www.artirev.com

User Manual: Artirev V1.0.0

January 2021

Maintained by BIBREV (Non-Profit Association)
Association N°W061014843
bibrev.asso@gmail.com

Contents

I. Introduction:

ARTIREV (**ART**ificial **I**ntelligence and Literature **REV**iew) is a tool that can assist you in conducting a literature review with the help of bibliometrics and artificial intelligence. This tool downloads metadata directly from SCOPUS (Elsevier), and groups relevant texts in clusters. It creates visual maps to help you understand and analyze your research field.

If you would like to know how this tool was constructed, you can read Walsh & Renaud (2017), which served as basis for it (Walsh, I., & Renaud, A. (2017). Reviewing the literature in the IS field: Two bibliometric techniques to guide readings and help the interpretation of the literature. Systèmes d'Information & Management, 22(3), 75-115).

As prerequisites, to run ARTIREV, you need:

- Java Runtime Environment version 8 (1.8) minimum

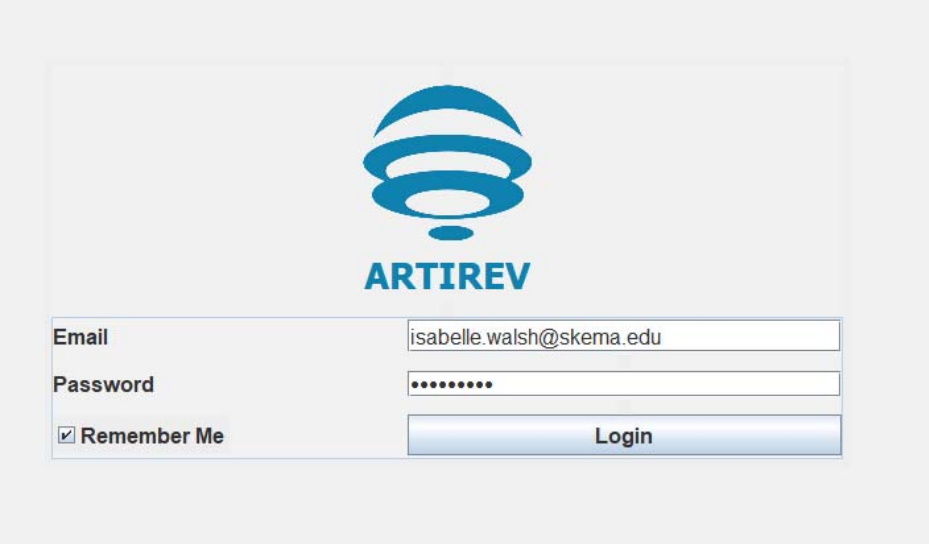
You will find the JRE for Windows, Mac OS X et Linux at the following link: https://www.java.com/fr/download/manual.jsp

- A web browser of your choice

You will be able to download ARTIREV, once you have received the relevant link:

- Download the .zip file
- Unzip the file at he place you want it on your computer. You can rename this file after unzipping. It is important that the /www subfile is in the same file as the .jar subfile.
- Execute (double click) "artirev-standalone_ jar_1.0 "

To log in (see below), you will need you identification/email and password, which should have been forwarded to you by your institutional administrator.



II. Workflow

After logging in, there are 4 steps/modules in the worklow: extraction (Extracts meta data from Scopus: section II-1), cleaning (Automatical cleaning of downloaded data: section II-2), Matrices (generating reference co-citaiton and document bibliographic coupling matrices: section II-3), and clustering, layout and sensemaking (section II-4). To make it easier to use, we show a picture of the screen at each step based on a simple keyword search: “serendipity”.

Step	Module	Description
	Login	▪ Authentification
1	Extraction	▪ Extraction of documents with their bibliographic notices from Scopus (Elsevier)
2	Cleaning	▪ Automatic cleaning of meta data and cited references ▪ Complementary semi-automatic cleaning of references with a thesaurus
3	Matrices	▪ Filtering of references and documents to include in the analysis ▪ Calculation of co-citation and bibliographic coupling matrices.
4	Clustering, layout & sensemaking	▪ Clustering of cited references and citing documents ▪ Calculation of their positions in the mappings. ▪ Layout with Leiden or Louvain algorithms ▪ Calculation of various indices (e.g., local hubness) ▪ Generation of mappings and word clouds to support interpretation (sensemaking)

