# Bibliometric analysis

## of a scientific journal based on OpenAlex data

**Vladimir Batagelj**
IMFM, UP IAM, FMF

**Applied statistics**
Koper, 21-23. september 2025

# Outline







**Vladimir Batagelj**: vladimir.batagelj@fmf.uni-lj.si
**Current version of slides (September 21, 2025 at 02:11):** slides PDF
https://github.com/bavla/OpenAlex

# OpenAlex2Pajek



OpenAlex is a fully open catalog of the global research system. It's named after the ancient Library of Alexandria and was made by the nonprofit OurResearch. OpenAlex launched in January 2022 with a free API and data snapshot [1]. It is a free alternative to commercial bibliographic services such as Web of Science and Scopus. Through its API, it provides programming access to bibliographic data and enables complex analyses and the development of higher-order bibliographic services.

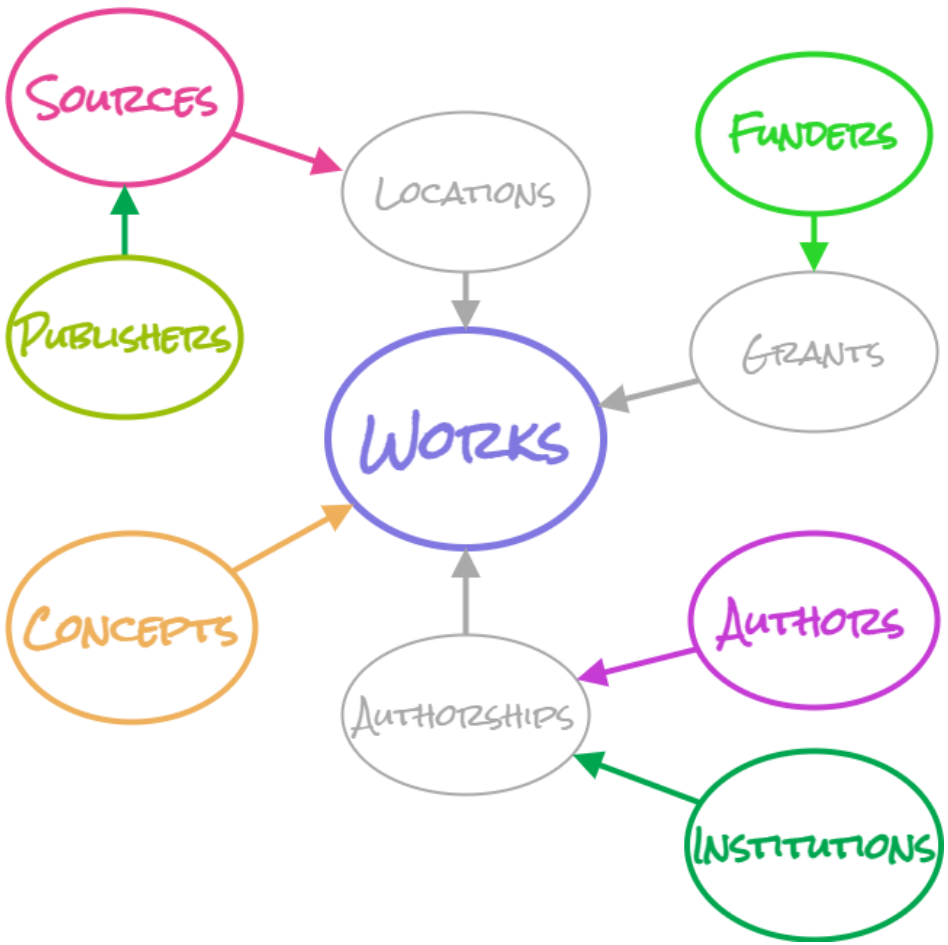
We are developing an R package, `OpenAlex2Pajek`, for creating bibliographic networks from the OpenAlex database. The basic package version supports the collection of data on selected topics.

Usually, the first step is to prepare a list of interesting works. In the second step, we create corresponding bibliographic networks for the works from this list. We continue with the analysis of the obtained networks.

# OpenAlex

## Types of bibliographic units

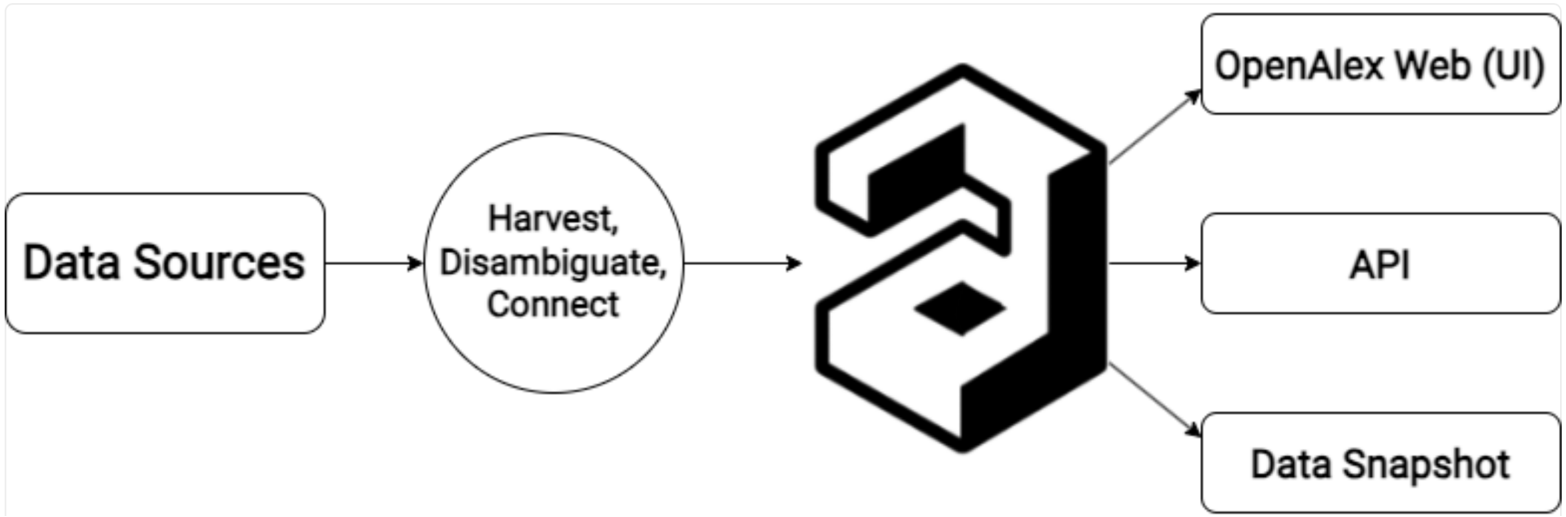
OpenAlex is based on 7 types of units (entities): **W**(ork), **A**(uthor), **S**(ource), **I**(nstitution), **C**(oncept), **P**(ublisher), or **F**(under)



# OpenAlex

## Comments






OpenAlex solves several important questions for the analysis of bibliographic data:

1. identification of bibliographic units (IDs, disambiguation)
2. free access (share derived data, Download to your machine)
3. improving content through user participation (Submit a request)

# OpenAlex2Pajek / Sources

In this contribution, we present an extension, the function `OpenAlexSources`, that creates networks related to a selected journal (all papers published by the chosen journal and all works citing/cited by these papers). Since in networks the units (works, authors, sources, keywords, etc.) are identified by their OpenAlex IDs, another function, `unitsInfo`, provides the user with additional information about the units appearing in the results of analyses.

