

## Repozitoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

Publishing

# Repozitoriji shrambe znanja

**Vladimir Batagelj**  
UP FAMNIT Koper in IMFM Ljubljana

**1354. and 1355. sredin seminar**  
Ljubljana, November 13 and 20, 2024

# Outline

Repositoriji

V. Batagelj

Prevod

Razvoj

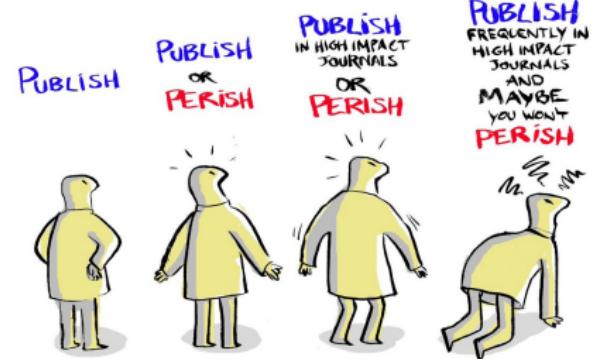
Digitalization

Repositories

Publishing

- 1 Prevod
- 2 Razvoj
- 3 Digitalization
- 4 Repositories
- 5 Publishing

## THE EVOLUTION OF ACADEMIA



[facebook.com/pediatrics](http://facebook.com/pediatrics)

Vladimir Batagelj: [vladimir.batagelj@fmf.uni-lj.si](mailto:vladimir.batagelj@fmf.uni-lj.si)  
Current version of slides (November 20, 2024 at 15:46): [PDF](#)

Prevod

Repozitorij

V. Batagelj

Prevod

Razvoj

Digitalization

## Repositories

Publishing

slovar	primeri
Prevodi za repository v slovarju angleško » slovenski (Skoči na slovensko » angleški)	
<b>re·pos·i·tory</b> [rɪ'pozɪtɔri] SAM, form	i
<b>1.</b> <i>repository (place):</i>	
repository	skladišče n
fig. my diary is a <b>repository</b> for all my secret thoughts	moj dnevnik je zatočišče vseh mojih skritih misli
<b>2.</b> <i>repository (container):</i>	
repository	shramba ž...spol
repository	zbirališče n
repository fig.	vir m...spol
repository fig.	zakladnica ž...spol

# Prevod

## Repozitoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

Publishing

Repozitorij

Skladišče

Zbirka (collection)

**Shramba**

Knjižnica (library)

Zbirališče

Vir

Zakladnica

Zatočišče

Odložišče

Kako bi stvari rekli, če bi si jo sami izmislili?

# Pripombe

Repositoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

Publishing

Predstavitve prvega dela tega seminarja nisem uspel posneti. Zato sam po spominu dodal nekatere dodatke iz ustne predstavitve.

Besede grškega ali latinskega izvora naj bi bile načeloma sprejemljive. Sam bi raje, kjer le gre, ohranjal slovensko osnovo. Razpad pogosto poteka po drobnih delcih.

Slovensko jezikoslovje se (po moje) preveč nagiba na stran beleženja jezikovnih dogajanj in je premalo dejavno v poskusih vplivanja nanje (priporočila, predlogi; ne predpisovanje). Jezik določa njegova raba, a jezikoslovci niso le opazovalci temveč tudi del uporabnikov.

# Ohranjanje znanja skozi čas

Repozitoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

Publishing

- iz roda v rod; vedenje skupnosti
- Grki razpravljanje
- izum pisave; prenos znanja v prostoru in času (jezik)
  - zapisi posamičnih vedenj in znanj: astronomska opažanja, postopki izračuna
  - povezovanje vedenj, logika, заметки disciplin – Evklid geometrija (Curry)
  - problem obstojnosti – več izvodov zapisa
  - problem razmnoževanja zapisov – prepisovanje
  - problem dostopnosti – knjižnice (Aleksandrija); заметки поšte
  - problem jezikov – prevajanje ali skupni jezik (grščina, latinščina, arabščina, angleščina)
- tekmovanja v reševanju kubične enačbe – Cardano (~1535); dopisovanje med znanstveniki (Pascal, Descartes, ...)

