



RepOD

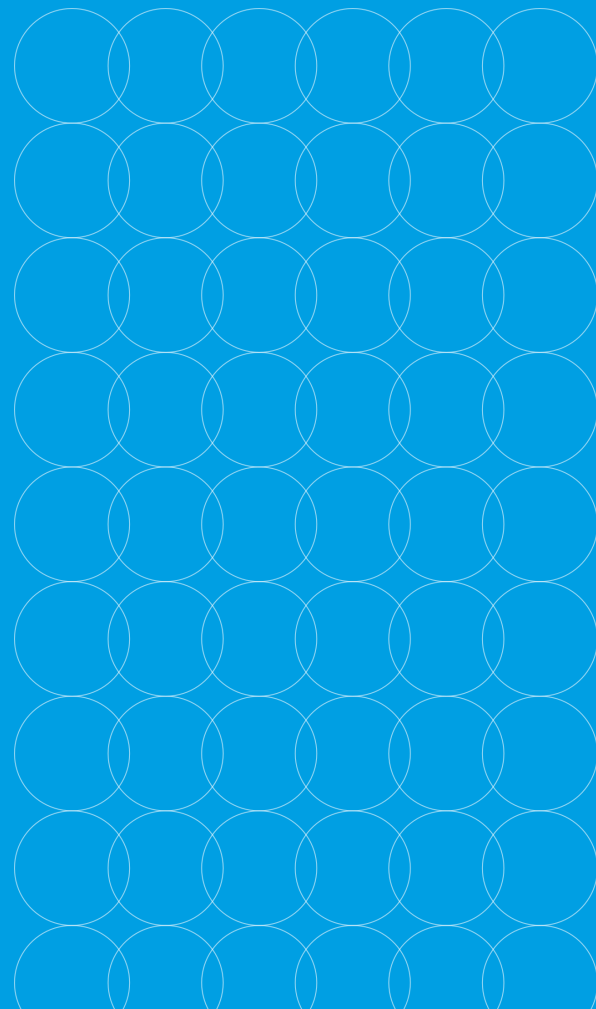
Repozytorium Otwartych Danych

Repozytorium Otwartych Danych Badawczych RepOD

reporod.icm.edu.pl



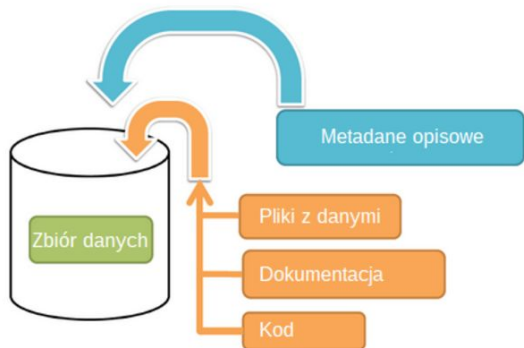
UNIwersytet
Warszawski



Dane badawcze

→ To, co potrzebne do uzyskania wyników badań naukowych i ich oceny, na przykład:

- wyniki pomiarów
- notatki z obserwacji, wyniki kwerend
- komentarze do analizowanych tekstów, streszczenia, wypisy
- bazy motywów literackich, wykresy, mapy, analizy cyfrowe
- kwestionariusze, wyniki ankiet
- nagrania audio i wideo, grafiki
- modele matematyczne
- oprogramowanie
- opisy metodologiczne...



repositor.ict.edu.pl/guides/pl/1.0.0-SNAPSHOT/user/dataset-management.html

Dlaczego należy (i warto) właściwie zarządzać danymi badawczymi i udostępniać je w sposób otwarty?

→ [Poradnik, s. 5–8](#)

Zbiór danych – przykład

Helman-Waźny, Agnieszka; Ramble, Charles, 2024, "Drangsong 239: La 'dos[mdos] gyer la stang[btang] ste 'di skad do/", <https://doi.org/10.58132/F3VQ3L>, Dane Badawcze UW, V1

Drangsong Manuscripts

Dane Badawcze UW > Wydział Dziennikarstwa, Informacji i Bibliologii > Drangsong Manuscripts > Drangsong 239: La 'dos[mdos] gyer la stang[btang] ste 'di skad do/

Statystyki 0 Pobrań

Drangsong 239: La 'dos[mdos] gyer la stang[btang] ste 'di skad do/

Helman-Waźny, Agnieszka; Ramble, Charles, 2024, "Drangsong 239: La 'dos[mdos] gyer la stang[btang] ste 'di skad do/", <https://doi.org/10.58132/F3VQ3L>, Dane Badawcze UW, V1

Opis

Drangsong 239: La 'dos[mdos] gyer la stang[btang] ste 'di skad do/.

