

Publikovanie vedeckých dát

Matej Harvát

Odbor podpory otvorenej vedy

CVTI SR



Investícia do Vašej budúcnosti



EURÓPSKA ÚNIA
Európske štrukturálne a investičné fondy
OP Integrovaná infraštruktúra 2014 – 2020



MINISTERSTVO
DOPRAVY A VÝSTAVBY
SLOVENSKEJ REPUBLIKY



MINISTERSTVO
ŠKOLSTVA, VEDY,
VÝSKUMU A ŠPORTU
SLOVENSKEJ REPUBLIKY



Prečo publikovať dáta?

Ako publikovať dáta?



Výhody publikovania dát

- objektívna potreba zabezpečenia dlhodobého, trvalého prístupu k vedeckým údajom
- potreba jednoznačnej identifikácie tvorcu zdroja/vedeckého súboru údajov/dát
- V mnohých prípadoch spočíva najväčšia hodnota súboru údajov v jeho zdieľaní, nie nevyhnutne v poskytovaní interpretácie alebo analýzy.
- Čoraz viac financovateľov očakáva, že údaje získané v rámci výskumných projektov, ktoré financujú, budú vyhľadateľné, dostupné a čo najotvorenejšie.



Prečo publikovať výskumné dáta



Pomáhanie vede: objavovanie dát a prístup k nim posilňuje otvorené vedecké skúmanie – reprodukovateľnosť a transparentnosť



Rozpoznanie a uznanie kreditu: Môže poskytnúť priame zásluhy výskumníkovi



Zvyšuje dosah a viditeľnosť výskumu



Propaguje výskum a preukazuje jeho využiteľnosť a význam

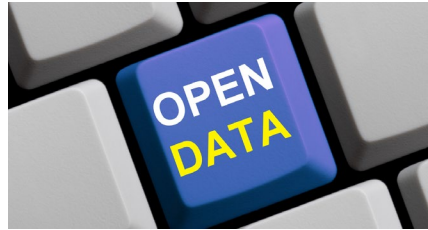


Znižuje náklady vynakladané na zber duplikovaných údajov



Zverejňovaním výskumných údajov sa realizujú politiky financovateľov výskumu a inštitúcií

Politiky, mandáty, dobrá prax



Politiky zdieľania otvorených dát:

- Vydavateľstvá (napr. Wiley, Sage, Springer)
- Financovatelia výskumu (Open Research Europe)
- Inštitúcie a organizácie

Immediate open access to research data: a federal mandate and much debate

CATEGORIES

Open access / open source,
Publication planning,
Transparency

KEY TAKEAWAYS

- From the end of 2025, US government policy will mandate immediate open access to federally funded research data and publications, eliminating the current 12-month embargo.
- While further details on implementation are worked out, publishers, funders, and researchers grapple with the best ways to fund open access.

Mandáty zdieľania otvorených dát

- zatiaľ ojedinelé
- (príliš) komplexná problematika správy dát (porov. „data friction“)

Fakty a čísla o otvorených výskumných dátach podľa Európskej komisie:

https://research-and-innovation.ec.europa.eu/strategy/strategy-2020-2024/our-digital-future/open-science/open-science-monitor/facts-and-figures-open-research-data_en

zdieľanie a
zverejnenie
dát

Dobrá prax otvorenej vedy

Kolaboratívny výskum a vedecké publikovanie si vyžaduje zdieľanie údajov ako jednu z povinností integrity výskumu.

