

Otwartość dla sprawiedliwości klimatycznej
Biblioteka Uniwersytetu Łódzkiego
28.10.2022 r.



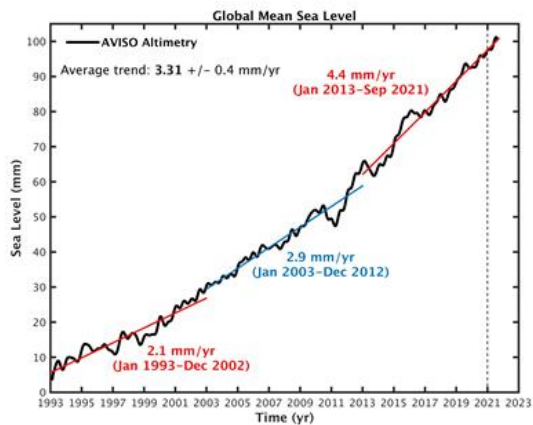
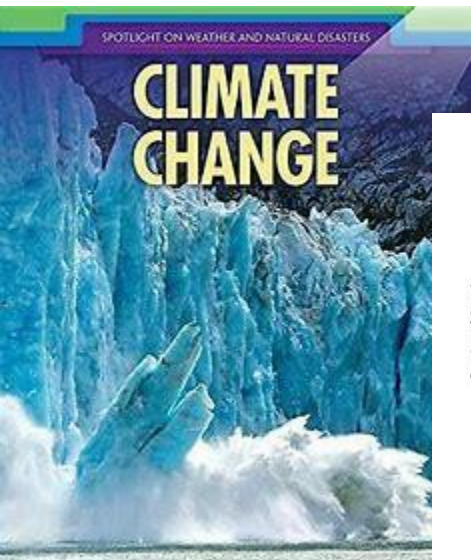
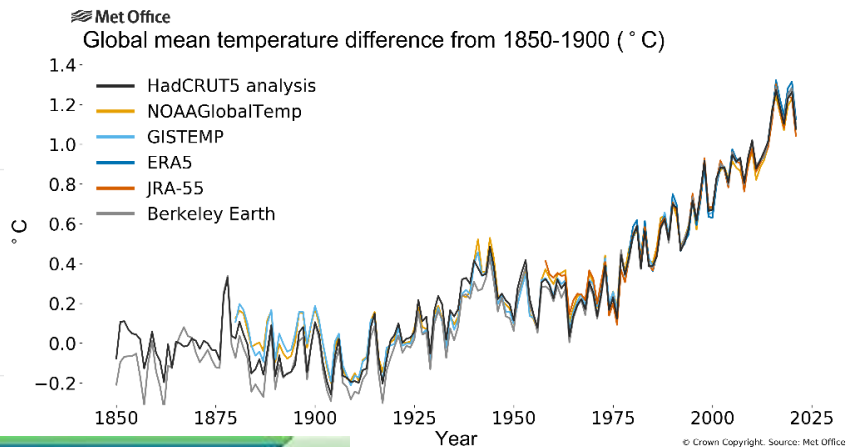
Globalne wyzwania a otwarte dane badawcze w polityce NCN

Natalia Galica
Samodzielne stanowisko ds. Otwartej Nauki

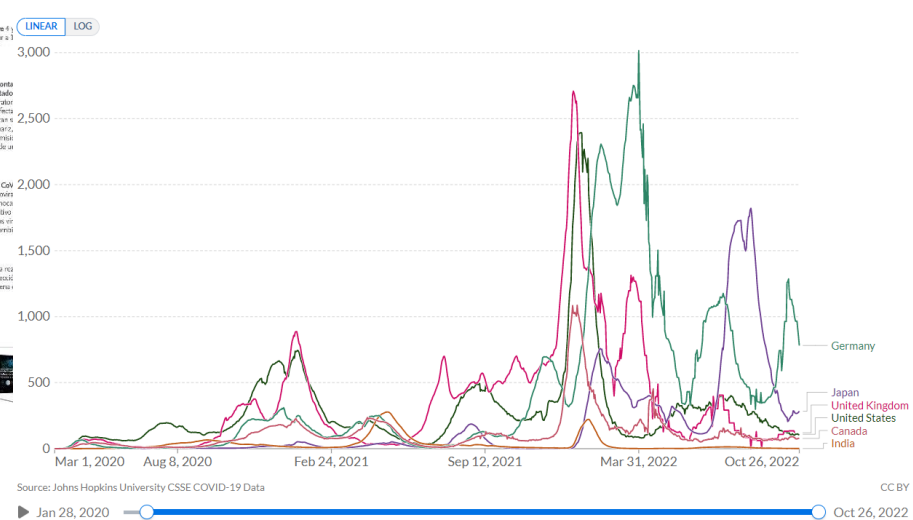
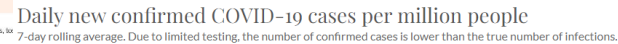
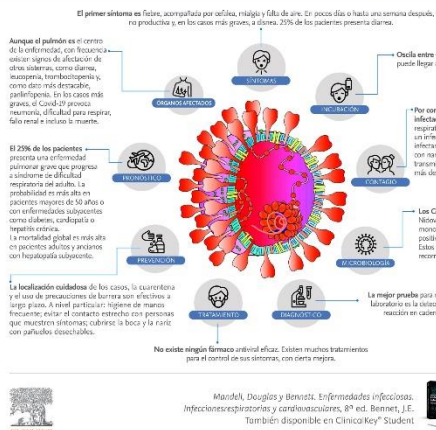
Plan spotkania

