INTRODUCTION TO BIBLIOMETRICS

Stéphane Goria – Université de Lorraine

Summer school in digital humanities, Metz, 2023

Purposes

- *Having a basic vocabulary about bibliometrics*
- *Know the main blibliometric variables and indicators*
- *Have an idea of its usefulness and biases*
- *Know a few sources for bibliometric research*
- *Being able to prepare research aimed at building reference corpora*
- *Knowing some tools and techniques that can be implemented*
- *Being able to apply these techniques to one's own research topic*

For a PhD student

- ***Bibliometric*** **methods can be highly valuable for PhD students in several ways:**

- Literature Review: It helps in conducting a comprehensive literature review by identifying key authors, journals, and seminal works in the research area. This ensures that you are building on existing knowledge effectively.
- Research Gap Identification: By analyzing the literature, you can identify gaps in existing research, which can guide the direction of your PhD work. You can identify areas that need further exploration or where there is limited research.
- Publication Strategy: It can inform your publication strategy by helping you target journals or conferences that are most relevant and influential in your field.
- Career Development: Understanding bibliometric concepts (indicators) can be useful for your academic career, as it's a skill that is increasingly valued in academia.

Notions and Definitions

- ***Scientometrics***

The quantitative methods of the research on the development of science as an informational process which proposed by Vasiliy Nalimov and Zinaida Mul'chenko (1969). These methods focus on the analysis of scientific fields or domains in relation to each other.

- ***Bibliometrics***

The application of mathematics and statistical methods to books and other media of communication which proposed by Alan Pritchard (1969).

Bibliometrics is a subset of scientometrics, specifically focusing on the bibliographic aspects of scholarly literature, such as citations and publication patterns, while scientometrics takes a more comprehensive view of the scientific research process, including the social and organizational aspects of science. Both fields use quantitative methods to gain insights into the world of academic research and scholarship.

Sources:

Mingers, J., & Leydesdorff, L. (2015). A review of theory and practice in scientometrics. *European journal of operational research*, *246*(1), 1-19.

Nalimov, V. V. & Mul'chenko, Z. M. Наукометрия, Изучение развития Науки как информационного процесса [Naukometriya, the Study of the Development of Science as an Information Process] (in Russian). Moscow: Nauka, 1969.

Nalimov, V., & Mulcjenko, B. (1971). *Measurement of Science: Study of the Development of Science as an Information Process*. Washington DC: Foreign Technology Division.

Pritchard, A. (1969). Statistical bibliography or bibliometrics? *Journal of Documentation*, 25, 348-349.

Rousseau, R. (2021). Naukometriya, Nalimov and Mul'chenko. *Collnet Journal of Scientometrics and Information Management*, *15*(1), 213-224.

Notions and Definitions

- ***Bibliometrics***

Bibliometrics is a set of quantitative methods used, notably in library and information science, as well as in the field of research evaluation. It involves the statistical analysis of various aspects of written publications, such as books, articles, and other forms of academic or scientific literature.

One goal of bibliometrics is to assess and measure the impact, influence, and productivity of authors, journals, institutions, or research fields based on the analysis of citation patterns, publication counts, and other bibliographic data. This field of study is widely used in academia and research to evaluate the significance and reach of scholarly work and to inform decision-making in academic publishing, funding, and research assessment. Among these methods there are citation analysis, content analysis, multidimensional analysis and network analysis.

Bibliometrics can be applied in other professional domains. For example, patent bibliometrics is similar with scholar bibliometrics. It is a valuable tool for analyzing the patent landscape, understanding technological innovation, and making informed decisions related to intellectual property and innovation strategies. It allows for data-driven insights into the world of patents and their role in advancing technology and innovation.

There are exist too: webometrics / cybermetrics, infometrics, econometrics, ...

Some bibliometrics applications for an university

Used in conjunction with other qualitative measures, bibliometrics can provide a comprehensive assessment of research and academic activities as:

- ***Research evaluation:***

- Faculty Assessment: Universities can use bibliometrics to evaluate the research productivity and impact of faculty members. This can inform promotion and tenure decisions.
- Departmental and Institutional Rankings: Bibliometric data can help universities assess the research performance of departments and the institution as a whole. This information can be used for strategic planning and benchmarking.

- ***Resource allocation:***

- Library Collection Management: Bibliometrics can guide libraries in making informed decisions about which journals, books, or databases to subscribe to based on their relevance and impact.
- *Research Funding Allocation*: Universities can use bibliometric data to allocate research funding to departments or projects based on their research track records.

Some bibliometrics applications for an university

- ***Research Collaboration and Networking:***

- Identifying Collaboration Opportunities: Bibliometrics can help researchers identify potential collaborators by analyzing co-authorship networks and research interests.
- Mapping Research Landscape: It can be used to map the research landscape in various disciplines, helping universities identify emerging areas and interdisciplinary research opportunities.

- ***Institutional Repositories:***

- Managing Institutional Repositories: Universities can use bibliometrics to track the usage and impact of materials in their institutional repositories, such as theses, dissertations, and open-access publications.
- Promoting open Access Initiatives: Bibliometrics can support universities in their efforts to promote open access publishing by tracking the impact of open access research.

Some bibliometrics applications for an university

- ***Monitoring and Reporting:***

- Research Impact Reporting: Universities can generate reports on the impact of their research output for various stakeholders, including funding agencies, alumni, and the public.
- Benchmarking: Bibliometrics can be used to benchmark a university's research performance against peer institutions regionally or globally.
- Publication Trends: Researchers can analyze publication trends in academic journals or conference proceedings. An abrupt increase in the number of publications in a particular area, especially if they are from a diverse set of authors and institutions, can be a sign of a new scientific domain gaining prominence.
- Citation Patterns: Citation analysis is a powerful tool in bibliometrics. When publications in a specific field start to receive more citations, it can indicate growing interest and recognition in that field (see the main bibliometrics laws).
- Keyword Analysis: Bibliometric analysis often involves examining the keywords used in academic publications. Researchers can track the appearance and frequency of new keywords or terms related to a specific scientific domain. An increase in the use of these keywords over time may indicate the emergence of a new field.

Some controverses and critics about bibliometrics

- ***Overreliance on Quantitative Metrics:***

➢ The citation counts and the use of journal impact factors to assess research quality and productivity are controversed. Critics argue that these metrics do not capture the full spectrum of research impact and can lead to skewed incentives and biases in academia.

- ***Inequality and Bias:***

➢ Bibliometrics can exacerbate existing inequalities in academia. Researchers from underrepresented groups or in less-resourced institutions may be at a disadvantage in terms of citation counts and access to high-impact journals, leading to systemic bias.

- ***Quality vs. Quantity:***

➢ Some argue that bibliometrics place too much emphasis on quantity (e.g., the number of publications) over quality (e.g., the significance and impact of research findings). This can discourage deep and innovative research in favor of producing more publications.

- ***Negative Impact on Scientific Culture:***

➢ Critics contend that the emphasis on bibliometrics can erode the scientific culture by promoting a competitive "publish or perish" mentality rather than fostering collaboration, open science, and creativity.

Some controverses and critics about bibliometrics

- ***Disciplinary Differences:***

➢ Different academic disciplines have unique publication and citation patterns. Using a one-size-fits-all approach to bibliometrics can disadvantage researchers and fields with different norms and practices.

- ***Language Differences:***

➢ Publications in English are cited far more than those in other languages, with no relation to their quality.

- ***Source Bias:***

➢ Results will vary depending on the sources (journals, conferences, books…) and data used in the calculations. The platforms that offer bibliometrics indicators do so on the basis of the sources they are able to interrogate.

- ***Manipulation and influence by a gamification / pointification system:***

➢ The use of bibliometrics in performance evaluation has led to concerns about researchers and institutions are being encouraged to game the system to their own advantage. These include abuse through self-citation, excessive authorship and other practices aimed at artificially inflating citation counts and other metrics.

How to use bibliometrics with ethics

- ***Transparent Methodologies:***

➢ Clearly document the methodologies and data sources used for bibliometric analyses. Make this information publicly available so that others can assess the validity of the analysis.

- ***Responsible Metrics Framework:***

➢ Encourage the use of multiple metrics and sources to provide a more holistic view of research impact, including qualitative assessments and alternative metrics (altmetrics).

➢ Implement measures to mitigate biases in bibliometric data, such as adjusting for field-specific differences in citation practices or considering the impact of research collaborations.

- ***Research on Metrics:***

➢ Invest in research that examines the impact and ethics of bibliometric practices. Encourage studies that evaluate the effectiveness of different assessment approaches.

- ***Expert Evaluation:***

➢ Encourage experts to consider the quality, significance, and societal impact of research, in addition to bibliometric metrics.

Principal bibliometric laws

- *(Alfred) Lotka's law (1926)*

➢ Production and recognition systems in science follow a Pareto distribution, with **20% of researchers producing 80% of the publications.**

- *(Samuel) Bradford's law of scattering (1934)*

➢ The references of documents follow a Pareto distribution, with **20% of the journals publishing 80% of the most cited articles.**

- *(George Kingsley) Zipf and (Benoit) Mandelbrot's law (1935 + 1952) : Long tail*

➢ The distribution of words in a text or texts by the same author follows a unique Pareto law.

Some examples of altmetrics

- ***Article Views and Downloads:***

➢ Tracking the number of times a research article is accessed or downloaded from a publisher's website or a preprint server can provide insights into its popularity and interest.

- ***Online Reviews and Comments:***

➢ The presence of user-generated content, such as reviews and comments on platforms like ResearchGate or academic social networks, can reflect the engagement of the research community.

- ***Usage in Online Courses:***

➢ Inclusion of research papers or materials in online courses, syllabi, or educational resources can demonstrate their influence on teaching and learning.

☝ These other metrics and inicators also have their limits and shortcomings, and can be subject to manipulation.

Main bibliometric indicators

- ***Journal Impact Factor:***

➢ It assesses the influence of a journal by calculating the average number of citations received per paper published in that journal during a specific time period. It's primarily used to evaluate journals, not individual papers or authors.

- ***CiteScore (proposed by Scopus):***

➢ It measures the average number of citations received per document published in a journal over a specific time frame.

- ***Citation Count:***

➢ The number of times a publication has been cited by other scholarly works. It's one of the most common and widely used bibliometric indicators.

- ***Author Productivity:***

➢ It measures the total number of publications by an author, which can be used to gauge their research output.

- ***Co-authorship Network Analysis:***

➢ Analyzing co-authorship networks can reveal collaboration patterns and highlight influential researchers within a field.

Main bibliometric indicators

- ***H-index:***

➢ It measures both the productivity and impact of a researcher's work. An author has an h-index of h if they have published h papers that have each been cited at least h times.

- ***HA-index (proposed by Harzing's Publish or Perish tool):***

➢ This index divides the citation count of each paper by the age of that paper.

- ***I10-index (proposed by Google Scholar):***

➢ It represents, for one author, the number of papers with at least 10 citations.

- ***Egghe's G-Index:***

➢ It corresponds to the greatest number of articles for which all g articles received at least g^2 citations.

- ***Collaboration index:***

➢ It is a measure of the degree to which an author collaborates with others in their research. It is calculated by dividing the number of co-authored publications by the total number of publications authored by the individual.

Main variables used

- **Researcher Identifier**: a unique identifier for researchers, such as ORCID, to track their publications and citations accurately
- **Researcher's affiliation** and work country,
- **Publication year** and publication count,
- **Publication type** and **journal** name, **conference** name…
- Publication **language**,
- **Disciplinary field** of the journal or conference
- **Authorship position** and author's average position,
- **Co-authors** (and their affiliations),
- **Citation** count and self-citation count,
- Cited references,
- Words used in **titles** or **summaries** of researcher’s publications, **keyword** indexed…

Some scientific platforms presenting bibliometric indicators or variables directly

Google Scholar : indicators for authors consulted one by one

Google Scholar

Ajouter une photo
Complétez votre profil.
AJOUTER

Ajouter les coauteurs
Des suggestions de coauteurs sont disponibles.
AJOUTER

Stéphane Goria

SUIVRE

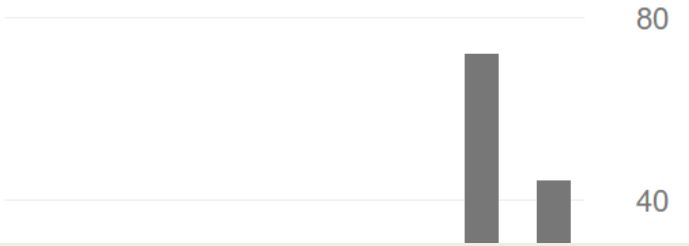
Centre de Recherche sur les Médiations (CREM) - Université de Lorraine
Adresse e-mail validée de univ-lorraine.fr

veille stratégique innovation information scientifique et t… visualisation de données gamification

TITRE	CITÉE PAR	ANNÉE
A deck of cards to help track design trends to assist the creation of new products S Goria International Journal of Technology, Innovation and Management (IJTIM) 2 (2 …	49	2022