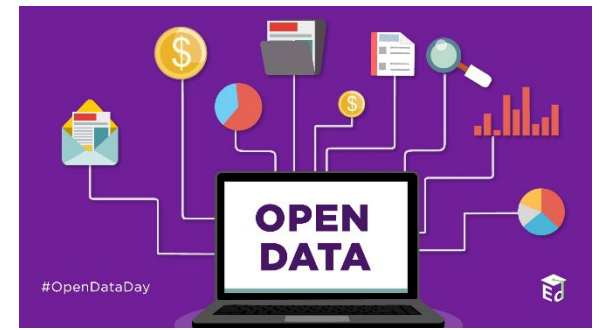
Otvorený prístup k výstupom výskumu zvyšuje efektívnosť vedeckého systému tým, že znižuje duplicitu a náklady na tvorbu a opätovné použitie dát, umožňuje väčší počet výskumov na základe tých istých dát a znásobuje možnosti účasti verejnosti na výskumnom procese.

archivácia
dát

Jeden súbor dát, viacero možností publikovania

Dáta je možné zverejniť rôznymi spôsobmi. Možnosti, ktoré budú uprednostené, môžu závisieť od existujúcich postupov v tom-ktorom odbore alebo od očakávaní poskytovateľa finančných prostriedkov.

- Uverejňovanie dát ako doplnkového materiálu spojeného s výskumným článkom, zvyčajne s dátovými súbormi, ktorých hosting zabezpečuje vydavateľ článku
- Uloženie dát do úložiska (repozitára), ktoré bolo vyvinuté na podporu zverejňovania údajov, napr. Dataverse, Dryad, Figshare, Zenodo.
- Publikovanie článku o súbore údajov (data paper), ktorý môže byť uverejnený ako:
 - 1. preprint,
 - 2. v (klasickom) časopise alebo
 - 3. v dátovom časopise, ktorý je určený na publikovanie článkov o údajoch. Dáta môžu byť umiestnené v časopise alebo samostatne v dátovom úložisku. Príkladmi dátových časopisov sú Scientific Data (vydavateľstvo SpringerNature) a Data Science Journal (vydavateľstvo






Intelligence
Volume 40, Issue 2, March–April 2012, Pages 73-76



Publish (your data) or (let the data) perish! Why not publish your data too?

Jelte M. Wicherts , , Marjan Bakker

Show more 

+ Add to Mendeley  Share  Cite

<https://doi.org/10.1016/j.intell.2012.01.004>

[Get rights and content](#)

Abstract

The authors argue that upon publication of a paper, the data should be made available through online archives or repositories. Reasons for not sharing data are discussed and contrasted with advantages of sharing, which include abiding by the scientific principle of openness, keeping the data for posterity, increasing one's impact, facilitation of secondary analyses and collaborations, prevention and correction of errors, and meeting funding agencies' increasingly stringent stipulations concerning the dissemination of data. Practicing what they preach, the authors include data as an online appendix to this

AKÝM SPÔSOBOM SA BUDÚ ZDIELAŤ DÁTA, KTORÉ MAJÚ BYŤ ZVEREJNENÉ?

5 CIEST PUBLIKOVANIA DÁT

([CESSDA](#))

- [Dátový časopis](#)
- [Inštitucionálny dátový repozitár](#)
- [Všeobecný dátový repozitár](#)
- [Odborový dátový repozitár](#)
- [Dôveryhodný odborový dátový repozitár](#)

DÁTOVÝ REPOZITÁR

obrovská databázová

infraštruktúra, ktorá

zhromažďuje, spravuje a ukladá

rôzne súbory dát na analýzu,

distribúciu a tvorbu

požadovaných výstupov.

TYPY DÁTOVÝCH REPOZITÁROV (PRE OPEN DATA)

- Data Warehouse
- Data Lake
- Data Marts
- Data Cubes
- Metadata Repositories

Dátové repozitáre

Online úložiská výskumných údajov

- veľké databázové infraštruktúry vytvorené na správu, zdieľanie, prístupňovanie a archiváciu súborov údajov výskumných pracovníkov.
- Repozitáre môžu byť špecializované a určené na zhromažďovanie údajov z jednotlivých odborov alebo všeobecnejšie, zhromažďujúce údaje z väčších oblastí poznania, ako sú prírodné alebo spoločenské vedy.
- Repozitáre môžu tiež agregovať údaje od expertov na globálnej alebo lokálnej úrovni, zhromažďovať údaje výskumníkov univerzity alebo konzorcia univerzít pre vzájomný prospech.