- Kontekst: globalne wyzwanie społeczne i naukowe
- Dlaczego ważne jest zapewnienie otwartego dostępu do danych badawczych
- Polityka otwartego dostępu NCN
- Podsumowanie

Globalne wyzwania/zagrożenia: klimat, pandemia, bezpieczeństwo...



El nuevo coronavirus SARS-CoV-2 y su enfermedad, Covid-19, ¿a qué nos enfrentamos?



<https://ourworldindata.org/coronavirus>

[Upload](#)
[Communities](#)

Featured communities

[Need help uploading?](#) [Contact us](#)

Coronavirus Disease Research Community - COVID-19

This community collects research outputs that may be relevant to the Coronavirus Disease (COVID-19) or the SARS-CoV-2. Scientists are encouraged to upload their outputs in this collection to facilitate sharing and discovery of information. Although Open Access articles and datasets are...

Curated by: Covid19_Team_OpenAIRE

Recent uploads

[January 29, 2022 \(y2\)](#)
[Software](#)
[Open Access](#)

Forecasting the propagation of pandemic shocks with a dynamic input-output model

[Pichler, Antonio](#),
 [Pangallo, Marco](#),
 [del Rio-Chanona, R. Maria](#),
 [Lafond, François](#),
 [Farmer, J. Doyné](#)

Reproduction code and data for Pichler, Pangallo, del Rio-Chanona, Lafond, Farmer (2022). Forecasting the propagation of

Need help?

Zenodo prioritizes all requested related to the COVID-19 outbreak.

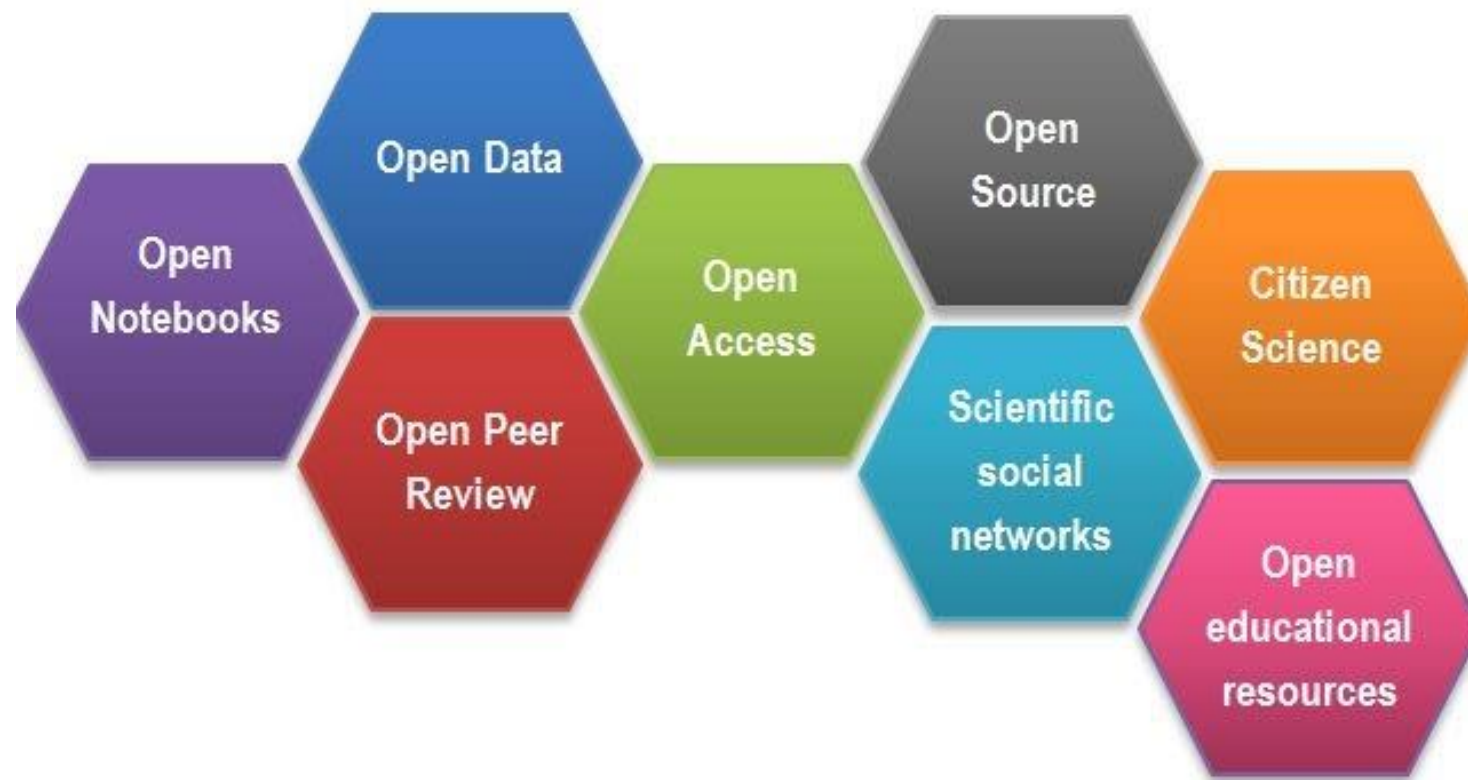
We can help with:



Globalny kryzys nauki

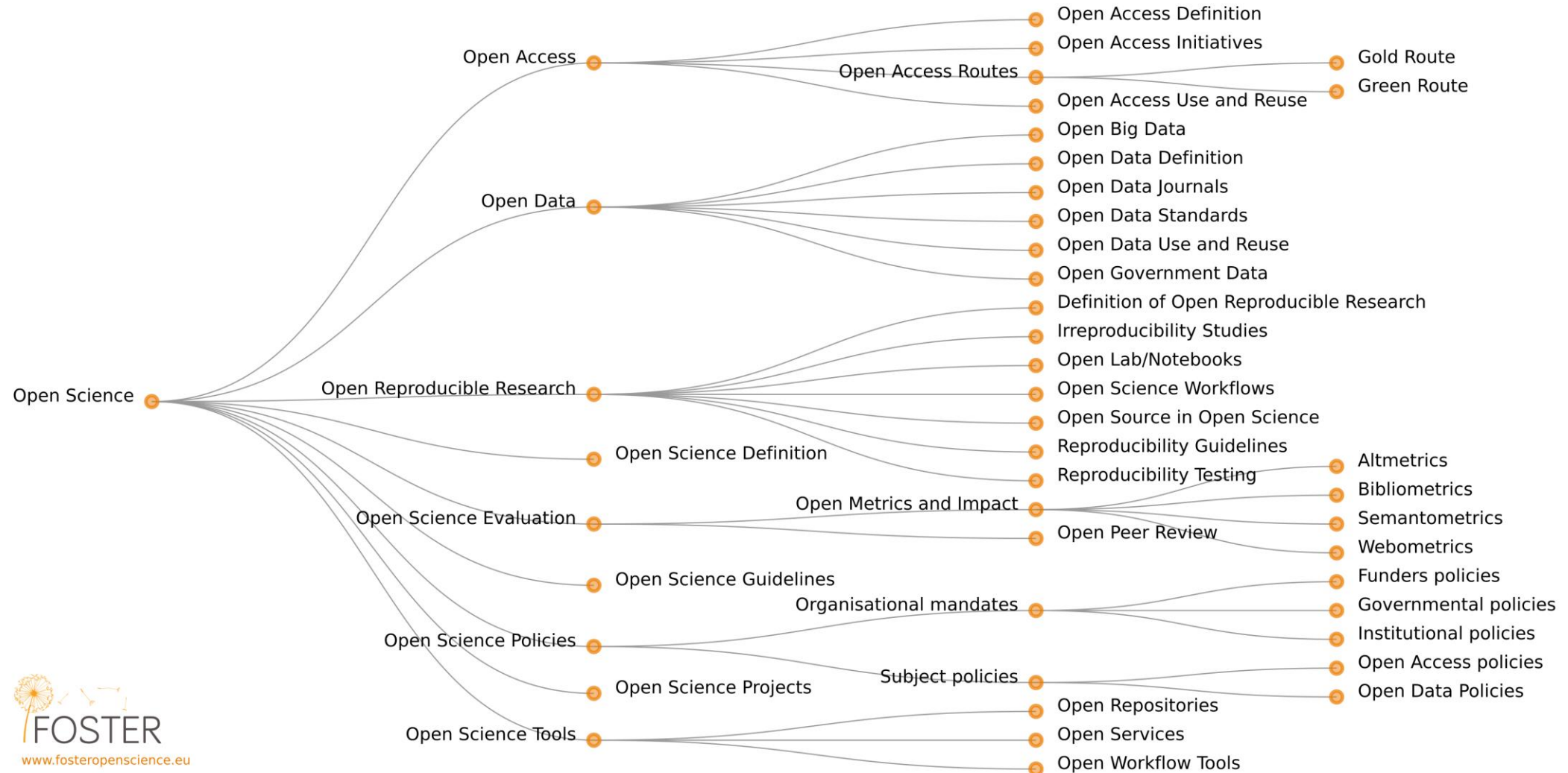
- Problemy z transparentnością procesu badawczego
- Problemy z rzetelnością badawczą i replikowalnością
- Nierówny i dyskryminujący dostęp do wyników badań (system subskrypcyjny)
- Działania nieetyczne
- Problemy z odpowiedzialnością naukową i użytecznością społeczną badań
- Ewaluacja naukowa oparta na wskaźnikach ilościowych a nie jakościowych
- Nieefektywne wykorzystanie środków publicznych [wielkokrotne finansowanie tych samych badań, powtarzanie tych samych błędów; brak dostępu do danych już pozyskanych]
- Globalne wyzwania i zagrożenia spowodowały konieczność dostosowania nauki do potrzeb społecznych
- Utrata zaufania społecznego

