

A decorative graphic in the top right corner featuring a grid of thin red lines and several red circles of varying sizes scattered across it.

Otwarte dane badawcze w polityce i praktyce Narodowego Centrum Nauki

A decorative graphic on the right side consisting of two horizontal red lines and a single red circle positioned between them.

Natalia Galica
Narodowe Centrum Nauki



Rynek danych w Europie

Projected figures 2025



530%

increase of global
data volume

From 33 zettabytes
in 2018 to 175
zettabytes



**€829
billion**

value of data
economy in the
EU27

From €301 billion
(2.4% of EU GDP)
in 2018



**10.9
million**

data
professionals in
the EU27

From 5.7 million in
2018



65%

Percentage of EU
population with
basic digital skills

From 57% in 2018

Poziom europejski

2019: Directive no 2019/1024 of the EP and of the Council (EU) of 20 June 2019 on open data and the re-use of public sector information;

2021: Memorandum of Understanding for the Co-programmed European Partnership for the European Open Science Cloud (EOSC)

2022: Strategic Research and Innovation Agenda (SRIA) of EOSC;

2022: European Strategy for Data, Data Act, Data Governance Act, Digital Services Act, Digital Markets Act...

Poziom krajowe (wybrane przykłady):

National OS Plan 2021-2027 (IT),
National OS Plan 2021-2024 (FR)

Nelson Memorandum 2022 (USA)



European Strategy for Data, 2022

Europejska strategia w zakresie danych ma na celu stworzenie jednolitego rynku danych, który zapewni Europie:

- wiodącą pozycję w globalnej gospodarce opartej na danych,
- globalną konkurencyjność,
- suwerenność danych.

Wspólne europejskie przestrzenie danych zapewnią dostęp do większej ilości danych do wykorzystania w gospodarce i społeczeństwie, przy jednoczesnej kontroli nad podmiotami, które generują dane.

eosc

Position of EOSC according to the European Commission

Taken from EC slides



“EOSC is the basis for a science, research and innovation data space that will bring together data resulting for research and deployment programmes and will be connected and articulated with the sectoral data spaces”

(European Data Strategy, COM(2020) 66 final)



Dane badawcze: definicja

(...) dane badawcze udostępnia się do ponownego wykorzystywania w celach komercyjnych lub niekomercyjnych zgodnie z rozdziałami III i IV, w zakresie, w jakim dane te są finansowane ze środków publicznych i zostały już publicznie udostępnione przez naukowców, organizacje prowadzące badania naukowe lub organizacje finansujące badania naukowe za pośrednictwem repozytorium instytucjonalnego lub tematycznego. W tym kontekście uwzględnia się uzasadnione interesy handlowe, działania związane z transferem wiedzy oraz istniejące już prawa własności intelektualnej

[DYREKTYWA PARLAMENTU EUROPEJSKIEGO I RADY (UE) 2019/1024 z dnia 20 czerwca 2019 r. w sprawie otwartych danych i ponownego wykorzystywania 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.

[USTAWA z dnia 11 sierpnia 2021 r. o otwartych danych i ponownym wykorzystywaniu informacji sektora publicznego]

Dlaczego ważne jest zapewnienie otwartego dostępu do danych badawczych?

Kwestie naukowe

- cytowalność danych
- widzialność krajowa/międzynarodowa w oparciu o dane poprzez repozytoria danych
- współpraca krajowa/międzynarodowa w oparciu o dane
- możliwość (ponownego) wykorzystania danych
- unikanie wcześniej popełnionych błędów badawczych (pokazywanie błędów a nie tylko sukcesów)
- transparentność procesu badawczego prowadząca do zwiększenia rzetelności naukowej, wiarygodność i zaufania wobec nauki
- replikowalność badań
- odpowiedzialność naukowa wobec społeczeństwa

Kwestie instytucjonalne i polityki narodowe

- efektywność wydawania środków publicznych na badania (unikanie finansowania tych samych badań i tych samych błędów badawczych; efektywne wykorzystanie już istniejących danych)
- premiowanie publikacji danych (polityka instytucjonalna, krajowa i UE)
- wymagania agencji finansujących badania
- nowy typ konkursów KE (prognoza)
- publikowanie wyników badań w czasopismach naukowych z obowiązkiem prezentowania danych (polityka wydawnictw)
- ewaluacja naukowa/ocena naukowa

Udostępnianie danych badawczych: polityka i praktyka NCN



Otwarta nauka w NCN

OTWARTE PUBLIKACJE

- 2018: NCN sygnatariuszem Planu S i członkiem cOAlition S
- 2018: NCN sygnatariuszem DORA
- 2020: Polityka NCN dot. otwartego dostępu do publikacji naukowych (obowiązuje dla wszystkich umów od 1.1.21)

OTWARTE DANE

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

Otwarty dostęp / Publikacje i dane powiązane

POLITYKA

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

UMOWA, par. 5

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



Joint Declaration of Data Citation Principles (FORCE11, 2014)

1. Znaczenie

Dane należy uznać za dające się cytować produkty badań. Cytowanie danych powinno mieć taką samą wagę w dorobku naukowym, jak cytowanie innych obiektów badawczych, takich jak publikacje.