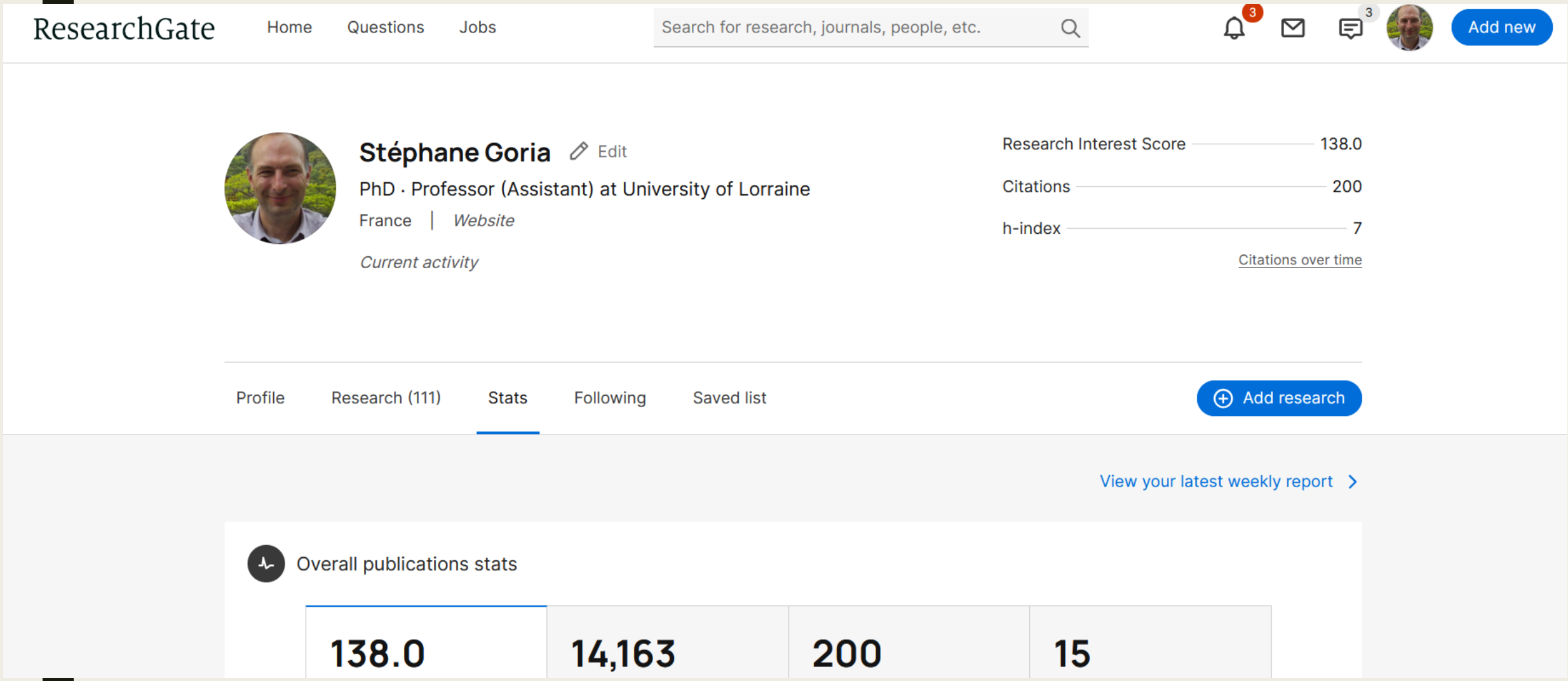
Citée par TOUT AFFICHER

	Toutes	Depuis 2018
Citations	425	196
indice h	10	7
indice i10	12	4



Some scientific platforms presenting bibliometric indicators or variables directly

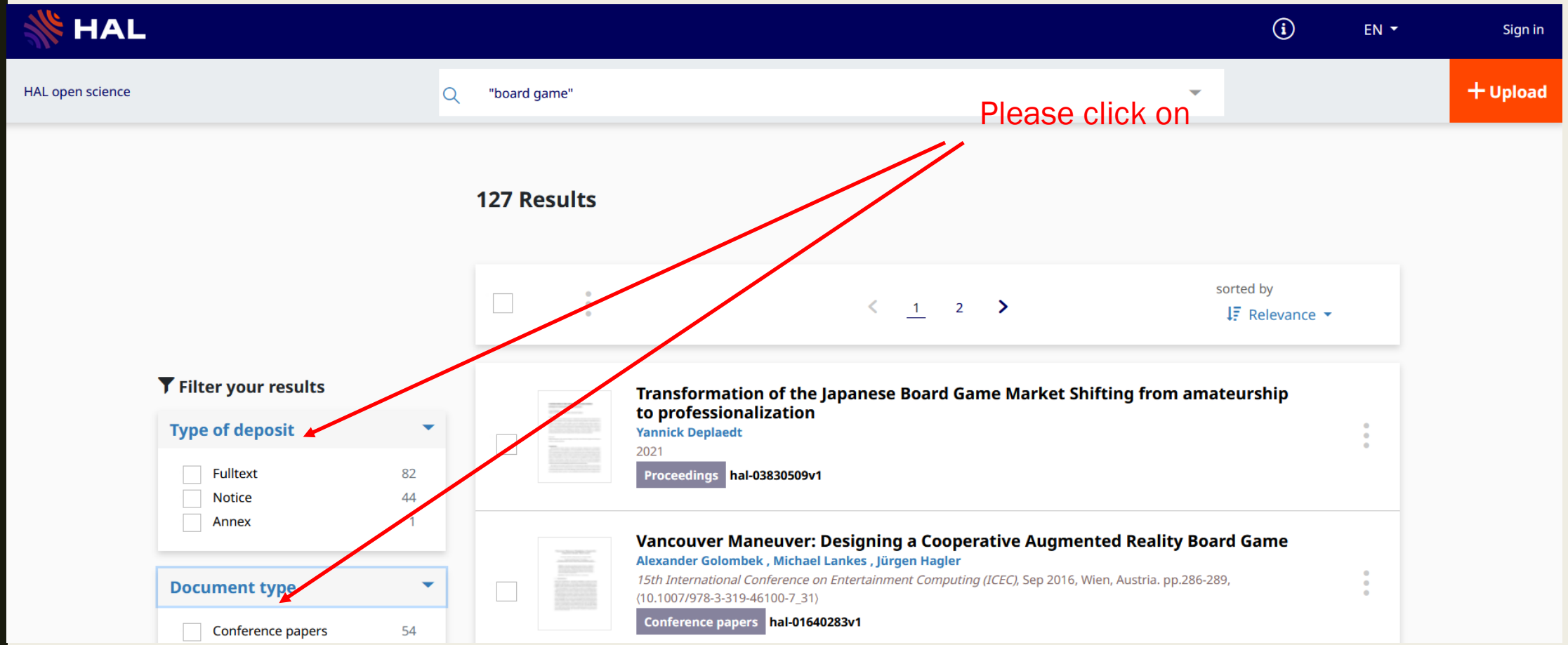
Researchgate : indicators for authors consulted one by one



Some scientific platforms presenting bibliometric indicators or variables directly

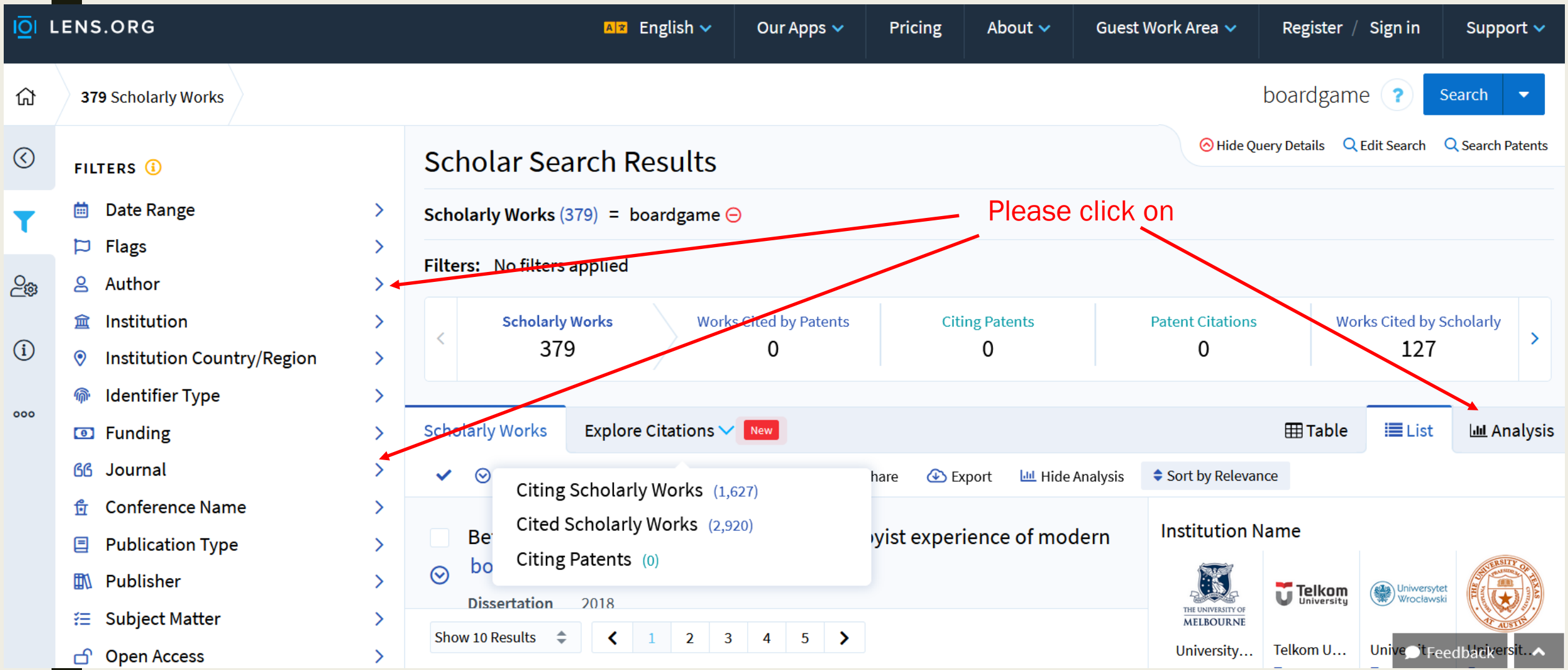
Hal.science : indicators for many variables, but not h-index.

N.B : This platform is interesting if you want to get a better idea of French academic research than many other more general platforms.



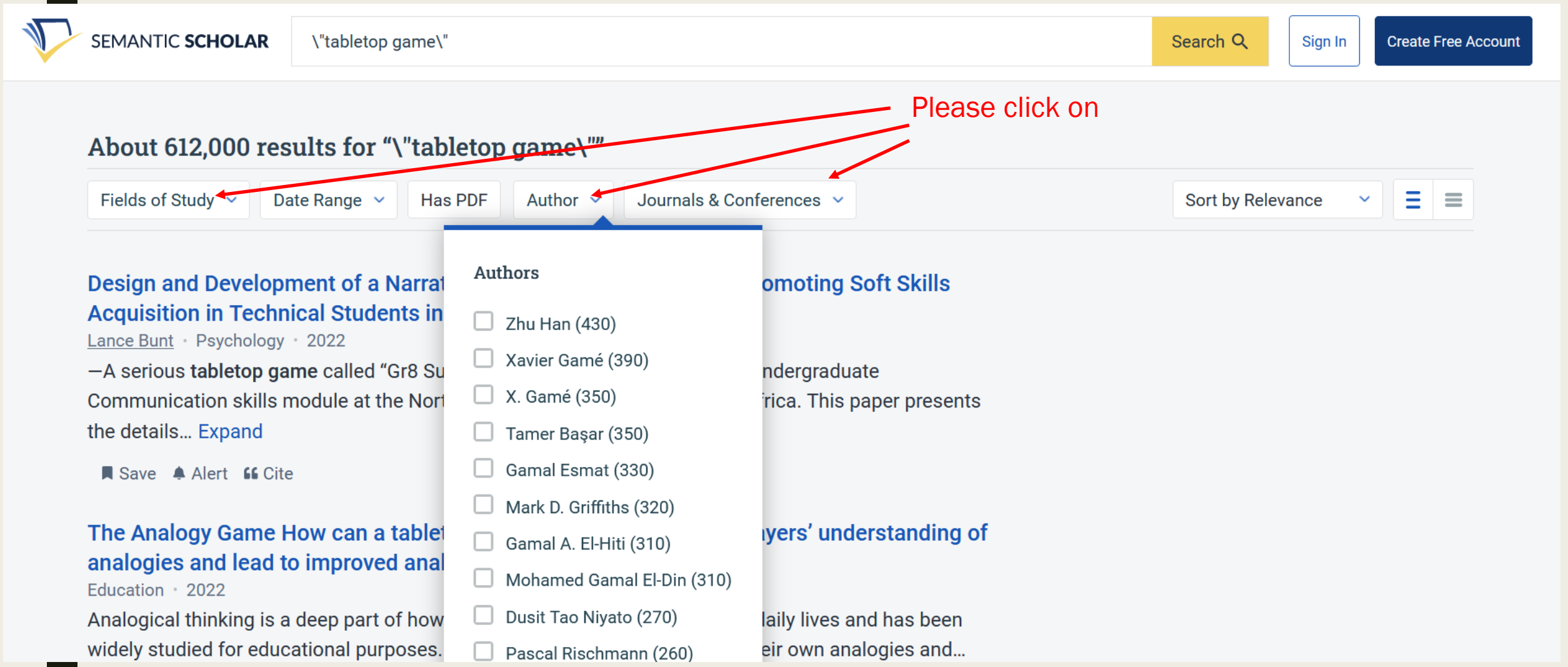
Some scientific platforms presenting bibliometric indicators or variables directly

Lens.org : indicators for many variables + h-index, but only limited to user account