We applied the new functions to create bibliographic networks for the journals Metodološki zvezki – Advances in Methodology and Statistics (S4210169332, AMS) and Ars Mathematica Contemporanea (S61442588, AMC).

# Collecting works

If you open these examples in a web browser, they will look much better if you have a browser plug-in such as JSONVue installed.

**A**. List of works published by a given source (journal) (call) – in a single line

```
https://api.openalex.org/works?
filter=primary_location.source.id:S4210169332&
select=id,title,type,cited_by_count,publication_year
```

**B**. List of citing works of a given work (call)

```
https://api.openalex.org/works?filter=cites:W4206962290&
select=id,title,type,cited_by_count,publication_year&
per_page=200&page=1
```

**C**. List of works cited by a given work (call)

```
https://api.openalex.org/works?filter=openalex:
W4205437711|W4206962290|W2096252182|W4206003933&
select=id,title,publication_year,referenced_works
```

# Collecting works

## Creating the set of relevant works W and networks

**D**. Let $j$ be the selected source (journal). Determine (**A**) the set $W_j$ of works published in the journal $j$. Now we can determine

- the set $W_{in}$ of works citing some work from $W_j$ – for each $k \in W_j$ determine (**B**) the set $W_k$ of works citing the work $k$. The set $W_{in}$ is the union of all $W_k$s.
- the set $W_{out}$ of works cited from some work from $W_j$ – for each $k \in W_j$ determine (**C**) the set $W_k$ of works cited by the work $k$. The set $W_{out}$ is the union of all $W_k$s.
- the set of relevant works is $W = W_{in} \cup W_j \cup W_{out}$. To get networks, apply the procedure `OpenAlex2PajekAll` on $W$.

Note that for sources different from $j$ only the citations from/to $j$ are complete. Other citations consider only cases where at least one end-node is related to a work from the source $j$. The obtained networks can be used to determine the set of important sources $J$.

# Collecting works
# Creating the set of relevant works W and networks

**E**. For each important source $j$ from $J$, we determine (**D**) the corresponding set of relevant works. The union of these sets $W_J$ is used in the procedure `OpenAlex2PajekAll` to create networks. Now, the citation data are complete for all sources from $J$ (but not for the other sources).

The size of the set $W_J$ can be very large. To reduce it, we can consider some restrictions, such as the interval of considered years of publication, the type of publication, etc.

The programming of support for the collection of the selected source data resulted in two functions OpenAlexSources and unitsInfo.

# Collecting works

## Creating the set of relevant works W and networks

To build networks for a selected source `sID` is now simple. First, we create a vector $R$ of all works from `sID`, works citing them, and works cited by them. -

```
> setwd(wdir <- "C:/test/OpenAlex/sources")
> library(httr); library(jsonlite)
> source("https://raw.githubusercontent.com/bavla/Rnet/master/R/Pajek.R")
> source("https://raw.githubusercontent.com/bavla/OpenAlex/main/code/OpenAlex2Pajek.R")
> sID <- "s4210233660"
> R <- OpenAlexSources(sID,step=250)
OpenAlex2Pajek / Sources Tue May  6 19:32:47 2025
...
2522 source s4210233660 works collected Tue May  6 19:32:50 2025
...
 4092 citing works collected Tue May  6 19:38:59 2025
..........
 14515 cited works collected Tue May  6 19:39:10 2025
17642 different works Tue May  6 19:39:10 2025
```

We save the vector R in a file.

```
> csv <- file("worksTest.csv","w",encoding="UTF-8")
> write(R,sep="\n",file=csv)
> close(csv)
```

To get the networks, we apply 'OpenAlex2PajekAll' to R

```
> setwd(wdir <- "C:/Users/vlado/docs/papers/2025/AS/MZ")
> library(httr); library(jsonlite)
> source("https://raw.githubusercontent.com/bavla/Rnet/master/R/Pajek.R"
> source("https://raw.githubusercontent.com/bavla/OpenAlex/main/code/Ope
> sID <- "S4210169332"
> R <- OpenAlexSources(sID,step=250)
OpenAlex2Pajek / Sources Mon May 26 05:08:49 2025
  238 source S4210169332 works collected Mon May 26 05:08:50 2025
 1423 citing works collected Mon May 26 05:10:29 2025
 4490 cited works collected Mon May 26 05:10:31 2025
 5323 different works Mon May 26 05:10:31 2025
> csv <- file("worksMZ.csv","w",encoding="UTF-8")
> write(R,sep="\n",file=csv)
> close(csv)
> OpenAlex2PajekAll(NULL,name="MZ",listF="worksMZ.csv")
OpenAlex2Pajek / All - Start Mon May 26 05:12:37 2025
*** OpenAlex2Pajek / All - Process Mon May 26 05:12:37 2025
...
*** OpenAlex2Pajek / All - Data Collected Mon May 26 05:46:36 2025
hits: 5323 works: 157256 authors: 10268 anon: 120 sources: 1776
>>> Citation Cite
>>> publication year, type of publication, language of publication
>>> cited by count, countries distinct count, referenced works
>>> Authorship WA, Sources WJ, Countries WC, Keywords WK
*** OpenAlex2Pajek / All - Stop Mon May 26 05:47:50 2025
```

# Analysis

## Works citing/cited_by the journal MZ

We first clean the networks **Ci**, **WA**, **WJ**, . . ., removing multiple links and loops.

```
Network/Create new network/Transform/Remove/Multiple lines [Single line]
Network/Create new network/Transform/Remove/Loops [No]
```

$|W| = 157256, |J| = 1776, |A| = 10268$, ..., $m_{Ci} = 233191$, $m_{WJ} = 5294$, $m_{WA} = 13491$,... Initially, in the citation network, there were 1481615 arcs.
The product $\mathbf{u} = \mathbf{A} \cdot \mathbf{v}$ of the network $\mathbf{A}$ with the vector $\mathbf{v}$ is defined as

$$u_i = \sum_{j:(i,j)\in L} A_{ij} \cdot v_j$$

We need the index $j$ of the node representing MZ in the set of journals J. We can get it from $\mathbf{JW} = \mathbf{WJ}^T$

```
select WJ as the First network
Network/2-mode/Transpose
File/Network/Change label [JW]
```

# Analysis

## Works citing/cited_by the journal MZ

We apply the command `Info/Vertex label -> Vertex number [S4210169332]` on the network **JW**. We get $j = 142$ – the index of the node representing MZ.
We start with the set $W_j$ of all works published by the journal $j$.

$$W_j = \{w : WJ[w, j] > 0\}$$

Let $\mathbf{w}_j$ be its characteristic vector. Then $\mathbf{w}_j = \mathbf{WJ} \cdot [j]$ where $[j]$ is a vector over $J$ having 1 at the $j$th place. We create the vector $[j]$

```
Vector/Create constant vector [1776, 0]
using vector editor (magnifying glass icon for vectors),
  change the 142nd value to 1
File/Vector/Change label [ [j] ]
select WJ as the First network
Operations/Network+Vector/Network*Vector
File/Vector/Change Label [Wj]
```