Transformacja nauki w naukę otwartą



Źródło: www.fosteropenscience.eu

Open Science Taxonomy



Kwestie regulacyjne

2013: San Francisco Declaration on Research Assessment (DORA); NCN sygnatariuszem DORA od 2018

2018: PLAN S; NCN sygnatariuszem-założycielem cOAlition S

2019: Dyrektywa Parlamentu Europejskiego i Rady (UE) 2019/1024 z dnia 20 czerwca 2019 r. w sprawie otwartych danych i ponownego wykorzystywania informacji sektora publicznego

2020: Polityka Narodowego Centrum Nauki dotycząca otwartego dostępu do publikacji naukowych (2020) (obowiązuje dla wszystkich umów od 1.1.21)

2021: Ustawa z dnia 11 sierpnia 2021 r. o otwartych danych i ponownym wykorzystywaniu informacji sektora publicznego.

2022: Polityka Naukowa Państwa

Nowy paradygmat: NAUKA – DOBRO PUBLICZNE

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.

[art. 10, 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]

Zarządzanie otwartymi danymi badawczymi w polityce NCN

- 2019: **Plan Zarządzania Danymi** obowiązkowym elementem każdego projektu badawczego [wytyczne Science Europe]
- Styczeń 2021 (umowy): dane powiązane z publikacjami muszą być obowiązkowo udostępniane

Dane powiązane
(podstawowy zestaw danych)

Wiarygodne i zaufane repozytorium

Metadane

CC0

Plan Zarządzania Danyami

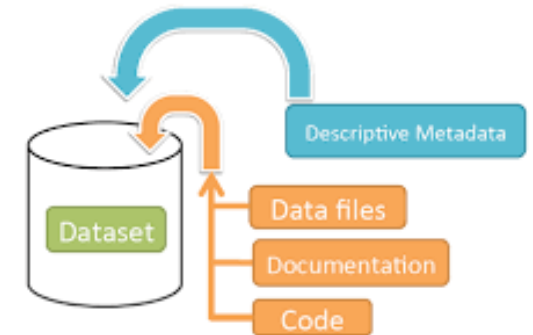
W rozumieniu Planu Zarządzania Danyami, dane to **wszelkiego rodzaju informacje** (ilościowe i jakościowe) wytwarzane, zbierane, przetwarzane, poddawane analizie w ramach projektu badawczego, obejmujące **wszystkie możliwe formaty** zarówno cyfrowe jak i nie cyfrowe, w tym:

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

Na czym polega otwarty dostęp do danych:

- Dane powinny być tak otwarte, jak to możliwe i na tyle zamknięte, na ile to jest konieczne.
- Udział w udostępnianiu bez embarga.
- 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.
- **Zasady FAIR:** zasady udostępniania danych badawczych, ale to nie są sztywne zasady, a raczej wytyczne do właściwej pracy z danymi, aby uczynić dane dostępnymi zarówno dla użytkowników jak i oprogramowania komputerowego przeszukującego bazy danych bez udziału człowieka.

Schematic Diagram of a **Dataset** in Dataverse 4.0



Container for your data, documentation, and code.

Źródło: <https://repod.icm.edu.pl/>

Zasady FAIR

Findable: odnajdywalne

Metadane; PID powiązany z metadanymi; metadane zawierają osobny PID danych (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 i ontologie, odnośniki do powiązanych zbiorów.

Reusable: możliwe do ponownego wykorzystania

prawna interoperacyjność (licencje) – odczytywane przez ludzi i maszyny, precyzyjne informacje dotyczące cytowania (np. gdy dane pozyskane od innych) – kogo należy cytować; dostosowanie do standardów obowiązujących w danej dziedzinie

Dokumentacja

Dokumentacja (np. metodologia oraz sposoby pozyskiwania i organizacji danych):

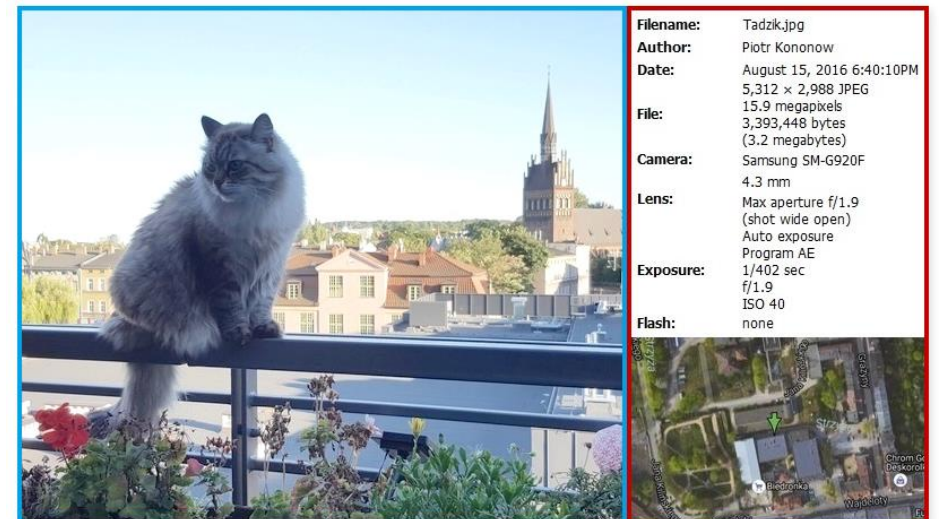
1. Opis kontekstu badania.
2. Opis metodologii: metody pozyskiwania danych, metody analizy danych, użyte oprogramowanie i sprzęt.
3. Struktura i relacje między folderami: chronologicznie i tematycznie.
4. Sposób kontroli jakości danych.
5. Informacja o otwartości danych.
6. Słowniczek.

