Bibliographic
units
identification

V. Batagelj

Scopus

Resources

Conclusions

References

Bibliographic units identification problem part 3

Vladimir Batagelj

UP FAMNIT Koper in IMFM Ljubljana

1339. sredin seminar

Ljubljana, 15. november 2023

Outline

Vladimir Batagelj: `vladimir.batagelj@fmf.uni-lj.si`
Current version of slides (November 20, 2023 at 23 : 40): PDF
https://github.com/bavla/biblio/

Scopus

My grandson Mark is a handball player. He is also a student of Informatics at the Faculty of Social Sciences. To combine his interests we collected from WoS and Scopus all papers on handball.

He presented his analyses based on WoS data at Sunbelt 2022 and EuSN 2023. We intended to combine the WoS data with the Scopus data by first converting the Scopus descriptions into WoS descriptions and using WoS2Pajek afterward to produce the networks
Scopus2WoS
but the task is not as simple as we expected.

Some time ago, Nataliya Matveeva sent me data on the publications of young universities obtained from Scopus. I was pleasantly surprised to notice that several types of bibliographic units in Scopus have identifiers associated with them. This makes Scopus an interesting resource for building high-quality bibliographic networks.
vladowiki/scopus

It turned out that our decision to base Scopus2WoS on exported RIS files (similar to WoS format) was the worst possible because RIS files don't contain the IDs and references.

Examples of bibliographic files exported from Scopus Truss.ris and Truss.csv on the topic of k-truss.

Scopus
Unusual distributions in Nataliya's data
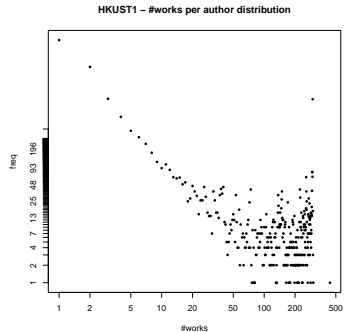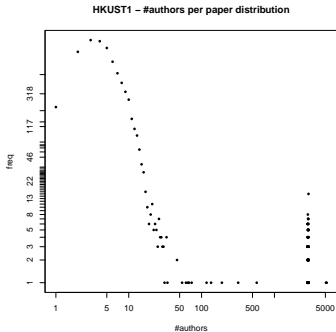
The Hong Kong University of Science and Technology (HKUST) 2017–2019

5215 co-authors/participation in very large international projects.: The ATLAS collaboration, & The CMS collaboration (2019). Combinations of single-top-quark production cross-section measurements and $|f\ LV\ Vtb|$ determinations at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS experiments. Journal of High Energy Physics, 2019(5), Article 88. https://doi.org/10.1007/JHEP05(2019)088 / Springer

Multi-persons or hyperproductive labs?

A selection of some additional resources on the web that may come in handy in solving the problem of identifying bibliographic units.

DBLP computer science bibliography.
DBLP authors; How does dblp handle homonyms and synonyms?; dblp and ORCID 2020; Name disambiguation suffixes in dblp

Similar for mathematics **MathSciNet** and **zbMATH-Open**.
The Identification of Authors in the MR Database+2011; Author disambiguation.

VIAF Virtual International Authority File.
VIAF authors

Wikidata the free knowledge base.
Wikidata authors; How does Wikidata work?

schema.org is a collaborative, community activity with a mission to create, maintain, and promote schemas for structured data on the web.
Person; Book;

Crossref makes research objects easy to find, cite, link, assess, and reuse.
crossref/rest-api; Crossref community

SCImago is both a Spanish consulting company (Scimago Lab) and a network of research groups.
Scimago journals; Assessing universities and other research-focused institutions

Citations from **Google** – use "
Google scholar; Google books; Citing

Conclusions

- importance of unique identifiers

- different publication cultures in different disciplines

Acknowledgments

Bibliographic
units
identification

V. Batagelj

Scopus

Resources

Conclusions

References

Anne-Wil Harzing: harzing.com

Anne-Wil Harzing: Health warning: Might contain multiple personalities. The problem of homonyms in Thomson Reuters Essential Science Indicators. Scientometrics 105(3):2259-2270. paper