2. Uznanie autorstwa

Cytowanie danych powinno ułatwiać uznanie autorstwa oraz normatywnego i prawnego przypisania danych wszystkim osobom, które przyczyniły się do ich powstania.

3. Dowody

W literaturze naukowej zawsze i wszędzie tam, gdzie twierdzenie opiera się na danych, należy przytoczyć odpowiednie dane.

4. Niepowtarzalna identyfikacja

Cytowanie danych powinno zawierać trwałą metodę identyfikacji, którą można zastosować maszynowo, unikalną na skalę światową i szeroko stosowaną przez społeczność / PID.

5. Dostęp

Cytowanie danych powinno ułatwiać dostęp do samych danych oraz do powiązanych metadanych, dokumentacji, kodu i innych materiałów, które są niezbędne zarówno ludziom, jak i maszynom do świadomego wykorzystania danych, do których się odnoszą.

6. Trwałość

Unikalne identyfikatory i metadane opisujące dane powinny być trwałe — nawet po upływie czasu życia danych.

7. Specyfika i weryfikowalność

Cytowanie danych powinno ułatwiać identyfikację, dostęp i weryfikację konkretnych danych.

8. Interoperacyjność i elastyczność

Metody cytowania danych powinny być wystarczająco elastyczne, aby uwzględnić różne praktyki stosowane w społecznościach naukowych, ale nie powinny różnić się na tyle, aby zagrażać interoperacyjności praktyk.

TOP Guidelines: Transparency and Openness Promotion Guidelines (Centre for Open Science, 2015)

Osiem standardów transparentności:

- Standardy cytowania
Cytowanie artykułu jest rutynowe i dobrze sformułowane. Podobne standardy można zastosować do cytowania danych, kodu i innych wyników badań, **aby rozpoznać i uznać oryginalny wkład intelektualny.**
- Transparentność/Przejrzystość danych
- Przejrzystość metod analitycznych (kodów)
- Przejrzystość materiałów badawczych
- Przejrzystość projektu i analizy
- Wstępna rejestracja badań
- Wstępna rejestracja planów analiz
- Replikacja

Trzy poziomy zaawansowania.



TOP Factor: „TOP Factor is a rating, not a ranking”:

The screenshot shows the journal page for 'Religion' on the TOP Factor website. The page includes a journal description, a table of guidelines, and a sidebar with disciplines and ISSN numbers.

JOURNAL DESCRIPTION
"Religion is an internationally recognized, Thomson Reuters-listed journal, publishing original scholarly research in the comparative and interdisciplinary study of religion. Religion is committed to the publication of significant, novel research, as well as review symposia and responses. In addition, the journal includes book reviews and discussions of important venues for the publication of scholarly work in the study of religion. Religion has European and North-American editors, a multi-national Editorial Board, and is committed to publishing work from scholars of religion around the globe, including occasional translations of important papers."

GUIDELINE	LEVEL	SUMMARY	JUSTIFICATION
Total	2		
Data Citation	1	Journal describes citation of data in guidelines to authors with clear rules and examples.	Provides example of data citation.
Data Transparency	1	Article states whether data are available, and, if so, where to access them.	" if there is a data set associated with the paper, please provide information about where the data supporting the results or analyses presented in the paper can be found."
Analysis Code Transparency	0	Journal encourages code sharing, or says nothing.	No mention.
Materials Transparency	0	Journal encourages materials sharing, or says nothing.	No mention.

DISCIPLINES
Sociology and Political Science
Religious Studies

ISSN NUMBERS
0048-721X

ISSN NUMBERS
1096-1151

LAST UPDATED
2022-11-02

[Journal Page](#)

[Search Journals Now](#)

The screenshot shows the 'All Journals' page on the TOP Factor website. It features a search bar and a list of journals with their respective TOP Factor ratings.

JOURNAL	TOTAL TOP FACTOR
1 Meta-Psychology	27
2 Advances in Methods and Practices in Psychological Science	25
3 Peer Community In Registered Reports	25
4 Personality Science	24
5 Comprehensive Results in Social Psychology	23
6 Cortex	23
7 European Journal of Personality	23
8 Collabra	20
9 Journal of Personality	20
10 Journal of Personality and Social Psychology	20

Rows per page: 10 | 1-10 of 1575





FAIR Data Standards

W polityce NCN:

Open data = FAIR data

As open as possible, as closed as necessary.

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

Możliwość ograniczenia lub wykluczenia otwartego dostępu: eksploatacja przemysłowa, ograniczenia wynikające z klauzuli poufność, 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...

Zasady FAIR

Findable: odnajdywalne

Metadane oraz unikalny elektroniczny identyfikator (ang.PID); dane oraz metadane posiadają PID (np. DOI).

Accessible: dostępne

dostęp do (meta)danych z wykorzystaniem standardowych protokołów komunikacyjnych (otwarte, bezpłatne i powszechnie stosowane), metadane są przechowywane dłużej niż dane (jest to tańsze i łatwiejsze).