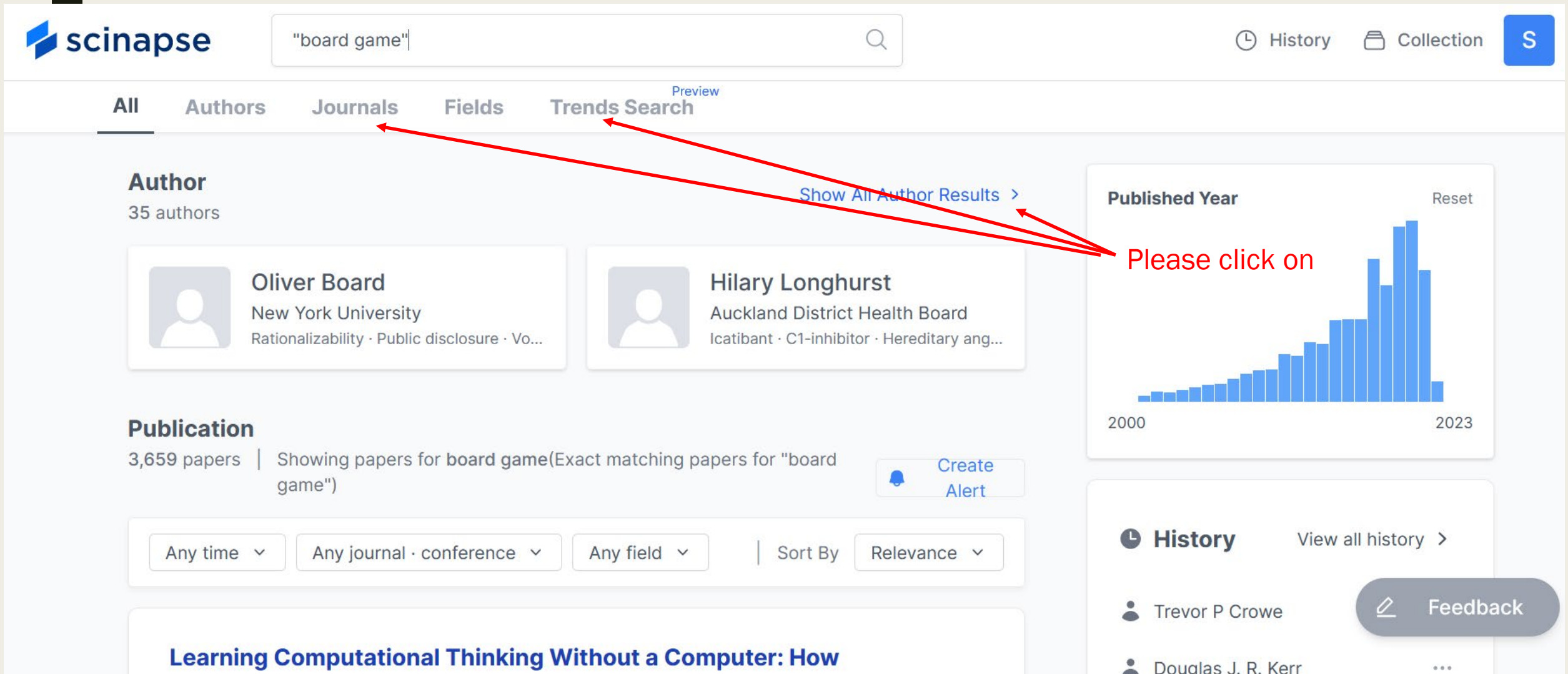
Some scientific platforms presenting bibliometric indicators or variables directly

Semanticscholar.org : indicators for many variables + h-index, but only limited to user account



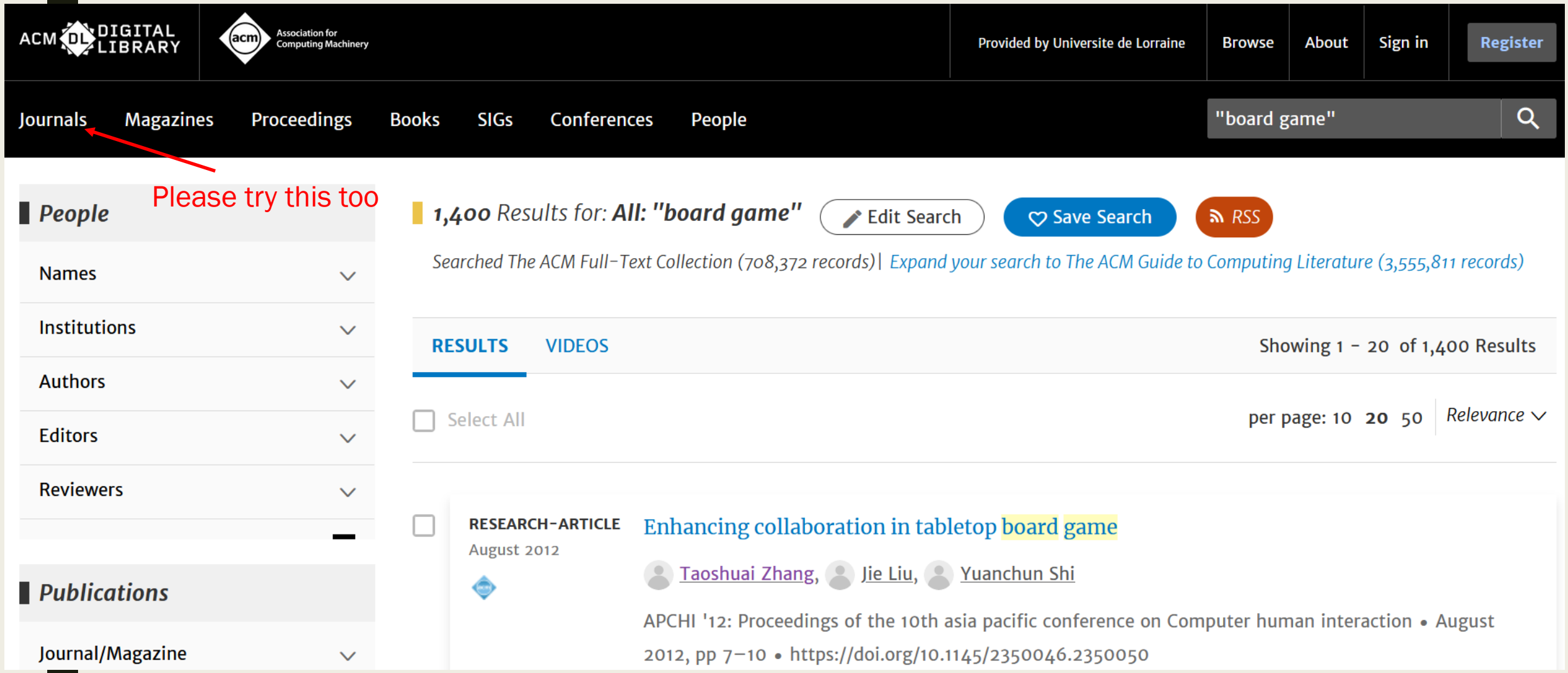
Some scientific platforms presenting bibliometric indicators or variables directly

Scinapse.io : indicators for many variables + h-index for author, impact factor for journals, etc.



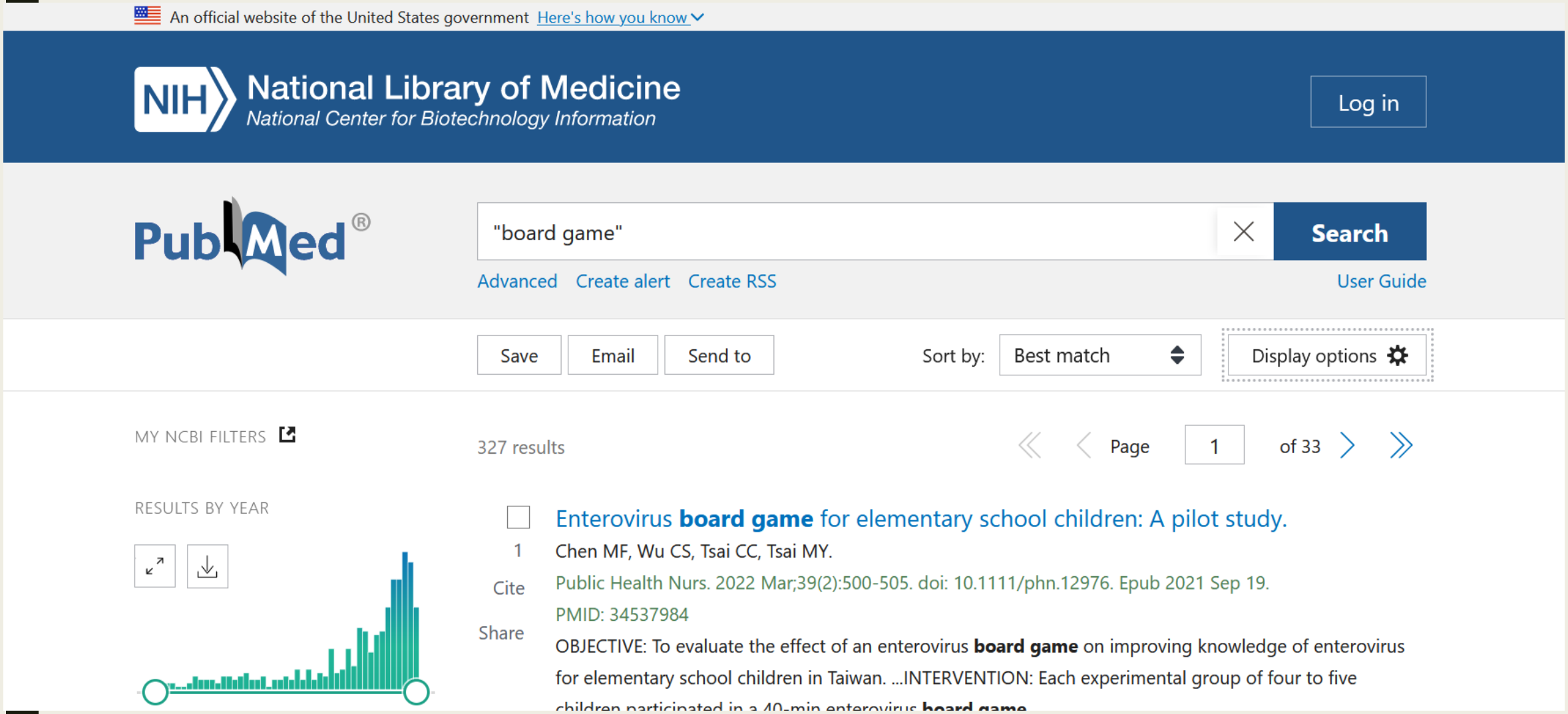
Some scientific platforms presenting bibliometric indicators or variables directly

ACM Digital Library : indicators for many variables



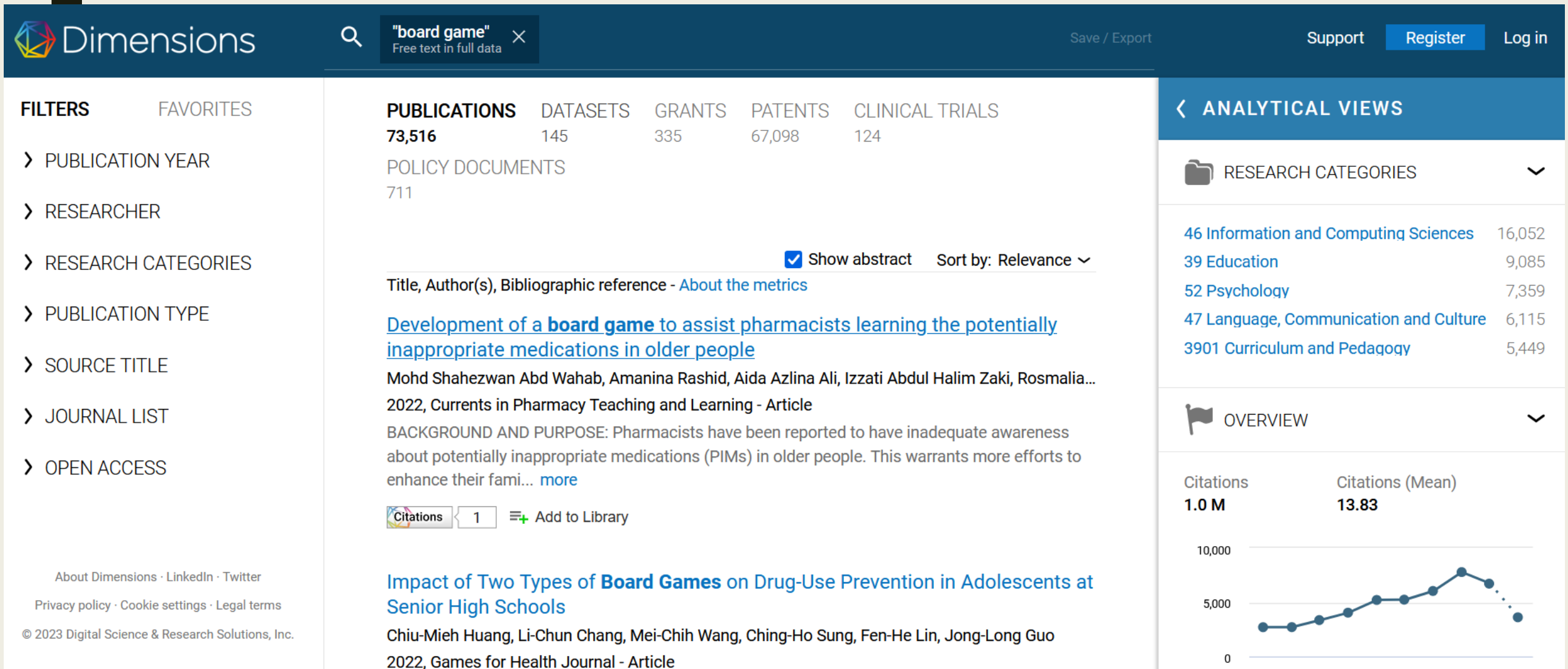
Some scientific platforms presenting bibliometric indicators or variables directly