Repozitár umožňuje skúmanie, dokazovanie, hodnotenie, transparentnosť a overovanie výsledkov výskumníka inými odborníkmi **nad rámec publikovaného recenzovaného akademického článku.**

- ❑ Umiestnenie výskumných údajov online umožňuje okamžitý prístup globálne rozptýlenej skupine výskumných pracovníkov, ktorí sa môžu podeliť o výsledky, pochopiť ich a syntetizovať.
- ❑ Dátové úložiská umožňujú aj zverejnenie predtým skrytých negatívnych údajov, v podstate experimentov, ktoré nefungovali. To umožňuje ostatným výskumníkom vyhnúť sa predchádzajúcim slepým uličkám tých, ktorí vyskúšali cestu pred nimi, aby lepšie našli cestu k úrodnejšiemu územiu.
- ❑ Globálna/lokálna komunita odborníkov profituje z online zdieľania a agregácie výskumných údajov.



Dátové repozitáre



VÝBER DÁTOVÉHO REPOZITÁRA

- Existuje tematická databáza/databáza komunity?
- Je dlhodobo udržateľný?
- Ukladá dáta bezpečným spôsobom?
- Budú dáta vyhľadateľné (napr. pomocou PID), prístupné a opakovaná použiteľné?
- Popisuje dáta štandardným spôsobom? Používa prijaté štandardy metadát?
- Môže vkladateľ určiť licenciu, ktorou sa riadi prístup k dátam a ich opakované použitie?

AKO NÁJŠŤ VHODNÝ OA REPOZITÁR?

Medzinárodné registre repozitárov s otvoreným prístupom

- **Registry of Open Access Repositories, ROAR**, (<http://roar.eprints.org/>)
- **Open Directory of Open Access Repositories, OpenDOAR** (<https://v2.sherpa.ac.uk/opensoar/>)
- **re3data - Registry of Research Data Repositories**, (<https://www.re3data.org/>)
- **Registry of Open Access Repository Mandates, ROARMAP**, (politiky, <http://roarmap.eprints.org/>)
- **OpenAire odporúčania** <https://www.openaire.eu/find-trustworthy-data-repository-certified-repositories>
- **Core Trust Seal certifikáty** <https://www.coretrustseal.org/>



Overenie dôveryhodnosti repozitára

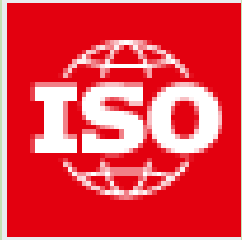
Referenčný model OASIS pre architektúru zameranú na služby

OASIS: nezisková organizácia, vznikla v r. 1993

Repozitár by mal spĺňať 10 podmienok:

1. Musí prijať **záväzok dlhodobej a udržateľnej správy** digitálnych objektov pre cieľovú komunitu.
2. Musí **preukázať spôsobilosť a zdroje** pre zabezpečenie udržateľnosti.
3. Musí mať **jasné formálne a legislatívne rámce** ako východisko svojich záväzkov.
4. Musí mať efektívnu a udržateľnú rámcovú **stratégiu**.
5. Musí definovať a prijať **inštitucionálne interné politiky** pre získavanie a ukladanie digitálnych objektov.
6. Musí byť schopný poskytnúť **zabezpečenie integrity**, autenticity a dostupnej použiteľnosti digitálnych objektov.
7. **Vytvára a uchováva relevantné metadáta** o digitálnych objektoch, obsahu, ako aj udalostiach súvisiacich s objektmi, vrátane podmienok prístupňovania a kontextu použitia.
8. Musí byť schopný zabezpečiť splnenie požiadaviek na **sprístupňovanie** digitálnych objektov konkrétnej komunite.
9. Musí mať spracovanú **stratégiu dlhodobej ochrany** vrátane riešenia bezpečnostných konfliktov a havárií systému.
10. Musí mať **technickú infraštruktúru adekvátnu udržateľnej údržbe** a zaistenia digitálnych objektov.

