

Szkolenie z przygotowywania planu zarządzania danymi badawczymi (PZD)



Dr Maciej Bisaga Bogumiła Gołek
Katowice 9.05.2025

Dane badawcze

USTAWA

z dnia 11 sierpnia 2021 r.

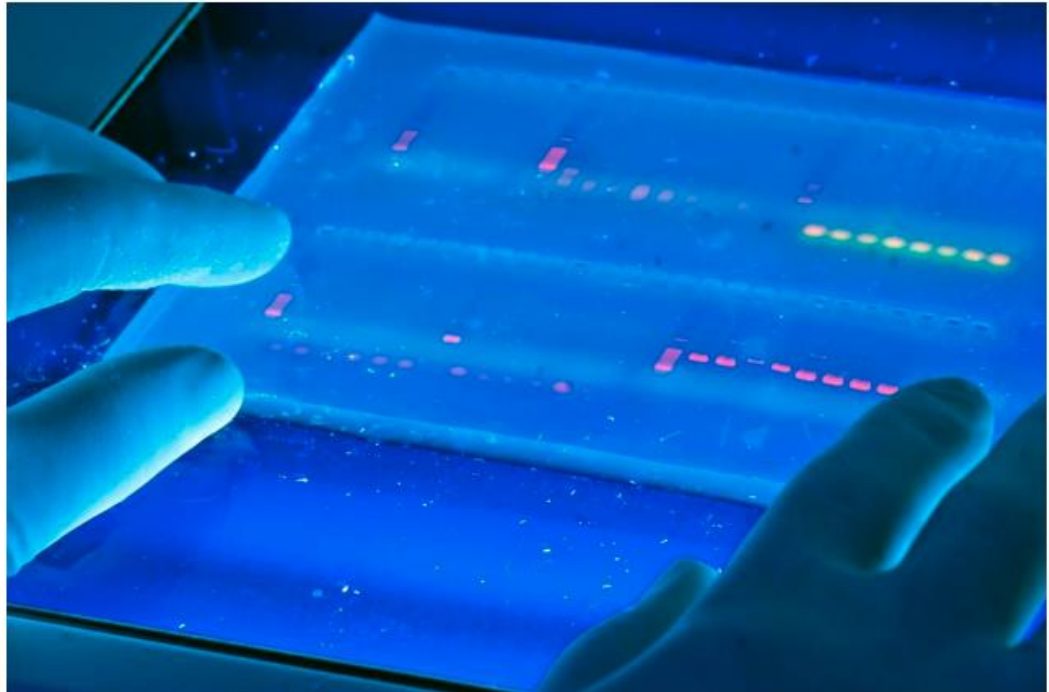
o otwartych danych i ponownym
wykorzystywaniu informacji
sektora publicznego

dane badawcze – informacje sektora publicznego utrwalone w postaci elektronicznej, inne niż publikacje naukowe, które zostały wytworzone lub zgromadzone w ramach działalności naukowej w rozumieniu art. 4 ustawy z dnia 20 lipca 2018 r. – Prawo o szkolnictwie wyższym i nauce (Dz. U. z 2021 r. poz. 478, 619 i 1630) i są wykorzystywane jako dowody w procesie badawczym lub służą do weryfikacji poprawności ustaleń i wyników badań.

Dane badawcze

W rozumieniu Planu Zarządzania Danymi, **dane** to **wszelkiego rodzaju informacje** (ilościowe i jakościowe) wytwarzane, zbierane, przetwarzane, poddawane analizie w ramach projektu badawczego, obejmujące **wszystkie możliwe** formy zarówno cyfrowe jak i nie cyfrowe, w tym:

- dokumenty tekstowe
- dane liczbowe
- wynik ankiety lub kwestionariuszy
- nagrania audio i wideo
- zdjęcia, rysunki, obrazy
- zawartość bazy danych (wideo, audio, teksty, obrazy)
- modele matematyczne, algorytmy
- oprogramowanie (skrypty, pliki wejściowe)
- wyniki symulacji komputerowych
- raporty laboratoryjne
- próbki, artefakty, przedmioty



Kwestie prawne

Polityka NCN dot. otwartego dostępu do publikacji

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#.XqfeHf0zbIU>) i zawierać adnotację o finansowaniu ze środków projektu (Narodowe Centrum Nauki, numer projektu).

Umowa, § 5. Efekty realizacji projektu

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).

Jakie dane udostępniamy

1. To zależy.

decyzja i odpowiedzialność naukowca
w jakim zakresie i formie udostępni dane

2. Dane powinny być tak otwarte, jak to możliwe i na tyle zamknięte, na ile to jest konieczne.

ang.: as open as possible, as closed as necessary.

Domyślnie uznajemy, że dane powiązane muszą zostać otwarte. Jednak mogą być sytuację, w których możemy uznać **ograniczenie dostępu** (ang. opt-out criteria) lub **embargo** (opóźniony czas otwarcia danych). Ograniczenia muszą być rzetelnie uzasadnione.

Ograniczenie dostępu w przypadku: eksploatacja przemysłowa, ograniczenia wynikające z klauzuli poufności, zasady bezpieczeństwa i dobra publicznego, prawa własności intelektualnej, w tym ochrony patentowej czy tajemnicy handlowej, kwestie etyczne i ochrony danych osobowych, dane na licencji strony trzeciej...

Co to jest dataset

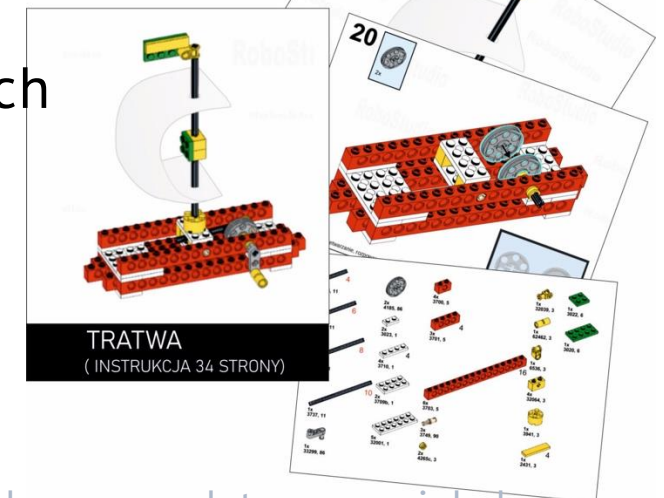
Dane badawcze najczęściej udostępniane są w postaci tzw. **datasetów**, czyli zbiorów stanowiących pewną odrębną całość i zawierających dane powiązane z jedną publikacją, projektem naukowym, bądź eksperymentem.

Metadane



DATASET

Pliki z danymi
Dokumentacja
Kod do analizy danych



Metadane

Metadane – ustrukturyzowane informacje stosowane do opisu zasobów informacji lub obiektów informacji, dostarczające szczegółowych danych, dotyczących atrybutów zasobów lub obiektów informacji, w celu ułatwienia ich znalezienia, identyfikacji, a także zarządzania tymi zasobami



Data

Filename:	Tadzik.jpg
Author:	Piotr Kononow
Date:	August 15, 2016 6:40:10PM
File:	5,312 × 2,988 JPEG 15.9 megapixels 3,393,448 bytes (3.2 megabytes)
Camera:	Samsung SM-G920F 4.3 mm
Lens:	Max aperture f/1.9 (shot wide open) Auto exposure Program AE
Exposure:	1/402 sec f/1.9 ISO 40
Flash:	none

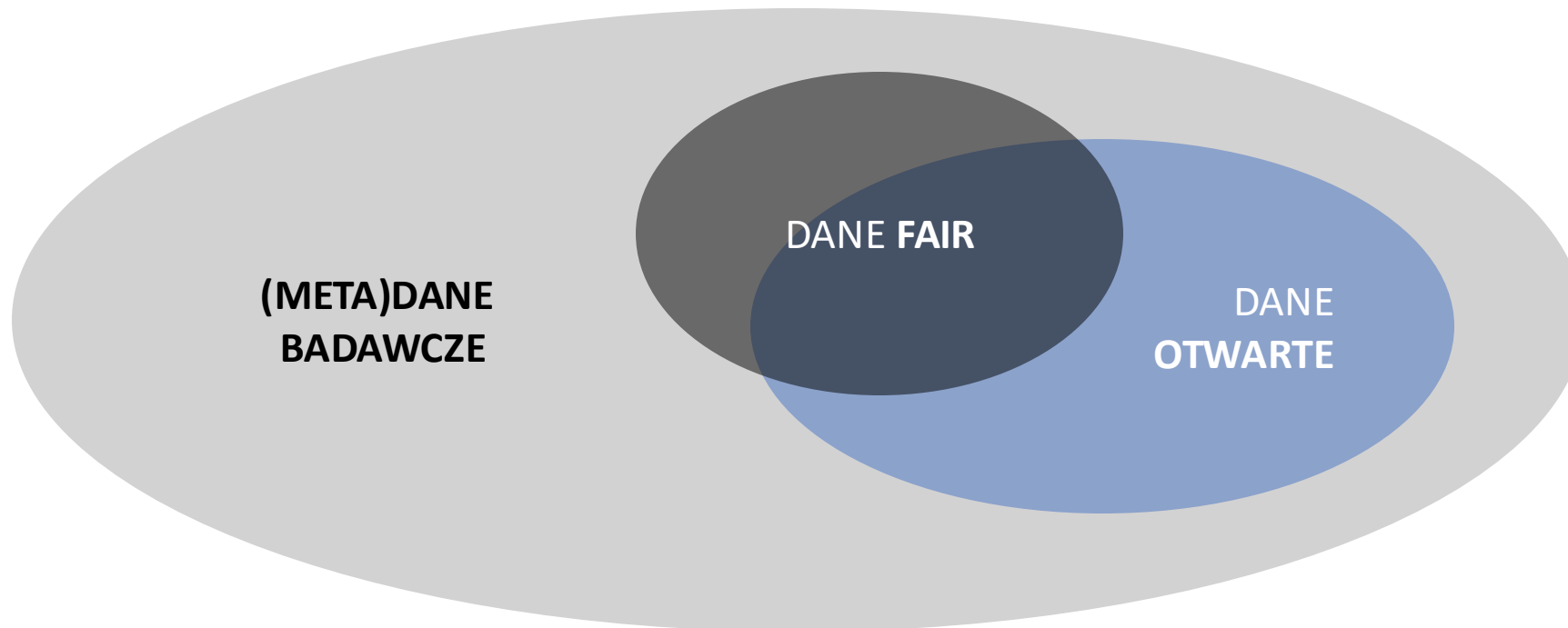


Metadata

FAIR i OPEN DATA

W polityce NCN: **Open data** = FAIR data

Dane otwarte muszą być zgodne z zasadami FAIR



FAIR-yfikacja danych

F

A

I

R



Findable



Accessible



Interoperable



Reusable

https://tools.niehs.nih.gov/srp/news/view.cfm?newsitem_ID=3130

Comment | [Open access](#) | Published: 15 March 2016

The FAIR Guiding Principles for scientific data management and stewardship

[Mark D. Wilkinson](#), [Michel Dumontier](#), [IJsbrand Jan Aalbersberg](#), [Gabrielle Appleton](#), [Myles Axton](#), [Arie Baak](#), [Niklas Blomberg](#), [Jan-Willem Boiten](#), [Luiz Bonino da Silva Santos](#), [Phillip E. Bourne](#), [Jildau Bouwman](#), [Anthony J. Brookes](#), [Tim Clark](#), [Mercè Crosas](#), [Ingrid Dillo](#), [Olivier Dumon](#), [Scott Edmunds](#), [Chris T. Evelo](#), [Richard Finkers](#), [Alejandra Gonzalez-Beltran](#), [Alasdair J.G. Gray](#), [Paul Groth](#), [Carole Goble](#), [Jeffrey S. Grethe](#), ... [Barend Mons](#)  [+ Show authors](#)

[Scientific Data](#) **3**, Article number: 160018 (2016) | [Cite this article](#)

837k Accesses | **2260** Altmetric | [Metrics](#)



w ogromnym uproszczeniu:



Findability

Resource and its metadata are easy to find by both, humans and computer systems. Basic machine readable descriptive metadata allows the discovery of interesting data sets and services.

