

**Tydzień Otwartej Nauki**

**It matters how we open knowledge: building structural equity**

**Warszawa, 29 października 2021 r.**

# **Dane badawcze i plany służące zarządzaniu nimi. Wprowadzenie do tematyki**

**Agnieszka Adamiec**

**Samodzielne Stanowisko  
ds. Otwartych Zasobów Nauki  
agnieszka\_adamiec@sggw.edu.pl  
tel. 59 35 782**



**SZKOŁA GŁÓWNA  
GOSPODARSTWA  
WIEJSKIEGO**

**Biblioteka Główna  
im. Władysława  
Grabskiego**

 OPEN ACCESS WEEK 2021

# *It Matters How We Open Knowledge:* Building Structural Equity

OCTOBER 25-31



SZKOŁA GŁÓWNA  
GOSPODARSTWA  
WIEJSKIEGO

**Biblioteka Główna  
im. Władysława  
Grabskiego**

# Program:



- **Definicja danych badawczych**
- **Korzyści z planowania zarządzania danymi**
- **Wybrane elementy planu zarządzania danymi badawczymi (na przykładzie grantów NCN)**



SZKOŁA GŁÓWNA  
GOSPODARSTWA  
WIEJSKIEGO

Biblioteka Główna  
im. Władysława  
Grabskiego



# Dane badawcze



**(ang. research data) dokumenty w formie cyfrowej, inne niż publikacje naukowe, które są gromadzone lub produkowane w ramach działalności badawczo-naukowej i są wykorzystywane jako dowody w procesie badawczym, bądź też są powszechnie akceptowane w środowisku naukowym jako konieczne do weryfikacji poprawności ustaleń i wyników badań.**

**Źródło:** *Dyrektywa Parlamentu Europejskiego i Rady z 20 czerwca 2019 r. w sprawie otwartych danych i ponownego wykorzystania informacji sektora publicznego (wersja przekształcona) (Dz.U.UE.L.2019.172.56)*



# Dane badawcze



**(ang. research data) wszelkie informacje gromadzone, przechowywane i przetwarzane w celu uzyskania i potwierdzenia oryginalnych wyników badań. Dane mogą być wykorzystywane do udowodnienia lub obalenia teorii, wzmocnienia twierdzeń zgłoszonych w badaniach lub do poszerzenia wiedzy na temat określonego tematu lub problemu.**

**Źródło: DeWitt Wallace Library *Data Module #1: What is Research Data?***  
**<https://libguides.macalester.edu/data1>**



SZKOŁA GŁÓWNA  
GOSPODARSTWA  
WIEJSKIEGO

Biblioteka Główna  
im. Władysława  
Grabskiego

# Kategorie danych badawczych ze względu na sposób ich stworzenia lub zgromadzenia:



- **obserwacyjne** (ang. *observational data*), np. **odczyty z czujników, wyniki ankiet, dane telemetryczne lub obserwacje laboratoryjne**
- **eksperymentalne** (ang. *experimental data*), np. **sekwencje genów, mikroskopia i spektroskopia**



# Kategorie danych badawczych ze względu na sposób ich stworzenia lub zgromadzenia cd.:



- **symulacyjne (ang. *simulation data*), np. modele klimatyczne, ekonomiczne, aktywność sejsmiczna, reakcje chemiczne**
- **pochodne lub skompilowane (ang. *derived / compiled data*), np. modele 3D, bazy danych, informacje pochodzące z eksploracji tekstów**



SZKOŁA GŁÓWNA  
GOSPODARSTWA  
WIEJSKIEGO

Biblioteka Główna  
im. Władysława  
Grabskiego



# Kategorie danych badawczych ze względu na sposób ich stworzenia lub zgromadzenia cd.:



**W niektórych źródłach uwzględnia się piątą kategorię:**

- **dane referencyjne lub kanoniczne (ang. *reference or canonical*), np. bazy danych z sekwencjami genów, struktury chemiczne lub dane dotyczące spisu ludności**

