

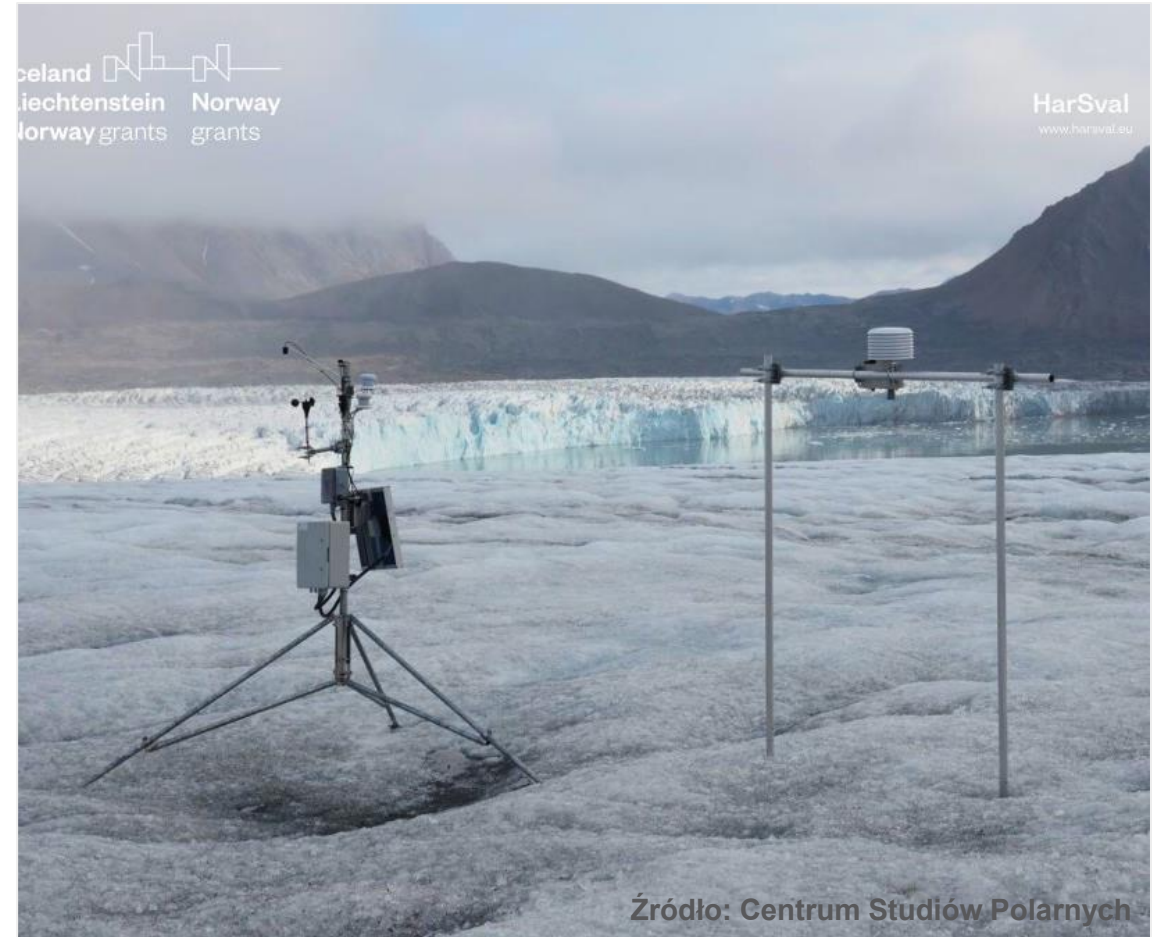


Warsztaty HarSval: FAIR Data Principles and Data Format

Zespół ds. Otwartej Nauki
Narodowe Centrum Nauki

Gabriela Czarny
Natalia Galica

Centrum Analiz Geoprzestrzennych i Obliczeń
Satelitarnych, Józefosław
20 czerwca 2024 r.

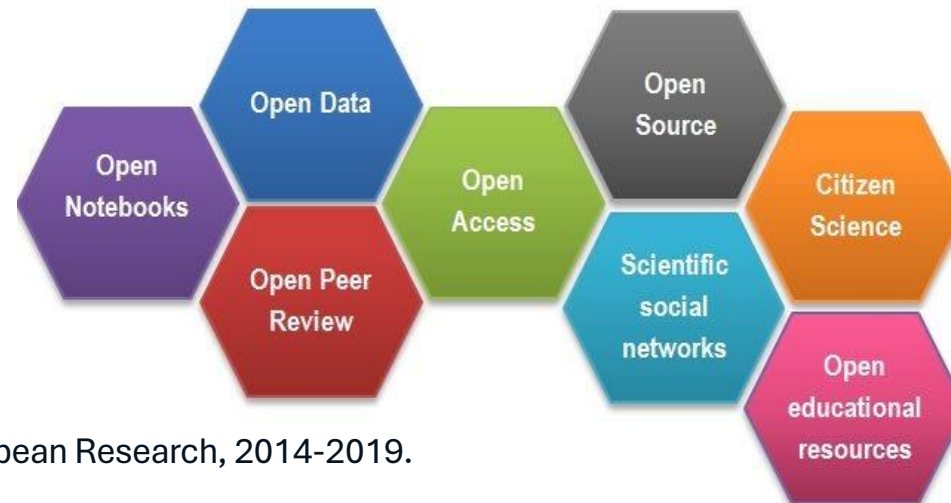


Otwarta nauka a otwarte dane badawcze

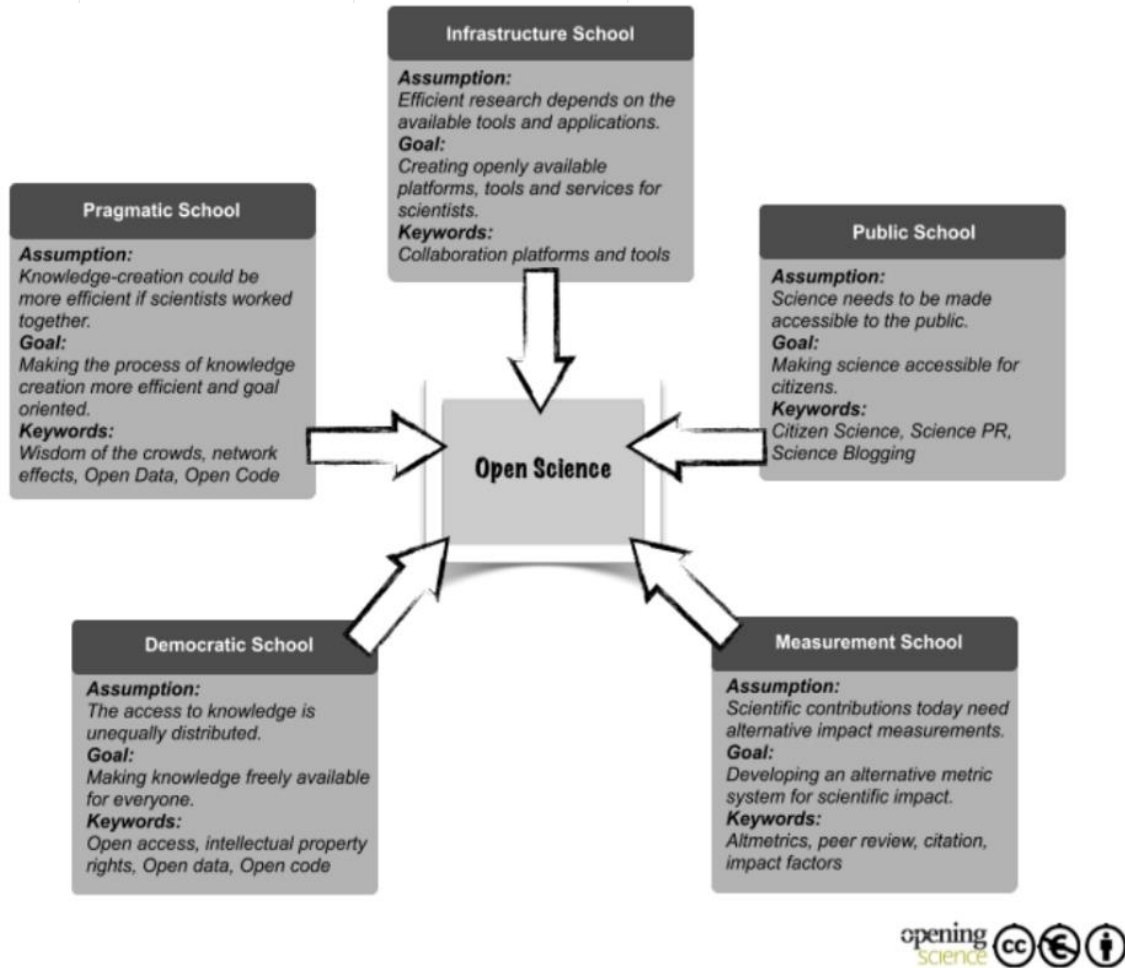
Otwarta nauka (ang. *Open Science*) – to koncepcja, która łączy różne ruchy i praktyki mające na celu udostępnienie wiedzy naukowej, metod, danych i dowodów badawczych, zwiększenie współpracy naukowej, dzielenie się informacjami z korzyścią dla nauki i społeczeństwa oraz otwarcie procesu tworzenia i obiegu wiedzy naukowej z uwzględnieniem podmiotów społecznych spoza zinstytucjonalizowanej społeczności naukowej.

