

OPEN DATA DAY

Wrocław, 3.03.2018 r.



HR EXCELLENCE IN RESEARCH



Politechnika Wroclawska

Open Data Day 2018

- coroczne międzynarodowe święto otwartych danych
- obchodzone w tym roku po raz ósmy
- organizowane aby promować politykę otwartości danych publicznych oraz wskazywać korzyści ze swobodnego dostępu do nich i ponownego wykorzystania dla społeczeństwa, biznesu i administracji publicznej
- w tym roku organizatorzy proponują skupienie się na czterech głównych obszarach:



**WYDATKOWANIE
FINANSÓW PUBLICZNYCH**



OTWARTE MAPY/PLANY

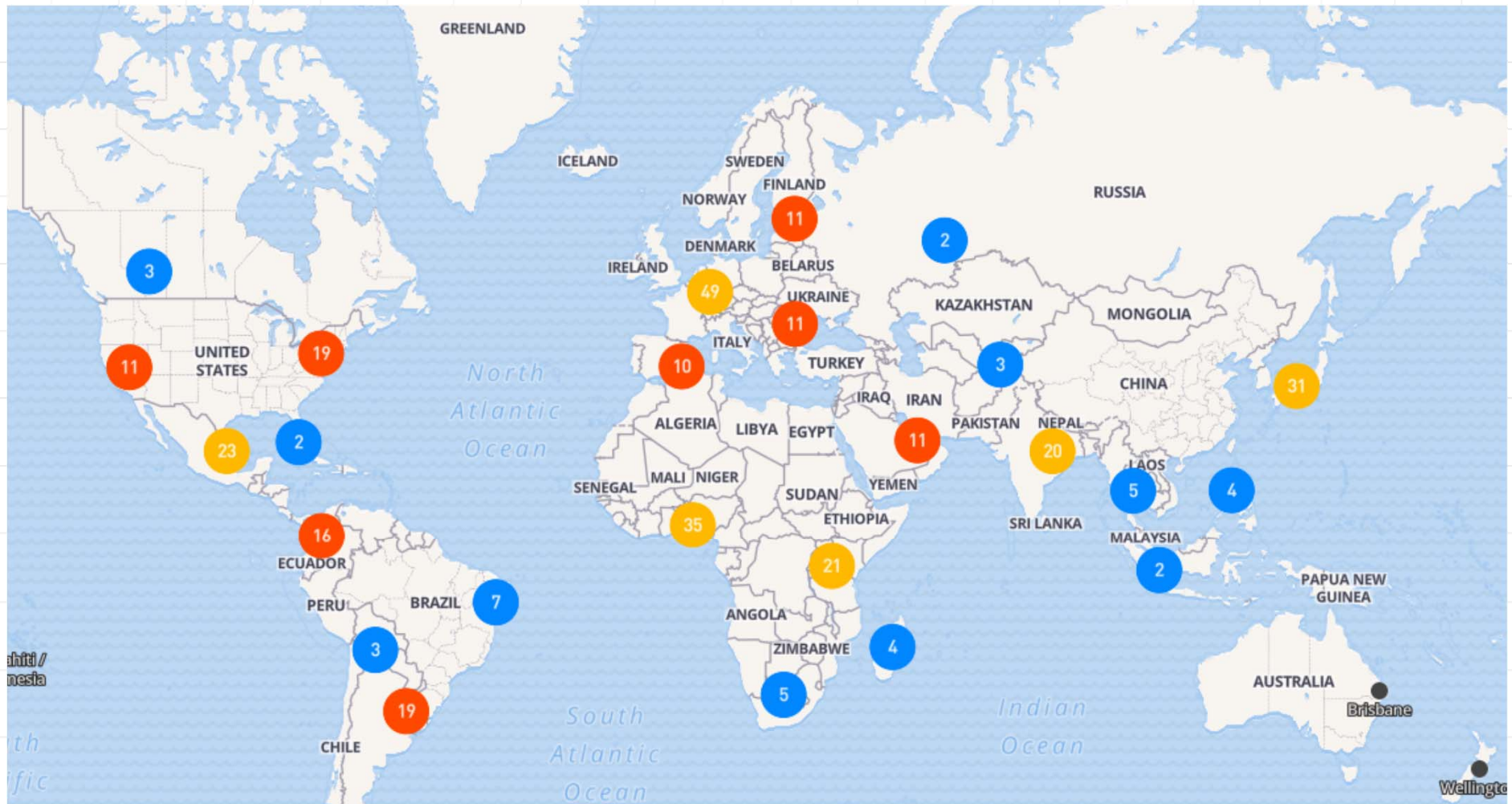


**DANE DLA
ZRÓWNOWAŻONEGO
ROZWOJU**



**OTWARTE DANE
BADAWCZE**

<http://opendataday.org/>



DOSTĘP DLA WSZYSTKICH
(BEZ DYSKRYMINACJI)

INTEROPERACYJNOŚĆ

OTWARTY FORMAT

ŁATWOŚĆ
WYSZUKIWANIA

OTWARTOŚĆ

BEZPŁATNY DOSTĘP
W SIECI

DOWOLNOŚĆ UŻYTKOWANIA
(Z OZNACZENIEM
AUTORSTWA DANYCH)

Dane publiczne

„(...) materiały finansowane ze środków publicznych powinny być dostępne dla wszystkich do swobodnego wykorzystania, gdyż stanowią nasze dobro wspólne”



Otwartość zasobów publicznych.

Online: <https://centrumcyfrowe.pl/projekty/otwartosc-zasobow-publicznych/>

[dostęp z dn. 02.02.2018]

Rodzaje otwartych danych



FINANSE (dane budżetowe,
dane podatkowe)



NAUKA (dane będące
wytworem procesu
badawczego)



DANE STATYSTYCZNE
(demografia, edukacja,
ochrona zdrowia, gospodarka)