Interoperable: interoperacyjne

(meta)dane doczytywane i przetwarzane przez ludzi i maszyny bez konieczności wykorzystywania tłumaczeń, specjalistycznych algorytmów; format (zapewniający odczyt maszynowy i przez ludzi), powszechnie stosowane słowniki dostosowanie do standardów obowiązujących w danej dziedzinie, odnośniki do powiązanych zbiorów.

Reusable: możliwe do ponownego wykorzystania

prawna interoperacyjność; licencję określającą jednoznacznie warunki ponownego wykorzystania i przetwarzania danych np. CC BY, CC 0; odczytywane przez ludzi i maszyny, precyzyjne informacje dotyczące cytowania;

FAIR DATA PRINCIPLES

AH!



FINDABLE



ACCESSIBLE

HOW DO YOU OPEN A .XZQ FILE?



INTEROPERABLE

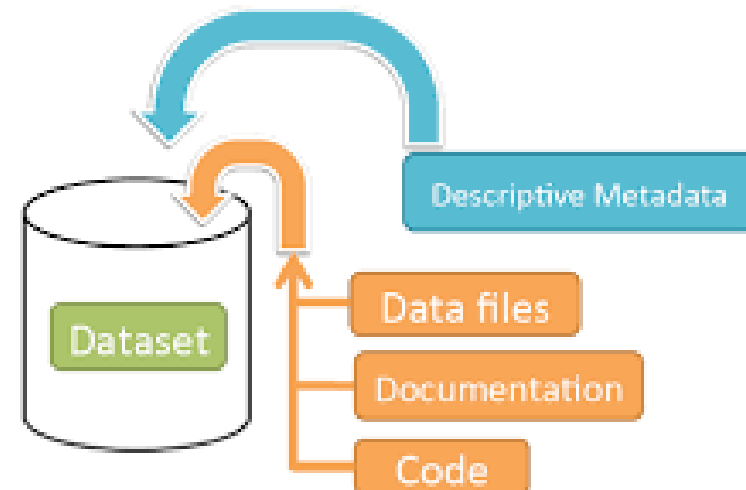


REUSABLE

Dataset

1. 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.
2. Zasady FAIR
3. Dokumentacja
4. Metadane
5. Trwały unikalny elektroniczny identyfikator (PID)

Schematic Diagram of a Dataset in Dataverse 4.0



Container for your data, documentation, and code.

<https://repod.icm.edu.pl/>



Dokumentacja i jakość danych

1. Opis kontekstu badania.
2. Opis metodologii: metody pozyskiwania danych, metody analizy danych, użyte oprogramowanie i sprzęt.
3. Struktura i relacje między folderami: chronologicznie i tematycznie.
4. Sposób kontroli jakości danych.
5. Informacja o otwartości danych.
6. Słowniczek.

ReadMe.txt file lub **URL address** – z wyjaśnieniem jak używać danych

Przykład standardu Dublin Core Standard

- tytuł zestawu danych/datasetu
- nazwisko/numer ID/ORCID
- opis danych
- data zebrania danych
- obszar pozyskania danych (krajowy, regionalny, globalny)
- format
- język
- lokalizacja danych
- DOI
 - status dostępu&embargo
- licencja (np. CC BY)
- Instytucja finansująca badanie
- powiązane zestawy danych
- sposób cytowania danych



Data

Metadata

Źródło: www.dataedu.com, Autor: Piotr Kononow

Plan Zarządzania Danymi

Ogólne Informacje



Plan Zarządzania Danymi / Co to jest Plan Zarządzania Danymi

- to **część planu badań**
- jego celem jest uzupełnienie planu badań o **techniczny opis** sposobu zarządzania danymi

Plan Zarządzania Danymi / Dlaczego jest ważny?

- zabezpieczenie danych przed ich utratą lub nieuprawnionym wykorzystaniem
- pomaga monitorować postępy w realizacji badań
- ułatwia współpracę pomiędzy naukowcami pracującymi w danym projekcie
- wzmacnia kontrolę dostępu do danych i jakości danych
- pomaga zorganizować i skoordynować realizację badań
- zabezpiecza osobowe i wrażliwe dane np. przed ich nieuprawnionym wykorzystaniem
- wspiera organizację długotrwałego przechowywania danych

Plan Zarządzania Danymi / Ogólne cechy DMP w NCN

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

1. Opis danych oraz pozyskiwanie lub ponowne wykorzystanie dostępnych danych

Plan Zarządzania Danymi / Opis danych

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

- **Nowe dane:** metody pozyskiwania i analizowania danych i oprogramowanie, które posłużą do pozyskiwania lub wytwarzania nowych danych
- **Dane istniejące:** własne lub będące w posiadaniu strony trzeciej
- sposób udokumentowania pochodzenia danych
- sposób organizacji plików i zarządzanie ich różnymi wersjami
- kontrola jakości i spójności (kalibracja, walidacja, recenzje...)
- informacja czy konieczna będzie digitalizacja danych analogowych lub wydanych w formie papierowej (np. mapy, fotografie, teksty)

