



# Warsztaty z zarządzania danymi badawczymi

Natalia Gruenpeter

CC BY 4.0

Treść licencji: <https://creativecommons.org/licenses/by/4.0/legalcode.pl>





# Platforma Otwartej Nauki



Interdyscyplinarne Centrum  
Modelowania Matematycznego  
i Komputerowego Uniwersytetu  
Warszawskiego



## REPOZYTORIUM

*Prace udostępniane przez polskich naukowców*



## BIBLIOTEKA NAUKI

*Czasopisma w otwartym dostępie*



## OTWÓRZ KSIĄŻKĘ

*Książki w otwartym dostępie*



## Aktualności

26.04.2019

### Platforma Polskich Publikacji Naukowych

Jednym z projektów realizowanych przez nas w ICM UW jest Platforma Polskich...

[Czytaj więcej](#)

09.04.2019

### Warsztaty z zarządzania danymi badawczymi w Warszawie

Rozpoczęliśmy wczoraj serię warsztatów, których celem jest przybliżenie...

[Czytaj więcej](#)

[Zobacz więcej](#) →



Conferences

Łódź  
Warsztaty z zarządzania  
danymi badawczymi **11**  
06.2019



O Platformie  
Otwartej Nauki



Wideo



Czasopisma  
w otwartym dostępie



Nasze  
publikacje



Narzędzia  
i rozwiązania



Materiały  
informacyjne i promocyjne



Kontakt  
i współpraca



# OpenAIRE

- europejska infrastruktura otwartej nauki,  
m.in. repozytorium Zenodo

The Zenodo logo is displayed in white lowercase letters on a blue rectangular background. The letters are stylized and rounded.

Funded by





# Guiding you in Open Science

## RESOURCES

Open Science Primers

Guides

Factsheets

Use cases

## HELPDESK

FAQs

Ask a Question

## TRAINING

Webinars

Workshops

Community of  
Practice

## Researchers

How to comply with H2020 mandates - publications

How to comply with H2020 mandates - research data

Research Data Management costs in H2020 projects

How to make your data FAIR

---

**W moich badaniach mam do  
czynienia z danymi...**

---

**Co to są dane badawcze?  
Dlaczego należy i warto  
je udostępniać?**



## Dane badawcze

Zarejestrowane materiały o charakterze faktograficznym, powszechnie uznawane przez społeczność naukową za niezbędne do oceny wyników badań naukowych.

Dane zebrane, zaobserwowane lub wytworzone jako materiał do analizy, w celu uzyskania oryginalnych wyników naukowych.

Wszystko co zostało wyprodukowane lub wytworzone w ramach prowadzonych badań.





## Dane badawcze

- dane liczbowe,
- dokumenty tekstowe, notatki,
- kwestionariusze, wyniki badań ankietowych,
- nagrania audio i wideo, fotografie,
- zawartość baz danych,
- oprogramowanie,
- wyniki symulacji komputerowych,
- protokoły laboratoryjne, opisy metodologiczne
- ...



## Dlaczego nas to dotyczy?

- polityka grantodawców
- wymogi czasopism publikujących artykuły naukowe
- przepisy prawa, dostęp do informacji publicznej

Zbigniew Blocki  
Dyrektor  
Narodowego Centrum Nauki  
ZKA.071.1.2019

Kraków, 03-04-2019

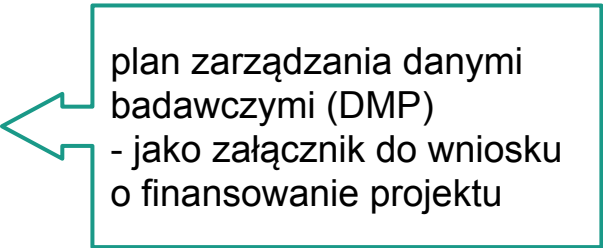
Szanowni Państwo,

mając na uwadze fakt, że Komisja Europejska w swoich *Zaleceniach z dnia 17 lipca 2012 r. w sprawie dostępu do informacji naukowej i jej ochrony* wezwwała kraje członkowskie, aby jasno określiły strategie rozpowszechniania i otwartego dostępu do publikacji naukowych, Ministerstwo Nauki i Szkolnictwa Wyższego w przyjętym w 2015 roku dokumencie zatytułowanym *Kierunki rozwoju otwartego dostępu do publikacji i wyników badań naukowych w Polsce*<sup>1</sup> sformułowało wytyczne dotyczące wprowadzenia otwartego dostępu przez jednostki naukowe i uczelnie poprzez:

1. przyjęcie instytucjonalnej polityki otwartego dostępu;
2. wyznaczenie przez kierowników jednostek naukowych i uczelni pełnomocników ds. *Open Access* (OA);
3. wspieranie modelu otwartego czasopism oraz
4. korzystanie z doświadczeń i potencjału bibliotek prowadzących otwarte repozytoria.

Mając powyższe na uwadze, wraz z innymi agencjami finansującymi badania naukowe zrzeszonymi w Science Europe, podjęliśmy działania mające na celu wypracowanie wspólnych europejskich wytycznych dotyczących zarządzania danymi naukowymi i sposobu ich udostępniania w tzw. otwartym dostępie. Praktyczny przewodnik dot. ujednoczonych europejskich praktyk związanych z zarządzaniem danymi naukowymi został opublikowany w styczniu br.<sup>2</sup>

W oparciu o wytyczne zawarte w tym przewodniku planujemy jeszcze w tym roku wprowadzić do formularza wniosku o finansowanie projektu załącznik, w którym wnioskodawca przedstawi skrócony plan zarządzania danymi badawczymi (z ang. *Data Management Plan*), które powstaną w ramach realizacji projektu. Plan ten będzie uzupełniany na etapie składania raportu końcowego i oceniany po zakończeniu realizacji projektu.



plan zarządzania danymi  
badawczymi (DMP)  
- jako załącznik do wniosku  
o finansowanie projektu



Zbigniew Blocki  
Dyrektor  
Narodowego Centrum Nauki  
ZKA.071.1.2019

Kraków, 03-04-2019

Szanowni Państwo,

mając na uwadze fakt, że Komisja Europejska w swoich *Zaleceniach z dnia 17 lipca 2012 r. w sprawie dostępu do informacji naukowej i jej ochrony* wezwwała kraje członkowskie, aby jasno określiły strategię rozpowszechniania i otwartego dostępu do publikacji naukowych, Ministerstwo Nauki i Szkolnictwa Wyższego w przyjętym w 2015 roku dokumencie zatytułowanym *Kierunki rozwoju otwartego dostępu do publikacji i wyników badań naukowych w Polsce*<sup>1</sup> sformułowało wytyczne dotyczące wprowadzenia otwartego dostępu przez jednostki naukowe i uczelnie poprzez:

**Będzie to pierwszy etap wprowadzania przez Narodowe Centrum Nauki polityki otwartego dostępu do danych naukowych wytworzonych/wykorzystanych w ramach realizacji projektów finansowanych przez Centrum, a w niedalekiej perspektywie również do publikacji naukowych.**

europejskich praktyk związanych z zarządzaniem danymi naukowymi został opublikowany w styczniu br.<sup>2</sup>

W oparciu o wytyczne zawarte w tym przewodniku planujemy jeszcze w tym roku wprowadzić do formularza wniosku o finansowanie projektu załącznik, w którym wnioskodawca przedstawi skrócony plan zarządzania danymi badawczymi (z ang. *Data Management Plan*), które powstaną w ramach realizacji projektu. Plan ten będzie uzupełniany na etapie składania raportu końcowego i oceniany po zakończeniu realizacji projektu.

plan zarządzania danymi  
badawczymi (DMP)  
- jako załącznik do wniosku  
o finansowanie projektu

**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. 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
<b>1. Opis danych oraz pozyskiwanie lub ponowne wykorzystanie dostępnych danych</b>	
1.1 W jaki sposób będą pozyskiwane lub wytwarzane nowe dane lub ponownie wykorzystywane dane już istniejące? Warto rozważyć następujące kwestie: <ul style="list-style-type: none"><li>- Jakie standardy, metody i oprogramowanie posłużą do pozyskiwania lub wytwarzania nowych danych?</li><li>- Jakie zostaną wdrożone procesy kontroli jakości?</li><li>- Jakie dane już istniejące (własne lub będące w posiadaniu strony trzeciej) zostaną ponownie wykorzystane?</li><li>- W jaki sposób zostanie udokumentowane pochodzenie danych?</li><li>- Jak wyglądać będzie organizacja plików i zarządzanie ich różnymi wersjami?</li></ul>	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: <ul style="list-style-type: none"><li>- Jaki rodzaj, format i objętość danych planują Państwo pozyskać, wytworzyć lub ponownie wykorzystać?</li></ul>	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



