

OPIS ZAŁOŻEŃ PROJEKTU INFORMATYCZNEGO

Tytuł projektu	Cyfrowy Bank Genów Roślin		
Wnioskodawca	Minister Rolnictwa i Rozwoju Wsi		
Beneficjent	Instytut Hodowli i Aklimatyzacji Roślin - Państwowy Instytut Badawczy		
Partnerzy	Yellow's Sp. z o.o.		
Źródło finansowania	FERC, Działanie 02.03 Cyfrowa dostępność i ponowne wykorzystanie informacji budżet państwa: część 27 „Informatyzacja”		
Całkowity koszt projektu	16 921 000,00 zł		
Planowany okres realizacji projektu	09-2026 do 11-2029		
Osoba kontaktowa	Maja Boczkowska	m.boczkowska@ihar.edu.pl	607547959

1. POWODY PODJĘCIA PROJEKTU

1.1. Identyfikacja problemu i potrzeb

Projekt odpowiada na problem braku możliwości szerokiego, zautomatyzowanego i interoperacyjnego udostępniania cyfrowych informacji o sekwencji DNA roślinnych zasobów genowych zgromadzonych w Krajowym Centrum Roślinnych Zasobów Genowych. Obecnie zasoby te mają strategiczne znaczenie dla nauki, hodowli roślin, ochrony agrobioróżnorodności i bezpieczeństwa żywnościowego, lecz ich potencjał informacyjny nie jest dostępny w formie pozwalającej na maszynowe przetwarzanie, integrację i ponowne wykorzystanie.

Stan obecny odbiega od stanu oczekiwanego, w którym dane naukowe finansowane ze środków publicznych powinny być udostępniane on-line, w sposób uporządkowany, wyszukiwalny, interoperacyjny i zgodny z zasadami FAIR. Dotychczas cyfrowe informacje o sekwencji DNA obejmują jedynie około 3 000 obiektów, 60 mln loci i 5,7 mld nukleotydów, co oznacza bardzo niski poziom digitalizacji zasobu. Dane sekwencyjne nie są przechowywane w jednolitym repozytorium, a dane paszportowe, fenotypowe i metadane funkcjonują w rozproszonych strukturach. Brak API ogranicza automatyczny dostęp do danych oraz ich integrację z systemami zewnętrznymi.

Projekt zakłada wytworzenie systemu teleinformatycznego Cyfrowy Bank Genów Roślin, repozytorium danych sekwencyjnych i metadanych, publicznego API oraz mechanizmów wyszukiwania i pobierania danych. Produkty te umożliwią digitalizację, integrację, standaryzację i udostępnienie danych o zasobach KCRZG, eliminując obecną lukę między znaczeniem zasobu a jego dostępnością cyfrową.

Interesariusz	Zidentyfikowany problem	Szacowana wielkość grupy
Pracownicy naukowcy z kraju i ze świata - nauki rolnicze, biologiczne i pokrewne	- Brak jednolitego repozytorium danych sekwencyjnych, paszportowych i fenotypowych KCRZG ogranicza wykorzystanie zasobów genowych w badaniach nad zmiennością genetyczną,	300 000

Interesariusz	Zidentyfikowany problem	Szacowana wielkość grupy
	<p>odpornością roślin, adaptacją do stresów i ochroną agrobioróżnorodności.</p> <p>- Niewystarczający poziom digitalizacji sekwencji DNA utrudnia prowadzenie analiz genomowych, porównawczych i populacyjnych.</p> <p>-Ograniczona standaryzacja metadanych oraz brak API zmniejszają możliwość automatycznego pobierania, przetwarzania i integracji danych z krajowymi i międzynarodowymi zasobami naukowymi.</p>	
Hodowcy roślin z kraju i ze świata	<p>- Brak jednolitego dostępu do danych sekwencyjnych, paszportowych i fenotypowych KCRZG ogranicza możliwość identyfikacji materiałów przydatnych w hodowli roślin.</p> <p>- Niewystarczający poziom digitalizacji sekwencji DNA utrudnia wykorzystanie zasobów genowych w hodowli wspomaganej markerami, selekcji genomowej oraz poszukiwaniu źródeł odporności i tolerancji na stresy środowiskowe.</p> <p>- Ograniczona interoperacyjność danych oraz brak API zwiększają czas i koszty pozyskiwania informacji oraz utrudniają ich automatyczne wykorzystanie w narzędziach hodowlanych.</p>	50 000
Pracownicy administracji publicznej i instytucji międzynarodowych	<p>- Brak jednego źródła danych o cyfrowych informacjach sekwencyjnych zasobów KCRZG.</p> <p>- Niewystarczająca interoperacyjność danych utrudniająca raportowanie i analizy strategiczne.</p> <p>- Ograniczony dostęp do danych maszynowo odczytywalnych, wymaganych do integracji z systemami krajowymi i międzynarodowymi.</p>	20000
Pracownicy przedsiębiorstw biotechnologicznych i podmiotów rozwijające usługi oparte na danych z kraju i ze świata	<p>- Brak API umożliwiającego automatyczny dostęp do danych genomowych i metadanych.</p> <p>- Niewystarczająca standaryzacja danych ograniczająca ich wykorzystanie w analizach bioinformatycznych i biotechnologii.</p> <p>- Ograniczona dostępność zbiorów</p>	100000

Interesariusz	Zidentyfikowany problem	Szacowana wielkość grupy
	referencyjnych do tworzenia narzędzi analitycznych i usług opartych na danych.	
Studenci, doktoranci i kadra dydaktyczna z kraju i ze świata	<ul style="list-style-type: none"> - Brak łatwo dostępnych, rzeczywistych, krajowych zbiorów danych typu big data do celów edukacyjnych. - Niewystarczająca dostępność danych sekwencyjnych i metadanych do nauczania genomiki, bioinformatyki i analizy danych. - Ograniczona możliwość pracy z danymi zgodnymi z zasadami FAIR i otwartej nauki. 	100000
Banki genów i instytucje zarządzające zasobami genetycznymi roślin	<ul style="list-style-type: none"> - Brak interoperacyjnych danych sekwencyjnych umożliwiających porównywanie zasobów i identyfikację duplikatów. - Niewystarczająca standaryzacja metadanych utrudniająca wymianę informacji między bankami genów. - Ograniczona możliwość integracji danych KCRZG z krajowymi i międzynarodowymi systemami informacji o zasobach genetycznych. 	20000
Ministerstwo Rolnictwa i Rozwoju Wsi	<ul style="list-style-type: none"> - Brak jednego, interoperacyjnego źródła danych o cyfrowych zasobach genowych ogranicza możliwość wykorzystania tych danych w politykach publicznych dotyczących rolnictwa, bezpieczeństwa żywnościowego, ochrony bioróżnorodności i adaptacji do zmian klimatu. - Niewystarczająca dostępność danych utrudnia monitorowanie i raportowanie w obszarze roślinnych zasobów genetycznych. 	1
Instytut Hodowli i Aklimatyzacji Roślin – Państwowy Instytut Badawczy	<ul style="list-style-type: none"> - Brak systemu umożliwiającego gromadzenie, integrację i udostępnianie cyfrowych informacji o sekwencji DNA ogranicza efektywność zarządzania zasobami KCRZG. - Niewystarczająca automatyzacja i brak API utrudniają udostępnianie danych użytkownikom zewnętrznym oraz integrację z innymi systemami. 	1

1.2. Opis stanu obecnego

Obecnie procesy gromadzenia, ewidencji i udostępniania danych o roślinnych zasobach genowych w Krajowym Centrum Roślinnych Zasobów Genowych realizowane są z wykorzystaniem kilku odrębnych systemów informatycznych o ograniczonej interoperacyjności. Kluczową rolę pełni system EGISET, oparty na rozwiązaniu klasy ERP, utrzymywany na serwerze z systemem Windows Server 2008, którego baza danych funkcjonuje na oddzielnej jednostce. System obsługuje ewidencję zasobów oraz podstawowe funkcje wyszukiwania i udostępniania informacji o kolekcji. Równolegle wykorzystywany jest system SGMS działający w środowisku Linux i Docker, wspierający zarządzanie kolekcją, jednak bez funkcjonalności integracji danych i ich zautomatyzowanego udostępniania. Infrastruktura obejmuje serwery fizyczne i maszyny wirtualne, a część usług działa w środowisku kontenerowym. W obecnym modelu przetwarzane są głównie dane paszportowe, ewidencyjne oraz dane związane z obsługą udostępniania materiału biologicznego. Istotnym ograniczeniem jest brak infrastruktury do gromadzenia, integracji i udostępniania cyfrowych zapisów sekwencji DNA (DSI). Dane te nie są przechowywane w jednolitym repozytorium ani udostępniane użytkownikom, co uniemożliwia ich wykorzystanie w analizach genomowych. Dodatkowo brak standaryzacji metadanych oraz interfejsów API uniemożliwia maszynowy dostęp do danych i ich integrację z systemami zewnętrznymi. Poziomą digitalizację w zakresie danych sekwencyjnych pozostaje bardzo niski i szacowany jest na około 2%. Realizacja projektu zakłada rozwój infrastruktury umożliwiającej wytwarzanie, gromadzenie i integrację danych sekwencyjnych oraz ich powiązanie z istniejącymi zasobami. Obecnie nie funkcjonuje system teleinformatyczny umożliwiający gromadzenie, walidację, integrację i publiczne udostępnianie cyfrowych informacji o sekwencji DNA wraz z metadanymi. Brak API uniemożliwia maszynowy dostęp do danych oraz automatyczną integrację z systemami zewnętrznymi.