II.1.Extraction

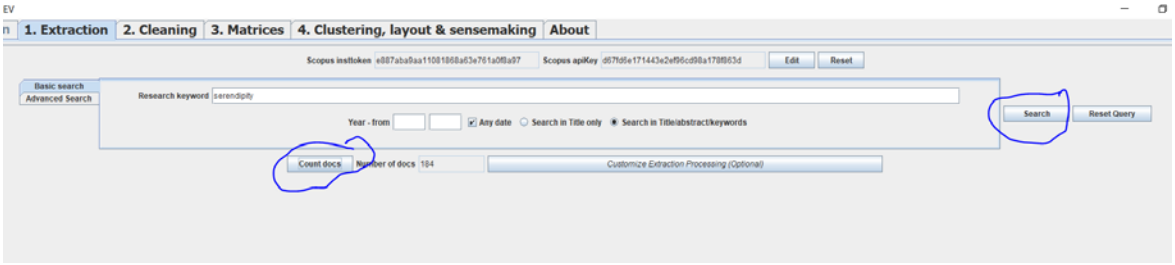
To extract relevant data, you need a subscription to Scopus. If your institution does not have such a subscription and you would like to use ARTIREV, please contact BIBREV (bibrev.asso@gmail.com).

First, you extract the metadata directly from SCOPUS in Artirev, instead of downloading it first on SCOPUS and then uploading it to the software. There are two possibilities: Basic search or Advanced search.

II.1.a. Basic search

In the basic search, you just have to enter your keyword(s). Check on the number of documents by clicking on “Count docs”. It is recommended not to go beyond 1300 documents, unless you have a very powerful computer with a lot of RAM, as otherwise your computer would not be able to process the data. If you find that you reach more than 1300 documents, readjust your keywords.

Once you have finalized your keywords, click on “search”. The software will download the necessary data from Scopus.



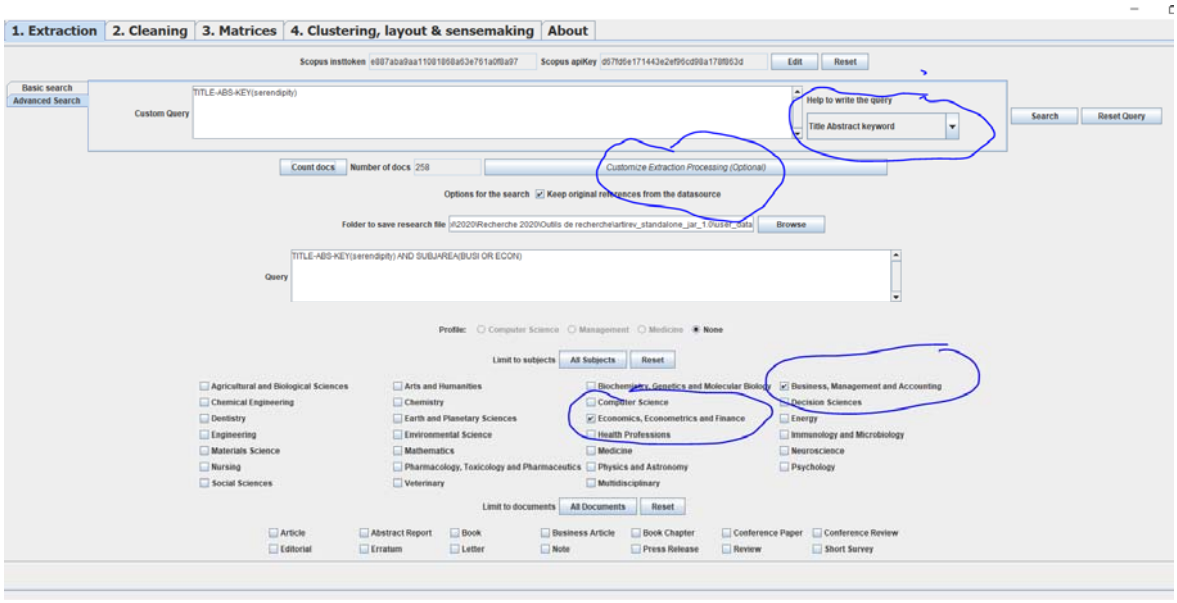
II.2.b. Advanced search

This option helps you to precise your research.

a. Insert your keywords (click on “Help to write the query”)

Click on “*Customize Extraction Processing (Optional)*” to restrict your search. You limit your search by checking the boxes you need.