PubMed : indicators for many variables + co-citations



Some scientific platforms presenting bibliometric indicators or variables directly

Dimensions.ai : indicators for many variables

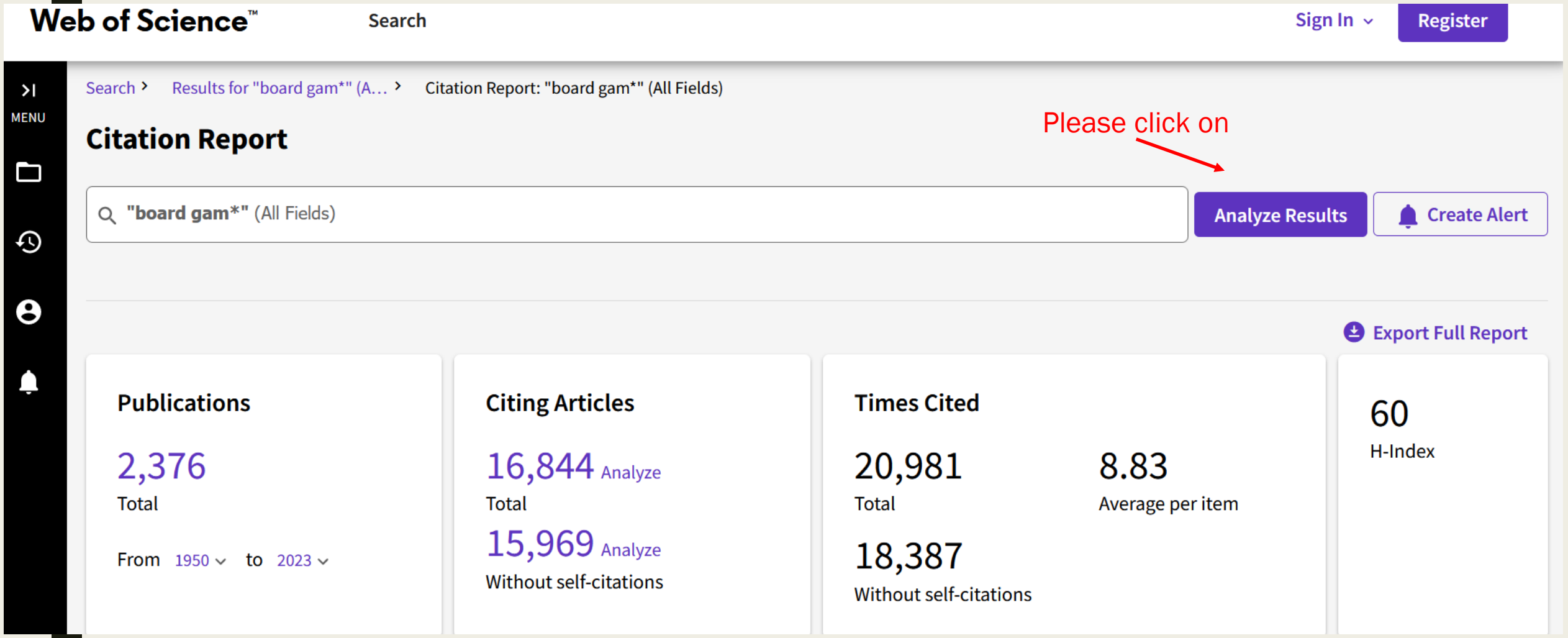


Some scientific platforms presenting bibliometric indicators or variables directly

Scopus : indicators for many variables & h-index for authors & CiteScore for journals

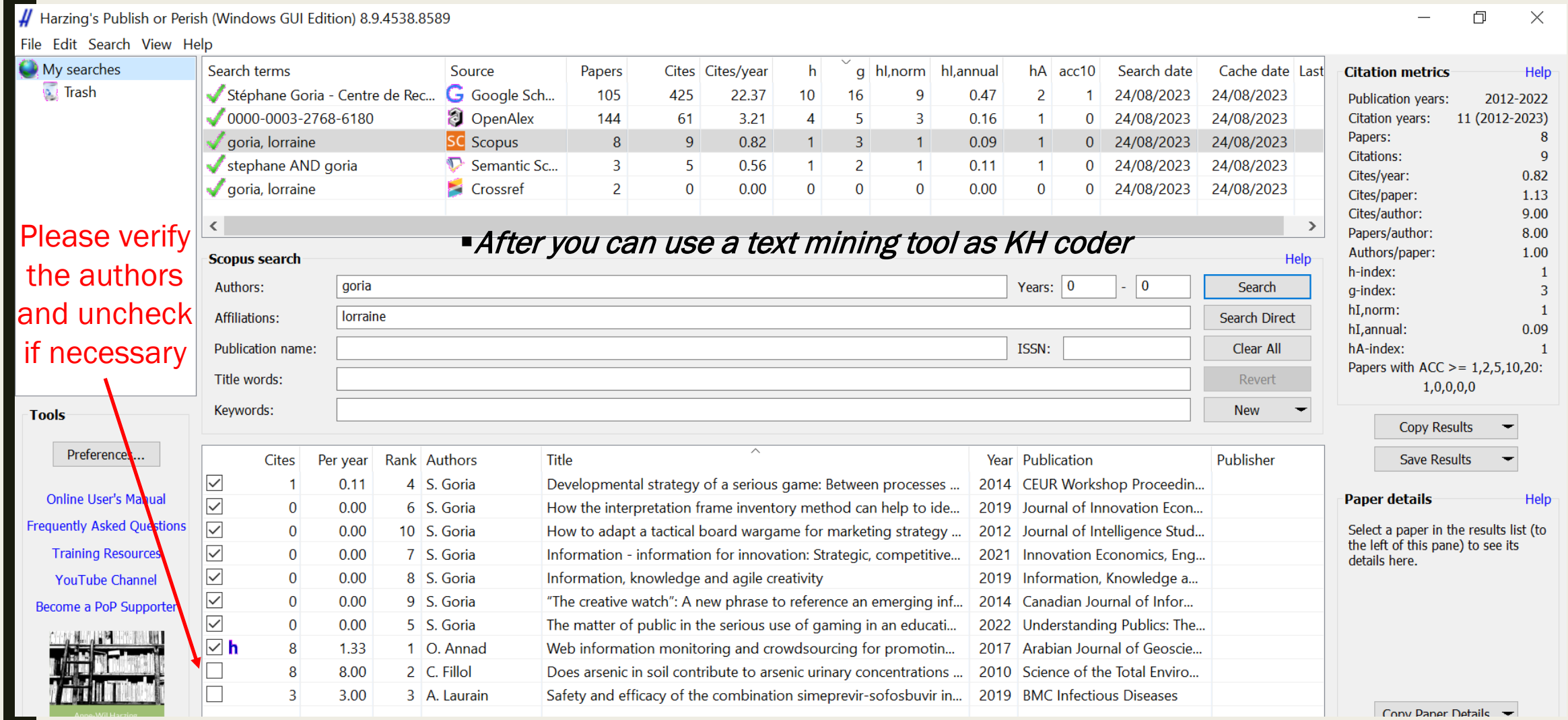
Some scientific platforms presenting bibliometric indicators or variables directly

Web of Science (WoS) : indicators for many variables & h-index for authors

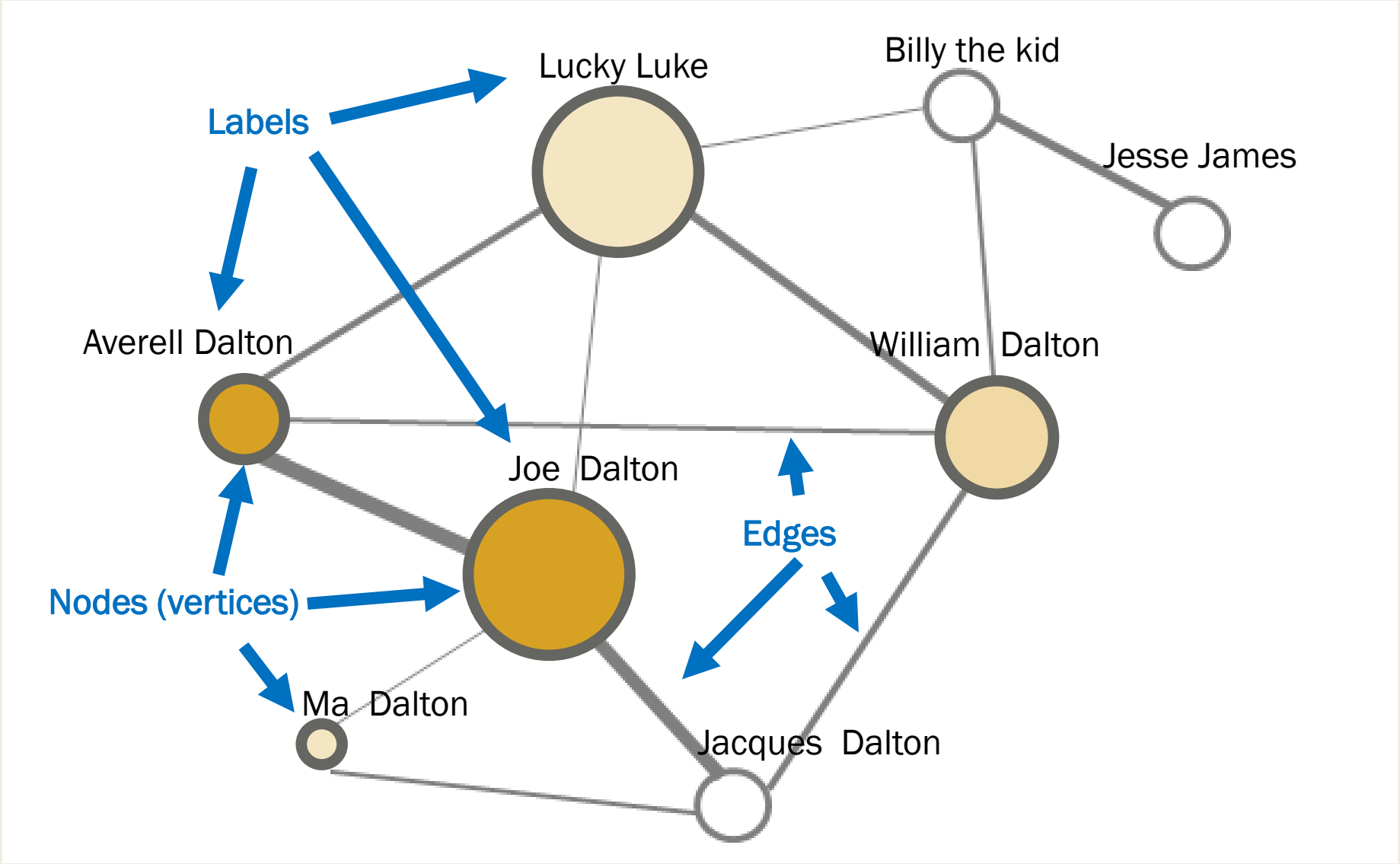


Some scientific platforms presenting bibliometric indicators or variables directly

Harzing's Publish or Perish tool (PoP): h-index and G-index for authors and others scores



Principal notions of network analysis



Principal notions of network analysis

- ***Entities and network:***

- The identification and analysis of a network of entities implies the existence of at least one **relationship between the entities** under study. In bibliometrics, they are co-authors, co-citations and word co-occurrences.
- The relationships highlighted may be **oriented**, bidirectional or **non-oriented**.
- When drawing a network, certain rules must be followed. Software for creating or highlighting networks includes, at least by default, association rules.

- ***Network density:***

- The density of a network is the ratio of the number of existing arcs (oriented links) or edges (non-oriented links) to the maximum number of possible arcs or edges.

Principal notions of network analysis

- ***Connectivity:***

➢ The connectivity of a graph designates the absence of vertices (nodes) isolated from others.

➢ A graph is said to be connected if every pair of vertices in the graph is connected.

- ***Degree:***

➢ For a non-oriented graph, its degree corresponds to the number of links attached to a node X.

➢ For a directed graph, it's the number of links pointing to a node Y or leaving a node Y..

- ***Centrality:***

➢ Centrality measures the relative position of nodes within a set.

Network construction and analysis

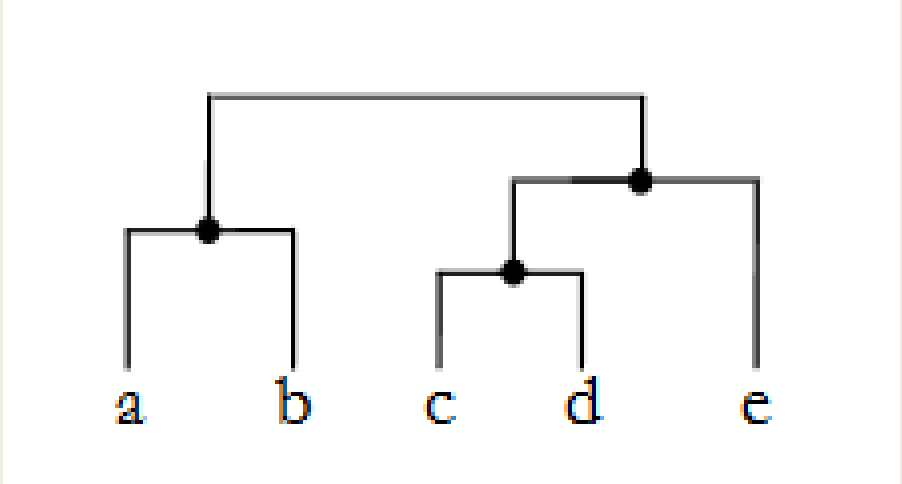
- *Construction and analysis of the drawn network should be based on a number of questions, such as:*

- Which entities are the most important?
- Which entities are most closely linked?
- Which entities are most peripheral?
- Are there any groupings of entities?
- Do certain entities play the role of unavoidable pivots?
- Which entities have the strongest links?
- Why are these entities so closely linked?
- What is the true nature of these links?