2. EFEKTY PROJEKTU

2.1. Cele i korzyści wynikające z projektu

Cel - 1	Zwiększenie cyfrowej dostępności zasobów nauki Krajowego Centrum Roślinnych Zasobów Genowych poprzez ich publiczne udostępnienie on-line w formie umożliwiającej dalsze wykorzystanie.
Cel strategiczny	<p>1) Fundusze Europejskie na Rozwój Cyfrowy 2021–2027, Priorytet II „Zaawansowane usługi cyfrowe”, Działanie 02.03 „Cyfrowa dostępność i ponowne wykorzystanie informacji” – cel: zwiększenie cyfrowej dostępności i jakości informacji sektora publicznego oraz ich ponownego wykorzystania.</p> <p>2) Program otwierania danych na lata 2021–2027 – cele: zwiększenie dostępności danych, poprawa interoperacyjności i jakości danych oraz wzrost wykorzystywania i wymiany danych.</p> <p>3) Polityka Naukowa Państwa – cel: wspieranie otwartej nauki, w tym otwartego dostępu do publikacji naukowych i otwierania danych badawczych, oraz zwiększenie wykorzystania wyników badań w społeczeństwie i gospodarce.</p> <p>4) Krajowy plan działania do programu polityki „Droga ku cyfrowej dekadzie” do 2030 r. – cel: realizacja krajowych działań służących osiągnięciu celów cyfrowych Unii Europejskiej do 2030 r., w tym rozwojowi cyfrowych usług publicznych, infrastruktury danych i interoperacyjności.</p> <p>5) Drugi Globalny Plan Działań FAO dla roślinnych zasobów genetycznych dla żywienia i rolnictwa – cel: dokumentacja, charakterystyka, ocena i</p>

	<p>zrównoważone użytkowanie zasobów genetycznych</p> <p>6) Konwencja o różnorodności biologicznej – cele: ochrona różnorodności biologicznej, zrównoważone użytkowanie jej elementów oraz sprawiedliwy i równy podział korzyści wynikających z wykorzystania zasobów genetycznych.</p> <p>7) Międzynarodowy Traktat o Zasobach Genetycznych Roślin dla Wyżywienia i Rolnictwa – cele: ochrona i zrównoważone użytkowanie roślinnych zasobów genetycznych dla wyżywienia i rolnictwa oraz sprawiedliwy i równy podział korzyści wynikających z ich wykorzystania.</p>
Korzyść:	<ul style="list-style-type: none"> - Zwiększenie dostępności cyfrowych danych o zasobach nauki KCRZG. - Poprawa użyteczności danych sekwencyjnych, paszportowych i fenotypowych. - Zwiększenie przejrzystości i wartości publicznych zasobów nauki. - Usprawnienie dostępu do danych dla nauki, administracji, hodowli i podmiotów wykorzystujących dane.
KPI:	<p>KPI1 - Liczba podmiotów wspartych w zakresie rozwoju usług, produktów i procesów cyfrowych</p> <p>KPI2 - Liczba podmiotów, które udostępniły informacje sektora publicznego/dane prywatne on-line</p> <p>KPI3 - Liczba udostępnionych on-line dokumentów zawierających informacje sektora publicznego/dane prywatne</p> <p>KPI4 - Liczba zdigitalizowanych dokumentów zawierających informacje sektora publicznego/dane prywatne</p> <p>KPI5 - Rozmiar udostępnionych on-line informacji sektora publicznego/prywatnego</p> <p>KPI6 - Liczba loci udostępnionych w systemie</p> <p>KPI7 - Rozmiar zdigitalizowanych informacji sektora publicznego/danych prywatnych [TB]</p>
Wartość aktualna i docelowa KPI:	<p>KPI1: - wartość aktualna: 0 - wartość docelowa: 2</p> <p>KPI2: - wartość aktualna: 0 - wartość docelowa - 1</p> <p>KPI3: - wartość aktualna: 0 - wartość docelowa –35 000</p> <p>KPI4: - wartość aktualna: 3000 - wartość docelowa: 35 000</p> <p>KPI5: - wartość aktualna- 0.025 TB</p>

	<ul style="list-style-type: none"> - wartość docelowa - 5.025 TB <p>KPI6:</p> <ul style="list-style-type: none"> - wartość aktualna - 0 - wartość docelowa - 1 600 000 000 <p>KPI7:</p> <ul style="list-style-type: none"> - wartość aktualna- 0.025 TB - wartość docelowa - 5.025 TB
<p>Metoda pomiaru KPI</p>	<p>KPI1</p> <ul style="list-style-type: none"> - Metoda i sposób pomiaru: pomiar ilościowy na podstawie analizy dokumentacji projektowej oraz stopnia realizacji zakresu rzeczowego projektu. - Źródło danych: umowa o dofinansowanie, dokumentacja projektowa, protokoły odbioru, raporty z realizacji projektu. - Częstotliwość pomiaru: jednorazowo na zakończenie rzeczowej realizacji projektu (pomiar wartości docelowej). <p>KPI2</p> <ul style="list-style-type: none"> - Metoda i sposób pomiaru: pomiar jakościowo-ilościowy polegający na weryfikacji uruchomienia platformy oraz publikacji zasobów. - Źródło danych: protokół uruchomienia platformy, dokumentacja wdrożeniowa, raport publikacji danych, potwierdzenie dostępności systemu. - Częstotliwość pomiaru: jednorazowo na zakończenie rzeczowej realizacji projektu. <p>KPI3</p> <ul style="list-style-type: none"> - Metoda i sposób pomiaru: pomiar ilościowy na podstawie inwentaryzacji zasobów oraz liczby opublikowanych on-line obiektów. - Źródło danych: baza danych systemu, repozytorium danych, raport wdrożeniowy, statystyki systemowe. - Częstotliwość pomiaru: jednorazowo na zakończenie rzeczowej realizacji projektu (pomiar wartości docelowej). <p>KPI4</p> <ul style="list-style-type: none"> - Metoda i sposób pomiaru: pomiar ilościowy na podstawie inwentaryzacji zasobów oraz liczby zdigitalizowanych obiektów. - Źródło danych: baza danych systemu, repozytorium danych, raport wdrożeniowy, statystyki systemowe. - Częstotliwość pomiaru: półroczna oraz jednorazowo na zakończenie rzeczowej realizacji projektu (pomiar wartości docelowej). <p>KPI5</p> <ul style="list-style-type: none"> - Metoda i sposób pomiaru: pomiar ilościowy wielkości danych udostępnionych on-line, wyrażony w TB, na podstawie sumarycznego rozmiaru plików (FASTQ raw, danych po obróbce, metadanych i raportów jakości). - Źródło danych: repozytorium danych, statystyki systemowe, raport wdrożeniowy. - Częstotliwość pomiaru: jednorazowo na zakończenie rzeczowej realizacji projektu (pomiar wartości docelowej). <p>KPI6</p>

	<p>- Metoda i sposób pomiaru: pomiar ilościowy na podstawie liczby zdigitalizowanych i udostępnionych obiektów oraz przypisanej liczby loci. - Źródło danych: baza danych systemu, repozytorium danych, raport wdrożeniowy, statystyki systemowe. - Częstotliwość pomiaru: jednorazowo na zakończenie rzeczowej realizacji projektu (pomiar wartości docelowej).</p> <p>KPI7</p> <p>- Metoda i sposób pomiaru: pomiar ilościowy polegający na określeniu łącznego rozmiaru danych zdigitalizowanych w ramach projektu, wyrażonego w TB. - Źródło danych: repozytorium danych, statystyki systemowe, raport wdrożeniowy. - Częstotliwość pomiaru: jednorazowo na zakończenie rzeczowej realizacji projektu (pomiar wartości docelowej).</p>
Cel - 2	Rozwój interoperacyjnych usług i narzędzi cyfrowych umożliwiających techniczne ponowne wykorzystanie zasobów nauki Krajowego Centrum Roślinnych Zasobów Genowych.
Cel strategiczny	<p>1) Fundusze Europejskie na Rozwój Cyfrowy 2021–2027, Priorytet II „Zaawansowane usługi cyfrowe”, Działanie 02.03 „Cyfrowa dostępność i ponowne wykorzystanie informacji” – cel: zwiększenie cyfrowej dostępności i jakości informacji sektora publicznego oraz ich ponownego wykorzystania.</p> <p>2) Program otwierania danych na lata 2021–2027 – cele: zwiększenie dostępności danych, poprawa interoperacyjności i jakości danych oraz wzrost wykorzystywania i wymiany danych.</p> <p>3) Polityka Naukowa Państwa – cel: wspieranie otwartej nauki, w tym otwartego dostępu do publikacji naukowych i otwierania danych badawczych, oraz zwiększenie wykorzystania wyników badań w społeczeństwie i gospodarce.</p> <p>4) Krajowy plan działania do programu polityki „Droga ku cyfrowej dekadzie” do 2030 r. – cel: realizacja krajowych działań służących osiągnięciu celów cyfrowych Unii Europejskiej do 2030 r., w tym rozwojowi cyfrowych usług publicznych, infrastruktury danych i interoperacyjności.</p> <p>5) Europejska strategia w zakresie danych – cel: budowa europejskiej gospodarki opartej na danych oraz zwiększenie dostępności, wymiany i ponownego wykorzystania danych.</p> <p>6) Unijna strategia na rzecz bioróżnorodności 2030 „Przywracanie przyrody do naszego życia” – cel: ochrona przyrody, odwrócenie degradacji ekosystemów oraz wprowadzenie europejskiej różnorodności biologicznej na ścieżkę regeneracji do 2030 r.</p> <p>7) Drugi Globalny Plan Działań FAO dla roślinnych zasobów genetycznych dla żywienia i rolnictwa – cel: dokumentacja, charakterystyka, ocena i zrównoważone użytkowanie zasobów genetycznych</p> <p>8) Konwencja o różnorodności biologicznej – cele: ochrona różnorodności biologicznej, zrównoważone użytkowanie jej elementów oraz sprawiedliwy i</p>

	<p>równy podział korzyści wynikających z wykorzystania zasobów genetycznych.</p> <p>9) Międzynarodowy Traktat o Zasobach Genetycznych Roślin dla Wyżywienia i Rolnictwa – cele: ochrona i zrównoważone użytkowanie roślinnych zasobów genetycznych dla wyżywienia i rolnictwa oraz sprawiedliwy i równy podział korzyści wynikających z ich wykorzystania.</p>
Korzyść:	<ul style="list-style-type: none"> - Zwiększenie interoperacyjności danych KCRZG. - Usprawnienie technicznego dostępu do danych przez API. - Poprawa jakości zarządzania danymi i metadanymi. - Zwiększenie możliwości integracji danych z systemami zewnętrznymi.
KPI:	<p>KPI1 - Liczba baz danych udostępnionych on-line poprzez API</p> <p>KPI2 - Liczba platform udostępniania dokumentów zawierających informacje sektora publicznego/dane prywatne</p> <p>KPI3 - Liczba uruchomionych systemów teleinformatycznych</p> <p>KPI4 - Liczba utworzonych API</p> <p>KPI5 - Liczba rozwiązań wykorzystujących informacje sektora publicznego/dane prywatne</p> <p>KPI6 - Liczba pracowników IT objętych wsparciem szkoleniowym</p> <p>KPI7 - Liczba pracowników nie będących pracownikami IT, objętych wsparciem szkoleniowym</p> <p>KPI8 - Liczba wygenerowanych kluczy API</p> <p>KPI9 - Wartość usług, produktów i procesów cyfrowych opracowanych dla przedsiębiorstw</p>
Wartość aktualna i docelowa KPI:	<p>KPI1: - wartość aktualna: 0 - wartość docelowa: 1</p> <p>KPI2: - wartość aktualna: 0 - wartość docelowa: 1</p> <p>KPI3: - wartość aktualna: 0 - wartość docelowa: 1</p> <p>KPI4: - wartość aktualna: 0 - wartość docelowa: 1</p> <p>KPI5: - wartość aktualna: 0 - wartość docelowa: 1</p> <p>KPI6: - wartość aktualna: 0</p>

