

OPIS ZAŁOŻEŃ PROJEKTU INFORMATYCZNEGO

Tytuł projektu	Genomowa mapa polski w otwartym dostępie - digitalizacja zasobów biomolekularnych pracowni Blobank UŁ		
Wnioskodawca	Minister Nauki i Szkolnictwa Wyższego		
Beneficjent	Uniwersytet Łódzki, Wydział Biologii i Ochrony Środowiska		
Partnerzy			
Źródło finansowania	Centrum Projektów Polska Cyfrowa, Poddziałanie 2.3.1 „Cyfrowe udostępnienie informacji sektora publicznego ze źródeł administracyjnych i zasobów nauki”. Budżet Państwa część 27.		
Całkowity koszt projektu	10 000 000,00 zł		
Planowany okres realizacji projektu	08-2020 do 01-2023		
Osoba kontaktowa	Błażej Marciniak	blazej.marciniak@biol.uni.lodz.pl	600936417

1. POWODY PODJĘCIA PROJEKTU

1.1. Identyfikacja problemu i potrzeb

Od momentu zakończenia HGP w naukowej przestrzeni publicznej dostępne są dane na temat pełnych genomów ludzkich np. 1000 Genome Project. Zróżnicowanie genetyczne i fenotypowe występujące pomiędzy populacjami sprawia, że w wielu badaniach dostęp do referencyjnej grupy genomów lokalnych jest wymagany. Ustalenie częstych wariantów charakterystycznych dla Polaków, pomocne będzie w interpretacji polimorfizmów w kontekście medycyny spersonalizowanej oraz badań populacyjnych.

Kolejne kraje realizują projekty mające na celu poznanie genomów swoich obywateli (m.in. Genomics England, Estonian Genome Project, FinnGen). Jedyną próbą jaka została podjęta i zrealizowana w Polsce to przedsięwzięcie Polgenom – repozytorium genomów ok. 100 Polaków. Dostęp do danych jest płatny, zarówno dla naukowców jak i przedsiębiorstw.

We wszystkich wymienionych projektach zdigitalizowane dane genetyczne zostały udostępnione dla nauki w zależności od przyjętej polityki danego państwa. Regulacje te wahają się od bardzo liberalnych w przypadku 1000 Genome Project (dane są w pełni publiczne) do bardziej restrykcyjnych, w przypadku polityki projektu FinnGen (dane genetyczne obywateli nie mogą opuszczać kraju). W Polsce przechowywanie i udostępnianie danych genetycznych nie jest jeszcze uregulowane prawnie.

Istniejące systemy archiwizacji danych genetycznych (np. EGA) nie gwarantują, że standardy ochrony informacji i ich dostępu będą zgodne z przyjętymi w przyszłości krajowymi standardami legislacyjnymi. Powstaje konieczność utworzenia krajowego repozytorium danych gnomicznych, które będzie mogło być na bieżąco modyfikowane zgodnie z nowymi regulacjami.

Rozwiązaniem może być proponowana lokalna wersja European Genome Archive, w której dane będą przechowywane w Kronik@. Realizacja projektu wytworzy referencyjne dane genomiczne dla polskiej populacji do których dostęp będzie bezpłatny i licencjonowany oraz narzędzie do

bezpiecznego przechowywania tych informacji i kontroli dostępu.

Interesariusz	Zidentyfikowany problem	Szacowana wielkość grupy
Pracownicy Naukowi	<ul style="list-style-type: none"> - Brak dostępnych w otwartym dostępie/bezpłatnie wysokiej jakości danych genomicznych umożliwiających prowadzenie badań podstawowych w dziedzinie Genetyki populacji Polskiej lub medycynie spersonalizowanej - Niewystarczająca ilość ośrodków wyposażonych w odpowiednią infrastrukturę i kompetencje umożliwiające digitalizację danych genetycznych człowieka w Polsce (przekładający się na terminy i koszty usług) – konieczność posilkowania się usługami za granicą - niewystarczająca ilość centrów eksperckich wspierających procesy digitalizacji tego typu danych w Polsce - brak odpowiednio zabezpieczonych publicznie dostępnych przestrzeni składowania wrażliwych danych osobowych 	105284 (system POLON).
Personel B+R	<ul style="list-style-type: none"> - Niewystarczająca ilość ośrodków wyposażonych w odpowiednią infrastrukturę i kompetencje umożliwiające digitalizację danych genomicznych człowieka w Polsce (przekładający się na terminy i koszty usług) – konieczność posilkowania się usługami za granicą - Niewystarczająca ilość centrów eksperckich wspierających procesy digitalizacji tego typu danych w Polsce - Brak dostępnych w otwartym dostępie/darmowych wysokiej jakości danych umożliwiających prowadzenie badań 	239 300 (Działalność badawcza i rozwojowa w Polsce w 2017 roku - GUS)
Lekarze	<ul style="list-style-type: none"> - Niewystarczająca ilość ośrodków wyposażonych w 	148 903 (Zestawienie liczbowe lekarzy i lekarzy

