



Otwarte dane badawcze w polityce NCN

Dr Aneta Pazik-Aybar

Seminarium MOST DANYCH

jak skutecznie motywować naukowców do udostępniania danych badawczych w Otwartym Dostępie?

Spotkanie grupy roboczej DSCC-IN PL

30 -31 maja 2022



Otwarte dane w projektach NCN / Plan prezentacji

- Otwarty dostęp do danych badawczych: wprowadzenie
- Korzyści z udostępniania danych badawczych
- Struktura Planu Zarządzania Danymi (DMP) stanowiącego część wniosku o finansowanie projektów przez NCN
- Ocena Planów Zarządzania Danymi w NCN
- European Open Science Cloud - EOSC

Wprowadzenie



Otwarte dane w projektach NCN / Co to są dane badawcze

Dane obejmują **wszelkiego rodzaju informacje** (ilościowe i jakościowe) wytwarzane, zbierane, przetwarzane, poddawane analizie w ramach projektu badawczego, obejmujące **wszystkie możliwe formaty** zarówno cyfrowe jak i niecyfrowe (np. próbki, wypełnione papierowe kwestionariusze, nagrania dźwiękowe na dyktafonie, papierowe dokumenty archiwalne itd.)

statystyki, wyniki eksperymentów, wyniki analizy treści, pomiary, obserwacje wynikające z prac terenowych, wyniki ankiet, nagrania wywiadów, obrazy, modele matematyczne, algorytmy, oprogramowanie, notatki

Otwarte dane w projektach NCN / Na czym polega otwarty dostęp do danych

Open data = **FAIR data**

“As open as possible, as closed as necessary”

Korzyści wynikające z udostępniania danych badawczych



TRANSAPRENTOŚĆ BADAŃ – wymogi czasopism

- Wzmocnienie wiarygodności realizowanych badań
- Wymagane coraz częściej w publikacjach (standardy COPE), w tym również umieszczanie surowych danych przed złożeniem manuskryptu
- Transparency and Openness Promotion (TOP) Guidelines (Springer Nature, American Association for the Advancement of Science)
- Niektórzy redaktorzy naczelni wymagają przesłania surowych danych

REPLIKOWALNOŚĆ (ang. *reproducibility*)

- **w ostatnich lat: trudności z replikowalnością wcześniej przeprowadzonych badań – jako wynik niewłaściwych praktyk badawczych (p -hacking, HARKing)**
- Zapewnienie informacji (w tym danych), które pozwalają na odtworzenie badań
- Możliwość testowania badań
- Zwiększenie wiarygodności badań i badacza

POSZERZENIE ZAKRESU BADAŃ

- Możliwość wykorzystania danych z badań, których nie można odtworzyć ze względu na szczególne warunki
- Ocena jakości przeprowadzonych badań
- Lepsze możliwości przeprowadzania meta-analiz
- Zwiększenie dynamiki procesu badawczego

WZMOCNIENIE WIDOCZNOŚCI NAUKOWCÓW

- Naukowiec zyskuje widoczność nawet gdy praca nie została opublikowana lub uzyskano negatywne wyniki
- Prace, które zawierają linki do danych, posiadają o 25% wyższy wskaźnik cytowań

Źródło: G. Colavizza, I. Hrynaszkiewicz, I. Staden, K. Whitaker, B. McGillivray, The citation advantage of linking publications to research data, „Computer Science”, arXiv:1907.02565, DOI: 10.48550/arXiv.1907.02565

ODPOWIEDZIALNA GOSPODARNOŚĆ ŚRODKAMI PUBLICZNYMI

- Udostępnione dane – w lepszym stopniu informują środowisko naukowe (i społeczeństwo) o badaniach, które zostały sfinansowane ze środków publicznych
- Zabezpieczenie danych przed ich utratą lub ograniczonym wykorzystaniem
- Dostęp do danych pozwala na sformułowanie nowych hipotez, założeń badawczych bez konieczności ponoszenia dodatkowych środków
- Możliwość rozszerzenia badań (np. o aspekty porównawcze) bez konieczności ponoszenia dodatkowych środków

Otwarte dane badawcze/ Korzyści

EDUKOWANIE W ZAKRESIE NAJWYŻSZYCH STANDARDÓW METODOLOGICZNYCH

- Udostępnione dane wraz z dokumentacją – ważną informacją dla naukowców rozpoczynających karierę naukową (analiza technik, metod badawczych)
- Możliwość poznania nowych technik badawczych