# Analysis

## Works citing/cited_by the journal MZ

Next, for the set $W_j$, we determine the set $W_I$ of citing works and the set $W_O$ of cited works.

$$W_I = \{w : \exists z \in W_j : Ci[w, z] > 0\} \quad \text{and}$$

$$W_O = \{w : \exists z \in W_j : Ci[z, w] > 0\}$$

The vectors $\mathbf{d}_I = \mathbf{Ci} \cdot \mathbf{w}_j$ and $\mathbf{d}_O = \mathbf{Ci}^T \cdot \mathbf{w}_j$

$$d_I(i) = \sum_k Ci[i, k] \cdot w_j(k) \quad \text{and}$$

$$d_O(i) = \sum_k Ci^T[i, k] \cdot w_j(k) = \sum_k Ci[k, i] \cdot w_j(k)$$

count: $d_I(i)$ - how many works from $W_j$ are cited by the work $i$; and $d_O(i)$ - how many works from $W_j$ are citing the work $i$.

```
select Ci as the First network
select Wj as the First vector
Operations/Network+Vector/Network*Vector
File/Vector/Change Label [dI]
```

# Analysis

## Works citing/cited_by the journal MZ

Inspect the vector *dI*. We list the largest 20 nodes [+20]. Extract the selected top lines and copy them in a text file in TextPad. Remove the Rank and Vertex columns. We get a table
4.0000 W2995900225
...

Save it to a CSV file `dI.csv`. We will collect in R from OpenAlex the additional information about the selected works.
It turns out that the authors' names are not directly accessible as a data field – they are contained inside the field `authorships`. To extract them, we use the function

```
authors <- function(L) {
  A <- L$authorships; k <- length(A); N <- rep("",k)
  for (i in 1:k) N[i] <-
    paste(A[i][[1]]$author$display_name,collapse=", ")
  return(N)
}
```

# Analysis

## Works citing/cited_by the journal MZ

Now we are ready to get the information about the selected works. Some data (`authors` and `title`) can be very long. To get a readable report we truncate them.

```
> LI <- read.table("dI.csv",head=FALSE,sep="")
> selW <- paste0("id,language,countries_distinct_count,cited_by_count,",
+    "relevance_score,publication_year,title,authorships")
> RW <- unitsInfo(IDs=LI$V2,units="works",select=selW,order="input")
> rep <- data.frame(id=RW$id,cdc=RW$countries_distinct_count,
+ cby=RW$cited_by_count,dI=LI$V1,year=RW$publication_year,
+ authors=substr(authors(RW),1,35),title=substr(RW$title,1,45))
> rep
```

# The top works citing the journal MZ

| | id | cdc | cby | dI | year | authors | title |
|---|---|---|---|---|---|---|---|
| 1 | W2995900225 | 3 | 3 | 4 | 2019 | C Nordlund, A Žiberna | Blockmodeling of Valued Networks |
| 2 | W2184576020 | 1 | 9 | 4 | 2009 | T Kogovšek, V Hlebec | Stability of typologies produced on the |
| 3 | W2240460603 | 3 | 15 | 4 | 2015 | L Prota, P Doreian | Finding roles in sparse economic hierarc |
| 4 | W2185659002 | 2 | 2 | 4 | 2012 | V Hlebec, T Kogovšek, + | Measurement quality of social support su |
| 5 | W4205564009 | 0 | 0 | 4 | 2012 | A Žnidaršič | Impact of fixed choice design on blockmo |
| 6 | W4205097511 | 1 | 2 | 3 | 2008 | T Kogovšek, V Hlebec | Measuring ego-centered social networks |
| 7 | W2807932250 | 1 | 0 | 3 | 2018 | JHP Hoffmeyer-Zlotnik, + | Introduction, Problem, and Research Ques |
| 8 | W4400362431 | 0 | 0 | 3 | 2024 | R González, E Aedo-Muño | Nonresponse in name generators across co |
| 9 | W4206030688 | 1 | 1 | 3 | 2010 | T Kogovšek, M Mrzel, + | "Please name the first two people you wo |
| 10 | W3129803968 | 3 | 4 | 3 | 2021 | M Cugmas, A Ferligoj, + | Global structures and local network mech |
| 11 | W2498378275 | 1 | 2 | 3 | 2011 | V Hlebec, M Mrzel, + | The Comparability of Event-Related and G |
| 12 | W2166296097 | 1 | 9 | 3 | 2011 | V Hlebec, M Mrzel, + | Assessing social support networks in cro |
| 13 | W2071989778 | 1 | 98 | 3 | 2008 | V Vehovar, KL Manfreda | Measuring ego-centered social networks c |
| 14 | W4206130755 | 1 | 2 | 3 | 2011 | V Hlebec, T Kogovšek | How (not) to measure social support netw |
| 15 | W2083653613 | 1 | 114 | 3 | 2014 | DE Eagle, RJ Proeschold | Methodological considerations in the use |
| 16 | W2024234381 | 1 | 27 | 3 | 2012 | V Hlebec, T Kogovšek | Different approaches to measure ego-cent |
| 17 | W2995396414 | 3 | 5 | 2 | 2019 | A Žnidaršič, P Doreian,+ | An Treating Missing Network Data Before |
| 18 | W2335008118 | 1 | 189 | 2 | 2016 | A Mrvar, V Batagelj | Analysis and visualization of large netw |
| 19 | W2741577583 | 1 | 2 | 2 | 2017 | D Döring, BE Haberla | Consistency in behavior: Evaluation of b |
| 20 | W2308243582 | 1 | 2 | 2 | 2014 | Selvarangam | SELECTING PERFECT INTERESTINGNESS MEASUR |

# Analysis

## Works citing/cited_by the journal MZ

Some improvements: add source; in names, use only the last name, or initials + last name; ...
We can check selected works – for example W3129803968.

```
Network/Create new network/Transform/Transpose [yes]
File/Network/Change label [CiT]
select Wj as the First vector
Operations/Network+Vector/Network*Vector
File/Vector/Change Label [dO]
```