UNESCO Recommendation on Open Science, przyjęta na 41. sesji Konferencji Generalnej UNESCO w listopadzie 2021 r.

<https://unesdoc.unesco.org/ark:/48223/pf0000379949.locale=en>



Otwarta nauka: przesłanki filozoficzne a rzeczywistość



Dlaczego ważne jest otwarty dostęp do danych?

Perspektywa badacza i 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 i replikacja procesu badawczego prowadząca do zwiększenia rzetelności naukowej, wiarygodność i zaufania wobec naukowca i nauki,
- odpowiedzialność nauki wobec społeczeństwa.

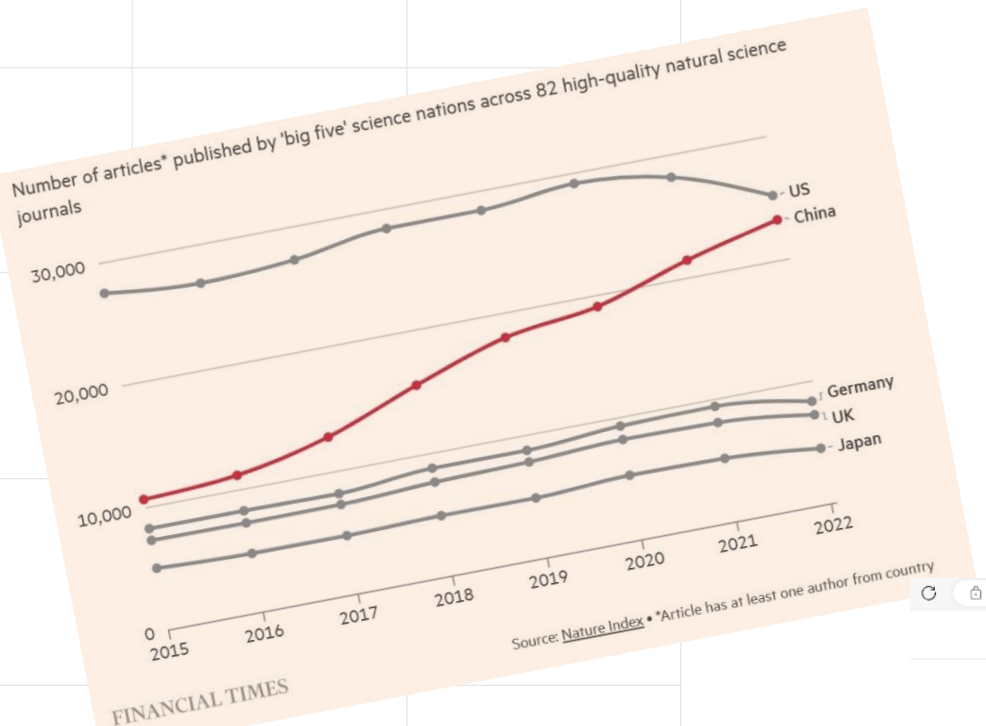
Perspektywa publiczna, kwestie instytucjonalne i polityki narodowe:

- efektywność wydatkowanie środków publicznych na badania (unikanie finansowania tych samych badań i tych samych błędów badawczych; efektywne wykorzystanie już istniejących danych),
- premiowanie publikowania danych w otwartym dostępie (polityka instytucjonalna, krajowa i UE),
- wymagania agencji finansujących badania,
- ewaluacja naukowa/ocena naukowa,
- wspieranie innowacyjności rynkowej i rozwoju społecznego,
- zaufanie do nauki i naukowców,
- szukanie efektywnych rozwiązań na globalne i lokalne wyzwania.

Wydawcy:

- publikowanie wyników badań w czasopismach naukowych z obowiązkiem prezentowania danych (polityka wydawnictw),
- obawa o reputację „czasopisma drapieżnego”.

Otwarte dane: przesłanki



Liczba artykułów naukowych

2022: 16,08 miliona

- 4,4 miliona w USA
- 3,7 miliona w Chinach

Źródło: Institute for Scientific Information, 2022; FT, 2023

Liczba artykułów wycofanych:

2013: 1 000

2022: 4 000

2023: 10 000

Źródło: Nature 2023

Scientific research [+ Add to myFT](#)

'Open science' advocates warn of widespread academic fraud

Scandals at Stanford and Harvard show manipulation of research remains an issue despite growth of 'data detectives'



The Observer Peer review and scientific publishing

'The situation has become appalling': fake scientific papers push research credibility to crisis point

Last year, 10,000 sham papers had to be retracted by academic journals, but experts think this is just the tip of the iceberg

Robin McKie

Sat 3 Feb 2024 17:00 CET



Otwarte dane: przesłanki ekonomiczne

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

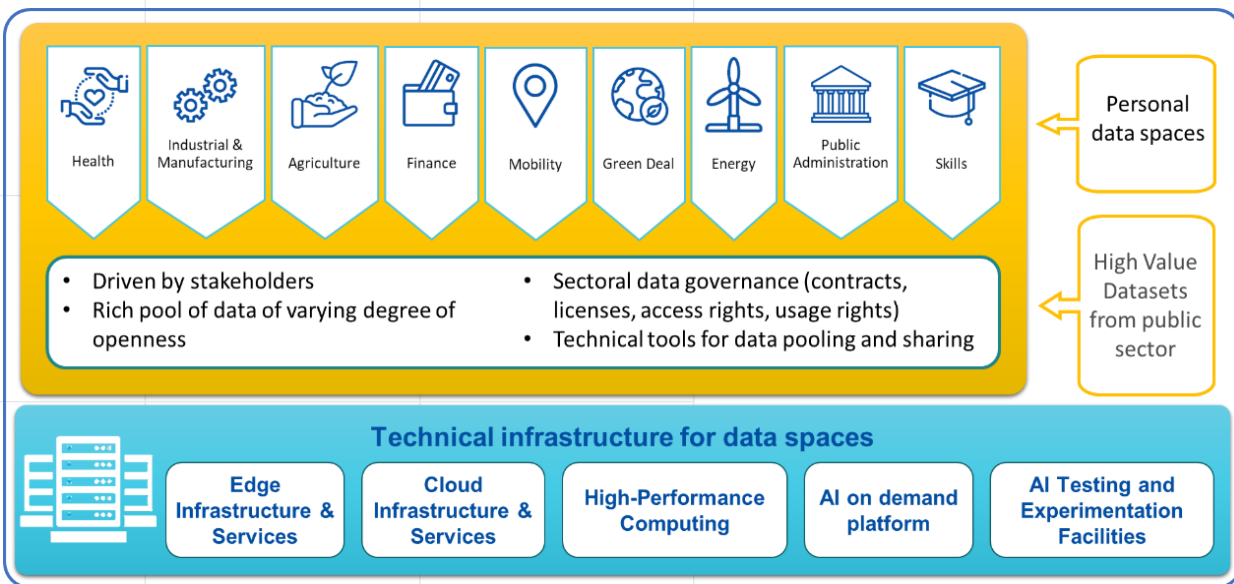
From 57% in 2018

Koszt niestosowania
zasad FAIR w UE

**10,2 mld
euro/rok**

Źródło: Komisja Europejska 2018

Regulacje międzynarodowe



Poziom globalny

2004: OECD Declaration on Access to Research Data From Public Funding

2010: Singapore Statement on Research Integrity, 2nd Conference on Research Integrity

2013: San Francisco Declaration on Research Assessment (DORA)

2021: UNESCO Recommendation on Open Science.

Poziom europejski

2002: Budapest Open Access Initiative Declaration

2016: Rada Unii Europejskiej (*Council Conclusions on the Transition towards Open Science System*),

2017: European Research Council (*Guidelines on Implementation of Open Access to Scientific Publications and Research Data*),

2018: Komisja Europejska (*Commission recommendation on access to and preservation of scientific information 2018*), Koalicji S (*Plan S*).

2019: Dyrektywa Parlamentu Europejskiego i Rady (UE)

2019/1024 z dn. 20 czerwca 2019 r. w sprawie otwartych danych i ponownego wykorzystywania informacji sektora publicznego

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; European Strategy for Data, Data Act, Data Governance Act, Digital Services Act, Digital Markets Act...



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)



N A R



Regulacje krajowe

2013: Konferencja Rektorów Akademickich Szkół Polskich (KRASP): Stanowisko Prezydium KRASP i Prezydium PAN w sprawie zasad otwartego dostępu do treści publikacji naukowych i edukacyjnych.

2015: Kierunki rozwoju otwartego dostępu do publikacji i wyników badań naukowych w Polsce, MEiN

2020: Polityka NCN dot. otwartego dostępu do publikacji naukowych

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

2022: Polityka Naukowa Państwa, MEiN (obecnie MNiSW)

2024 (prognoza) Polityka dotycząca Otwartych Danych Badawczych, MNiSW



Definicja

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 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

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.

