

**Raport z postępu rzeczowo-finansowego projektu informatycznego  
za IV kwartał 2020 roku**

<b>Tytuł projektu</b>	Dziedzinowe Repozytoria Otwartych Danych Badawczych
<b>Wnioskodawca</b>	Minister Nauki i Szkolnictwa Wyższego
<b>Beneficjent</b>	Uniwersytet Warszawski
<b>Partnerzy</b>	Instytut Filozofii i Socjologii Polskiej Akademii Nauk, Uniwersytet im. Adama Mickiewicza w Poznaniu
<b>Źródło finansowania</b>	Program Operacyjny Polska Cyfrowa na lata 2014-2020 Oś Priorytetowa nr 2 „E-administracja i otwarty rząd” Działanie nr 2.3 „Cyfrowa dostępność i użyteczność in- formacji sektora publicznego” Poddziałanie nr 2.3.1 „Cyfrowe udostępnienie informacji sektora publicznego ze źródeł administracyjnych i zasobów nauki (typ projektu: cyfrowe udostępnienie zasobów nauki)” Budżet państwa: część budżetowa 27 - Informatyzacja
<b>Całkowity koszt projektu</b>	4 998 889,- PLN
<b>Całkowity koszt projektu - wydatki kwalifikowalne</b>	4 998 889,- PLN
<b>Okres realizacji projektu</b>	01.08.2018 – 29.10.2021 (wcześniej 01.08.2018 - 31.07.2021)

### 1. Otoczenie prawne <maksymalnie 1000 znaków>

Kwestia otwartych danych badawczych znajduje odzwierciedlenie w dokumentach zarówno na poziomie Unii Europejskiej (*Zalecenie KE w sprawie dostępu do informacji naukowej oraz jej ochrony* oraz wymogi dotyczące deponowania danych badawczych w ramach programu Horyzont 2020), jak i poziomie krajowym (przyjęte przez Ministerstwo Nauki i Szkolnictwa Wyższego w 2015 r. „Kierunki rozwoju otwartego dostępu do publikacji i wyników badań naukowych w Polsce”). Sprawia to, iż projekt wpisuje się istniejące tendencje dotyczące prawnych zaleceń i wymogów dotyczących otwierania danych badawczych.

## 2. Postęp finansowy

Czas realizacji projektu	Wartość środków wydatkowanych	Wartość środków zaangażowanych
74,36 %	1. 63,28 % *  2. 63,28 %          *Wg sprawozdanych wniosków o płatność zawierających wydatki do 11.2020	79,00 %

### 3. Postęp rzeczowy <maksymalnie 5000 znaków>

## Kamienie milowe

Nazwa	Powiązane wskaźniki projektu	Planowany termin osiągnięcia	Rzeczywisty termin osiągnięcia	Status realizacji kamienia milowego
Wersja wstępna wymagań funkcjonalnych i schematu metadanych		11-2018	11-2018	Osiągnięty

Nazwa	Powiązane wskaźniki projektu	Planowany termin osiągnięcia	Rzeczywisty termin osiągnięcia	Status realizacji kamienia milowego
Wersja finalna wymagań funkcjonalnych i schematu metadanych		10-2019	10-2019	Osiągnięty
Opracowanie procedur długoterminowego przechowywania		04-2019	05-2019	Osiągnięty. Przyczyna przekroczenia planowanego terminu osiągnięcia: konieczność przeprowadzenia dodatkowych konsultacji i ustaleń mających na celu pełne uzgodnienie procedur długoterminowego przechowywania, archiwizacji i zapewnienia bezpieczeństwa danych opracowanych w ramach projektu z istniejącymi procedurami instytucjonalnymi w ICM.
Wdrożenie procedur i narzędzi długoterminowego przechowywania		12-2020	12-2020	Osiągnięty
Zakończenie prac nad oprogramowaniem repozytorium ogólnego	4 (3 szt. łącznie z kamieniem milowym „Zakończenie prac nad oprogramowaniem repozytoriów dziedzinowych)	03-2020	03-2020	Osiągnięty
Zakończenie prac nad oprogramowaniem repozytoriów dziedzinowych	4 (3 szt. łącznie z kamieniem milowym „Zakończenie prac nad oprogramowaniem repozytorium ogólnego”)	06-2020	06-2020	Osiągnięty
Opracowanie danych krystalograficznych		08-2020	09-2020	Osiągnięty. Przyczyna przekroczenia planowanego terminu osiągnięcia: problemy techniczne na ostatnich etapach procesu opracowywania danych krystalograficznych, związane z pracą zdalną na skutek sytuacji epidemiologicznej.

