronicznej Profil mObywatel. System integruje się z krajowymi systemami teleinformatycznymi. Wykorzystanie w projekcie CDD: System mObywatel będzie wykorzystywany jako uzupełniający środek uwierzytelniania mobilnego dla użytkowników korzystających z Portalu CDD z urządzeń mobilnych, w szczególności dla studentów, doktorantów i pracowników WUM, oraz do potwierdzania tożsamości użytkowników przy wnioskowaniu o dostęp do zbiorów danych chronionych. Węzeł Krajowy pozostaje podstawowym mechanizmem uwierzytelniania dla dostępu z przeglądarki internetowej.		
18	Węzeł Krajowy	Ministerstwo Cyfryzacji	Węzeł Krajowy to system wspierający uwierzytelnianie użytkowników systemów teleinformatycznych korzystających z usług online. Celem systemu jest zapewnienie wspólnego mechanizmu identyfikacji elektronicznej obywateli w usługach administracji publicznej oraz innych podmiotów wykorzystujących krajowe środki identyfikacji elektronicznej. System nie prowadzi	Istniejący	

Lp.	Nazwa systemu	Gestor systemu	Opis systemu	Status	Krótki opis ewentualnej zmiany
			rejestru publicznego. Główne grupy funkcjonalności obejmują: • obsługę procesów uwierzytelniania użytkowników, • pośrednictwo pomiędzy systemami identyfikacji elektronicznej a systemami udostępniającymi usługi online, • obsługę atrybutów identyfikacyjnych użytkowników, • udostępnianie identyfikatorów logowania, • obsługę środków identyfikacji elektronicznej notyfikowanych w trybie krajowym. System integruje się z krajowymi systemami identyfikacji elektronicznej (m.in. profil zaufany, profil osobisty z e-dowodu, Profil mObywatel) oraz z Węzłem transgranicznym umożliwiającym współpracę z systemami identyfikacji elektronicznej innych państw członkowskich Unii Europejskiej.		
19	Węzeł transgraniczny	Ministerstwo Cyfryzacji	Węzeł transgraniczny to system wspierający wzajemne uwierzytelnianie obywateli i podmiotów gospodarczych pomiędzy państwami członkowskimi Unii Europejskiej. Celem systemu jest umożliwienie korzystania z krajowych	Modyfikowany	Modyfikowany poza Projektem CDD

Lp.	Nazwa systemu	Gestor systemu	Opis systemu	Status	Krótki opis ewentualnej zmiany
			notyfikowanych środków identyfikacji elektronicznej w usługach elektronicznych dostępnych w innych państwach członkowskich Unii Europejskiej oraz korzystania ze środków identyfikacji elektronicznej z innych państw członkowskich w polskich e-usługach. System nie prowadzi rejestru publicznego. Główne grupy funkcjonalności obejmują: • obsługę żądań identyfikacji transgranicznej, • przekazywanie danych identyfikacyjnych pomiędzy notyfikowanymi systemami identyfikacji elektronicznej państw członkowskich, • zapewnienie interoperacyjności krajowych infrastruktur identyfikacji elektronicznej, • obsługę zakresu danych obowiązkowych (imię, nazwisko, identyfikator, data urodzenia) wymaganego w ramach uwierzytelniania transgranicznego, • realizację obowiązków wynikających z rozporządzenia eIDAS (electronic IDentification, Authentication and trust Services). System integruje się z krajowymi systemami teleinformatycznymi (w tym z Węzłem Krajowym) oraz z zagranicznymi systemami identyfikacji		

