



RepOD

Repozytorium Otwartych Danych

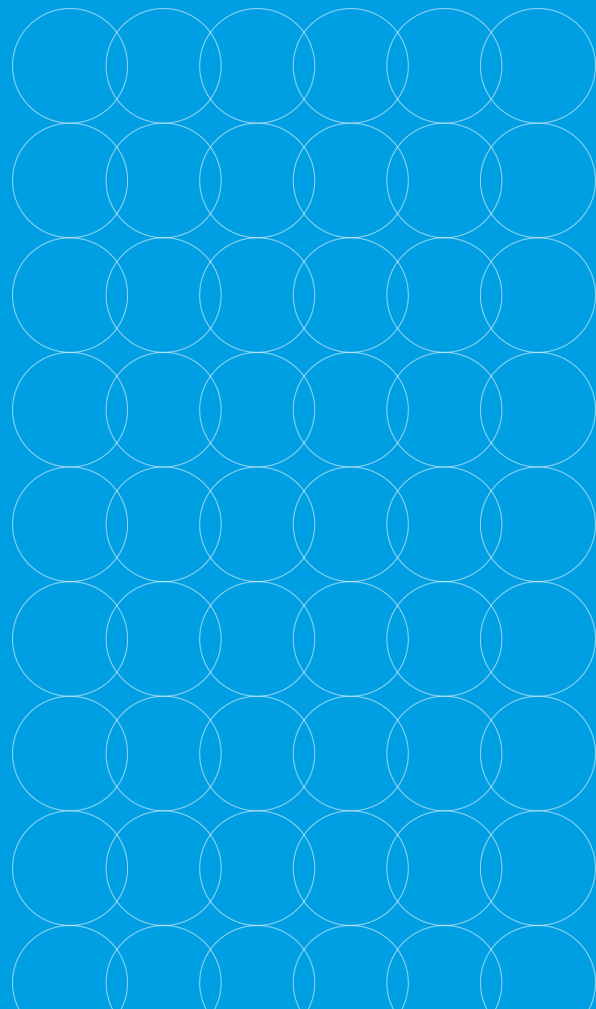
Dziedzinowe Repozytoria Otwartych Danych Badawczych