	<p>- wartość docelowa: 3</p> <p>KPI7: - wartość aktualna: 0 - wartość docelowa: 50</p> <p>KPI8: - wartość aktualna: 0 - wartość docelowa: 1</p> <p>KPI9: - wartość aktualna: 0 - wartość docelowa: 7 658 878</p>
Metoda pomiaru KPI	<p>KPI1 - Metoda i sposób pomiaru: pomiar ilościowy na podstawie weryfikacji liczby baz danych udostępnionych poprzez działające interfejsy API. - Źródło danych: dokumentacja wdrożeniowa, rejestr API, protokół odbioru systemu, raport wdrożeniowy. - Częstotliwość pomiaru: kwartalna; wartość docelowa mierzona na zakończenie rzeczowej realizacji projektu.</p> <p>KPI2 - Metoda i sposób pomiaru: pomiar ilościowy poprzez identyfikację i weryfikację uruchomionych platform spełniających wymagania projektu. - Źródło danych: dokumentacja projektowa, protokół odbioru systemu, raport wdrożeniowy. - Częstotliwość pomiaru: Częstotliwość pomiaru: kwartalna; wartość docelowa mierzona na zakończenie rzeczowej realizacji projektu.</p> <p>KPI3 - Metoda i sposób pomiaru: pomiar ilościowy na podstawie liczby systemów wdrożonych i uruchomionych produkcyjnie. - Źródło danych: dokumentacja wdrożeniowa, protokoły odbioru, raport z realizacji projektu. - Częstotliwość pomiaru: kwartalna; wartość docelowa mierzona na zakończenie rzeczowej realizacji projektu.</p> <p>KPI4 - Metoda i sposób pomiaru: pomiar ilościowy na podstawie liczby utworzonych i udokumentowanych interfejsów API udostępniających dane. - Źródło danych: dokumentacja techniczna systemu, rejestr API, raport wdrożeniowy. - Częstotliwość pomiaru: wartość docelowa mierzona na zakończenie rzeczowej realizacji projektu.</p> <p>KPI5 - Metoda i sposób pomiaru: pomiar ilościowy na podstawie identyfikacji rozwiązań wykorzystujących dane (np. integracje, narzędzia analityczne), uzupełniony analizą wykorzystania API. - Źródło danych: statystyki systemowe, raporty wykorzystania API, ankiety/ interakcje użytkowników, dokumentacja projektowa. - Częstotliwość pomiaru: roczna; wartość docelowa mierzona do 12 miesięcy po zakończeniu rzeczowej realizacji projektu.</p>

	<p>KPI6</p> <ul style="list-style-type: none"> - Metoda i sposób pomiaru: pomiar ilościowy na podstawie liczby osób, które ukończyły szkolenia i uzyskały potwierdzenie uczestnictwa. - Źródło danych: listy obecności, certyfikaty ukończenia szkoleń, raporty szkoleniowe. - Częstotliwość pomiaru: kwartalna; wartość docelowa mierzona na zakończenie rzeczowej realizacji projektu. <p>KPI7</p> <ul style="list-style-type: none"> - Metoda i sposób pomiaru: pomiar ilościowy na podstawie liczby osób, które ukończyły szkolenia i uzyskały potwierdzenie uczestnictwa. - Źródło danych: listy obecności, certyfikaty ukończenia szkoleń, raporty szkoleniowe. - Częstotliwość pomiaru: kwartalna; wartość docelowa mierzona na zakończenie rzeczowej realizacji projektu. <p>KPI8</p> <ul style="list-style-type: none"> - Metoda i sposób pomiaru: pomiar ilościowy na podstawie liczby unikalnych kluczy API wygenerowanych i aktywowanych w systemie Cyfrowy Bank Genów Roślin dla użytkowników lub podmiotów korzystających z dostępu maszynowego do danych. - Źródło danych: rejestr kluczy API, logi systemowe, dokumentacja administracyjna systemu. - Częstotliwość pomiaru: roczna; wartość docelowa mierzona na zakończenie rzeczowej realizacji projektu. <p>KPI9</p> <ul style="list-style-type: none"> - Metoda i sposób pomiaru: Pomiar ilościowy na podstawie wartości kosztów projektu z wykluczeniem kosztów wynagrodzeń, szkoleń, informacji i promocji i kosztów pośrednich. - Źródło danych: budżet projektu, umowa partnerska, dokumentacja księgową i raporty z realizacji projektu. - Częstotliwość pomiaru: pomiar półroczny oraz jednorazowo na zakończenie rzeczowej realizacji projektu.
Cel - 3	- Zwiększenie wykorzystania cyfrowo udostępnionych zasobów nauki KCRZG przez użytkowników końcowych oraz podmioty ponownie wykorzystujące dane.
Cel strategiczny	<p>Fundusze Europejskie na Rozwój Cyfrowy 2021–2027, Priorytet II „Zaawansowane usługi cyfrowe”, Działanie 02.03 „Cyfrowa dostępność i ponowne wykorzystanie informacji” – cel: zwiększenie cyfrowej dostępności i jakości informacji sektora publicznego oraz ich ponownego wykorzystania.</p> <p>Program otwierania danych na lata 2021–2027 – cele: zwiększenie dostępności danych, poprawa interoperacyjności i jakości danych oraz wzrost wykorzystywania i wymiany danych.</p> <p>Polityka Naukowa Państwa – cel: wspieranie otwartej nauki, w tym otwartego dostępu do publikacji naukowych i otwierania danych badawczych, oraz zwiększenie wykorzystania wyników badań w społeczeństwie i gospodarce.</p> <p>Krajowy plan działania do programu polityki „Droga ku cyfrowej dekadzie” do</p>

	<p>2030 r. – cel: realizacja krajowych działań służących osiągnięciu celów cyfrowych Unii Europejskiej do 2030 r., w tym rozwojowi cyfrowych usług publicznych, infrastruktury danych i interoperacyjności.</p> <p>Europejska strategia w zakresie danych – cel: budowa europejskiej gospodarki opartej na danych oraz zwiększenie dostępności, wymiany i ponownego wykorzystania danych.</p> <p>Unijna strategia na rzecz bioróżnorodności 2030 „Przywracanie przyrody do naszego życia” – cel: ochrona przyrody, odwrócenie degradacji ekosystemów oraz wprowadzenie europejskiej różnorodności biologicznej na ścieżkę regeneracji do 2030 r.</p> <p>Drugi Globalny Plan Działań FAO dla roślinnych zasobów genetycznych dla żywienia i rolnictwa – cel: dokumentacja, charakterystyka, ocena i zrównoważone użytkowanie zasobów genetycznych</p> <p>Konwencja o różnorodności biologicznej – cele: ochrona różnorodności biologicznej, zrównoważone użytkowanie jej elementów oraz sprawiedliwy i równy podział korzyści wynikających z wykorzystania zasobów genetycznych.</p> <p>Międzynarodowy Traktat o Zasobach Genetycznych Roślin dla Wyżywienia i Rolnictwa – cele: ochrona i zrównoważone użytkowanie roślinnych zasobów genetycznych dla żywienia i rolnictwa oraz sprawiedliwy i równy podział korzyści wynikających z ich wykorzystania.</p>
Korzyść:	<p>Zwiększenie liczby użytkowników cyfrowych zasobów KCRZG. Zwiększenie skali ponownego wykorzystania danych. Poprawa wykorzystania danych w badaniach, edukacji, administracji i gospodarce. Wzmocnienie widoczności KCRZG jako krajowego źródła danych referencyjnych.</p>
KPI:	KPI1 - Użytkownicy nowych i zmodernizowanych publicznych usług, produktów i procesów cyfrowych
Wartość aktualna i docelowa KPI:	<p>KPI1</p> <ul style="list-style-type: none"> - wartość aktualna: 0 - wartość docelowa: 15 000
Metoda pomiaru KPI	<p>KPI1</p> <p>Metoda i sposób pomiaru: pomiar ilościowy na podstawie analizy liczby unikalnych użytkowników korzystających z zasobów cyfrowych systemu (np. liczba unikalnych adresów IP, kont użytkowników lub sesji), uzupełniony analizą aktywności (pobrania danych, zapytania do systemu, wykorzystanie zasobów).</p> <p>Źródło danych: statystyki systemowe (logi serwera, narzędzia analityczne), baza danych systemu, raporty wykorzystania zasobów, dokumentacja projektowa i raport końcowy.</p> <p>Częstotliwość pomiaru: pomiar kwartalny po zakończeniu realizacji projektu; pomiar wartości docelowej – 12 miesięcy po zakończeniu rzeczowej realizacji projektu.</p>

2.2. Udostępnione e-usługi

Lp.	Nazwa e-usługi	Typ	Zakres oddziaływania	Poziom dojrzałości e-usługi
1	Wyszukiwanie i przeglądanie zasobów KCRZG	A2C A2B A2A	Pracownicy naukowcy z kraju i ze świata - nauki rolnicze, biologiczne i pokrewne Hodowcy roślin z kraju i ze świata Pracownicy administracji publicznej i instytucji międzynarodowych Pracownicy przedsiębiorstw biotechnologicznych i podmiotów rozwijające usługi oparte na danych z kraju i ze świata Studenci, doktoranci i kadra dydaktyczna z kraju i ze świata Banki genów i instytucje zarządzające zasobami genetycznymi roślin Ministerstwo Rolnictwa i Rozwoju Wsi Instytut Hodowli i Aklimatyzacji Roślin – Państwowy Instytut Badawczy (rocznie ok 50000 transakcji)	Dwustronna interakcja
2	Pobieranie danych genomowych i zbiorów danych przez API	A2B A2A	Banki genów i instytucje zarządzające zasobami genetycznymi roślin Pracownicy przedsiębiorstw biotechnologicznych i podmiotów rozwijające usługi oparte na danych z kraju i ze świata Hodowcy roślin z kraju i ze świata Pracownicy administracji publicznej i instytucji międzynarodowych Ministerstwo Rolnictwa i Rozwoju Wsi	Personalizacja