Interesariusz	Zidentyfikowany problem	Szacowana wielkość grupy
	odpowiednią infrastrukturę i kompetencje umożliwiające digitalizację danych genetycznych człowieka w Polsce (przekładający się na terminy i koszty usług) – konieczność posilkowania się usługami za granicą - Brak dostępnych w otwartym dostępie/darmowych wysokiej jakości danych umożliwiających prowadzenie badań	dentystów wg przynależności do okręgowej izby lekarskiej i tytułu zawodowego - Naczelna Izba Lekarska)
Przemysł farmaceutyczny	- badania wykazały, że niektóre nowoczesne leki działają skuteczniej w grupach/populacjach pacjentów posiadających określone polimorfizmy – brak danych służących do budowy modeli oddziaływania leków w populacji Polskiej. - Niewystarczająca ilość ośrodków wyposażonych w odpowiednią infrastrukturę i kompetencje umożliwiające digitalizację danych genetycznych człowieka w Polsce (przekładający się na terminy i koszty usług) – konieczność posilkowania się usługami za granicą - niewystarczająca ilość centrów eksperckich wspierających procesy digitalizacji tego typu danych w Polsce	313 podmiotów (Rocznik statystyczny przemysłu - GUS) Zatrudnienie pośrednio Ok. 100 000 osób (https://www.producencilekow.pl/krajowy-przemysl-farmaceutyczny/)

1.2. Opis stanu obecnego

Najważniejsze zasoby pracowni Biobank to:

1. Kolekcja POPULOUS (10000 próbek mat. Biol. od dawców z terenu całej Polski). Częściowo zdigitalizowana (dla 6000 dawców uzyskano cyfrowe zapisy dotyczące 0,5 mln. polimorfizmów - dane mikromacierzowe).
2. Infrastruktura Techniczno-Informatyczna umożliwiającą digitalizację DNA i dalsze analizy.
3. Wysoko wykwalifikowany i doświadczony personel projektu mający za sobą uzyskanie dziesiątek tysięcy cyfrowych zapisów DNA.

Wymienione wyżej czynniki stawiają Beneficjenta w absolutnej czołówce tego typu instytucji w

Polsce. O pozycji Beneficjenta może świadczyć ilość zrealizowanych projektów zarówno naukowych (3 duże projekty na łączną kwotę ok. 15 mln złotych) jak i komercyjnych (43 umowy na wykonanie usług naukowo-badawczych z zakresu genomiki i NGS, na łączną kwotę ok. 1,7 mln złotych). Od początku swojego istnienia (2014) Pracownia Biobank nastawiona jest na współpracę i dzielenie się swoimi zasobami z partnerami naukowymi. Pracownia Biobank UŁ, jest członkiem założycielem Polskiej Sieci Biobanków, członkiem Konsorcjum BBMRI.pl i sieci Biobanków ESBB, BCNet.

Współczesne analizy DNA wykonuje się jedynie w oparciu o wersje cyfrowe – uzyskanie tych danych jest podstawą każdego projektu w tej dziedzinie. Z uwagi na znaczny koszt uzyskania pełnych genomów/WES (whole exome sequencing) do badań z reguły wykorzystuje się fragmenty Genomów i na takich danych opiera się większość prac Beneficjenta. Niestety w pewnych obszarach i precyzyjnych analizach jest to często niewystarczające.

Beneficjent i tak jest w szczęśliwym położeniu, iż posiada środki techniczne i kompetencje umożliwiające digitalizację DNA. Nie wszystkie ośrodki w kraju mogą pochwalić się takim możliwościami a brak wystarczającej infrastruktury w Kraju zmusza je do korzystania, przy przetwarzaniu wrażliwych danych dot. mieszkańców Polski, z usług zagranicznych kontrahentów (często również po za granicami Unii Europejskiej).