Overenie dôveryhodnosti repozitára – certifikáty



ISO 16363 Audit and certification of trustworthy digital repositories

- komplexný externý audit + certifikácia
- repozitáre s rôznym obsahom, on-site

CoreTrustSeal

- dátové repozitáre
- menej metrík, kratšia doba trvania, každé tri roky, off-site



Nestor Seal

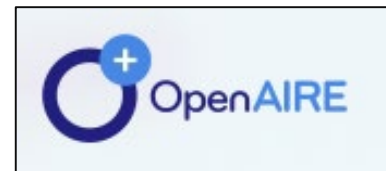
podľa nemeckej normy DIN 31644



Európska komisia a projekty financované EK:

FAIR dáta v dôveryhodných repozitároch

<https://www.openaire.eu/find-trustworthy-data-repository>



Dátové repozitáre – príklady



Image: Ainsley Seago. "To deposit or not to deposit, that is the question."
doi:10.1371/journal.pbio.1001779. CC-BY License.

Repository Name	Information on fees/costs	Size limits	Integrated with <i>Scientific Data's</i> manuscript submission system	Re3data / FAIRsharing entry
Dryad Digital Repository	\$120 USD for first 20 GB, and \$50 USD for each additional 10 GB	None stated	Yes ✓	view FAIRsharing entry
figshare	100 GB free per <i>Scientific Data</i> manuscript.	1 TB per dataset	Yes ✓ - To qualify for the 100 GB of free storage, data must be uploaded to figshare via our submission system. Download instructions.	view FAIRsharing entry
Harvard Dataverse	Contact repository for datasets over 1 TB	2.5 GB per file, 10 GB per dataset	No	view re3data entry
Open Science Framework	Free of charge	5 GB per file, multiple files can be uploaded	No	view FAIRsharing entry
Zenodo	Donations towards sustainability encouraged	50 GB per dataset	No	view re3data entry
Science Data Bank	Free of charge	8 GB per file, no limit to dataset size	No	view FAIRsharing entry

[Data Repository Guidance – Nature](#)