KULTURA I SPORT
(wydarzenia, imprezy, kolekcje
bibliotek, archiwów, muzeów,
galerii, obiekty sportowe)



ŚRODOWISKO (ochrona,
zanieczyszczenia, klimat,
pogoda)



GEODANE

Otwarte dane badawcze

„Dane badawcze to informacja, w szczególności zebrane fakty, liczby, które mogą posłużyć badaniom i być traktowane jako podstawa do dalszego wnioskowania, dyskusji lub obliczeń.”

Za: Kierunki rozwoju otwartego dostępu do publikacji i wyników badań naukowych w Polsce, s. 18.
Online: http://www.nauka.gov.pl/g2/oryginal/2015_10/9f62cc350837b942e51ae23dd1f23df8.pdf
[dostęp z dn. 02.02.2018]

wytwarzane przez/dla
administracji publicznej
(pierwotnie nie
przeznaczone do celów
naukowych)

generowane
przez instytucje
naukowo-badawcze

Otwarte dane badawcze - przykłady

dokumenty tekstowe,
notatki

dane liczbowe
(w formie
nieprzetworzonej
lub czytelnej dla
komputera)

kwestionariusze
ankiet, wyniki badań
ankietowych

nagrania audio
i wideo, zdjęcia

modele
matematyczne,
algorytmy, schematy

oprogramowanie
(skrypty, pliki
wejściowe)

wyniki symulacji
komputerowych

protokoły
laboratoryjne, opisy
metodologiczne

próbki, artefakty,
obiekty

metadane (opisy)

E. Strzelczyk, *Otwarte dane badawcze – kolejny krok do otwierania nauki*. „Materiały Konferencyjne EBIB” 2017, nr 25.
Online: http://open.ebib.pl/ojs/index.php/Mat_konf/article/view/599/735 [dostęp z dn. 02.02.2018]