Nazwa	Powiązane wskaźniki projektu	Planowany termin osiągnięcia	Rzeczywisty termin osiągnięcia	Status realizacji kamienia milowego
Opracowanie zasobu danych społecznych	2 (1 szt.)	09-2021		W trakcie realizacji
Uruchomienie repozytoriów		04-2019	05-2019	Osiągnięty. Przyczyna przekroczenia planowanego terminu osiągnięcia: dłuższy niż pierwotnie zakładano czas niezbędnych prac w zakresie adaptacji i rozwoju oprogramowania repozytoryjnego Dataverse, których wykonanie było koniecznym warunkiem uruchomienia repozytoriów.
Raport z funkcjonowania repozytoriów	1 (3 szt.) 3 (620 szt.) 5 (3 szt.)	09-2021		Planowany
Uruchomienie strony internetowej Projektu		10-2018	10-2018	Osiągnięty
Przeprowadzenie szkoleń		09-2021		W trakcie realizacji
Finalne posiedzenie komitetu sterującego		09-2021		Planowany
Nabycie środków trwałych		11-2018	10-2019	Osiągnięty. Przyczyna przekroczenia planowanego terminu osiągnięcia: przedłużająca się procedura zakupu środków trwałych oraz wyższa niż zakładana we wniosku o dofinansowanie cena środków trwałych (laptopy).

**Wskaźniki efektywności projektu (KPI)**

<b>Nazwa</b>	<b>Jedn. miary</b>	<b>Wartość docelowa</b>	<b>Planowany termin osiągnięcia</b>	<b>Wartość osiągnięta od początku realizacji projektu (nastająco)</b>
Liczba podmiotów, które udostępniły on-line informacje sektora publicznego	szt.	3,00	09-2021	0
Liczba zdigitalizowanych dokumentów zawierających informacje sektora publicznego	szt.	1,00	09-2021	0
Liczba udostępnionych on-line dokumentów zawierających informacje sektora publicznego	szt.	620,00	09-2021	0
Liczba utworzonych API	szt.	3,00	06-2020	3,00
Liczba baz danych udostępnionych on-line poprzez API	szt.	3,00	09-2021	0
Liczba pobrań/ odtworzeń dokumentów zawierających informacje sektora publicznego	szt./rok	37 200,00	10-2022	0
Rozmiar udostępnionych on-line informacji sektora publicznego	TB	3,36	09-2021	0
Liczba zbiorów udostępnionych na wolnych licencjach	szt.	220	09-2021	0

#### 4. E-usługi A2A, A2B, A2C <maksymalnie 2000 znaków>

Nazwa	Plano- wana data wdroże- nia	Rzeczy- wista data wdroże- nia	Opis zmian
nie dotyczy			

#### 5. Udostępnione informacje sektora publicznego i zdigitalizowane zasoby <maksymalnie 2000 znaków>

Nazwa	Plano- wana data wdroże- nia	Rzeczy- wista data wdroże- nia	Opis zmian
620 zbiorów danych badawczych	06-2021		