- ✓ F1. Resource is uploaded to a public repository.
- ✓ F2. Metadata are assigned a globally unique and persistent identifier.



Accessibility

Resource and metadata are stored for the long term such that they can be easily accessed and downloaded or locally used by humans and ideally also machines using standard communication protocols.

- ✓ A1. Resource is accessible for download or manipulation by humans and is ideally also machine readable.
- ✓ A2. Publications and data repositories have contingency plans to assure that metadata remain accessible, even when the resource or the repository are no longer available.



Interoperability

Metadata should be ready to be exchanged, interpreted and combined in a (semi)automated way with other data sets by humans as well as computer systems.

- ✓ I1. Resource is uploaded to a repository that is interoperable with other platforms.
- ✓ I2. Repository meta- data schema maps to or implements the CG Core metadata schema.
- ✓ I3. Metadata use standard vocabularies and/or ontologies.



Reusability

Data and metadata are sufficiently well-described to allow data to be reused in future research, allowing for integration with other compatible data sources. Proper citation must be facilitated, and the conditions under which the data can be used should be clear to machines and humans.

- ✓ R1. Metadata are released with a clear and accessible usage license.
- ✓ R2. Metadata about data and datasets are richly described with a plurality of accurate and relevant attributes.

PID

(np. ORCiD, DOI, ROR)

otwarte, bezpłatne **formaty**,
metadane „żyją” dłużej

odczyt maszynowy i przez ludzi,
dziedzinowe **standardy**
metadanych i **słowniki**
kontrolowane

licencje, informacje dotyczące
cytowania.

Formaty plików

	FORMAT PREFEROWANY	FORMAT AKCEPTOWALNY
DANE TEKSTOWE	.odt, .ods	.doc, .docx, .pdf, .xml, .htm, .html, .rtf, .xlsx, .epub
DANE LICZBOWE	.csv, .tsv, .spss, .por	.xls, .sav, .dta, .mdb/.accdb
DANE OBRAZU	.tiff, .jpeg2000, .png, .svg	.gif, .jpg, .ai, .cgm
DANE GEOPRZETRZENNE	.shp, .shx, .dbf, .sbn, .sbx, .prj, .xml	PostGIS, .tif, .tfw, .fdg, .adf, .dat, .nit
PLIKI AUDIO	.wav, .aif, .aiff, .flac	.mp3, .m4p, .m4a, .mid, .midi, .ogg
PLIKI WIDEO	.avi	.mov, .wmv, .mpg
PREZENTACJE	.pdf, .opd	.pptx

Otwarte formaty są bardziej trwałe i łatwiejsze do odczytu w przyszłości.

Zastrzeżone formaty mogą stać się nieczytelne z powodu zmian technologicznych.

Konwersja do standardowych formatów zapewnia długoterminowy dostęp do danych.

Dane źródłowe

- Zebrane/wyprodukowane „surowe dane”
- Ponownie wykorzystane dane z bazy danych/repozytorium

Przetwarzanie danych

- Transformacja surowych danych

Analiza danych

- Interpretacja danych
- Generowanie wyników i rezultatów

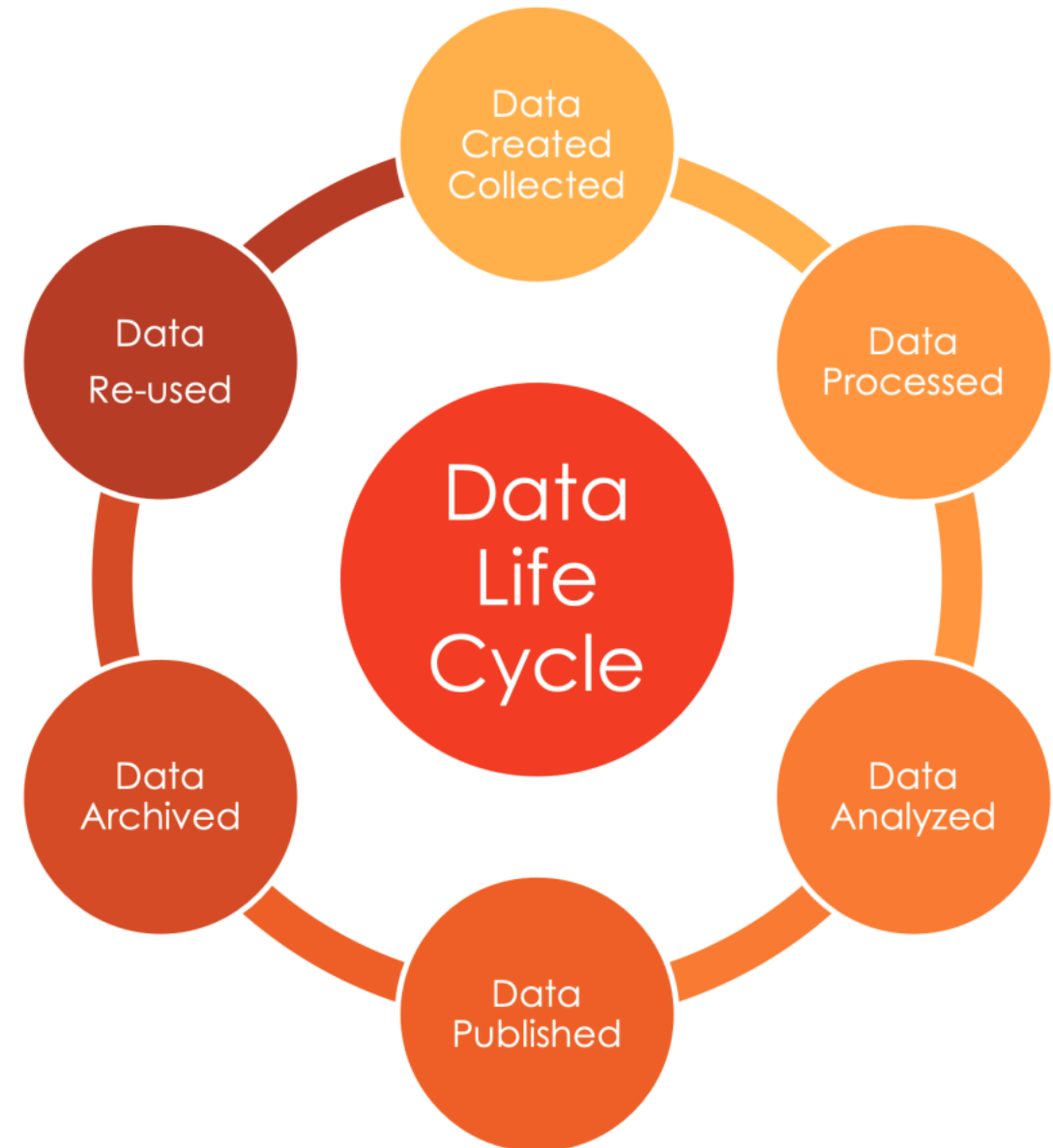
Publikacja danych (artykuł naukowy)

- Rękopis + informacje uzupełniające

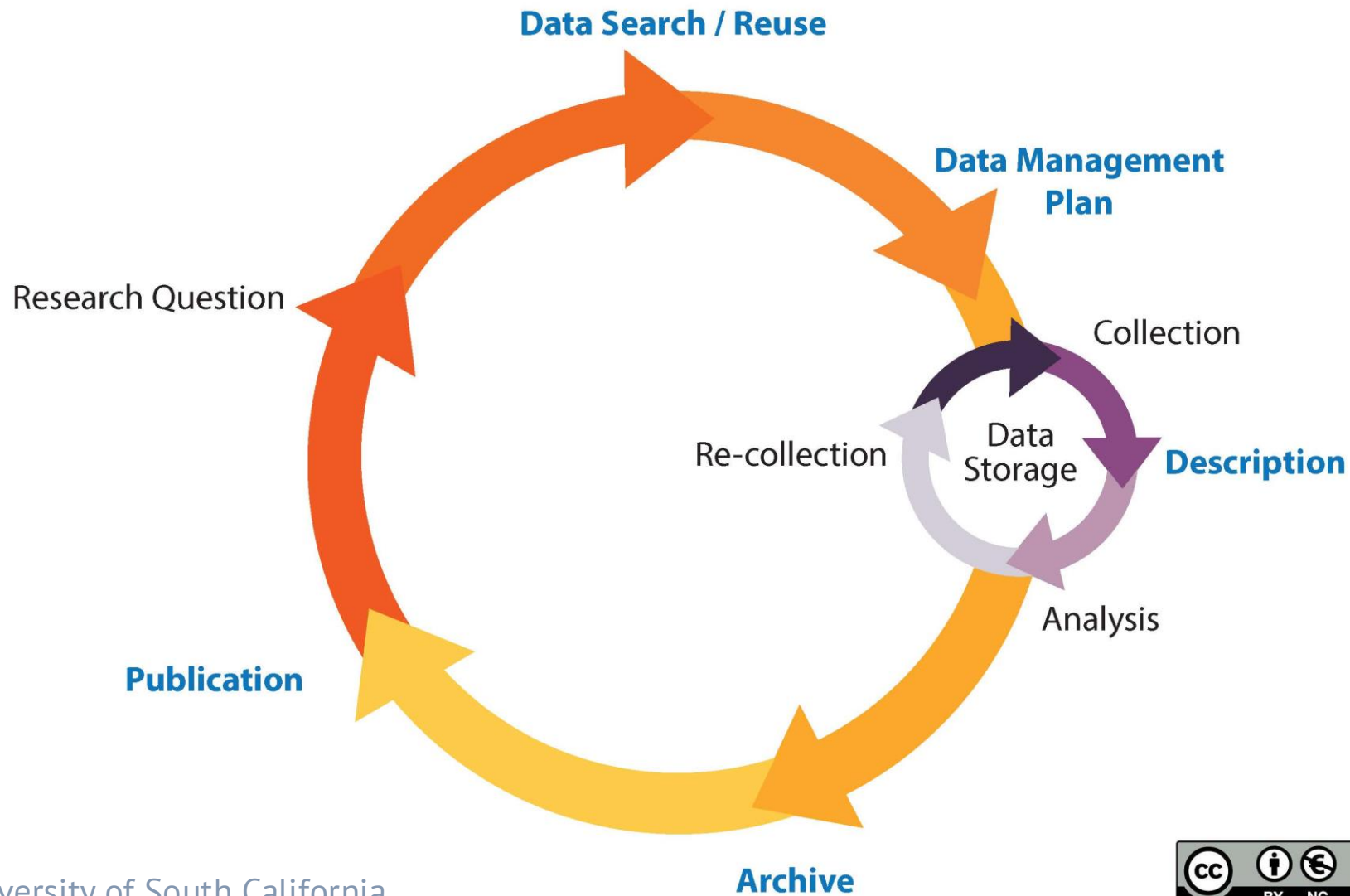
Archiwizacja danych (bazy danych, repozytoria, archiwa)

