

## Repozitoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

# Repozitoriji shrambe znanja

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

**1354. sredin seminar**  
Ljubljana, November 13, 2024

# Outline

Repozitoriji

V. Batagelj

Prevod

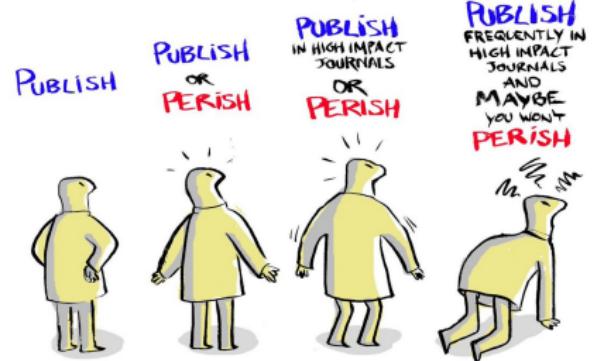
Razvoj

Digitalization

Repositories

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

## THE EVOLUTION OF ACADEMIA



Vladimir Batagelj: [vladimir.batagelj@fmf.uni-lj.si](mailto:vladimir.batagelj@fmf.uni-lj.si)  
Current version of slides (November 13, 2024 at 06:16): [PDF](#)

Prevod

Repozitorij

V. Batagelj

Prevod

Razvoj

Digitalization

## Repositories

slovar	primeri
Prevodi za repository v slovarju angleško » slovenski (Skoči na slovensko » angleški)	
<b>re·posi·tory</b> [rɪ'pozɪtɔri] SAM, form	<span>i</span>
<b>1.</b> <i>repository (place):</i>	
repository	skladišče <i>n</i>
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 <i>ž...spol</i>
repository	zbirališče <i>n</i>
repository fig.	vir <i>m...spol</i>
repository fig.	zakladnica <i>ž...spol</i>

# Prevod

## Repozitoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

Repozitorij

Skladišče

Zbirka (collection)

**Shramba**

Knjižnica (library)

Zbirališče

Vir

Zakladnica

Zatočišče

Odložišče

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

# Ohranjanje znanja skozi čas

Repositoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

- 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



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.

## Pictures

Repozitoriji

V. Batagelj

Prevod

Razvoj

Digitalization

## Repositories



## 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



Papyri Merer, 2589-66 BC



Phaistos Disc A, 2 mill BC



Medical instruments, Kom Ombo, 1 cent BC



Rhind papyrus, ~1550 BC

# Pictures

Repositoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories



Hammurabi code, Plato and Aristotle,  
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

- 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

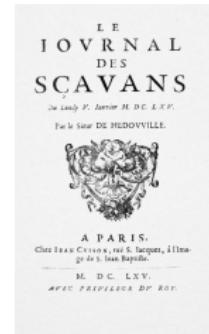
Repositories



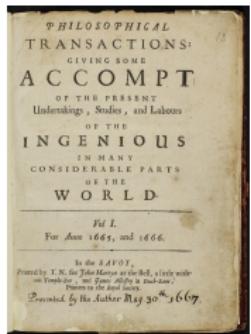
Gutenberg, 1393?–1468



Elements, 1482



Jd ščavans, 1665 PT, 1665



Remington-2, 1878



Xerox 914, 1959



IBM PC, 1981

# Pictures

Repozitoriji

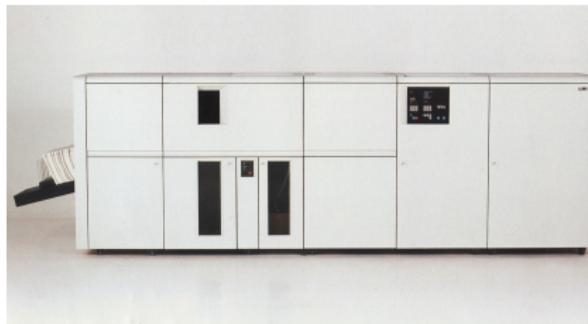
V. Batagelj

Prevod

Razvoj

Digitalization

Repositories



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

- © 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

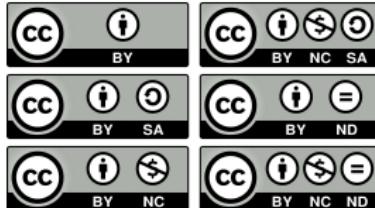
Repositories



Richard Stallman, 2012 GNU, 1989

Lawrence Lessig, 2004

 **creative commons**

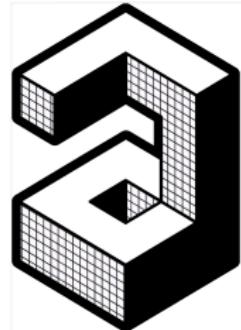


CC, 2001

My **pictures**.



Ars Electronica, 2004



OpenAlex, 2022

# Digitalization

## Repozitoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

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 the capabilities of printing (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.).