# Program Horyzont 2020

- publikacje naukowe - wymóg otwartego udostępniania publikacji
- dane badawcze - pilotażowy program udostępniania danych wedle określonych zasad



- **“open by default” - domyślnie dane powinny zostać udostępnione jak najszybciej co najmniej w zakresie koniecznym do weryfikacji ustaleń zawartych w publikacjach,**
- **opcja “opt out” - w kilku uzasadnionych sytuacjach,**
- **“as open as possible, as closed as necessary”,**
- **zasady FAIR**



# Program Horyzont 2020

- możliwość “opt out” w uzasadnionych sytuacjach,
- można z niej skorzystać na każdym etapie

## Opting out – partially or entirely

By extending the pilot, open access becomes the default setting for research data generated in Horizon 2020.

However, not all data can be open. Projects can therefore opt out at any stage (either before or after signing the grant) and so free themselves retroactively from the obligations associated with the conditions – if:

- participation is incompatible with the obligation to protect results that can reasonably be expected to be commercially or industrially exploited
- participation is incompatible with the need for confidentiality in connection with security issues
- participation is incompatible with rules on protecting personal data
- participation would mean that the project's main aim might not be achieved
- the project will not generate / collect any research data or
- there are other legitimate reasons (you can enter these in a free-text box at the proposal stage).

The Commission's approach can therefore be described as “as open as possible, as closed as necessary”.



# Dane sektora publicznego

- dyrektywa zatwierdzona przez Parlament Europejski
- dane badawcze jako dane sektora publicznego (badania finansowane ze środków publicznych),
- zasady udostępniania - na wzór programu Horyzont 2020

## Article 10

### **Research data**

1. Member States shall support the availability of research data by adopting national policies and relevant actions aiming at making publicly funded research data openly available ('open access policies'), *following the principle of “open by default” and compatible with the FAIR principles. In that context, concerns relating to intellectual property rights, personal data protection and confidentiality, security and legitimate commercial interests, shall be taken into account in accordance with the principle of “as open as possible, as closed as necessary”. Those open access policies shall be addressed to research performing organisations and research funding organisations.*
2. Without prejudice to point (c) of Article 1(2), research data shall be re-usable for commercial or non-commercial purposes in accordance with Chapters III and IV, insofar as they are publicly funded and *researchers, research performing organisations or research funding organisations have already made them publicly available* through an institutional or subject-based repository. In that context, legitimate commercial interests, *knowledge transfer activities* and pre-existing intellectual property rights shall be taken into account.



## Komisja Europejska - Oświadczenie

### Jednolity rynek cyfrowy: Komisja z zadowoleniem przyjmuje wynik głosowania Parlamentu Europejskiego w sprawie nowych przepisów dotyczących udostępniania danych sektora publicznego

Bruksela, 4 kwietnia 2019 r.

**Parlament Europejski zatwierdził dzisiaj dyrektywę w sprawie otwartych danych i informacji dotyczących sektora publicznego, która znacznie poprawi dostępność i innowacyjne wykorzystanie danych publicznych i finansowanych ze środków publicznych. A to z kolei przyczyni się do rozwoju technologii wymagających dużych ilości danych, takich jak sztuczna inteligencja.**

Wiceprzewodniczący do spraw jednolitego rynku cyfrowego Andrus **Ansip** i komisarz do spraw gospodarki cyfrowej i społeczeństwa cyfrowego Marija **Gabriel** we wspólnym oświadczeniu z zadowoleniem przyjęli ten wynik głosowania:

*Jesteśmy bardzo zadowoleni z powszechnego poparcia dla dyrektywy w sprawie otwartych danych wśród posłów do Parlamentu Europejskiego.*