This dataset contains photographs of the manuscript pages with handwritten text from Mustang, Nepal.

The manuscript Drangsong 239 contains praying to a *btsan* deity who resides in the sky for assistance in fulfilling wishes.

It is composed of 2 folios.

File naming convention within Drangsong_239.zip archive:

Drangsong_239_005r.jpg

Drangsong_239 → collection name and number

001 → page number

r / v → recto / verso

METADATA

DATASET FILES

The dataset contains files:

- Drangsong_239.zip (master scans of the manuscript pages with color checker),
- Drangsong_239.pdf (containing presentation images of the manuscript pages),
- Drangsong_239_005r.png (manuscript page sample),
- Drangsong_239_backlites.zip (backlit samples),
- Drangsong_239_XRF_analysis.png (X-ray fluorescence spectrometry).

DESCRIPTION OF THE MANUSCRIPT

Detailed information on writing, materials, format, and layout is provided in a Readme.txt file attached to this dataset. (2024)

Obszar tematyczny ?

Sztuka i nauki humanistyczne

Słowo kluczowe ?

Tibetan Nepal Mustang, Codicology Nepal Mustang, Nepal District History, Bön (Tibetan religion), Bön literature, Buddhist literature, Tibetan Manuscripts, Material analysis, Archaeometry

Licencja ?

Różne licencje dla poszczególnych plików

Pliki Metadane Księga gości Wersje

Szukaj w zbiorze danych... Szukaj

1 do 6 z 6 Plików

Pobierz wszystkie -

Readme.txt	Plain Text - 1,1 KB - 2024-12-05 - 0 Pobrań MDS: 8d4ebf826a660540bb50925fc17346a9 Licencja: CC BY - Creative Commons Uznanie Autorstwa 4.0 Detailed information on writing, materials, format, and layout.	Podgląd Pobierz
Drangsong_239.zip	Archiwum ZIP - 28,3 MB - 2024-12-05 - 0 Pobrań MDS: 0d2ed74b1e32c562fd37ce3375eb5129 Licencja: CC0 - Creative Commons Zero 1.0 Master scans of the manuscript pages with color checker.	Pobierz

Metadane

Element dokumentacji

Plik z danymi

Pliki – przykłady

A		B		C	D	E	F	
1	year book published	author		title	fiction/non-fiction (F/NF)	animal species	additional notes	
2	1894	Rudyard Kipling		The Jungle Book	F	wolves, panther, bear		
3	1900	Abbie Farwell Brown		"Saint Keneth of the Gulls"	F	gulls, doe		
4	1900	Abbie Farwell Brown		"The Wolf Mother of Saint Ailbe"	F	wolves		
5							a series of 24 books, the final one published in 1966 (some of them were either co-written by Burroughs, or officially authorized by his estate)	
6	1912	Edgar Rice Burroughs		Tarzan of the Apes	F	gorillas		
7	1919	Olaf Baker		Shasta of the Wolves	F	wolves	Native American baby raised by wolves	
8	1926	George Bird Grinnell		"Black Wolf and His Fathers"	F	wolves	the folk tale of the Cheyenne	
9	1942	Robert M. Zingg		Wolf Children and Feral Man	NF	wolves	a highly controversial book, proven to be unreliable	
10	1949	Rutherford G. Montgomery		Kidgee House	F			
11	1958	Peter Matthiessen		"The Wolves of Aguila"	F	wolves		
12	1959	Jean Craighead George		My Side of the Mountain	F	falcon	not exactly a feral child, a 12-year-old boy chooses a solitary life	
13	1959	David Severn		Foxy-Boy	F	foxes		
14	1960	Scott O'Dell		The Island of the Blue Dolphins	F	dolphins	inspired by a true story	
15	1967	Roy Meyers		Dolphin Boy	F (fantasy)	dolphins	a series of books, Dolphin Trilogy	
16	1972	Jean Craighead George		Julie of the Wolves	F	wolves		
17	1974	Edgar Rice Burroughs, Gene Wolf, Philip José Farmer		Mother Was a Lovely Beast: A Feral Man Anthology	F (some claim to be NF)	various		
18	1994	Cormac McCarthy						
19	1995	Douglas Keith Candland						