Polityka Naukowa Państwa (MEiN/obecnie MNiSW 2022)

Polityka Naukowa Państwa

MINISTERSTWO EDUKACJI I NAUKI

Warszawa 2022

Spis treści	
1. Wprowadzenie	3
2. Cele Polityki Naukowej Państwa (PNP)	4
3. System szkolnictwa wyższego i nauki - szanse i wyzwania	5
4. Priorytety Polityki Naukowej Państwa	8
4.1 Rozwój systemu szkolnictwa wyższego i nauki	8
Rozwój instytucjonalny i promowanie innowacyjności	8
Podnoszenie jakości realizowanych badań	10
Zasoby ludzkie	13
Infrastruktura badawcza	17
Współpraca z otoczeniem	18
Szkolnictwo wyższe – stabilny rozwój	21
4.2 Rola Polski w rozwoju globalnym	22
Synergia działań badawczych i edukacyjnych w UE	24
Umiejdzynarodowienie i mobilność	27
4.3 Zasoby naturalne i środowisko	28
Zasoby i bioróżnorodność	30
Żywność i biogospodarka	32
Energia i klimat	35
4.4 Technologie cyfrowe w gospodarce i w społeczeństwie	38
Transformacja cyfrowa	38
Sztuczna inteligencja	39
Technologie kwantowe	40
Open Science	41
Inteligentne miasta (smart cities)	42
4.5 Społeczeństwo i jakość życia	42
Demografia i rynek pracy	42
Wyzwania społeczne i migracje	43
Znaczenie badań w naukach humanistycznych i społecznych	44
Zdrowie	45
4.6 Wspólnota, kultura, tradycja	47
4.7 Obronność i bezpieczeństwo państwa	49
5. Podsumowanie	51

4.4 Technologie cyfrowe w gospodarce i w społeczeństwie

Open Science

Wyzwaniem pozostaje rozbudowa infrastruktury repozytorium danych oraz działania edukacyjne skierowane do badaczy. Szczególnie istotne na tym etapie jest (...) kształcenie pracowników naukowych i doktorantów w zakresie OD oraz budowanie i rozwijanie repozytoriów cyfrowych.

Ważny dla wsparcia i realizacji założeń Open Science jest aktywny udział krajowych jednostek naukowych w rozwoju infrastruktury Europejskiej Chmury Otwartej Nauki (European Open Science Cloud, EOSC) na poziomie europejskim oraz krajowym. Zgodnie z założeniami Komisji Europejskiej, w perspektywie do 2027 r., system usługowo-narzędziowy EOSC odegra kluczową rolę dla otwartości, interoperacyjności i dostępności danych naukowych oraz procesów ich przetwarzania.



Polityka NCN dot. otwartego dostępu do publikacji (2020)

Zarządzenie nr 38/2020
Dyrektora Narodowego Centrum Nauki
w sprawie

ustalenia POLITYKI NARODOWEGO CENTRUM NAUKI DOTYCZĄCEJ OTWARTEGO DOSTĘPU
DO PUBLIKACJI
z dnia 27-05-2020

Na podstawie art. 34 ust. 2 pkt 4 oraz pkt 6 ustawy z dnia 30 kwietnia 2010 r. o Narodowym Centrum Nauki (t.j. Dz.U. z 2019 r. poz. 1384) w zw. z § 2 ust. 3 Regulaminu organizacyjnego Biura Narodowego Centrum Nauki i zakresu zadań Koordynatorów Dyscyplin Narodowego Centrum Nauki, którego tekst jednolity stanowi załącznik nr 1 do zarządzenia Dyrektora Narodowego Centrum Nauki nr 26/2020 z dnia 21 kwietnia 2020 r., zarządza się, co następuje:

§ 1.

Ustala się Politykę Narodowego Centrum Nauki dotyczącą otwartego dostępu do publikacji będących efektem realizacji projektów badawczych, stypendiów i staży naukowych oraz działań naukowych finansowanych lub współfinansowanych przez Narodowe Centrum Nauki.

§ 2.

Polityka, o której mowa w § 1 obowiązuje dla projektów badawczych, stypendiów naukowych, staży i działań naukowych wyłonionych do finansowania od konkursów ogłoszonych w dniu 16 czerwca 2020 r., do których umowy o realizację i finansowanie zostaną podpisane po 1 stycznia 2021 r.

§ 3.

Polityka, o której mowa w § 1 stanowi załącznik nr 1 do niniejszego zarządzenia.

§ 4.

Zarządzenie wchodzi w życie z dniem podpisania.

Dyrektor
Narodowego Centrum Nauki

Załącznik nr 1 do
Zarządzenie nr 38/2020
Dyrektora Narodowego Centrum Nauki
w sprawie
ustalenia POLITYKI NARODOWEGO
CENTRUM NAUKI DOTYCZĄCEJ
OTWARTEGO DOSTĘPU
DO PUBLIKACJI
z dnia 27-05-2020

POLITYKA NARODOWEGO CENTRUM NAUKI DOTYCZĄCA OTWARTEGO DOSTĘPU DO PUBLIKACJI

Powszechny i bezpłatny dostęp do aktualnej wiedzy stanowi jeden z fundamentów funkcjonowania i rozwoju demokratycznego społeczeństwa. Koncepcja otwartego dostępu (z ang. *Open Access*) dotyczy udostępniania w postaci cyfrowej w Internecie naukowych i wyników badań finansowanych ze środków publicznych w celu bezpłatnego ich wykorzystania przez naukowców, studentów, przedsięwzięcia społeczne. Otwarty dostęp do publikacji naukowych jest korzystny dla nauki przez większą dostępność danych przyczynia się do zwiększenia rozpoznania w danej dziedzinie naukowej.

Zgodnie z polityką Ministerstwa Nauki i Szkolnictwa Wyższego przyjętą w 2017 r. i opisaną w dokumencie pt.: *Kierunki rozwoju otwartego dostępu do publikacji naukowych w Polsce* oraz założeniami [Planu S](#), podpisanego 15 czerwca 2020 r., którego członkiem jest Narodowe Centrum Nauki, przedstawiamy politykę do publikacji naukowych, określającą zasady publikowania w otwartym dostępie (głównie dotyczy to artykułów w recenzowanych czasopiśmie recenzowanych materiałów konferencyjnych, czy też oryginalnych i będących efektem realizacji projektów badawczych finansowanych lub ze środków publicznych na naukę. Polityka ta nie obejmuje jeszcze monografii, rozdziałów monografii i recenzowanych utworów zebranych o finansowanie i realizację zostaną podpisane po 1 stycznia 2021 r.

Centrum definiuje otwarty dostęp do publikacji, tzw. *open access* jako dostęp do publikacji w internecie, ale także możliwość powielania i dowolnego wykorzystania treści przez czytelnika, zgodnie z licencjami autorskich CC-BY.

 NARODOWE CENTRUM NAUKI

Zbigniew Błocki
p.o. Dyrektora
Narodowego Centrum Nauki
ZON.070.4.2023

Kraków, 11-10-2023

Pismo w sprawie złagodzenia zapisów dotyczących "Polityki Narodowego Centrum Nauki dot. otwartego dostępu do publikacji"

Narodowe Centrum Nauki wprowadzając „Politykę dotyczącą otwartego dostępu do publikacji” (zwaną dalej Polityką) począwszy od konkursów ogłoszonych w dniu 15 czerwca 2020 r., miało na celu:

1. podkreślenie, że wydawnictwa nie mogą ograniczać praw naukowców do swobodnego dysponowania rezultatami ich prac,
2. dążenie do wzmocnienia wpływu badań realizowanych przez polskie jednostki naukowo-badawcze,
3. dążenie do zwiększenia widoczności i cytowalności naukowców afiliowanych w polskich jednostkach naukowo-badawczych,
4. dążenie do zapewnienia szerokiego międzynarodowego dostępu do wyników badań finansowanych z krajowych środków publicznych,
5. podkreślenie braku akceptacji dla praktyk związanych z tzw. *double-dipping*, czyli osiągania przez wydawnictwa dochodów z dwóch źródeł środków publicznych.

Jednocześnie, Centrum działa na rzecz zapewnienia, by efekty badań finansowanych przez Centrum, były publikowane w wydawnictwach o zasięgu międzynarodowym, spełniających najwyższe standardy publikacyjne. Wprowadzenie Polityki nie spowodowało zmian w stosowaniu tej zasady, która jest odzwierciedlona w zapisach umowy o realizację i finansowanie projektu badawczego.

Polityka NCN dot. otwartego dostępu do publikacji (2020)

OTWARTE PUBLIKACJE/ang. Open Access/OA

- 2018: NCN sygnatariuszem Planu S i członkiem-założycielem Koalicji S (ang. cOAlition S)
- 2020: Polityka NCN dot. otwartego dostępu do publikacji naukowych na podstawie Zarządzenia nr 38 i 40 Dyrektora NCN
- 2021: OA obowiązuje dla wszystkich umów od 1 stycznia 2021

OTWARTE DANE/ ang. Open Research Data

- 2019: Plan Zarządzania Danymi obowiązkowym elementem każdego projektu badawczego (zgodnie z wytycznymi i rekomendacjami Science Europe)
- 2020: Polityka NCN dot. otwartego dostępu do publikacji naukowych na podstawie Zarządzenia nr 38 i 40 Dyrektora NCN
- 2021: dane powiązane z publikacjami są obligatoryjnie udostępniane

Polityka i umowa o dofinansowanie projektu badawczego a dane badawcze (przykład NCN)

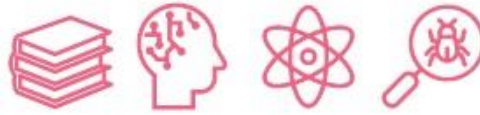
POLITYKA

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](#)).