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

Metadane

Przykład Dublin Core Standard:

- tytuł zestawu danych/datasetu/
- źródło danych
- nazwisko/numer ID/
ORCID
- opis danych
- data zebrania danych
- obszar pozyskania
danych (krajowy, regionalny,
globalny)
- format
- język
- lokalizacja danych
- PID, np. DOI
- status dostępu&embargo
- licencja (np. CC BY, CC0)
- Instytucja finansująca badanie
- powiązane zestawy danych
- sposób cytowania danych



Data

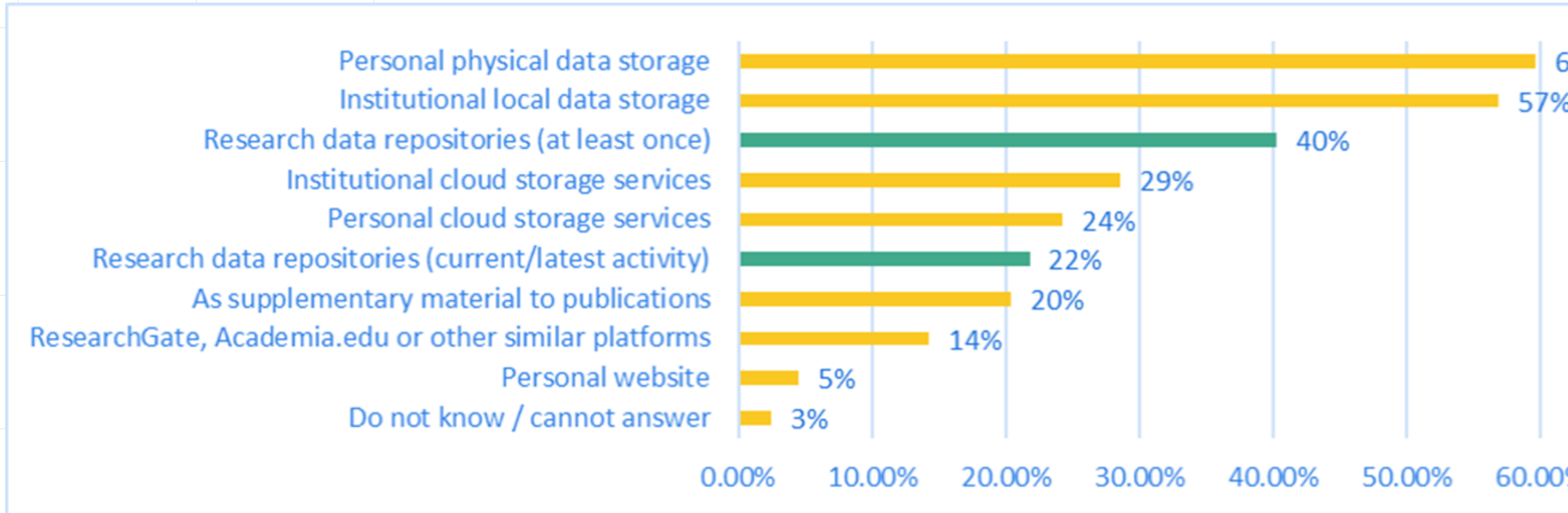
Metadata



NARODOWE CENTRUM NAUKI

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

Przechowywanie danych

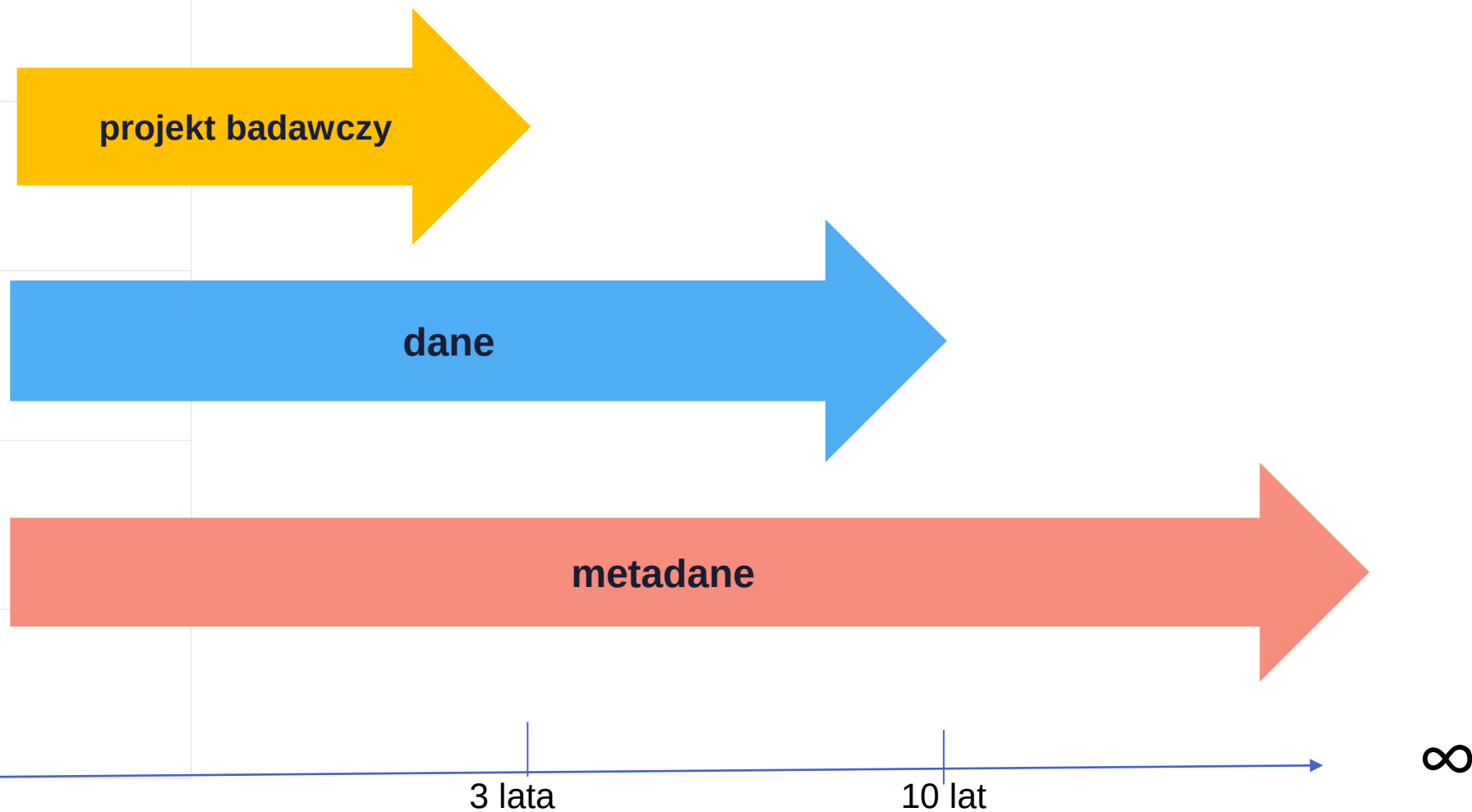


Source: European Research data landscape study 2022 commissioned by the European Commission Elaboration by the study performers based on unweighted researchers' survey data. Total N=10,914.

Przechowywanie danych

- bezpieczeństwo danych przed utratą lub zniszczeniem: reguła 3+2+1
- komercyjne vs niekomercyjne lokalizacje
- częstotliwość wykonywania kopii zapasowych...
- osoba odpowiedzialna za wykonywanie kopii zapasowych: opiekun/kurator danych=data steward, kierownik badania, inna osoba
- aktualizacja kopii zapasowych, autoryzacja
- zabezpieczenie danych przez zastosowanie programów antywirusowych i haseł dostępu do plików/folderów/komputera/chmury