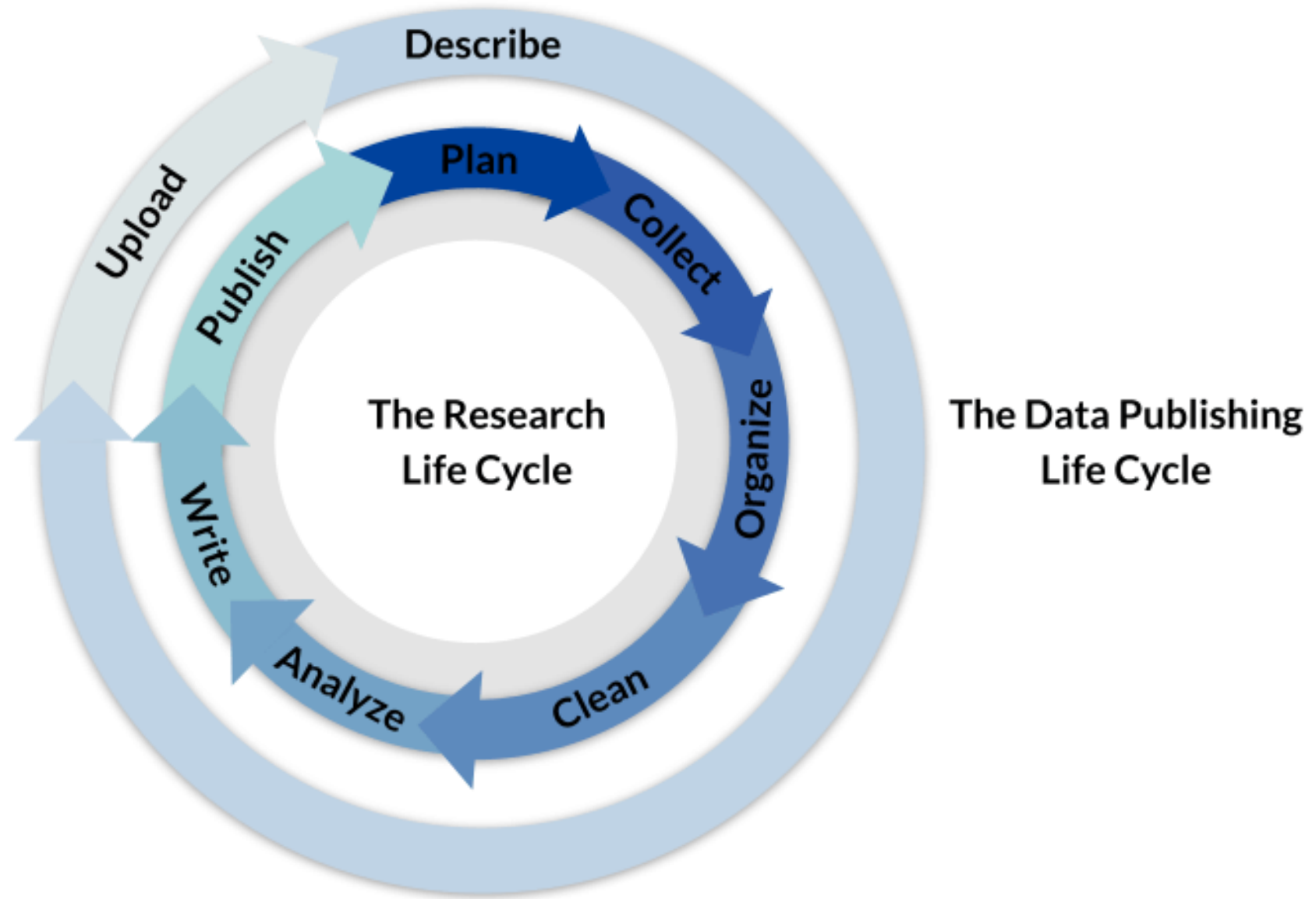
- Dane stanowiące podstawę publikacji
- Oddzielne zestawy danych

Ponowne wykorzystanie danych (rejestry, repozytoria)



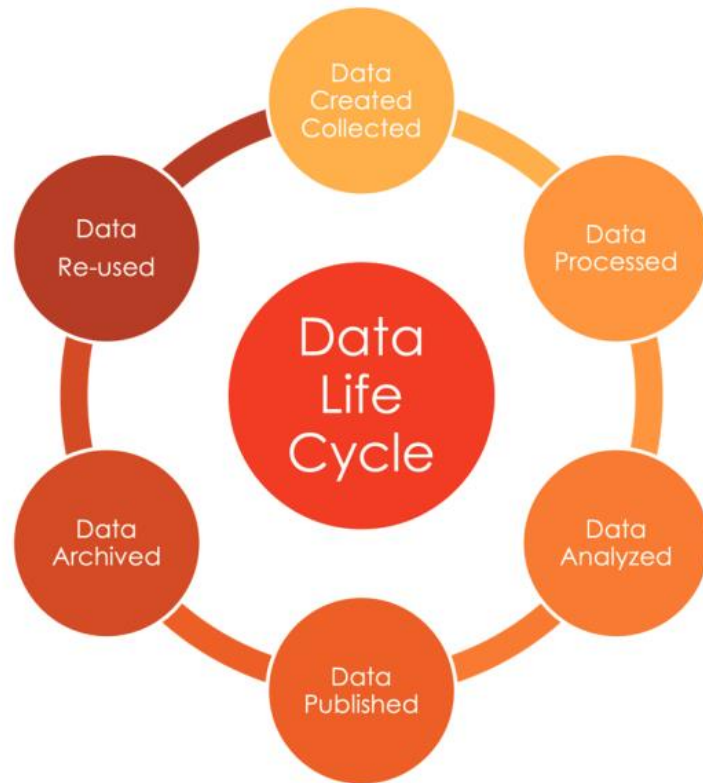
The Research Data Management Lifecycle





1.

- Generowanie pomysłów
- Projektowanie badań
- Wniosek o finansowanie



1.5. Kwestie do przemyślenia

Organizacja

- Struktura katalogów
- Format, nazwy, wersje

Dokumentacja

- Opis danych
- Szczegóły eksperymentalne
- Podjęte decyzje
- Metadane

Przechowywanie

- Kopia zapasowa
- Długoterminowa archiwizacja

Dostęp do danych

- Prawa dostępu (otwarty, ograniczony)
- Licencje

Zasady agencji finansujących

- Polityka otwartego dostępu
- Plan zarządzania danymi

Wymogi prawne i etyczne

- Krajowe i europejskie przepisy prawne
- Ramy etyczne dla naukowców
- Ochrona danych osobowych
- Prawa własności intelektualnej
- Komercyjne wykorzystanie danych

Zasady instytucjonalne

- Polityka zarządzania danymi badawczymi (RDM)
- Kodeksy postępowania i etyki
- Ochrona danych
- Umowy o współpracy

Zasady czasopism i wydawców

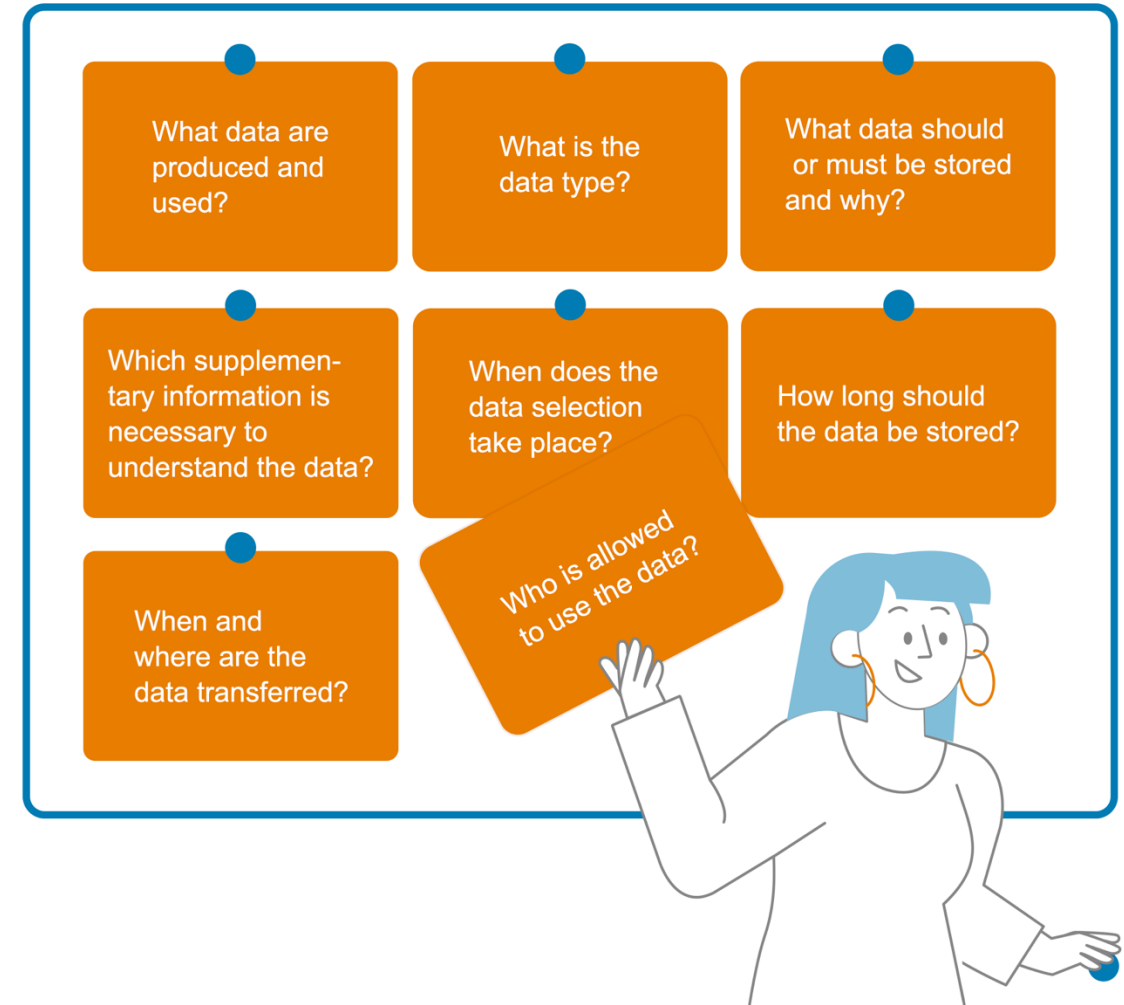
- Polityka udostępniania danych

Plan zarządzania danymi (DMP) to formalny dokument, który określa, w jaki sposób badacz zamierza zarządzać swoimi danymi badawczymi w trakcie i po zakończeniu projektu.