#### 6. Produkty końcowe projektu (inne niż wskazane w pkt 4 i 5) <maksymalnie 2000 znaków>

Nazwa produktu	Planowana data wdroże- nia	Rzeczywista data wdrożenia	Komplementarność względem produktów innych projektów
Repozytorium ogólnego przeznaczenia, repozytorium danych społecznych, repozytorium danych krystalograficznych	06-2021		

#### 7. Ryzyka <maksymalnie 2000 znaków>

Ryzyka wpływające na realizację projektu

Nazwa ryzyka	Siła oddziaływania	Prawdopodobieństwo wystąpienia ryzyka	Sposób zarządzania ryzykiem
Brak wprowadzenia polityki otwartych danych przez MNiSW na szczeblu krajowym.	mała	wysokie	<p>Projekt przewiduje własną działalność informacyjną i edukacyjną wśród naukowców, dzięki której minimalizowane jest ryzyko, że zainteresowanie naukowców udostępnianiem danych będzie niewielkie - nawet gdyby MNiSW nie prowadziło aktywnej polityki w zakresie udostępniania danych na szczeblu krajowym..</p> <p>W stosunku do poprzedniego okresu sprawozdawczego nie nastąpiła zmiana w zakresie ryzyka.</p>
Decyzja partnerów o innym, pozaprojektowym wykorzystaniu zadeklarowanych danych.	duża	niskie	<p>W celu zapobieżenia innemu wykorzystaniu danych podpisana została umowa o partnerstwie z instytucjami, które będą udostępniać swoje dane badawcze w ramach projektu. Część danych została już udostępniona.</p> <p>W stosunku do poprzedniego okresu sprawozdawczego nie nastąpiła zmiana w zakresie ryzyka.</p>

Nazwa ryzyka	Siła oddziaływania	Prawdopodobieństwo wystąpienia ryzyka	Sposób zarządzania ryzykiem
Pojawienie się alternatywnych narzędzi o zbliżonej funkcjonalności, które zyskają większą popularność wśród użytkowników niż te powstałe w ramach projektu.	mała	średnie	<p>Aby zmniejszyć ryzyko związane z pojawieniem się polskich narzędzi alternatywnych, prowadzona będzie aktywna promocja projektu. Środowiska naukowe w Polsce są od początku prac informowane o powstających repozytoriach, uwagi zgłaszane przez naukowców w trakcie wydarzeń edukacyjnych i informacyjnych były w miarę możliwości uwzględniane, tak aby powstające narzędzia w największym możliwym stopniu odpowiadały ich potrzebom.</p> <p>W stosunku do poprzedniego okresu sprawozdawczego nie nastąpiła zmiana w zakresie ryzyka.</p>



Nazwa ryzyka	Siła oddziaływania	Prawdopodobieństwo wystąpienia ryzyka	Sposób zarządzania ryzykiem
Zakończenie przez Uniwersytet Harvarda projektu Dataverse.	średnia	niskie	<p>W celu zminimalizowania ewentualnych skutków wycofania się twórców z utrzymywania oprogramowania, na którym projekt bazuje, wybrane zostało takie oprogramowanie, które nie tylko jest w pełni otwarte, ale też jest tworzone przez stabilną instytucję. W przypadku Dataverse, tworzonego przez Uniwersytet Harvarda, oprogramowanie jest zarazem intensywnie wykorzystywane przez inne duże instytucje akademickie (np. holenderski DANS - Digital Archiving and Networking Services), które mogą wspomóc lub przejąć utrzymywanie i rozwój oprogramowania. Obecnie ICM rozwija własną gałąź oprogramowania Dataverse.</p> <p>W stosunku do poprzedniego okresu sprawozdawczego nie nastąpiła zmiana w zakresie ryzyka.</p>