Źródła: Parland-von Essen, J. et al. (2018) *Supporting FAIR data: categorization of research data as a tool in data management*. Informaatiotutkimus, 37(4). <https://doi.org/10.23978/inf.77419>; DeWitt Wallace Library *Data Module #1: What is Research Data?* <https://libguides.macalester.edu/c.php?g=527786&p=3608643>







# Kategorie danych badawczych ze względu na stopień ich przetworzenia:

- **dane surowe**
- **dane przetworzone**

## Przykłady danych badawczych:

**dane liczbowe, dokumenty tekstowe, notatki, kwestionariusze, wyniki ankiet, nagrania audio i wideo, fotografie, bazy danych, oprogramowanie, wyniki symulacji komputerowych, protokoły metodologiczne, obserwacje laboratoryjne**



# Otwarte dane badawcze



**(ang. *open research data*) dane, uzyskane w procesie badawczym i użyte w pracy naukowej, do których każdy ma powszechny, zdalny dostęp z możliwością ich bezpłatnego ponownego wykorzystania z poszanowaniem praw autorskich.**



SZKOŁA GŁÓWNA  
GOSPODARSTWA  
WIEJSKIEGO

Biblioteka Główna  
im. Władysława  
Grabskiego



# Korzyści płynące z otwartego udostępniania danych badawczych:



- **wymiana informacji pomiędzy specjalistami reprezentującymi różne dyscypliny nauki;**
- **wzrost liczby cytowań zarówno danych badawczych, jak i powiązanych z nimi publikacji;**
- **uniknięcie czasochłonnego i często kosztownego powtarzania badań;**
- **wzmocnienie wiarygodności autora – naukowca;**
- **możliwość przeprowadzenia oceny rzetelności badań;**
- **perspektywa skorzystania z już istniejących danych.**



# FAIR data



- **możliwe do znalezienia** (*Findable*)
- **dostępne** (*Accessible*)
- **interoperacyjne** (*Interoperable*)
- **nadające się do ponownego wykorzystania** (*Reusable*)

Więcej informacji: GO FAIR <https://www.go-fair.org/fair-principles/>



SZKOŁA GŁÓWNA  
GOSPODARSTWA  
WIEJSKIEGO

Biblioteka Główna  
im. Władysława  
Grabskiego



# Podstawowa zasada udostępniania danych badawczych



**„tak otwarte, jak jest to możliwe, tak zamknięte, jak jest to niezbędne”**

**(ang. as open as possible, as closed as necessary)**



SZKOŁA GŁÓWNA  
GOSPODARSTWA  
WIEJSKIEGO

Biblioteka Główna  
im. Władysława  
Grabskiego





# Korzyści płynące z planowania zarządzania danymi badawczymi:

*Data Management Plans (DMPs) are a key element of good data management | European Commission, 2016*

- przydatne narzędzie pomagające w planowaniu;
- pozwala na łatwe zarządzanie projektami, jako punkt dostępu do informacji z nimi związanych;
- pomaga określić potrzebny budżet;
- przyczynia się o tego, że dane są FAIR;
- pokazuje odpowiedzialność.

**Źródło: CESSDA Training <https://www.cessda.eu/Training/Training-Resources/Library/Data-Management-Expert-Guide/1.-Plan/Benefits-of-data-management>**



# Europejskie i krajowe inicjatywy



- **Dyrektywa Parlamentu Europejskiego i Rady z dnia 20 czerwca 2019 r. w sprawie otwartych danych i ponownego wykorzystywania informacji sektora publicznego**
- **Zarządzenie Ministra Edukacji i Nauki z dnia 20 kwietnia 2021 r. w sprawie powołania Zespołu doradczego do spraw otwartych danych naukowych**



SZKOŁA GŁÓWNA  
GOSPODARSTWA  
WIEJSKIEGO

Biblioteka Główna  
im. Władysława  
Grabskiego

# Europejskie i krajowe inicjatywy



- **[Ustawa z dnia 11 sierpnia 2021 r. o otwartych danych i ponownym wykorzystaniu sektora publicznego](#)**
- **[Plan S](#) – Making full and immediate Open Access a reality. Principles and implementation z 2018 r.**
- **[Zarządzenie Dyrektora Narodowego Centrum Nauki](#) w sprawie ustalenia polityki Narodowego Centrum Nauki dotyczącej otwartego dostępu do publikacji z dnia 27.05.2020 r.**