Plan Zarządzania Danymi / Kontekst

Kontekst międzynarodowy

- SCIENCE EUROPE
- EOSC
- COALITION S
- Program Horyzont Europa



Plan Zarządzania Danyymi / Ogólne informacje

Practical Guide to the International Alignment of Research Data Management



Otwarte dane w projektach NCN / Otwarta nauka w NCN

OTWARTE DANE

- 2019: Plan Zarządzania Danymi obowiązkowym elementem każdego projektu (działania, komponentu) badawczego; publikacja wytycznych (Science Europe)
- Styczeń 2021 (umowy): dane powiązane z publikacjami muszą być obowiązkowo udostępniane

Otwarte dane w projektach NCN / Publikacje i dane powiązane

POLITYKA

Dane powiązane (podstawowy zestaw danych) z opublikowanymi artykułami powinny być udostępniane w otwartym repozytorium, tam gdzie to możliwe, zgodnie z warunkami licencji *Creative Commons Public Domain* (licencja [CC0](#)), ze standardami cytowania danych zawartych w [Declaration of Data Citation Principles by FORCE 11](#) oraz na zasadach zawartych w [TOP Guidelines](#). Wszystkie publikowane metadane muszą spełniać wytyczne podane przez OpenAIRE⁷ (<https://zenodo.org/record/6918#.XqfeHf0zblU>) i zawierać adnotację o finansowaniu ze środków projektu (Narodowe Centrum Nauki, numer projektu).

UMOWA, par. 5

6. Dane będące podstawą publikacji naukowych stanowiących efekt realizacji projektu badawczego, o których mowa w ust. 2 i ust. 3, powinny być rzetelnie udokumentowane w sposób spełniający zasady maszynowego lub manualnego wyszukiwania, dostępności, interoperacyjności i ponownego użycia (tzw. **FAIR Data**). Powinny być one udostępniane w **repozytorium**, tam gdzie to możliwe, zgodnie z warunkami licencji *Creative Commons Public Domain* (licencja **CC0**).

Plan Zarządzania Danymi

Ogólne Informacje



Plan Zarządzania Danymi / Co to jest Plan Zarządzania Danymi

- to **część planu badań**
- jego celem jest uzupełnienie planu badań o **techniczny opis** sposobu zarządzania danymi
- w wypełnieniu formularza powinni pomóc pracownicy biblioteki/odpowiedniego repozytorium/działu IT/specjaliści ds. RODO

NCN dopuszcza, że w ramach niektórych projektów nie będą wytwarzane, na nowo wykorzystywane, ani poddawane analizie żadne dane badawcze ani inne podobne materiały. W takich wypadkach wymagane jest jednak krótkie uzasadnienie.

Plan Zarządzania Danymi / Ogólne cechy DMP w NCN

- stosowany dla wszystkich konkursów i naborów
- oparty na stałym formacie (ale DMP musi być dostosowane do projektu!)
- w formie kwestionariusza składającego się z sześciu głównych sekcji
- dostępny we wniosku w systemie OSF i w ogłoszeniu konkursowym
- to dokument dynamiczny – może (a nawet powinien) zmieniać się w trakcie projektu
- podlega ocenie merytorycznej
- dostępne wytyczne: zawierają pytania – mają pomóc w wypełnieniu formularza

Plan Zarządzania Danymi / Opis danych

cz. 1. Opis danych oraz pozyskiwanie lub ponowne wykorzystanie dostępnych danych

1.1 Sposób pozyskiwania lub wytwarzania nowych danych lub ponownego wykorzystywania danych istniejących

1.2 Jakie dane (tj. rodzaje, formaty, objętości) będą pozyskiwane lub wytwarzane w projekcie?

cz. 2 Dokumentacja i jakość danych

2.1 Jakie metadane i dokumentacja (np. metodologia oraz sposoby pozyskiwania i organizacji danych) będą towarzyszyć danym w projekcie?

2.2 Jakie planują Państwo zastosować środki kontroli jakości?

Plan Zarządzania Danymi / Przechowywanie i kopie zapasowe

cz. 3. Przechowywanie i tworzenie kopii zapasowych podczas badań

3.1 W jaki sposób w trakcie projektu będą przechowywane dane i metadane?

3.2 W jaki sposób zostanie zapewnione bezpieczeństwo i ochrona danych wrażliwych w okresie trwania projektu?

Plan Zarządzania Danymi / Wymogi prawne, kodeksy postępowania

cz. 4. Wymogi prawne, kodeksy postępowania

4.1 Jeżeli będzie miało miejsce przetwarzanie danych osobowych, w jaki sposób zostanie zapewniona zgodność z przepisami dotyczącymi danych osobowych oraz ich ochrony?