Korzyści otwierania danych badawczych



możliwość weryfikacji
wyników



obniżenie kosztów
pracy naukowej



wzbogacanie materiału
badawczego (łączenie
danych badawczych)



możliwość nawiązania
współpracy między
badaczami



większa
wyszukiwalność



szerszy krąg odbiorców



rozwój innowacji



szybszy przepływ
informacji



archiwizacja



jawność badań
finansowanych
ze środków publicznych

Zagrożenia/obawy związane z udostępnianiem danych badawczych



utrata autorstwa



utrata potencjału
dochodowego



brak ochrony danych
osobowych i danych
wrażliwych



brak korzyści
z publikowania danych
badawczych przy ocenie
naukowców



ryzyko wykrycia błędów
i nieprawidłowości



dodatkowa praca poświęcona
na przygotowanie danych

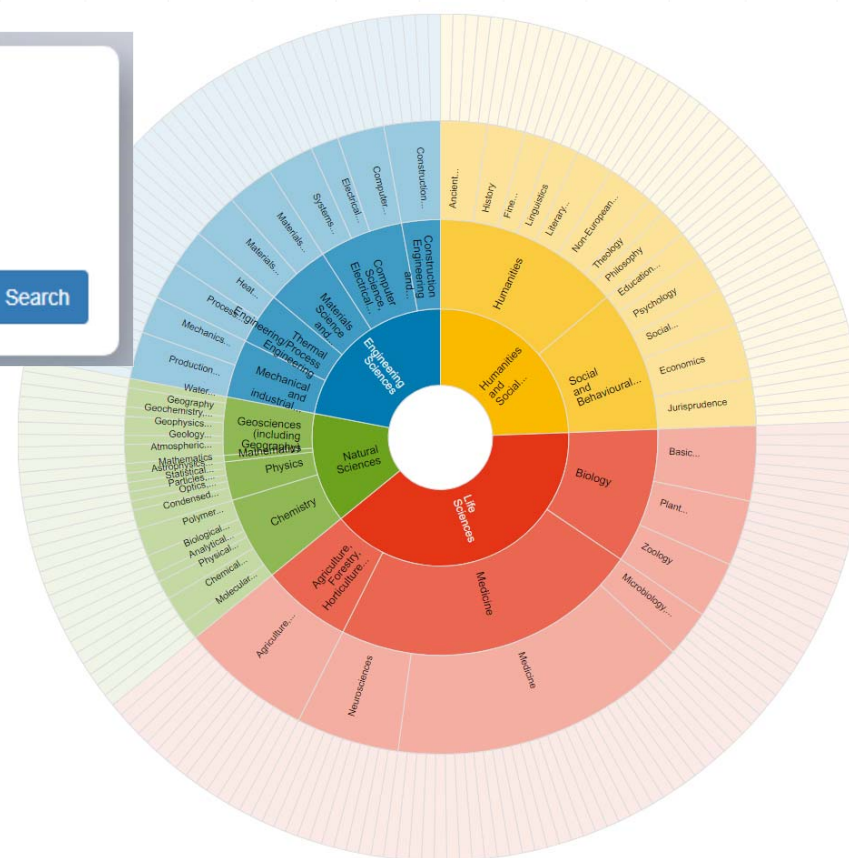


brak regulacji prawnych

Repozytoria otwartych danych badawczych



Wyszukiwarka repozytoriów
danych badawczych
re3data.org



Repozytoria otwartych danych badawczych – ogólne



Po co
udostępniać
dane
badawcze?

Po co udostępniać dane badawcze w otwartym repozytorium?

Otwarte udostępnianie danych badawczych może przynieść korzyści zarówno rozwojowi nauki, jak i karierze badacza:

- udostępnienie danych umożliwia ich **ponowną analizę** i zachęca do **nowych interpretacji**.