Lp.	Nazwa e-usługi	Typ	Zakres oddziaływania	Poziom dojrzałości e-usługi
			(rocznie ok 1000 transakcji)	
3	Zamawianie próbek materiału genetycznego (nasiona i DNA)	A2C A2B A2A	Pracownicy naukowcy z kraju i ze świata - nauki rolnicze, biologiczne i pokrewne Hodowcy roślin z kraju i ze świata Pracownicy administracji publicznej i instytucji międzynarodowych Pracownicy przedsiębiorstw biotechnologicznych i podmiotów rozwijające usługi oparte na danych z kraju i ze świata Studenci, doktoranci i kadra dydaktyczna z kraju i ze świata Banki genów i instytucje zarządzające zasobami genetycznymi roślin (rocznie ok 250 transakcji)	Transakcja

2.3. Udostępnione informacje sektora publicznego i zdigitalizowane zasoby

Rodzaj informacji/zasobów	Planowana data udostępnienia	Szacowana liczba obiektów objętych digitalizacją (udostępnianiem informacji)
Zdygitalizowane i udostępnione cyfrowe informacje o sekwencji	30-11-2029	35000
Udostępnione dane paszportowe obiektów w kolekcji zasobów genowych	30-11-2029	90000
Udostępnione historyczne dane fenotypowe	30-11-2029	20000

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	10-2026
Interfejs API CBGR (REST API, specyfikacja OpenAPI)	10-2028
Raport z badań UX / testów użyteczności - wynik pozytywny	11-2028
Ocena Skutków dla Ochrony Danych (DPIA) wraz z RCPD	04-2029
Raport z testów wydajności - wynik pozytywny	06-2029
Raport z testów bezpieczeństwa - wynik pozytywny (pentest)	06-2029
Materiały szkoleniowe (instrukcje, nagrania, ćwiczenia)	08-2029
Raport ze szkoleń użytkowników systemu CBGR	08-2029
Dokumentacja techniczna, użytkowa, administracyjna i integracyjna systemu CBGR	08-2029
System teleinformatyczny Cyfrowy Bank Genów Roslin (kompletna platforma produkcyjna)	08-2029
Rejestr publiczny Cyfrowy Bank Genów Roslin	08-2029
Raport końcowy projektu	11-2029
Raport z migracji danych z EGISET do CBGR	06-2029

3. KAMIENIE MIŁOWE

Kamienie milowe	Planowany termin osiągnięcia
Przeprowadzony inicjalny test prywatności	2026-10-31
Ogłoszono przetargi na zakup aparatury laboratoryjnej do digitalizacji sekwencji DNA	2026-12-31
Ogłoszono przetarg na wykonanie usługi digitalizacji sekwencji DNA	2026-12-31
Zatwierdzona analiza wymagań, model procesów biznesowych i koncepcja architektury systemu	2027-04-30
Uruchomiona aparatura laboratoryjna do digitalizacji sekwencji DNA	2027-05-31
Wybrany dostawca usługi digitalizacji sekwencji DNA	2027-06-30
Zatwierdzona architektura, schemat bazy danych oraz pilotażowa migracja danych	2027-10-31
Zdigitalizowane 30% kolekcji	2028-01-31
Utworzone repozytorium genotypów wraz z wdrożonymi podstawowymi funkcjonalnościami systemu Cyfrowy Bank Genów Roślin	2028-04-30
Zdigitalizowane 60% kolekcji	2028-08-31

Kamienie milowe	Planowany termin osiągnięcia
Dostarczona i odebrana wyszukiwarka genotypów	2028-10-31
Uzyskany pozytywny wynik testów badan UX systemu Cyfrowy Bank Genów Roslin	2028-11-30
Wdrożony moduł obsługi zamówień nasion i sekwencji genetycznych	2029-04-30
Zdigitalizowane 90% kolekcji	2029-05-31
Zakończone testy integracyjne i akceptacyjne systemu Cyfrowy Bank Genów Roślin	2029-06-30
Uzyskany pozytywny wynik testów wydajności systemu Cyfrowy Bank Genów Roślin	2029-06-30
Uzyskany pozytywny wynik testów bezpieczeństwa systemu Cyfrowy Bank Genów Roślin	2029-06-30
Zakończona pełna migracja danych z systemu EGISET do systemu CBGR	2029-06-30
Uruchomiony system Cyfrowy Bank Genów Roślin wraz z dokumentacją i przeprowadzonymi szkoleniami użytkowników	2029-08-31
Zdigitalizowane 100% kolekcji	2029-09-30
Zakończona stabilizacja produkcyjna, ewaluacja działania systemu Cyfrowy Bank Genów Roślin	2029-10-31
Przygotowany raport końcowy	2029-11-30
Raport z oceny skutków dla ochrony Danych (DPIA) wraz z RCPD	2029-04-30

4. KOSZTY

4.1. Koszty ogólne projektu wraz ze sposobem finansowania

Całkowity koszt projektu (netto oraz brutto), w tym	Netto 15 328 390,24 zł Brutto 16 921 000,00 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)	2026	Netto 1 291 186,99 zł Brutto 1 452 000,00 zł
	2027	Netto 5 744 910,57 zł Brutto 6 448 000,00 zł
	2028	Netto 4 876 918,70 zł Brutto 5 399 000,00 zł
	2029	Netto 3 415 373,98 zł Brutto 3 622 000,00 zł

4.2. Wykaz poszczególnych pozycji kosztowych

Nazwa pozycji kosztowej		Przewidywany koszt brutto	Uzasadnienie pozycji kosztowej (przeznaczenie)
Oprogramowanie	<p>1) Wytworzenie systemu teleinformatycznego o CBGR (5 komponentów + ok. 3 e-usługi) w metodyce Agile: analiza, projektowanie, implementacja, testy, migracja danych z EGISET, dokumentacja i wsparcie wdrożeniowe. Wykonawca: Yellow's Sp. z o.o. (Partner technologiczny).</p> <p>2) Przygotowanie materiałów do dygitalizacji - Koszty zatrudnienia zespołu przygotowującego</p>	13 722 483,00 zł	<p>1) Koszty obejmują:</p> <ul style="list-style-type: none"> - Analiza wymagań, projekt architektury i modelu danych - 233645 PLN - Wytworzenie 5 komponentów systemu CBGR: Repozytorium danych, Wyszukiwarka publiczna, API Gateway, Platforma administracyjna, Moduł wymiany - 1261682 PLN - Wytworzenie modułu obsługi zamówień próbek nasion i sekwencji (e-usługa L4 z autoryzacją) - 233645 PLN - Migracja danych z systemu EGISET - 93458 PLN - Konfiguracja środowiska chmurowego (CI/CD, konteneryzacja, monitoring, backupy, DR) - 121495 PLN - Testy deweloperskie – jednostkowe, integracyjne, akceptacyjne (UAT z udziałem IHAR-PIB) - 186916 PLN - Dokumentacja techniczna, użytkowa, administracyjna i

Nazwa pozycji kosztowej		Przewidywany koszt brutto	Uzasadnienie pozycji kosztowej (przeznaczenie)
	<p>materiał biologiczny do digitalizacji, zakup odczynników i materiałów zużywalnych niezbędnych do izolacji DNA do digitalizacji, zakup niezbędnej aparatury laboratoryjnej</p> <p>3) Dygitalizacja sekwencji nici DNA - sekwencjonowanie metodą DArTseq</p>		<p>integracyjna systemu - 121495 PLN</p> <p>- Wsparcie wdrożeniowe i stabilizacja produkcyjna (ok 60 dni) - 93458 PLN</p> <p>- Korekty wymagań wynikające z UAT i zmiany regulacyjne - 3738 PLN</p> <p>2) Koszty obejmują:</p> <p>- Zespół przygotowujący materiał biologiczny do digitalizacji: pracownicy naukowci (2,1 FTE - 1214927 PLN), pracownicy inżynieryjno-techniczni (7,2 FTE - 2509425 PLN)</p> <p>- Odczynniki i materiały zużywalne potrzebne do izolacji DNA (zestawy do izolacji - 1121495 PLN, odczynniki chemiczne - 93458 PLN, plastiki i szkło laboratoryjne - 280374 PLN, drobny sprzęt laboratoryjny - 103738 PLN) - 1599065 PLN</p> <p>- Aparatura laboratoryjna porzeba do izolacji DNA: spektrofotometry i fluorymetry (350467 PLN), pompy i zamykarki próżniowe (112150 PLN), automatyczna stacja do izolacji DNA (233645 PLN), lizer do tkanek/młyn (142991 PLN), liofilizator (93458 PLN), zestaw do wizualizacji żeli (205607 PLN), wirówka laboratoryjna z wyposażeniem (102804 PLN), wagi laboratoryjne (84112 PLN).</p> <p>3) Zlecenie usługi sekwencjonowania (digitalizacja sekwencji DNA) w technologii DArTseq 32 000 próbek DNA z obiektów przechowywanych w KCRZG (1.2 mln odczytów/ próbkę, 150 par zasad), koszty transportu próbek - 4724299 PLN</p> <p>Koszty oszacowano na podstawie kosztów pracy Partnera, regulaminu IHAR-PIB oraz ofert z rozeznania rynku.</p>

Nazwa pozycji kosztowej		Przewidywany koszt brutto	Uzasadnienie pozycji kosztowej (przeznaczenie)
Infrastruktura	brak	0,00 zł	Projekt zakłada wykorzystanie systemu ZUCH - system Zapewniania Usług Chmurowych, bez budowy lokalnej infrastruktury serwerowej,
Koszty UX i grafiki	Projektowanie interfejsu i doświadczeń użytkownika (UX/UI) systemu CBGR oraz badania użyteczności z użytkownikami końcowymi. Zgodność z WCAG 2.2 AA.	285 047,00 zł	<ul style="list-style-type: none"> - Makiety lo-fi (wireframes – min. 45 ekranów) i prototypy hi-fi w Figma - 98131 PLN - Testy użyteczności z użytkownikami końcowymi (2 rundy × 6 uczestników, sesje moderowane) - 34579 PLN - Projekty graficzne ekranów produkcyjnych (60-100 widoków, końcowa warstwa wizualna) - 34579 PLN - Badania użytkowników i opracowanie person (ok. 6 grup docelowych; wywiady, customer journey maps) - 28037 PLN - System projektowy / Design system (komponenty, design tokens, biblioteka UI) - 18692 PLN - Audyt dostępności WCAG 2.2 AA (raport + 2 rundy poprawek) - 14953 PLN - Raport końcowy z badań UX (produkt końcowy - KM obligatoryjny RA IT) - 9346 PLN - Dodatkowe iteracje poprawek graficznych - 46730 PLN <p>Koszty skalkuowano na podstawie kosztów pracy Partnera/Yellows Sp. z o.o.</p>
Bezpieczeństwo	Niezależne testy bezpieczeństwa systemu (SAST/DAST, pentesty zewnętrzne), DPIA wraz z RCPD oraz wdrożenie środków ochrony danych. Pozycja niezależna od testów wewnętrznych zespołu deweloperskiego	240 187,00 zł	<ul style="list-style-type: none"> - Testy penetracyjne aplikacji webowej i API - 2 iteracje + retest (podmiot zewnętrzny CREST / OSCP) - 65421 PLN - Dodatkowe iteracje remediacji, aktualizacje SAST/DAST - 46729 PLN - Ocena Skutków dla Ochrony Danych (DPIA) wraz z RCPD - 26168 PLN - Centralne logowanie SIEM + playbook reagowania na incydenty - 23364 PLN - Wdrożenie SAST i DAST w pipeline CI/CD (SonarQube/