2. EFEKTY PROJEKTU

2.1. Cele i korzyści wynikające z projektu

Cel - 1	Digitalizacja zasobów nauki
Cel strategiczny	Wzmocnienie cyfrowych fundamentów dla rozwoju kraju poprzez digitalizację i udostępnienie on-line zasobów - Program Operacyjny Polska Cyfrowa Budowa gospodarki opartej na danych – założeniami do strategii rozwoju AI w Polsce Rozwój kompetencji biotechnologicznych – Regionalna Strategia Innowacji dla Województwa Łódzkiego „LORIS 2030”, "Strategia Rozwoju Polski Centralnej do roku 2020 z perspektywą 2030"
Korzyść:	Zwiększenie potencjału badawczego polskich uczelni, obniżenie kosztów prowadzenia prac badawczo rozwojowych, badań podstawowych, możliwość prowadzenia badań nad skutecznością leków w odniesieniu do populacji polskiej, rozwój biotechnologii
KPI:	Rozmiar zdigitalizowanych danych
Wartość aktualna i docelowa KPI:	0 TB 15 TB
Metoda pomiaru KPI	Metoda: Zsumowanie objętości wszystkich udostępnionych zasobów wyrażonej w TB, Sposób pomiaru: narzędzie systemowe szacujące objętość plików np. polecenie dh lub protokoły odbioru produktów projektu zaakceptowane przez Kierownika projektu. Częstotliwość pomiaru zgodnie z harmonogramem raportów okresowych - zgodnie z DIP
Cel - 2	Utworzenie infrastruktury IT pozwalającej na udostępnienie danych o Ludzkim

	DNA z terytorium RP w bezpieczny sposób
Cel strategiczny	Wzmocnienie cyfrowych fundamentów dla rozwoju kraju poprzez digitalizację i udostępnienie on-line zasobów - Program Operacyjny Polska Cyfrowa Budowa gospodarki opartej na danych – założeniami do strategii rozwoju AI w Polsce
Korzyść:	Budowa gospodarki opartej na wiedzy
KPI:	Liczba podmiotów, które udostępniły on-line informacje sektora publicznego
Wartość aktualna i docelowa KPI:	0 szt. 1 szt.
Metoda pomiaru KPI	metoda pomiaru: Oszacowanie ilości jednostek naukowych, które wytworzyły infrastrukturę umożliwiającą digitalizację zasobów. sposób pomiaru: Potwierdzenie odbioru i przyjęcie do eksploatacji przez kierownika projektu infrastruktury służącej udostępnianiu danych
Cel - 3	Udostępnienie wysokiej jakości danych cyfrowych na potrzeby polskiej nauki i sektora B+R
Cel strategiczny	Wzmocnienie cyfrowych fundamentów dla rozwoju kraju poprzez digitalizację i udostępnienie on-line zasobów - Program Operacyjny Polska Cyfrowa Budowa gospodarki opartej na danych – założeniami do strategii rozwoju AI w Polsce Rozwój kompetencji biotechnologicznych – Regionalna Strategia Innowacji dla Województwa Łódzkiego „LORIS 2030”, "Strategia Rozwoju Polski Centralnej do roku 2020 z perspektywą 2030"
Korzyść:	Budowa gospodarki opartej na wiedzy, możliwość doboru skuteczniejszych leków/terapii, obniżenie kosztów prac B+R.
KPI:	Liczba pobrań/odtworzeń dokumentów zawierających informacje sektora publicznego w ciągu 12 miesięcy od daty zakończenia realizacji projektu
Wartość aktualna i docelowa KPI:	0 szt. 3750 szt.
Metoda pomiaru KPI	metoda pomiaru: Statystyki odwiedzin strony/pobrań dokumentów sposób pomiaru: wybrane narzędzie do zliczania statystyk odwiedzin strony np. Google Analytics. Częstotliwość pomiaru zgodnie z harmonogramem raportów okresowych zgodnie z DIP

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)
Pełne genomy	31-01-2023	300
Egzomy	31-01-2023	450

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

2.4. Produkty końcowe projektu

Nazwa produktu	Planowana data wdrożenia
Infrastruktura IT - Repozytorium danych naukowych LEGA	01-2022
zdigitalizowane zasoby nauki	05-2022
Infrastruktura digitalizacji ludzkiego DNA	02-2021

3. KAMIENIE MIŁOWE

Kamienie milowe	Planowany termin osiągnięcia
Zakupienie infrastruktury niezbędnej do procesu digitalizacji.	2021-01-31
Zakup materiałów niezbędnych w procesie digitalizacji	2022-01-31
Szkolenie personelu laboratoryjnego	2021-01-31
Digitalizacja pełnych genomów - stan 10%	2021-03-31
Digitalizacja pełnych genomów - stan 40%	2021-06-30
Digitalizacja pełnych genomów - stan 70%	2021-11-30
Digitalizacja pełnych genomów - stan 100%	2022-04-30
Digitalizacja egzomów - stan 10%	2021-03-30
Digitalizacja egzomów - stan 40%	2021-06-30
Digitalizacja egzomów - stan 70%	2021-11-30
Digitalizacja egzomów - stan 100%	2022-04-30
Utworzenie ścieżek bioinformatycznych	2021-02-28
Walidacja ścieżek bioinformatycznych	2021-05-31
Realizacja ścieżek bioinformatycznych	2022-12-31