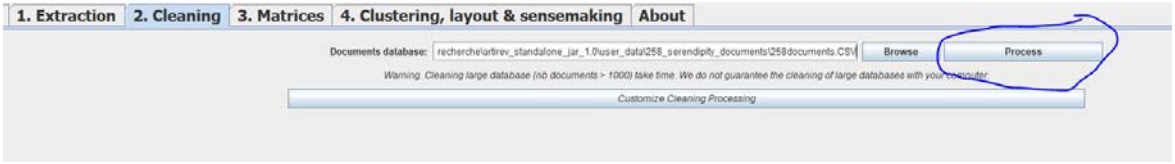
Don’t forget to click on “count” to check on the number of documents before you download the data. If this number of documents suits you, you can now click on “*Search*” to extract the data.



The .csv file containing the data, named “XXX documents”, will automatically upload in the folder “*user_data*” in Artirev’s folder on your computer and will automatically be entered in the Cleaning step.

II.2. Cleaning

Click on “Process”



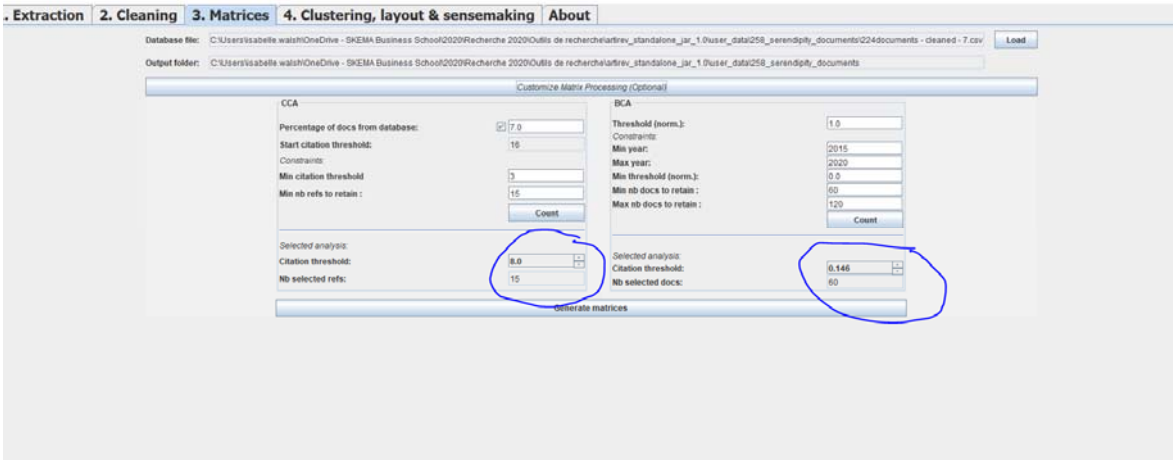
Artirev will automatically clean the data downloaded from Scopus.
If you have downloaded many documents, this may take quite some time. Let the program run on your computer until it has finished the cleaning.

Expert users can effect further cleaning with a semi-automatic thesaurus when clicking on "Customize Cleaning Processing".

II.3. Generate Matrices

After the software has finished cleaning the data in the previous step, we now may begin setting parameters to restrict our metadata to certain relevant documents and create co-citation (CCA) and bibliographic coupling matrices (BCA).
The clean database is automatically loaded. ARTIREV will choose some relevant citation thresholds by default. If the user wishes to use these thresholds, she will only have to click on "Generate matrices"

If the user would like to set her own citation thresholds, she can click on "Customize Matrix Processing (Optional)" to expand the menu and new fields will appear (see below), where the default settings can be modified. For CCA, the citation threshold is the raw number of citations. For BCA, the citation threshold is the normalized number of citations. The final field at the bottom of each column will display the total number of selected references/documents.