Nazwa pozycji kosztowej		Przewidywany koszt brutto	Uzasadnienie pozycji kosztowej (przeznaczenie)
			<p>Semgrep + OWASP ZAP) - 20561 PLN</p> <ul style="list-style-type: none"> - Konfiguracja WAF i ochrony anti-DDoS w środowisku chmurowym -18692 PLN - Polityka bezpieczeństwa systemu - 14953 PLN - Analiza zagrożeń i ocena ryzyka - 13084 PLN - Inicjalny test prywatności – Raport (produkt końcowy - 11215 PLN <p>Koszty skalkuowano na podstawie kosztów pracy Partnera/Yellows Sp. z o.o.</p>
Wydajność rozwiązań	Niezależne testy wydajnościowe i optymalizacja przetwarzania dużych zbiorów danych genomowych (5 TB). Pozycja niezależna od testów wewnętrznych zespołu deweloperskiego	214 953,00 zł	<ul style="list-style-type: none"> - Dodatkowe iteracje optymalizacji, testy regresji wydajności - 61682 PLN - Optymalizacja wydajności: cache (Redis), query tuning, kompresja odpowiedzi API - 35514 PLN - Testy wydajności bazy danych i repozytorium obiektowego (35 000 obiektów, 5 TB) - 29907 PLN - Testy obciążeniowe publicznego API: peak load, rate limiting, stress test, soak test - 28037 PLN - Implementacja skryptów testów dla 3 e-usług + scenariusza masowego pobierania plików genomowych - 20561 PLN - Konfiguracja monitoringu APM (Grafana, dashboardy SLA, alerting progowy, Zabbix) - 16822 PLN - Plan testów wydajności – kryteria akceptacji: latency p95 < 500 ms, 200 RPS przy 500 users, SLA- 11215 PLN - Raport końcowy z testów wydajności - 11215 PLN <p>Koszty skalkulowano na podstawie kosztów pracy Partnera/Yellows Sp. z o.o.</p>
Szkolenia	Program szkoleniowy dla administratorów IT	100 000,00 zł	<ul style="list-style-type: none"> - Materiały szkoleniowe – instrukcje, nagrania (5 modułów), ćwiczenia praktyczne - 28972

Nazwa pozycji kosztowej		Przewidywany koszt brutto	Uzasadnienie pozycji kosztowej (przeznaczenie)
	(3 osoby) i redaktorów zasobów IHAR-PIB (50 osób) wraz z materiałami szkoleniowymi i wsparciem poszkoleniowym. Podstawa do osiągnięcia KPI 2.3		<p>PLN</p> <ul style="list-style-type: none"> - Wsparcie poszkoleniowe – helpdesk 3 miesiące (~2 h/tydz.) - 26168 PLN - Koszty logistyczne szkoleń w 4 oddziałach IHAR-PIB (Radzików, Bonin, Bydgoszcz, Młochów): transport, noclegi specjalistów, delegacje - 18692 PLN - Program szkoleniowy dla administratorów systemu – min. 3 osoby IT, 2 dni (platforma Chmurowa, backupy, monitoring, incydenty) - 9346 PLN - Raport ze szkoleń użytkowników systemu - 9346 PLN - Program szkoleniowy dla redaktorów zasobów - 7477 PLN - Koszty skalkuowano na podstawie kosztów pracy Yellow's i kosztów dojazdów i noclegów - szkolenia w Oddziałach IHAR-PIB.
Działania informacyjno-promocyjne	Działania informacyjno-promocyjne obejmujące opracowanie i realizację kampanii informacyjnej, przygotowanie materiałów promocyjnych, udział w wydarzeniach naukowych i branżowych oraz organizację spotkań informacyjnych dotyczących projektu i jego rezultatów w kraju i na świecie.	300 000,00 zł	<ul style="list-style-type: none"> - Udział w 6 konferencjach naukowych krajowych i zagranicznych (10 osob łącznie: opłaty konferencyjne, podróże, noclegi, materiały konferencyjne) - 168225 PLN - Spotkanie informacyjno-promocyjne CBGR (wynajem sali, catering, materiały, prelegenci) - 56075 PLN - Materiały promocyjne i informacyjne (broszury, ulotki, plakaty, gadżety projektowe) - 32710 PLN - Usługi graficzne i poligraficzne (identyfikacja wizualna, skład materiałów, druk) - 23364 PLN - Tablice i banery informacyjne projektu – wymóg FERC dot. promocji środków UE - 14018 PLN - Strona promocyjna projektu i obecność w mediach społecznościowych - 5608 PLN
Koszty zarządzania i wsparcia (w tym wynagrodzenia)	1. Zarządzanie projektem IT - kierownictwo i	2 058 330,00 zł	<p>1. Pozycja obejmuje:</p> <ul style="list-style-type: none"> - Kierownik projektu IT (39 mies., ok. 0,5 FTE): nadzór

Nazwa pozycji kosztowej		Przewidywany koszt brutto	Uzasadnienie pozycji kosztowej (przeznaczenie)
personelu wspomagającego)	<p>wsparcie</p> <p>2. Zarządzanie projektem w IHAR-PIB</p> <p>3. Koszy analizy finansowej do studium wykonalności</p> <p>4. Koszty pośrednie</p>		<p>harmonogramu, budżetu, ryzyk, raportowanie do Beneficjenta - 448598 PLN</p> <p>- Scrum Master (30 mies., fazy I-III wytwarzania): planowanie sprintów, retro, usuwanie blokerów - 121495 PLN</p> <p>- Analityk biznesowy / Product Owner (24 mies., fazy I-II): priorytetyzacja backlogu, prioryteżacja baclogu, kryteria odbioru kryteria odbioru - 98130 PLN</p> <p>- Wsparcie po wdrożeniu i formalne przekazanie systemu do utrzymania IHAR-PIB (60 dni) - 23 366</p> <p>2. Pozycja obejmuje</p> <p>- Kierownik projektu (39 mies., ok. 0,3 FTE): działania związane z planowaniem, koordynacją, monitoringiem postępu, kontrolą harmonogramu, nadzorem nad budżetem, przygotowaniem dokumentacji, raportowaniem, obsługą zamówień, rozliczaniem projektu oraz komunikacją z partnerem, wykonawcami i instytucjami zaangażowanymi w ocenę oraz finansowanie projektu - 249480 PLN</p> <p>3. Koszty przygotowania analizy finansowej przez firmę zewnętrzną - 10280 PLN</p> <p>4. Koszty pośrednie projektu - 7% kosztów bezpośrednich zgodnie z dokumentacją FERC - 1 106 981 PLN</p> <p>Koszty oszacowano na podstawie kosztów pracy Partnera, regulaminu IHAR-PIB oraz ofert z rozeznania rynku.</p>

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

Całkowity koszt utrzymania trwałości projektu (brutto)	1 240 000,00 zł		Źródło finansowania
Podział całkowitego kosztu utrzymania trwałości projektu na poszczególne lata (netto oraz brutto)	2030	248 000,00 zł (brutto) (201 626,02 zł netto)	krajowe środki publiczne - budżet państwa
	2031	248 000,00 zł (brutto) (201 626,02 zł netto)	krajowe środki publiczne - budżet państwa
	2032	248 000,00 zł (brutto) (201 626,02 zł netto)	krajowe środki publiczne - budżet państwa
	2033	248 000,00 zł (brutto) (201 626,02 zł netto)	krajowe środki publiczne - budżet państwa
	2034	248 000,00 zł (brutto) (201 626,02 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
Niewystarczająca jakość, kompletność i standaryzacja danych źródłowych	Duża	Niskie	Przeprowadzenie audytu danych przed wdrożeniem, opracowanie reguł walidacji i czyszczenia danych, etapowe ładowanie zasobów do repozytorium, kontrola jakości po każdej migracji oraz wyznaczenie właścicieli danych odpowiedzialnych za akceptację rekordów.
Niepełna lub błędna analiza	Duża	Średnie	Realizacja szczegółowej analizy przedwdrożeniowej z udziałem

Nazwa ryzyka	Siła oddziaływania	Prawdopodobieństwo wystąpienia ryzyka	Sposób zarządzania ryzykiem
wymagań funkcjonalnych i technicznych			użytkowników końcowych, specjalistów merytorycznych, IT i prawnych; stosowanie prototypów, warsztatów walidacyjnych i odbiorów etapowych; zarządzanie zmianą wymagań w trybie kontrolowanym.
Problemy z integracją i migracją danych z systemów źródłowych	Duża	Średnie	Wczesne przygotowanie mapowania danych, testowe migracje na próbach reprezentatywnych, wdrożenie środowiska testowego, dokumentacja interfejsów oraz rezerwa czasowa na korekty integracyjne.
Opóźnienia harmonogramu wynikające z niedostępności wykonawców lub kluczowych zasobów kadrowych	Duża	Średnie	Podział projektu na mierzalne etapy i kamienie milowe, bieżący monitoring postępu, jasny podział odpowiedzialności, plan zastępstw dla kluczowych ról oraz uruchamianie działań naprawczych po przekroczeniu progów ostrzegawczych.
Ryzyko incydentów bezpieczeństwa lub awarii środowiska wdrożeniowego	Duża	Niskie	Zastosowanie środowisk separowanych, kopii zapasowych, kontroli dostępu, szyfrowania, testów bezpieczeństwa, monitoringu zdarzeń oraz procedur odtworzeniowych. Działania te ograniczą ryzyko utraty danych i opóźnień wdrożenia.