Usmernenia + odkazy na odborové repozitáre

<https://www.nature.com/sdata/policies/repositories>

a všeobecné repozitáre

<https://www.nature.com/sdata/policies/repositories#general>

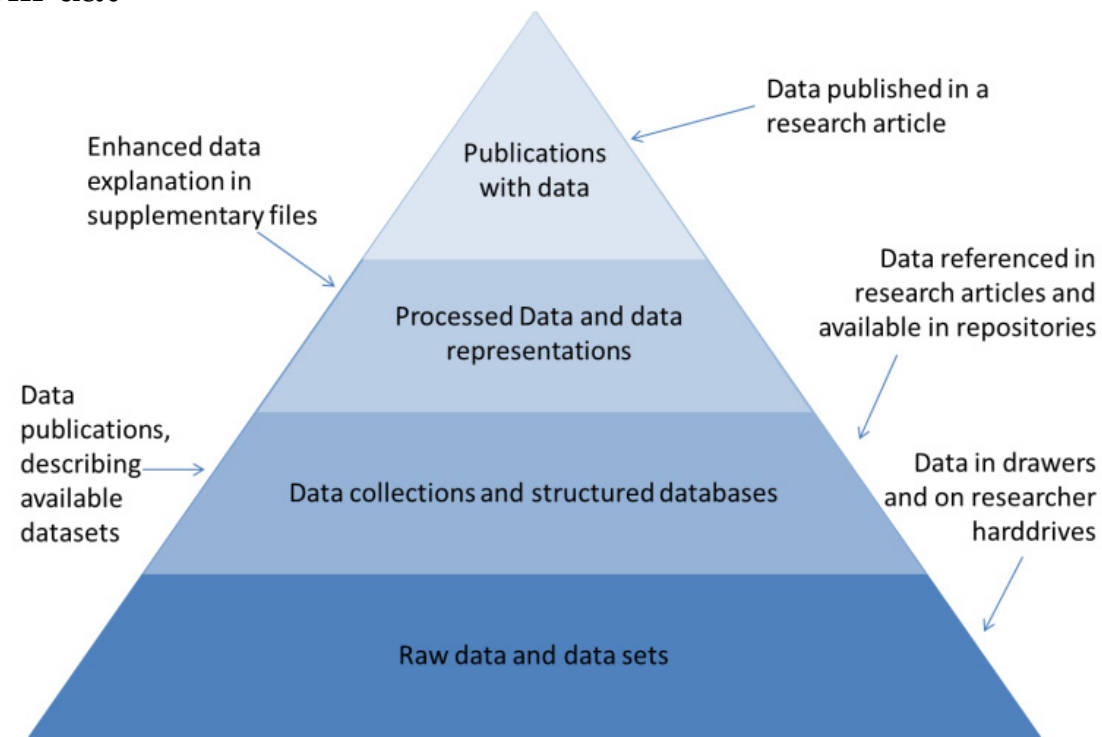
Dátové časopisy

- publikácie, ktorých hlavným cieľom je zverejňovať súbory údajov.

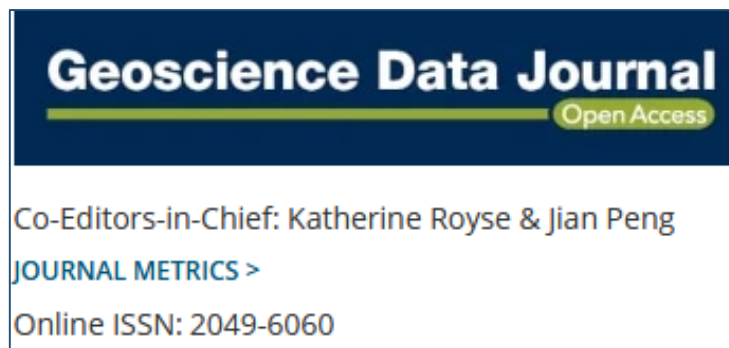
Ciele dátových časopisov:

- podporovať opakované použitie výsledkov výskumu
- Zlepšiť transparentnosť vedeckých metód a výsledkov
- Podporovať dobrú prax správy dát
- Poskytnúť prístupnú, trvalú a prehľadnú cestu k súborom dát

Autor sa v dátových časopisoch môže sústrediť na samotné údaje, a nie na rozsiahlu analýzu týchto údajov, ktorá sa vyskytuje v tradičných časopisoch.



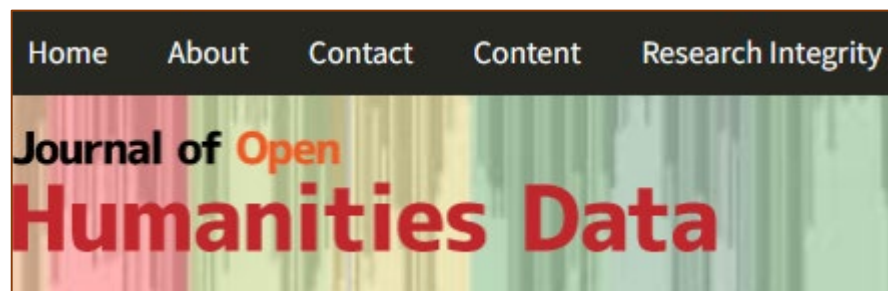
Dátové časopisy – příklady



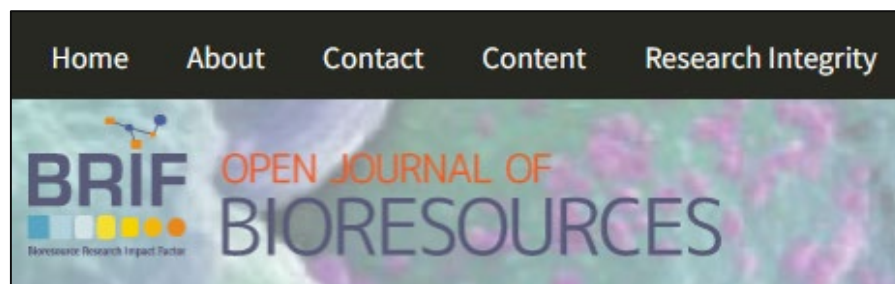
<https://academic.oup.com/ije>

<https://brill.com/view/journals/rdj/rdj-overview.xml>

<https://rmets.onlinelibrary.wiley.com/journal/20496060>



<https://openhumanitiesdata.metajnl.com/>



<https://openbioresources.metajnl.com/>

Viac na:
<https://www.wiki.ed.ac.uk/display/dashare/Sources+of+dataset+peer+review>