Using the same approach as for $\mathbf{d}_I$ we get

# The top works cited by the journal MZ

| | id | cdc | cby | d0 | year | authors | title |
|---|---|---|---|---|---|---|---|
| 1 | W2582743722 | 0 | 303929 | 19 | 2014 | R Core Team | R: A language and environment for st |
| 2 | W2061901927 | 1 | 18568 | 10 | 1994 | S Wasserman, K Faust | Social Network Analysis: Methods and |
| 3 | W4285719527 | 0 | 0 | 9 | 1955 | | Deleted Work |
| 4 | W977705565 | 2 | 473 | 9 | 2004 | P Doreian, V Batagelj, + | Generalized Blockmodeling |
| 5 | W2023723604 | 2 | 128 | 9 | 1992 | V Batagelj, A Ferligoj, + | Direct and indirect methods for stru |
| 6 | W2116814842 | 3 | 68 | 8 | 2002 | T Kogovšek, A Ferligoj, + | Estimating the reliability and valid |
| 7 | W1987455866 | 1 | 258 | 7 | 1973 | PW Holland, S Leinhardt | The structural implications of measu |
| 8 | W2017099446 | 1 | 1642 | 7 | 1971 | F Lorrain, HC White | Structural equivalence of individual |
| 9 | W2001947224 | 2 | 105 | 7 | 1992 | V Batagelj, P Doreian, + | An optimizational approach to regula |
| 10 | W4205589890 | 0 | 29 | 6 | 2004 | K L Manfreda, V Vehovar | Collecting ego-centred network data |
| 11 | W2182840202 | 0 | 15 | 6 | 2005 | V Hlebec, T Kogovšek | Hypothetical versus actual support p |
| 12 | W2186748749 | 0 | 16 | 6 | 2006 | P Ziherl, H Iglič, + | Research Groups' Social Capital: A C |
| 13 | W1893868194 | 0 | 14 | 6 | 2006 | L Coromina Soler | Social networks and performance in k |
| 14 | W2151243887 | 1 | 82 | 6 | 2006 | A Žiberna | Generalized blockmodeling of valued |
| 15 | W2054720216 | 1 | 112 | 6 | 1999 | A Ferligoj, V Hlebec | Evaluation of social network measure |
| 16 | W1556604050 | 1 | 1120 | 6 | 1989 | A Vaux | Social support: theory, research, an |
| 17 | W2133011836 | 1 | 361 | 6 | 1983 | DR White, K Reitz | Graph and semigroup homomorphisms on |
| 18 | W1981385379 | 2 | 201 | 6 | 2004 | P Doreian, V Batagelj, + | Generalized blockmodeling of two-mod |
| 19 | W2109278577 | 1 | 1404 | 6 | 1999 | V Batagelj, A Mrvar | Pajek - Program for Large Network An |
| 20 | W1873057782 | 1 | 4378 | 5 | 2000 | R Tourangeau, LJ Rips, + | The Psychology of Survey Response |

# The top authors by the number of works in the journal MZ

$$\mathbf{a}_j = \mathbf{W}\mathbf{A}^T \cdot \mathbf{w}_j$$

$a_j(a) = \#$ of works in the journal $j$ co-authored by the author $a$.

| | id | orcid | wc | cbc | papers | name |
|---|---|---|---|---|---|---|
| 1 | A5038897789 | 0000-0002-3691-7959 | 160 | 2391 | 18 | Valentina Hlebec |
| 2 | A5049753566 | 0009-0001-4355-8608 | 60 | 642 | 14 | Tina Kogovšek |
| 3 | A5029499420 | 0000-0002-3682-6854 | 139 | 2740 | 12 | Anuška Ferligoj |
| 4 | A5039511070 | 0000-0002-5204-6882 | 153 | 3696 | 10 | Germà Coenders |
| 5 | A5040950908 | <NA> | 65 | 2024 | 8 | Katarina Košmelj |
| 6 | A5083575454 | 0000-0002-3253-7959 | 145 | 4129 | 6 | Vasja Vehovar |
| 7 | A5023248667 | <NA> | 59 | 217 | 6 | Uwe Warner |
| 8 | A5010863389 | 0000-0003-1534-6971 | 41 | 443 | 6 | Aleš Žiberna |
| 9 | A5068940001 | <NA> | 51 | 3063 | 5 | Katja Lozar Manfreda |
| 10 | A5025019965 | 0000-0001-6461-3007 | 86 | 661 | 5 | Rosalia Castellano |
| 11 | A5002890522 | 0000-0001-7851-6216 | 109 | 678 | 4 | Jana Mali |
| 12 | A5044693419 | 0000-0003-0769-0633 | 79 | 1179 | 4 | Lluís Coromina |
| 13 | A5019207040 | 0000-0002-2564-8781 | 70 | 3835 | 4 | Janez Stare |
| 14 | A5033311124 | 0000-0001-8557-4692 | 100 | 8835 | 4 | Andrej Mrvar |
| 15 | A5001676164 | 0000-0002-0240-9446 | 271 | 13374 | 4 | Vladimir Batagelj |
| 16 | A5046373528 | 0000-0002-2248-1517 | 68 | 376 | 4 | Irena Ograjenšek |
| 17 | A5041301436 | 0000-0003-3069-9863 | 66 | 743 | 4 | Nataša Kejžar |
| 18 | A5022627222 | <NA> | 160 | 1316 | 4 | Jürgen H. P. Hoffmeyer-Zlotnik |
| 19 | A5052875930 | 0000-0001-7906-0580 | 1024 | 13664 | 4 | Darío Gregori |
| 20 | A5102781233 | 0000-0002-5550-7007 | 131 | 401 | 4 | Malgorzata Graczyk |
| 21 | A5025045918 | 0000-0002-5395-1593 | 113 | 280 | 4 | Bronislaw Ceranka |
| 22 | A5011534481 | <NA> | 28 | 102 | 3 | Anton Cedilnik |
| 23 | A5065490876 | 0000-0002-3301-7840 | 182 | 5385 | 3 | Patrick Doreian |
| 24 | A5055592225 | 0000-0001-6861-9553 | 68 | 579 | 3 | Gennaro Punzo |
| 25 | A5084724910 | 0000-0003-4261-8928 | 89 | 1241 | 3 | Giuseppe Scandurra |

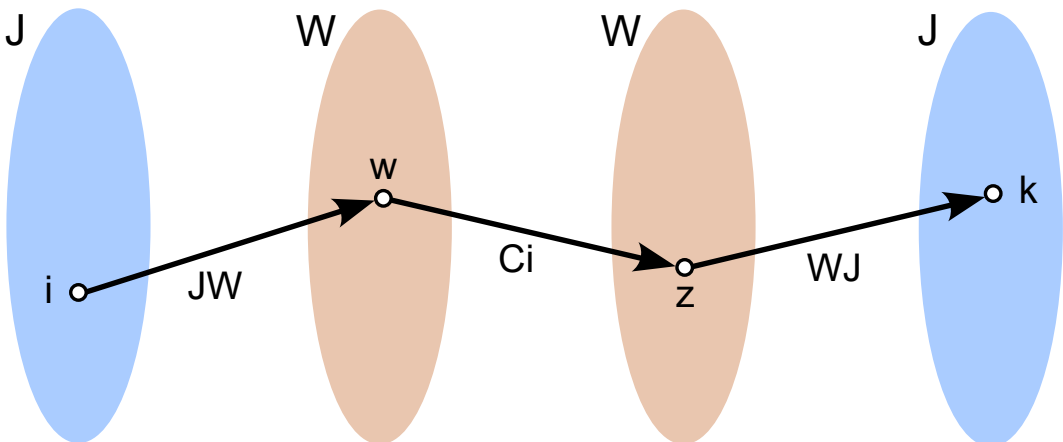
# Analysis

## Citations between authors (and journals)

[2–4]

$$\mathbf{ACiA} = \mathbf{WA}^T \cdot \mathbf{Ci} \cdot \mathbf{WA}$$

$ACiA[a, b]$ = # of times author $a$ cites author $b \equiv$
# of citations of a work of author $a$ to a work of author $b$.




$$\mathbf{JJ} = \mathbf{WJ}^T \cdot \mathbf{Ci} \cdot \mathbf{WJ}$$

$JJ[i, k]$ = # of times journal $i$ cites journal $k \equiv$
# of citations of a work from journal $i$ to a work from journal $k$.