Okres przechowywania (meta)danych w polityce NCN

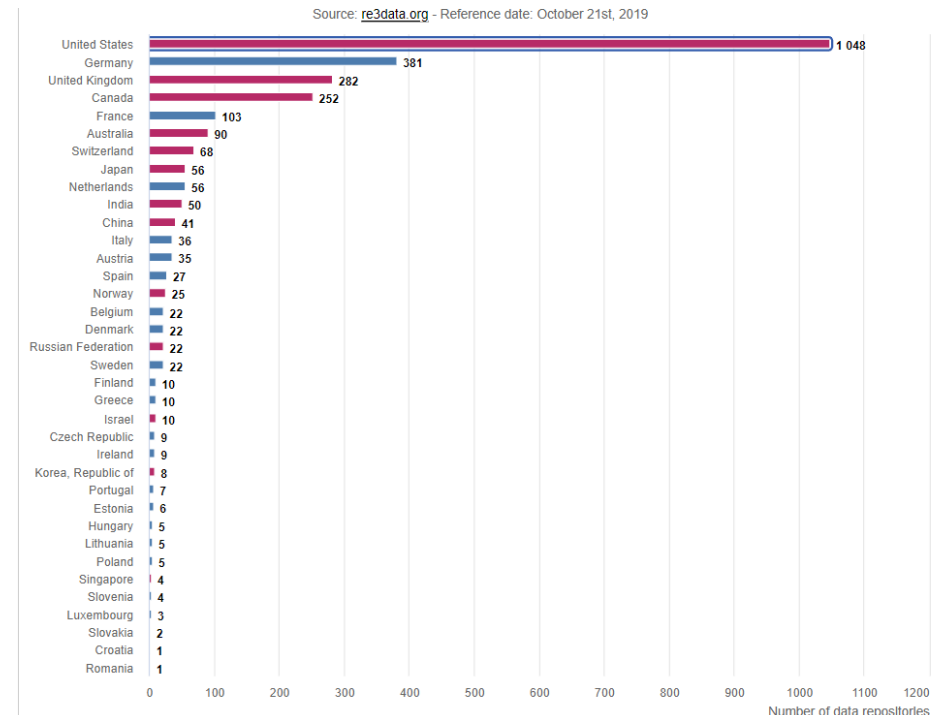
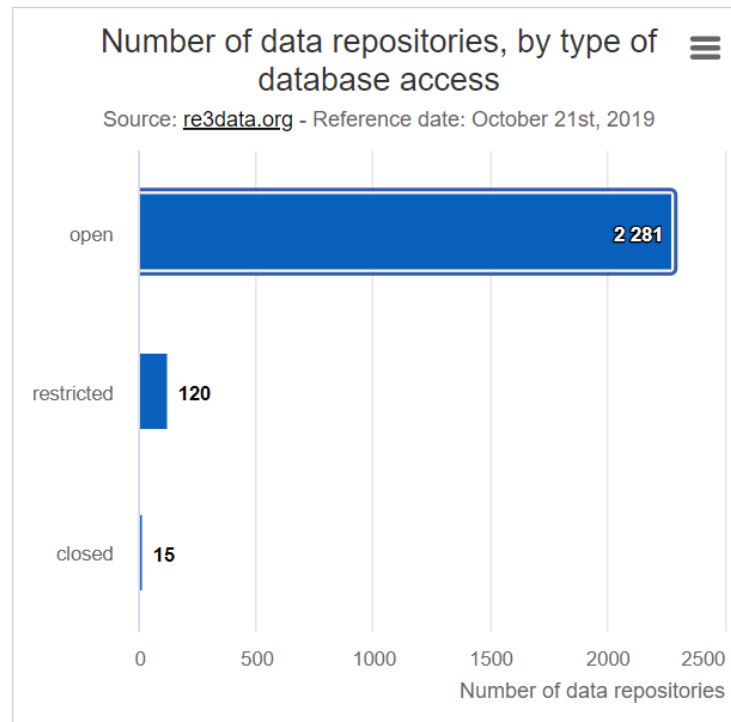
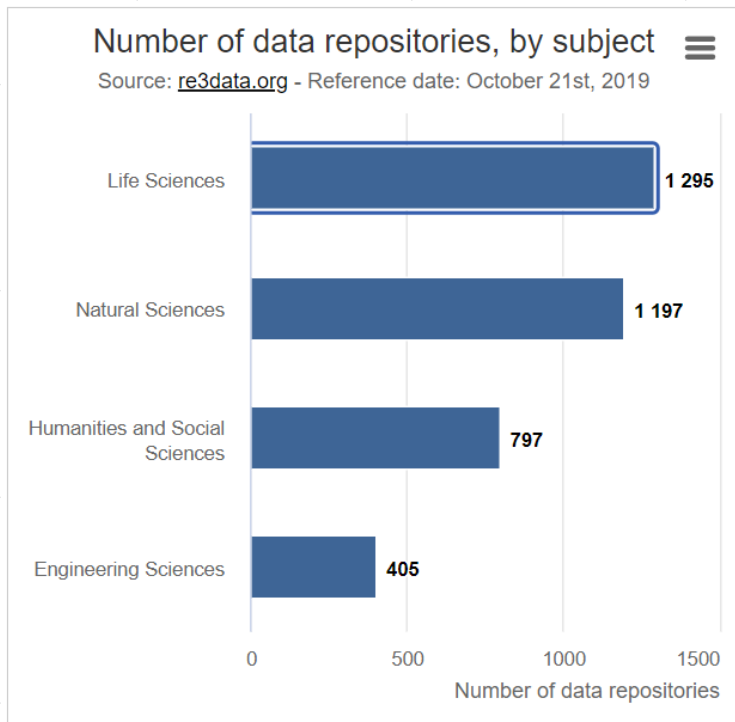