Instrukcja Polityka NCN dot. otwartego dostępu do publikacji



Instrukcja

Polityka NCN dotycząca otwartego dostępu do publikacji

Wersja 1.0

Październik 2021



W sytuacji jeśli, po wyczerpaniu wszelkich możliwych działań po stronie autora, wydawnictwo będzie próbowało zastosować praktyki mające na celu wymuszenie przeniesienia praw autorskich, prosimy o [kontakt](#) z NCN w celu wypracowania optymalnych rozwiązań z punktu widzenia realizacji projektu, a także interesów autora / autorów.

15. Obowiązek udostępniania podstawowego zestawu danych powiązanych z publikacjami

W umowach grantowych zawieranych po 1 stycznia 2021 r.³ dane będące podstawą publikacji naukowych stanowiących efekt realizacji projektów finansowanych przez NCN powinny być rzetelnie udokumentowane w sposób spełniający zasady maszynowego lub manualnego wyszukiwania, dostępności, interoperacyjności i ponownego użycia (tzw. FAIR Data). Tam gdzie to możliwe, dane te powinny być udostępniane w repozytorium, zgodnie z warunkami licencji Creative Commons Public Domain ([licencja CC0](#)). Należy zastosować standardy cytowania danych zawarte w „[Declaration of Data Citation Principles by FORCE 11](#)” oraz na zasadach zawartych w „[TOP Guidelines](#)”. Metadane opisujące zestaw danych muszą spełniać wytyczne [OpenAIRE](#).

16. Kontakt

Dr Aneta Pazik-Aybar
Mail: aneta.pazik-aybar@ncn.gov.pl
Tel. + 48 12 341 91 78

17. Przydatne linki

1. [Polityka NCN dotycząca otwartego dostępu do publikacji](#)
2. [Journal Checker Tool](#) – narzędzie służące sprawdzaniu zgodności czasopisma z Polityką NCN dot. otwartego dostępu
3. [FAQ: Open access – najczęściej zadawane pytania](#)
4. [Plan S](#)

Wytyczne dla wnioskodawców do uzupełnienia Planu Zarządzania Danymi



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: <ul style="list-style-type: none"> - 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: <ul style="list-style-type: none"> - Jaki rodzaj, format i objętość danych planują Państwo pozyskać, wytworzyć lub ponownie udokumentować? 	Odpowiedź powinna uwzględnić rodzaj, format i zawartość każdego planowanego zbioru danych. Należy również przedstawić ich szacunkową objętość. Prosimy o uwzględnienie formatu, tj. sposobu w

Wytyczne dla wnioskodawców do uzupełnienia Planu Zarządzania Danymi

(ang. Data
Management
Plan/DMP)

	<ul style="list-style-type: none"> - W jaki sposób o Państwa danych dowiedzą się potencjalni użytkownicy? - Jak długo będą one przechowywane? - Czy istnieją jakieś ograniczenia i przeszkody uniemożliwiające ich pełne lub częściowe udostępnienie? - Czy wydawcy czasopism będą wymagać składowania danych na poparcie ustaleń publikacji? - Czy udostępnianie danych wymaga zgody uczestników badania? 	<p>danych zostanie ocenione i stwierdzone przez innych naukowców. Kiedy zostaną udostępnione dane? Prosimy o uzasadnienie długości ich przechowywania². Należy również wskazać oczekiwaną datę udostępnienia i określić, czy ulegnie ono opóźnieniu lub ograniczeniu np. w celu publikacji, ochrony własności intelektualnej, czy ubiegania się o patenty. Czy umowa o poufności będzie w stanie zapewnić odpowiednią ochronę danych poufnych?</p>
	<p>5.2 Jak będzie wyglądać selekcja danych przeznaczonych do utrwalenia i gdzie będą one długoterminowo przechowywane (np. w repozytorium danych, archiwum)? Warto rozważyć następujące kwestie:</p> <ul style="list-style-type: none"> - Które dane trzeba zachować, a które zniszczyć z przyczyn wynikających z umów, przepisów prawnych lub regulacji? - W jaki sposób zostanie podjęta decyzja o tym, które dane zachować? - Jaka będzie procedura selekcji przeznaczonych do utrwalenia danych? - Z jakiego repozytorium będą Państwo korzystać? Czy przestrzega ono zasad FAIR Data³? - Czy Państwa instytucja tworzy regularnie kopie zapasowe? 	<p>Prosimy o określenie, jak i w jakim repozytorium⁴ zostaną udostępnione dane. Prosimy wziąć pod uwagę koszty depozytu i pamięci. Jaka pojemność pamięci będzie niezbędna w okresie trwania projektu, a jaka po jego zakończeniu? Czy zdecydowali się Państwo na cyfrowe repozytorium prowadzone przez organizację non-profit?</p>
	<p>5.3 Jakie metody lub oprogramowanie umożliwiają dostęp do danych i korzystanie z</p>	<p>Metody udostępniania danych będą zależec</p>

² Dane surowe i przetworzone powinny być przechowywane przez okres odpowiedni dla danej dyscypliny i zastosowanej metodologii. W rozumieniu NCN uzasadniony okres przechowywania danych to minimum 10 lat.

³ Zasady FAIR Data określają szereg atrybutów, jakie musi posiadać opublikowany zbiór danych, aby był łatwy do odnalezienia, dostępny, interoperacyjny i nadający się do wielokrotnego użytku (patrz: Wilkinson et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship, *Scientific Data* 3, doi:10.1038/sdata.2016.18) oraz <https://www.force11.org/fairprinciples>.

⁴ Istnieje szereg międzynarodowych systemów certyfikacji określających poziom wiarygodności poszczególnych repozytoriów danych. Pośród nich najbardziej podstawowy zestaw kryteriów to Data Seal of Approval certyfikowany przez Core Trust Seal. Zaufane repozytoria cyfrowe ze znakiem jakości obejmują repozytoria z certyfikatem Data Seal of Approval, DIN-31644-, ISO-16363- lub WDS/ICSU. Przegląd takich repozytoriów można znaleźć pod niniejszym [adresem](#). Inne przydatne wykazy to: Registry of Research Data Repositories <https://www.re3data.org/>. Niektóre repozytoria, takie jak Zenodo, OpenAIRE czy CERN, pozwalają na przechowywanie zarówno danych, jak i publikacji, zapewniając również odpowiednie narzędzia dostępu. Zaleca się korzystanie w pierwszej kolejności z repozytoriów certyfikowanych lub szeroko uznanych w danej dyscyplinie, a jeżeli takie repozytoria nie istnieją, kierować się w swoim wyborze kryteriami zawartymi w dokumencie *Practical Guide to the international alignment of research data management*, https://www.scienceurope.org/wp-content/uploads/2018/12/SE_RDM_Practical_Guide_Final.pdf.



Pojęcia ogólne



NARODOWE CENTRUM NAUKI



DANE: 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 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.

W praktyce to **wszelkiego rodzaju dane** 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

- pierwotne
- wtórne

- ilościowe
- jakościowe

- obserwacyjne
- eksperymentalne
- symulacyjne



Zakres polityki NCN

„Dane powiązane z opublikowanymi artykułami...” (Polityka NCN)

„Dane będące podstawą publikacji naukowej...” (Umowa z NCN)

Zakres polityki: „wszystkie rodzaje prac z wyjątkiem: monografii, rozdziałów w pracach zbiorowych, recenzowanych utworów zebranych” = artykuły naukowe i recenzowane materiały konferencyjne.





Formaty cyfrowych danych a proces/cykl badawczy

➤ **Etap realizacji badania (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,