# Analysis

## Citations between authors (and journals)

```
select Ci as the First network
Network/Create new network/Transform/1-mode to 2-mode
select WA as the First network
Network/2-mode/Transpose
File/Network/Change label [AW]
select 2-mode Ci as the Second network
Networks/Multiply networks
select WA as the Second network
Networks/Multiply networks [yes]
File/Network/Change label [ACiA]
```

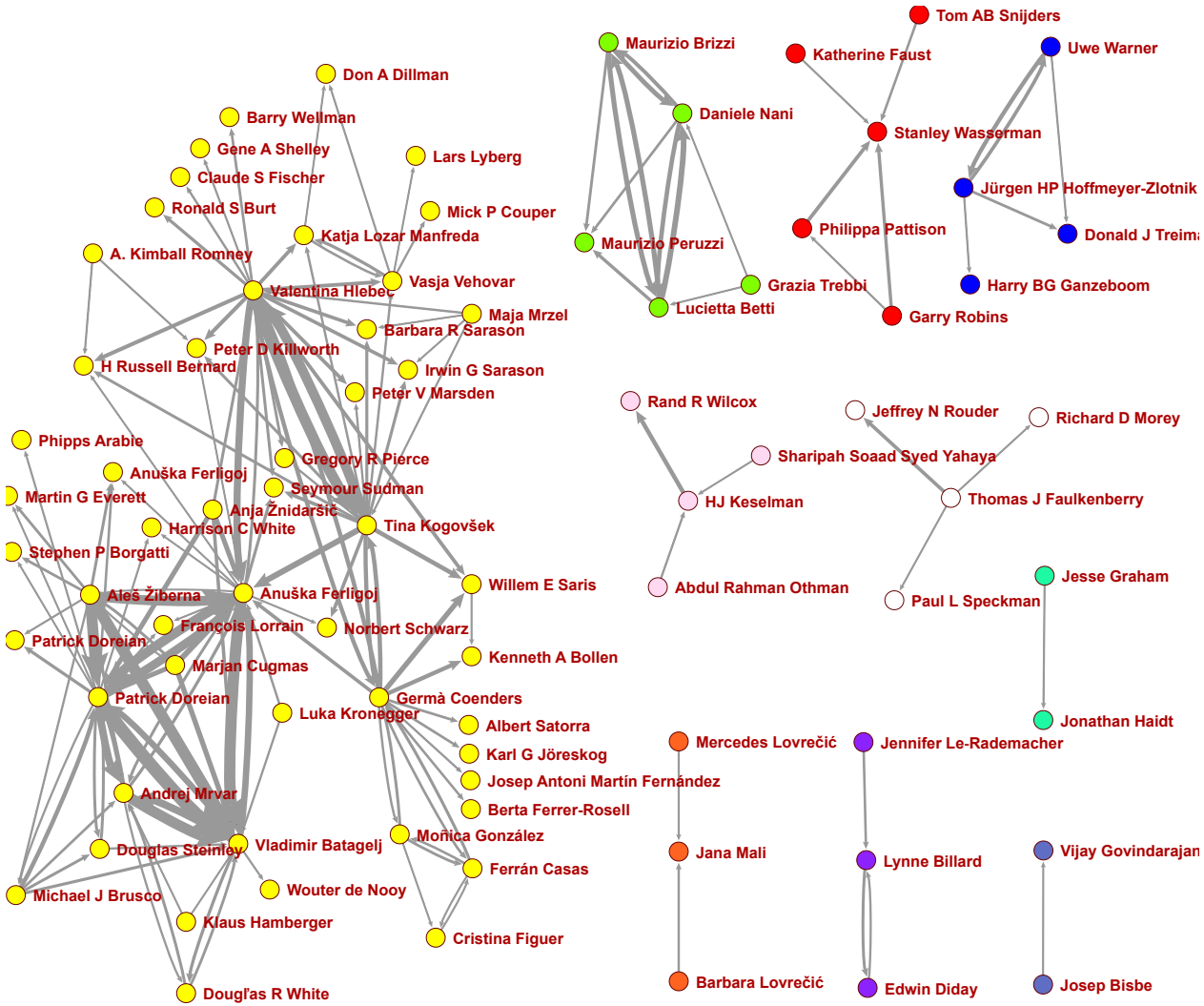
$n_{ACiA} = 10268$, $m^A_{ACiA} = 119301$, and 701 loops.
Using `Network/Info/Line values` we select the threshold $t = 15$. We make a link cut at level $t$.

For journals we get $n_{JJ} = 1776$, $m^A_{JJ} = 8382$, and 141 loops.
Using `Network/Info/Line values` we select the threshold $t = 30$. We make a link cut at level $t$.