PREZENTACJA PROJEKTU

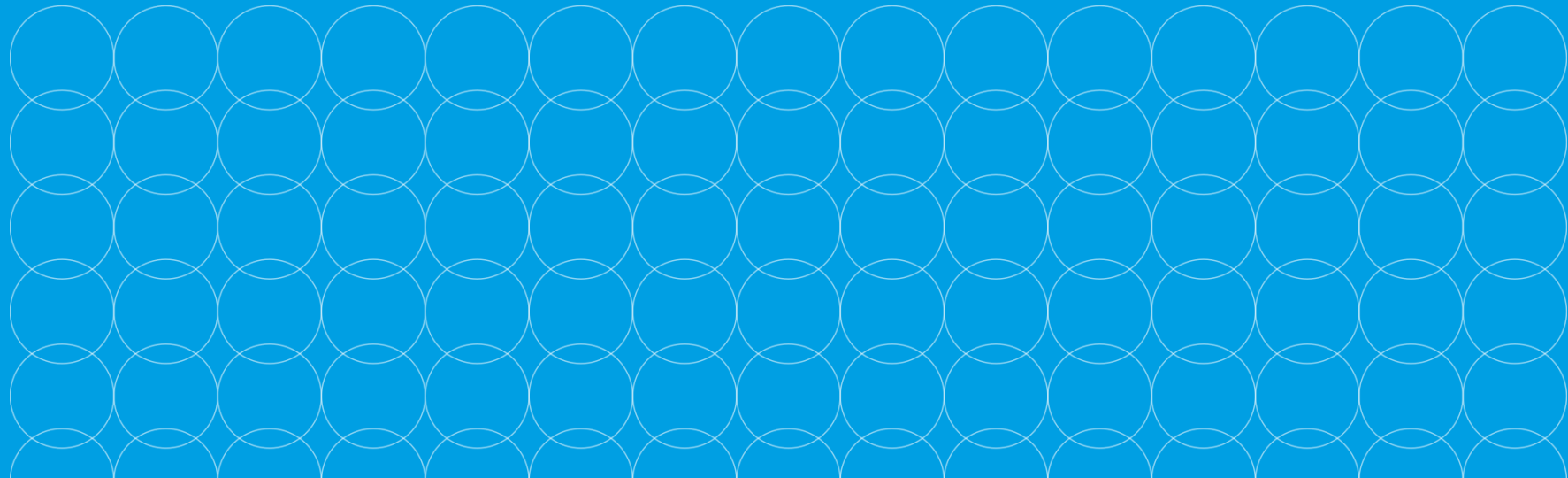


Rzeczpospolita
Polska

Unia Europejska
Europejski Fundusz
Rozwoju Regionalnego



Informacje o projekcie



Ogólne informacje

Beneficjent projektu: Uniwersytet Warszawski,
jednostki realizujące: ICM UW, ISS UW

Partnerzy projektu:
Instytut Filozofii i Socjologii PAN,
Uniwersytet Adama Mickiewicza w Poznaniu

Okres realizacji: 1 sierpnia 2018 r. - 31 lipca 2021 r.

Kwota dofinansowania: 4 998 889 PLN
(w tym UE 4 230 559,76 PLN,
budżet państwa 768 329,24 PLN).



RepOD

Repozytorium Otwartych Danych

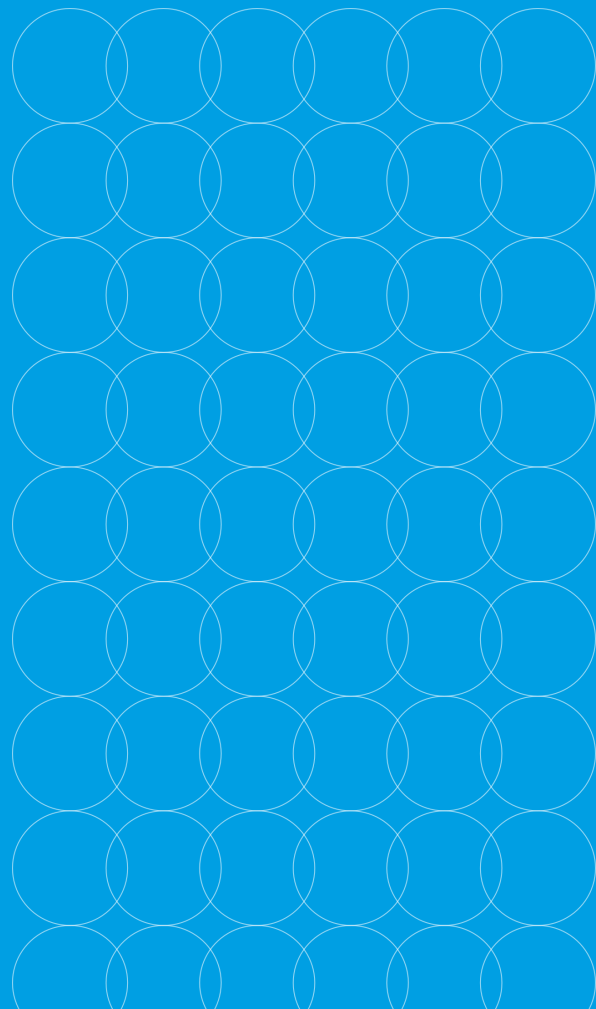
RepOD

- Repozytorium otwarte dla wszystkich zainteresowanych użytkowników: repod.icm.edu.pl
- Repozytorium ogólnego przeznaczenia
- Zastąpiło działającą pilotażową wersję repozytorium