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
- inny

Plan Zarządzania Danymi / Opis danych

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

- **format (przechowywanie) i szacunkowa objętość** danych wytworzonych, pozyskanych, ponownie wykorzystanych



Formaty (przykłady)

- Pliki tekstowe: .txt, .rtf, .docx, .odt,
- Ankiety online: html
- Nagrania: .wav, .aif, .flac, .wv, .apl, .mp3
- Zdjęcia, obrazy: .bmp, .gif, .jpg, .png
- Bazy danych: .xls, .spv, .dbs,
- Pliki charakterystyczne dla urządzeń pomiarowych/badawczych, np. mikroskopu

Otwarte formaty!

kluczowy dla dostępności

dla użytkowników format szeroko stosowany

Przewidywana ilość danych ogółem:KB/MB/TB



2. Dokumentacja i jakość danych

Plan Zarządzania Danymi / Dokumentacja i jakość

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

- zakres metadanych
- międzynarodowe standardy lub schematy służące do organizacji metadanych (Dublin Core, DDI)
- jakie informacje będą podawane, aby potencjalni użytkownicy byli w stanie w przyszłości odczytać i zinterpretować zebrane dane
- sposób utworzenia odpowiedniej dokumentacji (np. struktura folderów, wersjonowanie)

Plan Zarządzania Danymi / Dokumentacja i jakość

2.2 Jakie planują Państwo zastosować środki kontroli jakości?

W projekcie należy wykazać wysokiej jakości dane.

- wpływ metod pozyskiwania, przetwarzania i analizy danych
- wpływ instrumentów badawczych i infrastruktury badawczej na jakość danych
- metody weryfikacji jakości danych
- eliminacja błędów pomiarowych i stronniczości
- minimalizacja ryzyka dotyczącego poprawności danych (zabezpieczenie przed nieuprawnioną modyfikacją)

3. Przechowywanie i tworzenie kopii zapasowych podczas badań.

Plan Zarządzania Danymi / Przechowywanie i kopie zapasowe

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

- bezpieczeństwo danych
- pojemność pamięci i miejsce przechowywania danych
- sposób i procedury tworzenia kopii zapasowych: zasada 3+2+1
- odzyskiwanie danych w przypadku utraty
- bezpieczeństwo przepływu danych między członkami zespołu: autoryzowany dostęp lub log-book
- częstotliwość wykonywania kopii
- zabezpieczenie danych przez zastosowanie programów antywirusowych i haseł dostępu do plików/folderów/komputera/chmury
- przenoszenia danych z urządzeń mobilnych, stanowisk terenowych lub sprzętu domowego na główny serwer w miejscu pracy

Plan Zarządzania Danymi / Przechowywanie i kopie zapasowe

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

- miejsce przechowywania danych wrażliwych/osobowych
- sposób odzyskiwania danych utraconych w wyniku incydentu
- informacja kto będzie miał dostęp do danych w czasie trwania projektu i jak będzie przebiegać kontrola dostępu do danych (zwłaszcza w przypadku współpracy między kilkoma partnerami)

4. Wymogi prawne, kodeksy postępowania.



Plan Zarządzania Danymi / Wymogi prawne, kodeksy postępowania

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

- Informacja czy w procesie pozyskiwania danych niezbędna będzie ich **anonimizacja/pseudonimizacja lub inne rozwiązanie?**
- Powołanie się na odpowiednie regulacje
- W przypadku pełnej anonimizacji zgoda na udostępnianie danych nie jest wymagana, ale jest to elementem kultury badawczej i dobrej praktyki aby zarówno informacja jak i zgoda uwzględniały:
 - informację o archiwizowaniu danych badawczych i osobowych (dane osobowe będą bezpiecznie przechowywane/będą niejawne i ostatecznie zniszczone a zanonimizowane dane badawcze będą jawne, udostępniane innym i przechowywane przez czas określony w wytycznych organizacji finansującej badanie);
 - informację o długoterminowym przechowywaniu i wykorzystaniu zanonimizowanych danych;
 - oświadczenie uczestnika badania potwierdzające świadomą zgodę na udział.

Dane osobowe: regulacje

- Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2016/679 z dnia 27 kwietnia 2016 r. w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych i swobodnym przepływem takich danych oraz uchylające dyrektywę 95/46 / WE (ogólne rozporządzenie o ochronie danych) - Dz.U. L 119 z 4.5.2016
- Dyrektywa Parlamentu Europejskiego i Rady (UE) 2019/1024 z dnia 20 czerwca 2019 r. w sprawie otwartych danych i ponownego wykorzystywania informacji sektora publicznego PE / 28/2019/REV/1
- Etyka i ochrona danych (dokument sporządzony przez panel ekspertów na wniosek Komisji Europejskiej (DG ds. Badań Naukowych i Innowacji), 14 listopada 2018 r.
- Polityka ochrony danych instytucji, Polityka bezpieczeństwa informacji instytucji, Instytucjonalna polityka otwartości/otwartego dostępu (Zarządzenie nr...)