4.2 W jaki sposób planują Państwo zapewnić zgodność z innymi przepisami, takimi jak prawa własności intelektualnej i prawa własności? Jakie przepisy znajdują w tym przypadku zastosowanie?

Plan Zarządzania Danyymi / Udostępnianie i długotrwałe przechowywanie

cz. 5. Udostępnianie i długotrwałe przechowywanie danych

5.1 Kiedy i w jaki sposób będą udostępniane dane z projektu? Czy istnieją ewentualne ograniczenia i zakazy dotyczące ich udostępniania?

5.2 Jak będzie wyglądać selekcja danych przeznaczonych do utrwalenia i gdzie będą one długoterminowo przechowywane?

5.3 Jakie metody lub oprogramowanie umożliwiają dostęp do danych i korzystanie z danych?

5.4 W jaki sposób zagwarantują Państwo stosowanie unikalnego i trwale przypisanego identyfikatora (np. takiego jak cyfrowy identyfikator dokumentu elektronicznego (DOI)) dla każdego zbioru danych?

Plan Zarządzania Danymi / Zadania związane z zarządzaniem danymi oraz zasoby

cz. 6. Zadania związane z zarządzaniem danymi oraz zasoby

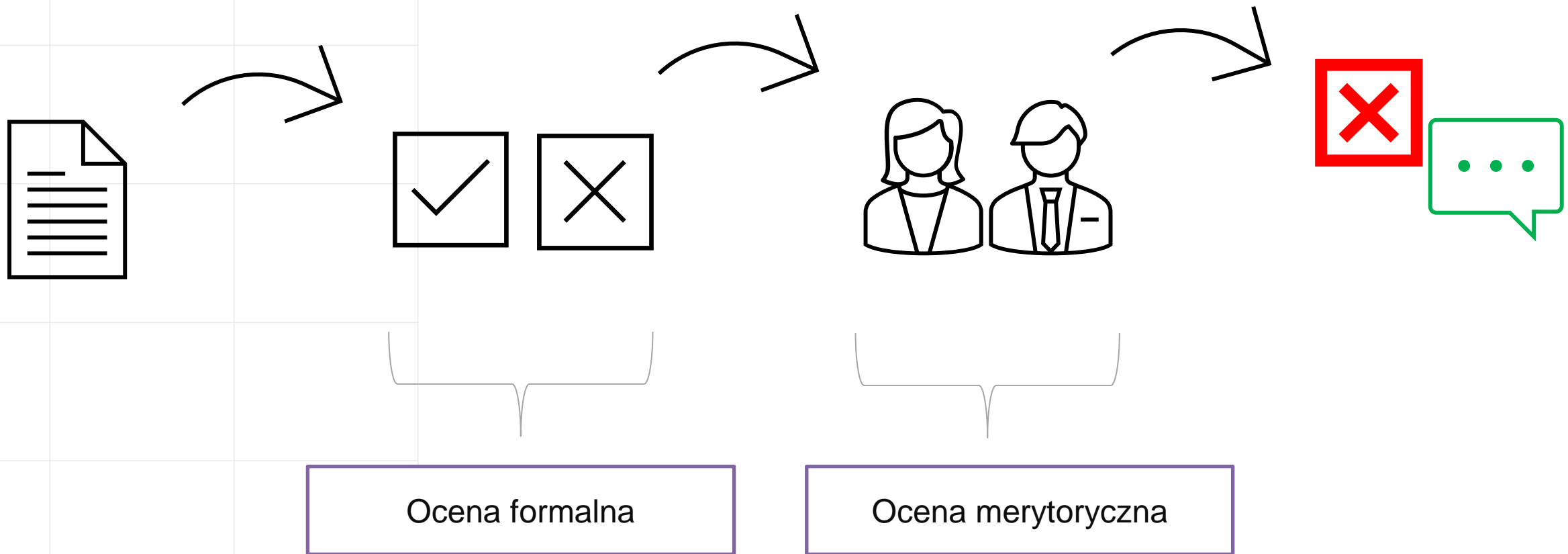
6.1 Kto będzie odpowiadał za zarządzanie danymi (tj. kto będzie ich opiekunem)?