Rzeczpospolita
Polska

Unia Europejska
Europejski Fundusz
Rozwoju Regionalnego





RDS

Repozytorium
Danych Społecznych

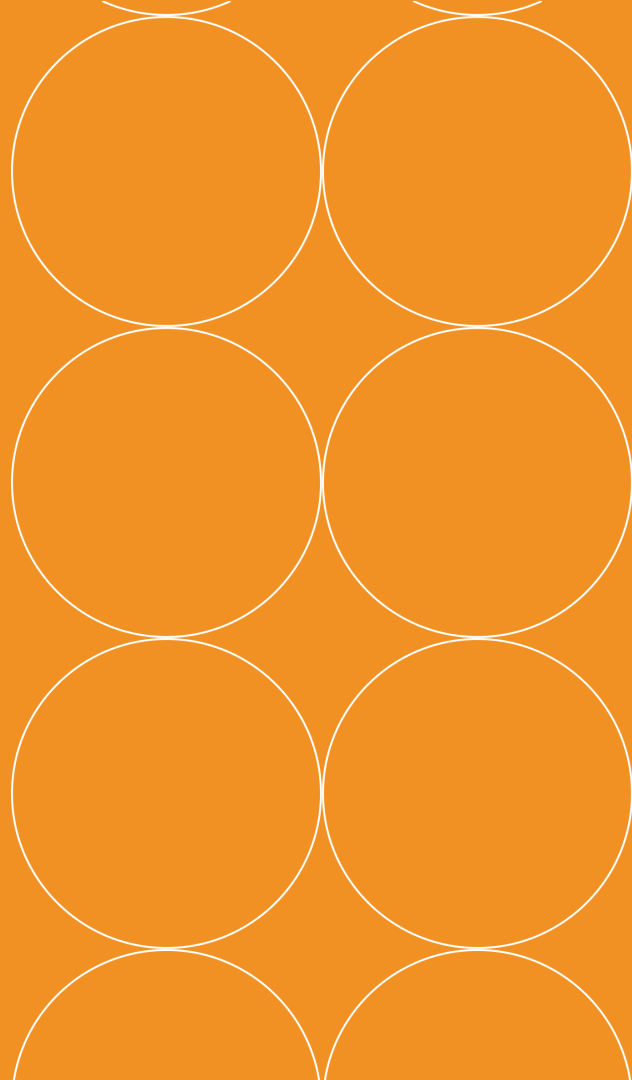
RDS: Repozytorium Danych Społecznych

- Repozytorium otwarte dla wszystkich zainteresowanych użytkowników: rds.icm.edu.pl
- 400 zbiorów danych (jakościowych i ilościowych) udostępnionych w ramach projektu



Rzeczpospolita
Polska

Unia Europejska
Europejski Fundusz
Rozwoju Regionalnego





MX-RDR

Macromolecular Xtallography
Raw Data Repository

MX-RDR: Repozytorium Danych Krystalograficznych

- Repozytorium otwarte dla wszystkich zainteresowanych użytkowników: mxrdr.icm.edu.pl
- 200 zbiorów surowych danych krystalograficznych udostępnionych w ramach projektu



UNIWERSYTET
IM. ADAMA MICKIEWICZA
W POZNANIU



Rzeczpospolita
Polska

Unia Europejska
Europejski Fundusz
Rozwoju Regionalnego





RepOD

Repozytorium Otwartych Danych

Jak korzystać z Repozytorium Otwartych Danych RepOD?

PRZEWODNIK

CC-BY

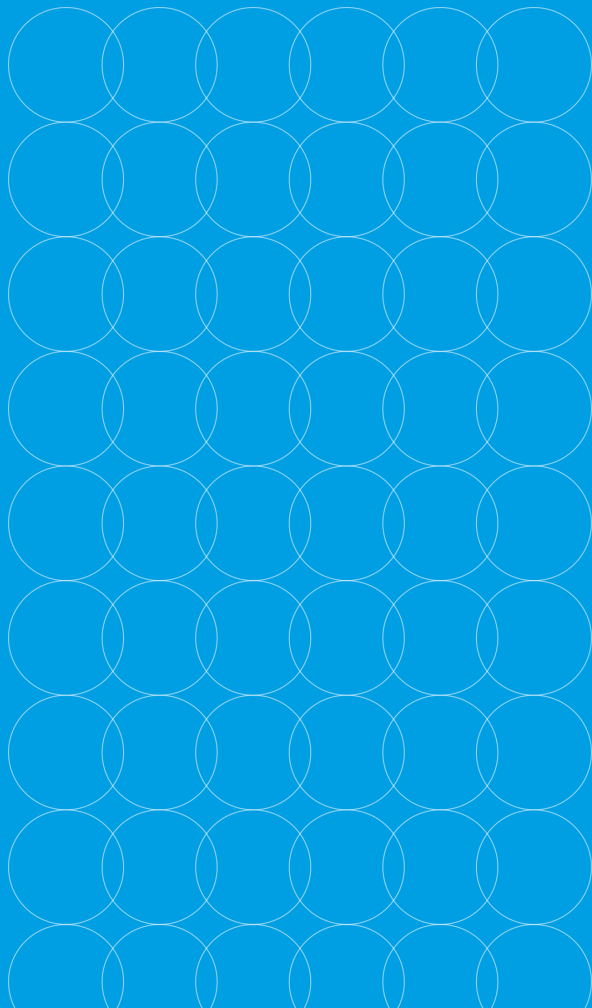
Treść licencji dostępna na stronie:

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

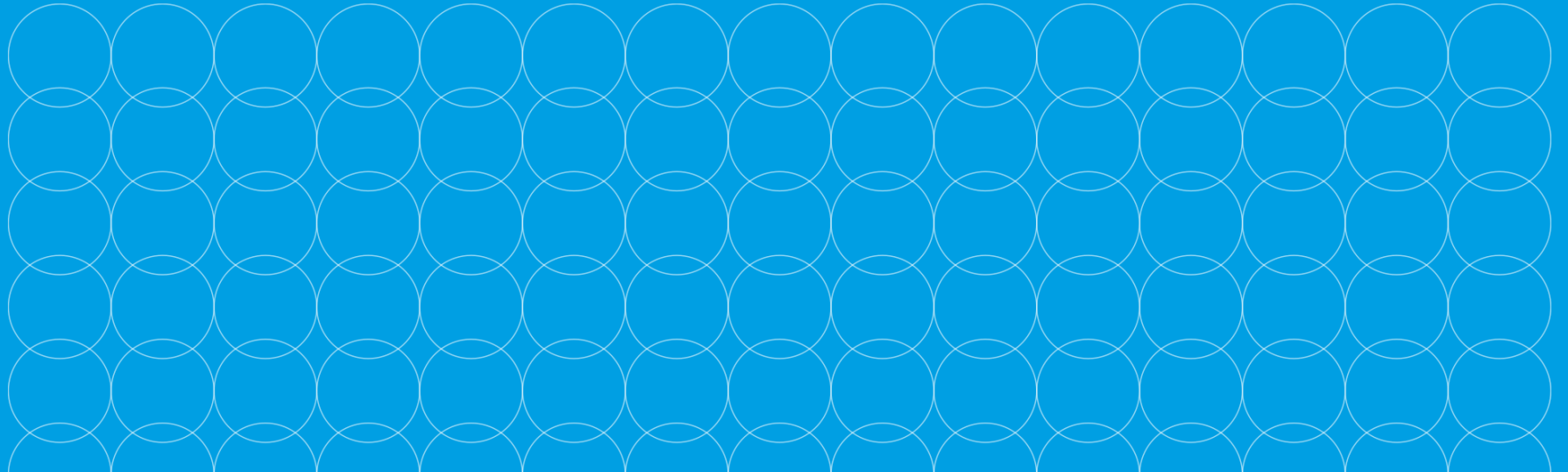


Rzeczpospolita
Polska

Unia Europejska
Europejski Fundusz
Rozwoju Regionalnego



○ repozytorium



Podstawowe informacje

Repozytorium Otwartych Danych RepOD umożliwia:

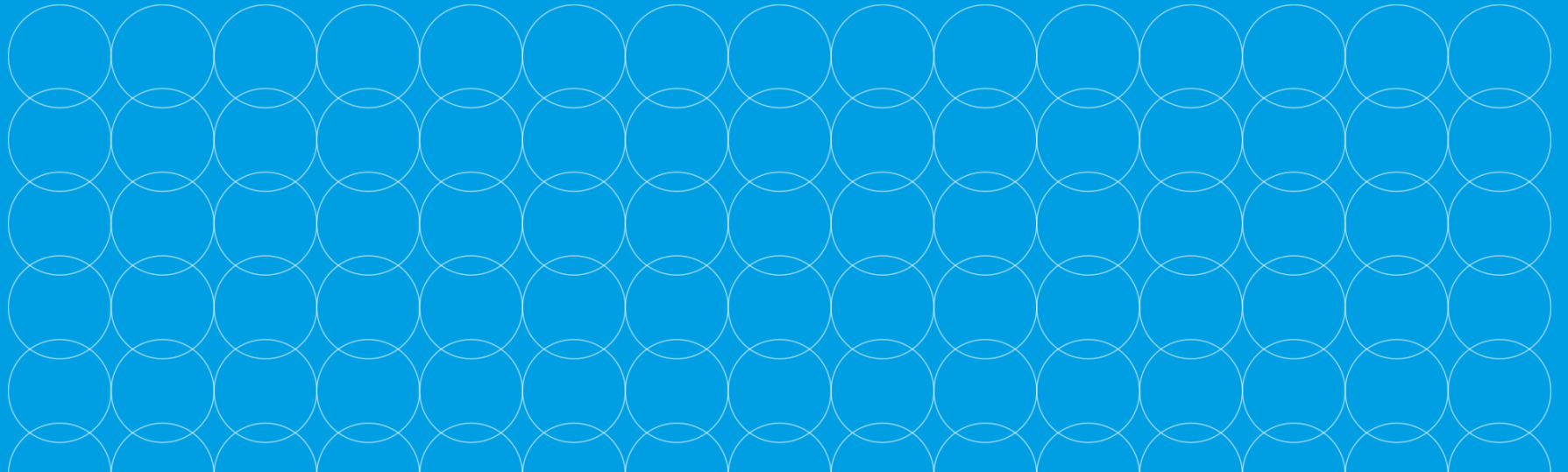
- przechowywanie danych badawczych;
- publiczne udostępnianie danych badawczych;
- przeglądanie danych badawczych;
- pobieranie danych badawczych.

Zdeponowane dane otrzymują numer DOI.

Korzystanie z repozytorium jest darmowe i otwarte dla wszystkich zainteresowanych.

Niektóre funkcje (np. przechowywanie i publiczne udostępnianie danych) wymagają rejestracji użytkownika.

Deponowanie danych badawczych w Repozytorium Otwartych Danych RepOD



Deponowanie danych

[film]

Metadane: wymagane

Wymagane metadane w głównej kolekcji RepOD:

- tytuł,
- autor/autorzy,
- osoba do kontaktu,
- opis,
- obszar tematyczny (wybór z listy).

Metadane → licencja CC0 (Przekazanie do domeny publicznej)

<https://creativecommons.org/publicdomain/zero/1.0/deed.pl>

Metadane: opcjonalne

Metadane opcjonalne → dlaczego warto je dodać?

- niektóre metadane mogą być widoczne dopiero po zapisaniu wersji roboczej, w trakcie edycji,
- bogate metadane ułatwiają wyszukiwanie i identyfikację poszukiwanych informacji,
- zgodność z zasadami FAIR (F - Findable),
- możliwość automatycznego powiązania zbioru danych z innymi obiektami cyfrowymi (publikacjami, zbiorami danych)

Metadane: opcjonalne

Słowa kluczowe

- możliwość wskazania kontrolowanego słownika słów kluczowych,
- w repozytorium RDS wybrane słowniki kontrolowane zaimplementowane jako osobne pola

Metadane: opcjonalne

Powiązana publikacja

- różne rodzaje powiązań,
- wskazanie trwałego identyfikatora powiązanej publikacji,
- umożliwia włączenie zbioru w inne ekosystemy (np. OpenAIRE).

Rodzaj powiązania ? Typ identyfikatora ?

Wybierz ▼ Wybierz ▼

Zdeponowany zbiór danych jest suplementem powiązanej publikacji

Powiązana publikacja jest suplementem zdeponowanego zbioru danych

Zdeponowany zbiór danych jest częścią powiązanej publikacji

Powiązana publikacja odnosi się do zdeponowanego zbioru danych bez jego bezpośredniego cytowania

Powiązana publikacja stanowi dokumentację zdeponowanego zbioru danych

Zdeponowany zbiór danych stanowi dokumentację powiązanej publikacji

Powiązana publikacja stanowi recenzję zdeponowanego zbioru danych

Zdeponowany zbiór danych jest materiałem źródłowym dla powiązanej publikacji

Metadane: opcjonalne

Informacje o grantcie

Institucja finansująca ?	Skrócona nazwa instytucji ?	+
European Commission	EC	
Program grantowy ?	Numer grantu ?	
H2020	123456789	

Obowiązek wypełnienia tego pola → może wynikać ze zobowiązań wobec instytucji finansującej badania i jej polityki otwartości, np. NCN - dane powiązane z publikacjami powinny być udostępniane z adnotacją o źródle finansowania.

https://www.ncn.gov.pl/sites/default/files/pliki/zarzadzenia-dyrektora/zarzadzenieDyr-38_2020.pdf

Zdeponowanie a publikacja

Zdeponowanie danych \neq publikacja danych

Numer DOI zostaje wstępnie przypisany do zestawu podczas zdeponowania, jeszcze przed publikacją.