5.2. Ryzyka wpływające na utrzymanie efektów

Nazwa ryzyka	Siła oddziaływania	Prawdopodobieństwo wystąpienia ryzyka	Sposób zarządzania ryzykiem
Ryzyko technologiczne (cyberbezpieczeństwo, awarie, utrata dostępności usług)	Duża	Niskie	Utrzymanie systemu w stabilnym środowisku chmurowym, regularne aktualizacje, monitoring bezpieczeństwa i wydajności, kopie zapasowe, procedury odtworzeniowe, testy przywracania oraz okresowe audyty bezpieczeństwa.
Ryzyko niewystarczającego finansowania utrzymania i rozwoju	Duża	Średnie	Zapewnienie źródeł finansowania trwałości w planie finansowym instytucji, uwzględnienie kosztów utrzymania systemu, hostingu, wsparcia technicznego i aktualizacji w budżetach

Nazwa ryzyka	Siła oddziaływania	Prawdopodobieństwo wystąpienia ryzyka	Sposób zarządzania ryzykiem
powdrożeniowego			po zakończeniu projektu oraz przypisanie odpowiedzialności organizacyjnej za utrzymanie rozwiązania.
Ryzyko niskiego zainteresowania użytkowników i niewystarczającego o wykorzystania zasobów	Średnia	Średnie	Prowadzenie aktywnych działań informacyjno-promocyjnych, szkoleń i warsztatów dla grup docelowych, bieżące badanie potrzeb użytkowników, rozwój funkcjonalności zgodnie z informacją zwrotną oraz monitorowanie statystyk użycia systemu.
Ryzyko regulacyjno-prawne ograniczające zakres udostępniania lub ponownego wykorzystania danych	Duża	Średnie	Stałe monitorowanie zmian legislacyjnych i szybkie dostosowanie systemu do nowych wymogów prawnych. Ścisła współpraca z odpowiednimi ministerstwami i instytucjami regulacyjnymi w zakresie interpretacji przepisów i dostosowania repozytorium do wymagań prawnych (RODO, protokół z Nagoi, ITPGRFA).
Ryzyko dezaktualizacji danych i osłabienia wiarygodności repozytorium	Duża	Znikome	Wprowadzenie procedur cyklicznej aktualizacji danych, walidacji nowych rekordów, przeglądu jakości metadanych i odpowiedzialności właścicieli danych za utrzymanie aktualności zasobu.
Ryzyko organizacyjne związane z utratą kompetencji personelu po zakończeniu projektu	Średnia	Średnie	Wyznaczenie stałych ról odpowiedzialnych za administrację, rozwój i utrzymanie systemu, przygotowanie dokumentacji technicznej i użytkowej, szkolenie personelu oraz zapewnienie przekazania wiedzy od zespołu wdrożeniowego do zespołu utrzymaniowego.

6. OTOCZENIE PRAWNE

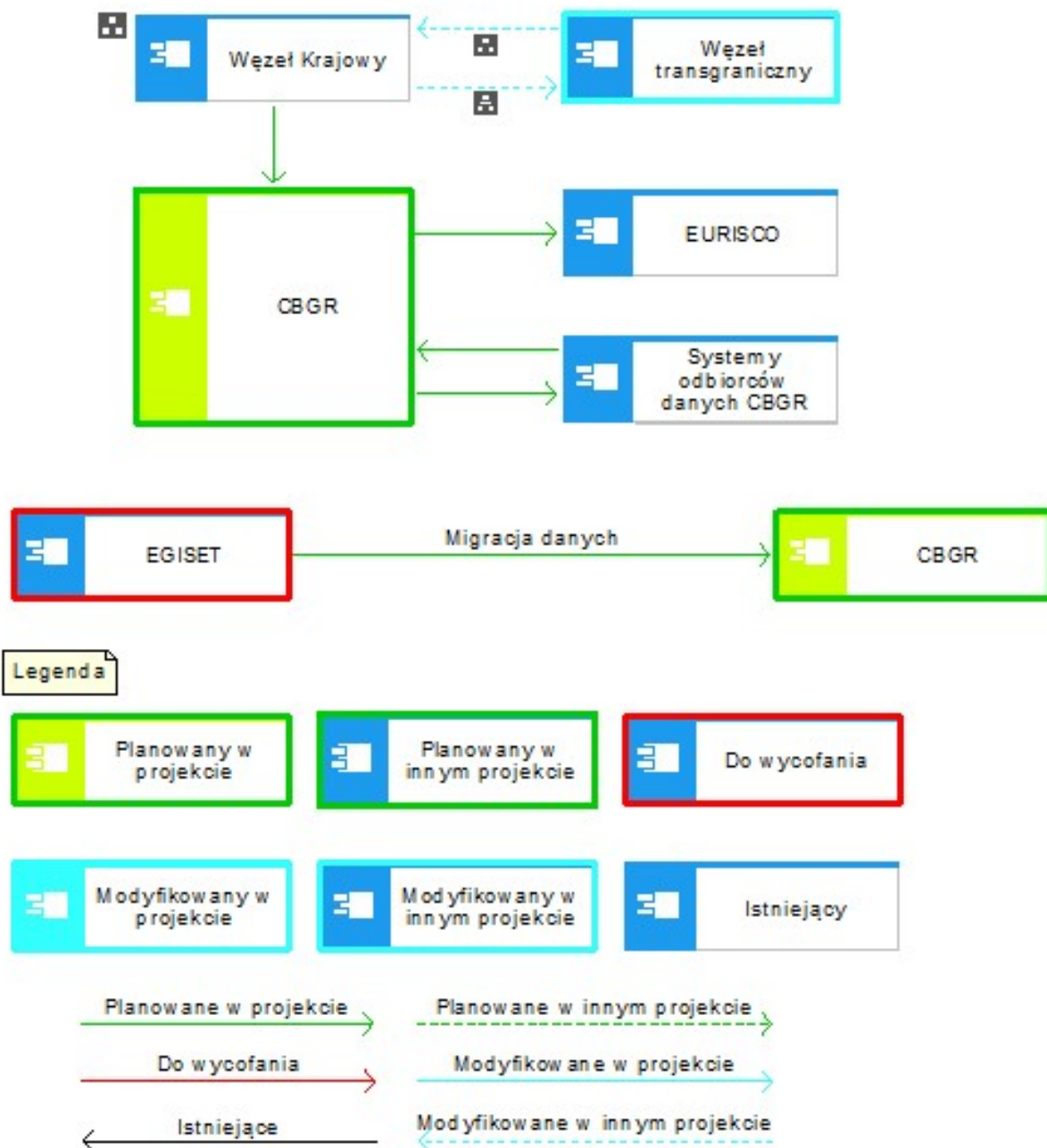
Lp.	Tytuł aktu prawnego	Czy wymaga zmian	Opis zmian (jeśli dotyczy)	Etap prac legislacyjnych (jeśli dotyczy)
1	Ustawa z dnia 11 sierpnia 2021 r. o otwartych danych i ponownym wykorzystywaniu	TAK/NIE		

Lp.	Tytuł aktu prawnego	Czy wymaga zmian	Opis zmian (jeśli dotyczy)	Etap prac legislacyjnych (jeśli dotyczy)
	informacji sektora publicznego			
2	Dyrektywa Parlamentu Europejskiego i Rady (UE) 2019/1024 z dnia 20 czerwca 2019 r. w sprawie otwartych danych i ponownego wykorzystywania informacji sektora publicznego	TAK/NIE		
3	Ustawa z dnia 17 lutego 2005 r. o informatyzacji działalności podmiotów realizujących zadania publiczne	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	TAK/NIE		
5	Ustawa z dnia 5 lipca 2018 r. o krajowym systemie cyberbezpieczeństwa	TAK/NIE		
6	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	TAK/NIE		
7	Ustawa z dnia 19 lipca 2016 r. o dostępie do zasobów genetycznych i podziale korzyści z ich wykorzystania	TAK/NIE		
8	Rozporządzenie Parlamentu Europejskiego i Rady (UE) nr 511/2014 z dnia 16 kwietnia 2014 r. w sprawie środków zapewniających zgodność użytkowników w Unii z wymogami wynikającymi z Protokołu z Nagoi dotyczącego dostępu do zasobów genetycznych oraz sprawiedliwego i równego podziału korzyści wynikających z wykorzystania tych zasobów w Unii	TAK/NIE		
9	Konwencja o różnorodności biologicznej, sporządzona w Rio de Janeiro dnia 5 czerwca 1992	TAK/NIE		
10	Protokół z Nagoi do Konwencji o różnorodności biologicznej dotyczący dostępu do zasobów genetycznych oraz sprawiedliwego i równego podziału korzyści wynikających z wykorzystania tych zasobów	TAK/NIE		
11	Międzynarodowy Traktat o zasobach genetycznych roślin dla żywienia i rolnictwa,	TAK/NIE		

Lp.	Tytuł aktu prawnego	Czy wymaga zmian	Opis zmian (jeśli dotyczy)	Etap prac legislacyjnych (jeśli dotyczy)
	sporządzony w Rzymie dnia 3 listopada 2001			
12	Ustawa z dnia 27 lipca 2001 r.o ochronie baz danych	TAK/NIE		
13	Ustawa z dnia 4 kwietnia 2019 r. o dostępności cyfrowej stron internetowych i aplikacji mobilnych podmiotów publicznych	TAK/NIE		
14	Ustawa z dnia 7 maja 2010 r. o wspieraniu rozwoju usług i sieci telekomunikacyjnych	TAK/NIE		
15	Ustawa z dnia 18 listopada 2020 r.o doręczeniach elektronicznych	TAK/NIE		
16	Ustawa z dnia 5 września 2016 r. o usługach zaufania oraz identyfikacji elektronicznej	TAK/NIE		
17	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		
18	Rozporządzenie Prezesa Rady Ministrów z dnia 20 lipca 2011 r. w sprawie podstawowych wymagań bezpieczeństwa teleinformatycznego	TAK/NIE		
19	Rozporządzenie Prezesa Rady Ministrów z dnia 14 września 2011 r. w sprawie sporządzania pism w formie dokumentów elektronicznych, doręczania dokumentów elektronicznych oraz udostępniania formularzy, wzorów i kopii dokumentów elektronicznych	TAK/NIE		
20	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		
21	Rozporządzenie Ministra Cyfryzacji z dnia 29 czerwca 2020 r. w sprawie profilu zaufanego i podpisu zaufanego	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	Cyfrowy Bank Genów Roślin (CBGR)	Instytut Hodowli i Aklimatyzacji	Cyfrowy Bank Genów Roślin (CBGR) to system teleinformatyczny	Planowany	Nie dotyczy