CCA and BCA parameters should be set in accordance with the nature of the topic and the purpose of the review. Users should recall that CCA is utilized to look into and understand the 'past' of a field, the theoretical and methodological grounding of this field (schools of thought), and BCA is utilized to understand the 'present' of the research field and current research themes. For CCA, ARTIREV proposes by default to look into the most cited references in the data base and highlights at least 15 such references. For CCA analysis,

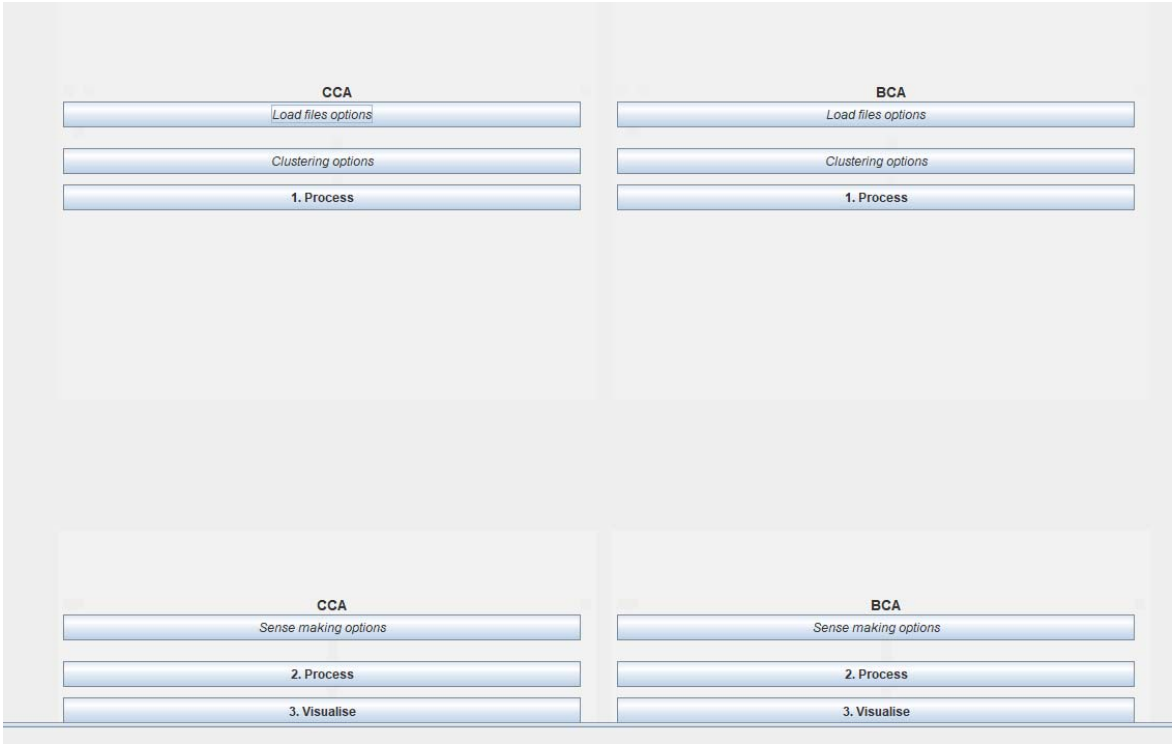
to obtain a comprehensive view of the past of the field (if it is a mature enough research field), the ideal total number of references is probably between 25 and 60. For BCA ARTIREV proposes by default searching the previous 5 years of publications. This is generally a good rule of thumb but can be extended, for example, if the topic calls for a more extended time frame and this can be justified. The minimum and maximum number of documents can be set, but by default we propose for BCA analysis between 60 and 120 documents. The citation thresholds are adjustable and will determine the final output.

If the user changes parameters, she should click “Count” and view the final number of output documents at the bottom, and decide if the number is suitable based on the nature of her study.

Once parameters are set and the user is satisfied with the numbers of total documents, click “Generate Matrices” for Artirev to generate the matrices to be used in the final step.