Dane osobowe: regulacje

- USTAWA z dnia 11 sierpnia 2021 r. o otwartych danych i ponownym wykorzystywaniu informacji sektora publicznego
- Ustawa z dnia 10 maja 2018 r. o ochronie danych osobowych (tekst jednolity - Dz.U. z 2019 r. Poz. 1781)
Ustawa z dnia 20 lipca 2018 r.
- Umowy bilateralne z instytucjami np. szkołami, JST, szpitalami, itp.
- Zgoda uczestnika badania: osoby fizyczne biorące udział w badaniu lub ich opiekunowie prawni, itp.
- Regulacje etyczne (Regulamin Komisji Etyki Badań)

Plan Zarządzania Danymi / Wymogi prawne, kodeksy postępowania

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

- właściciel danych (pierwotne: jednostka badawcza, wtórne: jednostka badawcza, osoba prywatna, instytucja publiczna, organizacja pozarządowa, firma, organizacja kościelna ...)
- zastosowane licencje (w tym Creative Commons)
- ograniczenia ponownego wykorzystania danych pochodzących od osób trzecich

Plan Zarządzania Danymi / Wymogi prawne, kodeksy postępowania (przekłady)

- Prawo autorskie
- Prawa własności intelektualnej
- Ustawa o ochronie baz danych
- Ustawa o zwalczaniu nieuczciwej konkurencji
- Umowy dwustronne
- Regulamin konkursu
- Umowa o grant
- Licencje (np. oprogramowanie / sprzęt)
- Regulamin korzystania z repozytorium

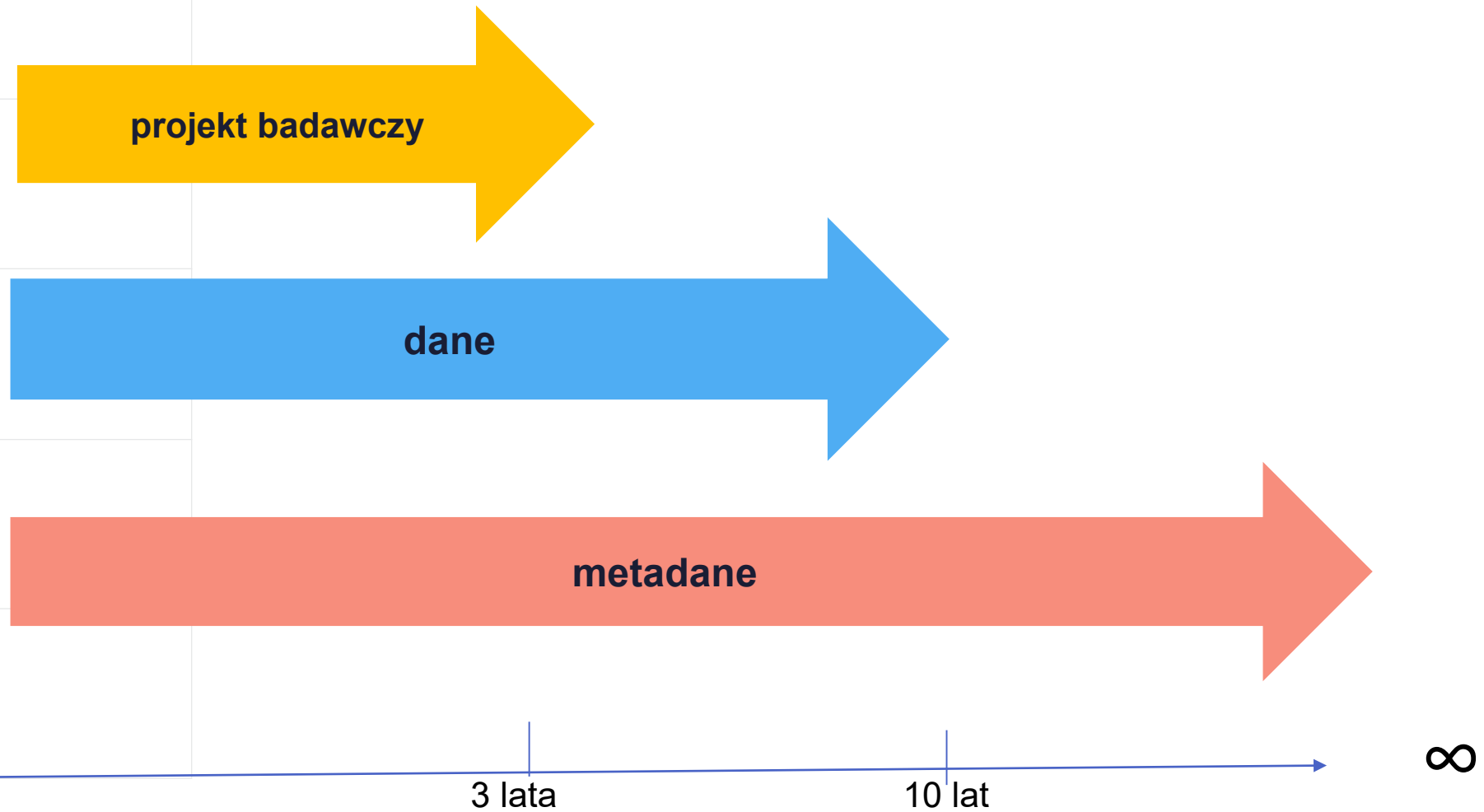
5. Udostępnianie i długotrwałe przechowywanie danych.