# The Fertile Crescent and Greece

Repozitoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

Publishing



In the Fertile Crescent and Greece, virtually every area of human knowledge was advanced by these people, including: Science and Technology, Writing and Literature, Religion, Agricultural Techniques, Mathematics and Astronomy, Astrology and the Development of the Zodiac, Domestication of Animals, Long-Distance Trade, Medical Practices (including dentistry), The Wheel, The Concept of Time, basic Measurements. [4, 3]

# Pictures

**Repozitoriji**

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

Publishing



Gobekli Tepe, 9600-8200 BC



Nabta Playa, 7500 BC



Cylinder seal, 1900 BC



Plimpton-322, 1800 BC

## Pictures

Repozitoriji

V. Batagelj

Prevod

Razvoj

Digitalization

## Repositories

Publishing



## Papyri Merer, 2589-66 BC



## Phaistos Disc A, 2 mill BC



## Medical instruments, Kom Ombo, 1 cent BC



Rhind papyrus, ~1550 BC

# Pictures

Repozitoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

Publishing



Making and Serving Beer, Hammurabi code, Plato and Aristotle,  
2700-2600 BC



1755–50 BC



~380 BC



Euclid's Elements Arabic, 1270



Antikythera mechanism, 2 cent, BC

# Ohranjanje znanja skozi čas

Repositoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

Publishing

- izum tiska (žigi, Faistos) Biblij, Evklid; bistveno poceni razmnoževanje in s tem dostopnost
- prve revije vezane na društva / akademije (London, Paris, St. Petersburg, ...)
- razvoj industrializacije je tesno povezan z raziskavami in vse širšim izobraževanjem
- pošta, fotoaparat, telegraf, telefon, gramofon, pisalni stroj, film, radio, televizija
- po drugi svetovni vojni založbe prevzemajo in ustanavljajo revije, naročnine
- računalnik, kopirni stroj (Xerox), laserski tiskalnik, multimedija, internet, mobilni telefon
- avtorski izvodi, prednatisi (preprint), poročila