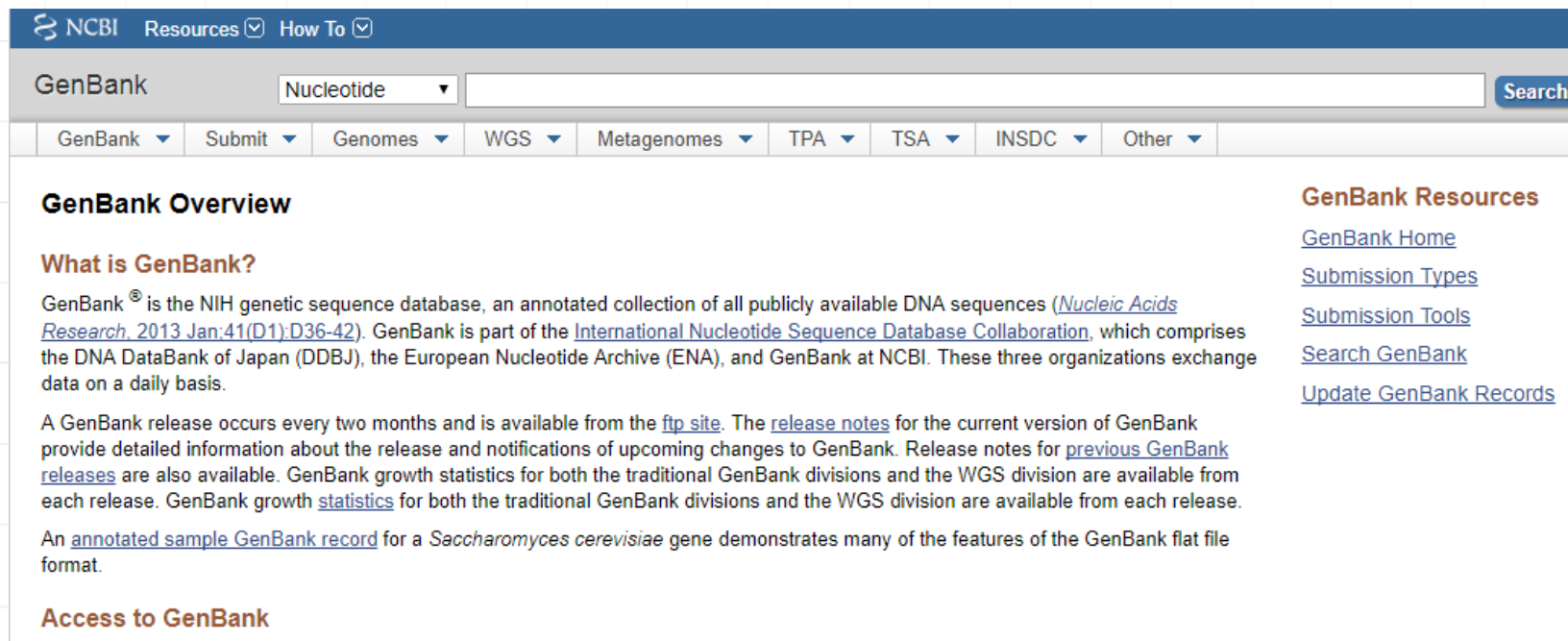
<https://repod.pon.edu.pl/pl/>

Repozytoria otwartych danych badawczych – ogólne

The screenshot displays the Zenodo website interface. At the top, there is a blue navigation bar with the Zenodo logo, a search bar, and links for 'Upload' and 'Communities'. On the right side of the navigation bar, there are 'Log in' and 'Sign up' buttons. Below the navigation bar, the main content area is titled 'Zenodo'. Underneath, there is a section for 'Recent uploads' which includes a search bar and a 'View' button. A specific upload is highlighted with a date of 'January 29, 2018 (2.1.0)', a 'Dataset' tag, and an 'Open Access' tag. The title of the upload is 'Genome assemblies for "Versatile genome assembly evaluation with QUASt-LG"'. The authors listed are Alla Mikheenko; Andrey Prjibelski; Vladislav Saveliev; Dmitry Antipov; and Alexey Gurevich. The description mentions 'De novo genome assemblies of Yeast_PB (S. cerevisiae, genome size: 12.1 Mb): Canu, FALCON, Flye, MaSuRCA (from Illumina pair-ends and PacBio SMRT) Yeast_NP (S. cerevisiae, genome size: 12.1 Mb): Canu, Flye, MaSuRCA (from Illumina pair-ends and Oxford Nanopores) Worm_PB (C. e'. It also notes the upload date as 'January 31, 2018' and that '2 more version(s) exist for this record'. To the right of the upload list, there is a green 'New upload' button and a grey box with the heading 'Want your upload to appear in this community?' and a list of three bullet points: 'Click the button above to upload straight to this community.', 'The community curator is notified, and will either accept or reject your upload (see community curation policy above).', and 'If your upload is rejected by the curator, it will still be available on Zenodo, just not in this community.'