Wg NCN:

- to **część planu badań**
- jego celem jest uzupełnienie planu badań o **techniczny opis** sposobu zarządzania danymi badawczymi
- podlega ocenie formalnej, merytorycznej i raportowaniu [od 1. stycznia 2021]

Data Management Plan THE ELEMENTS



Plan Zarządzania Danyymi

PLAN ZARZĄDZANIA DANYMI [w języku angielskim]	
Sekcja wypełniana w języku angielskim.	
<i>Przed wypełnieniem formularza należy zapoznać się z Wytycznymi do uzupełniania planu zarządzania danymi w projekcie badawczym dostępnymi w ogłoszeniu o konkursie.</i>	
<i>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, które należy zamieścić w ramach odpowiedzi na pytanie: 1.1. Sposób pozyskiwania i opracowywania nowych danych lub ponownego wykorzystania dostępnych danych. Pod każdym pytaniem pole tekstowe na opis (do 1000 znaków).</i>	
1.	Opis danych oraz pozyskiwanie lub ponowne wykorzystanie dostępnych danych
1.1	Sposób pozyskiwania i opracowywania nowych danych i/lub ponownego wykorzystania dostępnych danych
1.2	Pozyskiwane lub opracowywane dane (np. rodzaj, format, ilość)
2.	Dokumentacja i jakość danych
2.1	Metadane i dokumenty (np. metodologia lub pozyskiwanie danych oraz sposób porządkowania danych) towarzyszące danym
2.2	Stosowane środki kontroli jakości danych
3.	Przechowywanie i tworzenie kopii zapasowych podczas badań
3.1	Przechowywanie i tworzenie kopii zapasowych danych i metadanych podczas badań
3.2	Sposób zapewnienia bezpieczeństwa danych oraz ochrony danych wrażliwych podczas badań
4.	Wymogi prawne, kodeks postępowania
4.1	Sposób zapewnienia zgodności z przepisami dotyczącymi danych osobowych i bezpieczeństwa danych w przypadku przetwarzania danych osobowych
4.2	Sposób zarządzania innymi kwestiami prawnymi, np. prawami własności intelektualnej lub własnością. Obowiązujące przepisy
5.	Udostępnianie i długotrwałe przechowywanie danych
5.1	Sposób i termin udostępnienia danych. Ewentualne ograniczenia w udostępnianiu danych lub przyczyny embarga
5.2	Sposób wyboru danych przeznaczonych do przechowania oraz miejsce długotrwałego przechowywania danych (np. repozytorium lub archiwum danych)
5.3	Metody lub narzędzia programowe umożliwiające dostęp do danych i korzystanie z danych
5.4	Sposób zapewniający stosowanie unikalnego i trwałego identyfikatora (np. cyfrowego identyfikatora obiektu (DOI)) dla każdego zestawu danych
6.	Zadania związane z zarządzaniem danymi oraz zasoby
6.1	Osoba (np. funkcja, stanowisko i instytucja) odpowiedzialna za zarządzanie danymi (np. <i>data steward</i>)
6.2	Środki (np. finansowe i czasowe) przeznaczone do zarządzania danymi i zapewnienia możliwości odnalezienia, dostępu, interoperacyjności i ponownego wykorzystania danych

Plan Zarządzania Danymi / Dlaczego jest ważny na etapie planowania badań?

Stworzenie PZD pomaga:

- podejmować decyzje dotyczące zarządzania danymi będąc świadomymi ich konsekwencji
- zidentyfikować zasoby i narzędzia potrzebne do badań
- oszacować koszty zasobów i narzędzi oraz odpowiednio zaplanować budżet
- przewidywać i mitygować problemy (np. utratę danych, powielanie wysiłków, naruszenia bezpieczeństwa) z odpowiednim wyprzedzeniem
- określić zasady i sposób dostępu do danych na etapie realizacji badań
- wspierać członków grupy badawczej w rozumieniu, co dzieje się z danymi.

Plan Zarządzania Danymi / Dlaczego jest ważny w pracy badawczej?

Aktualizowany, dostępny, wykonalny i wdrożony PZD:

- pomaga monitorować postępy w realizacji badań
- pomaga zorganizować i skoordynować realizację badań
- ułatwia współpracę pomiędzy naukowcami pracującymi w danym projekcie
- wzmacnia kontrolę dostępu do danych i jakości danych
- wspiera bezpieczeństwo danych osobowych i wrażliwych np. zabezpieczając przed ich nieuprawnionym wykorzystaniem
- wspiera organizację długotrwałego przechowywania danych

Plan Zarządzania Danymi/Ogólne cechy DMP w NCN

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

Plan Zarządzania Danymi/Wytyczne NCN



Wytyczne dla wnioskodawców do uzupełnienia PLANU ZARZĄDZANIA DANYMI w projekcie badawczym

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.

Pytania	Pytania pomocnicze
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? Warto rozważyć następujące kwestie: – Jakie standardy, metody i oprogramowanie posłużą do pozyskiwania lub wytwarzania nowych danych? – Jakie dane już istniejące (własne lub będące w posiadaniu strony trzeciej) zostaną ponownie wykorzystane? – W jaki sposób zostanie udokumentowane pochodzenie danych? – Jak wyglądać będzie organizacja plików i zarządzanie ich różnymi wersjami? – Czy konieczna będzie digitalizacja danych analogowych lub wydanych w formie papierowej (takich jak mapy, fotografie, teksty)?	Prosimy o wyjaśnienie, w jaki sposób będą pozyskiwane lub wytwarzane nowe dane lub ponownie wykorzystywane dane już istniejące. Należy również pokrótce opisać ich rodzaj, np. teksty, obrazy, zdjęcia, pomiary, statystyki, próbki fizyczne czy kody. Prosimy o określenie, w jaki sposób planują Państwo kontrolować i dokumentować ich jakość i spójność: procesy kalibracji, powtórne pobieranie próbek lub ponowne wykonywanie pomiarów, standardy zapisywania/nagrywania i utrwalania danych, wykorzystanie narzuconego słownictwa, walidacja wprowadzanych danych, specjalistyczne recenzje, itd.
1.2 Jakie dane (tj. rodzaje, formaty, objętości) będą pozyskiwane lub wytwarzane w projekcie? Warto rozważyć następujące kwestie: – Jaki rodzaj, format i objętość danych planują Państwo pozyskać, wytworzyć lub ponownie wykorzystać?	Odpowiedź powinna uwzględnić rodzaj, format i zawartość każdego planowanego zbioru danych. Należy również przedstawić ich szacunkową objętość. Prosimy o wyszczególnienie formatu, tj. sposobu, w jaki dane będą kodowane na potrzeby



6.2 Jakie zasoby zostaną przeznaczone na cele zarządzania danymi i zagwarantowanie przestrzegania zasad FAIR ⁵ ? Warto rozważyć następujące kwestie: – Czy potrzebne są dodatkowe zasoby do zarządzania danymi, takie jak osoby, czas, sprzęt lub oprogramowanie? – Jakiekolwiek koszty związane będą z zapewnieniem standardów FAIR w projekcie? – W jaki sposób zostaną opłacone?	Prosimy o wyjaśnienie, w jaki sposób koszty niezbędne do przygotowania danych do ich udostępniania i przechowywania będą oszacowane? Czy do przygotowania danych do przechowywania i archiwizowania niezbędne będą dodatkowe zasoby? Jeżeli tak, jak wysokie będą to koszty i w jaki sposób zostaną opłacone? Rozważ konieczność wydatkowania środków na długoterminowe przechowywanie danych?
--	--

Dodatkowe informacje:

Jeżeli składają Państwo wniosek w konkursie międzynarodowym organizowanym przez NCN w oparciu o procedurę agencji wiodącej, w którym rolę agencji wiodącej pełni zagraniczna instytucja partnerska, to bez względu na to, czy ta agencja wymaga przedłożenia planu zarządzania danymi czy nie, nadal mają Państwo obowiązek uzupełnienia go w formularzu wniosku krajowego składanym w systemie ZSUN/OSF. Treść planu dotyczy tylko danych badawczych, które powstaną lub będą ponownie wykorzystywane w trakcie realizacji projektu przez polski zespół badawczy.

Jeżeli to NCN pełni rolę agencji wiodącej to wymagane jest przedłożenie planu zarządzania danymi dotyczącego całości projektu. Ewentualne koszty związane z przechowywaniem, archiwizacją i udostępnianiem danych mogą dotyczyć jednak tylko danych powstałych w trakcie realizacji projektu przez polski zespół badawczy.

Jeżeli składają Państwo wniosek w konkursie międzynarodowym organizowanym przez NCN we współpracy dwu- i wielostronnej, to bez względu na to, czy na poziomie międzynarodowym wymagane jest przedłożenie planu zarządzania danymi czy nie, nadal mają Państwo obowiązek uzupełnienia go w formularzu wniosku krajowego składanym w systemie ZSUN/OSF. Treść planu dotyczy tylko danych badawczych, które powstaną lub będą ponownie wykorzystywane w trakcie realizacji projektu przez polski zespół badawczy.

Plany zarządzania danymi mają charakter bardzo indywidualny. Istnieje wiele ich rodzajów, a ich organizacja może się różnić. Różnorodność tę ilustrują [przykłady przedstawione na stronie Digital Curation Centre](#) (UK).

⁵ Istnieje szereg narzędzi pozwalających samodzielnie ocenić stopień, w jakim dane spełniają kryteria FAIR, np.





<https://doi.org/10.5281/zenodo.4915862>

Additional guidance to the Science Europe DMP assessment rubric

Authors
Maaike Verborg and Marjan Grootveld
(DANS)

Data Archiving and Networked Services
DANS

Related FAIRsFAIR recommendation¹

Recommendation B1:

Formalise and support appropriate data management plans (DMPs) for FAIR data
Researchers, data stewards and academic journals can use FAIR data criteria to review whether a DMP includes steps needed to make the data FAIR.

Introduction

One way to support data management planning for FAIR data is to incorporate FAIR criteria more explicitly in data management plan (DMP) templates. As DMPs hold an important position in the planning phase of a research project, the inclusion of FAIR data criteria at this stage will facilitate a greater understanding of the necessary steps needed to make data FAIR as well as an increase in the number of FAIR datasets produced as a result. In this output of the FAIRsFAIR project, FAIR guidance has been added to the Science Europe DMP evaluation rubric to help researchers and data stewards better plan for FAIR data. This guidance is based on FAIR-Aware², the FAIR learning tool developed in the FAIRsFAIR project.