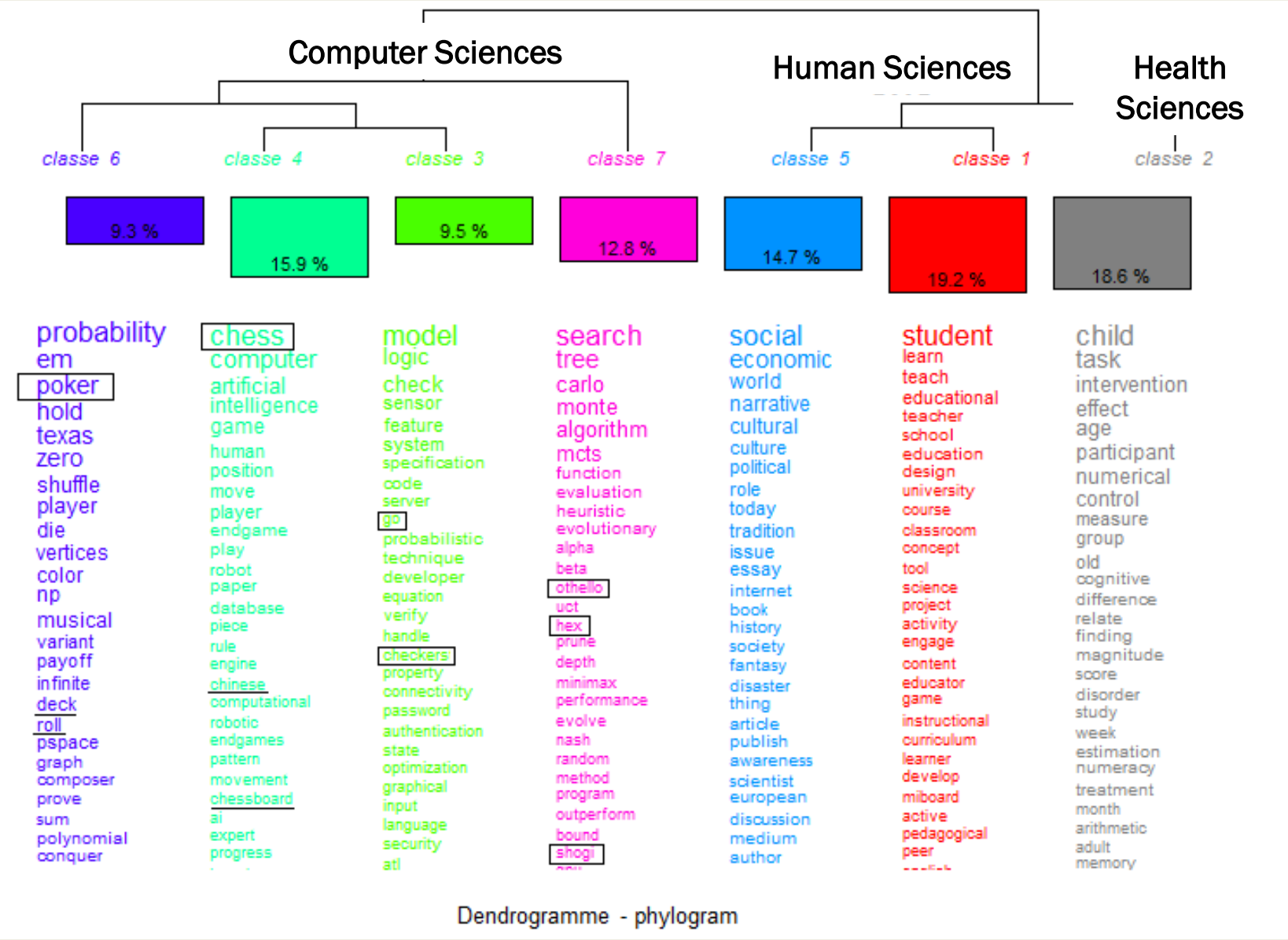
Dendogram and hierarchical clustering

- *Caracteristics:*

- Most often, the dedicated algorithm involves creating, at each step, a partition obtained by aggregating pairwise the closest elements. These are unsupervised clustering algorithms. They are sensitive to the initially defined partitions, but some propose an "average" ranking based on multiple iterations of clustering on the same dataset. They have default parameters, but they are adjustable, especially the number of main branches one wishes to obtain. Tools like Iramuteq (and R), KH coder, or Orange data mining can be used to generate them.
- As a result, the algorithm provides a hierarchy of partitions in the form of classification trees also known as dendrograms.



Dendogram and hierarchical clustering



(Iramuteq classification from the PoP boardgame reference corpora)

Self Organizing Map and classification

- ***Caracteristics:***

- Kohonen maps, also known as Self-Organizing Maps (SOM), are generated using an unsupervised classification method, but one that does not rely on points (the positions of elements in an n-dimensional space) like dendrograms do. Instead, they are based on a tiling of enclosures around these points, using others called neurons.
- The "classic" structure of the charts obtained by this type of algorithm is a tiling in hexagons. The result of this type of classification is somewhat like a tree seen from above, where only the main elements are displayed, with the others located one or more levels below in function of areas colored.
- Tools like KH coder, Orange data mining or XLSTAT

can be used to generate them.

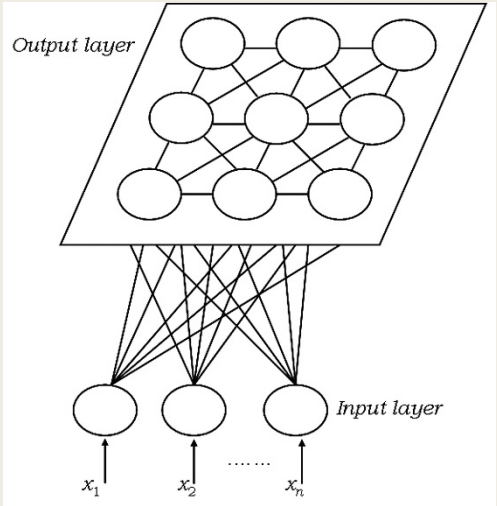


Image source:
https://cybernetist.com/2017/01/13/self-organizing-maps-in-go/

Dendogram and hierarchical clustering

(KH coder classification from the PoP boardgame reference corpora)

Proposition for the methodology begining

- ***Define the subject and the goal***

➢ One must decide what is going to be explored, define the goal of the work (exploring a topic, examining publications from an institution or a group of authors, conducting a literature review, etc), and the kind of data and/or documents.

For my example case, I have chosen to explore a research topic in order to identify the most significant authors and the distribution of publications. The topic is the academic research about board games, the period is 2018-2022. The kind of data are bibliographic references (including: authors, years and titles) and the type of document is journal publications.

- ***Identify keywords and expressions to use***

Next, we need to define the vocabulary at the heart of this topic and the time period to be studied. To do this, it can be helpful to rephrase the problem in one or two different ways and to use a tool for this.

With my example, I use ChatGPT 3,5 to have two new problem/topic expressions. The result is:

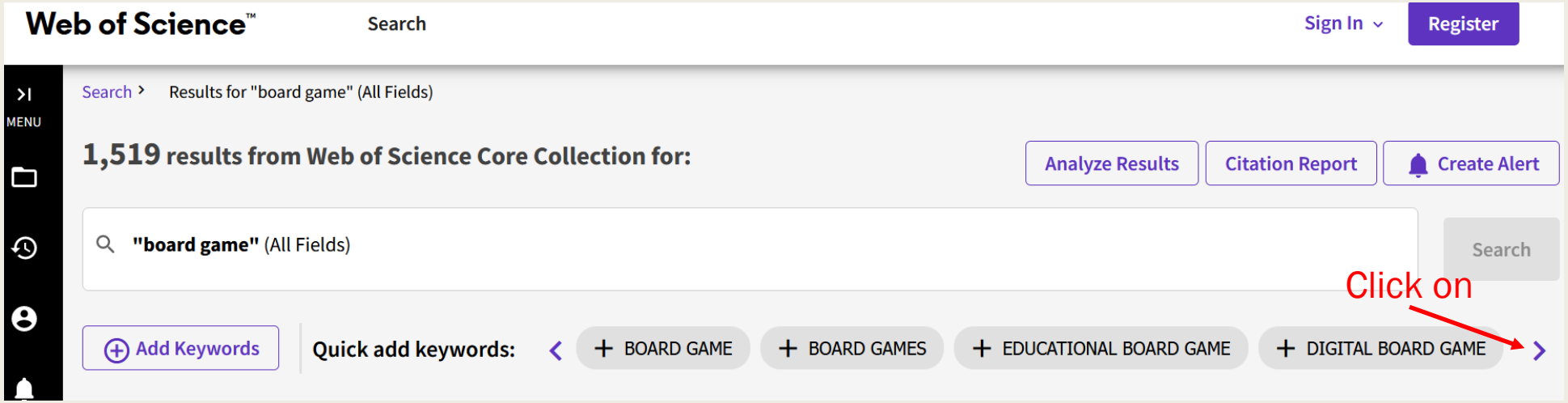
1. "Academic Inquiries into Board Games"
2. "Scholarly Investigations into Board Games"

Proposition for the methodology begining

- *Identify keywords and expressions to use*

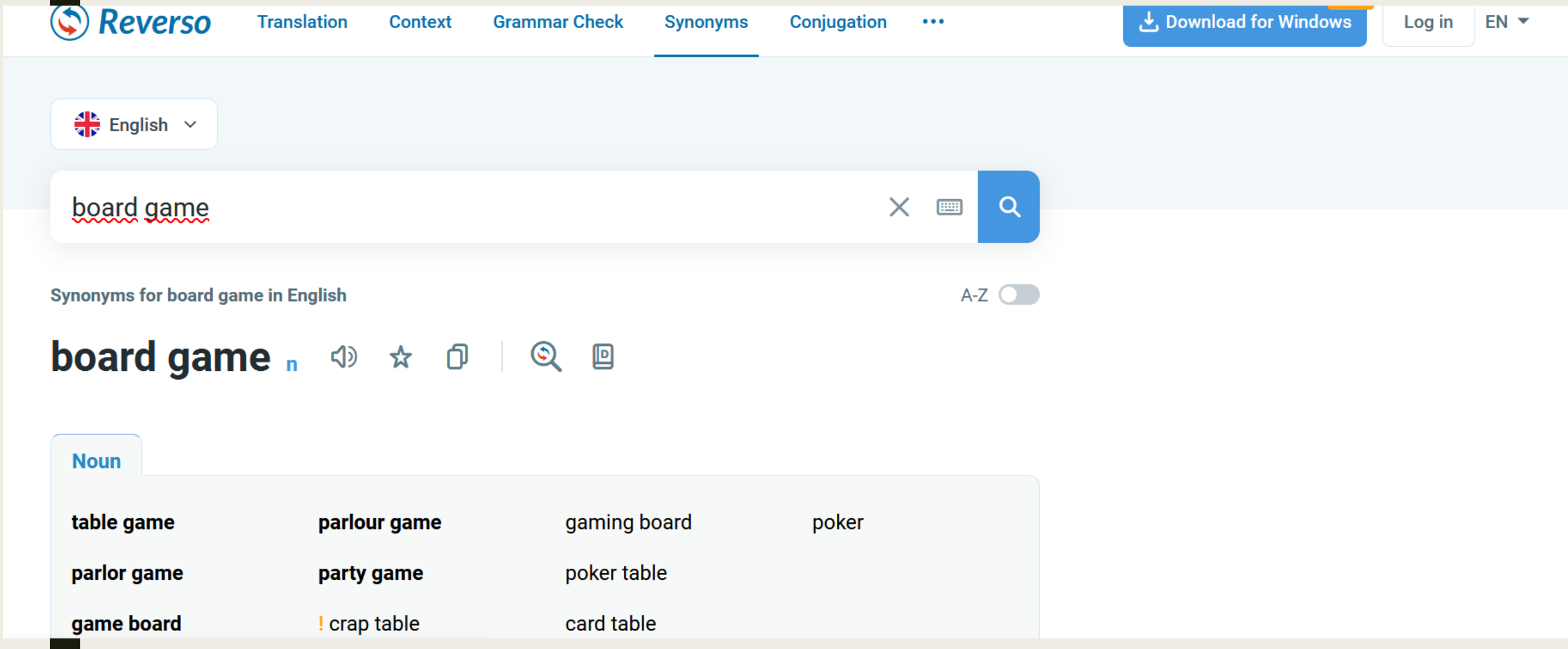
ChatGpt didn't help me in this example case, but sometimes, it is very useful to reconsider a topic.

For the keywords and other expressions, I'm focus on "board game" expression. I need to found few tools for this operation. I can use the Web of Science platform to select few keywords with it proposes to me. Used this way, this tool allows you to add a few terms to your list, but it is rarely entirely satisfying, even though it already helps. With my example, thus I add to my keywords list: table top game. I'll use two others online tool.