Dzięki temu możliwe jest powiązanie danych z publikacją (poprzez wskazanie w publikacji odpowiedniego DOI) bez konieczności udostępniania danych przed ukazaniem się publikacji.

Formaty plików

- Zalecamy otwarte formaty plików.
- Limit przesyłania plików wynosi 5.0 GB na plik.
- Nie ma limitu wielkości zbioru.

Wersje

- Jeśli zauważymy błąd w opublikowanych danych lub metadanych, możemy go poprawić i opublikować nową wersję zbioru.
- Zmiana w metadanych może zostać opublikowana jako “mała” (1.1; 1.2; 2.1 etc.) lub “duża” (2.0; 3.0 etc) wersja zbioru.
- Zmiana w danych (plikach) zawsze musi być opublikowana jako “duża” wersja.
- Poprzednie wersje zbioru pozostają dostępne – być może ktoś już użył tych danych lub je zacytował.

Wybór licencji

- Metadane i zestawienie zbioru - na mocy regulaminu zawsze na licencji CC0.
- Możliwość wybrania warunków użycia dla poszczególnych plików w zbiorze.
 - Licencja (do wyboru z listy, w tym popularne licencje Creative Commons).
 - Dozwolony użytek.
- Szczegółowa lista licencji i warunków użycia zależy od repozytorium.

F
indable



A
ccessible



I
nteroperable



R
eusable



Ilustracja: Sungya Pundir, Wikimedia Commons,
CC BY-SA 4.0

Zasady FAIR

- Polityka Komisji Europejskiej, Horyzont 2020 (zasady FAIR uwzględnione w DMP)
- wytyczne Narodowego Centrum Nauki do sporządzania DMP
- Dyrektywa ws. otwartych danych i ponownego wykorzystywania informacji sektora publicznego



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?

The screenshot shows the DCC website with a red header containing the logo and the tagline 'because good research needs good data'. A search bar is located in the top right. Below the header is a navigation menu with links: Home, Digital curation, About us, News, Events, Resources, Training, Projects, Community, and Tailored support. The main content area is titled 'Home > Resources > Metadata Standards'. On the left, there is a sidebar with a list of links under the heading 'In this section', including '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', and 'Standards'. The 'Standards' section is expanded to show 'Disciplinary Metadata' and 'DIFFUSE'. Below this is a 'Search by Discipline' section with six categories, each with a representative image and a label: 'Biology' (microscope), 'Earth Science' (coastline), 'General Research Data' (circuit board), 'Physical Science' (molecular model), and 'Social Science & Humanities' (hands raised). The 'Disciplinary Metadata' section contains text explaining the importance of metadata for research data access and re-use, and provides links to information about standards and broader metadata standards.

<http://www.dcc.ac.uk/resources/metadata-standards>

Accessible

Które dane zostaną udostępnione w sposób otwarty?

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

Czy warunki dostępu będą jasno określone?

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

The screenshot shows a web interface for a self-assessment tool. The 'Accessible' section is active, with a dropdown menu open showing the following options: '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', and 'No access to data or metadata'. Below the dropdown, there are three questions: 'How accessible is the data?', 'Is the data available online without requiring specialised protocols or tools once access has been approved?', and 'Will the metadata record be available even if the data is no longer available?'. Each question has a corresponding input field. Below these questions are sections for 'Interoperable' and 'Reusable', each with an information icon. At the bottom, there is a 'Total across F.A.I.R.' section with an input field.