Plan Zarządzania Danymi / Udostępnianie i długotrwałe przechowywanie

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

- **sposób przekazywania informacji** o Państwa danych potencjalnym użytkownikom
- kiedy dane zostaną udostępnione
- **okres** przechowywania danych
- **ograniczenia i przeszkody** uniemożliwiające ich pełne lub częściowe udostępnienie
- informacja **czy wydawcy czasopism będą wymagać składowania danych** na poparcie ustaleń publikacji
- informacja czy **udostępnianie danych wymaga zgody uczestników badania**

Plan Zarządzania Danymi / Okres przechowywania (meta)danych



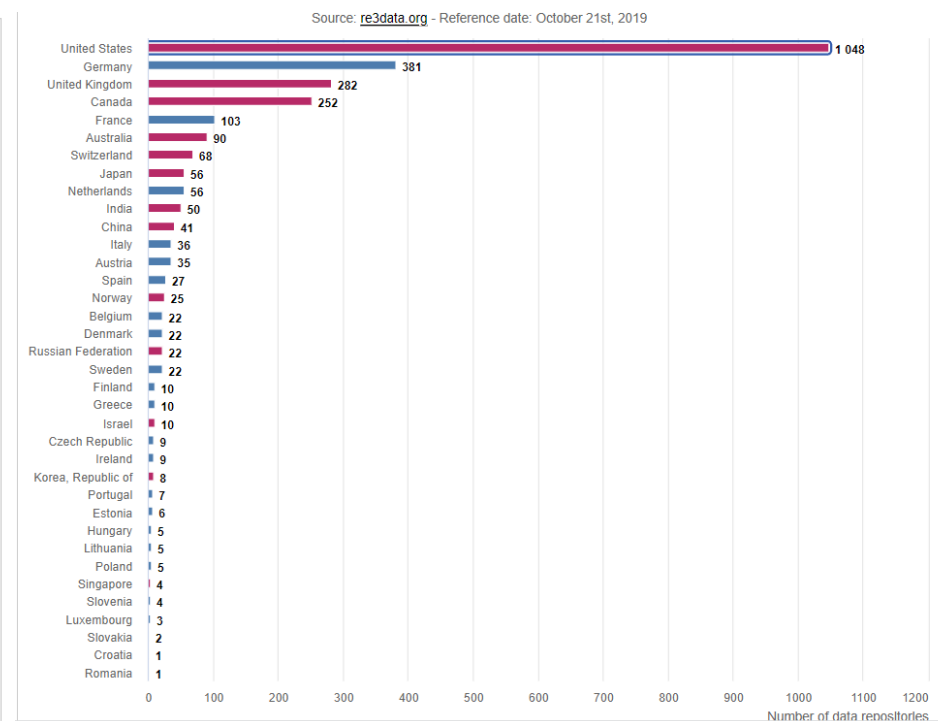
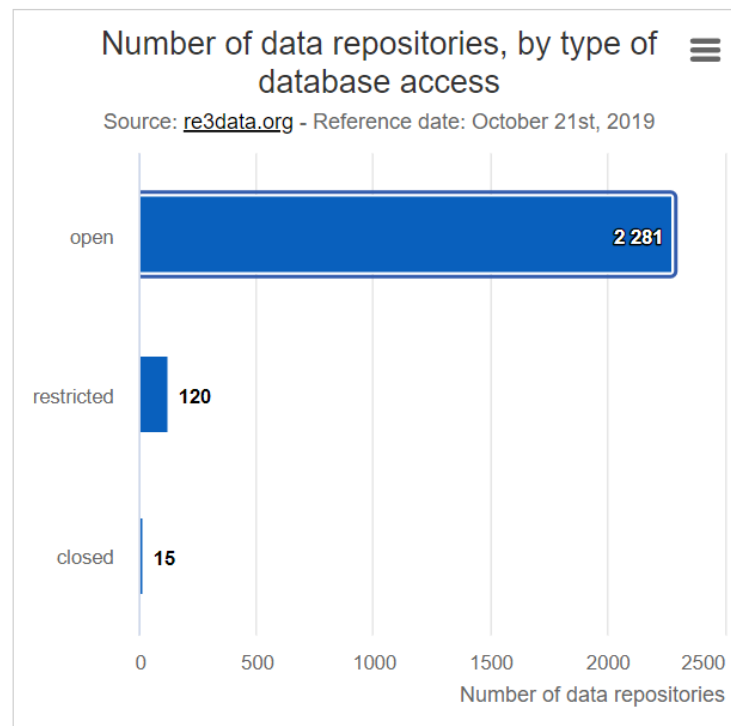
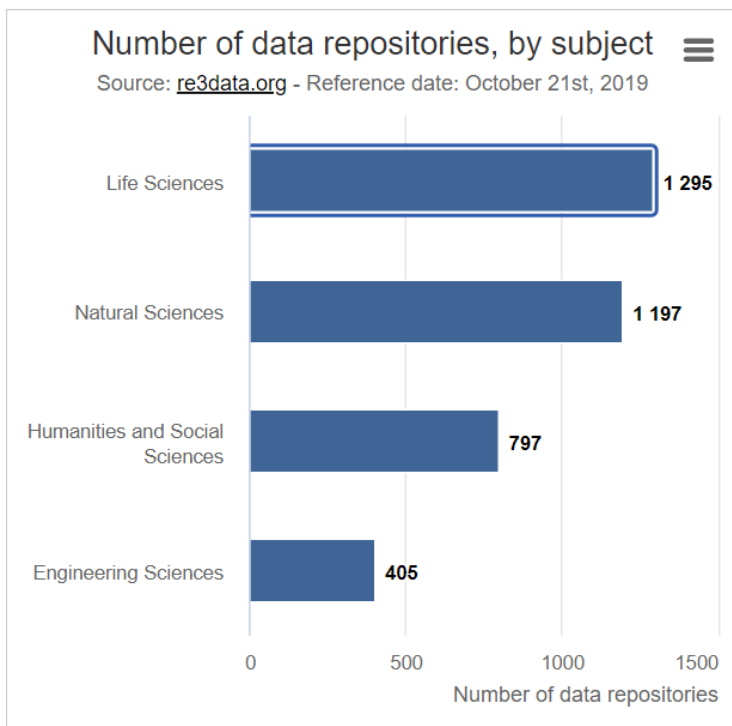
Plan Zarządzania Danymi / Udostępnianie i długotrwałe przechowywanie