## Pictures

Repozitoriji

V. Batagelj

Prevod

Razvoj

## Digitalization

## Repositories

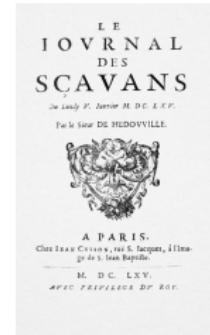
Publishing



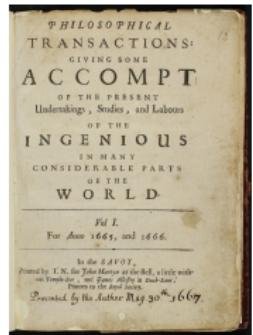
Gutenberg, 1393?–1468



Elements, 1482



Jd sçavans, 1665 PT, 1665



PT, 1665



Remington-2, 1878



Xerox 914, 1959



IBM PC, 1981

# Pictures

Repozitoriji

V. Batagelj

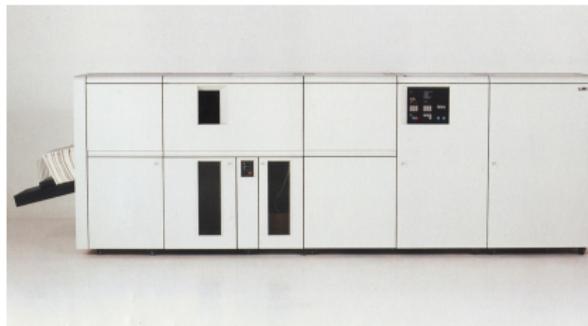
Prevod

Razvoj

Digitalization

Repositories

Publishing



IBM 3800, 1976



Donald E Knuth



The Art + TeX



Apple LaserWriter, 1985



Adobe PostScript

1985



PDF

1992

# Ohranjanje znanja skozi čas

Repozitoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

Publishing

- © za knjige in programsko opremo
- Creative Commons – ureditev pravic za neprogramje.
- Richard Stallman GNU
- “Če je nekaj ustvarjeno za denar davkoplačevalcev, naj bo javno dostopno”. Mosaic → Netscape, Firefox; Explorer
  - SKJ – če želimo spodbuditi razvoj jezikovnih storitev, je pomembno, da so jezikovni podatki prosto dostopni
  - OpenAlex – bibliografske storitve
- V pogodbe o projektih določilo o dostopnosti

# Pictures

Repozitoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

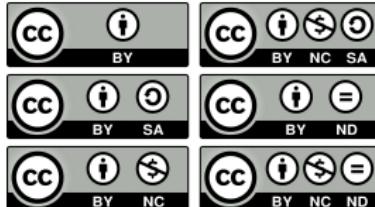
Publishing



Richard Stallman, 2012 GNU, 1989

Lawrence Lessig, 2004

 **creative commons**

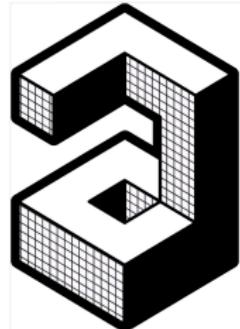


CC, 2001

My **pictures**.



Ars Electronica, 2004



OpenAlex, 2022

# Digitalization

Repozitoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

Publishing

IT has had a profound impact on scientific publications

- digital works are straightforward to duplicate (copyright issues).
- software tools for creating texts (TeX, Word, etc.) allow the author to prepare the work for publication fully.
- works in digital format can exceed printing capabilities (multimedia, code, external links),
- sustainability problem.

Due to the rapid (exponential) growth in the number of new works, the current approach to publishing is becoming inadequate (growth in the number of journals, mega-journals, the problem of obtaining reviewers, etc.).

Formats preserving the structure of textual data: SGML, XML, JSON.

# Avtorske pravice

Repositoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

Publishing

## Copyright (WP), basics; Fair use (WP).

- *Javno / Public Domain, (WP)*
- *Prosto / GNU (Copyleft), The FreeBSD Copyright*
- *Odprto / Open Source Initiative, Open Content, Open Access, Open Educational Resources (OER), SourceForge*
- *Preizkusno / Association of Shareware Professionals, cNet*
- *Tržno / U.S. Copyright Office*

Javna, prosta in odprta gradiva omogočajo – uporabo doma, varnostni vpogled, učenje z vpogledom v 'drobovje', slovenjenje, spremembo ali dopolnitev zmogljivosti in druge prilagoditve, ...

# Vannevar Bush (11.mar 1890 - 28.jun 1974)

Repozitoriji

V. Batagelj

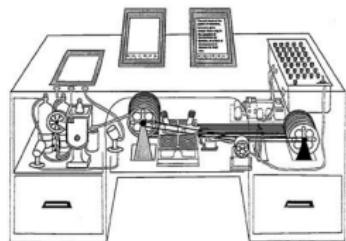
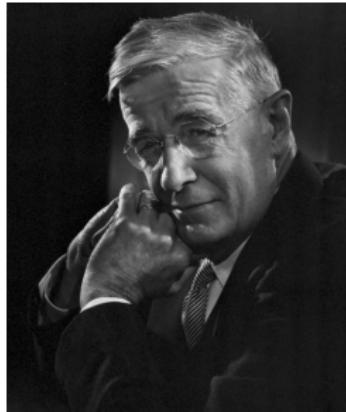
Prevod

Razvoj

Digitalization

Repositories

Publishing



V tridesetih letih je razvil na mikrofilmih temelječ sestav *Memex*, ki je omogočal hitro pregledovanje gradiv.