As the association representing public research performing and funding organizations throughout Europe, Science Europe offers resources to accommodate the implementation of data management and FAIR data practices as one of their priorities. One of the resources offered by Science Europe is a 'Practical Guide to the International Alignment of Research Data Management'. This resource offers targeted guidance for organisations, scientific communities, and individual researchers, to organise research data and preserve it appropriately. The FAIRsFAIR team has taken this practical guide as a basis and proposed additional guidance related to FAIR data based on FAIR-Aware. This guidance material is meant to complement, rather than replace the guidelines of the original guide, wherever possible.



<https://doi.org/10.5281/zenodo.6088215>

1. FAIRsFAIR D3.4 Recommendations on practice to support FAIR data principles <https://doi.org/10.5281/zenodo.5357329>
2. FAIR-Aware <https://www.fairsfair.eu/fair-aware>
3. Science Europe, Practical Guide to the International Alignment of Research Data Management - Extended Edition. <https://doi.org/10.5281/zenodo.4915861> In addition, Science Europe published a Practical Guide to Sustainable Research Data. <https://doi.org/10.5281/zenodo.4769702>



W razie niejasności – jesteśmy do dyspozycji ;)

Wspólnota Data Stewardów na Uniwersytecie Śląskim

Wydział Nauk Ścisłych i Technicznych

Bogumiła Gołek
Marta Urbaniak



Wydział Nauk Przyrodniczych

Dariusz Ignatiuk
Łukasz Małarzewski
Agnieszka Bugaj-Nawrocka



Wydział Nauk Społecznych

Aleksandra Sokoła



Wspólnota Data Stewardów UŚ



Wydział Prawa i Administracji

Monika Łyczba



Wydział Teologiczny

Dagmara Miłek



Administracja Centralna

Maciej Bisaga



Wydział Humanistyczny

Anna Dziadek

Plan Zarządzania Danymi / Błędy ogólne

- DMP w j. polskim
- informacja, że dane będą opracowywane papierowo
- deklaracja, że projekt nie generuje żadnych danych, ani nie wykorzystuje istniejących danych (a projekt przewiduje np. badania eksperymentalne)
- brak **indywidualnego** DMP – schemat powielany przez naukowców z tej samej uczelni – tymczasem DMP jest zawsze oceniane w kontekście danego projektu badawczego
- pojawiają się ogólnikowe stwierdzenia – brak precyzyjnych informacji (np. zostanie zastosowany najbardziej odpowiedni format dla tego rodzaju danych)

Plan Zarządzania Danymi / Najczęstsze błędy

Pozyskiwanie/wykorzystywanie pozyskanych danych

- brak informacji na temat formatu i rodzaju danych, ich objętości
- nie daje się pierwszeństwa standardowym otwartym formatom

Dokumentacja i jakość

- brak informacji o zakresie/standardzie metadanych, jaki zostanie zastosowany
- brak informacji w jaki sposób dane będą organizowane (foldery i subfoldery, nazwy plików, wersje)
- brak informacji w jakim miejscu znajdzie się opis dokumentacji danych

Plan Zarządzania Danymi/Najczęstsze błędy

Przechowywanie i kopie zapasowe

- brak informacji gdzie dane będą przechowywane lub podanie firmy, z którą Uniwersytet Śląski nie ma sformalizowanej umowy (np. DropBox)
- dane nie są przechowywane zgodnie z zasadą **3-2-1**

Wymogi prawne, kodeksy postępowania

- podawanie bardzo ogólnikowych informacji („będzie przestrzegane RODO”, „będą przestrzegane regulacje uczelniane”)
- brak podania informacji, kto jest właścicielem danych
- brak informacji na temat licencji jakie zostaną zastosowane dla udostępnionych danych (Creative Commons) lub podanie takiej informacji dla artykułów, a nie dla danych badawczych

The 3-2-1 Rule



Have at Least Three Copies of Your Data



Store Two Copies on Different Storage Types



Keep One Copy Off-Site

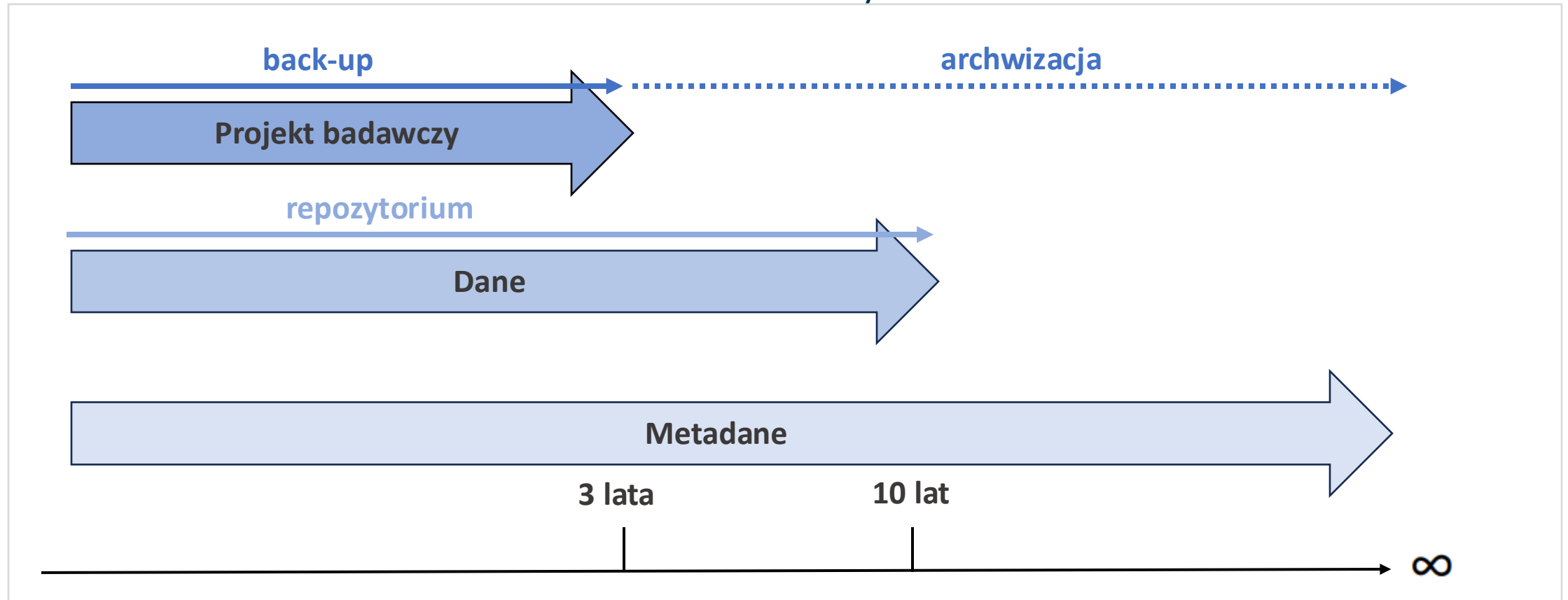
3 kopie plików
(w tym jeden oryginał)

na **2** różnych nośnikach

1 kopia poza miejscem
przetwarzania danych
(inny serwer, chmura)

voonami.com

Przechowywanie danych/Okres przechowywania danych i metadanych



Plan Zarządzania Danymi/Najczęstsze błędy

Przechowywanie i kopie zapasowe

- brak informacji gdzie dane będą przechowywane lub podanie firmy, z którą Uniwersytet Śląski nie ma sformalizowanej umowy (np. DropBox)
- dane nie są przechowywane zgodnie z zasadą **3-2-1**

Wymogi prawne, kodeksy postępowania

- podawanie bardzo ogólnikowych informacji („będzie przestrzegane RODO”, „będą przestrzegane regulacje uczelniane”)
- brak podania informacji, kto jest właścicielem danych
- brak informacji na temat licencji jakie zostaną zastosowane dla udostępnionych danych (Creative Commons) lub podanie takiej informacji dla artykułów, a nie dla danych badawczych

Plan Zarządzania Danymi /Najczęstsze błędy

Długotrwałe przechowywanie i udostępnianie danych

- nie podanie repozytorium lub podanie repozytorium, które nie spełnia warunków otwartości zgodnie z zasadami FAIR (RE-BUŚ/OPUS, PAAD!!)
- brak deklaracji dotyczącej udostępniania danych lub wyjaśnienia dlaczego dane nie będą udostępniane (puste pola)

Zadania związane z zarządzaniem danymi oraz zasoby

- kierownik projektu jako jedyna osoba odpowiedzialna za wszystkie działania związane z zarządzaniem danymi



<https://us.edu.pl/nauka-i-badania/otwarta-nauka/>



European Open Science School (EOSS)

EOSS: szkoła zarządzania danymi badawczymi zgodnie z zasadami otwartej nauki



https://us.edu.pl/nauka-i-badania/otwarta-nauka/do_pobrania

Pliki do pobrania

Dziedzinowe Plany Zarządzania Danymi
Instrukcje Narodowego Centrum Nauki PL/ENG

-  [Wytyczne dla wnioskodawców do uzupełnienia PLANU ZARZĄDZANIA DANYMI w projekcie badawczym \(122 b\)](#)
-  [Zarządzanie jakościowymi danymi badawczymi \(122 b\)](#)

Instrukcja Deponowania Danych Badawczych w RepOD

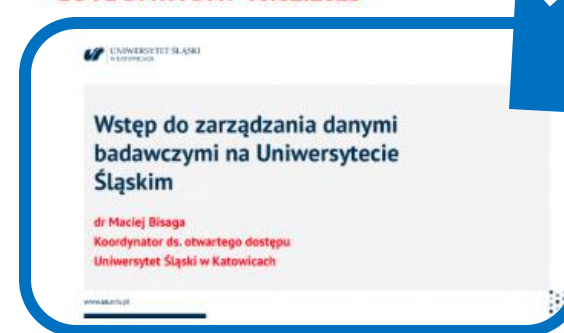


 [Instrukcja deponowania danych badawczych \(3,44 MB\)](#)
w repozytorium RePOD dla pracowników
Uniwersytetu Śląskiego w Katowicach

z czym wiąże się
otwieranie nauki

Prezentacje Ze Szkoleń Prowadzonych Wewnątrz

LOVE DATA DAY 10.02.2025

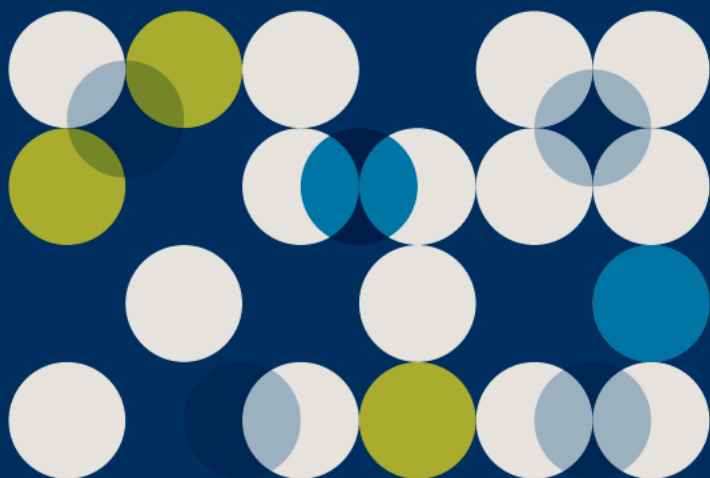


 „Wstęp do zarządzania danymi badawczymi na Uniwersytecie Śląskim” – Maciej Bisaga (5,41 MB)

Pomocne materiały

PORADNIK DLA NAUKOWCÓW I DATA STEWARDÓW

Zarządzanie danymi badawczymi



Poradnik przedstawia wytyczne i regulacje dotyczące danych badawczych, analizuje proces zarządzania danymi badawczymi w całym cyklu ich życia i omawia istotne zagadnienia prawne. Przeznaczony zarówno dla naukowców, jak i data stewardów.