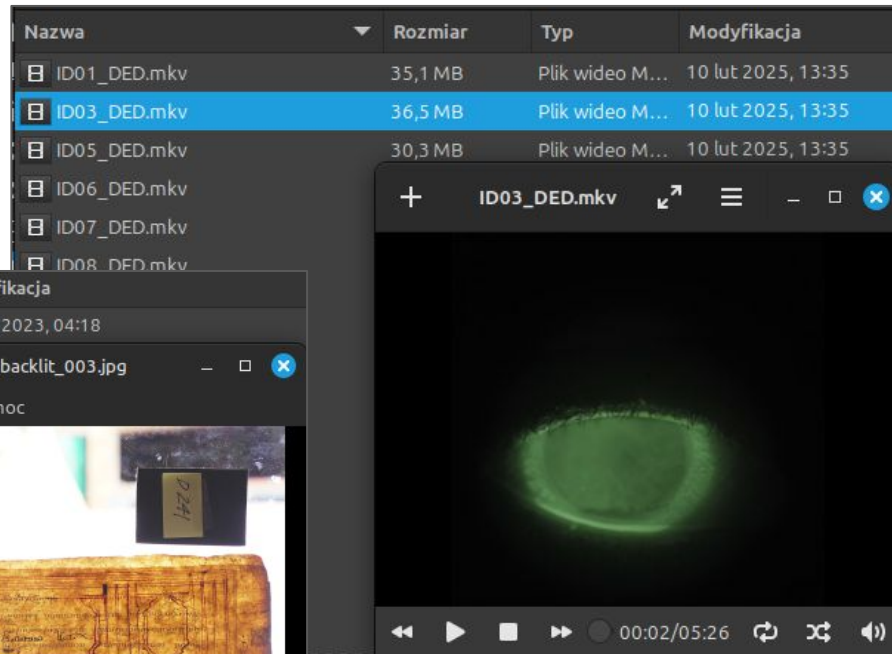
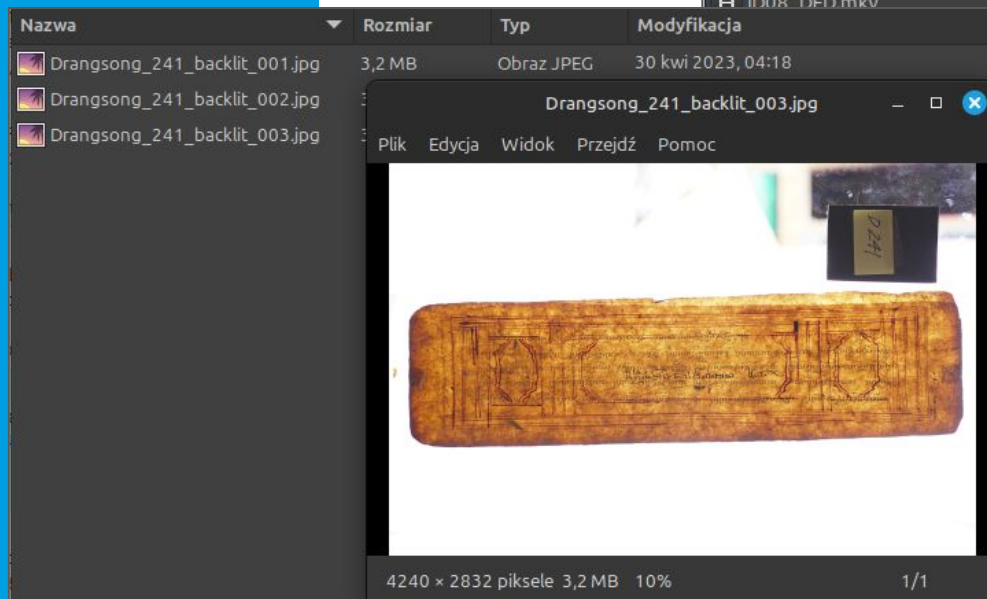
	ID#	Gender	Teachexp	DegreeProfProm	PCC1Cod1	TSC1Cod1	RAPC1Cod1	BMC1Cod1	PC1Cod1	NC*C1Cod1
1	1.0	1.0	2.0	1.0	5.0	4.0	4.0	6.0	7.0	1.0
2	2.0	2.0	11.0	2.0	4.0	3.0	1.0	5.0	5.0	1.0
3	3.0	1.0	7.0	2.0	6.0	6.0	5.0	7.0	7.0	1.0
4	4.0	1.0	18.0	3.0	4.0	5.0	2.0	7.0	6.0	1.0
5	5.0	2.0	35.0	3.0	4.0	4.0	2.0	6.0	2.0	1.0
6	6.0	1.0	39.0	3.0	2.0	1.0	3.0	4.0	5.0	3.0
7	7.0	1.0	21.0	3.0	2.0	2.0	1.0	6.0	5.0	1.0
8	8.0	1.0	25.0	3.0	4.0	5.0	2.0	6.0	6.0	1.0
9	9.0	3.0	35.0	3.0	4.0	4.0	1.0	6.0	2.0	1.0
10	10.0	2.0	3.0	1.0	5.0	4.0	1.0	7.0	7.0	1.0
11	11.0	2.0	4.0	1.0	6.0	5.0	3.0	3.0	2.0	1.0
12	12.0	1.0	1.0	1.0	4.0	4.0	2.0	7.0	7.0	1.0
13	13.0	2.0	15.0	3.0	4.0	5.0	1.0	6.0	7.0	1.0
14	14.0	2.0	17.0	3.0	3.0	4.0	1.0	7.0	5.0	1.0
15	15.0	2.0	4.0	1.0	4.0	4.0	4.0	7.0	5.0	1.0