# MZ citations between authors

## link cut at level 15, loops removed

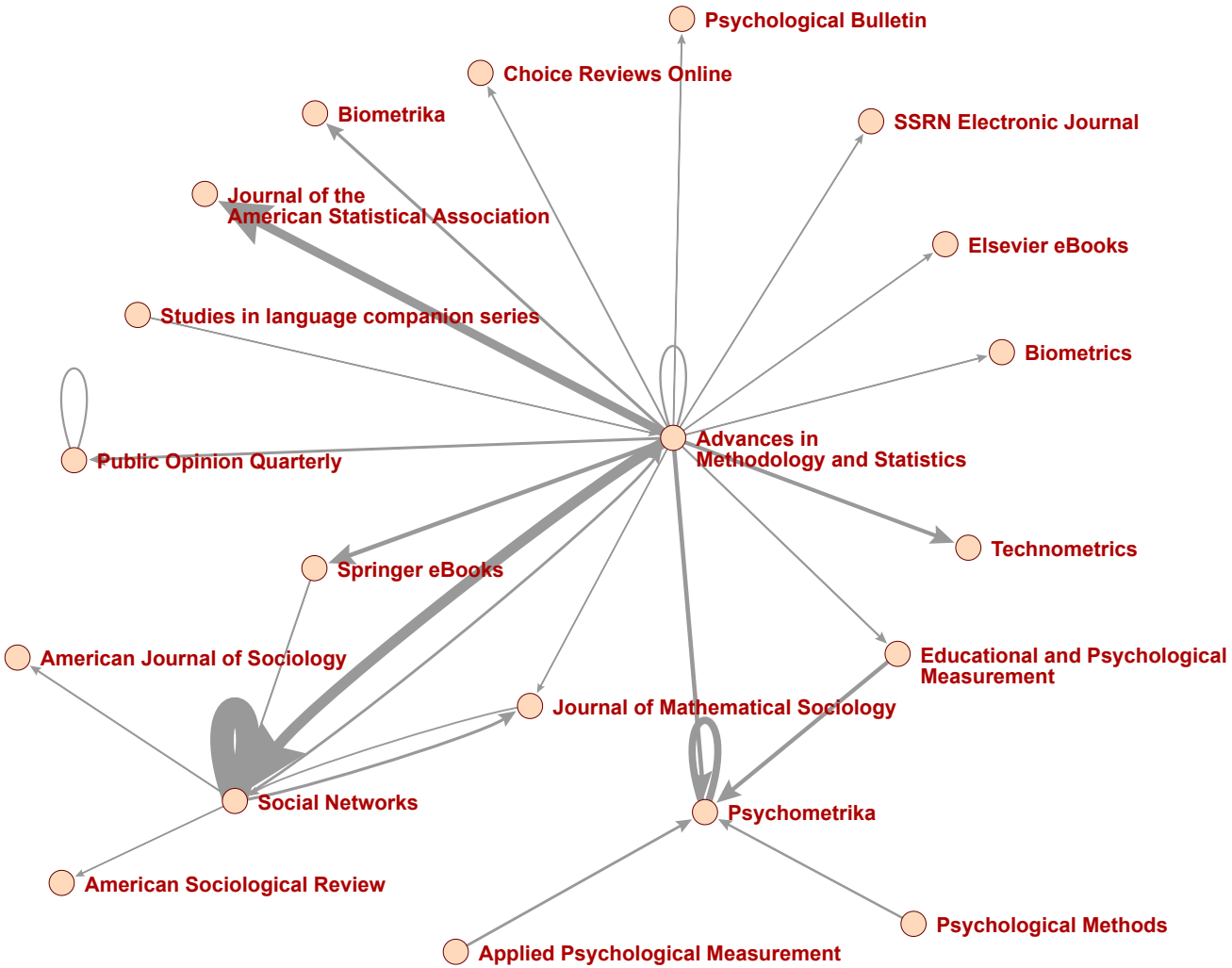




# MZ citations between journals

## link cut at level 30, `Sunknown` removed





# Creating networks

## AMC – Ars Mathematica Contemporanea

```
> OpenAlex2PajekAll(NULL,name="AMC",listF="worksAMC.csv")
OpenAlex2Pajek / All - Start Mon May 26 06:10:05 2025
*** OpenAlex2Pajek / All - Process Mon May 26 06:10:05 2025
...
*** OpenAlex2Pajek / All - Data Collected Mon May 26 07:20:44 202
hits: 12758 works: 137751 authors: 10849 anon: 185 sources: 1192
*** OpenAlex2Pajek / All - Stop Mon May 26 07:21:54 2025
```

# The top works citing the journal AMC

| | id | cdc | cby | dI | year | authors | title |
|---|---|---|---|---|---|---|---|
| 1 | W4395099390 | 2 | 3 | 7 | 2023 | A Pasotti, JH Dinitz | A Survey of Heffter Arrays |
| 2 | W1082115497 | 2 | 16 | 6 | 2015 | D Cvetković, P Rowlinson + | Graphs with least eigenvalue -2: Ten y |
| 3 | W1846554597 | 1 | 7 | 6 | 2015 | A Malnič, R Požar | On the split structure of lifted group |
| 4 | W2894294252 | 2 | 40 | 6 | 2022 | X Liu, S Zhou | Eigenvalues of Cayley Graphs |
| 5 | W2885665086 | 3 | 13 | 5 | 2018 | JL Gross, IF Khan + | Calculating genus polynomials via stri |
| 6 | W3187989029 | 0 | 0 | 5 | 2021 | FG Yin, YQ Feng + | Prime-valent Symmetric graphs with a q |
| 7 | W4311499018 | 1 | 0 | 5 | 2022 | FG Yin, YQ Feng + | Prime-valent symmetric graphs with a q |
| 8 | W3158572985 | 1 | 3 | 5 | 2021 | JH Koolen, MY Cao + | Recent Progress on Graphs with Fixed S |
| 9 | W3095252518 | 1 | 0 | 5 | 2020 | J Bokowski | Punkt-Geraden-Konfigurationen |
| 10 | W3022932399 | 1 | 3 | 5 | 2020 | DW Morris, K Wilk | Cayley graphs of order kp are hamilton |
| 11 | W2884984121 | 0 | 2 | 5 | 2018 | T Short | The saturation number of carbon nanoco |
| 12 | W4284887997 | 1 | 0 | 5 | 2022 | M Grech, A Kisielewicz | Asymmetric edge-colorings of graphs wi |
| 13 | W4404725349 | 1 | 0 | 5 | 2024 | M Bachratý, M Hagara | Observations about skew morphisms of c |
| 14 | W2501786760 | 0 | 2 | 5 | 2013 | R Nedela, M Škoviera | Maps |
| 15 | W2981706891 | 1 | 6 | 5 | 2019 | T Dobson, Š Miklavič + | On Automorphism Groups of Deleted Wrea |
| 16 | W2969633698 | 2 | 12 | 5 | 2019 | GHJ Lanel, HK Pallage + | A survey on Hamiltonicity in Cayley gr |
| 17 | W3031054242 | 3 | 7 | 4 | 2020 | Y Chen, JL Gross + | Recurrences for the genus polynomials |
| 18 | W2964137422 | 2 | 4 | 4 | 2014 | JL Gross, M Kotrbčik | Genus distributions of cubic series-pa |
| 19 | W4225266799 | 1 | 2 | 4 | 2022 | L Lu, Z Lou, Q Hua | Mixed graphs with smallest eigenvalue |
| 20 | W4280594573 | 1 | 0 | 4 | 2022 | K Kutnar, D Marušič | Symmetries in graphs via simplicial au |

# The top works cited by the journal AMC

| | id | cdc | cby | dO | year | authors | title |
|---|---|---|---|---|---|---|---|
| 1 | W1558273801 | 1 | 1537 | 45 | 2003 | T Pisanski, P Potočnik + | Topological Graph Theory |
| 2 | W1976677460 | 1 | 7335 | 45 | 1997 | W Bosma, J Cannon + | The Magma Algebra System I: The User La |
| 3 | W4285719527 | 0 | 0 | 33 | 1955 | | Deleted Work |
| 4 | W4298236575 | 2 | 2074 | 29 | 1989 | AE Brouwer, AM Cohen + | Distance-Regular Graphs |
| 5 | W4238591275 | 2 | 6085 | 27 | 2001 | C Godsil, G Royle | Algebraic Graph Theory |
| 6 | W2490805901 | 0 | 1023 | 25 | 2011 | RH Hammack, W Imrich + | Handbook of Product Graphs |
| 7 | W51037165 | 2 | 445 | 22 | 2002 | P McMullen, E Schulte | Abstract Regular Polytopes |
| 8 | W1480793893 | 1 | 165 | 22 | 2009 | B Grünbaum | Configurations of Points and Lines |
| 9 | W3138922280 | 0 | 252 | 21 | 2013 | | Topological Graph Theory |
| 10 | W2917893419 | 0 | 1321 | 21 | 1964 | | Finite Permutation Groups |
| 11 | W2051170661 | 0 | 1736 | 21 | 1966 | F Haimo, H Wielandt | Finite Permutation Groups. |
| 12 | W2060397425 | 0 | 335 | 20 | 1978 | GA Jones, D Singerman | Theory of Maps on Orientable Surfaces |
| 13 | W2798943694 | 0 | 1534 | 20 | 1996 | B Mortimer, JD Dixon | Permutation Groups |
| 14 | W247697463 | 0 | 5066 | 17 | 2009 | C Godsil, G Royle | Algebraic Graph Theory |
| 15 | W1855706715 | 1 | 297 | 16 | 1996 | MO Albertson, KL Coll | Symmetry Breaking in Graphs |
| 16 | W2117862969 | 1 | 280 | 16 | 1998 | M Xu | Automorphism groups and isomorphisms of |
| 17 | W2490728539 | 1 | 5093 | 16 | 1967 | B Huppert | Endliche Gruppen I |
| 18 | W4234295943 | 1 | 811 | 16 | 1996 | JD Dixon, B Mortimer | Permutation Groups |
| 19 | W1515707356 | 1 | 3110 | 16 | 1974 | N Biggs | Algebraic Graph Theory |
| 20 | W4300303686 | 0 | 3115 | 15 | 1977 | E Kay, JA Bondy + | Graph Theory with Applications |