<https://ardc.edu.au/resources/working-with-data/fair-data/fair-self-assessment-tool/>

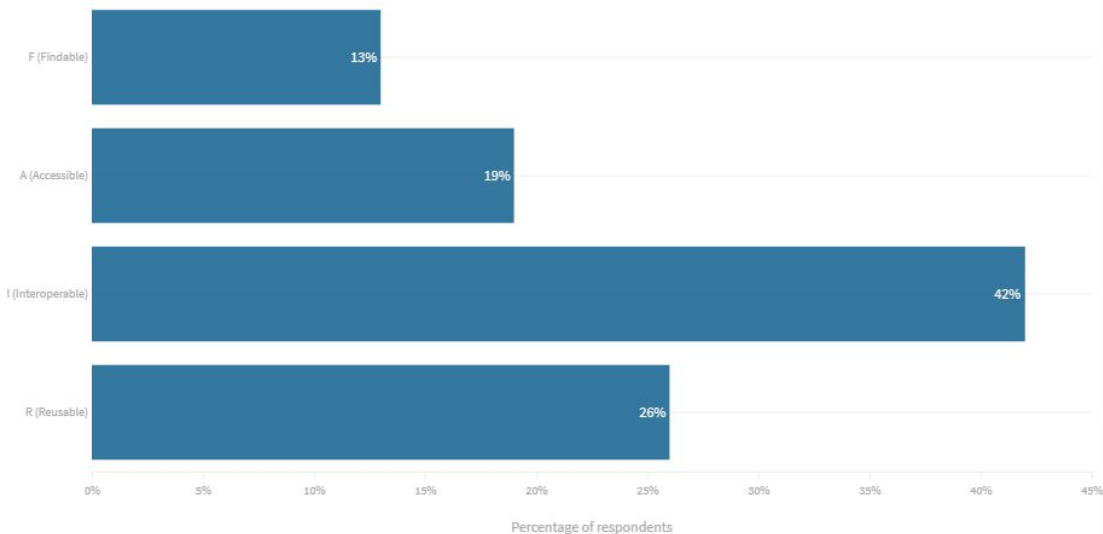
Interoperable

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

Czy w opisie danych zastosowane zostaną standardowe rozwiązania, np. słowniki kontrolowane?

Which of the FAIR principles do you think most needs better definition?

Interoperability is the least understood FAIR principle. Some 42% of the 187 respondents who answered this question felt that it needed further clarification.



Source: [State of Open Data](#)

Made with Flourish

<https://www.natureindex.com/news-blog/what-scientists-need-to-know-about-fair-data>

Reusable

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

Czy sposób pozyskiwania danych oraz kontekst, w jakim zostały wytworzone zostały dostatecznie dobrze opisane?

Licencje Creative Commons - trzy warstwy:

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

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

Reusable



Which of the following best describes the license/usage rights attached to the data?