throughout the ages contrasted with male"

Wilde, Julia, 2024, "Figurations of interspecies harmony: Feral children fiction and non-fiction originally written in English, from 1894 onwards", <https://doi.org/10.58132/BJGWC3>, Dane Badawcze UW, V1

Firek, Wiesław, 2024, "Students' Emotional Support in Physical Education", <https://doi.org/10.18150/PJO5NF>, RepOD, V1

Pliki – przykłady



Miejsce dla danych

„Zobacz supplementary materials”.



Repozytorium
danych



FAIR

Czym są zasady FAIR? → [Poradnik, s. 12](#)

„Dane udostępniane na
żądanie. Proszę wysłać
e-mail”.

Dropbox, Google Drive...

www.wordpress.com/wydzial/mojprojekt/dane



RepOD
Repozytorium Otwartych Danych

Na podstawie: Harvard Longwood Medical Area Research Data Management Working Group. "Data Discussions: Let's Find a Data Repository". osf.io/tmpxn



UNIWERSYTET
WARSZAWSKI



Rodzaje repozytoriów danych badawczych

- Repozytoria specjalistyczne (np. [Protein Data Bank](#))
- Repozytoria dziedzinowe (np. [Repozytorium Danych Społecznych](#))
- Repozytoria instytucjonalne
- Repozytoria ogólnego przeznaczenia

re3data.org

Rodzaje repozytoriów danych badawczych

- Repozytoria specjalistyczne
- Repozytoria dziedzinowe
- Repozytoria instytucjonalne (np. [Dane Badawcze UW](#), RepOD)
- Repozytoria ogólnego przeznaczenia (np. RepOD)

Korzystanie z repozytorium

RepOD umożliwia bezpłatne:

- deponowanie i przechowywanie danych badawczych,
- publiczne udostępnianie danych badawczych,
- wyszukiwanie, przeglądanie i pobieranie danych badawczych,
- prowadzenie [kolekcji instytucjonalnych](#):
 - dla każdej instytucji,
 - na mocy krótkiego porozumienia,
 - na elastycznych zasadach,
- działanie w zgodzie z zasadami FAIR i wymogami instytucji finansujących oraz wydawców.