*Dane są kluczowym atutem gospodarki cyfrowej, a znaczna ich część pochodzi z sektora publicznego. Ważne jest, aby obywatele i przedsiębiorstwa w UE mieli łatwy dostęp do danych publicznych i finansowanych ze środków publicznych, aby w pełni wykorzystać potencjał innowacyjny takich danych. Dotyczy to w szczególności technologii sztucznej inteligencji, których rozwój zależy od dostępu do ogromnych ilości danych.*

*Wzmocnienie przepisów UE w zakresie ponownego wykorzystywania informacji sektora publicznego jest konieczne, aby dostosować je do szybko zmieniających się technologii i usunąć bariery, które uniemożliwiają MŚP i przedsiębiorstwom typu start-up wdrażanie innowacji z wykorzystaniem publicznych zasobów danych.*

*Dzięki przyjętej w dniu dzisiejszym dyrektywie w sprawie otwartych danych dane sektora publicznego w czasie rzeczywistym, a także dane pochodzące od organizacji prowadzących badania i przedsiębiorstw publicznych staną się łatwiej dostępne i będą z nich mogły korzystać zarówno maszyny, jak i ludzie.*


*Pragniemy podziękować posłom do Parlamentu Europejskiego, państwom członkowskim i zainteresowanym stronom za zapewnienie konstruktywnego i szybkiego procesu legislacyjnego, który doprowadził do dzisiejszego znaczącego głosowania.*

### Dalsze działania

Tekst przyjęty dziś przez Parlament Europejski musi zostać formalnie przyjęty przez Radę UE. Następnie państwa członkowskie będą musiały wdrożyć zmienione przepisy w ciągu dwóch lat, po czym wejdą one w życie. Komisja rozpocznie współpracę z państwami członkowskimi w celu ustalenia zbiorów danych o wysokiej wartości, które zostaną określone w akcie wykonawczym.

6.06.2019 - dyrektywa  
zatwierdzona przez  
Radę Unii Europejskiej

dane w formie nadającej  
się do odczytu  
maszynowego



## Dane badawcze w dyrektywie dot. udostępniania danych sektora publicznego

„Dane badawcze” zdefiniowano jako „dokumenty w formie cyfrowej, inne niż publikacje naukowe, które są gromadzone lub opracowywane 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”.



## Dlaczego należy udostępniać dane badawcze?

- Bo ułatwia to prowadzenie badań i pozwala skorzystać z danych w przyszłości.
- Bo pozwala to na przeprowadzenie nowych analiz i zachęca do nowych interpretacji.
- Bo pozwala to ocenić rzetelność naszych danych i uniknąć konieczności kilkukrotnego wytwarzania tych samych danych.
- Bo to uczciwe wobec publicznego fundatora badań.



## Dlaczego warto udostępniać dane badawcze?

- Bo udostępnione dane mogą być zacytowane.
- Bo dostępność danych zwiększa szanse na zacytowanie naszych publikacji.
- Bo dane to bardzo dobre źródło informacji o tym, czym i w jaki sposób się zajmujemy.

**„W sytuacji braku jasnego i wiarygodnego sygnału, że naukowcy udostępniający zgromadzone przez siebie dane otrzymają za to uznanie, zrozumienie oraz nagrodę, trudno będzie podjąć ten wysiłek w sposób spontaniczny.”**

Bernard Rentier, *Open Science, the Challenge of Transparency* (2019)

---

---


# Jak i gdzie udostępniać dane badawcze?



# Gdzie udostępniać dane badawcze?

data journals

niektóre data journals  
publikują listy  
rekomendowanych  
repozytoriów



- czasopisma publikujące opisy zestawów danych badawczych,
- artykuły są recenzowane, czasopisma działają na wzór tradycyjnych czasopism publikujących artykuły,
- zestawy danych zwykle deponowane są w repozytoriach, czasami załączane do artykułu i publikowane w czasopiśmie



# Data in Brief

Open access



Editors-in-Chief: Hao-Ran Wang, Ganhui Lan

[View editorial board](#)

[View aims and scope](#)

## Explore journal content

- [Latest issue](#)
- [Articles in press](#)
- [All issues](#)

 [Sign in to set up alerts](#)

 [RSS](#)

## Latest issues

[Volume 24](#)  
*In progress* (June 2019)