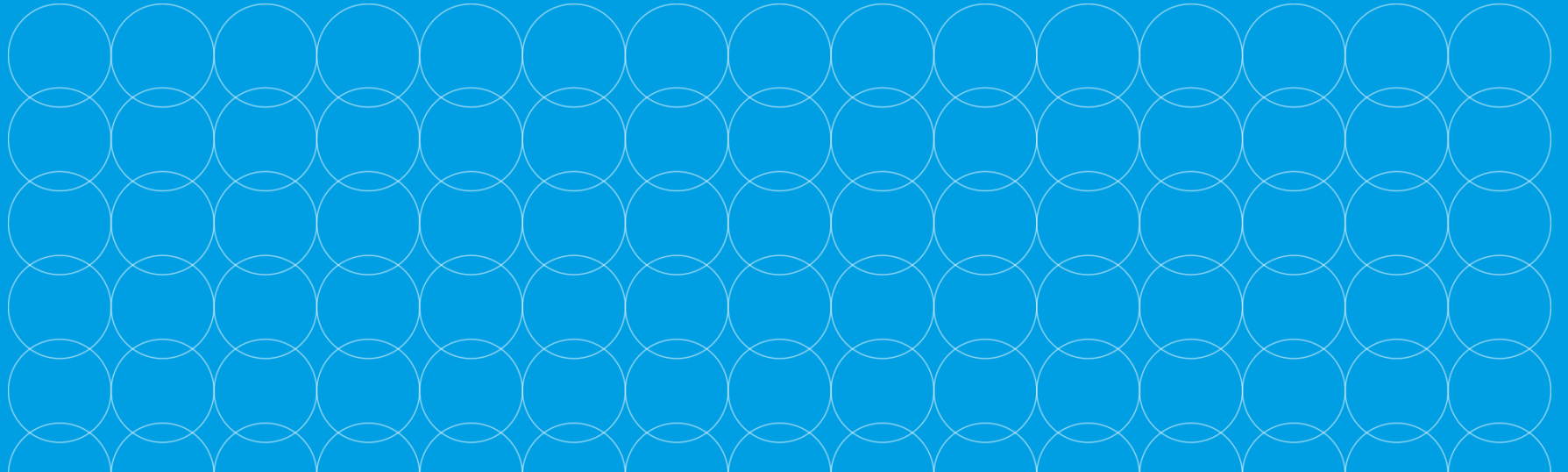
No license

How much provenance information has been captured to facilitate data reuse?

No provenance information is recorded

<https://ardc.edu.au/resources/working-with-data/fair-data/fair-self-assessment-tool/>

Korzystanie z danych badawczych w Repozytorium Otwartych Danych RepOD



Wyszukiwanie

- Wyszukiwanie proste.
- Wyszukiwanie zaawansowane - po wartościach wybranych pól metadanych.
- Wyszukiwanie przez API.
- W najbliższym czasie: możliwość odnalezienia danych z RepOD w innych serwisach harwerstujących metadane (np. OpenAIRE).

Pobieranie

- Dla plików dostępnych publicznie istnieje możliwość pobrania plików przez interfejs graficzny (“z przeglądarki”) oraz API.
- Jeśli zbiór zawiera wiele plików, można je pobrać za jednym zamachem spakowane do formatu ZIP.
- W przypadku plików tabelarycznych istnieje możliwość pobrania plików w kilku tworzonych automatycznie formatach.

Cytowanie

- Sugerowane cytowanie znajduje się na stronie zbioru danych.
- Możliwość pobrania danych bibliograficznych zbioru w formatach EndNote XML, RIS i BibTEX (umożliwia import np. do menedżera cytowań).
- Szczególnie ważne: numer DOI oraz wersja zbioru.

Strony internetowe

[Regulamin korzystania z repozytorium RepOD](#)

[Witryna informacyjna repozytorium RepOD](#)

[Strona projektu Dziedziny Repozytoria Otwartych Danych Badawczych](#)

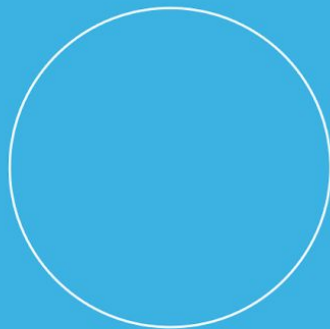
Broszury informacyjne w zakładce [Pliki do pobrania](#):

- Selekcja i przygotowanie danych badawczych do udostępniania | [pdf](#)
- Jak korzystać z zasobów w repozytoriach danych | [pdf](#)
- Prawne aspekty otwierania danych badawczych – poradnik | [pdf](#)

Broszury informacyjne

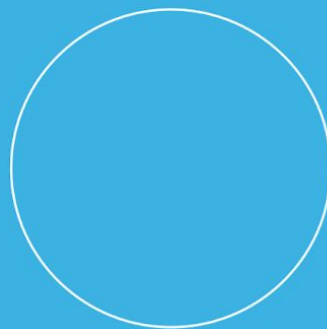
Prawne aspekty otwierania danych badawczych – poradnik

Wersja 1.0



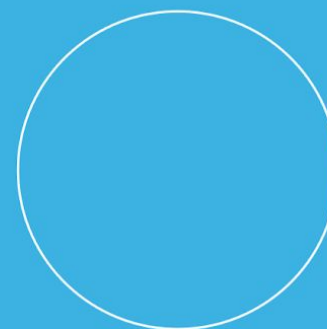
Selekcja i przygotowanie danych badawczych do udostępnienia

Wersja 1.0



Jak korzystać z zasobów w repozytoriach danych

Wersja 1.0



<https://drodb.icm.edu.pl/materialy-2/>



Rzeczpospolita
Polska

Unia Europejska
Europejski Fundusz
Rozwoju Regionalnego



Dziękujemy!

Kontakt: drodb@icm.edu.pl



Fundusze Europejskie
Polska Cyfrowa



Rzeczpospolita
Polska

Unia Europejska
Europejski Fundusz
Rozwoju Regionalnego