Zgodność z wymogami

amt.coretrustseal.org/certificates

- Bezterminowe przechowywanie danych.
- Przechowywanie kopii danych w innej lokalizacji.
- Weryfikowanie integralności plików (MD5).
- Zgodność z wymogami WCAG.
- Rozbudowany formularz metadanych, różne schematy.
- Automatyczne przypisywanie identyfikatora DOI każdemu zbiorowi danych.
- Metadane udostępniane na licencji CC0.
- Możliwość wyboru licencji dla każdego pliku (w tym licencji Creative Commons).
- Możliwość tworzenia kolejnych wersji zbioru.
- Weryfikacja zbiorów.
-  Certyfikat Core Trust Seal.

Dostępność

repod.icm.edu.pl

- Zasoby otwarte, dostępne dla niezalogowanych użytkowników – nie trzeba podawać danych osobowych „w zamian” za dostęp do danych badawczych.
- Logowanie dla deponentów.
- Lista kolekcji widoczna w karuzeli.
- Instrukcje, FAQ → *O repozytorium*.

Jak deponować?

- Krok 1: wypełnij formularz metadanych — pola obowiązkowe (tylko sześć) i dodatkowe; im więcej, tym lepiej.

Metadane cytowania

Tytuł* ?

Tłumaczenie tytułu ?

Treść ?

Język tłumaczenia ? Wybierz... +

Autor* ?

Nazwisko i imię autora* ?

System identyfikacji ? Wybierz... +

Identyfikator ?

Afiliacja ?

ROR instytucji afiliującej ?

Osoba do kontaktu* ?

Nazwisko i imię ?

Afiliacja ?

E-mail* ?

Opis* ?

Treść* ?

Jak deponować?

- Krok 2: dodaj pliki — dane, dokumentację (np. plik *readme* – [więcej](#)) i wybierz licencję.
 - Zalecane otwarte formaty plików → [lista](#)
 - Maksymalna wielkość pojedynczego pliku: 5 GB; transzy: 20 GB; zbioru: ~2 TB.
 - Automatyczna analiza plików tabelarycznych.

Pliki

Więcej informacji na temat wspieranych formatów plików można znaleźć w "Poradniku użytkownika".

Prześlij za pomocą HTTP przez swoją przeglądarkę ^

Wybierz pliki lub przeciągnij je i upuść do widżetu przesyłania. Limit wielkości pojedynczego pliku wynosi 5,0 GB. W jednej transzy można przesłać pliki o łącznej wielkości 20,0 GB. W skład pojedynczej transzy wchodzi pliki, które zostały przesłane, ale nie zostały jeszcze zapisane.

+ Wybierz pliki do dodania

Przeciągnij i upuść pliki w to miejsce

Jak deponować?

- Krok 3: Zapisz zbiór, **prześlij go do weryfikacji** i czekaj na informację zwrotną.



Analiza plików tabelarycznych

- Automatyczna analiza plików tabelarycznych

→ Celem analizy jest wydobycie zawartości plików i zarchiwizowanie jej w neutralnym dla aplikacji, łatwym do odczytania formacie. Pomyślna analiza upewnia nas, że dane spełniają standardy FAIR. Zwłaszcza R – *reusability*.

→ Komercyjne aplikacje używają do kodowania plików swoich formatów własnościowych. Nie są to formaty właściwe do długotrwałej archiwizacji.

→ Oprogramowanie repozytorium pozwala przechowywać surowe dane pobrane z takich plików w postaci plików tekstowych rozdzielanych tabulatorami. Informacje o metadanych opisujących zawartość pliku są przechowywane oddzielnie, w relacyjnej bazie danych. Można je wyeksportować w postaci tekstowych plików XML przy użyciu znormalizowanego, otwartego formatu DDI Codebook. [więcej](#)

→ Analiza plików CSV/TSV: [więcej](#)

→ Analiza plików XLSX: [więcej](#)

Analiza plików – przykład

Aby zapewnić pomyślną analizę z trafnymi wynikami, plik danych tabelarycznych powinien:

- składać się z jednego tylko arkusza danych,
- zawierać tylko jedną tabelę w układzie pionowym, tj. każda użyta kolumna powinna zawierać jedną zmienną, a poszczególne wiersze – wartości zmiennych dla jednego przypadku,
- nie zawierać komórek scalonych (program analizujący dokona próby rozdzielenia scalonych komórek, ale zrobi to źle),
- nie zawierać pustych kolumn ani wierszy, w tym „marginesów” górnych ani lewych: tabela powinna zaczynać się od kolumny A i od wiersza 1,
- w wierszu 1. zawierać wyłącznie kompletne nazwy zmiennych,
- od wiersza 2. aż do końca zawierać tylko wartości zmiennych,
- nie zawierać komentarzy, objaśnień stosowanych jednostek i skrótów, opisów warunków pomiaru itp. – te powinny się znaleźć w opisie pliku, opisie zbioru i/lub w specjalnie przygotowanym pliku readme.txt ([wzór](#)).