<https://pon.edu.pl/dane-materialy>



Wskazówki praktyczne i pomocne materiały



Szukaj...

[Strona główna](#)[Katalog kursów](#)[Moje kursy](#)

PL

Moje konto

FILTRUJ

wyczyść filtry

WYŚWIETLANIE 4 KURSÓW

SORTUJ: Domyślnie

Kategoria

Systemy dla nauki i szkolnictwa wyższego 4

Poziom trudności

początkujący 2

średnio zaawansowany 2

Język

polski 4

Organizator

Uniwersytet Warszawski



NOWY KURS

Zarządzanie danymi badawczymi dla naukowców - kurs średnio zaawansowany

Uniwersytet Warszawski

Rozpoczęcie: 2 paź 2023



NOWY KURS

CELE KURSU

Zi

bi

ki

Ur

Rc

- Zdobędziesz podstawową wiedzę na temat otwartej komunikacji naukowej i polityk otwartości instytucji finansujących badania naukowe.
- Dowiesz się, jak przygotować plan zarządzania danymi badawczymi zgodnie z wytycznymi Narodowego Centrum Nauki i Komisji Europejskiej.
- Poznasz dobre praktyki związane z zarządzaniem danymi i ich otwartym udostępnianiem.
- Zdobędziesz umiejętności potrzebne do tego, aby udostępnić dane zgodnie z zasadami FAIR, czyli tak, aby były łatwe do znalezienia, dostępne, interoperacyjne i możliwe do ponownego wykorzystania.
- Poznasz różne typy repozytoriów danych badawczych, dowiesz się jakie są ich funkcje i jak należy z nich korzystać.
- Zapoznasz się z podstawowymi prawnymi aspektami pracy z danymi badawczymi i ich udostępniania oraz dowiesz się, jak korzystać z wolnych licencji.
- Poznasz różne rodzaje danych, baz danych, trwałych identyfikatorów i metadanych.
- Dowiesz się, jaką rolę pełni w instytucji naukowej data steward.

<https://navoica.pl/courses>

Wskazówki praktyczne i pomocne materiały

<https://www.ncn.gov.pl/finansowanie-nauki/otwarta-nauka>

<https://www.ncn.gov.pl/finansowanie-nauki/otwarta-nauka/szkolenia-i-wydarzenia>



Otwarta nauka

o NCN

o NCN

» Zadania NCN

» Dyrekcja

» Rada NCN

» Koordynatorzy dyscyplin

» Akty prawne

» Struktura

» Otwarta nauka

» Wspieranie równości płci

» Praca w NCN

» Zamówienia publiczne

» Wydarzenia

» Materiały

Otwarta nauka to sposób uprawiania badań i komunikowania wyników w przejrzystości, replikowalności, weryfikowalności. Są one kluczowe dla rosnących najwyższymi standardami rzetelności badawczej. Otwarta nauka to również społeczność, zapewnia bowiem równy, szeroki dostęp do najnowszej, zwięzłej.

Zasady otwartej nauki zostały wdrożone przez wiele agencji finansujących badania, w tym m.in. przez Komisję Europejską. Są one również promowane przez Komisję Europejską odpowiedzialną za kształtowanie polityki naukowej (np. Science Europe, UNESCO).

Otwarta nauka dotyczy wielu aspektów procesów i artefaktów badawczych: publikacji, danych, protokołów, oprogramowania, czy też kodów źródłowych. W zakresie otwartego dostępu do publikacji i danych badawczych, a w raportach uwzględniono zasady DORA – *San Francisco Declaration on Research*.

[Pismo Dyrektora Narodowego Centrum Nauki w sprawie otwartego dostępu do danych badawczych](#)

Rozwiń wszystko →



Otwarty dostęp do publikacji



Otwarty dostęp do danych badawczych

Dokumenty i informacje dodatkowe:

- [POLITYKA NARODOWEGO CENTRUM NAUKI DOTYCZĄCA OTWARTEGO DOSTĘPU DO PUBLIKACJI](#)
- [Instrukcja: Polityka NCN dotycząca otwartego dostępu do publikacji](#)
- [Otwarty dostęp do publikacji – najczęściej zadawane pytania](#)



Otwarty dostęp do danych badawczych

W 2019 roku NCN wdrożył do wniosków o finansowanie projektów badawczych, działań naukowych Plan Zarządzania Danymi – Data Management Plan (DMP). Formularz został oparty na wytycznych Science Europe. Jednocześnie, NCN opublikowało wytyczne wskazujące jak należy wypełnić Plan. DMP podlega ocenie eksperckiej na etapie oceny wniosku.

DMP może być zmieniane na etapie realizacji projektu. Nie trzeba informować NCN o każdorazowo wprowadzonych do niego zmianach. Dopiero w raporcie końcowym należy zamieścić informacje dotyczące zakresu zmian, a także uzasadnić konieczność ich wprowadzenia, co podlega ocenie ekspertów oceniających raporty końcowe.

Więcej informacji dotyczących DMP znajduje się w linkach poniżej.

Dokumenty i informacje dodatkowe:

- [Wytyczne dla wnioskodawców do uzupełnienia planu zarządzania danymi w projekcie badawczym](#)
- [Otwarty dostęp do danych – najczęściej zadawane pytania](#)

Znaczenie prawidłowego* zarządzania danymi

- ◆ Ochrona integralności danych i rzetelności badań
- ◆ Zgodność z wymaganiami NCN
- ◆ Ułatwienie (współ)pracy i ponownego wykorzystania danych



*stworzenia dobrego i wykonalnego planu

Część praktyczna

Walidacja planu zarządzania danymi

Przełożenie podstawowych wymagań na szablon DMP

Informacje ogólne

1. Opis danych, gromadzenie lub ponowne wykorzystanie istniejących danych

2. Dokumentacja i jakość danych

Informacje administracyjne

1a. W jaki sposób będą gromadzone lub tworzone nowe dane i/lub w jaki sposób istniejące dane będą ponownie wykorzystywane?

1b. Jakie dane (na przykład rodzaje, formaty i ilości) będą gromadzone lub generowane?

2a. Jakie metadane i dokumentacja (na przykład metodologia gromadzenia danych i sposób ich organizacji) będą towarzyszyć danym?

2b. Jakie środki kontroli jakości danych będą stosowane?

Przełożenie podstawowych wymagań na szablon DMP

3. Przechowywanie i tworzenie kopii zapasowych podczas procesu badawczego

3a. W jaki sposób dane i metadane będą przechowywane i archiwizowane podczas procesu badawczego?

3b. W jaki sposób zostanie zapewnione bezpieczeństwo danych i ochrona danych wrażliwych podczas badania?

4. Wymogi prawne i etyczne, kodeksy postępowania

4a. Jeśli dane osobowe są przetwarzane, w jaki sposób zostanie zapewniona zgodność z przepisami dotyczącymi danych osobowych i bezpieczeństwa danych?

4b. W jaki sposób będą zarządzane inne kwestie prawne, takie jak prawa własności intelektualnej i prawo własności? Jakie przepisy mają zastosowanie?

4c. W jaki sposób będą brane pod uwagę ewentualne kwestie etyczne i przestrzegane kodeksy postępowania?

Przełożenie podstawowych wymagań na szablon DMP

5. Udostępnianie i długoterminowa ochrona danych

5a. Jak i kiedy dane będą udostępniane? Czy istnieją możliwe ograniczenia w udostępnianiu danych lub powody embarga?

5b. W jaki sposób dane do zachowania będą wybierane i gdzie będą przechowywane długoterminowo (np. w repozytorium danych lub archiwum)?

5c. W jaki sposób zapewnione zostanie zastosowanie unikalnego i trwałego identyfikatora (takiego jak identyfikator obiektu cyfrowego (DOI)) do każdego zestawu danych?

5d. Czy będą potrzebne metody lub narzędzia programowe do dostępu i wykorzystania danych?

6. Obowiązki i zasoby związane z zarządzaniem danymi

6a. Kto (na przykład rola, stanowisko i instytucja) będzie odpowiedzialny za zarządzanie danymi (tj. Data Steward)?

6b. Jakie zasoby (np. finansowe i czasowe) zostaną przeznaczone na zarządzanie danymi i zapewnienie, że dane będą FAIR (możliwe do znalezienia, dostępne, interoperacyjne, nadające się do ponownego użycia)?

Zadanie: Walidacja planu zarządzania danymi

Twoim zadaniem jest dokładne sprawdzenie i zatwierdzenie przygotowanego planu zarządzania danymi (DMP). Celem jest upewnienie się, że DMP spełnia wszystkie wymagane standardy i wytyczne, dokładnie odzwierciedla praktyki zarządzania danymi i odnosi się do wszystkich istotnych aspektów, takich jak gromadzenie, przechowywanie, udostępnianie i bezpieczeństwo danych. Pracując w zespołach/parach/indywidualnie, zidentyfikuj wszelkie niespójności, luki lub obszary wymagające poprawy w ramach planu.

Przełącz konstruktywne opinie i sugestie dotyczące ulepszeń, aby upewnić się, że DMP jest solidny, kompleksowy i gotowy do wdrożenia w rzeczywistych projektach. Ćwiczenie to ma na celu pogłębienie zrozumienia zasad zarządzania danymi i poprawę jakości zarządzania danymi w organizacji.



https://dv-website.s3.amazonaws.com/uploads/2022/08/kf_writingadatamanagementplan_aug22.jpg

Plan zarządzania danymi: Identyfikacja genów odporności na suszę w genomie *Avena sativa*



1. Opis danych i gromadzenie lub ponowne wykorzystanie istniejących danych

- 1a. W jaki sposób będą gromadzone lub tworzone nowe dane i/lub w jaki sposób istniejące dane będą ponownie wykorzystywane?

Dane zostaną zebrane przez zespół projektowy w trakcie eksperymentów laboratoryjnych oraz na podstawie obserwacji wzrostu roślin w warunkach suszy. Planujemy także wyszukiwać w internecie dane o owiesie, które mogą się przydać. Gdy znajdziemy jakieś ciekawe informacje w artykułach naukowych albo na stronach instytucji badawczych, to po prostu je skopiujemy i dodamy do naszej analizy. Nie będziemy prowadzić szczegółowej ewidencji, skąd pochodzą konkretne dane – ważne, że będą użyteczne.