Data Journals

Eine Liste mit Data Journals, die durch die Community ergänzt und gepflegt werden kann, wird

Einige Beispiele für Data Journals werden hier gelistet:

- [Atomic Data and Nuclear Data Tables](#) (Elsevier)
- [Biodiversity Data Journal](#) (Pensoft Publishers)
- [Biomedical Data Journal](#) (Procon Ltd.)
- [BMC Research Notes](#) (Biomed Central)
- [Chemical Data Collections](#) (Elsevier)
- [Data](#) (MDPI)
- [Data in Brief](#) (Elsevier)
- Rubrik [Data Observer](#) der Jahrbücher für Nationalökonomie und Statistik (DeGruyter)
- [Earth System Science Data - ESSD](#) (Copernicus Publications)
- [Ecological Archives](#) (Ecological Society of America - ESA/Wiley)
- [E1000Research](#) (E1000 Research)

- Hauptseite
- Letzte Änderungen
- Zufällige Seite
- Hilfe zu MediaWiki
- Werkzeuge
- Links auf diese Seite
- Änderungen an verlinkten Seiten
- Spezialseiten
- Druckversion
- Permanenter Link
- Seiteninformationen



Zenodo.org will be unavailable for 2 hours on October 13 from 06:00-08:00 UTC - see announcement and learn what's new and what's changed.

November 3, 2022

List of data journals

Kindling, Maxi; Strecker, Dorothea

This document describes a dataset that aggregates information about 135 data journals. Data journals focus on the publication of data papers – a specialized publication type describing datasets, their collection and reuse potential that is peer-reviewed, citable and indexed.

This dataset includes a comprehensive list of data journals that was compiled by aggregating existing sources, as well as an overview of these sources.

The list is continually updated on GitHub, where additional information on data journals (URLs of data journal homepages) is provided: <https://github.com/MaxiKi/data-journals>

issn	journal_title	publisher	data_journal_type
2363-4952	RIDE - A review journal for digital editions and resources	Institut für Dokumentologie und Editorik	pure
2059-481X	Journal of Open Humanities Data	Ubiquity Press	pure
2296-7745	Frontiers in Marine Science	Frontiers	mixed
2054-345X	Human Genome Variation	Springer Nature	mixed
2032-6378	Journal of Astronomical Data	Vrije Universiteit Brussel	pure
2603-431X	Viticulture Data Journal	Pensoft Publishers	pure
2661-8850	BMC Molecular and Cell Biology	BioMed Central (Springer Nature)	mixed
1472-6823	BMC Endocrine Disorders	BioMed Central (Springer Nature)	mixed
0092-640X	Atomic Data and Nuclear Data Tables	Elsevier	pure

- [Open Journal of Bioresources](#) (Ubiquity Press)

2,753 views
1,488 downloads

[See more details...](#)

Indexed in



Publication date:
November 3, 2022

DOI:
[10.5281/zenodo.7082126](https://doi.org/10.5281/zenodo.7082126)

Keyword(s):
[research data](#) [data journal](#)

Communities:
[re3data - Registry of Research Data Repositories](#)

License (for files):
[Creative Commons Zero v1.0 Universal](#)

Versions

Version 1.0
10.5281/zenodo.7082126
Nov 3, 2022

(TUjournals)

ess)

stronomical Society)

lications)