# Zalecenia polityki otwartego dostępu NCN



- dane powinny być udostępnione w otwartym repozytorium, tam gdzie to możliwe, zgodnie z warunkami licencji Creative Commons (najlepiej Public Domain CC0)
- metadane powinny spełniać wytyczne podane przez [OpenAIRE](#), w tym powinny zawierać adnotację o finansowaniu ze środków projektu (Narodowe Centrum Nauki, numer projektu)
- standardy cytowania danych powinny być zgodne z wytycznymi [Declaration of Data Citation Principles by FORCE 11](#) oraz [TOP Guidelines](#) (przykład: autor(zy), rok, tytuł danych, repozytorium, wersja, stały identyfikator)



# Wytyczne NCN dla wnioskodawców do uzupełnienia PLANU ZARZĄDZANIA DANymi



W tej części formularza pod pojęciem „dane” NCN rozumie zarówno dane zebrane i dotąd nieprzetworzone, jak i dane wytworzone i poddane analizie, inne niż publikacje naukowe. Definicja ta obejmuje wszystkie możliwe formaty, zarówno cyfrowe, jak i niecyfrowe (np. próbki, wypełnione kwestionariusze, nagrania dźwiękowe, itd.). Plan zarządzania danymi można uznać za część planu badań; jego celem jest uzupełnienie tego ostatniego o techniczny opis sposobu zarządzania danymi. 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.

W wypełnieniu poniższej części formularza pomogą Państwu pracownicy biblioteki/ odpowiedniego repozytorium/działu IT Państwa instytutu lub uniwersytetu.



SZKOŁA GŁÓWNA  
GOSPODARSTWA  
WIEJSKIEGO

**Biblioteka Główna  
im. Władysława  
Grabskiego**

# Podstawowe elementy Planu Zarządzania Danymi badawczymi



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

1.1. W jaki sposób będą pozyskiwane lub wytwarzane nowe dane lub ponownie wykorzystywane dane już istniejące?

Briefly describe what types of data you are collecting or producing. Also explain what kinds of already existing data you will use. **For example**, the types of texts, images, photographs, measurements, statistics, physical samples or codes.

Describe how you plan to control and document the consistency and quality of the collected data, **for example**: calibration processes, repeated samples or measurements, data recording and capture standards, usage of controlled vocabularies, data entry validation, data peer review, etc.



# Podstawowe elementy Planu Zarządzania Danymi badawczymi



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

Recommended data formats, **for example**:

- textual documents: .odt, .ods
- tabular data: .csv, .tsv, .spss, .por
- geo-spatial data: .shp, .shx, .dbf, .sbn, .sbx, .prj, .xml
- audio: .wav, .aif, .aiff, .flac
- video: .avi
- images: .tiff, .jpg2000, .png, .svg



# Podstawowe elementy Planu Zarządzania Danymi badawczymi



## 2. Dokumentacja i jakość danych.

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

Metadata standards, **for example**: Dublin Core, DataCite, DDI (Data Documentation Initiative).

2.2. Jakie planują Państwo zastosować środki kontroli jakości?

**For example**: The data will be cataloged fulfilling the requirements of FAIR standards. The data available in a open repository will have DOI assigned and will be positioned to ensure its accessibility.



# Podstawowe elementy Planu Zarządzania Danymi badawczymi



3. Przechowywanie i tworzenie kopii zapasowych podczas badań.

3.1. W jaki sposób w trakcie projektu będą przechowywane dane i metadane? W jaki sposób będą tworzone ich kopie zapasowe?

Please consider that data storage on laptops or hard drives, for example, is risky. Storage through IT teams is safer. Please specify your back-up procedure (frequency of updates, responsibilities, automatic/manual process, security measures, etc.).