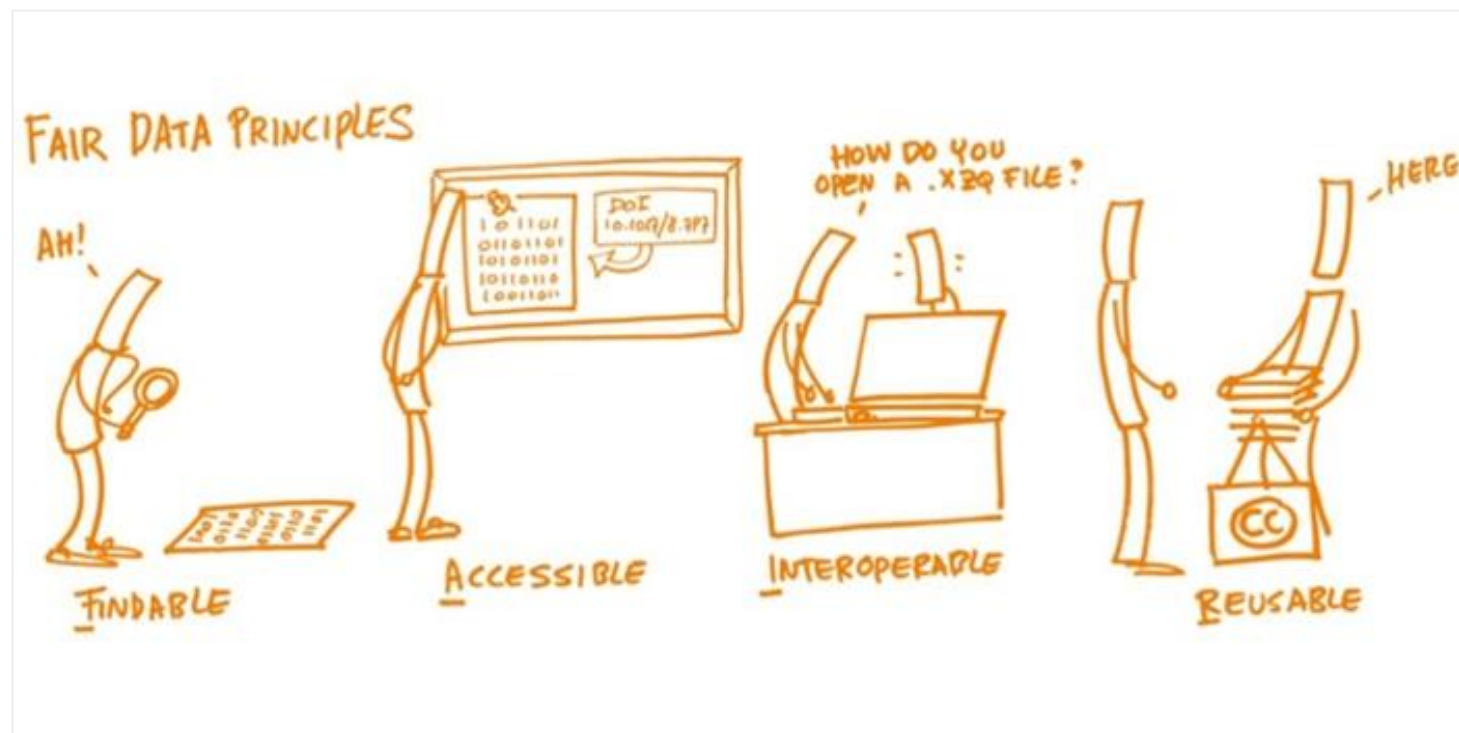
Pliki charakterystyczne dla urządzeń pomiarowych

➤ **Etap udostępnianie:**

Kluczowy dla dostępności i możliwości ich ponownego wykorzystania są formaty otwarte.



Dane otwarte a dane FAIR (ang. FAIR Data Principles)





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.

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; odczytywane przez ludzi i maszyny, precyzyjne informacje dotyczące cytowania;

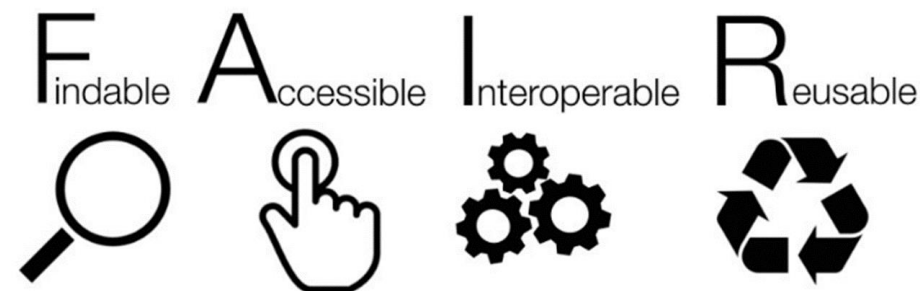
FAIR: Findable

- F1. (meta)dane mają przypisany stały i wyjątkowy identyfikator
- F2. Dane zawierają bogate metadane (opisane w R - poniżej)
- F3. metadane w oczywisty sposób zawierają

identyfikator danych, które opisują

- F4. (meta)dane są zarejestrowane lub zaindeksowane w lokalizacji w której są do odnalezienia

(Metadane i dane są zazwyczaj w oddzielnych plikach, więc jest kluczowe żeby metadane zawierały oddzielny odnośnik do danych)



FAIR: Accessible

- A1. (meta)dane są do odnalezienia przy wykorzystaniu standardowego protokołu komunikacyjnego
- A1.1 protokół ten jest darmowy, otwarty i może być uniwersalnie implementowany)
- A1.2. wybrany protokół pozwala, w razie konieczności na autentyfikację i autoryzację
- A2. metadane są dostępne, nawet jeśli dane już nie są (Np. http(s) albo ftp)

FAIR: Interoperable

- I1. (meta)dane wykorzystują formalne, dostępne, wspólne i szeroko stosowane terminy dla przedstawienia wiedzy
- I2. (meta)dane wykorzystują słowniki, które są w zgodności z zasadami FAIR
- I3. (meta)dane zawierają uściślające odwołanie do innych (meta)danych

FAIR Data: Reusable

- R1. (meta)dane są bogate opisane i zawierają liczne konkretne i istotne atrybuty danych
- R1.1 (meta)dane są udostępnione z jasną i ogólnodostępną licencją na wykorzystanie
- R1.2. (meta)dane są związane z pochodzeniem danych
- R1.3. (meta)dane są zgodne ze standardami w danej społeczności naukowej

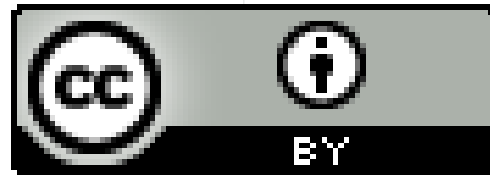
FAIRification: narzędzia online

- <https://fair-impact.eu/fair-assessment-tools>

LICENCJE dla danych badawczych



CC0: Przeniesienie do domeny publicznej.



CC-BY: Uznanie autorstwa 4.0

Licencja ta pozwala na kopiowanie, zmienianie, rozprowadzanie, przedstawianie i wykonywanie utworu jedynie pod warunkiem oznaczenia autorstwa. Jest to licencja gwarantująca najszersze swobody licencjobiorcy.



Źródło: <https://bg.uph.edu.pl>



Zasady otwierania danych badawczych

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 powinny zostać „otwarte”. Jednak mogą być sytuację, w których musimy uznać ograniczenie dostępu (ang. opt-out criteria) lub embargo (opóźniony czas otwarcia danych). Ograniczenia takie muszą być rzetelnie uzasadnione.

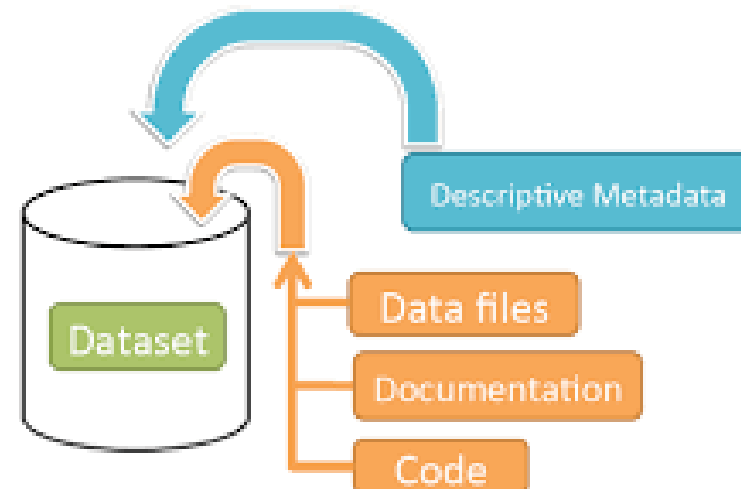
Ograniczenie dostępu w przypadku: eksploatacja przemysłowa, ograniczenia wynikające z klauzuli poufność, zasady bezpieczeństwa i dobra publicznego, prawa własności intelektualnej, w tym ochrony patentowej czy tajemnicy handlowej, dane na licencji strony trzeciej...

Anonimizacja, ochrona danych osobowych i kwestie etyczne.

Zestaw danych (ang. 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. Dokumentacja
3. Metadane

Schematic Diagram of a Dataset in Dataverse 4.0



Container for your data, documentation, and code.

Źródło: <https://repod.icm.edu.pl/>

1. Opis kontekstu badania.
2. Opis metodologii: metody pozyskiwania danych, metody analizy danych, użytego oprogramowanie i sprzęt, narzędzia badawcze.
3. Opis próby badawczej.
4. Wzory dokumentów np. zgoda uczestnika na udział w badaniach, kwestionariusza, itp..
5. Struktura i relacje między folderami: chronologicznie, geograficznie, tematycznie...
6. Sposób kontroli jakości danych.
7. Informacja o zasadach korzystania i otwartości danych.
8. Słownik/Kontrolowane słownictwo.
9. Ontologie
10. Inne...

Niezbędna do replikowania badania ...

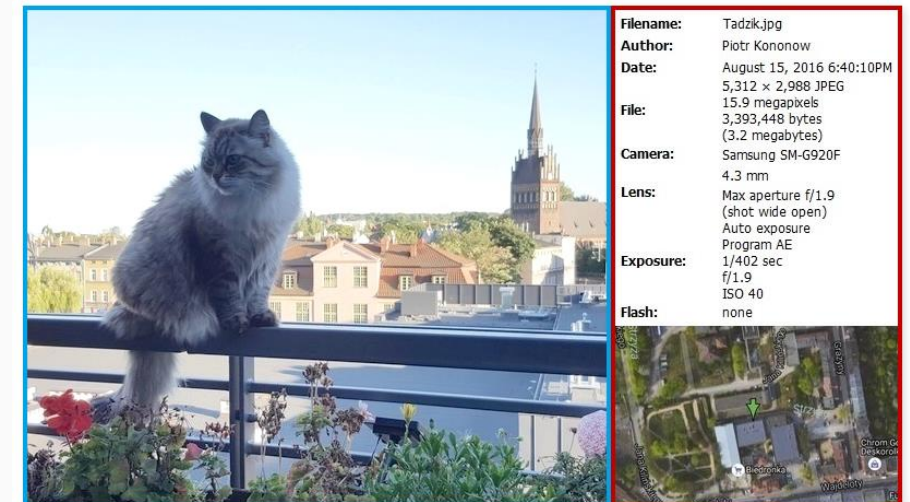
Często streszczenie dot. korzystania z danych znajduje się w pliku ReadMe.txt lub pod adresem URL.