Kamienie milowe	Planowany termin osiągnięcia
Przygotowanie danych do udostępnienia	2022-12-31
Udostępnienie danych	2023-01-31
Konfiguracja infrastruktury IT	2021-01-31
Instalacja i uruchomienie LEGA	2021-07-31
Testy i stabilizacja systemu	2021-09-30
Uruchomienie produkcyjne i oddanie do eksploatacji	2023-01-31

4. KOSZTY

4.1. Koszty ogólne projektu wraz ze sposobem finansowania

Całkowity koszt projektu (netto oraz brutto), w tym	Netto 8 363 873,49 zł Brutto 10 000 000,00 zł	
Procent dofinansowania ze środków UE (brutto)	85%	
Procent środków z budżetu państwa (brutto)	15%	
Podział całkowitego kosztu projektu na poszczególne lata (netto oraz brutto)	2020	Netto 5 527 857,93 zł Brutto 6 745 175,00 zł
	2021	Netto 2 306 777,15 zł Brutto 2 701 330,00 zł
	2022	Netto 475 418,41 zł Brutto 499 675,00 zł
	2023	Netto 53 820,00 zł Brutto 53 820,00 zł

4.2. Wykaz poszczególnych pozycji kosztowych

Nazwa pozycji kosztowej		Przewidywany koszt brutto	Uzasadnienie pozycji kosztowej (przeznaczenie)
Oprogramowanie	Oprogramowanie	0,00 zł	Beneficjent zakłada wykorzystanie oprogramowania dostępnego na licencji otwartej
Infrastruktura	Urządzenia (6 200 000,00) i Materiały (2 400 000,00) niezbędne do	8 600 000,00 zł	Aktualnie posiadana przez Beneficjenta infrastruktura sprzętowa jest wykorzystywana niemal 90% co oznacza brak

Nazwa pozycji kosztowej		Przewidywany koszt brutto	Uzasadnienie pozycji kosztowej (przeznaczenie)
	przeprowadzenia procesu digitalizacji		możliwości realizacji tak ambitnego projektu bez rozbudowy o nowy sprzęt – sekwenator nowej generacji
Koszty UX i grafiki	Koszty UX i grafiki	20 000,00 zł	Koszty Audytów UX i audytów WCAG 2.0
Bezpieczeństwo	Audyty bezpieczeństwa informacji	60 000,00 zł	Audyt bezpieczeństwa zaimplementowanych rozwiązań IT
Wydajność rozwiązań			
Szkolenia	Szkolenia, kursy, warsztaty	19 720,00 zł	Szkolenia dla personelu w celu niwelacji zdiagnozowanych w trakcie realizacji projektu luk kompetencyjnych.
Działania informacyjno-promocyjne	Promocja projektu	50 000,00 zł	Udział w konferencjach, organizacja własnych konferencji, hackathonów, udział w festiwalach nauki, wytwarzanie materiałów promocyjnych.
Koszty zarządzania i wsparcia (w tym wynagrodzenia personelu wspomagającego)	Koszty pośrednie	1 250 280,00 zł	Koszty obsługi biura projektu, personelu administracyjnego oraz pozostałe koszty pośrednie - 163 080; Wynagrodzenie personelu realizującego projekt wraz z kosztami pracodawcy - 1 087 200;

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

Całkowity koszt utrzymania trwałości projektu (brutto)	330 000,00 zł		Źródło finansowania
Podział całkowitego kosztu utrzymania trwałości projektu na poszczególne lata (netto oraz brutto)	2023	70 000,00 zł (brutto) (68 130,00 zł netto)	krajowe środki publiczne - budżet państwa
	2024	80 000,00 zł (brutto) (78 130,00 zł netto)	krajowe środki publiczne - budżet państwa
	2025	80 000,00 zł (brutto) (78 130,00 zł netto)	krajowe środki publiczne - budżet państwa
	2026	100 000,00 zł (brutto) (94 390,25 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
Ryzyka techniczne obejmują spory zakres zagadnień związany z realizacją projektu i są to: Brak możliwości digitalizacji wybranych zasobów z uwagi na stan lub brak adekwatnych narzędzi. Brak adekwatnej/zakup złej infrastruktury informatycznej pozwalającego na udostępnienie we właściwy sposób zinwentaryzowanych	Duża	Znikome	Redukowanie ryzyka – Personel Beneficjenta dokłada wszelkich starań na etapie doboru sprzętu. Tworzenie opisu przedmiotu zamówienia poprzedzane jest dokładnym rozpoznaniem rynku, analizą istniejących technologii oraz wszędzie, gdzie jest to możliwe testami.