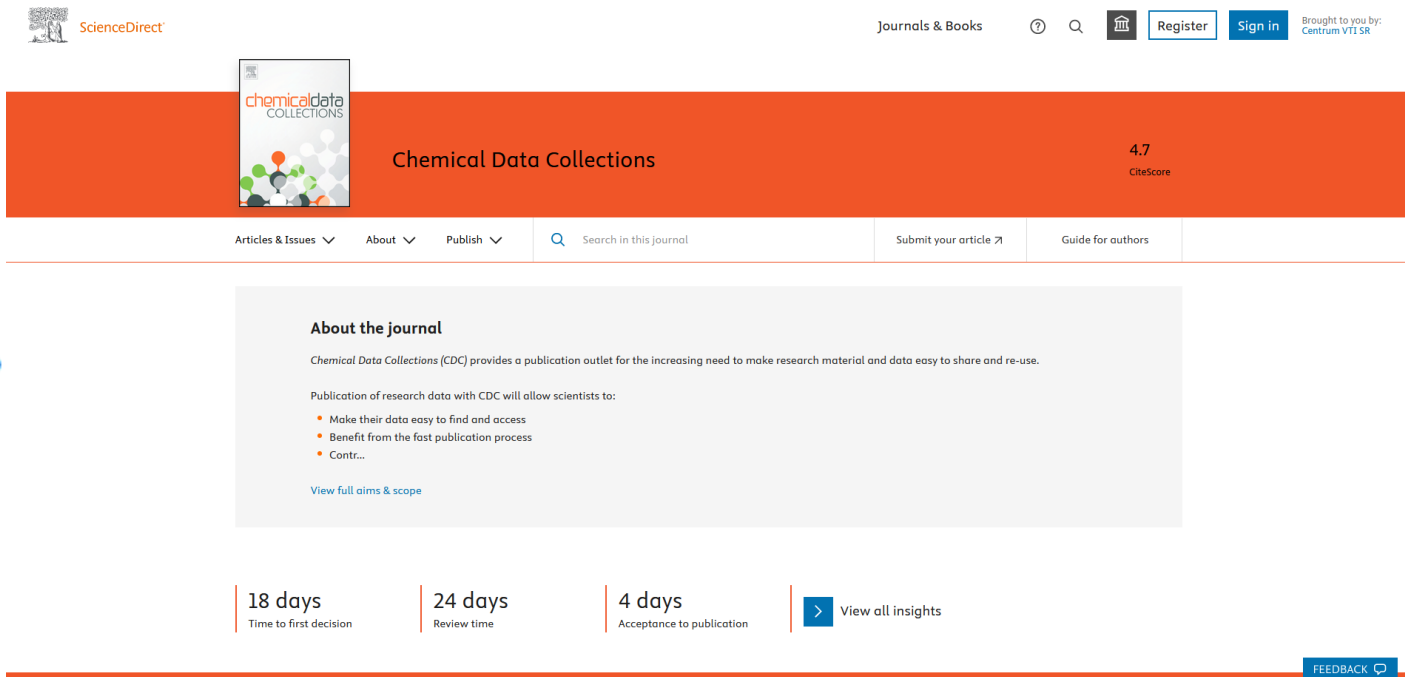
)

)

lishing)

Zoznamy dátových časopisov:

- [Wiki](#)
https://www.forschungsdaten.org/index.php/Data_Journals
- [Zenodo](#) <https://zenodo.org/record/7082126>
- [Vydavateläia](#) (Elsevier, Wiley etc.)



ScienceDirect Journals & Books Register Sign in Brought to you by: Centrum VTI SR

Chemical Data Collections 4.7 CiteScore

Articles & Issues About Publish Search in this journal Submit your article Guide for authors

About the journal

Chemical Data Collections (CDC) provides a publication outlet for the increasing need to make research material and data easy to share and re-use.

Publication of research data with CDC will allow scientists to:

- Make their data easy to find and access
- Benefit from the fast publication process
- Contr...

[View full aims & scope](#)

18 days Time to first decision | 24 days Review time | 4 days Acceptance to publication | [View all insights](#)

FEEDBACK