Nazwa ryzyka	Siła oddziaływania	Prawdopodobieństwo wystąpienia ryzyka	Sposób zarządzania ryzykiem
Wystąpienie niespodziewanych trudności w przygotowaniu danych do udostępnienia.	mała	średnie	<p>Zadania związane z przygotowaniem danych do udostępnienia zostały rozpoczęte na wczesnym etapie projektu, tak aby w razie opóźnienia nie wpłynęły na żadne inne zadania. Część tych zadań została już z powodzeniem ukończona.</p> <p>Dodatkowo, przez cały czas trwania projektu środowiska naukowe są zachęcane do deponowania we wszystkich repozytoriach dodatkowych, niezinwentaryzowanych w ramach projektu danych, tak by repozytoria były jak najszerzej wykorzystywane i wszystkie wskaźniki projektu zostały osiągnięte.</p> <p>W stosunku do poprzedniego okresu sprawozdawczego nie nastąpiła zmiana w zakresie ryzyka.</p>

Nazwa ryzyka	Siła oddziaływania	Prawdopodobieństwo wystąpienia ryzyka	Sposób zarządzania ryzykiem
Brak zainteresowania wśród naukowców deponowaniem danych.	mała	średnie	<p>Aby środowiska naukowe były dobrze poinformowane o możliwości deponowania danych i o korzyściach, jakie płyną z deponowania danych, projekt obejmuje odpowiednie działania informacyjne i edukacyjne. W przypadku wystąpienia zainteresowania niższego niż spodziewane, działania te mogą zostać zintensyfikowane.</p> <p>W stosunku do poprzedniego okresu sprawozdawczego nie nastąpiła zmiana w zakresie ryzyka.</p>
Brak zainteresowania korzystaniem z danych przez potencjalnych odbiorców z różnych grup.	mała	średnie	<p>Aby środowiska naukowe i społeczeństwo były dobrze poinformowane o możliwości korzystania z danych, projekt obejmuje odpowiednie działania informacyjne i edukacyjne. W przypadku wystąpienia zainteresowania niższego niż spodziewane, działania te mogą zostać zintensyfikowane.</p> <p>W stosunku do poprzedniego okresu sprawozdawczego nie nastąpiła zmiana w zakresie ryzyka.</p>

**Ryzyka wpływające na utrzymanie efektów projektu**

Nazwa ryzyka	Siła oddziaływania	Prawdopodobieństwo wystąpienia ryzyka	Sposób zarządzania ryzykiem
Brak wprowadzenia polityki otwartych danych przez MNiSW na szczeblu krajowym.	mała	wysokie	<p>Projekt przewiduje własną działalność informacyjną i edukacyjną wśród naukowców, dzięki której minimalizowane jest ryzyko, że zainteresowanie naukowców udostępnianiem danych będzie niewielkie - nawet gdyby MNiSW nie prowadziło aktywnej polityki w zakresie udostępniania danych na szczeblu krajowym.</p> <p>W stosunku do poprzedniego okresu sprawozdawczego nie nastąpiła zmiana w zakresie ryzyka.</p>

Nazwa ryzyka	Siła oddziaływania	Prawdopodobieństwo wystąpienia ryzyka	Sposób zarządzania ryzykiem
Decyzja partnerów o innym, pozaprojektowym wykorzystaniu zadeklarowanych danych.	duża	niskie	<p>W celu zapobieżenia innemu wykorzystaniu danych podpisana została umowa o partnerstwie z instytucjami, które będą udostępniać swoje dane badawcze w ramach projektu. Część danych została już udostępniona.</p> <p>W stosunku do poprzedniego okresu sprawozdawczego nie nastąpiła zmiana w zakresie ryzyka.</p>
Pojawienie się alternatywnych narzędzi o zbliżonej funkcjonalności, które zyskują większą popularność wśród użytkowników niż te powstałe w ramach projektu.	mała	średnie	<p>Aby zmniejszyć ryzyko związane z pojawieniem się polskich narzędzi alternatywnych, prowadzona będzie aktywna promocja projektu. Środowiska naukowe w Polsce są od początku prac informowane o powstających repozytoriach, uwagi zgłaszane przez naukowców w trakcie wydarzeń edukacyjnych i informacyjnych były w miarę możliwości uwzględniane, tak aby powstające narzędzia w największym możliwym stopniu odpowiadały ich potrzebom.</p> <p>W stosunku do poprzedniego okresu sprawozdawczego nie nastąpiła zmiana w zakresie ryzyka.</p>