Pliki niespełniające tych warunków nie będą odrzucane w procesie weryfikacji, ale mogą go znacząco wydłużyć. Pliki, które nie przeszły analizy, są trudniejsze w odbiorze i trudniejsze do ponownego wykorzystania//

[więcej](#)

:-(

	A	B	C	D	E	F	G	H	I
1	Figure 1								
2	C_data_raw								
3									
4			Region	Rep	Item	Units	UnitCost		Total
5		OrderDate	East	Jones	Pencil	95	1,99		189,05
6		2019-06-02	Central	Kiyell	Binder	50	19,99		999,5
7		2019-06-03	Central	Jardine	"	36	4,99		179,64
8		2019-06-04	Centralny	Gill	Pen	27	19,99		539,73
9									
10		2019			Pencil	56	2,99		167,44
11		2019-06-06	East	Jones	Binder	60	4,99		299,4
12		2019-06-07	Central	Andrews	Pencil		1,99		149,25

:-)

	A	B	C	D	E	F	G
1	OrderDate	Region	Rep	Item	Units	UnitCost	Total
2	2019-06-01	East	Jones	Pencil	95	1,99	189,05
3	2019-06-02	Central	Kiyell	Binder	50	19,99	999,5
4	2019-06-03	Central	Jardine	Pencil	36	4,99	179,64
5	2019-06-04	Central	Gill	Pen	27	19,99	539,73
6	2019-06-05	West	Sorvino	Pencil	56	2,99	167,44
7	2019-06-06	East	Jones	Binder	60	4,99	299,4
8	2019-06-07	Central	Andrews	Pencil	75	1,99	149,25
9	2019-06-08	Central	Jardine	Pencil	90	4,99	449,1
10	2019-06-09	West	Thompson	Pencil	32	1,99	63,68

Nazwa pliku Fig_1_C_podf.xlsx

MS Excel (XLSX)
MD5: ad66693511062c0336384bb845f7cd90;

Opis Dodaj opis pliku

Analiza ☒ Przeanalizuj plik

Analiza plików – przykład



Fig_1_C_pdf.tab

Dane tabelaryczne - Format oryginalny: MS Excel (XLSX) - 24,0 KB - 2025-05-15 - 0

Pobrań

Licencja: CC BY - Creative Commons Uznanie Autorstwa 4.0

22 Zmienne, 160 Obserwacje - UNF:6:yKkEp1rtqbryZDSu7503uw==

Data used to generate subpanel C of Fig. 1

 Podgląd

 Pobierz ▾

Oryginalny format pliku (MS Excel (XLSX))
Format Tab (dane oddzielone tabulatorem)
Format RData
Metadane zmiennych
Cytowanie pliku danych

Podgląd

Metadane

Wersje

Otwórz w nowym oknie... 

	Time	Rat1SALinf	Rat2SALinf	Rat3SALinf	Rat4SALinf	Rat5SALinf	Rat6SALinf	Rat7SALinf
1	0.5	6.4E-5	8.2E-5	2.8E-5	2.4E-5	3.2E-5	1.8E-5	2.6E-5
2	1.0	6.8E-5	8.5E-5	3.9E-5	1.5E-5	4.5E-5	1.9E-5	3.2E-5
3	1.5	8.4E-5	1.1E-4	3.1E-5	4.1E-5	4.2E-5	2.5E-5	6.0E-5
4	2.0	6.6E-5	8.6E-5	2.5E-5	1.7E-5	4.3E-5	2.8E-5	2.7E-5
5	2.5	7.9E-5	8.8E-5	3.8E-5	1.8E-5	5.2E-5	2.1E-5	2.5E-5
6	3.0	6.0E-5	8.1E-5	4.3E-5	1.9E-5	5.3E-5	3.6E-5	4.3E-5
7	3.5	5.0E-5	1.6E-4	1.8E-5	1.8E-5	4.4E-5	2.5E-5	2.6E-5
8	4.0	6.6E-5	7.2E-5	3.6E-5	3.4E-5	6.2E-5	2.2E-5	3.1E-5
9	4.5	9.7E-5	1.1E-4	2.7E-5	2.0E-5	4.3E-5	2.0E-5	2.5E-5
10	5.0	5.6E-5	7.0E-5	1.8E-5	1.8E-5	4.1E-5	2.3E-5	3.1E-5
11	5.5	6.3E-5	1.0E-4	3.5E-5	4.9E-5	7.4E-5	2.6E-5	3.2E-5
12	6.0	5.9E-5	9.1E-5	2.5E-5	3.1E-5	6.4E-5	1.1E-4	3.2E-5
13	6.5	8.8E-5	6.7E-5	2.2E-5	2.0E-5	4.2E-5	2.3E-5	4.7E-5