Lp.	Nazwa systemu	Gestor systemu	Opis systemu	Status	Krótki opis ewentualnej zmiany
			elektronicznej państw członkowskich Unii Europejskiej.		
20	WIM System RCMC	Wojskowy Instytut Medyczny – Państwowy Instytut Badawczy	WIM System RCMC (System Regionalnego Centrum Medycyny Cyfrowej Wojskowego Instytutu Medycznego – Państwowego Instytutu Badawczego) to system wspierający gromadzenie, przetwarzanie i wymianę danych medycznych i badawczych w modelu współpracy międzyośrodkowej. Celem systemu jest umożliwienie współdzielenia danych medycznych i badawczych w ramach środowiska Regionalnego Centrum Medycyny Cyfrowej. System nie prowadzi rejestru publicznego. Główne grupy funkcjonalności obejmują: • gromadzenie danych medycznych i badawczych, • przetwarzanie danych w modelu współdzielenia międzyośrodkowego, • obsługę wymiany metadanych i danych zanonimizowanych, • udostępnianie uzgodnionych zbiorów współdzielonych, • wspieranie współpracy międzyinstytucjonalnej w obszarze medycyny cyfrowej. System integruje się z innymi systemami informatycznymi w obszarze ochrony zdrowia oraz z systemami	Planowany	

Lp.	Nazwa systemu	Gestor systemu	Opis systemu	Status	Krótki opis ewentualnej zmiany
			badawczymi i klinicznymi ośrodków współpracujących w ramach Regionalnego Centrum Medycyny Cyfrowej.		

Lista przepływów

Lp.	System źródłowy	System docelowy	Zakres wymienianych danych	Sposób wymiany danych	Typ modyfikacji	Typ interfejsu
1	CTMS UCWBK	System CDD	Dane operacyjne badań klinicznych, harmonogramy, statusy badań, identyfikatory oraz dane administracyjne	Integracja usługowa oraz procesy ETL	Krytyczne dla sukcesu projektu	JSON, XML
2	System CDD	dane.gov.pl	Metadane zbiorów danych, opisy zasobów oraz informacje o dostępności	Publikacja metadanych i integracja usługowa	Krytyczne dla realizacji celów projektu	JSON, XML
3	dane.gov.pl	System CDD	Informacje o publikacji, identyfikatory oraz statusy synchronizacji	Integracja usługowa i synchronizacja zwrotna	Istotne dla realizacji projektu	JSON, XML
4	System CDD	eCRF ABM	Dane raportowe, zestawienia badawcze oraz metadane projektów	Dane raportowe, zestawienia badawcze oraz metadane projektów	Istotne dla współpracy zewnętrznej	JSON, XML
5	eCRF ABM	System CDD	Dane referencyjne projektów, identyfikatory oraz informacje o	Dane referencyjne projektów, identyfikatory oraz informacje o	Integracja usługowa oraz import danych	JSON, XML

Lp.	System źródłowy	System docelowy	Zakres wymienianych danych	Sposób wymiany danych	Typ modyfikacji	Typ interfejsu
			statusach	statusach		
6	eCRF UCWBK	System CDD	Ustrukturyzowane dane badawcze, dane wizyt, pomiary, wyniki obserwacji i parametry kliniczne	Integracja usługowa oraz cykliczne przetwarzanie danych	Krytyczne dla sukcesu projektu	JSON, XML
7	EDM	System CDD	Indeksy elektronicznej dokumentacji medycznej, identyfikatory dokumentów, metadane oraz informacje o lokalizacji dokumentów w repozytoriach źródłowych	Integracja usługowa z wykorzystaniem standardów krajowych	Istotne dla interoperacyjności krajowej	XML, JSON
8	System CDD	EDM	Informacje zwrotne o wykorzystaniu wskazanych zasobów oraz dane referencyjne uzupełniające indeks elektronicznej dokumentacji medycznej	Integracja usługowa z wykorzystaniem standardów krajowych	Istotne dla interoperacyjności krajowej	XML, JSON
9	System CDD	eKZ	Informacje zwrotne dotyczące wykorzystania danych w analizach naukowych oraz dane referencyjne wspierające standaryzację dokumentów	Integracja usługowa z wykorzystaniem standardów krajowych	Istotne dla interoperacyjności krajowej	XML, JSON
10	eKZ	System	Zanonimizowa	Integracja	Istotne dla	XML, JSON