Nazwa ryzyka	Siła oddziaływania	Prawdopodobieństwo wystąpienia ryzyka	Sposób zarządzania ryzykiem
Zakończenie przez Uniwersytet Harvarda projektu Dataverse.	średnia	niskie	<p>W celu zminimalizowania ewentualnych skutków wycofania się twórców z utrzymywania oprogramowania, na którym projekt bazuje, wybrane zostało takie oprogramowanie, które nie tylko jest w pełni otwarte, ale też jest tworzone przez stabilną instytucję. W przypadku Dataverse, tworzonego przez Uniwersytet Harvarda, oprogramowanie jest zarazem intensywnie wykorzystywane przez inne duże instytucje akademickie (np. holenderski DANS - Digital Archiving and Networking Services), które mogą wspomóc lub przejąć utrzymywanie i rozwój oprogramowania. Obecnie ICM rozwija własną gałąź oprogramowania Dataverse.</p> <p>W stosunku do poprzedniego okresu sprawozdawczego nie nastąpiła zmiana w zakresie ryzyka.</p>

Nazwa ryzyka	Siła oddziaływania	Prawdopodobieństwo wystąpienia ryzyka	Sposób zarządzania ryzykiem
Wystąpienie niespodziewanych trudności w przygotowaniu danych do udostępnienia.	mała	średnie	<p>Zadania związane z przygotowaniem danych do udostępnienia zostały rozpoczęte na wczesnym etapie projektu, tak aby w razie opóźnienia nie wpłynęły na żadne inne zadania. Część tych zadań została już z powodzeniem ukończona.</p> <p>Dodatkowo, przez cały czas trwania projektu środowiska naukowe są zachęcane do deponowania we wszystkich repozytoriach dodatkowych, niezinwentaryzowanych w ramach projektu danych, tak by repozytoria były jak najszerzej wykorzystywane i wszystkie wskaźniki projektu zostały osiągnięte.</p> <p>W stosunku do poprzedniego okresu sprawozdawczego nie nastąpiła zmiana w zakresie ryzyka.</p>

Nazwa ryzyka	Siła oddziaływania	Prawdopodobieństwo wystąpienia ryzyka	Sposób zarządzania ryzykiem
Brak zainteresowania wśród naukowców deponowaniem danych.	mała	średnie	<p>Aby środowiska naukowe były dobrze poinformowane o możliwości deponowania danych i o korzyściach, jakie płyną z deponowania danych, projekt obejmuje odpowiednie działania informacyjne i edukacyjne. W przypadku wystąpienia zainteresowania niższego niż spodziewane, działania te mogą zostać zintensyfikowane.</p> <p>W stosunku do poprzedniego okresu sprawozdawczego nie nastąpiła zmiana w zakresie ryzyka.</p>
Brak zainteresowania korzystaniem z danych przez potencjalnych odbiorców z różnych grup.	mała	średnie	<p>Aby środowiska naukowe i społeczeństwo były dobrze poinformowane o możliwości korzystania z danych, projekt obejmuje odpowiednie działania informacyjne i edukacyjne. W przypadku wystąpienia zainteresowania niższego niż spodziewane, działania te mogą zostać zintensyfikowane.</p> <p>W stosunku do poprzedniego okresu sprawozdawczego nie nastąpiła zmiana w zakresie ryzyka.</p>

## 8. Wymiarowanie systemu informatycznego

nie dotyczy

**9. Dane kontaktowe:** Jakub Szprot, Interdyscyplinarne Centrum Modelowania Matematycznego i Komputerowego Uniwersytetu Warszawskiego, [jszprot@icm.edu.pl](mailto:jszprot@icm.edu.pl), 22 874 94 04