Lp.	Nazwa systemu	Gestor systemu	Opis systemu	Status	Krótki opis ewentualnej zmiany
		<p>ji Roślin - Państwowy Instytut Badawczy</p>	<p>wspierający gromadzenie, walidację, zarządzanie i publiczne udostępnianie danych o roślinnych zasobach genowych Krajowego Centrum Roślinnych Zasobów Genowych (KCRZG).</p> <p>Celem systemu jest zapewnienie uporządkowanego, interoperacyjnego i maszynowo dostępnego źródła danych sekwencyjnych, paszportowych, metadanych i wybranych danych fenotypowych.</p> <p>System prowadzi rejestr publiczny danych o roślinnych zasobach genowych, obejmujący m.in. cyfrowe informacje o sekwencji DNA (DSI), identyfikatory obiektów, metadane, dane paszportowe oraz wybrane dane fenotypowe.</p> <p>Główne grupy funkcjonalności:</p> <ul style="list-style-type: none"> (1) repozytorium danych genomowych i metadanych (2) portal publiczny i wyszukiwarka zasobów (3) publiczne API danych/ API Gateway (4) moduł importu, walidacji i standaryzacji danych (5) moduł administracyjny i zarządzania uprawnieniami (6) moduł wymiany danych z systemami zewnętrznymi, 		

Lp.	Nazwa systemu	Gestor systemu	Opis systemu	Status	Krótki opis ewentualnej zmiany
			<p>(7) obsługa zamówień próbek materiału genetycznego.</p> <p>System integruje się z systemami zewnętrznymi poprzez API oraz kanały wymiany danych, w szczególności z EURISCO oraz systemami odbiorców danych CBGR</p>		
2	System informacyjny EGISET	IHAR-PIB	<p>System informacyjny EGISET to system teleinformatyczny wspierający ewidencję, wyszukiwanie i obsługę kolekcji zasobów genowych roślin KCRZG.</p> <p>Celem systemu jest gromadzenie i udostępnianie informacji ewidencyjnych o obiektach kolekcji oraz wsparcie procesów związanych z dokumentacją i obsługą udostępniania materiału biologicznego.</p> <p>System nie prowadzi rejestru publicznych danych genetycznych roślin ani repozytorium cyfrowych informacji o sekwencji DNA.</p> <p>Główne grupy funkcjonalności obejmują ewidencję obiektów kolekcji, wyszukiwanie zasobów, zarządzanie podstawowymi danymi paszportowymi i dokumentacją udostępniania materiału.</p> <p>System stanowi źródło danych historycznych i</p>	Istniejący	System jest wycofywany z użytkowania po zakończeniu jednorazowej migracji danych do systemu Cyfrowy Bank Genow Roslin

Lp.	Nazwa systemu	Gestor systemu	Opis systemu	Status	Krótki opis ewentualnej zmiany
			<p>ewidencyjnych przenoszonych do CBGR w ramach jednorazowej migracji. Po zakończeniu migracji nie jest elementem bieżącej architektury produkcyjnej CBGR.</p>		
3	EURISCO	<p>Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gaterslebe n - utrzymanie EURISCO na rzecz ECPG</p>	<p>EURISCO (European Search Catalogue for Plant Genetic Resources) to europejski katalog wyszukiwawczy roślinnych zasobów genowych wspierający publikowanie i wyszukiwanie informacji o obiektach przechowywanych w europejskich bankach genów i kolekcjach ex situ.</p> <p>Celem systemu jest zapewnienie wspólnego punktu dostępu do danych paszportowych o zasobach genetycznych roślin oraz wspieranie ich ponownego wykorzystania przez naukę, hodowlę, administrację i banki genów.</p> <p>System agreguje dane przekazywane przez krajowe banki genów i punkty kontaktowe.</p> <p>Główne grupy funkcjonalności obejmują przyjmowanie danych z krajowych banków genów, ich walidację, wyszukiwanie, prezentację i eksport informacji o obiektach.</p>	Istniejący	Nie dotyczy

Lp.	Nazwa systemu	Gestor systemu	Opis systemu	Status	Krótki opis ewentualnej zmiany
			Integracja CBGR z EURISCO polega na cyklicznym lub operacyjnym przekazywaniu danych paszportowych kolekcji w uzgodnionym standardzie		
4	Systemy odbiorców danych CBGR	System wielokrotny : systemy odbiorców danych CBGR w tym banki genów, instytuty badawcze, uczelnie, administracja publiczna, organizacje międzynarodowe oraz podmioty biotechnologiczne	<p>System wielokrotny - Systemy odbiorców danych Cyfrowego Banku Genów Roślin (CBGR) to system wielokrotny obejmujący zewnętrzne systemy teleinformatyczne podmiotów korzystających z danych CBGR, w tym banków genów, instytutów badawczych, uczelni, administracji publicznej, organizacji międzynarodowych oraz podmiotów biotechnologicznych.</p> <p>Celem tej grupy systemów jest automatyczne pobieranie, przetwarzanie i ponowne wykorzystanie danych sekwencyjnych, paszportowych, fenotypowych i metadanych udostępnianych przez CBGR.</p> <p>Systemy odbiorców nie prowadzą rejestru publicznego CBGR, lecz mogą wykorzystywać dane pochodzące z rejestru publicznego Cyfrowy Bank Genów Roślin.</p> <p>Główne funkcjonalności obejmują pobieranie</p>	Istniejący	Nie dotyczy

Lp.	Nazwa systemu	Gestor systemu	Opis systemu	Status	Krótki opis ewentualnej zmiany
			<p>danych przez API, import do lokalnych baz, integrację z własnymi zasobami, analizy bioinformatyczne, raportowanie i wizualizację danych. Integracja z CBGR będzie realizowana przez publiczne API, mechanizmy pobierania zbiorów danych oraz standardowe formaty wymiany danych.</p>		
5	Węzeł Krajowy	Ministerstwo Cyfryzacji	<p>Węzeł Krajowy to system wspierający proces uwierzytelniania użytkowników korzystających z usług online świadczonych przez podmioty publiczne. System pełni rolę pośrednika pomiędzy systemami identyfikacji elektronicznej a systemami udostępniającymi usługi elektroniczne, umożliwiając bezpieczne przekazywanie potwierdzeń tożsamości. Celem systemu jest zapewnienie jednolitego, bezpiecznego i interoperacyjnego mechanizmu uwierzytelniania użytkowników usług online administracji publicznej poprzez wykorzystanie różnych środków identyfikacji elektronicznej.</p> <p>Główne grupy funkcjonalności</p> <ul style="list-style-type: none"> – Pośredniczenie w procesie uwierzytelniania użytkowników usług online 	Istniejący	Nie dotyczy

Lp.	Nazwa systemu	Gestor systemu	Opis systemu	Status	Krótki opis ewentualnej zmiany
			<ul style="list-style-type: none"> - Integracja z systemami identyfikacji elektronicznej - Przekazywanie potwierdzeń tożsamości do systemów świadczących usługi elektroniczne - Zapewnienie bezpiecznej wymiany danych identyfikacyjnych Integracje System jest zintegrowany z krajowymi systemami teleinformatycznymi		
6	Węzeł transgraniczny	Ministerstwo Cyfryzacji	Węzeł Transgraniczny to krajowy system wspierający wzajemne uwierzytelnianie obywateli oraz podmiotów gospodarczych pomiędzy państwami członkowskimi Unii Europejskiej w dostępie do usług elektronicznych. System działa jako pośrednik pomiędzy krajowymi notyfikowanymi systemami identyfikacji elektronicznej a systemami udostępniającymi usługi online w innych państwach UE oraz pomiędzy krajowymi systemami udostępniającymi usługi online a notyfikowanymi systemami identyfikacji elektronicznej w innych krajach UE. Celem systemu jest zapewnienie interoperacyjności notyfikowanych systemów identyfikacji elektronicznej państw członkowskich Unii	Istniejący	Nie dotyczy

Lp.	Nazwa systemu	Gestor systemu	Opis systemu	Status	Krótki opis ewentualnej zmiany
			<p>Europejskiej oraz umożliwienie korzystania z nich w transgranicznie udostępnionych usługach online.</p> <p>Główne grupy funkcjonalności</p> <ul style="list-style-type: none"> - Pośredniczenie w transgranicznym procesie uwierzytelniania użytkowników usług elektronicznych - Przekazywanie potwierdzeń tożsamości pomiędzy krajowymi notyfikowanymi systemami identyfikacji elektronicznej a systemami udostępniającymi usługi online w innych państwach UE - Przekazywanie potwierdzeń tożsamości pomiędzy notyfikowanymi systemami identyfikacji elektronicznej w innych krajach UE a krajowymi systemami udostępniającymi usługi online - Przekazywanie potwierdzeń tożsamości pomiędzy systemami krajowymi - Zapewnienie interoperacyjnej i bezpiecznej wymiany danych identyfikacyjnych <p>System jest zintegrowany z krajowymi oraz zagranicznymi systemami teleinformatycznymi</p>		

Lista przepływów

Lp.	System źródłowy	System docelowy	Zakres wymienianych danych	Sposób wymiany danych	Typ modyfikacji	Typ interfejsu
1	CBGR	Systemy odbiorców danych CBGR	Udostępniane (z CBGR do systemu partnerskiego) dane opisujące obiekty kolekcji oraz powiązane metadane.	tryb odwołań bezpośrednich	krytyczny dla sukcesu projektu	usługa REST
2	Węzeł Krajowy	CBGR	Potwierdzenie uwierzytelnienia użytkownika oraz minimalny zestaw danych identyfikacyjnych niezbędnych do zalogowania i obsługi czynności wymagających identyfikacji, np. identyfikator użytkownika, imię, nazwisko oraz identyfikator sesji/atributy potwierdzenia tożsamości.	tryb odwołań bezpośrednich	krytyczny dla sukcesu projektu	usługa REST
3	EGISET	CBGR	Jednorazowa migracja danych sekwencyjnych, metadanych obiektów oraz obiektów z dotychczasowego systemu ewidencyjnego do CBGR.	tryb odwołań bezpośrednich	krytyczny dla sukcesu projektu	ETL / import danych z systemu źródłowego
4	CBGR	EURISCO	Udostępniane (z CBGR do EURISCO) dane	tryb odwołań bezpośrednich	krytyczny dla sukcesu projektu	usługa REST

Lp.	System źródłowy	System docelowy	Zakres wymienianych danych	Sposób wymiany danych	Typ modyfikacji	Typ interfejsu
			opisujące obiekty kolekcji oraz powiązane metadane.			
5	Systemy odbiorców danych CBGR	CBGR	Udostępniane (z systemu partnerskiego do CBGR) dane opisujące obiekty kolekcji oraz powiązane metadane.	tryb odwołań bezpośrednich	krytyczny dla sukcesu projektu	usługa REST
6	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 umożliwiających korzystanie ze środków identyfikacji elektronicznej przyłączonych do Węzła Krajowego. Zakres danych zależy od danych w środku identyfikacji	tryb odwołań bezpośrednich	krytyczny dla sukcesu projektu	usługa REST