Nazwa ryzyka	Siła oddziaływania	Prawdopodobieństwo wystąpienia ryzyka	Sposób zarządzania ryzykiem
zasobów.			
Uszkodzenie infrastruktury świadczenia usług.	Duża	Niskie	Unikanie ryzyka – infrastruktura zgromadzona jest w dedykowanych pomieszczeniach z ograniczonym dostępem. Planowana jest korzystanie z produktów projektu Kronika@, niezależnie zaplanowana jest redundancja sprzętu i rozdzielenie na różne lokalizacje.
Pojawienie się nie zidentyfikowanych wcześniej kosztów niekwalifikowanych.	Średnia	Średnie	Unikanie ryzyka - Beneficjent na etapie przygotowania projektów dokonał szeregu analiz w tym technicznej, inwentaryzacji zasobów sprzętu oraz kompetencji. Co pozwoliło dość dobrze określić zakres niezbędnych do wykonania prac oraz zidentyfikować potrzebne do zakupu zasoby, materiały lub kompetencje. Dodatkowo dokonano rozeznania rynku i na tej podstawie zaplanowano budżet projektu. Jako pracowników Biura Projektu (wsparcia projektu) Beneficjent zamierza zatrudnić wysokiej klasy specjalistów – rozpoczyna się poszukiwanie kandydatów zostaną podpisane umowy warunkowe. W sytuacji gdy mimo podjętych starań pojawią się

Nazwa ryzyka	Siła oddziaływania	Prawdopodobieństwo wystąpienia ryzyka	Sposób zarządzania ryzykiem
			koszty niekwalifikowane Beneficjent pokryje je z środków własnych.
Z uwagi na trwający 30 miesięcy okres realizacji projektu mogą ulec zmianie ceny usług, środków trwałych i materiałów niezbędnych do realizacji projektu.	Mała	Średnie	Redukowanie ryzyka – przy zakupie materiałów i usług strategicznych z punktu widzenia realizacji projektu Beneficjent stosuje umowy gwarantujące stałą cenę dla sukcesywnych dostaw lub dokonuje jednorazowego zakupu.
Zmiana w zakresie rzeczowym, zmiana w harmonogramie.	Średnia	Średnie	Akceptacja ryzyka - Beneficjent zamierza realizować Projekt w oparciu o metodykę PRINCE-2. Ma to zapewnić m.in. regularne planowanie i raportowanie postępów prac – dzięki czemu niezbędne zmiany w zakresie lub harmonogramie projektu będą odpowiednio szybko identyfikowane pozostawiając wystarczającą ilość czasu na odpowiednie wdrożenie.
Przedłużające się procedury wyboru wykonawców/ dostawców.	Średnia	Wysokie	Redukowanie ryzyka – personel Beneficjenta dokłada wszelkich starań na etapie tworzenia SIWZ. Beneficjent planuje zatrudnienie dedykowanego personelu do obsługi administracyjnej projektu. Do dyspozycji zespołu realizującego projekt dostępne jest również Centrum Gospodarcze UŁ - kilkunastoosobowa

Nazwa ryzyka	Siła oddziaływania	Prawdopodobieństwo wystąpienia ryzyka	Sposób zarządzania ryzykiem
			<p>komórka zajmująca się obsługą Beneficjenta od strony zakupów wynikających z PZP</p> <p>Szacunek wartości ryzyka został przeprowadzony na podstawie, doświadczeń beneficjenta w realizacji dużych projektów Unijnych lub Ministerialnych. W kontekście ryzyka „Przedłużające się procedury wyboru wykonawców/ dostawców” warto zauważyć, iż w dużej większości będą to postępowania realizowane na samym początku projektu (Beneficjent rozpocznie pracę nad SIWZ jeszcze przed rozpoczęciem projektu a po uzyskaniu informacji o pozyskaniu informacji o przyznaniu dofinansowania w celu minimalizacji ryzyka.</p>
Obecnie rynek pracy jest dynamiczny i nie można zagwarantować, iż osoby zaangażowane w realizację projektu będą chciały kontynuować współpracę aż do zakończenia projektu.	Średnia	Wysokie	<p>Akceptacja ryzyka – Beneficjent oferuje adekwatne wynagrodzenie na tle sektora oraz możliwość osobistego rozwoju. Możliwość zdobycia doświadczenia i umiejętności pracy w nowoczesnym laboratorium z wykorzystaniem innowacyjnych technik jest niezwykle cenna dla pracowników planujących karierę naukową. Ponad to</p>