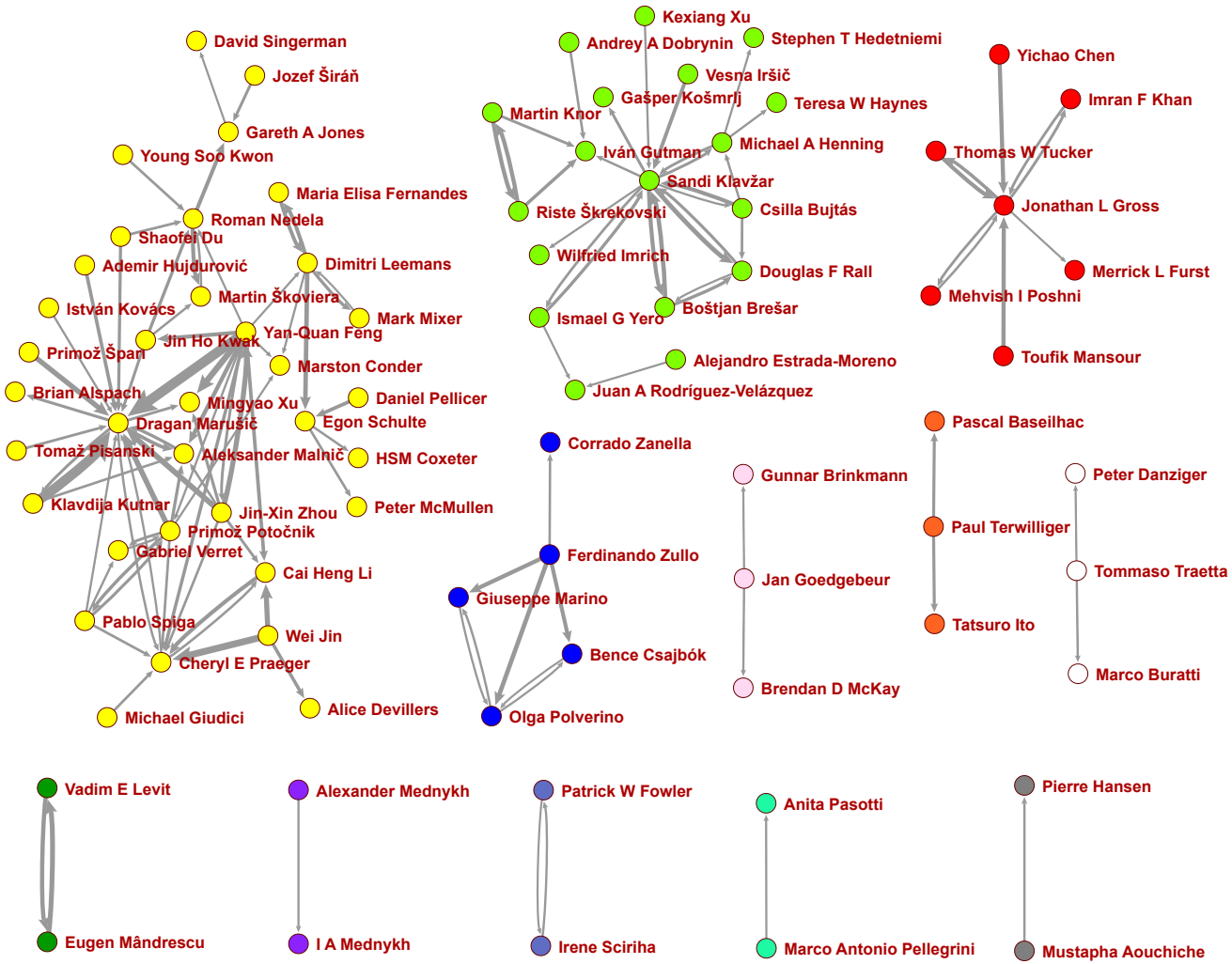
# The top authors by the number of works in the journal AMC

| | id | orcid | wc | cbc | papers | name |
|---|---|---|---|---|---|---|
| 1 | A5009207700 | 0000-0002-1257-5376 | 310 | 4639 | 38 | Tomaž Pisanski |
| 2 | A5028549991 | 0000-0002-8452-3057 | 235 | 5157 | 27 | Dragan Marušič |
| 3 | A5083674096 | 0000-0001-6851-3214 | 324 | 3542 | 17 | Riste Škrekovski |
| 4 | A5036402562 | 0000-0002-9836-6398 | 124 | 877 | 12 | Klavdija Kutnar |
| 5 | A5061875089 | 0000-0002-0256-6978 | 241 | 3016 | 12 | Marston Conder |
| 6 | A5011391021 | 0000-0002-8353-896X | 176 | 1412 | 11 | Jin-Xin Zhou |
| 7 | A5103154620 | 0000-0001-5028-3545 | 159 | 2987 | 10 | Primož Potočnik |
| 8 | A5081540165 | 0000-0002-3734-7230 | 74 | 427 | 9 | Monika Pilśniak |
| 9 | A5000602953 | 0000-0002-5901-7646 | 220 | 2813 | 9 | Jozef Širáň |
| 10 | A5078908489 | 0000-0002-4439-502X | 184 | 1130 | 9 | Dimitri Leemans |
| 11 | A5004612329 | 0000-0002-4949-4203 | 198 | 1037 | 9 | Zoran Stanić |
| 12 | A5025392880 | 0000-0002-0157-7405 | 343 | 1703 | 9 | Pablo Spiga |
| 13 | A5013070295 | 0000-0003-4245-9226 | 78 | 1472 | 9 | Aleksander Malnič |
| 14 | A5085576668 | 0000-0003-3214-0609 | 224 | 2309 | 8 | Yan-Quan Feng |
| 15 | A5101478915 | 0000-0002-2564-9530 | 162 | 5093 | 8 | Jonathan L. Gross |
| 16 | A5036070439 | 0000-0003-1766-4834 | 130 | 690 | 8 | Gabriel Verret |
| 17 | A5043525348 | 0000-0002-0881-7336 | 770 | 8480 | 8 | Cheryl E. Praeger |
| 18 | A5009892196 | 0000-0003-2416-669X | 147 | 818 | 7 | Joy Morris |
| 19 | A5089379545 | 0000-0002-6041-1106 | 239 | 1581 | 7 | Janez Žerovnik |
| 20 | A5039728806 | 0000-0002-5477-6803 | 92 | 927 | 7 | Irene Sciriha |
| 21 | A5101686757 | 0000-0002-7082-7025 | 226 | 3198 | 7 | Gareth A. Jones |
| 22 | A5083411825 | 0000-0002-8326-513X | 136 | 2103 | 7 | Tomislav Došlić |
| 23 | A5037896222 | 0000-0002-0475-9335 | 170 | 4124 | 7 | Wilfried Imrich |
| 24 | A5049148828 | 0000-0002-1556-4744 | 647 | 11134 | 7 | Sandi Klavžar |
| 25 | A5066245855 | 0000-0002-9826-704X | 151 | 1949 | 7 | Roman Nedela |
| 26 | A5059638892 | 0000-0002-2878-0745 | 98 | 556 | 7 | Štefko Miklavič |
| 27 | A5071017819 | 0000-0001-5412-4656 | 386 | 4954 | 6 | Michael Giudici |
| 28 | A5066088171 | 0000-0002-9082-7647 | 183 | 3747 | 6 | Juan A. Rodríguez-Velázquez |
| 29 | A5113798192 | <NA> | 134 | 759 | 6 | Dave Witte Morris |
| 30 | A5047753351 | 0000-0003-2106-1104 | 447 | 14010 | 6 | Patrick W. Fowler |
| 31 | A5022919726 | 0000-0002-7868-6925 | 104 | 2911 | 6 | Thomas W. Tucker |
| 32 | A5089473322 | 0000-0001-8185-067X | 735 | 10163 | 6 | Michael A. Henning |
| 33 | A5054789661 | 0000-0003-0935-5724 | 64 | 175 | 6 | Leah Wrenn Berman |
| 34 | A5075057406 | 0000-0002-1034-3993 | 129 | 2899 | 6 | Brian Alspach |
| 35 | A5033117907 | 0000-0002-2108-7518 | 139 | 1830 | 6 | Martin Škoviera |