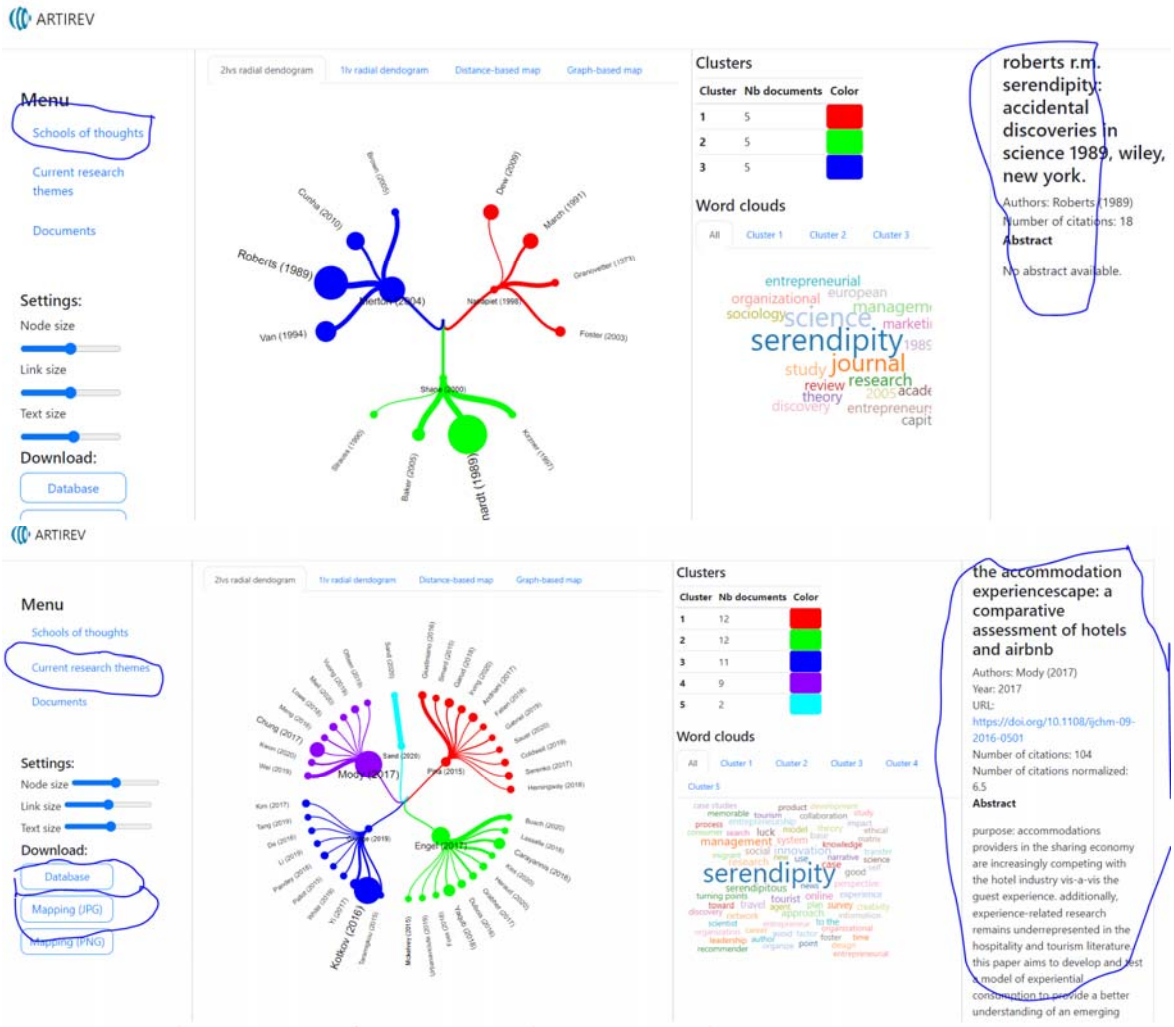
II.4. Clustering, layout and sensemaking

The matrices that have been generated serve as basis for the various maps proposed by ARTIREV. The network and map files should upload automatically.
The user should simply click in turn: 1. Process (Groups references and documents in clusters), 2. Process sense making (Creates word clouds for each cluster) and 3. Visualize (Opens your navigator with the bibliometric maps) for CCA and then for BCA.



A new internet window with the CCA maps: schools of thought (then the BCA maps: current research themes) to analyze appears. You obtain four maps (relation-based, distance-based, graph-based).

We will only comment on The 2 level relational map (radial dendrogram 2 levels: rd2lv) from Artirev: see below.



Nodes are grouped in clusters (schools of thought/ research themes) with different colors. When you click on a node in the mappings, you get relevant details on the right. In each cluster, the documents that have the most/strongest links with other texts in their cluster are at the center of the map. They give an indication of the content of each cluster, even though there are often subclusters. The thickness of the links describes their strengths. The bigger the node, the higher the citation count (raw for CCA and normalized for BCA), the more significant the document.

You can download the detailed contents of clusters by clicking on “Database” and you can download the mappings by clicking on “Mapping JPG”

You can then read the documents and analyze the clusters to schools of thought and research themes of the investigated field.