<https://zenodo.org>

Repozytoria otwartych danych badawczych - dziedzinowe



The screenshot shows the NCBI GenBank website. At the top, there is a navigation bar with 'NCBI Resources' and 'How To' links. Below this is a search bar with a dropdown menu set to 'Nucleotide' and a 'Search' button. A secondary navigation bar contains links for 'GenBank', 'Submit', 'Genomes', 'WGS', 'Metagenomes', 'TPA', 'TSA', 'INSDC', and 'Other'. The main content area is titled 'GenBank Overview' and includes a section 'What is GenBank?' with a paragraph explaining the database's scope and a list of resources on the right side.

GenBank Overview

What is GenBank?

GenBank[®] is the NIH genetic sequence database, an annotated collection of all publicly available DNA sequences (*Nucleic Acids Research*, 2013 Jan;41(D1):D36-42). GenBank is part of the [International Nucleotide Sequence Database Collaboration](#), which comprises the DNA DataBank of Japan (DDBJ), the European Nucleotide Archive (ENA), and GenBank at NCBI. These three organizations exchange data on a daily basis.

A GenBank release occurs every two months and is available from the [ftp site](#). The [release notes](#) for the current version of GenBank provide detailed information about the release and notifications of upcoming changes to GenBank. Release notes for [previous GenBank releases](#) are also available. GenBank growth statistics for both the traditional GenBank divisions and the WGS division are available from each release. GenBank growth [statistics](#) for both the traditional GenBank divisions and the WGS division are available from each release.

An [annotated sample GenBank record](#) for a *Saccharomyces cerevisiae* gene demonstrates many of the features of the GenBank flat file format.

Access to GenBank

GenBank Resources

- [GenBank Home](#)
- [Submission Types](#)
- [Submission Tools](#)
- [Search GenBank](#)
- [Update GenBank Records](#)

<https://www.ncbi.nlm.nih.gov/genbank/>

Repozytoria otwartych danych badawczych - dziedzinowe



NOFER INSTITUTE OF OCCUPATIONAL MEDICINE

ECNIS-NIOM Repository

Home Browse About Submit an Item Visualize

BROWSE ITEMS BY

- Communities [Show All](#)
- Title
- Authors
- Issue Date
- Submit Date
- Subjects
- Subject (MeSH)
- Researchers

WELCOME TO OUR OPEN ACCESS REPOSITORY OF ECNIS-NIOM RESEARCH!

The database has been developed to store, index, preserve and redistribute in digital format work of ECNIS-NIOM researchers and makes it freely available on the Web.

There are two principal ways of finding content within this repository:

- **Searching** (the Simple Search Box and Advanced Search appear at the top right hand corner on each page)
- **Browsing** (the various browse options are listed on the top left of the left-hand navigation bar)


Please use the [Quick Guides](#) on the right hand menu for further information. If you have any questions regarding this repository then please use the [Feedback form](#) to contact the database administrator.