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

- selekcja/procedura selekcji danych przeznaczonych do utrwalenia dot. **danych istotnych**; dane, które nie są dostępne publicznie; dane pozyskane za świadomą zgodą; dane nie zawierające danych personalnych lub wrażliwych; zgoda komisji etyki badań w krajach realizacji badania.
- informacja które dane trzeba zachować, a które zniszczyć z przyczyn wynikających z umów, przepisów prawnych lub regulacji
- organizacja danych zgodna z wytycznymi FAIR
- nazewnictwo plików i wersjonowanie plików
- wybór repozytorium

Gdzie udostępniać dane

- Repozytoria posiadające CoreTrust Seal
- Repozytoria w [Registry of Research Data Repositories](#)





Kryteria wyboru repozytorium

- regulamin repozytorium (warunki korzystania z usług, w tym w okresie przechowywania, opłata/lub jej brak za archiwizację danych)
- dyscyplinowe lub ogólne/sieroce
- „popularność” repozytorium wśród naukowców z naszej dyscypliny – wpływa na naszą widzialność
- popularyzacja dorobku naukowego: zasoby repozytorium znajdują się lub nie w bazach indeksujących repozytoria danych badawczych. przykłady baz indeksujących: data citation index, mendeley data czy google dataset search.
- kryteria formalne: zastosowanie zasad FAIR, przypisanie numeru PID; brak embargo

Plan Zarządzania Danymi / Udostępnianie i długotrwałe przechowywanie

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

- metody udostępniania danych (czynniki)
- informacja czy należy je przekształcić **do formatu standardowego lub otwartego**, aby dane mogły być przechowywane przez dłuższy okres i zachowały długi okres ważności
- informacja czy do **skanowania lub konwersji** niezbędny będzie **dodatkowy sprzęt lub oprogramowanie**
- mechanizm jaki posłuży do udostępniania danych (np. odpowiedzi na żądanie, repozytorium)

Plan Zarządzania Danymi / Udostępnianie i długotrwałe przechowywanie

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

- informacja czy zostanie uzyskany trwale przypisany identyfikator (PID), jeśli tak jaki? (np. cyfrowe identyfikatory dokumentu elektronicznego, numery dostępowe)

6. Zadania związane z zarządzaniem danymi oraz zasoby

Plan Zarządzania Danymi / Zadania związane z zarządzaniem danymi oraz zasoby

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

- Czy instytucja posiada personel odpowiedzialny za zarządzanie danymi w trakcie trwania projektu
 - Jeśli NIE: kto jest odpowiedzialny za archiwizację i długoterminowe zarządzanie danymi w jednostce
 - Jeśli TAK: jakie zajmuje stanowisko w obrębie jednostki
- Podział ról (kto odpowiada za weryfikację, korektę, kto wdraża DMP)

Plan Zarządzania Danymi / Zadania związane z zarządzaniem danymi oraz zasoby

Właściciel danych

Dokonuje bezpośredniej korekty danych.

Opiekun danych/data steward

Specjalista ds. jakości danych, ocenia ich jakość, podejmuje środki zaradcze.