# AMC citations between authors

## link cut at level 75, loops removed

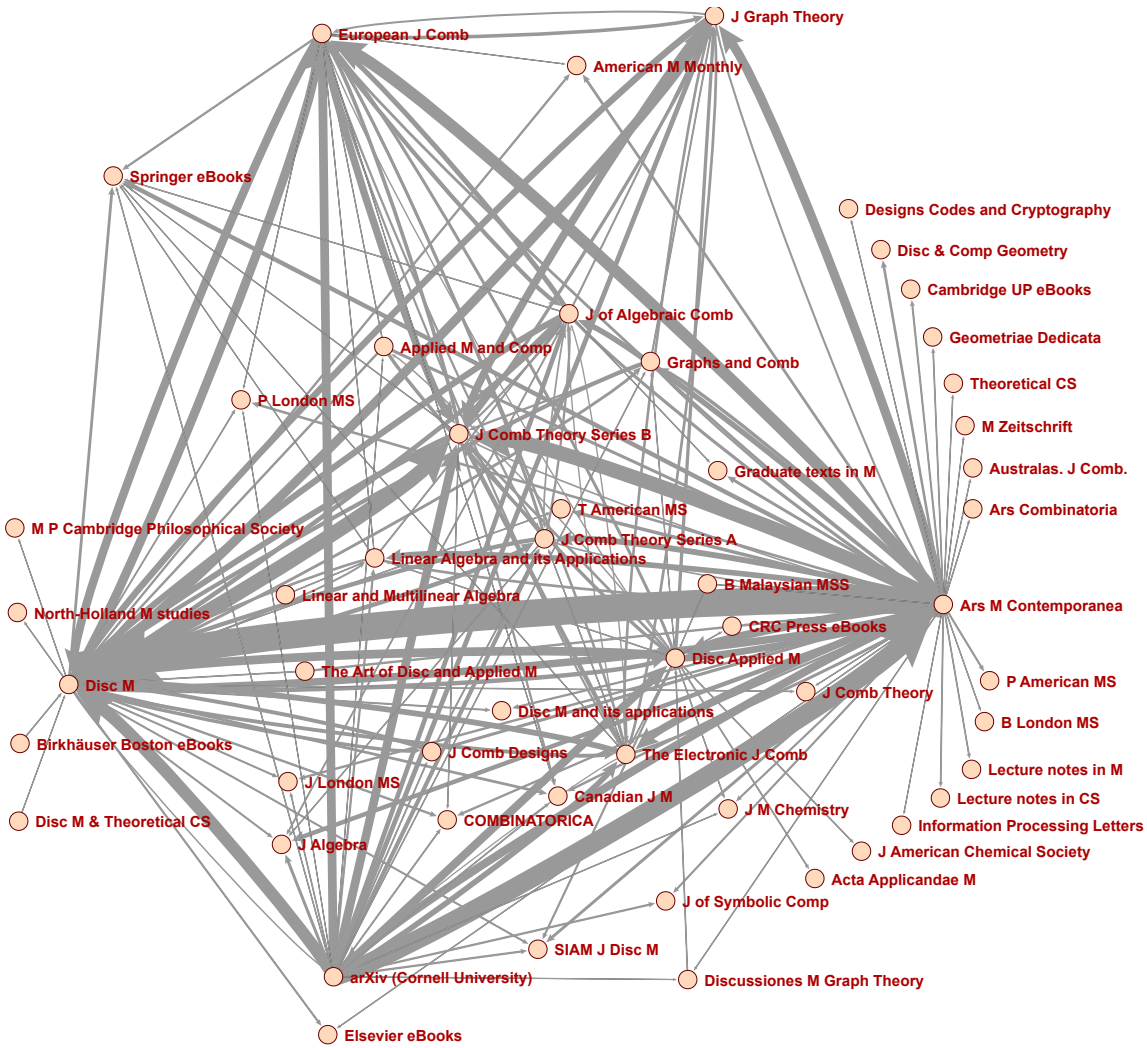




# AMC citations between journals

## link cut at level 50, loops and Sunknown removed





# Conclusions

1. The data in the OpenAlex database is not completely error-free. Most errors can be considered as noise – important units will float to the surface.
2. If the error is serious and is reflected in the final result, we correct it accordingly and repeat the analysis.
3. We can also contribute to the quality of the data in the OpenAlex database by informing the database maintainers about errors.

# Acknowledgments

The computational work reported in this paper was performed using the R [5] package `OpenAlex2Pajek` and the program Pajek for analysis of large networks [6]. The code and data are available at Github/Bavla/ OpenAlex.

This work is supported in part by the Slovenian Research Agency (research program P1-0294 and research project J5-4596), and prepared within the framework of the COST action CA21163 (HiTEc).

# References I

1. Priem, J., Piwowar, H. & Orr, R. OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts. *arXiv preprint arXiv:2205.01833* (2022).
2. Batagelj, V. & Cerinšek, M. On bibliographic networks. *Scientometrics* **96,** 845–864 (2013).
3. Batagelj, V. On fractional approach to analysis of linked networks. *Scientometrics* **123,** 621–633 (2020).
4. Batagelj, V. & Maltseva, D. Temporal bibliographic networks. *Journal of Informetrics* **14,** 101006 (2020).
5. Team, R. C. R: A language and environment for statistical computing. 2020. *R Foundation for Statistical Computing: Vienna, Austria* (2020).
6. De Nooy, W., Mrvar, A. & Batagelj, V. *Exploratory social network analysis with Pajek: Revised and expanded edition for updated software.* (Cambridge university press, 2018).