Lp.	System źródłowy	System docelowy	Zakres wymienianych danych	Sposób wymiany danych	Typ modyfikacji	Typ interfejsu
		CDD	ne dane statystyczne i epidemiologiczne dotyczące urodzeń i zgonów, dane referencyjne oraz słowniki na potrzeby badań naukowych	usługowa z wykorzystaniem standardów krajowych	interoperacyjności krajowej	
11	System CDD	e-learning WUM	Informacje zwrotne dotyczące wykorzystania zasobów dydaktycznych oraz dane referencyjne wspierające standaryzację, kategoryzację i opis materiałów edukacyjnych	Integracja usługowa i synchronizacja zwrotna	Istotne dla realizacji projektu	JSON, XML
12	e-learning WUM	System CDD	Metadane materiałów dydaktycznych, opisy zasobów edukacyjnych, klasyfikacje treści oraz identyfikatory obiektów dydaktycznych przeznaczonych do dalszego wykorzystania	Integracja usługowa i synchronizacja zwrotna	Istotne dla realizacji projektu	JSON, XML
13	eTMF UCWBK	System CDD	Metadane dokumentów badania, identyfikatory, statusy oraz informacje formalne	Integracja usługowa oraz wymiana metadanych	Istotne dla realizacji projektu	JSON, XML, PDF/A
14	System CDD	HealthData@EU	Metadane zbiorów	Integracja usługowa	Rozwojowe, strategiczne	JSON, XML

Lp.	System źródłowy	System docelowy	Zakres wymienianych danych	Sposób wymiany danych	Typ modyfikacji	Typ interfejsu
			danych oraz informacje o dostępności i warunkach wykorzystania	zgodna z EHDS		
15	HealthData@EU	System CDD	Informacje o zasobach, zapytania oraz dane katalogowe	Integracja usługowa zgodna z EHDS	Rozwojowe, strategiczne	JSON, XML
16	openCARD IO	System CDD	Zanonimizowane dane medyczne, dane badawcze, metadane zbiorów oraz dane historyczne	Integracja usługowa oraz import danych	Istotne dla realizacji projektu	JSON, XML, CSV
17	PPM	System CDD	Identyfikatory zasobów, statusy publikacji oraz informacje zwrotne	Integracja usługowa i synchronizacja metadanych	Istotne dla sukcesu projektu	JSON, XML
18	System CDD	PPM	Metadane publikacji, datasetów, opisy zasobów oraz identyfikatory	Integracja usługowa i synchronizacja zwrotna	Krytyczne dla sukcesu projektu	JSON, XML
19	PPM WUM	System CDD	Identyfikatory zasobów, statusy publikacji i datasetów	Integracja usługowa i synchronizacja zwrotna	Istotne dla realizacji projektu	JSON, XML
20	System CDD	PPM WUM	Metadane publikacji, datasetów, opisy zasobów naukowych oraz identyfikatory obiektów	Integracja usługowa i synchronizacja metadanych	Krytyczne dla sukcesu projektu	JSON, XML
21	REDCap	System CDD	Ustrukturyzowane dane badawcze,	Integracja usługowa, eksport	Krytyczne dla sukcesu projektu	JSON, XML, CSV

Lp.	System źródłowy	System docelowy	Zakres wymienianych danych	Sposób wymiany danych	Typ modyfikacji	Typ interfejsu
			metadane formularzy, identyfikatory rekordów oraz informacje o statusie danych	okresowy i synchronizacja danych		
22	System CDD	REDCap	Metadane zbiorów badawczych, identyfikatory projektów, opisy formularzy oraz informacje o zakresie danych przeznaczonych do dalszego wykorzystania	Integracja usługowa i wymiana metadanych badawczych	Istotne dla realizacji projektu	Istotne dla realizacji projektu JSON, XML
23	System CDD	SGZ	Dane zagregowane, wskaźniki oraz informacje zwrotne dotyczące integracji w obszarze zdarzeń medycznych	???????????????? ???????????????? ?	Istotne dla interoperacyjności krajowej	HL7 CDA, HL7 FHIR
24	SGZ	System CDD	Zanonimizowane dane o zdarzeniach medycznych, identyfikatory, dane referencyjne oraz słowniki	Integracja usługowa z wykorzystaniem standardów krajowych	Istotne dla interoperacyjności krajowej	HL7 CDA, HL7 FHIR
25	SSOZ	System CDD	Dane referencyjne, słowniki oraz informacje zwrotne	Integracja usługowa	Istotne dla interoperacyjności krajowej	XML, JSON
26	System CDD	SSOZ	Dane zagregowane, wskaźniki oraz informacje	Integracja usługowa oraz przekazywanie danych	Istotne dla interoperacyjności krajowej	XML, JSON