Nazwa ryzyka	Siła oddziaływania	Prawdopodobieństwo wystąpienia ryzyka	Sposób zarządzania ryzykiem
			wszystkie procedury badawcze zaakceptowane do stosowania w trakcie procesu digitalizacji są w formie pisemnej - odejście pracownika nie pozbawia Beneficjenta kompetencji w danym zakresie

5.2. Ryzyka wpływające na utrzymanie efektów

Nazwa ryzyka	Siła oddziaływania	Prawdopodobieństwo wystąpienia ryzyka	Sposób zarządzania ryzykiem
Istnieje ryzyko, iż udostępnione zasoby nie spotkają się z należyтым zainteresowaniem ze strony potencjalnych odbiorców. W takiej sytuacji uzyskanie wskaźników realizacji projektu może być zagrożone	Duża	Niskie	Redukcja ryzyka – Beneficjent przewidział udostępnienie popularnonaukowych treści mających na celu przyciągnięcie odbiorców (pasjonatów i hobbystów) do korzystania z produktów projektu. Zaplanowana jest promocja produktów na dziedzinowych konferencjach naukowych.
Ryzyko wycieku przechowywanych danych	Średnia	Średnie	Unikanie - Beneficjent zamierza dołożyć wszelkich starań aby należycie zabezpieczyć dane. Wybrany system informatyczny wymaga szyfrowania danych z wykorzystaniem PKI, komunikacja pomiędzy elementami systemu odbywa się w sposób zaszyfrowany z wykorzystaniem protokołów SLL i SSH, w 2020 po przeniesieniu do nowej siedziby Pracownia

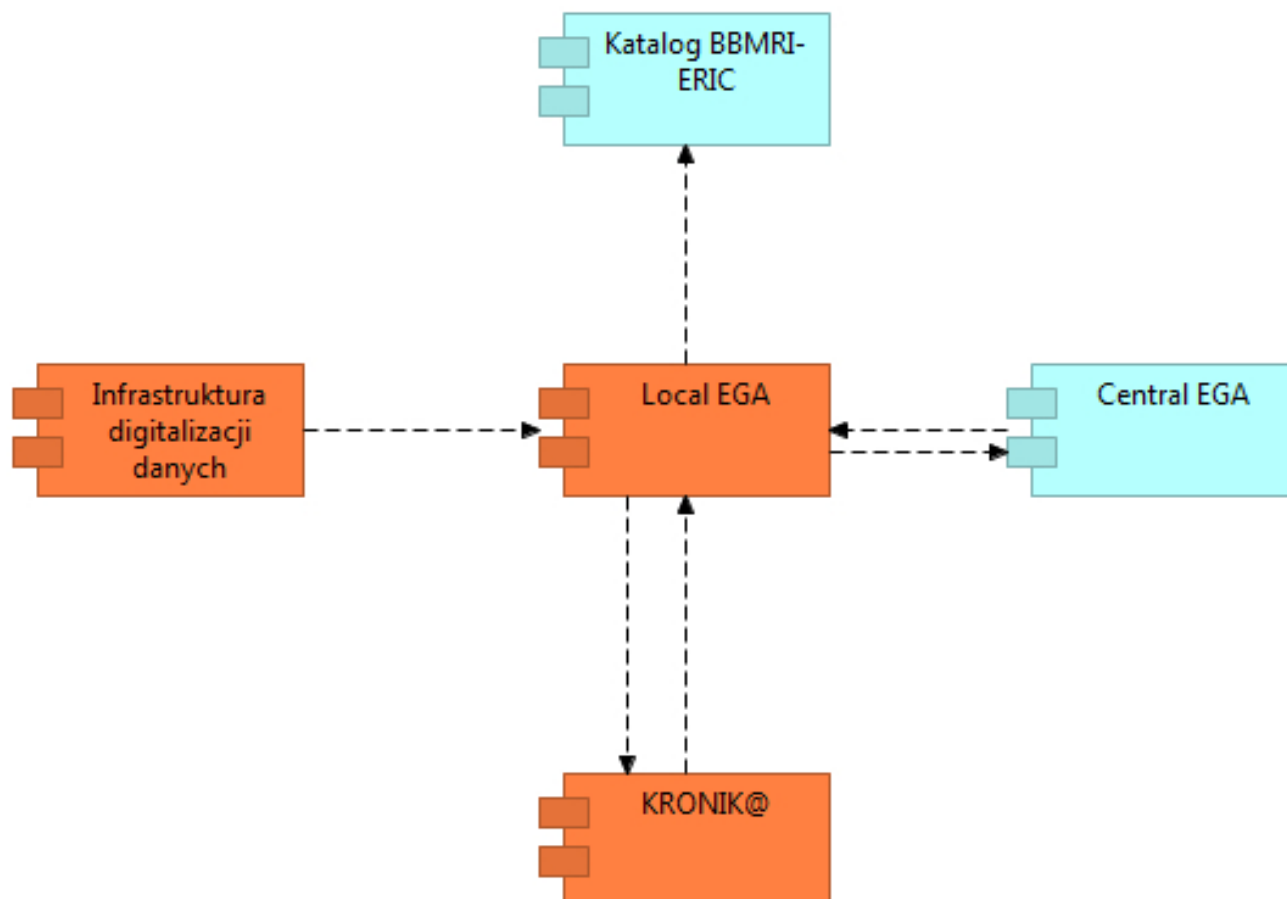
Nazwa ryzyka	Siła oddziaływania	Prawdopodobieństwo wystąpienia ryzyka	Sposób zarządzania ryzykiem
			Biobank (główny realizator projektu) rozpocznie przygotowania do uzyskania certyfikacji w zakresie ISO 27001.