UWAGA !!!

1. **Brak konkretnych metodologii i technologii gromadzenia danych.**
2. **Brak informacji na temat analizy genetycznej**
3. **Brak określonych źródeł istniejących danych genomowych do wykorzystania.**

Plan zarządzania danymi: Identyfikacja genów odporności na suszę w genomie *Avena sativa*



- 1a. W jaki sposób będą gromadzone lub tworzone nowe dane i/lub w jaki sposób istniejące dane będą ponownie wykorzystywane?

Dane pierwotne zostaną zgromadzone w wyniku eksperymentów laboratoryjnych, w tym sekwencjonowania genomów owsa zwyczajnego *Avena sativa** (odmiany Hendon, Masacni, Hi-Fi, Aspen) w warunkach kontrolowanych i stresowych (susza). Eksperyment suszy będzie przeprowadzony w szklarni na podstawie protokołu opisanego przez Bisaga et al. (2025) Dodatkowo wykorzystane zostaną dane (wtórne) genomowe z otwartych repozytoriów, takich jak NCBI (<https://www.ncbi.nlm.nih.gov>) i. Ensembl Plants <https://plants.ensembl.org/index.html>. Wszystkie ponownie użyte dane będą szczegółowo opisane z podaniem źródła, licencji oraz warunków użycia w pliku README.txt.

Plan zarządzania danymi: Identyfikacja genów odporności na suszę w genomie *Avena sativa*



- 1b. Jakie dane (na przykład rodzaje, formaty i ilości) będą gromadzone lub generowane?

Nie wiemy jeszcze dokładnie, ile danych zbierzemy, bo zależy to od wyników eksperymentów. Na pewno będą to jakieś pliki tekstowe, arkusze kalkulacyjne i wykresy. Może też trochę zdjęć albo plików z sekwencjami DNA, ale to się jeszcze okaże. Nie mamy ustalonych formatów zapisu ani konkretnych nazw plików – zobaczymy, jak będzie najwygodniej.

UWAGA !!!

- 1. Brak wskazania formatów danych genomowych.**
- 2. Brak oszacowania ilości danych**

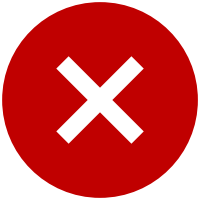
Plan zarządzania danymi: Identyfikacja genów odporności na suszę w genomie *Avena sativa*



- 1b. Jakie dane (na przykład rodzaje, formaty i ilości) będą gromadzone lub generowane?

Projekt wygeneruje dane sekwencyjne (FASTQ), dane opisowe: ilościowe i jakościowe (CSV, XLSX), pliki z analiz bioinformatycznych (VCF, BAM) oraz dokumentację (PDF, DOCX). Szacunkowa wielkość danych to ok. 300 GB.

Plan zarządzania danymi: Identyfikacja genów odporności na suszę w genomie *Avena sativa*



2. Dokumentacja i jakość danych

- 2a. Jakie metadane i dokumentacja będą towarzyszyć danym?

Nie przewidujemy tworzenia dodatkowej dokumentacji poza samymi plikami z danymi. Wszystko będzie nazwane intuicyjnie, więc nie powinno być problemów z odnalezieniem informacji. Każdy członek zespołu będzie miał swój folder, gdzie zapisze dane w sposób, który uzna za najwygodniejszy. Nie planujemy stosowania żadnych standardów metadanych – szkoda na to czasu.

UWAGA !!!

- 1. Brak informacji na temat przyjętych standardów metadanych**
- 2. Brak informacji, czy dla zbiorów danych zostanie utworzona dokumentacja w postaci pliku README.txt.**
- 3. Brak jasnej organizacji, struktury folderów i kontroli wersji.**

Plan zarządzania danymi: Identyfikacja genów odporności na suszę w genomie *Avena sativa*

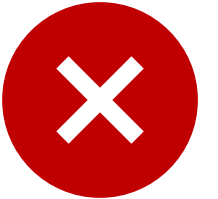


2. Dokumentacja i jakość danych

- 2a. Jakie metadane i dokumentacja będą towarzyszyć danym?

Każdy zbiór danych będzie posiadał opis metadanych zgodny ze standardem MIAPPE oraz Dublin Core. Zostanie udokumentowana metodologia eksperymentu, organizacja danych (struktura danych i wersja), warunki zbierania danych, oprogramowanie i wersje używane w analizach, licencje na których można wykorzystać dane, powiązane z danymi publikacje oraz słowa kluczowe w pliku README.txt

Plan zarządzania danymi: Identyfikacja genów odporności na suszę w genomie *Avena sativa*



- 2b. Jakie środki kontroli jakości danych będą stosowane?

Nie planujemy żadnych formalnych procedur weryfikacji danych. Uważamy, że skoro zespół składa się z doświadczonych badaczy, to wszystko będzie zrobione dobrze. W razie jakichś błędów będzie można poprawić dane ręcznie, kiedy już je zauważymy podczas pisania publikacji.

UWAGA !!!

1. Niespójne i niezdefiniowane środki kontroli jakości.

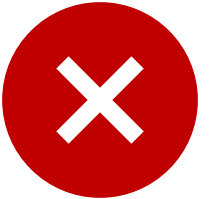
Plan zarządzania danymi: Identyfikacja genów odporności na suszę w genomie *Avena sativa*



- 2b. Jakie środki kontroli jakości danych będą stosowane?

Zostaną wdrożone procedury kontroli jakości, takie jak filtrowanie odczytów niskiej jakości (FastQC), walidacja formatów plików oraz testy replikacji biologicznej. Dane będą przeglądane przez co najmniej dwóch członków zespołu. Jakość danych będzie również kontrolowana poprzez odpowiedni dobór punktów pomiarowych, powtarzanie pomiarów, wykorzystanie profesjonalnego i licencjonowanego sprzętu badawczego, świeżych odczynników chemicznych, przeprowadzenie badania pilotażowego oraz kalibrację urządzeń pomiarowych przed rozpoczęciem badań.

Plan zarządzania danymi: Identyfikacja genów odporności na suszę w genomie *Avena sativa*



3. Przechowywanie i tworzenie kopii zapasowych podczas procesu badawczego

- 3a. W jaki sposób dane i metadane będą przechowywane i archiwizowane podczas procesu badawczego?

Dane będą przechowywane na prywatnych komputerach członków zespołu, głównie na laptopie kierownika projektu. Kopie zapasowe będą wykonywane nieregularnie – zależnie od potrzeb. Planujemy używać pendrive'ów i zewnętrznych dysków, bo to wygodne i szybkie. Dane będą organizowane według uznania każdego badacza – nie będziemy narzucać jednej struktury folderów.

UWAGA !!!

- 1. Danych badawczych nie przechowuje się na laptopie ani na dyskach zewnętrznych.**
- 2. Nieregularne tworzenie kopii zapasowych bez instytucjonalnych rozwiązań w zakresie przechowywania danych.**
- 3. Brak strategii organizacji danych badawczych**

Plan zarządzania danymi: Identyfikacja genów odporności na suszę w genomie *Avena sativa*

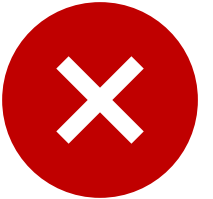


3. Przechowywanie i tworzenie kopii zapasowych podczas procesu badawczego

- 3a. W jaki sposób dane i metadane będą przechowywane i archiwizowane podczas procesu badawczego?

Dane badawcze będą zapisywane zgodnie z zasadą 3+2+1. Początkowo dane będą przechowywane na komputerach PC i laptopach badaczy (zabezpieczonych hasłami), sieciowych pamięciach masowych oraz chmurach (chmura Office 365 Education A1 z 1 TB na badacza - umowa z firmą Microsoft zapewnia bezpieczeństwo danych. Główny badacz będzie odpowiedzialny za comiesięczne tworzenie kopii zapasowych danych. Metadane wraz z danymi badawczymi które zostaną opublikowane będą przechowywane w repozytorium RePOD lub ZENODO.

Plan zarządzania danymi: Identyfikacja genów odporności na suszę w genomie *Avena sativa*



- 3b. W jaki sposób zostanie zapewnione bezpieczeństwo danych i ochrona danych wrażliwych podczas badania?

Dane będą przechowywane na urządzeniach zabezpieczonych hasłem. W razie awarii laptopa dane będzie można odzyskać z dysku zewnętrznego (o ile backup zostanie wykonany). Nie planujemy stosowania szyfrowania ani specjalnych procedur bezpieczeństwa, bo uważamy, że nasze dane nie są szczególnie wrażliwe.

UWAGA !!!

- 1. Niewystarczające środki bezpieczeństwa, brak szyfrowania.**
- 2. Brak konkretnego planu odzyskiwania danych.**

Plan zarządzania danymi: Identyfikacja genów odporności na suszę w genomie *Avena sativa*



- 3b. W jaki sposób zostanie zapewnione bezpieczeństwo danych i ochrona danych wrażliwych?

Dane badawcze będą zapisywane na serwerach uniwersyteckich, dostęp do danych na nich będzie zabezpieczony loginem i hasłem, dostęp do nich jest możliwy tylko z Intranetu (sieci lokalnej).

Dostęp do/z Internetu jest możliwy tylko przy użyciu Wirtualnej Sieci Prywatnej, a cały dostęp przez VPN jest logowany na routerze. Każdy użytkownik będzie miał własne dane logowania. W przypadku OwnCloud dostęp jest możliwy z dowolnego miejsca na świecie. Dostęp jest możliwy po wprowadzeniu loginu i hasła. W przypadku lokalnej utraty danych, zostaną one przywrócone z istniejących nośników kopii zapasowych (dyski zewnętrzne/chmura/NAS). W przypadku utraty danych w chmurze/NAS, dane zostaną przeniesione z nośnika kopii zapasowej lub lokalnej maszyny. PI będzie nadzorować procesy tworzenia i przywracania kopii zapasowych. Dane wrażliwe nie będą gromadzone ani przetwarzane.

Plan zarządzania danymi: Identyfikacja genów odporności na suszę w genomie *Avena sativa*



4. Wymagania prawne i etyczne, kodeksy postępowania

- 4a. Jak zapewniona zostanie zgodność z przepisami dotyczącymi danych osobowych?

Projekt nie obejmuje danych osobowych. W przypadku pojawienia się takich danych, zostanie uzyskana świadoma zgoda uczestników, a dane zostaną odpowiednio zanonimizowane zgodnie z RODO.

Na razie nie planujemy przetwarzać danych osobowych, ale jeśli zajdzie taka potrzeba, to podejmiemy odpowiednie kroki. Uczestnicy badań będą informowani ustnie o tym, że ich dane są zbierane. Dokumentacja zgód nie będzie konieczna, bo badania nie dotyczą ludzi bezpośrednio.

UWAGA !!!

- 1. Brak strategii anonimizacji lub pseudonimizacji.**
- 2. Niewystarczające traktowanie przepisów RODO**

Plan zarządzania danymi: Identyfikacja genów odporności na suszę w genomie *Avena sativa*



4. Wymagania prawne i etyczne, kodeksy postępowania

- 4a. Jak zapewniona zostanie zgodność z przepisami dotyczącymi danych osobowych?