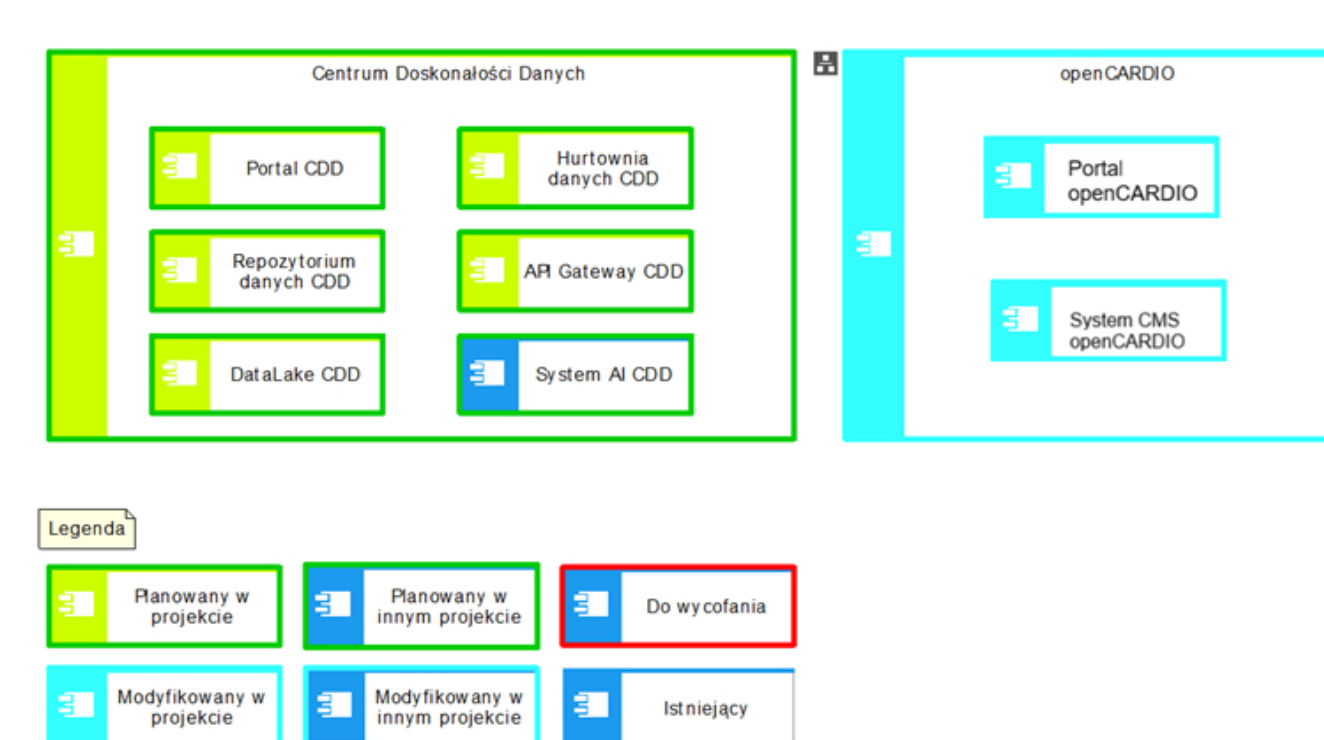
Lp.	System źródłowy	System docelowy	Zakres wymienianych danych	Sposób wymiany danych	Typ modyfikacji	Typ interfejsu
			statystyczne	raportowych		
27	System CDD	WIM System RCMC	Wybrane zbiory danych badawczych i medycznych, metadane oraz dane zanonimizowane	Integracja usługowa w uzgodnionym zakresie	Rozwojowe, docelowe	JSON, XML
28	Węzeł Krajowy	System CDD	Dane i atrybuty uwierzytelniające użytkownika oraz identyfikatory logowania	Integracja usługowa w zakresie uwierzytelniania i identyfikatory logowania	Krytyczne dla dostępu do usług	SAML, OpenID Connect
29	WIM System RCMC	System CDD	Dane zwrotne, metadane oraz informacje referencyjne	Integracja usługowa w uzgodnionym zakresie	Rozwojowe, docelowe	JSON, XML
30	System mObywatel	System CDD	Dane i atrybuty uwierzytelniające użytkownika oraz potwierdzenie tożsamości	Integracja usługowa w zakresie uwierzytelniania	Istotne dla zapewnienia dostępu	OAuth 2.0, OpenID Connect
31	Węzeł Krajowy	Węzeł transgraniczny	Obsługa żądania identyfikacji (z UE do PL). Dane identyfikujące osobę fizyczną w kontekście środka identyfikacji elektronicznej przekazywane za pośrednictwem Węzła Krajowego i dalej Węzła Transgranicznego do systemów zagranicznych	Integracja usługowa w zakresie uwierzytelniania	Istotne dla zapewnienia dostępu	SOAP, REST