SZKOŁA GŁÓWNA  
GOSPODARSTWA  
WIEJSKIEGO

**Biblioteka Główna  
im. Władysława  
Grabskiego**

# Podstawowe elementy Planu Zarządzania Danymi badawczymi



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

Consider data protection, particularly if your data is sensitive for example containing personal data, politically sensitive information, dual-use data. Explain which institutional data protection policies are in place.

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?

Consider anonymisation or pseudonymisation of personal data.



# Podstawowe elementy Planu Zarządzania Danymi badawczymi



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?

Licenses for data sharing: Creative Commons (for example CC0), [Open Data Commons](#) (Open Database License (ODbL) — Attribution Share-Alike for data/databases; Attribution License — Attribution for data/databases; Public Domain Dedication and License (PDDL) — Public Domain for data/databases).





# Podstawowe elementy Planu Zarządzania Danymi badawczymi



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?

Raw and processed data must be stored for a period appropriate for the discipline and methodology at issue. NCN considers a minimal period of 10 years reasonable.



SZKOŁA GŁÓWNA  
GOSPODARSTWA  
WIEJSKIEGO

**Biblioteka Główna  
im. Władysława  
Grabskiego**

# Podstawowe elementy Planu Zarządzania Danymi badawczymi



5.2. Jak będzie wyglądać selekcja danych przeznaczonych do utrwalenia i gdzie będą one długoterminowo przechowywane (np. w repozytorium danych, archiwum)?

It is always recommended to refer to broadly recognised discipline-specific or certified repositories ([CoreTrustSeal](#)) in the first place. In cases where no such a repository can be identified for selection of trustworthy repository please use criteria listed in [Practical Guide to the international alignment of research data management](#). Useful listings of repositories include: [Registry of Research Data Repositories](#).



# Podstawowe elementy Planu Zarządzania Danymi badawczymi



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

The methods applied to data sharing will depend on several factors such as type, size, complexity and sensitivity of data. Indicate whether potential users need specific tools to access and (re-)use the data.

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?

Persistent identifier should be applied so that data can be reliably and efficiently located and referred to. Persistent identifiers also help to track citations and reuse.



# Podstawowe elementy Planu Zarządzania Danymi badawczymi



6. Zadania związane z zarządzaniem danymi oraz zasoby.

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

Outline the roles and responsibilities for data management/stewardship activities. Indicate who is responsible for implementing the DMP, and ensuring it is reviewed and revised.



SZKOŁA GŁÓWNA  
GOSPODARSTWA  
WIEJSKIEGO

**Biblioteka Główna  
im. Władysława  
Grabskiego**



# Podstawowe elementy Planu Zarządzania Danymi badawczymi



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

Explain how the necessary resources (for example time) to prepare data for sharing/curation have been costed in.

There are a number of self-assessment tools that might be helpful to assess the FAIRness of your data, i.e. [FAIR self-assessment tool](#).



SZKOŁA GŁÓWNA  
GOSPODARSTWA  
WIEJSKIEGO

Biblioteka Główna  
im. Władysława  
Grabskiego



# Plany Zarządzania Danymi



**Plany zarządzania danymi mają charakter bardzo indywidualny. Różnorodność tę ilustrują przykłady przedstawione na stronie [Digital Curation Centre \(UK\)](#).**