Leta 1939 je objavil članek '*Mechanization and the Record*', v katerem je opozoril na to, da pristop kazal (indeksov) ne ustreza – "človeško razmišljanje ne deluje tako, s tekoče misli preskoči na naslednjo, z njo povezano v spletu celic človeških možganov". Napovedal je nastanek novih vrst enciklopedij.

Med drugo svetovno vojno je bil vodja za Raziskave in Razvoj pri ameriški vladi. No, že julija 1945 je v mesečniku *Atlantic Monthly* članek objavil v poljudni obliki pod naslovom '*As We May Think*'. Zelo pomembno je tudi njegovo poročilo, na osnovi katerega je bila ustanovljena NSF – *National Science Foundation*.

WP

# Virtual Software Library

Repozitoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

Publishing

In eighties FTP repositories (mainly software) **Simtel**, US, Finland.  
Later C|Net, WinSite, Tucows, Stroud's CWSApps, **4shared**.

C|net was based on **Virtual Software Library** developed by **Ziga Turk**.  
A meta-repository over several repositories (standardization).



## Other resources for searching

If you're looking for freeware or shareware (and they're not the same -- you're supposed to pay for shareware), try the Virtual Software Library, which allows searching and retrieval of more than 130,000 files. The library was started by Professor Ziga Turk at the University of Ljubljana, Slovenia, but has recently moved to a new site at a company that calls itself c|net: the computer network (<http://vsl.cnet.com>). [02 Mar 1997: now called Shareware.Com, <http://www.shareware.com>.]

from EEI Communications, November 1995

digitaltrends

# Repositories

## Repozitoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

Publishing

## Repositories:

- metadata (index; [Dublin Core Metadata](#));
  - works;
  - mixed.
- ① Software repository [c|Net](#), [CPAN](#) (Perl), [CTAN](#) ( $\text{\TeX}$ ), [CRAN](#) (R), [PyPI](#) (Python), [Google Play Apps](#), [SourceForge](#).
  - ② Papers
    - Bibliographies: [Web of science](#), [DBLP](#), [BibNet](#) ([Bib \$\text{\TeX}\$](#) ), [Scopus](#), [OpenAlex](#).
    - Papers: [ArXiv](#), [CiteSeerX](#), [ResearchGate](#), [Google Scholar](#), [Academia](#), [Papers with Code](#), [Bibsonomy](#).
  - ③ Books [Project Gutenberg](#), [Open Library](#), [Open book publishers](#), [Directory of Open Access Books](#), [Wikibooks](#),

# Repositories

## Repozitoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

Publishing

- ④ Data set repository [Kaggle](#), Microsoft Research Open Data, [Google Dataset Search](#), [Open Data Network / ex Socrata](#), [UCI Machine Learning Repository](#), [CERN Open Data](#), [Dryad](#), [Nature Scientific Data Journal](#).
- ⑤ Data Awesome public datasets, [AWS – Amazon Web Services Registry of Open Data](#), [fivethirtyeight](#), [FlowingData](#), [Data is plural](#), [TidyTuesday](#), [Harvard Dataverse](#), [PEW research](#), [European Social Survey](#), [The World Factbook](#), [data.gov](#), [OPSI-Odprti podatki Slovenije](#), [SiStat](#), [ICON](#), [Network Repository](#).
- ⑥ Multimedia [YouTube](#) (video), [Flickr](#) (photo), [Google photos](#) (ex [Picassa](#)), [Slideshare](#) (slides), [NASA](#).
- ⑦ Projektni [GitHub](#), [Figshare](#), [Zenodo](#).
- ⑧ General [Europeana](#).

## Repozitoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

Publishing



arXiv was founded by Paul Ginsparg in 1991 and is now maintained and operated by Cornell Tech. arXiv currently serves the fields of physics, mathematics, computer science, quantitative biology, quantitative finance, statistics, electrical engineering and systems science, and economics. [WP/arXiv](#)

[TechRxiv](#), [ChemRxiv](#), [bioRxiv](#), [medRxiv](#), [PsyArXiv](#), [SocArXiv](#), [HAL](#), [RUL](#), [zenodo](#), etc. [2]

# The number of papers in MAG over time

## Microsoft Academic Graph

Repozitoriji

V. Batagelj

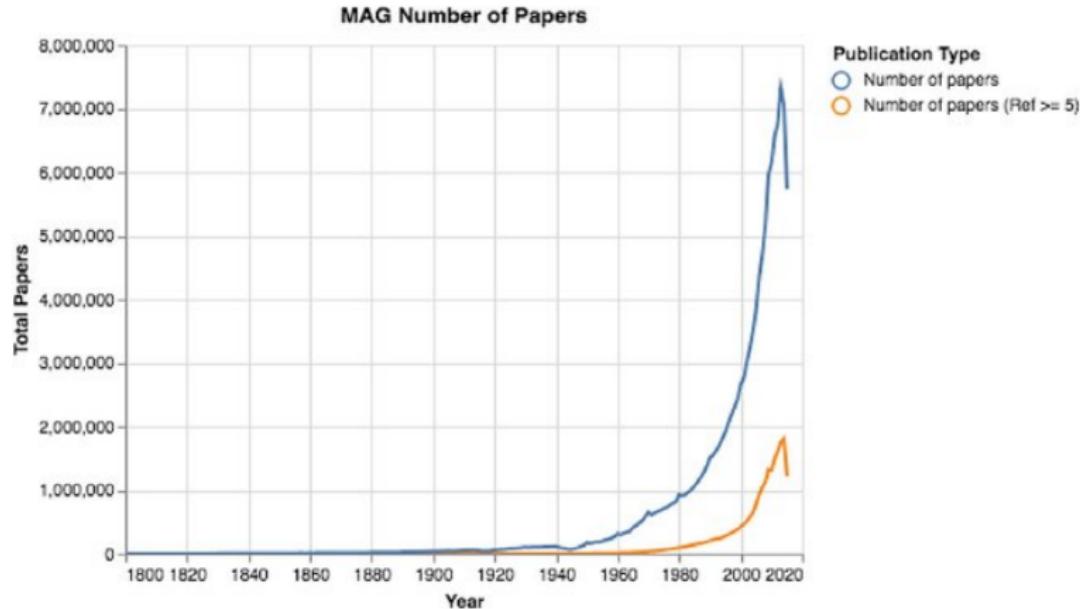
Prevod

Razvoj

Digitalization

Repositories

Publishing



The total number of papers has surged exponentially over the years. From Michael Fire, Carlos Guestrin: Over-optimization of academic publishing metrics: Observing Goodhart's Law in action. *GigaScience* 8(June 2019)6, 1–20.

OpenAlex. R.

# Top (first 10 months of) 2024 authors

LinkedIn

## Repozitoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

Publishing

Author	All Articles
Wiwanitkit, V.	492
Daungsupawong, H.	346
Bruze, M.	336
Api, A.M.	326
Lee, I.	323
Schultz, T.W.	322
Thakkar, Y.	320
Siddiqi, F.	320
Moustakas, H.	320
Lavelle, M.	320
Lapczynski, A.	320
Sullivan, G.	319
Sipes, I.G.	319
Ritacco, G.	319
Joshi, K.	319
Jones, L.	319
Dekant, W.	319
Chon, H.	319
Cancellieri, M.A.	319
Botelho, D.	319
Belsito, D.	319
Fryer, A.D.	313
Niyato, D.	312

Author	Just Articles
Taniguchi, T.	269
Niyato, D.	259
Cao, J.	256
Muhammad, T.	183
Xie, X.	182
Sayyed, M.I.	177
Watanabe, K.	173
Yang, H.J.	166
Yan, H.	163
Nisar, K.S.	160
Liu, J.B.	159
Watanabe, K.	158
Pawlak, T.M.	152
Imran, M.	152
Ahmad, H.	152
Wang, L.	151
Huang, W.	151
Jiao, L.	150
Pan, F.	149
Guizani, M.	148
Basem, A.	146
Ullah, R.	143
Ding, W.	143

Publishing with Integrity

# Papers with the largest number of co-authors

Repositoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

Publishing

Papers with the largest number of authors

ChatGPT listed as author on research papers

The Vancouver recommendations

Reproducibility and Replicability in Science

# Mega-journals and predatory journals

Repozitoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

Publishing

Mega-journals. PLOS / fees. Online Science Publishing.