Projekt nie obejmuje danych osobowych. W przypadku pojawienia się takich danych, zostanie uzyskana świadoma zgoda uczestników, a dane zostaną odpowiednio zanonimizowane zgodnie z RODO.

Projekt nie ma na celu gromadzenia danych osobowych. Dane osobowe mogą dotyczyć wyłącznie wykonawców projektu. Stosowane będą wówczas następujące wytyczne:

1. Dane osobowe dotyczące członków zespołu będą przetwarzane wyłącznie do celów administracyjnych, a nie jako dane badawcze.
2. Udział w szkoleniach na temat prawnych aspektów ochrony danych wrażliwych, otwartego dostępu i etyki badań jest zapewniany przez Wydział i Uniwersytet i/lub dostawców zewnętrznych.
3. Dane osobowe członków zespołu będą gromadzone, przechowywane i przetwarzane zgodnie z obowiązującymi przepisami prawa i wewnętrznymi politykami w celu zapewnienia bezpieczeństwa informacji wrażliwych.

Uczestnicy zostaną poinformowani, że w każdej chwili mogą wnieść skargę w związku z przetwarzaniem ich danych osobowych do polskiego i europejskiego Inspektora Ochrony Danych Osobowych; mają prawo do wycofania się z badań na każdym etapie.

Plan zarządzania danymi: Identyfikacja genów odporności na suszę w genomie *Avena sativa*



- 4b. W jaki sposób będą zarządzane inne kwestie prawne, takie jak prawa własności intelektualnej i prawo własności?

Zakładamy, że wszystkie dane i wyniki projektu będą należały do kierownika projektu, bo to on odpowiada za projekt i jego realizację. Nie przewidujemy spisywania umów dotyczących własności danych ani podziału praw autorskich między członków zespołu.

Plan zarządzania danymi: Identyfikacja genów odporności na suszę w genomie *Avena sativa*



4b. W jaki sposób będą zarządzane inne kwestie prawne, takie jak prawa własności intelektualnej i prawo własności?

Własność intelektualna wytworzona w ramach projektu, w tym wyniki badań, raporty, wynalazki i know-how, które mogą stanowić podstawę przyszłych zgłoszeń patentowych, będzie należała do Uniwersytetu Śląskiego w Katowicach jako pracodawcy naukowców, którzy ją wytworzyli zgodnie z:

- Prawo autorskie
- Prawo własności intelektualnej
- Ustawą o ochronie baz danych
- Regulaminem konkursu grantowego i umową z NCN
- Licencje na oprogramowanie/sprzęt/repozytorium
- Zarządzenie nr 38/2020 i 40/2020 Dyrektora Narodowego Centrum Nauki z dnia 27 maja 2020 r. w sprawie określenia Polityki Narodowego Centrum Nauki w zakresie otwartego dostępu do publikacji
- Polityka Bezpieczeństwa Informacji Uniwersytetu Śląskiego (Zarządzenie nr 153/2018 z późn. zm.)
- Instytucjonalna polityka otwartej nauki Uniwersytetu Śląskiego w Katowicach
- Zasady ochrony, korzystania i zarządzania wynikami pracy intelektualnej w Uniwersytecie Śląskim w Katowicach.

Plan zarządzania danymi: Identyfikacja genów odporności na suszę w genomie *Avena sativa*



5. Udostępnianie i długoterminowa ochrona danych

- 5a. Jak i kiedy dane będą udostępniane?

Dane zostaną udostępnione, jeśli ktoś o nie poprosi. Nie planujemy ich publikować w otwarty sposób. Być może niektóre dane zostaną dołączone jako załączniki do artykułów. Jeżeli uznamy, że dane są za cenne, możemy zdecydować, że nie będą upubliczniane.

UWAGA !!!

- 1. Natychmiastowe udostępnianie danych bez okresów embarga.**
- 2. Brak określonego repozytorium do długoterminowego przechowywania, można wybrać tylko jedno.**

Plan zarządzania danymi: Identyfikacja genów odporności na suszę w genomie *Avena sativa*



5. Udostępnianie i długoterminowa ochrona danych

- 5a. Jak i kiedy dane będą udostępniane?

Dane zebrane w ramach projektu i sfinansowane ze środków publicznych będą niezwłocznie udostępnione przez co najmniej 10 lat (zgodnie ze standardami NCN) (metadane bez ograniczeń czasowych) w repozytoriach takich jak Zenodo lub RepOD (wszystkie na licencji Creative Commons CC BY lub Creative Commons CC 0 4.0), co najmniej w momencie publikacji artykułu (artykuły i dane uzupełniające jako zbiory danych z DOI) (udostępniane publicznie w Open Access) bez ograniczeń wielkości i zgodnie z zasadą FAIR. Dane będą dostępne za pośrednictwem Repozytorium, ale także na żądanie. Nie przewiduje się embarga czasowego na dane wytworzone podczas projektu.

Plan zarządzania danymi: Identyfikacja genów odporności na suszę w genomie *Avena sativa*



- 5b. W jaki sposób dane do zachowania będą wybierane i gdzie będą przechowywane długoterminowo?

Po zakończeniu projektu wybierzemy te pliki, które wydadzą się najważniejsze. Planujemy nagrać je na płyty DVD i przechować w biurze. Nie zamierzamy korzystać z żadnych repozytoriów ani archiwów instytucjonalnych.

UWAGA !!!

- 1. Brak strategii przechowywania danych**
- 2. Zabronione są płyty DVD**

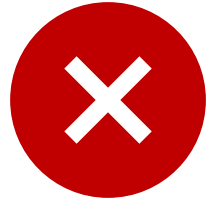
Plan zarządzania danymi: Identyfikacja genów odporności na suszę w genomie *Avena sativa*



- 5b. W jaki sposób dane do zachowania będą wybierane i gdzie będą przechowywane długoterminowo?

Do przechowywania danych wybierane będą wyłącznie zweryfikowane i zatwierdzone zbiory danych, aby zapewnić dokładność i wiarygodność przechowywanych informacji. Niepublikowane, nieistotne dane będą przechowywane na wewnętrznych serwerach i w chmurze. Dane te można również udostępnić po skontaktowaniu się z koordynatorem projektu. Gwarantuje to, że podczas gdy zweryfikowane dane są otwarcie dostępne, wszystkie inne dane są nadal dostępne do wewnętrznego przeglądu i potencjalnej przyszłej weryfikacji.

Plan zarządzania danymi: Identyfikacja genów odporności na suszę w genomie *Avena sativa*



- 5c. W jaki sposób zapewnione zostanie zastosowanie unikalnego i trwałego identyfikatora (np. DOI)?

Nie planujemy korzystania z identyfikatorów takich jak DOI, ponieważ nie wiemy, jak je nadać i nie są nam potrzebne.

UWAGA !!!

1. Data set musi mieć nadany numer DOI

Plan zarządzania danymi: Identyfikacja genów odporności na suszę w genomie *Avena sativa*



- 5c. W jaki sposób zapewnione zostanie zastosowanie unikalnego i trwałego identyfikatora (np. DOI)?

Każdy opublikowany zestaw danych otrzyma DOI nadany przez repozytorium danych badawczych (np. Zenodo lub RePOD).

Plan zarządzania danymi: Identyfikacja genów odporności na suszę w genomie *Avena sativa*



- 5d. Czy będą potrzebne metody lub narzędzia programowe do dostępu i wykorzystania danych?

Część danych może wymagać specjalistycznego oprogramowania, ale nie mamy jeszcze decyzji, czy udostępnimy informacje o tych narzędziach. Jeśli ktoś będzie chciał skorzystać z danych, to sam sobie poradzi.

UWAGA !!!

- 1. Brak określonego typu licencji CC**
- 2. Brak wskazania, jakie programy są potrzebne do analizy danych**

Plan zarządzania danymi: Identyfikacja genów odporności na suszę w genomie *Avena sativa*



- 5d. Czy będą potrzebne metody lub narzędzia programowe do dostępu i wykorzystania danych?

Tak, dane genomowe wymagają oprogramowania bioinformatycznego. Instrukcja korzystania z danych oraz lista wymaganych narzędzi zostanie udostępniona razem z danymi. Natomiast reszta danych nie będzie wymagała specjalistycznego oprogramowania i będzie udostępniona w otwartych formatach zgodnie z zasadami FAIR.

Plan zarządzania danymi: Identyfikacja genów odporności na suszę w genomie *Avena sativa*



6. Obowiązki i zasoby związane z zarządzaniem danymi

- 6a. Kto będzie odpowiedzialny za zarządzanie danymi?

Nie mamy osoby dedykowanej do zarządzania danymi. Kierownik projektu zajmie się wszystkim samodzielnie – w miarę dostępności czasu

UWAGA !!!

- 1. Niejasny podział obowiązków bez konkretnych ról.**
- 2. Brak dedykowanego Data stewarda**

Plan zarządzania danymi: Identyfikacja genów odporności na suszę w genomie *Avena sativa*

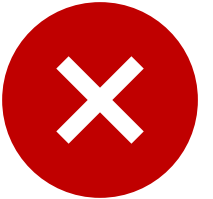


6. Obowiązki i zasoby związane z zarządzaniem danymi

6a. Kto będzie odpowiedzialny za zarządzanie danymi?

Kierownik projektu będzie odpowiedzialny za zarządzanie danymi i monitorowanie wdrażania planu zarządzania danymi przy wsparciu wydziałowego Data stewarda: Zespół Infrastruktury IT Kampusu Katowice
bezpieczeństwo i odzyskiwanie danych: Biuro Rzecznika Patentowego i Biuro Radców Prawnych
Uniwersytetu Śląskiego: prawo autorskie i prawo własności intelektualnej; Koordynator Otwartej Nauki:
zarządzanie danymi, przechowywanie i archiwizacja danych, aktualizacja i raportowanie DMP, regulacje i
polityka Otwartej Nauki.

Plan zarządzania danymi: Identyfikacja genów odporności na suszę w genomie *Avena sativa*



6b. Jakie zasoby zostaną przeznaczone na zarządzanie danymi i zapewnienie, że dane będą FAIR?

Nie przewidujemy dodatkowych kosztów ani zasobów na działania związane z danymi. Zarządzanie danymi nie jest głównym celem projektu, więc nie traktujemy go priorytetowo. Nie planujemy stosować zasad FAIR, bo to za bardzo komplikuje pracę.

UWAGA !!!

- 1. Brak oszacowania wymaganych zasobów (czasowych, finansowych, infrastrukturalnych).**
- 2. Błędne założenie, że nie są potrzebne żadne dodatkowe zasoby.**

Plan zarządzania danymi: Identyfikacja genów odporności na suszę w genomie *Avena sativa*



6b. Jakie zasoby zostaną przeznaczone na zarządzanie danymi i zapewnienie, że dane będą FAIR?

Nie przewiduje się dodatkowych środków finansowych aby gromadzić i przechowywać dane badawcze powstałe podczas projektu.