[Volume 23](#)  
*In progress* (April 2019)

[Volume 22](#)  
pp. 1–1094 (February 2019)

[Volume 21](#)  
pp. 1–2704 (December 2018)

[View all issues](#)

## Find out more

- [Submit your article](#)
- [Guide for authors](#)
- [About the journal](#)





## Research Data Journal for the Humanities and Social Sciences

*Contributing to transparency of research, accelerating dissemination and fostering reuse of scholarly data.*

*Research Data Journal for the Humanities and Social Sciences is soliciting new submissions, in particular related to the following domains:*

\*

[See More](#)

ISSN: 2452-3666

Publisher: Brill

[View PDF Flyer](#)

[Contents](#) [Submit an Article](#) [About](#)

Advance Articles

Volume: 3 (2018)



Volume: 2 (2017)



Search within Journal...



 Select Issue 

 Index Card

**Subjects:** General, History, General, Asian Studies



# Gdzie udostępniać dane badawcze?

repozytoria danych badawczych

- bezpieczne długoterminowe przechowywanie
- stały adres internetowy, możliwość uzyskania trwałego identyfikatora
- łatwość wyszukiwania
- często też informacje o tym, jak często dane były pobierane i oglądane
- łatwość cytowania



## Jak wybrać repozytorium?

1. repozytoria dziedzinowe
2. repozytorium instytucjonalne
3. repozytoria ogólnego przeznaczenia
4. wyszukiwanie w katalogu repozytoriów danych badawczych



---

# Plan zarządzania danymi badawczymi (DMP)



# DMP

Formalny dokument zawierający zarys tego, co będziemy robić z danymi w trakcie trwania projektu badawczego i po jego zakończeniu.



## Pozyskiwanie i wytwarzanie danych

Skąd pochodzą dane? Jak zostały pozyskane / wytworzone? Czy wykorzystane zostaną już istniejące dane?

Czy zbiory danych będą podlegać zmianom? Czy będą się powiększać?

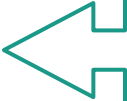
## Forma i formaty danych

Jaką formę mają dane? W jakich formatach zostaną zapisane? Jak nazwane będą pliki? Jak duże będą zbiory danych?



# Przechowywanie danych

Jakie będą zasady krótko- i długoterminowego przechowywania danych?

- 
- Do czego jesteśmy zobowiązani?
  - Jaka jest wartość naukowa lub historyczna danych?
  - Jak unikalne są nasze dane? Czy istnieje możliwość ich ponownego zebrania/wytworzenia?
  - Czy dane mogą wykorzystać inni?
  - Jakie koszty wiążą się z zarządzaniem i przechowywaniem danych?



## Zasady dostępu i korzystania z danych

Które dane zostaną udostępnione w sposób otwarty? Jeśli część danych nie może zostać udostępniona - dlaczego? Czy w takiej sytuacji udostępnione zostaną metadane? Czy dane zostaną udostępnione później (embargo)?

## Prawne i etyczne aspekty zarządzania danymi





## Jak udostępniać dane? Co wziąć pod uwagę?

- Jakie dane jesteśmy zobowiązani udostępnić? Jakie sytuacje zwalniają nas z tego obowiązku?
- Jakie dane chcemy i możemy udostępnić (niezależnie od zobowiązań)?
- Na jakich zasadach udostępniamy dane badawcze?



wymagania grantodawcy

kwestie prawne

kwestie etyczne

kontekst dziedzinowy (zwyczaje)

zasoby (czas, pieniądze, zespół)

...



Build your Data Management Plan

## Welcome to the DMPTool

Create data management plans that meet institutional and funder requirements.

[Get started](#)



### DMPTool by the Numbers



37 635



34 621



249

### Top Templates

Digital Curation Centre  
NSF-BIO: Biological Sciences  
USDA-NIFA: National Institute of Food and Agriculture

F  
Findable



A  
Accessible



I  
Interoperable



R  
Reusable




# Findable

Czy dane opatrzone zostaną metadanymi? Czy będą opisane zgodnie z przyjętymi standardami?

Czy dane będą posiadać trwałe identyfikatory (DOI)?

