



OpenAlex2Pajek

V. Batagelj

The saturation
approach

Institutions

Co-authorship
between
countries

OpenAlex2Pajek

Institutions, the saturation approach, and co-authorship between countries

Vladimir Batagelj
IMFM, UP IAM

1351. sredin seminar
Ljubljana, May 22, 2024

OpenAlex2Pajek

V. Batagelj

The saturation
approach

Institutions

Co-authorship
between
countries

- 1 The saturation approach
- 2 Institutions
- 3 Co-authorship between countries



Vladimir Batagelj: vladimir.batagelj@fmf.uni-lj.si

Current version of slides (May 22, 2024 at 05:33): [slides PDF](#)

<https://github.com/bavla/OpenAlex>

We continue the development of support for the conversion of **OpenAlex** data into Pajek's networks.

The saturation approach was split into two phases:

- the saturation phase dealing only with the citation network for the selection of the set of relevant works W
- creation of bibliographic networks for the selected set of relevant works W



OpenAlex2Pajek

saturation approach

OpenAlex2Pajek

V. Batagelj

The saturation
approach

Institutions

Co-authorship
between
countries

The set W is determined iteratively using the function `OpenAlex2PajekCite`.

- 1 Create the basic query Q and determine using `OpenAlex2PajekCite` the initial version of W ; list of old candidates C is empty
- 2 Analyze using Pajek macro `expNodes` the obtained citation network and identify new candidates N for relevant works. If N is empty **STOP**.
- 3 Save the list N in a CSV file. Using in R the command `joinLists("Cold.csv", "N.csv", "Cnew.csv")` join the old candidates and new candidates into the current list of candidates (removing duplicates).
- 4 Using `OpenAlex2PajekCite` determine the new version of W ; go to 2.

Creating a collection



OpenAlex2Pajek

saturation approach

OpenAlex2Pajek

V. Batagelj

The saturation
approach

Institutions

Co-authorship
between
countries

To create the collection we first change the parameter select in the query Q to selAll. Afterward, we run the function OpenAlex2PajekAll.

Currently, we get a collection of bibliometric networks:

```
>>> n Citation Cite
>>> c publication year
>>> c type of publication
>>> c language of publication
>>> c cited by count
>>> c countries distinct count
>>> c referenced works
>>> n Authorship WA
>>> n Sources WJ
>>> n Keywords WK
>>> n Countries WC
```



OpenAlex2Pajek

saturation approach

OpenAlex2Pajek

V. Batagelj

The saturation
approach

Institutions

Co-authorship
between
countries

I don't like the keywords provided by OpenAlex. In a future version of OpenAlex2Pajek I will provide an alternative based on words from the work's title (and abstract).

In phase one we could consider also other available properties of nodes (works).

On the **to do** list is to remove the use of Pajek from phase one and program the iterations in R.



OpenAlex2Pajek

converting dictionary into data frame

OpenAlex2Pajek

V. Batagelj

The saturation
approach

Institutions

Co-authorship
between
countries

Internal: **dict2DF**

```
dict2DF <- function(dict,ind) {  
  DF <- as.data.frame(do.call(rbind, as.list(dict)))  
  return(DF[order(unlist(unname(DF[[ind]])))]])  
}
```



OpenAlex2Pajek

Institutions

OpenAlex2Pajek

V. Batagelj

The saturation
approach

Institutions

Co-authorship
between
countries

In some cases, such as all works of researchers from a selected institution, the saturation phase is not needed.

Internal: **Young universities**

GitHub: **HKUST, IMFM**

OpenAlex2Pajek

IMFM co-authorship link cut at level 7

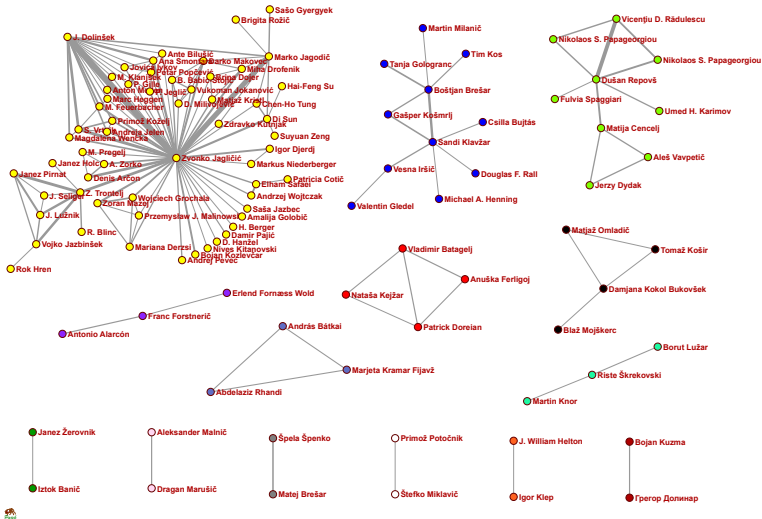
OpenAlex2Pajek

V. Batagelj

The saturation approach

Institutions

Co-authorship
between
countries



V. Batagelj

OpenAlex2Pajek



OpenAlex2Pajek

Co-authorship between countries

OpenAlex2Pajek

V. Batagelj

The saturation
approach

Institutions

Co-authorship
between
countries

We developed a function `coAuthorship` that creates a sequence of temporal networks describing the co-authorship between world countries.

Internal: **meaning**; **total**; **years**

GitHub: **pics**; **1-neighbors**

Problem: OpenAlex is using ISO 2-character country codes. Only currently existing countries are considered.



OpenAlex2Pajek

Total co-authorship between countries/1-neighbors

OpenAlex2Pajek

V. Batagelj

The saturation
approach

Institutions

Co-authorship
between
countries