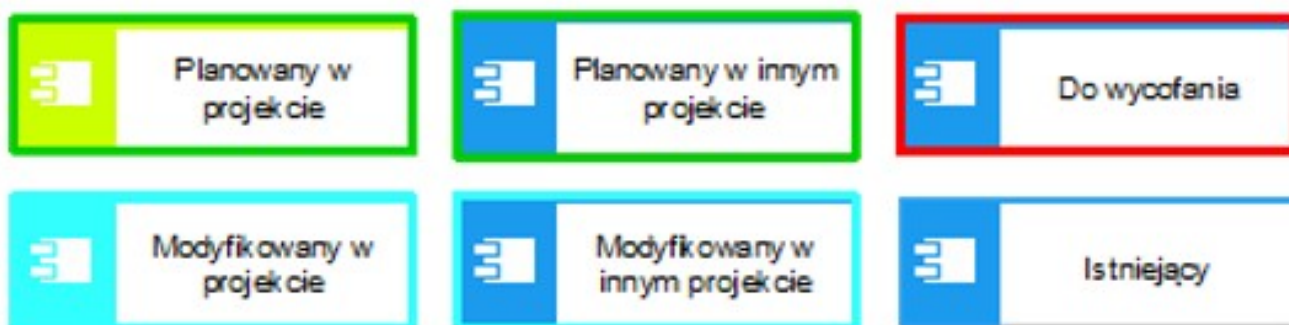
Lp.	System źródłowy	System docelowy	Zakres wymienianych danych	Sposób wymiany danych	Typ modyfikacji	Typ interfejsu
			<p>elektronicznej.</p> <p>Aktualnie Węzeł krajowy i transgraniczny przekazuje tylko zakres danych obowiązkowych:</p> <ul style="list-style-type: none"> * Imię * Nazwisko * Identyfikator * Data urodzenia <p>Dane mogą być wykorzystywane biznesowo, np. przez zagraniczne systemy udostępniające usługi online.</p>			
7	Węzeł transgraniczny	Węzeł Krajowy	<p>Obsługa żądania identyfikacji (z PL do UE). Dane identyfikujące osobę fizyczną (w szczególności cudzoziemca), w kontekście środka identyfikacji elektronicznej wystawionego przez system podłączony do węzła transgranicznego (w szczególności zagraniczny system identyfikacji</p>	tryb odwołań bezpośrednich	krytyczny dla sukcesu projektu	usługa REST

Lp.	System źródłowy	System docelowy	Zakres wymienianych danych	Sposób wymiany danych	Typ modyfikacji	Typ interfejsu
			<p>elektronicznej) przekazywane za pośrednictwem Węzła Transgranicznego przez Węzeł Krajowy do krajowych systemów udostępniających usługi online. Zakres danych zależy od danych w zagranicznym źródle identyfikacji elektronicznej.</p> <p>Aktualnie Węzeł transgraniczny przekazuje tylko zakres danych obowiązkowych:</p> <ul style="list-style-type: none"> * Imię * Nazwisko * Identyfikator * Data urodzenia <p>Mogą być wykorzystywane biznesowo, np. w systemach udostępniających usługi online, np. do wypełnienia wniosku.</p>			

7.2. Kluczowe komponenty architektury rozwiązania



Legenda



7.3. Przyjęte założenia technologiczne

Lp.	Obszar	Założenie technologiczne
1.	Infrastruktura	System Cyfrowy Bank Genów Roślin zostanie zbudowany jako rozwiązanie chmurowe utrzymywane w środowisku państwowej chmury obliczeniowej ZUCH lub równoważnym środowisku chmury spełniającym wymagania podmiotów publicznych. Architektura będzie oparta na separowanych środowiskach DEV/TEST /PROD, konteneryzacji usług aplikacyjnych oraz logicznym podziale na warstwę publicznego portalu, warstwę API, warstwę administracyjną, repozytorium danych i warstwę integracyjną. Dla zapewnienia skalowalności i ciągłości działania przewiduje się wykorzystanie elastycznych zasobów obliczeniowych, pamięci

Lp.	Obszar	Założenie technologiczne
		masowej dla dużych plików genomowych, mechanizmów backupu, odtwarzania po awarii, monitoringu oraz redundancji kluczowych komponentów. Dane źródłowe z systemu EGISET zostaną zmigrowane do nowego systemu, a EGISET po migracji nie będzie elementem bieżącej architektury produkcyjnej Cyfrowego Banku Genów Roślin.
2.	Sieć i bezpieczeństwo	System zostanie wdrożony w architekturze sieciowej z wydzielonymi strefami: publiczna, aplikacyjna oraz danych. W strefie publicznej będą udostępniane portal publiczny i publiczne API, natomiast strefa aplikacyjna i strefa danych będą dostępne wyłącznie dla uprawnionych komponentów systemu i administratorów. Komunikacja zewnętrzna i wewnętrzna będzie szyfrowana z wykorzystaniem TLS 1.2/1.3. Dostęp administracyjny będzie realizowany wyłącznie przez bezpieczne kanały komunikacji i ograniczony do uprawnionych użytkowników. Wdrożone zostaną mechanizmy kontroli dostępu RBAC, MFA dla kont administracyjnych, rejestrowania i audytu zdarzeń, centralnego logowania, monitoringu dostępności i bezpieczeństwa, backupów oraz procedur odtwarzania. Dla publicznego API przewiduje się limity zapytań, mechanizm kluczy API lub tokenów dostępowych, ochronne przed nadużyciem oraz walidację danych wejściowych. Rozwiązanie będzie projektowane z uwzględnieniem wymagań KRI, RODO, ustawy o krajowym systemie cyberbezpieczeństwa, wymagań dostępności cyfrowej oraz dobrych praktyk OWASP.
3.	Standardy wymiany danych	Integracja i udostępnianie danych będą realizowane przez REST API opisane w standardzie Open API/Swagger. Podstawowym formatem wymiany metadanych będzie JSON, a dla eksportów i ponownego wykorzystania danych przewiduje się również CSV oraz formaty właściwe dla danych genomowych, w szczególności FASTA, FASTQ i VCF oraz raporty jakościowe. Dane będą opisywane zestawem metadanych umożliwiających wyszukiwanie, filtrowanie, cytowanie i ponowne wykorzystanie zgodnie z zasadami FAIR oraz wymaganiami otwartych danych. API będzie wspierać paginację, filtrowanie, wersjonowanie, autoryzację dostępu i mechanizmy limitowania zapytań. Dokumentacja API będzie dostępna dla użytkowników technicznych i podmiotów ponownie wykorzystujących dane.
4.	Systemy operacyjne serwerowe	Środowiska serwerowe systemu będą oparte na systemach Linux klasy enterprise lub LTS, np. Ubuntu Server LTS, Debian albo równoważnych dystrybucjach wspieranych przez dostawcę środowiska chmurowego. Systemy operacyjne będą utrzymywane w aktualnych wersjach objętych wsparciem producenta, z regularnymi aktualizacjami bezpieczeństwa, centralnym zarządzaniem konfiguracją, ograniczeniem uruchomionych usług do niezbędnego minimum, monitoringiem zasobów i logowaniem zdarzeń.
5.	Bazy danych	Podstawowym rozwiązaniem bazodanowym systemu będzie relacyjna baza danych MariaDB w aktualnej, wspieranej wersji LTS lub rozwiązanie równoważne (PostgreSQL). Baza relacyjna

Lp.	Obszar	Założenie technologiczne
		będzie służyła do przechowywania metadanych, danych paszportowych, danych fenotypowych, informacji o użytkownikach i uprawnieniach, konfiguracji systemu, danych rejestrowych oraz referencji do plików przechowywanych poza bazą danych. Wyszukiwanie i filtrowanie danych publicznych będą wspierane przez mechanizmy indeksowania i wyszukiwania pełnotekstowego. W przypadku potrzeby zwiększenia wydajności systemu możliwe będzie zastosowanie mechanizmu pamięci podręcznej (np. Redis lub rozwiązania równoważnego).
6.	Serwery aplikacji	Warstwa aplikacyjna zostanie zbudowana jako zestaw konteneryzowanych (Docker) usług uruchamianych w środowisku chmurowym i orkiestrowanych. Serwerem HTTP/ reverse proxy będzie Apache HTTP Server lub Nginx albo rozwiązanie równoważne. Logika aplikacyjna będzie uruchamiana w środowisku właściwym dla przyjętej technologii backendowej (Java/Spring Boot lub Python/FastAPI albo równoważne), zaproponowanej przez wykonawcę i zaakceptowanej na etapie projektu technicznego. Usług aplikacyjne będą projektowane jako bezstanowe, możliwe do skalowania horyzontalnego. Wdrożenia będą realizowane przez proces CI/CD z automatycznymi testami, analiza jakości kodu, kontrola wersji i możliwością powtarzalnego odtworzenia środowisk.
7.	Portale	System będzie udostępniany użytkownikom przez publiczny portal internetowy Cyfrowego Banku Genów Roślin, obejmujący wyszukiwarkę zasobów, podgląd metadanych oraz funkcje pobierania danych. Platforma będzie zawierać również panel administracyjny dla uprawnionych pracowników Beneficjenta, umożliwiającą zarządzanie zasobami danych, walidację metadanych, publikację danych, zarządzanie uprawnieniami oraz monitorowanie procesów importu i udostępniania. Warstwa portalowa zostanie wykonana jako responsywna aplikacja webowa zgodna z wymaganiami dostępności cyfrowej WCAG 2.2 AA, z wykorzystaniem technologii HTML5, CSS3 oraz JavaScript/ TypeScript (React.js lub Vue.js albo równoważnych). Portal będzie zintegrowany z API oraz repozytorium danych systemu Cyfrowy Bank Genów Roślin.
8.	Inne	

7.4. Opis zasobów danych przetwarzanych w planowanym rozwiązaniu

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

Lp.	Tworzony rejestr publiczny	Opis
1	Cyfrowy Bank Genów Roślin	Rejestr Cyfrowy Bank Genów Roślin obejmuje dane o roślinnych zasobach genowych KCRZG, w tym cyfrowe informacje o

Lp.	Tworzony rejestr publiczny	Opis
		sekwencji DNA, identyfikatory obiektów, metadane, dane paszportowe oraz wybrane dane fenotypowe. Rejestr służy publicznemu udostępnianiu informacji sektora publicznego o charakterze naukowym oraz umożliwia ich ponowne wykorzystanie przez użytkowników końcowych i systemy zewnętrzne. Dane w rejestrze będą udostępniane on-line poprzez portal, wyszukiwarkę oraz API.

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~~