RepOD
Repozytorium Otwartych Danych



UNIwersytet
WARSAWSKI



Co jeszcze?

- Możliwość udostępnienia nieopublikowanego zbioru danych wybranym osobom ([prywatny link](#)).
- Możliwość ograniczenia dostępu do plików (embargo).
- Możliwość śledzenia liczby pobrań i cytowań.



Statystyki

13 Pobrań

1 Cytowanie

Wersjonowanie

- Wersjonowanie:
 - Raz opublikowane dane nie mogą zostać usunięte.
 - Błędy? Uzupełnienia? → Tworzymy nową wersję zbioru: należy wprowadzić zmiany i ponownie przesłać zbiór do weryfikacji.
 - Poprzednie wersje zbioru pozostają dostępne (mogły zostać pobrane i/lub zacytowane).
 - Numer wersji jest elementem cytowania.
 - Identyfikator DOI nie zmienia się.

Pliki	Metadane	Księga gości	Wersje
	Zbiór danych	Podsumowanie	
☐	2.0	Pliki (Zmienione licencje/warunki użycia: 1 plik); Wyświetl szczegóły	
☐	1.1	Metadane cytowania: Opis (1 Zmieniony); Autor (5 Dodany); Wyświetl szczegóły	
☐	1.0	To jest pierwsza opublikowana wersja	

Licencje CC – Creative Commons



Uznanie autorstwa. Wolno kopiować, rozprowadzać, przedstawiać i wykonywać objęty prawem autorskim utwór oraz opracowane na jego podstawie utwory zależne pod warunkiem, że zostanie przywołane nazwisko autora pierwowzoru.



Użycie niekomercyjne. Wolno kopiować, rozprowadzać, przedstawiać i wykonywać objęty prawem autorskim utwór oraz opracowane na jego podstawie utwory zależne jedynie do celów niekomercyjnych.



Na tych samych warunkach. Wolno rozprowadzać utwory zależne jedynie na licencji identycznej do tej, na jakiej udostępniono utwór oryginalny.



Bez utworów zależnych. Wolno kopiować, rozprowadzać, przedstawiać i wykonywać utwór jedynie w jego oryginalnej postaci – tworzenie utworów zależnych nie jest dozwolone.



Wszystko można (w granicach prawa)...



...ale proszę wskazać autora.



...ale proszę wskazać autora, a utwory zależne objąć tą samą licencją.



...ale proszę wskazać autora i nie korzystać z utworu w celach komercyjnych.



...ale proszę wskazać autora i nie tworzyć utworów zależnych.



...ale proszę wskazać autora, nie korzystać z utworu w celach komercyjnych, a utwory zależne objąć tą samą licencją.



...ale proszę wskazać autora, nie korzystać z utworu w celach komercyjnych i nie tworzyć utworów zależnych.



RepOD
Repozytorium Otwartych Danych



UNIwersytet
Warszawski



Jak przebiega proces deponowania danych badawczych?

linktr.ee/pon_icm

repod@icm.edu.pl

Magdalena Bielińska m.bielinska@icm.edu.pl

Agnieszka Cybulska-Phelan a.cybulska@icm.edu.pl

Wojciech Fenrich wf@icm.edu.pl

Ankieta ewaluacyjna: forms.gle/79AXxYatafUaZ99AA



RepOD

Repozytorium Otwartych Danych



UNIwersytet
Warszawski