Beall's list of Potential predatory scholarly open-access publishers.

Top sources (ISSNs) in 2023

List of basic info about each ISSN

Mega journals in OpenAlex (local)

# Decision on submission

Repozitoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

Publishing



Khalid Gul

Assistant Professor, National Institute of Technology,...

Follow

Recently, our manuscript was rejected by a well-regarded journal. The reason? After inviting 28 reviewers... [see more](#)

1 2 3 695 · 10 Comments

[Read more](#)

We regret to tell you we were not able to find the required number of reviewers to evaluate your manuscript; we don't want to delay the process for you any longer.

Decision on submission

# What should #Editors do?

Repozitoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

Publishing



Ronald Marquez • 3rd+

Assistant Professor at Universitat de Girona | PhD in Form...

5d • Edited •

+ Follow

What should #Editors do?

Yang, F. X., Zhu, Y. F., Cao, S., Wang, C. M., Ma, Y. J., Yang, R., & Hu, Q. M. (2024). Origin of the distinct site occupations of H atom in hcp Ti and Zr/Hf. International Journal of Hydrogen Energy, 91, 933-941. <https://lnkd.in/e9c9HV6W>

In this paper, first-principles calculations are performed to explore the physics underlying the distinct site-occupations of H in Ti and Zr/Hf. The solution energies of H in the T and O interstices of Ti, Zr, and Hf are calculated and compared to determine the stable occupation site of H. The lattice structures and electronic structures of the systems are analyzed, with which the distinct site occupations are discussed. The paper is arranged as follows. In Section 2, we describe the method and computational details. In Section 3, we present the solution energy, crystal structure, and electronic structure from the calculations. In Section 4, we discuss the physical origin of the distinct site occupations of H in Ti and Zr/Hf. Finally, we conclude our work in Section 5.

As strongly requested by the reviewers, here we cite some references [[35], [36], [37], [38], [39], [40], [41], [42], [43], [44], [45], [46], [47]] although they are completely irrelevant to the present work.

What should #Editors do?

# Academia holds monopoly over science

## Repozitoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

Publishing



**Swapnonil Banerjee, PhD** • 3rd+

Physicist / Engineer

6d • Edited •

+ Follow ...

Academia holds monopoly over science, as was conveyed in the comment made by Yann LeCun that if you don't publish papers, you are not doing science. I don't understand why. The question is particularly relevant since most PhD holders, with solid research experience, are outside academia. All the students I knew from my PhD days, who I thought were brilliant and for whom I have great respect, are outside academia. I think there should be a system parallel to academia where basic research happens. That parallel system, with its members free from the burden to publish run-of-the-mill papers in order to survive, can make truly significant contribution to science. Besides, the same sense of freedom can help them take science to the masses and to the would-be scientists with a lot more energy than those in academia. Like any other discipline, science requires good leadership, which can come from many talented individuals that are outside academia.

Academia holds monopoly over science

# Mieke Bal: Let's Abolish the Peer-Review System

[Repozitoriji](#)

[V. Batagelj](#)

Prevod

Razvoj

Digitalization

Repositories

Publishing



Mieke Bal: [Let's Abolish the Peer-Review System](#)  
Neil Levy: [Philosophy, Bullshit, and Peer Review](#)

# Scientific publishing

Repositoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

Publishing

## Expectations of the author-scientist

- instant publication and access,
- "registration" – timestamp,
- easy to find, promotion,
- unlimited access,
- permanent storage (longevity, permanence),
- readers' (community) response, recognition.

This can be achieved by prepublishing works in a repository.