Metadane (ang. metadata): różne standardy dla dyscyplin naukowych

Przykład: Wytyczne OpenAIRE (2013)

- tytuł zestawu danych/datasetu (M: mandatory)
- nazwisko/numer ID/ORCID (M)
- typ danych (M)
- data zebrania danych (M)
- obszar pozyskania danych (krajowy, regionalny, globalny)
- format (O: optional)
- język (R: recommended)
- lokalizacja danych
- PID (M)
- licencja (np. CC 0)
- instytucja finansująca badanie
- powiązane PID publikacji lub zestawów danych (MA: mandatory when applicable)

• sposób cytowania danych



Data

Metadata

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





Agriculture and Horticulture (Poznań University of Life Sciences)

RepOD > Poznań University of Life Sciences > Agriculture and Horticulture > Dataset of meteo data, soil and wa

Statystyki

148 Pobrań



Dataset of meteo data, soil and water properties and photos fr

Glina, Bartłomiej, 2022, "Dataset of meteo data, soil and water properties and photos from the Grojec Valley", <https://doi.org/10.18150/1T8PTC>, RepOD, V1

Opis

This dataset consists of meteorological data (temperature and rainfall) moisture, TOC, TN, DOC, DON) and water (pH, temperature, Eh7, DOC, D the Grojec Valley (52°17'00"N 18°23'31"E).

Obszar tematyczny

Nauki rolnicze; Nauki o Ziemi i środowisku

Słowo kluczowe

meteorological data, organic soils, groundwater, grasslands, Grojec Vall

Licencja

CC0 - Creative Commons Zero 1.0

Metadane cytowania

Trwały identyfikator zbioru danych

doi:10.18150/1T8PTC

Data publikacji

2022-05-18

Tytuł

Dataset of meteo data, soil and water properties and photos from the Grojec Valley

Autor

Glina, Bartłomiej – ORCID: 0000-0002-4593-649X – Department of Soil Science and Microbiology, Faculty of Agronomy, Horticulture and Bioengineering, Poznań University of Life Sciences – ROR: <https://ror.org/03tth1e03>

Osoba do kontaktu

Aby nawiązać kontakt, skorzystaj z linku "Kontakt" znajdującego się powyżej.

Glina, Bartłomiej (Department of Soil Science and Microbiology, Faculty of Agronomy, Horticulture and Bioengineering, Poznań University of Life Sciences)

Opis

This dataset consists of meteorological data (temperature and rainfall), monthly monitoring of soil (pH, temperature, moisture, TOC, TN, DOC, DON) and water (pH, temperature, Eh7, DOC, DON) properties in the period 2017-2019 within the Grojec Valley (52°17'00"N 18°23'31"E).

Obszar tematyczny

Nauki rolnicze; Nauki o Ziemi i środowisku

Słowa kluczowe

meteorological data

organic soils

groundwater

grasslands

Grojec Valley

fen peatland

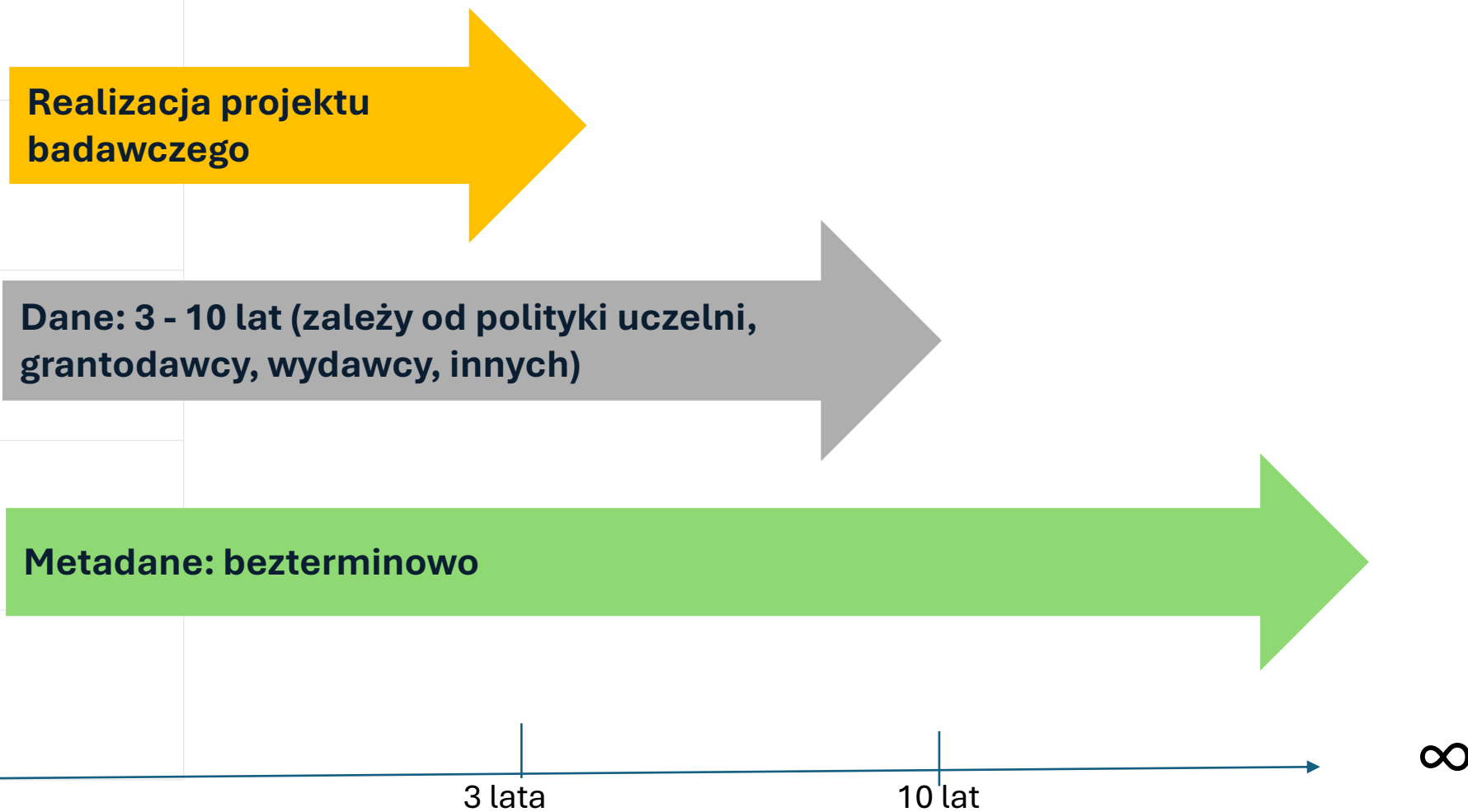
Informacje o granic

Ministry of Science and Higher Education (Poland) (MNiSW) – ROR: <https://ror.org/05dwvd537>

Deponujący

Gmerek, Krzysztof

Okres przechowywania (meta)danych



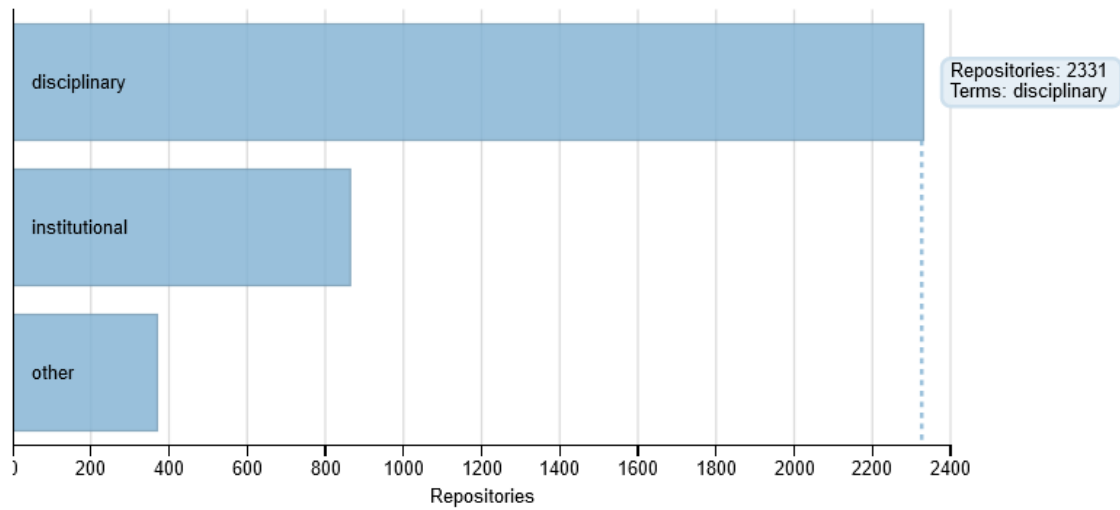
Gdzie udostępnić dane?