Top 5 viewed publications:

- [The challenge resulting from positive and negative effects of sunlight: how much solar UV exposure is appropriate to balance between risks of vitamin D deficiency and skin cancer? \(Reichrath, Jörg\)](#)
- [Toxic and metabolic effect of sodium butyrate on SAS tongue cancer cells: role of cell cycle deregulation and redox changes. \(Jeng, Jjiang-Huei et al\)](#)
- [Concentrations of resveratrol and derivatives in foods and estimation of dietary intake in a Spanish population: European Prospective Investigation into Cancer and Nutrition \(EPIC\)-Spain cohort. \(Zamora-Ros, Raul et al\)](#)

QUICK GUIDES

- Getting started
- Using search
- Browsing
- Submitting content



<http://ecnis.openrepository.com/ecnis/>



Otwieranie danych badawczych na Politechnice Wrocławskiej – Projekt AZON



<http://zasobynauki.pl>

Otwieranie danych badawczych na Politechnice Wrocławskiej - Projekt AZON - moduły

Biblio

- Skany 3D
- Nagrania Telewizji Politechniki Wrocławskiej
- Audycje radia LUZ
- Zasoby Dolnośląskiej Biblioteki Cyfrowej
- Baza DONA (publikacje i prace niepublikowane PWr)

Science

- Pliki wyjściowe i wejściowe różnych programów naukowych
- Pliki wyjściowe urządzeń typu firewall
- Baza przepływów ruchu sieciowego przez WCSS
- Dane rozliczeniowe klastra obliczeniowego
- Repozytorium kodów źródłowych
- Repozytorium danych IoT (Internet of Things)
- Programy i narzędzia udostępniane w ramach sieci CLARIN
- Baza danych z rekordami zawierającymi opisy dróg startowych lotnisk
- Dane eyetrackingowe, termograficzne itd.
- Ontologie językowe opisujące wiedzę z różnych dziedzin
- Repozytorium usług implementujących algorytmy detekcji anomalii w zbiorach danych
- Repozytorium kodu źródłowego autorskiego oprogramowania CFD/LBM do symulacji zjawisk transportu w mikro- i makrosystemach
- Baza danych powiązana z siecią telemetryczną

Otwieranie danych badawczych na Politechnice Wrocławskiej - Projekt AZON - moduły

Learning

- E-kursy z obszaru Analizy Matematycznej i Fizyki
- Wideo tutoriale dotyczące projektowania konstrukcji w programie AutoDesk Robot Structural Analysis
- Prezentacje i notatki z wykładów, ćwiczenia, instrukcje laboratoryjne, skrypty
- Instrukcje i ćwiczenia przygotowujące do samodzielnej nauki podstaw sieci komputerowych
- Wirtualne węzły sieciowe Akademii CISCO

Badania

- Bazy danych Ośrodka Współpracy Nauki z Gospodarką zawierające informacje o aparaturze, laboratoriach, ekspertach PWr, towarzystwach naukowych
- Baza danych zawierająca wyniki ankietowe, raporty z badań i artykułu recenzowane naukowo dotyczące badania *Mój Idealny Pracodawca* od 2014 roku

Statystyka

- Wybrane i zanonimizowane dane statystyczne Uczelni pochodzące z Zintegrowanego Systemu Informatycznego (ZSI), Jednolitego Systemu Obsługi Studenta (JSOS) oraz Repozytorium Agregowanych Danych o Uczelni (RADON)
- biblioteczne dane statystyczne generowane z systemu bibliotecznego Aleph

Otwieranie danych badawczych na Politechnice Wrocławskiej - Dolnośląska Biblioteka Cyfrowa

Dolnośląska Biblioteka Cyfrowa

(38 tys. publikacji):

- publikacje książkowe oraz czasopisma zawierające dane badawcze oraz dane statystyczne z różnych dziedzin wiedzy
- dane bibliograficzne
- dane o charakterze historycznym do prowadzenia badań porównawczych: książki adresowe, bibliografie, almanachy, raporty, inwentarze, spisy, kroniki
- raporty z prowadzonych badań

Przewidziane do udostępnienia po opracowaniu:

- kolekcja fotografii, zdigitalizowanych w ramach projektu z zakresu działalności upowszechniającej naukę (MNiSW), zawierająca zbiory pomiarów architektonicznych zabytków z terenu Dolnego Śląska (zwymiarowane obiekty)