SZKOŁA GŁÓWNA  
GOSPODARSTWA  
WIEJSKIEGO

Biblioteka Główna  
im. Władysława  
Grabskiego



## Mając w pamięci definicje danych badawczych, które z poniższych NIE kwalifikuje się jako ich przykład?

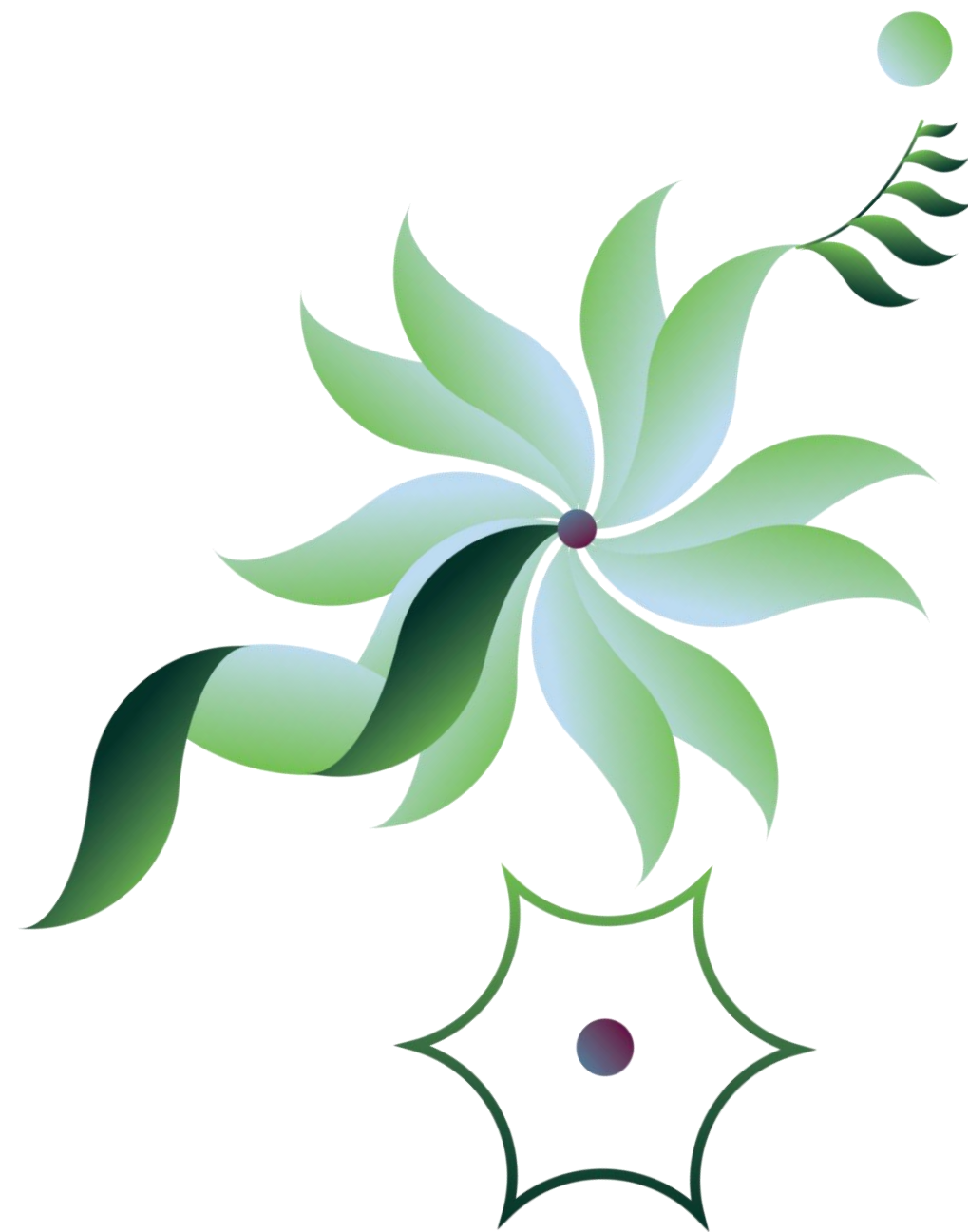
- Fotografie wydm i bagien Puszczy Kampinoskiej robione każdego roku w połowie czerwca, dokumentujące zmiany, które zaszły w okresie dwudziestu lat.
- Studia przypadków analizujących osiem unikalnych utworów muzycznych w celu zrozumienia wpływu technologii na proces twórczy komponowania muzyki.
- Nagrania audio przemówień wyborczych kandydatów na prezydenta w 2015 r.
- Suszone okazy grzybów pochodzące z doliny Wisły.
- Brak. Wszystkie powyższe są przykładami danych badawczych.

Na podstawie: DeWitt Wallace Library *Data Module #1: What is Research Data?*

<https://libguides.macalester.edu/c.php?g=527786&p=3608664>



# Dziękuję za uwagę



SZKOŁA GŁÓWNA  
GOSPODARSTWA  
WIEJSKIEGO

**Biblioteka Główna  
im. Władysława  
Grabskiego**