6. OTOCZENIE PRAWNE

Lp.	Tytuł aktu prawnego	Czy wymaga zmian	Opis zmian (jeśli dotyczy)	Etap prac legislacyjnych (jeśli dotyczy)
1	Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2016/679 z dnia 27 kwietnia 2016 r. w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych i w sprawie swobodnego przepływu takich danych oraz uchylenia dyrektywy 95/46/WE (ogólne rozporządzenie o ochronie danych)	TAK/NIE		

7. ARCHITEKTURA

7.1. Widok kooperacji aplikacji



Lista systemów wykorzystywanych w projekcie

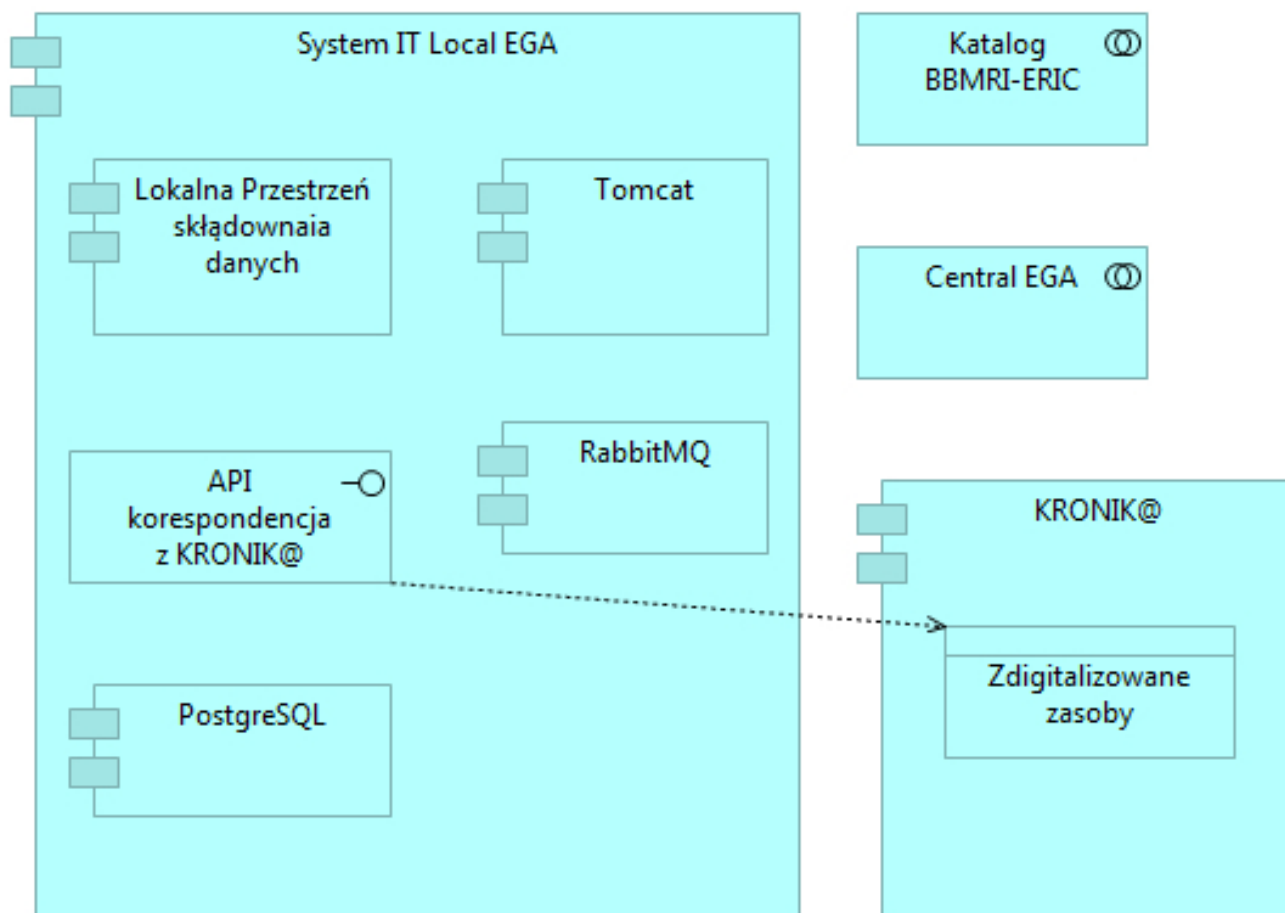
Lp.	Nazwa systemu	Gestor systemu	Opis systemu	Status	Krótki opis ewentualnej zmiany
1	EGA	EBI/CRG Barcelona	European Genome Archive – repozytorium danych genetycznych przechowujących i udostępniających dane dotyczące ludzi.	Istniejący	
2	Directory BBMRI-ERIC	BBMRI-ERIC (Uniwersytet Łódzki jest członkiem Konsorcjum BBMRI.pl)	System opisujący kolekcje materiału biologicznego w standardzie MIABIS 2.0	Istniejący	
3	Local EGA	UŁ	Lokalne repozytorium danych genetycznych połączone z infrastrukturą	Planowany	