Dolnośląska Biblioteka Cyfrowa



Otwieranie danych badawczych na Politechnice Wrocławskiej - CLARIN



CENTRUM TECHNOLOGII JĘZYKOWYCH
CLARIN-PL



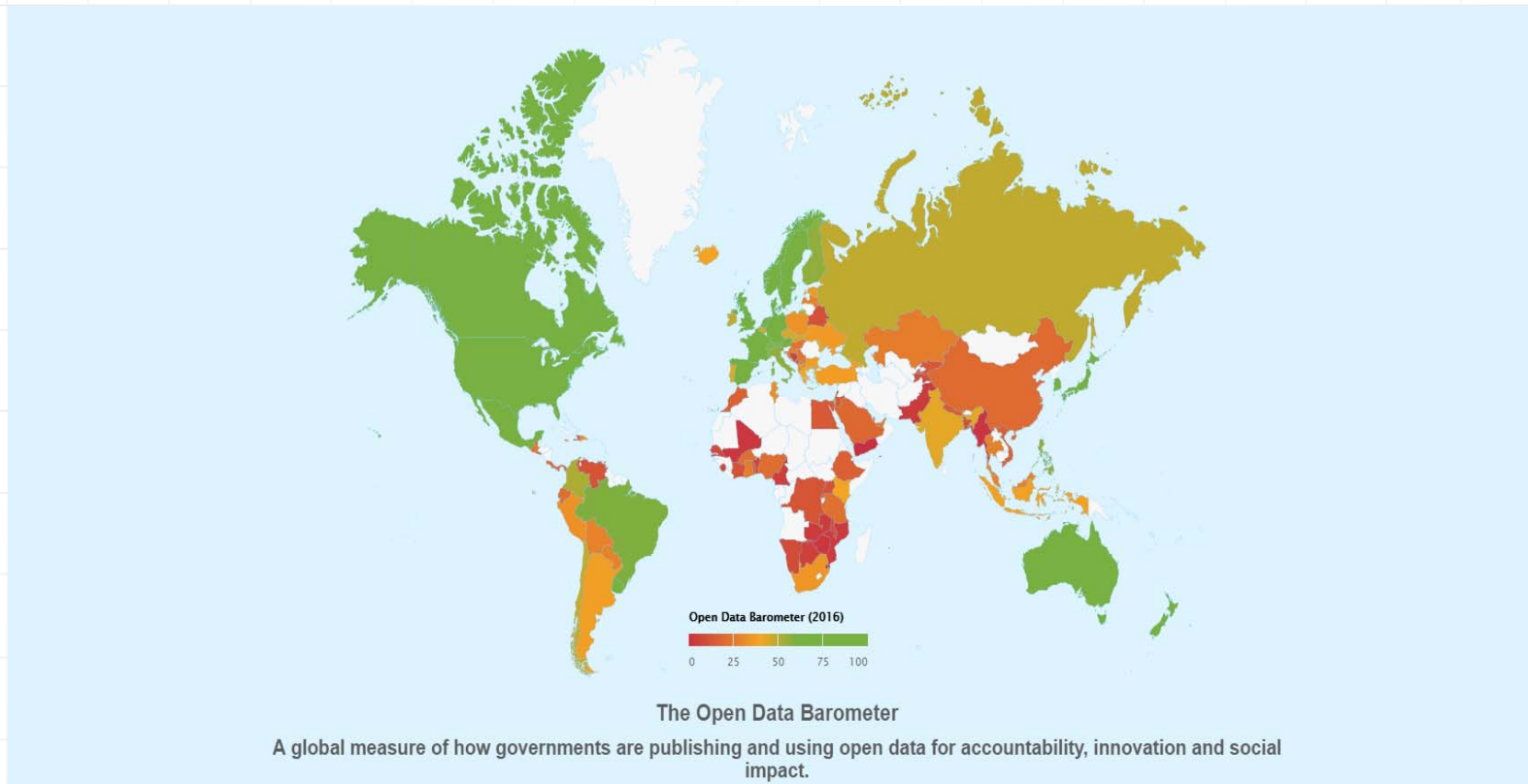
[Advanced Search](#)