Okres przechowywania danych

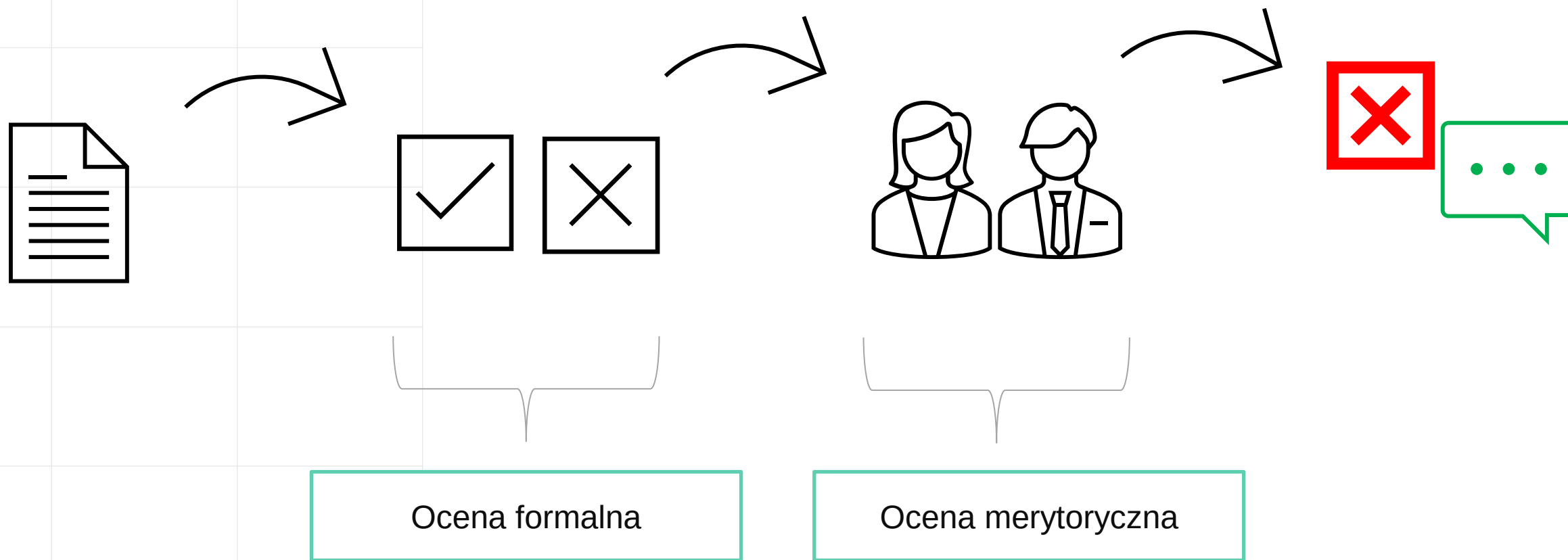
- 10 lat – kodeks rzetelności naukowej, umowa grantowa; gdzie: repozytoria danych otwartych spełniające wymagania grantodawcy
- do oceny badacza: jak długo dane posiadają wartość naukową?
- metadane – przechowywane bezterminowo
- repozytoria danych otwartych:
 - rejestr: www.re3data.org
 - najwyższa jakość to CoreTrustSeal

Gdzie możemy udostępniać dane?

Rejestr: www.re3data.org



Ocena DMP



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
- Podlega ocenie Zespołów Ekspertów
- Gdy brak kompletnego planu, błędy – wezwania / sankcje



Dlaczego warto otwierać dane badawcze – perspektywa naukowa

Kwestie naukowe

- cytowalność danych
- unikanie wcześniej popełnionych błędów badawczych
- rzetelność naukowa, wiarygodność i transparentność procesu badawczego
- replikowalność badań
- możliwość (ponownego) wykorzystania danych
- odpowiedzialność naukowa

Kwestie instytucjonalne i narodowe

- premiowanie publikacji danych (polityka instytucjonalna, krajowa. UE)
- publikowanie wyników badań w czasopismach naukowych (polityka wydawnictw)
- ewaluacja naukowa/ocena naukowa [np. Science Europe: Agreement on Reforming Research Assessment 2022]
- wymagania agencji finansujących badania – NCN, Program Horizon Europe itp.

Kwestie finansowe

- unikanie opłacania wiele razy tych samych badań – badania finansowane są przez podatnika
- efektywne wykorzystanie środków publicznych

Dlaczego warto otwierać dane badawcze – perspektywa społeczna

Biorąc pod uwagę pilność szukania rozwiązań wobec wyzwań globalnych, jakimi są kryzys klimatyczny, pandemia czy bezpieczeństwo, kraje i organizacje międzynarodowe nie mogą sobie pozwolić na stratę czasu lub brak skoordynowanych działań. Jeżeli działania rządów są oparte na rzetelnych, otwartych i weryfikowanych danych, to możliwe jest:

- **Ulepszone podejmowanie decyzji** od jednostek i społeczności lokalnych po działalność rządów, organizacji międzynarodowych: polityki oparte na wiedzy/danych; proces decyzyjny oparty na wiedzy/danych;
- **Wzmocnioną współpracę** pomiędzy krajami, regionami, społecznościami lokalnymi, naukowcami w celu podejmowania działań skierowanych na minimalizowania zagrożeń wynikających ze zmian klimatu, budowa zaufania w procesie np. współpracy między analitykami danych a osobami, które reprezentują potrzeby społeczności dotkniętych katastrofami
- **Ulepszone monitorowanie**, w tym m.in. wdrażania polityk i programów klimatycznych
- **Dostępne i odpowiedzialne modelowanie**, stanowiące podstawę do analiz i symulacji zmian klimatu, szacowania zagrożeń oraz przygotowywania prognoz.

Kontakt

Natalia Galica
Samodzielne stanowisko ds. otwartej nauki
Narodowe Centrum Nauki
natalia.galica@ncn.gov.pl