Czy (meta)dane będą zamieszczone lub indeksowane w serwisie, którego zasoby można przeszukiwać?

 because good research needs good data

Home | Digital curation | About us | News | Events | Resources | Training | Projects | Community | Tailored support

Home > Resources > Metadata Standards

## In this section

- Briefing Papers
- How-to Guides & Checklists
- Developing RDM Services
- Curation Lifecycle Model
- Curation Reference Manual
- Policy and legal
- Data Management Plans
- Tools
- Case studies
- Repository audit and assessment

## Standards

- Disciplinary Metadata**
- DIFFUSE
- Publications and presentations
- Roles
- Curation journals
- Informatics research
- External resources
- Online Store

## Disciplinary Metadata

While data curators, and increasingly researchers, know that good metadata is key for research data access and re-use, figuring out precisely what metadata to capture and how to capture it is a complex task. Fortunately, many academic disciplines have supported initiatives to formalise the metadata specifications the community deems to be required for data re-use. This page provides links to information about these disciplinary metadata standards, including profiles, tools to implement the standards, and use cases of data repositories currently implementing them.

For those disciplines that have not yet settled on a metadata standard, and for those repositories that work with data across disciplines, the General Research Data section links to information about broader metadata standards that have been adapted to suit the needs of research data.

Please note that a community-maintained version of [this directory](#) has been set up under the auspices of the Research Data Alliance.

## Search by Discipline



Biology



Earth Science



General Research Data



Physical Science



Social Science & Humanities

# Accesible

## Publicly accessible

Fully accessible to persons who meet explicitly stated conditions, e.g. ethics approval for sensitive data  
A de-identified / modified subset of the data is publicly accessible  
Embargoed access after a specified date  
Unspecified conditional access e.g. contact the data custodian for access  
Access to metadata only  
No access to data or metadata

Które dane zostaną udostępnione w sposób otwarty? Jeśli część danych nie może zostać udostępniona - dlaczego? Czy w takiej sytuacji udostępnione zostaną metadane?

W jaki sposób i gdzie dane zostaną udostępnione? Czy warunki dostępu będą jasno określone?

Accessible i

How accessible is the data?

Is the data available online without requiring specialised protocols or tools once access has been approved?

Will the metadata record be available even if the data is no longer available?

No access to data or metadata

**Publicly accessible**  
Fully accessible to persons who meet explicitly stated conditions, e.g. ethics approval for sensitive data  
A de-identified / modified subset of the data is publicly accessible  
Embargoed access after a specified date  
Unspecified conditional access e.g. contact the data custodian for access  
Access to metadata only  
No access to data or metadata

Interoperable i

Reusable i

Total across F.A.I.R



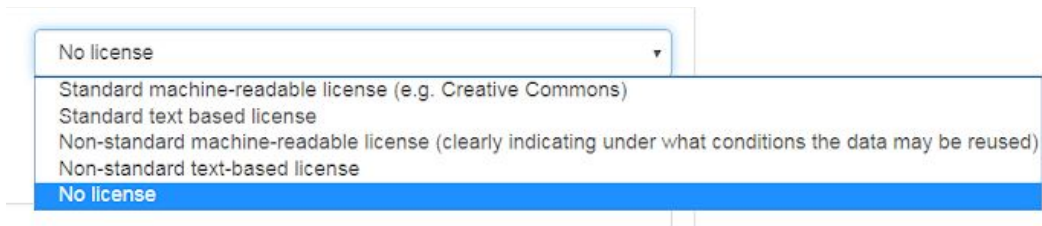
# Interoperable

Czy wykorzystane zostaną  
standardowe słowniki metadanych,  
standardowe metodologie?

Czy możliwe będzie połączenie  
danych z innymi zbiorami  
pochodzącymi z innych źródeł?

# Reusable

Czy dane zostaną opatrzone licencją, która pozwoli na ich ponowne wykorzystanie w stopniu tak szerokim jak to możliwe?



A screenshot of a dropdown menu for selecting a license. The menu is open, showing several options. The 'No license' option is highlighted in blue. The options are:

- No license
- Standard machine-readable license (e.g. Creative Commons)
- Standard text based license
- Non-standard machine-readable license (clearly indicating under what conditions the data may be reused)
- Non-standard text-based license
- No license

Licencje Creative Commons - trzy warstwy:

- tekst prawny,
- przystępne podsumowanie,
- dane (kod) do odczytu maszynowego.