Lp.	System źródłowy	System docelowy	Zakres wymienianych danych	Sposób wymiany danych	Typ modyfikacji	Typ interfejsu
			umożliwiających korzystanie ze środków identyfikacji elektronicznej przyłączonych do Węzła Krajowego. Zakres danych zależy od danych w środku identyfikacji elektronicznej. Aktualnie Węzeł krajowy i transgraniczny przekazuje tylko zakres danych obowiązkowych: * Imię * Nazwisko * Identyfikator * Data urodzenia Dane mogą być wykorzystywane biznesowo, np. przez zagraniczne systemy udostępniające usługi online.			
32	Węzeł transgraniczny	Węzeł Krajowy	Obsługa żądania identyfikacji (z PL do UE). Dane identyfikujące osobę fizyczną (w szczególności cudzoziemca), w kontekście	Integracja usługowa w zakresie uwierzytelniania	Istotne dla zapewnienia dostępu	SOAP, REST

Lp.	System źródłowy	System docelowy	Zakres wymienianych danych	Sposób wymiany danych	Typ modyfikacji	Typ interfejsu
			środku identyfikacji elektronicznej wystawionego przez system podłączony do węzła transgranicznego (w szczególności zagraniczny system identyfikacji elektronicznej) przekazywane za pośrednictwem Węzła Transgranicznego przez Węzeł Krajowy do krajowych systemów udostępniających usługi on-line. Zakres danych zależy od danych w zagranicznym środku identyfikacji elektronicznej. Aktualnie Węzeł transgraniczny przekazuje tylko zakres danych obowiązkowych: * Imię * Nazwisko * Identyfikator * Data urodzenia Mogą być wykorzystywane biznesowo,			

Lp.	System źródłowy	System docelowy	Zakres wymienianych danych	Sposób wymiany danych	Typ modyfikacji	Typ interfejsu
			np. w systemach udostępniających usługi online, np. do wypełnienia wniosku.			

7.2. Kluczowe komponenty architektury rozwiązania



7.3. Przyjęte założenia technologiczne

Lp.	Obszar	Założenie technologiczne
1.	Infrastruktura	Architektura wdrażana w modelu on-premise w oparciu o infrastrukturę WUM, z możliwością rozszerzenia do modelu hybrydowego. Wykorzystanie istniejących serwerowni WUM, sieci 10 Gb/s, istniejącej instancji DSpace/PPM oraz istniejących systemów klinicznych i badawczych. Zakup i wdrożenie brakujących elementów infrastrukturalnych niezbędnych do uruchomienia Systemu CDD, z wydzieleniem środowisk DEV/TEST/PROD.