Author	Subject	Language (ISO)
Marcińczuk, Michał (33)	corpus (13)	Polish (243)
Kocoń, Jan (23)	Polish (11)	English (12)
Piasecki, Maciej (23)	korpus (9)	Belarusian (1)
Radziszewski, Adam (19)	plWordNet (8)	Bulgarian (1)
Kędzia, Paweł (18)	syntax (7)	Lithuanian (1)
... View More	... View More	... View More

<https://clarin-pl.eu/dspace/>



Otwarte dane publiczne na świecie



[http:// opendatabarometer.org/](http://opendatabarometer.org/)

Otwarte dane publiczne - serwisy



Zbiory danych Aplikacje Dostawcy Baza wiedzy O serwisie

Zaloguj się

PL EN

Dane Publiczne

Dane o szczególnym znaczeniu dla rozwoju innowacyjności w państwie i rozwoju społeczeństwa informacyjnego w jednym miejscu

np. Baza teleadresowa



<http://danepubliczne.gov.pl/>

Otwarte dane publiczne - serwisy

- A+ | A

ZALOGUJ

otwartedane wrocław

Zbiory danych Organizacje Kategorie Wnioski Artykuły O serwisie

Unia Europejska

Dane publiczne i otwarte.
Wyszukuj w jednym miejscu.

Szukaj...

Plan zagospodarowania Bilety MPK Rozrywka Tablice rejestracyjne Internaty

Budżet Obywatelski Tramwaje Hot Spot Galerie sztuki

Adresy Biblioteki Gimnazja Internet

Kluby sportowe Komunikacja miejska Edukacja Szkoły

Autobusy Rekreacja Sport Ulice Pogoda

Adwokaci Parkingi Licea Rower miejski Mapy

Demografia Budżet Przystanki Inwestycje Infografiki Oświata

Samochody elektryczne Przedszkola Stacje meteorologiczne

<https://www.wroclaw.pl/open-data/>

Bibliografia

- Hoffman-Sommer M., *Otwieranie małych danych badawczych*.
Online: [http://repo.bg.pw.edu.pl/images/Otwieranie małych danych badawczych.pdf](http://repo.bg.pw.edu.pl/images/Otwieranie_ma%C5%82ych_danych_badawczych.pdf)
[dostęp z dn. 02.02.2018]
- Jachimczyk, A. *Otwarte dane badawcze. Casus polskich instytutów badawczych*.
„Zagadnienia Naukoznawstwa” 2015, r. 206, nr 4, s. 409–424.
- *Kierunki rozwoju otwartego dostępu do publikacji i wyników badań naukowych w Polsce*.
Online: http://www.nauka.gov.pl/g2/oryginal/2015_10/9f62cc350837b942e51ae23dd1f23df8.pdf
[dostęp z dn. 02.02.2018]
- *Otwartość zasobów publicznych*.
Online: <https://centrumcyfrowe.pl/projekty/otwartosc-zasobow-publicznych/>
[dostęp z dn. 02.02.2018]
- Strzelczyk E., *Otwarte dane badawcze – kolejny krok do otwierania nauki*. „Materiały Konferencyjne EBIB” 2017, nr 25.
Online: http://open.ebib.pl/ojs/index.php/Mat_konf/article/view/599/735 [dostęp z dn. 02.02.2018]

*Wszystkie ikony wykorzystane w prezentacji pochodzą z aplikacji *Icons by Noun Project* oraz *Icondrop* dla programu MS PowerPoint.

*Slajdy 17 i 18 pochodzą z prezentacji *Przygotowanie i udostępnienie zasobów nauki Politechniki Wrocławskiej* aut. Wojciecha Drzewińskiego.