## **prepublishing → publishing**

knjige: prodajna mreža, zaslužek

obvezni izvod v varni shrambi; po 70 letih postane javno  
objave nečlankov: programi, podatki, skladbe, slike, itd.

# Open Peer-Review

## Repozitoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

Publishing

“In scholarly publishing peer review is extremely important, but it comes with challenges: for instance, those around anonymity in the review process. One alternative to the widespread double-blind procedure is Open Peer Review, where scholars publicly and openly comment on their colleagues’ work: the manuscript is uploaded on a website and experts in the field are invited to comment within a specific time frame. This process is more transparent, less hierarchical, as well as faster and easier for everyone involved. As is always the case with peer review, criticism must be constructive, and a third party – the publisher for instance – must ensure that is the case for each comment before it becomes public.”

*Rabea Rittgerodt*

Alternative peer review models through the editor’s lens  
Exploring Alternative Models for Enhanced Academic Publishing  
Types, limitations, and possible alternatives of peer review  
Open peer review(WP); F1000; Proceedings RS A

# Improved repositories

## Repozitoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

Publishing

A possible answer is repositories that also support

- response from **registered** readers
  - stars
  - comments (+ response to them)
  - citation (implicit response)
  - tracing the use of work (implicit response)
- list of works citing the work
- powerful repository search (can) considers responses and the response author weights

This data can also be helpful in the evaluation of individual work.

## IFLA DDRS

# Improved repositories / problems

## Repozitoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

Publishing

## AI in shrambe

- splošno: dobri povzetki – izbrana gradiva nam ni potrebno prepbrati, da zvemo za kaj v njih gre
- podrobno: veliko natančneje najde gradiva (iz celotne shrambe), ki nas zanimajo (izboljšano stvarno kazalo)

Duplicated ID (submission in different repositories) problem.

Data quality in a repository. Who can insert, and give an opinion?

Implicit metadata - data obtained from the work itself (number of images) or collected about it by devices (EXIF for photos) or programs (number of material downloads from repository).

Composite works (structure).

Multilingualism.

# Glasovni prevajalnik PRO 6 za 179,00 €

Repositoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

Publishing

**REAL-TIME GLOBAL TRANSLATOR**  
Travel Around the World, Easy to Communicate



The advertisement features a central smartphone displaying the app's interface with various translation modes like Online translation, Offline translation, Smart recording, Photo translation, Drop translation, Language learning, Input translation, and Feature translation. Surrounding the phone are five smaller images illustrating its use in different situations: a woman shopping, a man and child reading a book (Language learning), two business people talking (business), a group of people at a restaurant (Menus), and a signpost (signposts).

Najnovejši Glasovni prevajalnik PRO 6 Odličen za direktno prevajanje 1 na 1. 138 jezikov tudi Slovenski jezik.

Glasovni prevajalnik PRO 6 - Slovenski jezik, odličen za direktno prevajanje 1 na 1. 138 jezikov

# Preprints

[Repositoriji](#)

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

Publishing

dLib: Preprint series

dLib: Preprint series, 2009-2014

COBISS: Preprint series

Jurišić A.: IMFM Preprints

IMFM: Publications

Preprint series (2014, letnik 52, številka 1199)

# Acknowledgments

**Repozitoriji**

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

Publishing

This work is supported in part by the Slovenian Research Agency (research program P1-0294, research program CogniCom (0013103) at the University of Primorska, and research projects J5-2557, J1-2481, and J5-4596), and prepared within the framework of the COST action CA21163 (HiTEc).

# References I

## Repositoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

Publishing



ARIS: [Sprememba metodologij nov24](#)



ARIS: [Zaupanja vredni repozitoriji](#)



Jones, R.T.: Presidential Address: Truth and error in scientific publishing. J. S. Afr. Inst. Min. Metall. vol.115 n.9 Johannesburg Sep. 2015.  
<http://dx.doi.org/10.17159/2411-9717/2015/v115n9a1> [PDF](#)



[World History Encyclopedia](#)