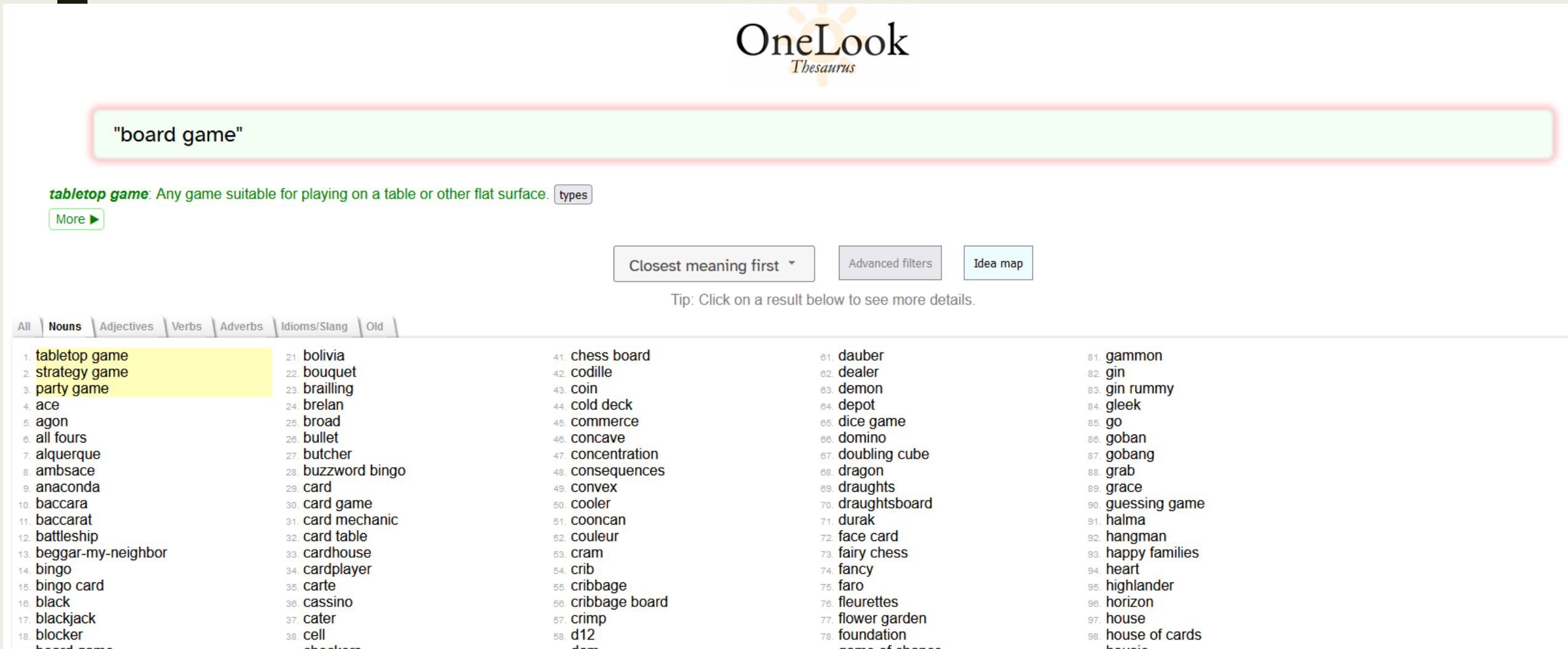
Few notions for methodology : keyword tools

Reverso



Few notions for methodology : keyword tools

OneLook Thesaurus



Proposition for the methodology begining

- *Identify keywords and expressions to use*

➢ Now, after selection, I have a principal list of keywords.

My **principal list** includes: board game, boardgame, card game, dice game, game board, gameboard, parlor game, parlour game, party game, table game, tabletop game.

But I have a second list specific to certain games: baccarat, blackjack, crap table, checkers, chess game, domino, draughts, gin rummy, goban, kriegspiel, majhong, mahjongg, poker, reversi, scopa, scrabble, shatranj, shogi. By reading it, we can add to this list other famous games that we are aware of: awale, oware, backgammon, clue game, cluedo, game of the goose, goose game, monopoly, stratego, trivial pursuit, Xiangqi.

And if I want a more complete study, I can you the fusion of the two last lists as my secondary list.

My **secondary list** includes: awale, oware, baccarat, backgammon, battleship, blackjack, clue game, cluedo, crap table, checkers, chess game, gameboard, draughtsboard, domino, draughts, game of the goose, goose game, gin rummy, goban, guess who, happy families, kriegspiel, majhong, mahjongg, monopoly, poker, reversi, risk game, scopa, scrabble, shatranj, shogi, stratego, trivial pursuit, wargame, Xiangqi.

It is preferable that the keywords in this second list also include the term game to limit response errors due to homonyms and surnames.

Few notions for methodology : sourcing

- For the next steps, we will focus on the main list. Anyway, the method used is exactly the same if we merge the two lists (main and secondary); only the number of collected references and duplicates will be higher.

- ***Sourcing (source and tool identification)***

Now, I need to identify the scientific platforms to query and/or specific tools to use. In any case, for each of these, I need to understand the syntax of their query language (possibly dedicated APIs) and determine if the interface accepts expressions in quotation marks, Boolean operators, truncation operator (especially for plurals), and fields that can be specifically queried. To do this, we first need to conduct tests.

As for the fields to be queried, these will include the fields: title, abstract, and keywords. Additionally, there will be the year/period field and the document type field.

All platforms allow for document export (often in CSV and/or BibTeX formats).

As specific tools or intermediaries that enable you to collect data through export or direct querying, you have the choice of:

- Bibliographic reference management tools, including **EndNote**, **Mendeley**, and **Zotero**.
- Tools for querying multiple platforms, including **Harzing's Publish or Perish** tool and **VosViewer**.

Few notions for methodology : investigation

- *Example with Hal*

Despite what one might think, this platform doesn't actually accept the truncation operator. Likewise, to specify the document type, it's better to select it from the right-hand menu, and do the same for the years. We need to add plurals to this main list as well as 'ing' forms that we find interesting.

Thus, with OR (Boolean operator) as a separator, the principal list begins :

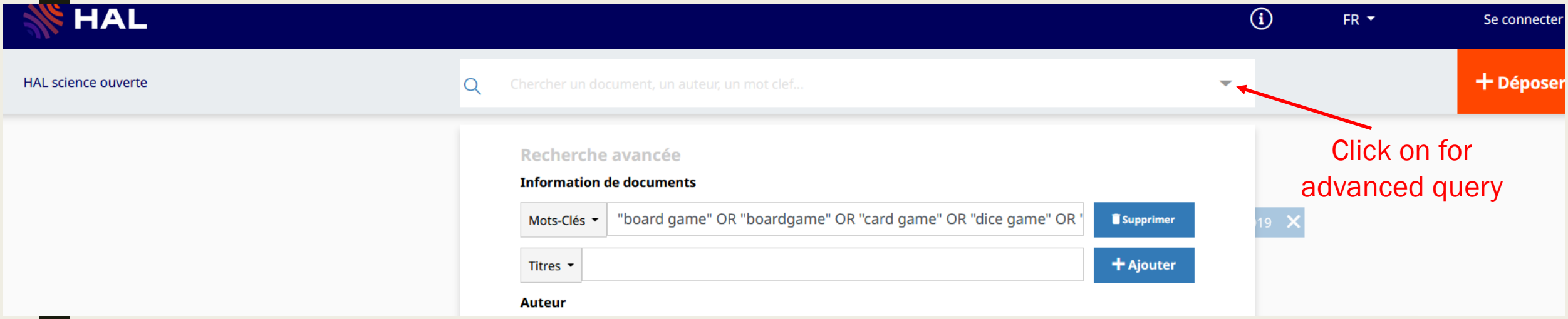
"board game" OR "boardgame" OR "card game" OR "dice game" OR "game board" OR "parlor game" OR "parlour game" OR "party game" OR "table game" OR "tabletop game" OR "board games" OR "boardgames" OR "card games" OR "dice games" OR "games board" OR "parlor games" OR "parlour games" OR "party games" OR "table games" OR "tabletop games" OR "board gaming" OR "boardgaming" OR "card gaming" OR "dice gaming" OR "game board" OR "parlor gaming" OR "parlour gaming" OR "party gaming" OR "table gaming" OR "tabletop gaming"

Few notions for methodology : investigation

- *Example with Hal*

And for the second keywords list :

(game OR games OR gaming) AND (awale OR oware OR baccarat OR backgammon OR Battleship OR blackjack OR Clue OR cluedo OR “crap table” OR checkers OR chess OR gameboard OR domino OR draughts OR “game of the goose” OR "guess who" OR “goose game” OR “gin rummy” OR “goban” OR "happy families" OR kriegsspiel OR majhong OR mahjongg OR monopoly OR poker OR reverse OR “risk game” OR scopa, scrabble OR shatranj OR shogi OR stratego OR trivial pursuit OR wargame OR xiangqi)



Few examples with tool

- *Hal.science = 1 platform*

It is also necessary to compile the results obtained from 3 queries focusing on the fields title, abstract, and keywords, respectively. Fortunately, this platform offers many APIs to save you time. These 3 equations will yield 3 result tables in CSV format. You will need to construct a file by adding and unbolting their results.

Query 1 : https://api.archives-ouvertes.fr/search/?q=**title_t**:("board game" OR "boardgame" OR "card game" OR "dice game" OR "game board" OR "parlor game" OR "parlour game" OR "party game" OR "table game" OR "tabletop game" OR "board games" OR "boardgames" OR "card games" OR "dice games" OR "games board" OR "parlor games" OR "parlour games" OR "party games" OR "table games" OR "tabletop games" OR "board gaming" OR "boardgaming" OR "card gaming" OR "dice gaming" OR "game board" OR "parlor gaming" OR "parlour gaming" OR "party gaming" OR "table gaming" OR "tabletop gaming")&fq=submittedDateY_i:[2018 TO 2022]&fq=docType_s:ART&wt=csv&fl=label_s

Query 2 : https://api.archives-ouvertes.fr/search/?q=**abstract_t**:("board game" ... OR "tabletop gaming")&fq=submittedDateY_i:[2018 TO 2022]&fq=docType_s:ART&wt=csv&fl=label_s

Query 3 : https://api.archives-ouvertes.fr/search/?q=**keyword_sci**:("board game" ... OR "tabletop gaming")&fq=submittedDateY_i:[2018 TO 2022]&fq=docType_s:ART&wt=csv&fl=label_s

Few examples with tool

- *Web of Science = 1 plateform*

With this platform, with advanced search, you can use the truncation operator. You can also check the 'Analyze Results' section.

Query: ((((TI=("board gam*" OR "boardgam*" OR "card gam*" OR "dice gam*" OR "game board*"OR "gameboard*" OR "parlor gam*" OR "parlour gam*" OR "party gam*" OR "table gam*" OR "tabletop gam*")) OR AB=("board gam*" OR "boardgam*" OR "card gam*" OR "dice gam*" OR "game board*"OR "gameboard*" OR "parlor gam*" OR "parlour gam*" OR "party gam*" OR "table gam*" OR "tabletop gam*")) OR KP=("board gam*" OR "boardgam*" OR "card gam*" OR "dice gam*" OR "game board*"OR "gameboard*" OR "parlor gam*" OR "parlour gam*" OR "party gam*" OR "table gam*" OR "tabletop gam*")) AND PY=(2018-2022)) AND DT=(Article)

However, for importing references (Citation Report -> Export Full Report), you will need to do them in groups of up to 1000 at a time and then compile them.

I have 1118 references with this query, so I need to export records 1-1000 + 1001-1118.

Few examples with tool

- ***Harzin's Publish or Perish (PoP) = 6-7 platform as sources***

If you can use Windows you can download this tool

Platforms for search:

- Crossref
- Google Scholar
- PubMed
- OpenAlex
- Scopus
- Semantic Scholar
- (Web of Science)

Few examples with tool

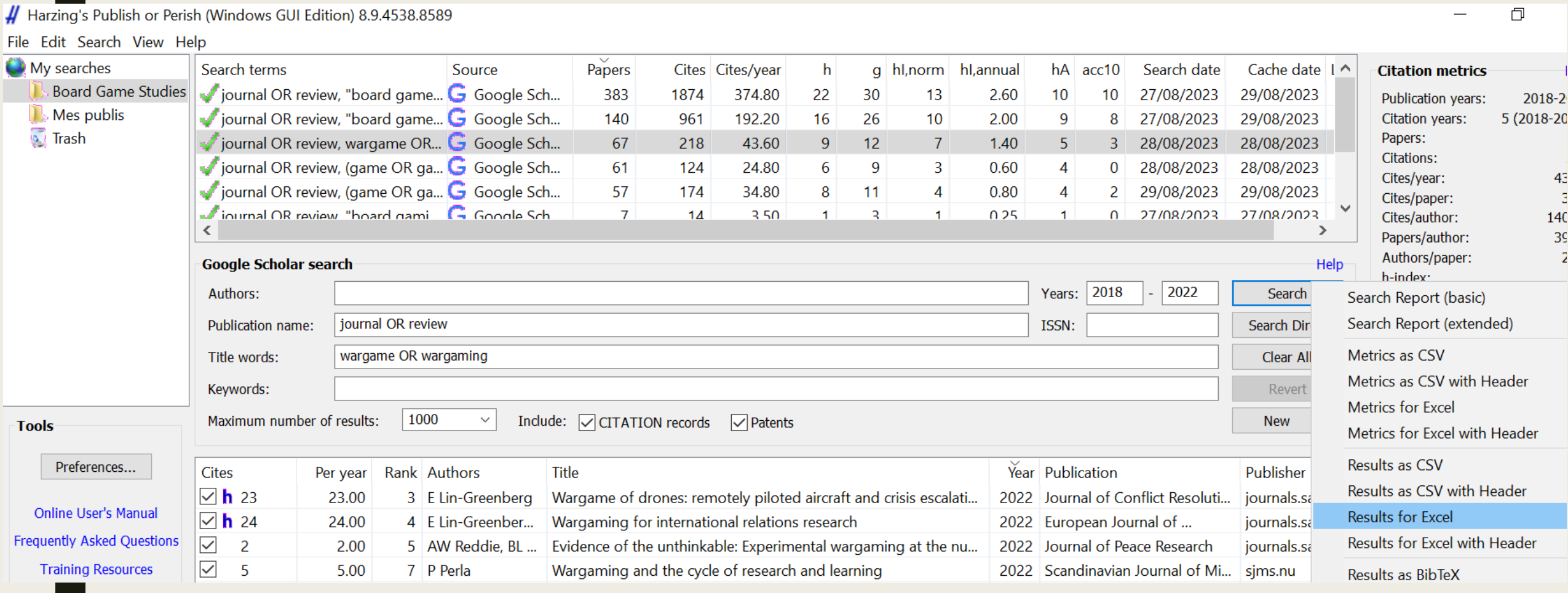
- ***Harzin's Publish or Perish (PoP) tool for Google Scholar***

For a large cover interrogate Google Scholar with PoP can be a good idea, but:

- Each query needs to be crafted based on the platform being queried.
- Its queries are limited query (title words and keywords fields) to 256 characters and imports are limited to 1000 if specified. Therefore, the search may require more than one query per platform, per year and per the field title words and keywords.
- To interrogate Google Scholar, you need to perform a first query on titles and a second one on keywords. As for the document type, we can work around it by specifying that the publication name must include the term 'journal' or 'review' .
- As this tool will generate multiple queries to query Google Scholar, it is also necessary to spread the queries over time (one per hour, for example) in order to avoid being rejected by Google as a robot.
- At the end, I can compile results of each query with a copy (select results and click on Copy Results (for Excel as an example).