- Repozytoria z certyfikatem CoreTrust Seal: [CoreTrustSeal – Core Trustworthy Data Repositories](#)
- Repozytoria w Registry of Research Data Repositories: www.re3data.org

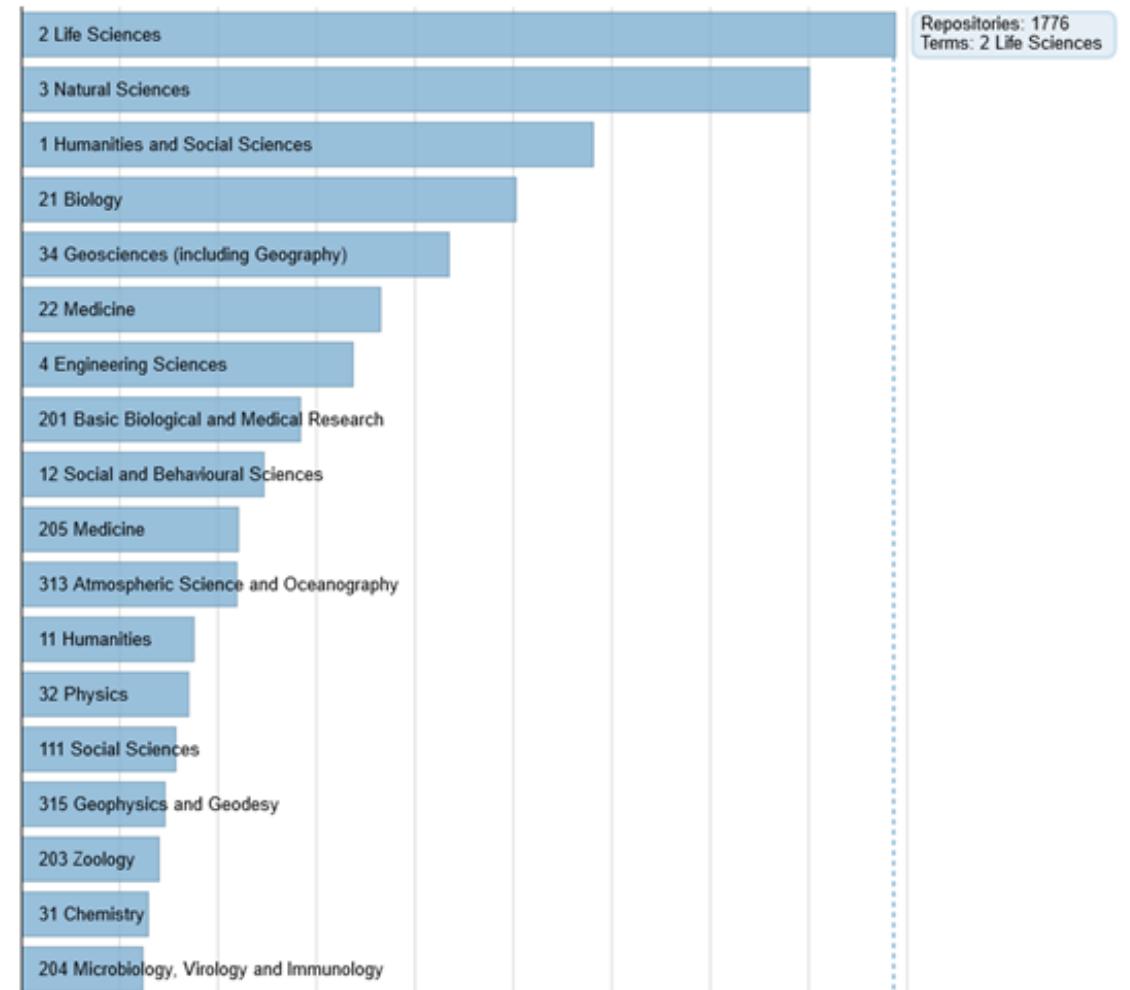




Repository types

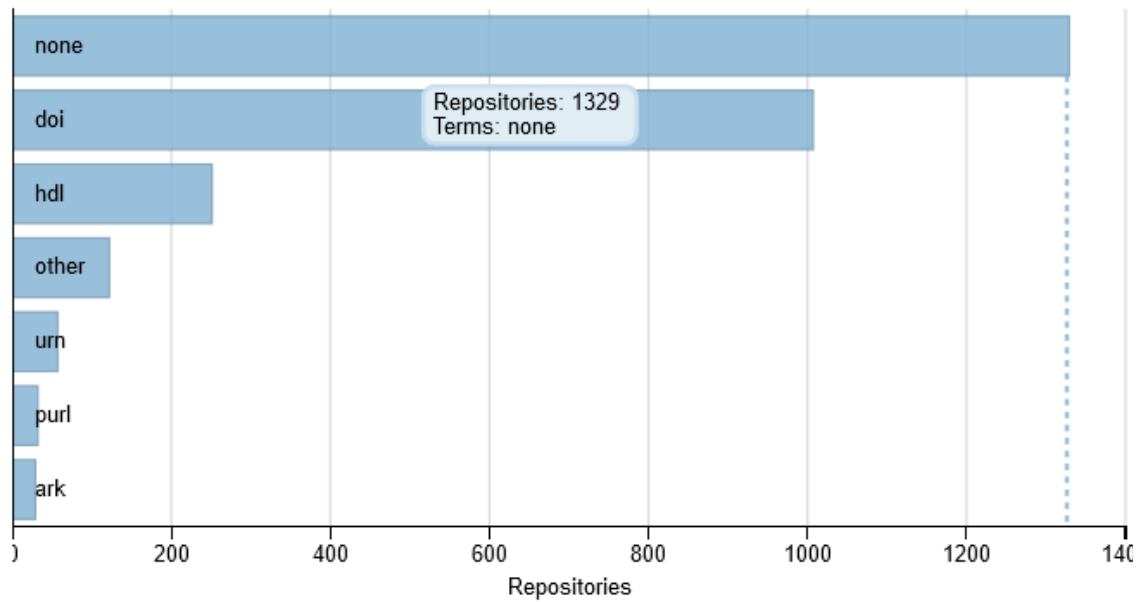


Subjects

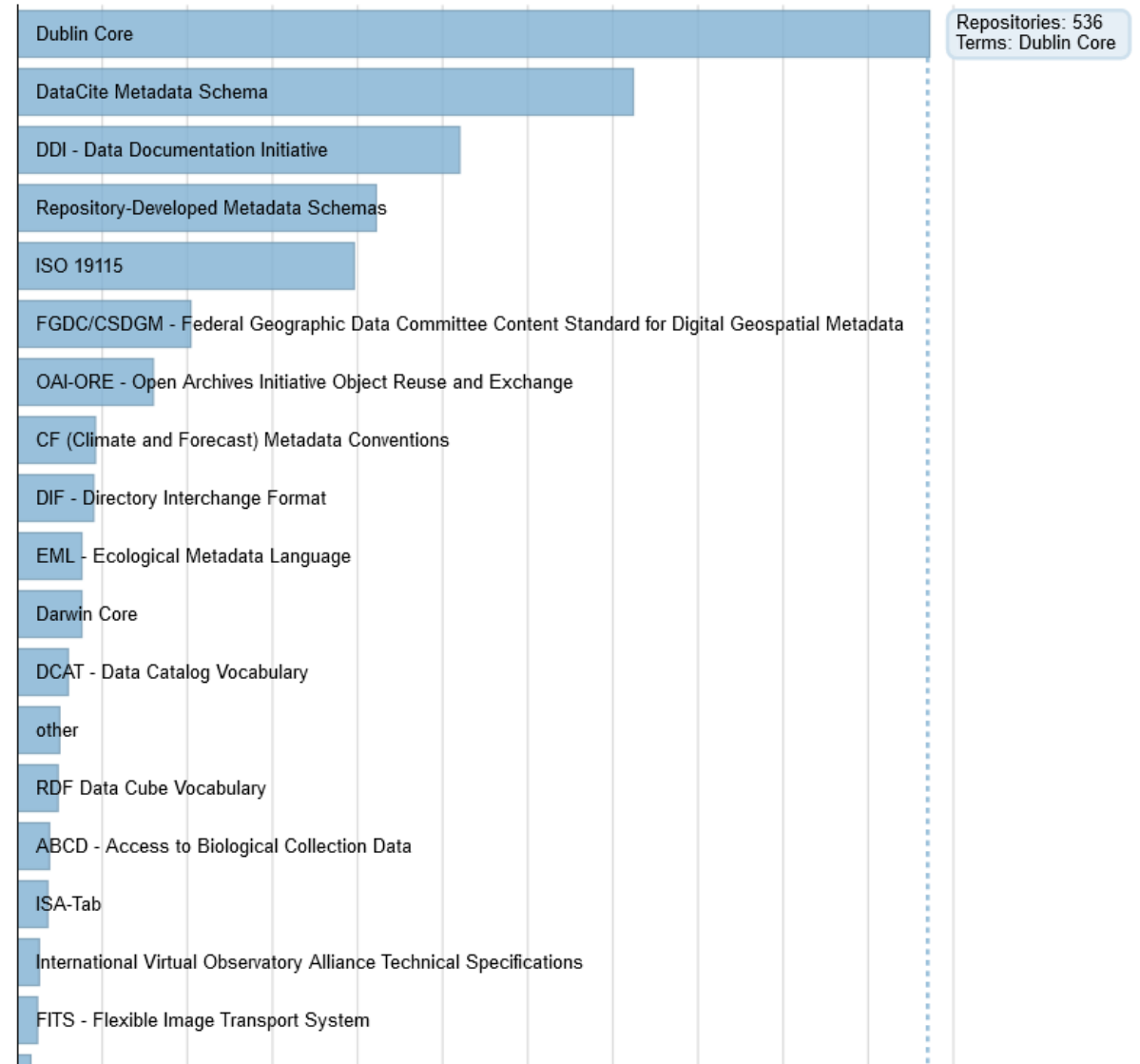




PID systems



Metadata standards



Repozytoria: kryteria wyboru i jakość zarządzania danymi badawczymi

Repository details

CLARIN-PL



General Institutions Terms Standards

Name of repository	CLARIN-PL
Additional name(s)	Language Technology Centre Centrum Technologii Jezykowych
Repository URL	https://clarin-pl.eu/dspace/
Subject(s)	Humanities and Social Sciences Linguistics Social and Behavioural Sciences Artificial Intelligence, Image and Language Processing Humanities Computer Science Computer Science, Electrical and System Engineering Engineering Sciences
Description	Polish CLARIN node – CLARIN-PL Language Technology Centre – is being built at Wrocław University of Technology. The LTC is addressed to scholars in the humanities and social sciences. Registered users are granted free access to digital language resources and advanced tools to explore them. They can also archive and share their own language data (in written, spoken, video or multimodal form).
Contact	clarin-pl@pwr.edu.pl
Content type(s)	Software applications Structured text Standard office documents Audiovisual data
Certificates and Standards	CoreTrustSeal CLARIN certificate B
Keyword(s)	digital archive electronic dictionaries language models semantic analyses speech recognition syntactic analyses

Previous 1 Next

Found 21 result(s)

Repozytorium Uniwersytetu Łódzkiego
Repozytorium Ut
Subject(s)
Content type(s)
Country
The University of Lodz Repository is an institutional repos enabling the document to be cited and indexed in scientifi

Open Science Resource Atlas 2.0 (AZON)
Atlas Zasobów Otwartej Nauki 2.0 (AZON)
Subject(s)
Content type(s)
Country
Repository "Open Science Resource Atlas 2.0" aims to in: and video recordings, photos, input and output files of va: important and unique - disabled, blind, visually impaired ar

Repository of the Jan Kochanowski Universi
Repozytorium Uniwersytetu Jana Kochanowskiego
Subject(s)
Content type(s)
Country
The purpose of the JKU Repository is to record and archi

Repozytorium IFJ PAN
The Repository of the Institute of Nuclear Physics Polish
Subject(s)
Content type(s)
Country
The repository of the Institute of Nuclear Physics Polish / the Institute.


Open Research Data at TUL
RDB.open
Subject(s)
Content type(s)
Country
TUL Open Research Data Repository (RDB.open) is a ser indicated by the funder or the scientists. The RDB.open is



Repozytoria: kryteria wyboru i jakość zarządzania danymi badawczymi

https://mostwiedzy.pl/pl/open-research-data/catalog

dla naukowca dla biznesu english zaloguj

 **MOST WIEDZY**

Strona główna > Dane Badawcze > Katalog

Katalog Danych Badawczych

Pomożemy Ci w:

- stworzeniu metadanych dla datasetu
- wyborze odpowiednich licencji
- przygotowaniu danych do udostępnienia
- dodaniu nowej wersji datasetu
- zgrupowaniu datasetów w kolekcji

Filtry wszystkich: 24171

Rok publikacji

- 2024
- 2023
- 2022
- 2021
- 2020
- 2019
- 2018
- 2017
- 2016
- 2015
- 2014
- 2013
- 2012
- 2008
- 2007
- 2006

Dziedzina

- Dziedzina nauk ścisłych i przyrodniczych
- Dziedzina nauk humanistycznych
- Dziedzina nauk inżynieryjno-technicznych
- Dziedzina nauk rolniczych
- Dziedzina nauk społecznych
- Dziedzina nauk medycznych i nauk o zdrowiu

wyświetlamy 1000 najlepszych wyników

Pokazuj dane z Wirtualnego Mikroskopu 1 2 3 ... 48 49 50 →

[Polish bibliological journals - publishing policy data set](#)
dane badawcze embargo J. Błasiok, K. Kokot-Kanikula
The file contains the results of an analysis of the publishing policies of Polish bibliological journals conducted by librarians of the Library of the Gdansk University of Technology. Among the elements studied were Open Access status, Creative Commons license type and self-archiving practices. The survey was conducted from 2018 to 2023 on the basis...

[Phylogenetic analysis of Oncidieae subtribe - matK plastid region](#)
dane badawcze open access P. Krajewski
The dataset contains alignment and consensus matrix files, and results of phylogenetic analysis of plastid matK region from 186 orchid species considered to belong to the subtribe Oncidieae. Sequences were assessed from NCBI GeneBank (list of all records with identifiers in the separate file)

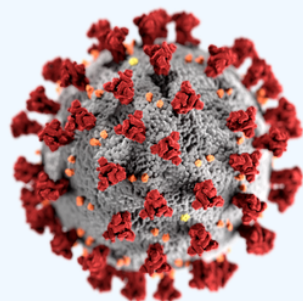