# Pravice

Repositoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

Copyright, basics; Fair use.

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

Javna, prosta in odprta gradiva omogočajo – uporaba doma, učenje z vpogledom v 'drobovje', slovenjenje, sprememba ali dopolnitev zmogljivosti in druge prilagoditve, ...

# The number of papers in MAG over time

Repozitoriji

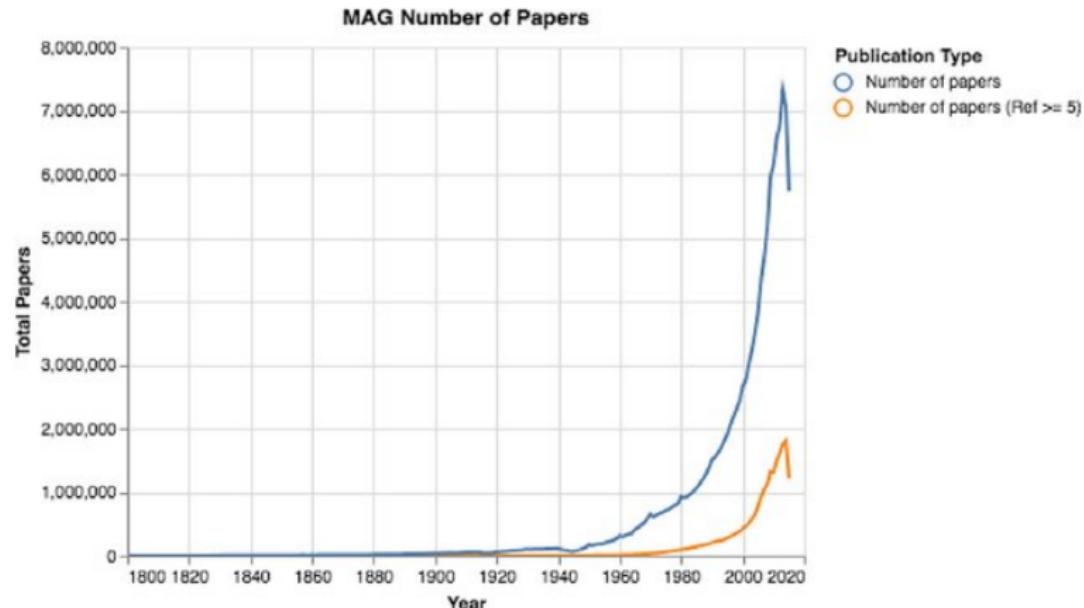
V. Batagelj

Prevod

Razvoj

Digitalization

Repositories



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

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

Repozitoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

## Papers with the largest number of authors

# Mega-journals and predatory journals

Repositoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

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

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

# From LinkedIn

## Repozitoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories



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](#)



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

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

[Repositoriji](#)

[V. Batagelj](#)

Prevod

Razvoj

Digitalization

Repositories



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

# Scientific publishing

Repozitoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

## Expectations of the author-scientist

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

This can be achieved with repositories.

**prepublishing → publishing**

# Virtual Software Library

Repozitoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

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 **Žiga 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

## Repositoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

### 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

- 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

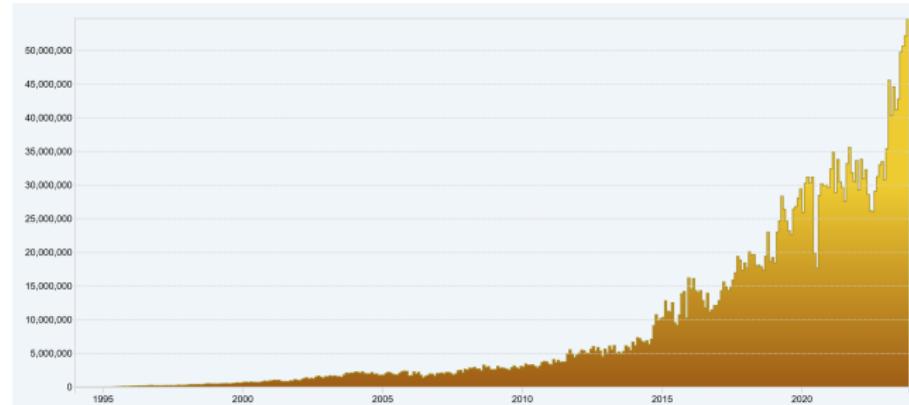
V. Batagelj

Prevod

Razvoj

Digitalization

Repositories



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]

# Improved repositories

## Repozitoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

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.

Duplicated ID (submission in different repositories) problem.

Preprints

Repozitoriji

V. Batagelj

Prevod

Razvoj

Digitalization

## Repositories

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

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 |

## Repozitoriji

V. Batagelj

Prevod

Razvoj

Digitalization

Repositories

-  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**
-  ARIS -- zaupanja vredni repozitoriji
-  World History Encyclopedia