6.2 Jakie zasoby zostaną przeznaczone na cele zarządzania danymi i zagwarantowanie przestrzegania zasad FAIR?

Ocena DMP na etapie wniosku i raportu końcowego



Plan Zarządzania Danymi / Kto ocenia DMP



Plan Zarządzania Danymi / Rozliczanie DMP

- DMP może zmieniać się w trakcie realizacji projektu (zalecane)
- Nie ma obowiązku informowania NCN o zmianach
- Raport końcowy – należy opisać stan faktyczny na koniec realizacji projektu
- Podlega ocenie Zespołów Ekspertów
- Gdy brak kompletnego planu, błędy – wezwania / sankcje





European Open Science Cloud

Otwarta nauka w NCN / Europejska Chmura Otwartej Nauki



**EUROPEAN OPEN
SCIENCE CLOUD**

Główny cel

stworzenie godnego zaufania, wirtualnego, federacyjnego internetowego środowiska, wykraczającego poza granice i dyscypliny naukowe, służącego do przechowywania, udostępniania, przetwarzania i ponownego użycia obiektów cyfrowych wykorzystywanych w badaniach naukowych zgodnie z **zasadami FAIR (ang. *Findable, Accessible, Interoperable and Reusable*)**

Cele, misja, KPI

Partnership Board

Steering Board

Komisja Europejska

Monitoring i ewaluacja
Konsultacje, doradztwo

EOSC Association
Members

Grupy robocze

HE WP; call topics
projekty

Advisory Groups
Task Forces

Działania dodatkowe

OTWARTOŚĆ I TRANSPARENTOŚĆ
UPOWSZECHNIANIE I KOMUNIKACJA
SPÓJNOŚĆ I KOORDYNACJA Z INNYMI PARTNERSTWAMI I POZA NIMI

Otwarta nauka w NCN / Europejska Chmura Otwartej Nauki

G01

nagradzanie i uczenie praktyk oraz umiejętności związanych z Otwartą Nauką – „new normal”

Percentage of publications from EOSC Association research-performing members that become immediate open access	70% do 2023
Number of national education systems that recognise European curricula for data stewardship	5 do 2025
Percentage of RPOs that are EOSC Association members that have data stewards to support their research	50% do 2025

G02

definicja standardów, rozwój narzędzi i usług pozwalających naukowcom na znajdowanie, dostęp, wykorzystywanie i kompilację wyników

Percentage of EOSC Association members that have policies which require FAIR to be implemented in project design via Data Management Plans	70% do 2023
Percentage of research data from EOSC Association members which is deposited in repositories that is made as open as possible	50% do 2025
A first generation of pan-European federation of infrastructures for preservation, management and sharing of research software is available	1 do 2025

G03

utworzenie zrównoważonej i federacyjnej infrastruktury pozwalającej na otwartą wymianę rezultatów badawczych

Percentage of the repositories in EOSC that will have a certification such as CoreTrustSeal	30% do 2025
Percentage of research disciplines that have documented standards and protocols for data sharing and reuse	60% do 2023
Percentage of the metadata related to publicly funded research datasets which are defined as Open Data that are discoverable through EOSC federated infrastructure	70% do 2025

EOSC Association:

- Ponad 230 członków i obserwatorów (RPOs, RFOs, service providers)
- z Polski: 7 organizacji (RPOs, service providers, an RFO)
- RPO, RFO, service provider może dołączyć do EOSC-A
- Roczna opłata członkowska (niższa opłata dla obserwatorów)
- Wkład do rozwoju standardów, infrastruktury, Rules of Participation, AAI, narzędzi, networking - Task Forces

www.eosc.eu

HORIZON-INFRA-2022-EOSC-01-01: Services and tools to underpin a research assessment system that incentivises open science practices

HORIZON-INFRA-2022-EOSC-01-02: Improving and coordinating technical infrastructure for institutional open access publishing across Europe

HORIZON-INFRA-2022-EOSC-01-03: FAIR and open data sharing in support of healthy oceans, seas, coastal and inland waters

HORIZON-INFRA-2022-EOSC-01-04: Support for initiatives helping to generate global standards, specifications and recommendations for open sharing of FAIR research data, publications and software



Access physical & eInfrastructures

Aggregators & Integrators

Processing & Analysis

Security & Operations

Sharing & Discovery

Training & Support

Help Desk

EOSC Portal - A gateway to information and resources in EOSC

The European Open Science Cloud (EOSC)

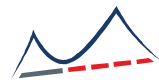
Providing 1.8 million European researchers and 70 million professionals in science and technology a virtual environment with open and seamless services for storage, management, analysis and re-use of research data, across borders and scientific disciplines.

<https://eosc-portal.eu/>

Dziękuję za uwagę!

Kontakt: aneta.pazik-aybar@ncn.gov.pl

Tel. 12 341 91 78



MOST WIEDZY
mostwiedzy.pl

Open Research Data

Projekt

*„MOST DANYCH. Multidyscyplinarny Otwarty System Transferu Wiedzy
– etap II: Open Research Data”*

współfinansowany jest

z Europejskiego Funduszu Rozwoju Regionalnego
w ramach Programu Operacyjnego Polska Cyfrowa na lata 2014-2020