[Phylogenetic analysis of Oncidieae subtribe - ITS1-5,8S-ITS2](#)
dane badawcze open access P. Krajewski
The dataset contains alignment and consensus matrix files, and results of phylogenetic analysis of ITS1-5,8S-ITS2 region from 186 orchid species considered to belong to the subtribe Oncidieae. Sequences were assessed from NCBI GeneBank (list of all records with identifiers in the separate file)

[The effect of pH change of ketoprofen solutions on the germination of Sorghum bicolor \(sorghum\) seeds](#)
dane badawcze open access M. Wiczerzak, B. Kudlak, J. Namieśnik - seria: *Impact of selected drugs and their binary mixtures on the germination of Sorghum bicolor (sorghum) seeds*
In these studies, an attempt was also made to determine whether (and, if so, to what extent) the change in pH affects the toxicity of ketoprofen will affect the germination of Sorghum bicolor (sorghum) seeds. As reported by Kudlak B, Wiczerzak M, Yotova G, Tsakovski S, Simeonov V, Namieśnik J (2016) Environmental risk assessment of Polish wastewater...

[Determination of the effective concentration of ketoprofen against the germination of Sorghum bicolor \(sorghum\) seeds](#)
dane badawcze open access M. Wiczerzak, B. Kudlak, J. Namieśnik - seria: *Impact of selected drugs and their binary mixtures on the germination of Sorghum bicolor (sorghum) seeds*
Research data includes an attempt to determine the effective concentration of ketoprofen that inhibits germination of Sorghum bicolor (sorghum) seeds.

[Elgold partial: Automotive blogs](#)
dane badawcze open access S. Olewniczak, J. Szymański - seria: *Elgold - partial*
The dataset contains 34 English texts scrapped from automotive blogs. In each text, the named entities are marked. Each name entity is linked to the

Featured communities

[Need help uploading? Contact us](#)

Coronavirus Disease Research Community - COVID-19

Browse

New upload

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 outcome 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 20, 2022 (v2)

Software

Open Access

View

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

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

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

Need help?

[Contact us](#)

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

We can help with:



Kryteria wyboru repozytorium

- wytyczne (kryteria formalne i merytoryczne) agencji finansującej badania, wydawcy, uczelni, np. : zastosowanie zasad FAIR, przypisanie numeru PID; zastosowanie metadanych, brak embargo, inne....
- 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, czy Google Dataset Search.

Otwartość danych o ocena dorobku naukowego oraz ewaluacja naukowa jednostki...

Publication Record	
<i>5-10 most important papers published or accepted for publication (letter of acceptance required) in the proposal submission year or over the period of 10 years prior to the proposal submission year. 1 to 3 most important papers presented above must be attached as PDF files (no more than 10 MB in size). For monographs, attach a file with at least the title page, copyright page, contents and a selection of fragments containing the author's chief theses.</i>	
<i>For research in art, 2-10 most important artistic achievements of international significance or works significant for the Polish culture and 1-10 most important papers published or accepted for publication (letter of acceptance required) in the proposal submission year or over the period of 10 years prior to the proposal submission year. Annex 1-3 publications (if any) as PDF files.</i>	
<i>The period of 10 years shall be extended by any career breaks specified in the form. Please note that the extensions do not apply to the statutory definition of a well-established and outstanding researcher.</i>	
Authors, title in the original publication language [with translation into English], journal, monograph title in the original publication language, year of publication, publisher, place of publication, volume/issue, pages, ISBN/ISSN, editors, open access (YES/NO), total number of citations without self-citations, publication status (published/accepted for publication), DOI Number of the publication, have the underlying data related to the publication been shared? (YES/NO), DOI (or another Persistent Identifier) of the data related to the publication	
PDF file	1-3 files
Letter of acceptance	if publication status:accepted for publication PDF file

Zespół ds. Otwartej Nauki
Narodowe Centrum Nauki

Kontakt:

otwarta.nauka@ncn.gov.pl