Few examples with tool

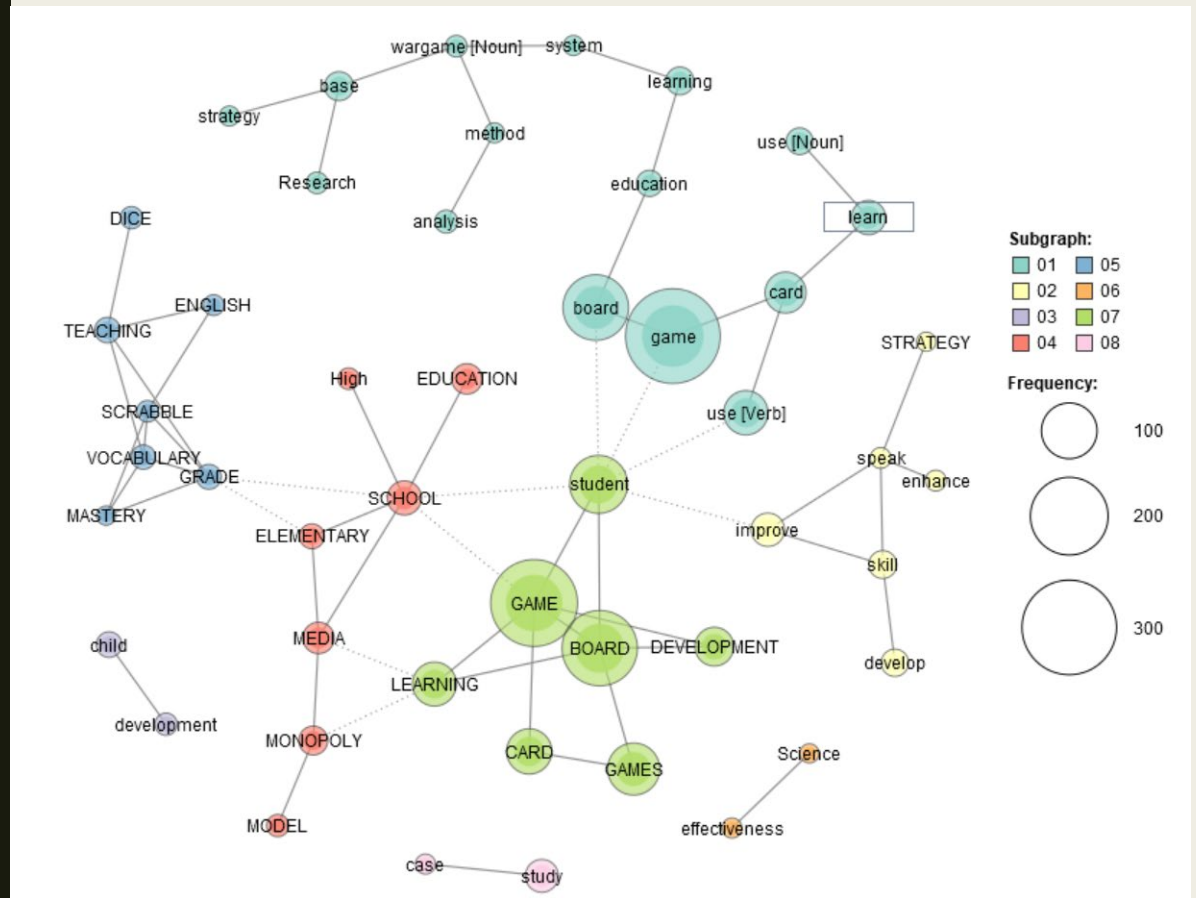
- *Harzin's Publish or Perish (PoP) tool : export results*

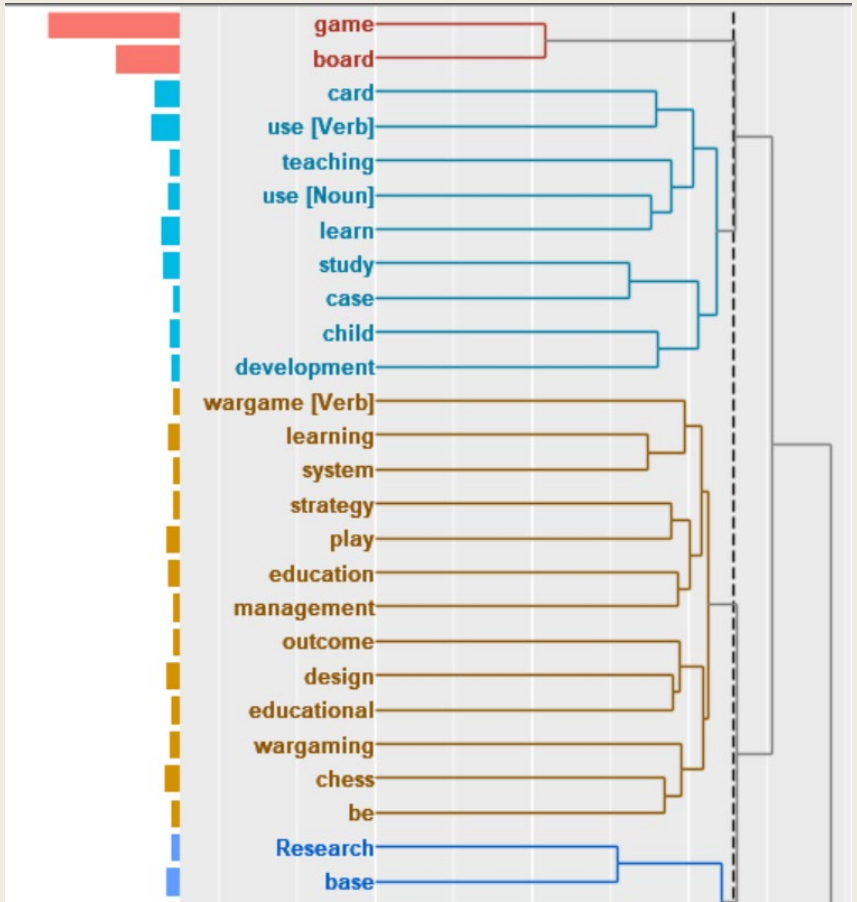


Few examples with tool

- *After you can adapt your data and use a text mining tool like KH coder*

N.B.: Data processing with this software can be very time-consuming.





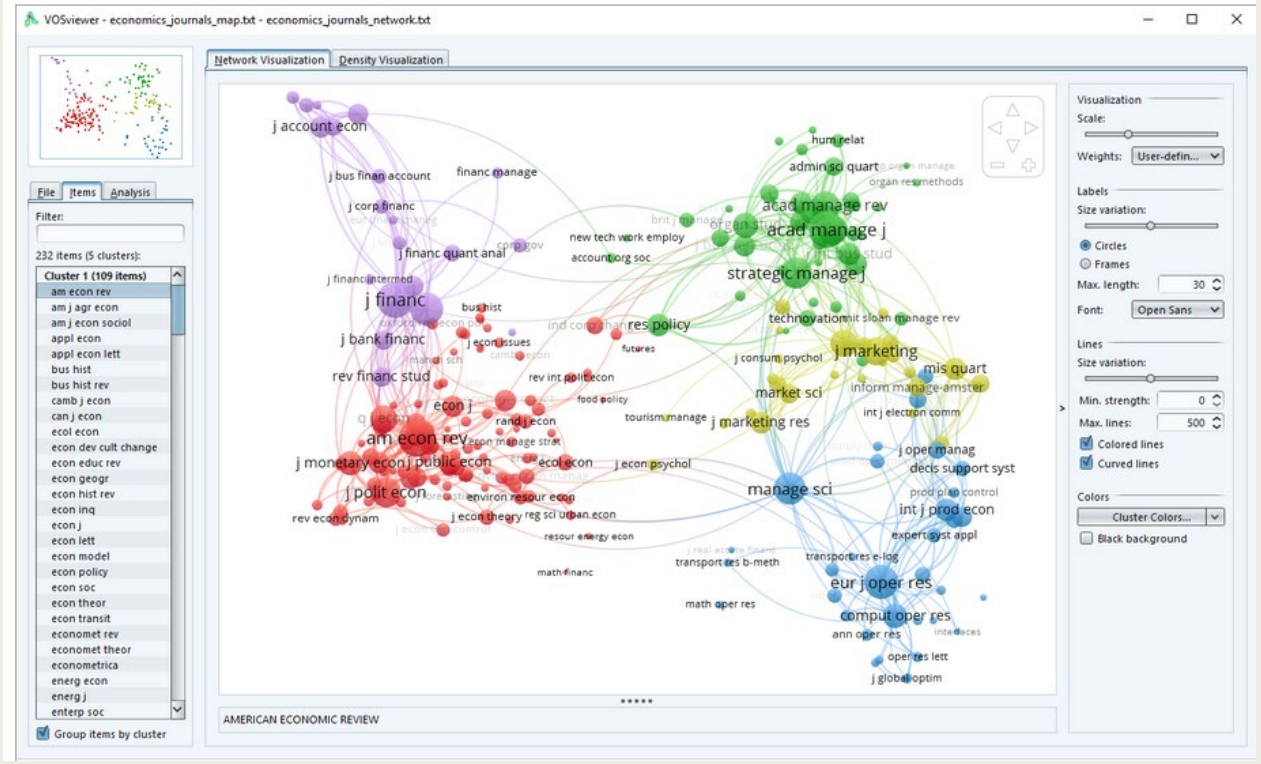
Few examples with tool

- *VOSviewer = a network visual analytics tool*

For Windows or MacOS, but it seems exist a package for Linux.

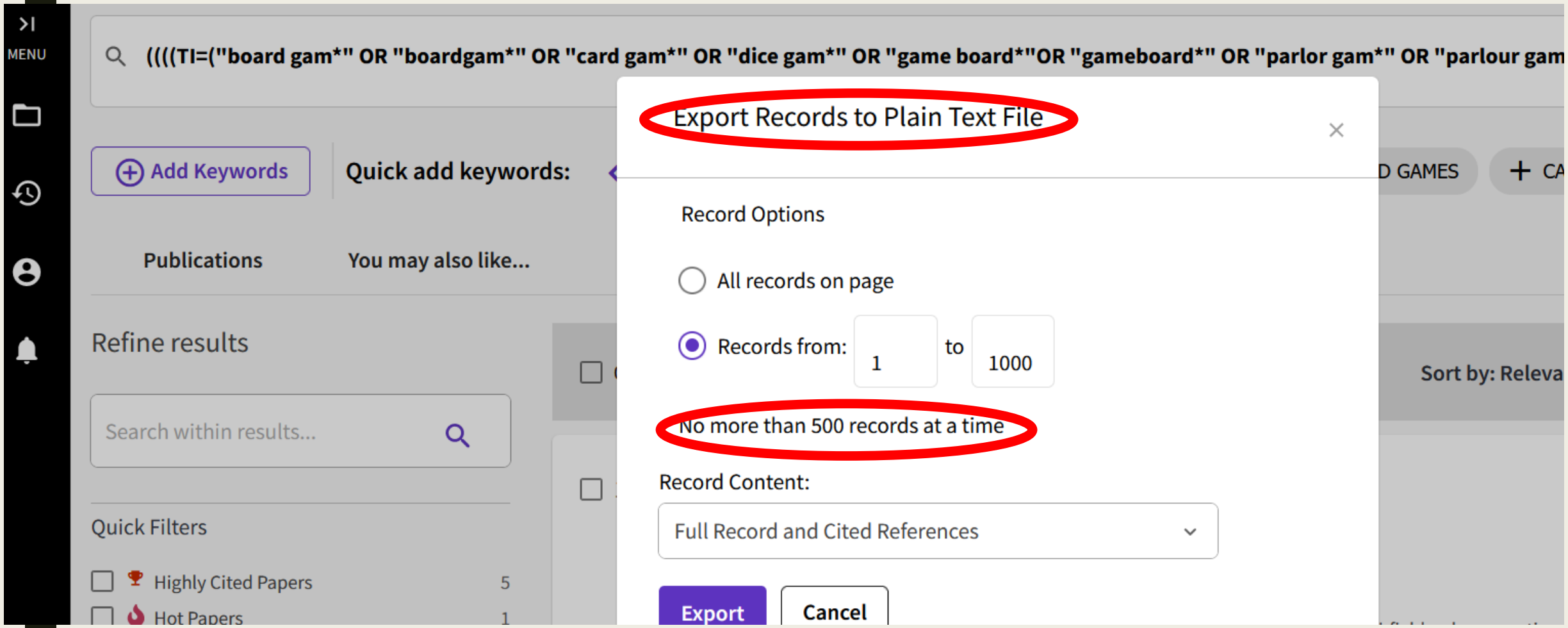
It ca use data from:

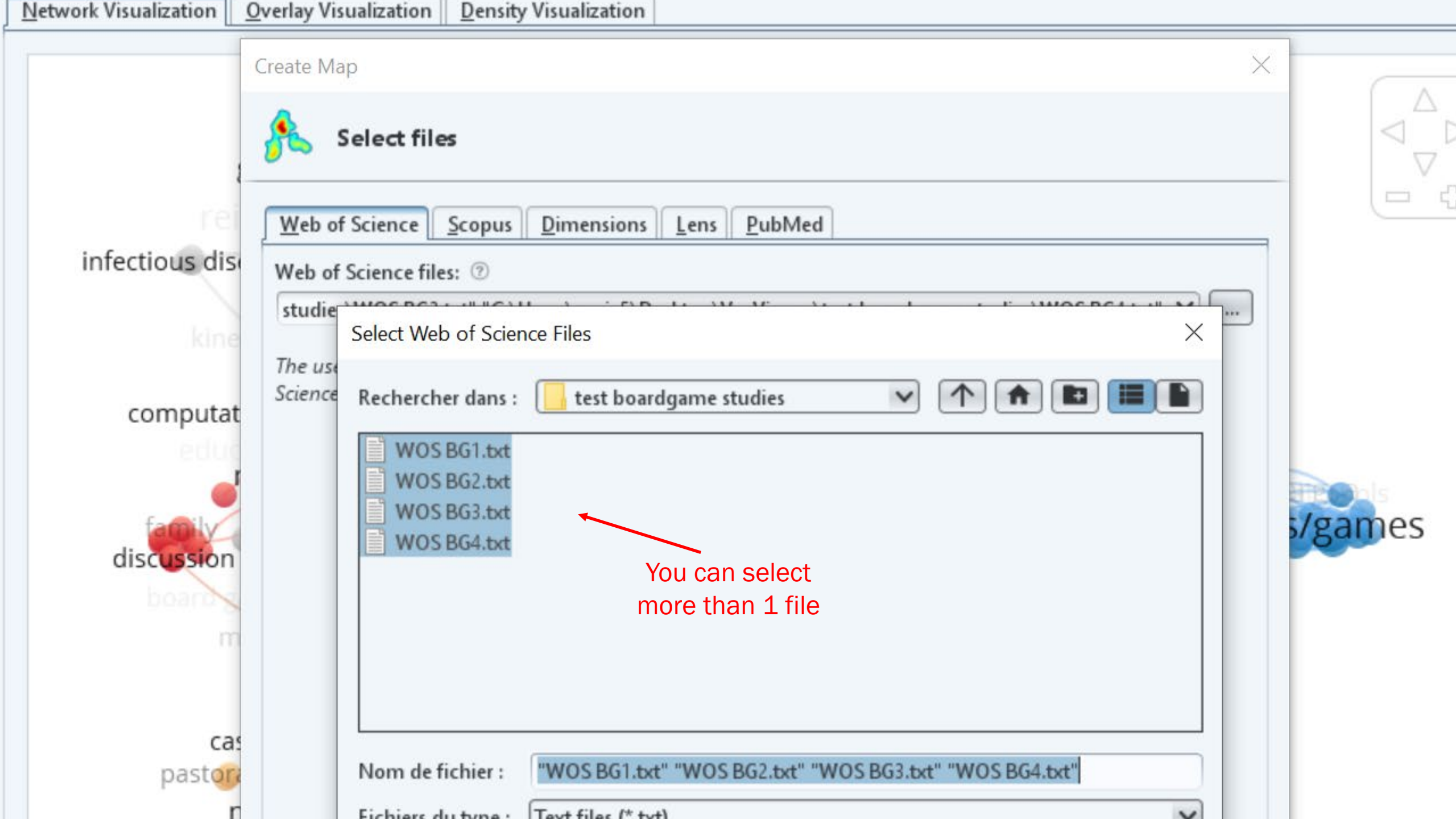
- Web of Science (plain text files)
- Scopus (csv files)
- Dimensions (if you have a user account)
- Lens (csv files)
- PubMed (Pubmed format files)
- Publish or Perish (Endnote or RIS files)
- Mendeley (RIS files)
- Zotero (RIS files) …



Few examples with tool

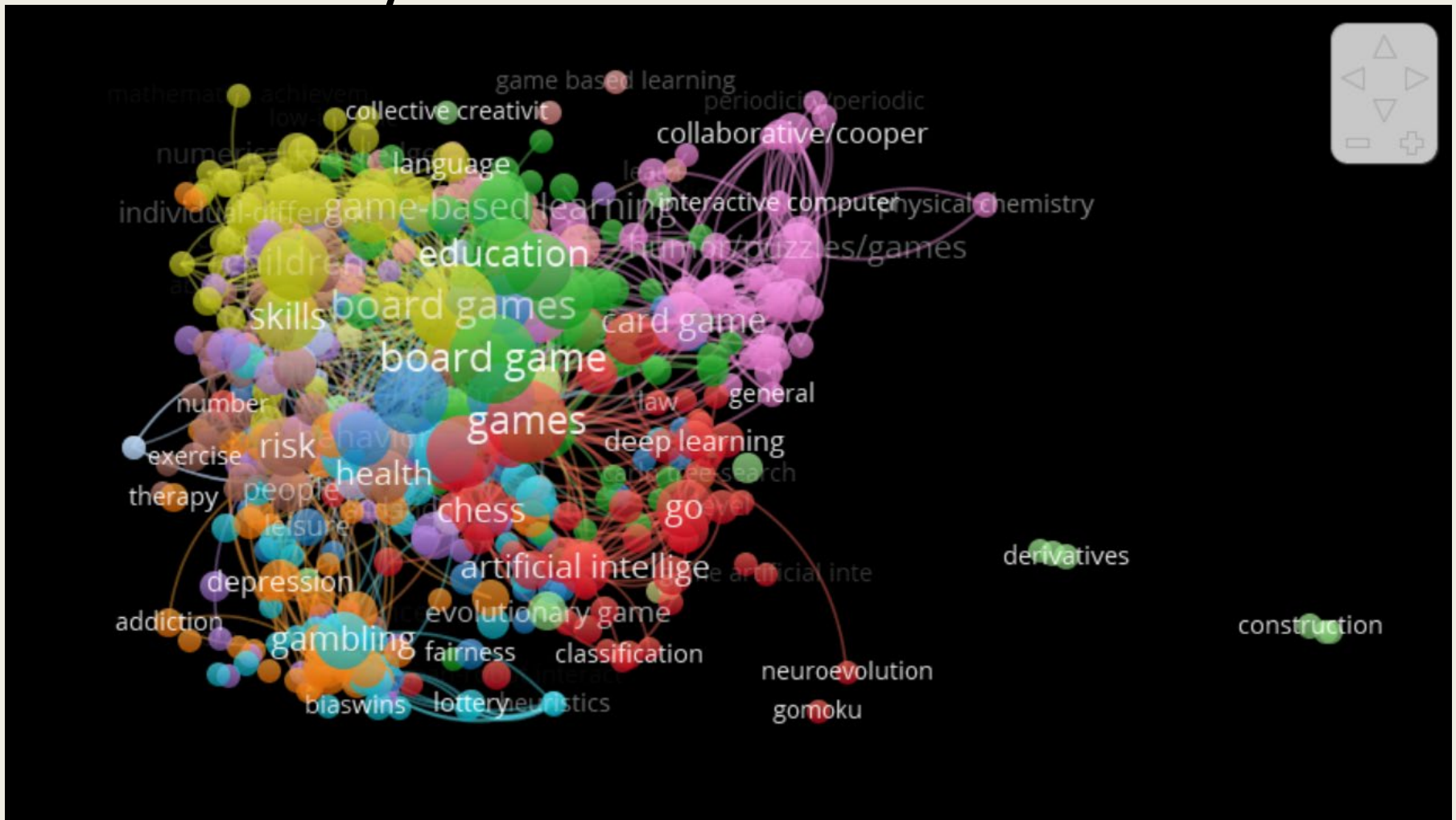
- *VOSviewer: example with fata from Web of Science*




Network Visualization
Overlay Visualization
Density Visualization
Create Map
Select files
Web of Science
Scopus
Dimensions
Lens
PubMed
Web of Science files:
Select Web of Science Files
Rechercher dans :
test boardgame studies
WOS BG1.txt
WOS BG2.txt
WOS BG3.txt
WOS BG4.txt
You can select
more than 1 file
Nom de fichier :
"WOS BG1.txt" "WOS BG2.txt" "WOS BG3.txt" "WOS BG4.txt"
infectious dis
computat
discussion
cas
pastor
s/games

Few examples with tool

- *VOSviewer: example with fata from Web of Science*

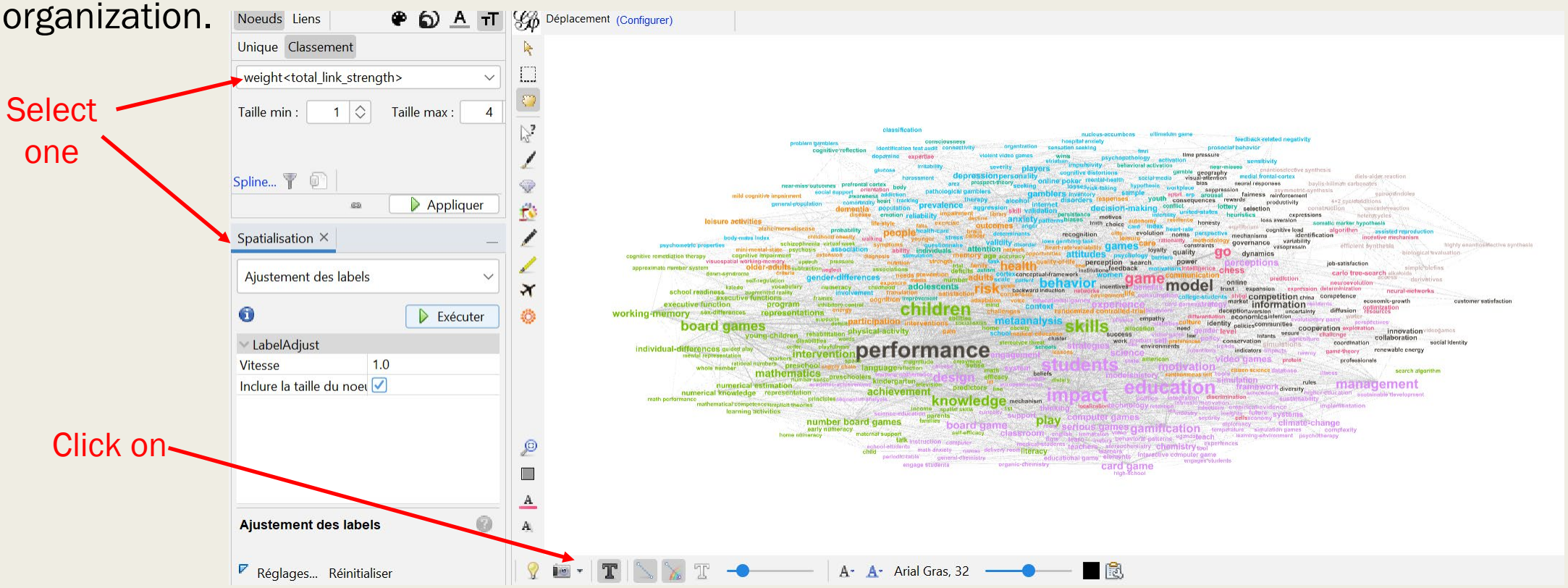


Few examples with tool

- ***From VOSviewer to Gephi :***

➢ save with VOSviewer as a GML file and open it with Gephi

➢ Duplicate id in Label column and chose an spatialization algorithm and attributes for visual organization.



Bibliography

Ball, R. (2020). *Handbook bibliometrics*. Walter de Gruyter GmbH & Co KG.

Ball, R. (2017). *An introduction to bibliometrics: New development and trends*. Chandos Publishing.

Ninkov, A., Frank, J. R., & Maggio, L. A. (2022). Bibliometrics: Methods for studying academic publishing. *Perspectives on medical education*, 11(3), 173-176.

Mingers, J., & Leydesdorff, L. (2015). A review of theory and practice in scientometrics. *European journal of operational research*, 246(1), 1-19.

Nalimov, V., & Mulcjenko, B. (1971). *Measurement of Science: Study of the Development of Science as an Information Process*. Washington DC: Foreign Technology Division.

Pritchard, A. (1969). Statistical bibliography or bibliometrics? *Journal of Documentation*, 25, 348-349.

Rostaing, H. (2021). *La bibliométrie au service des Sciences de l'Information et de la Communication*. In Captas, documentation et numérique.

Rousseau, R. (2021). Naukometriya, Nalimov and Mul'chenko. *Collnet Journal of Scientometrics and Information Management*, 15(1), 213-224.

Your turn

Now you can carry out a bibliometric analysis on your own research topic