Lp.	Nazwa systemu	Gestor systemu	Opis systemu	Status	Krótki opis ewentualnej zmiany
			EGA		
4	Kronik@	Ministerstwo Cyfryzacji	Repozytorium zdigitalizowanych obiektów nauki	Planowany	

Lista przepływów

Lp.	System źródłowy	System docelowy	Zakres wymienianych danych	Sposób wymiany danych	Typ modyfikacji	Typ interfejsu
1	Local EGA	Kronik@	Wszystkie zdigitalizowane zasoby udostępniane przez system	Kopiowanie	Realizowany inną metodą - alternatywna przestrzeń składowania danych	API
2	Kronik@	Local EGA	Wszystkie zdigitalizowane zasoby udostępniane przez system	Tryb odwołań bezpośrednich	Realizowany inną metodą - alternatywna przestrzeń składowania danych	API
3	Local EGA	Katalog BBMRI.pl	Metadane opisujące kolekcje	Kopiowanie danych	realizowalny inną metodą	API
4	EGA	Local EGA	po wcześniejszej pozytywnej weryfikacji celu udostępnienia przez DAC, żądania dostępu do danych	tryb odwołań bezpośrednich	Krytyczny	RabbitMQ

7.2. Kluczowe komponenty architektury rozwiązania



7.3. Przyjęte założenia technologiczne

Lp.	Obszar	Założenie technologiczne
1.	Infrastruktura	Beneficjent zamierza wykorzystać istniejącą lub planowaną do wytworzenia w ramach innych projektów infrastrukturę informatyczną (infrastrukturę informatyczną istniejącą w Biobanku, infrastrukturę informatyczną Central EGA – European Bioinformatics Institute i Center for Genomic Regulations, infrastrukturę projektu KJronik@)
2.	Sieć i bezpieczeństwo	Komunikacja z planowanymi do wdrożenia aplikacjami webowymi odbywać się będzie z wykorzystaniem protokołu SSL lub SSH. Zastosowane rozwiązania Informatyczne przejdą audyt bezpieczeństwa wykonany przez podmiot zewnętrzny.
3.	Standardy wymiany danych	Beneficjent zaimplementuje API zaproponowane przez projektu KRONIK@, w celu komunikacji z Central EGA w celu przesłania Metadanych wykorzystany zostanie broker RabbitMQ
4.	Systemy operacyjne serwerowe	Preferowane do stosowania przez Beneficjenta i twórców rozwiązania są systemy operacyjne klasy Linux.
5.	Bazy danych	Planowane do wdrożenia rozwiązanie wykorzystuje silnik bazy danych PostgreSQL
6.	Serwery aplikacji	LEGA został przygotowany w języku Java i wykorzystuje

Lp.	Obszar	Założenie technologiczne
		framework Spring. Uruchamiany jest na serwerze Tomcat. Wykorzystane technologie są szeroko wykorzystywane w systemach klasy biznes zapewniając bezpieczeństwo danych oraz wydajność aplikacji
7.	Portale	Zaproponowane przez wykonawców systemu IT
8.	Inne	

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

Obecnie nie ma wymogów dotyczących zabezpieczania danych objętych projektem. Nie mniej jednak Beneficjent zamierza wdrożyć Zintegrowany System Zarządzania w oparciu o standardy norm ISO9001 i ISO27001

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