Lp.	Obszar	Założenie technologiczne
2.	Sieć i bezpieczeństwo	Segmentacja sieci na strefy przetwarzania danych, publikacji i dostępu kontrolowanego. RBAC jako podstawowy model nadawania uprawnień z możliwością rozszerzenia o reguły atrybutowe. Keycloak jako centralny moduł IAM dla uwierzytelniania, autoryzacji, SSO oraz OAuth2/OIDC. Moduł DAC jako warstwa decyzji o dostępie do danych wrażliwych i zasobów controlled access. HashiCorp Vault do bezpiecznego przechowywania sekretów, TLS 1.3, WAF dla portalu publicznego, SIEM + ELK, audit logi, testy DR oraz zgodność organizacyjna z SZBI opartym o ISO 27001 i wymaganiami KSC (Krajowy System Cyberbezpieczeństwa)/NIS2.
3.	Standardy wymiany danych	Integracja i wymiana danych realizowana z użyciem HL7 FHIR R5 jako standardu preferowanego dla warstwy klinicznej, HL7 CDA dla dokumentów EDM, DICOM dla obrazów medycznych oraz JSON, XML, CSV i PDF/A dla danych badawczych, metadanych i dokumentów. Publikacja metadanych zgodna z DCAT-AP PL 1.2. Interfejsy API dokumentowane w OpenAPI 3.1. Mechanizmy OAI-PMH dla harvesting'u metadanych. Nadawanie trwałych identyfikatorów DOI przez integrację z DataCite.
4.	Systemy operacyjne serwerowe	Podstawową platformą serwerową są systemy Linux klasy enterprise lub community enterprise, w szczególności Debian, Ubuntu Server, Rocky Linux lub RHEL, wykorzystywane dla warstw backendowych, ETL/ELT, repozytoriów, AI i usług integracyjnych. Dopuszcza się użycie Windows Server dla komponentów wymagających technologii Microsoft lub integracji z istniejącymi usługami domenowymi i narzędziami administracyjnymi. Systemy objęte hardeningiem i politykami bezpieczeństwa.
5.	Bazy danych	Rdzeń danych Systemu CDD obejmuje trzy warstwy: Data Lake jako pierwszą warstwę gromadzenia danych badawczo-medyczno-dydaktycznych po pozyskaniu i anonimizacji, hurtownię danych jako drugą warstwę przechowującą dane ustrukturyzowane, zwalidowane jakościowo i zharmonizowane do wspólnych modeli danych, oraz bazę danych zweryfikowanych klinicznie jako trzecią warstwę przeznaczoną dla danych sprawdzonych ręcznie i dopuszczonych do wykorzystania przez klinicystów oraz usługi kliniczne. PostgreSQL 15+ HA stanowi podstawową relacyjną bazę danych dla rdzenia systemu i hurtowni danych, a Elasticsearch warstwę indeksowania i wyszukiwania.
6.	Serwery aplikacji	Warstwa serwerowa obejmuje serwery aplikacyjne, serwer bazy danych PostgreSQL HA, klastrer Elasticsearch, storage obiektowy S3-compatible oparty o MinIO lub Ceph, NAS archiwalny oraz serwery OCR i przetwarzania. Na tej infrastrukturze utrzymywane są moduły integracji, przetwarzania, repozytoria, usługi FHIR oraz komponenty analityczne i pomocnicze Systemu CDD.
7.	Portale	Warstwa portalowa obejmuje wspólny portal i API Gateway jako punkt ekspozycji danych, metadanych i usług CDD dla użytkowników końcowych, badaczy i systemów zewnętrznych. Frontend zakładany jest w technologii React/Next.js, backend w

Lp.	Obszar	Założenie technologiczne
		FastAPI/Python, a interfejsy API są dokumentowane w OpenAPI 3.1 i udostępniane przez Swagger UI. W tej kategorii mieszczą się także repozytorium danych badawczych, repozytorium publikacyjne rozwijane na bazie istniejącej instancji DSpace/PPM WUM oraz repozytorium obrazów, ponieważ zgodnie z architekturą stanowią one warstwy udostępniania zasobów i prezentacji ich użytkownikom.
8.	Inne	Pozostałe komponenty obejmują moduł integracji, ETL/ELT i anonimizacji oparty na Apache Airflow 2.x oraz adapterach FHIR R5, pipeline pseudonimizacji, tokenizacji, izolacji PHI i anonimizacji danych, Metadata Engine DCAT-AP PL 1.2, moduł DOI / DataCite, moduł udostępniania klinicznego i FHIR oraz warstwę AI i analityki danych wykorzystującą narzędzia takie jak TensorFlow, PyTorch, scikit-learn lub równoważne. Ta kategoria obejmuje komponenty przekrojowe, które nie są samodzielną infrastrukturą, bazą ani portalem, ale spinają architekturę funkcjonalnie.

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