Plan Zarządzania Danymi / Zadania związane z zarządzaniem danymi oraz zasoby

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

- **dodatkowe zasoby** do zarządzania danymi (osoby, czas, sprzęt lub oprogramowanie)
- **koszty** związane będą z zapewnieniem standardów FAIR w projekcie

Plan Zarządzania Danymi / Budżet

Na koszty pośrednie składają się:

- koszty pośrednie Open Access w wysokości do 2% kosztów bezpośrednich, które mogą być przeznaczone **wyłącznie** na koszty związane z udostępnieniem publikacji lub danych badawczych w otwartym dostępie;
- pozostałe koszty pośrednie w wysokości do 20% kosztów bezpośrednich, które mogą być przeznaczone na koszty pośrednio związane z projektem, **w tym** koszty udostępnienia publikacji lub danych badawczych w otwartym dostępie.

Maksymalna = nie oznacza, że musi tyle wynosić.

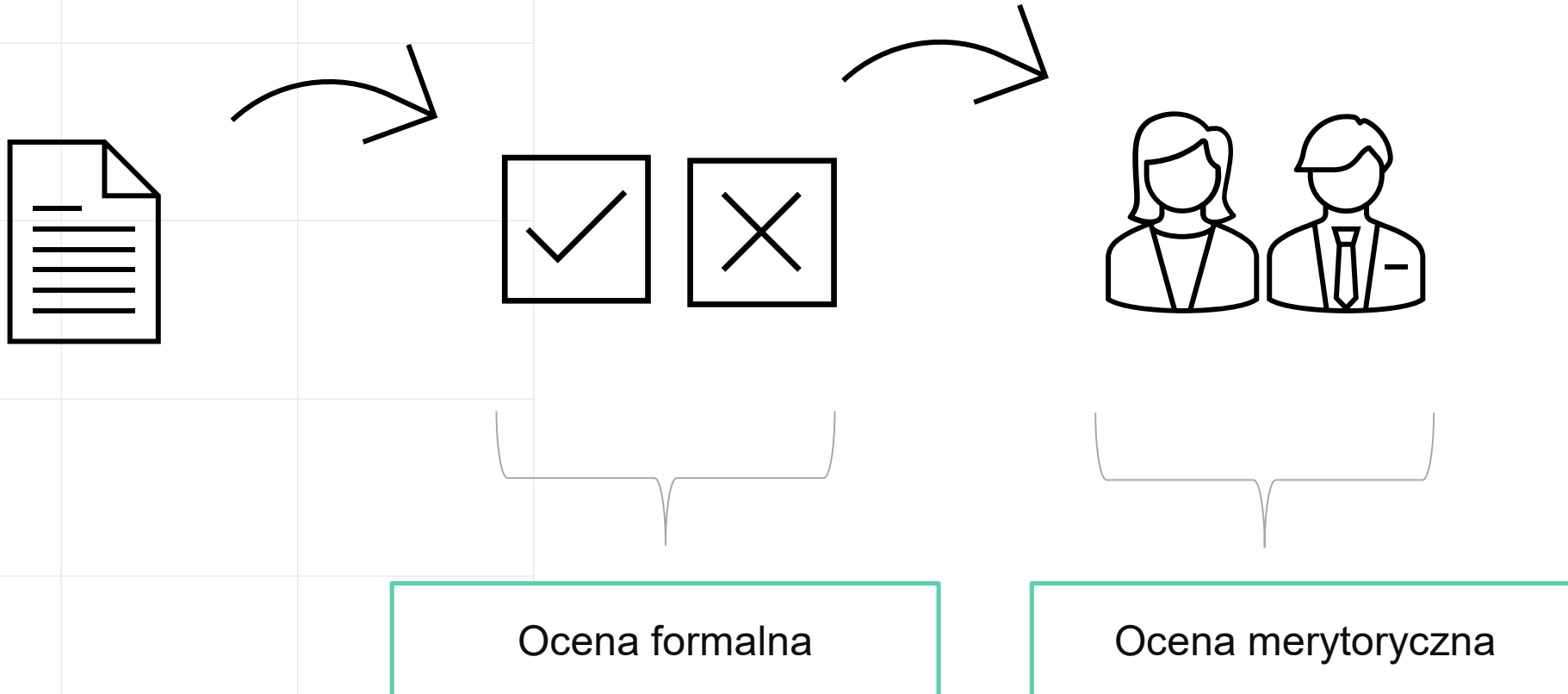
Na etapie realizacji projektu, podmiot realizujący jest **zobowiązany** do **uzgodnienia** z kierownikiem projektu zagospodarowania co najmniej 25% wartości kosztów pośrednich

Finansowanie po zakończeniu projektu (np.repozytorium)

Ocena Planów Zarządzania Danymi w NCN



Plan Zarządzania Danymi / Kto ocenia DMP



Plan Zarządzania Danymi / Rozliczanie DMP

- DMP może zmieniać się w trakcie realizacji projektu (zalecane)
- Nie ma obowiązku informowania NCN o zmianach
- Raport końcowy – należy opisać stan faktyczny na koniec realizacji projektu
- Raport końcowy – należy podać zestawy udostępnionych danych + metadane - nawet gdy dane nie zostały (jeszcze) udostępnione
- Podlega ocenie Zespołów Ekspertów
- Gdy brak kompletnego planu, błędy – wezwania / sankcje



Kontakt:

Natalia Galica

Narodowe Centrum Nauki

natalia.galica@ncn.gov.pl