# Reusable

Czy do danych dołączona zostanie dokumentacja? Czy wskazane zostanie źródło (pochodzenie) danych: kto i w jaki sposób je wytworzył? Jak były przetwarzane? Czy zawierają dane z innych źródeł?





## ***Unusable data***

„otwarty zbiór danych”

„dane publicznie dostępne”

„dane ogólnodostępne”

„brak ograniczeń”

„freely available”

„dane dostępne na licencji Creative Commons”





## FAIR self-assessment tool

Welcome to the ARDC FAIR Data self-assessment tool. Using this tool you will be able to assess the 'FAIRness' of a dataset and determine how to enhance its FAIRness (where applicable).

This self-assessment tool has been designed predominantly for data librarians and IT staff, but could be used by software engineers developing FAIR Data tools and services, and researchers provided they have assistance from research support staff.

You will be asked questions related to the principles underpinning Findable, Accessible, Interoperable and Reusable. Once you have answered all the questions in each section you will be given a 'green bar' indicator based on your answers in that section, and when all sections are completed, an overall 'FAIRness' indicator is provided.

Please be aware that additional explanatory information is provided within the tool. The (i) information button provides an overview of each of the FAIR high-level elements (Findable, Accessible, Interoperable and Reusable). Additionally, each question is hyperlinked, leading users to explanatory information and links to wider resources on related topics.

<b>Findable</b>	
<b>Does the dataset have any identifiers assigned?</b>	<input type="text" value="No identifier"/>
<b>Is the dataset identifier included in all metadata records/files describing the data?</b>	<input type="text" value="No"/>
<b>How is the data described with metadata?</b>	<input type="text" value="The data is not described"/>
<b>What type of repository or registry is the metadata record in?</b>	<input type="text" value="The data is not described in any repository"/>
<input type="text"/>	
<b>Accessible</b>	
<b>Interoperable</b>	
<b>Reusable</b>	



# Co poza danymi?

## metadane

podstawowe informacje stanowiące opis całego zbioru danych (autor, tytuł zbioru, data utworzenia, słowa kluczowe, źródło finansowania, licencja itp.)

## dokumentacja

informacje o charakterze metodologicznym, kontekst powstania, słowniki itp.



# Dokumentacja

## poziom projektu

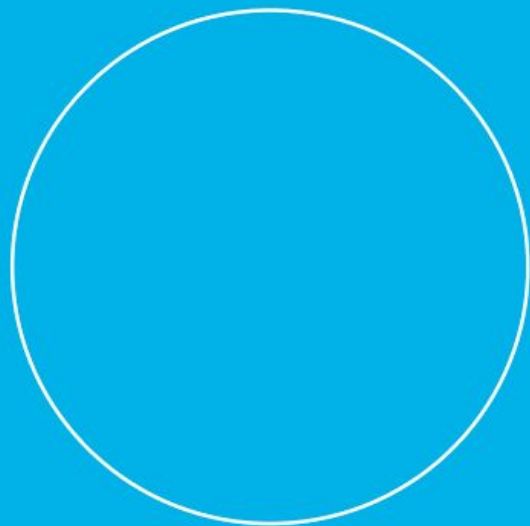
- opis celu i kontekstu badań
- metodologia, sposób pozyskiwania danych
- struktura danych, relacje między plikami
- linki do publikacji powiązanych z danymi
- zastosowane metody kontroli danych
- informacje o danych, których nie można udostępnić

## poziom danych

- nazwy i opisy zmiennych, skrótów, schematów klasyfikacyjnych,
- informacje o urządzeniach pomiarowych, ich ustawieniach, parametrach itp.
- oznaczenia brakujących wartości, dodatkowe informacje
- testy jakości danych

# Selekcja i przygotowanie danych badawczych do udostępnienia

Wersja 1.0



Dziedzinowe Repozytoria  
Otwartych Danych Badawczych  
[drodb.icm.edu.pl](http://drodb.icm.edu.pl)

**Dziękuję za uwagę.**

**Kontakt:**

**Natalia Gruenpeter**

**[n.gruenpeter@icm.edu.pl](mailto:n.gruenpeter@icm.edu.pl)**

