Instruction to generate figures of the paper:

A Deep Dive into OpenStreetMap Research Since its Inception (2008–2024): Contributors, Topics, and Future Trends

Prerequisite

Install software

Download and install the following software from the official websites, following the official guidelines to ensure correct installation and functionality.

Microsoft Excel: https://www.microsoft.com/en-us/microsoft-365/excel

Vosviewer: https://www.vosviewer.com/

Biblioshiny: https://www.bibliometrix.org/home/index.php/layout/biblioshiny

Python → jupyter notebook: https://jupyter.org/install

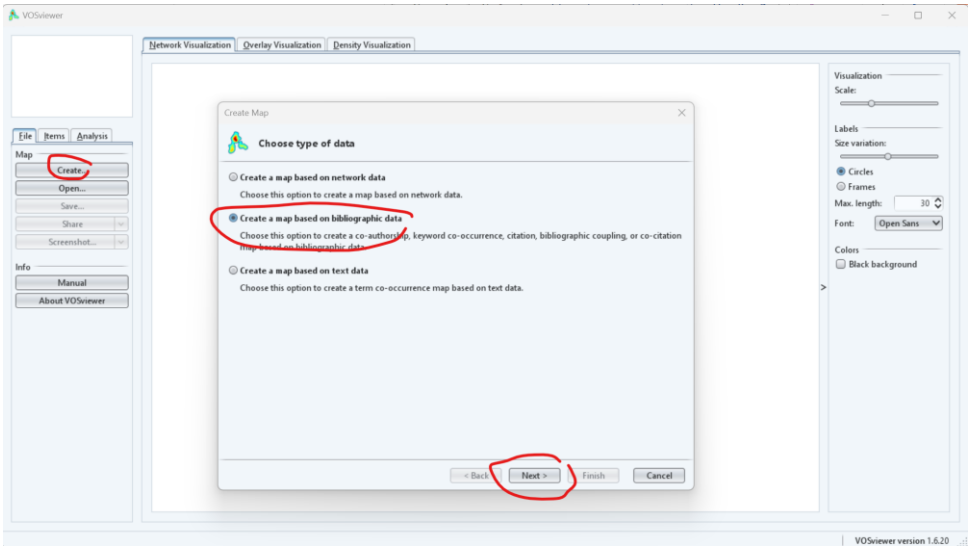
Software UI to start general analysis

Microsoft Excel

- Used for checking data and generating Figure 2.

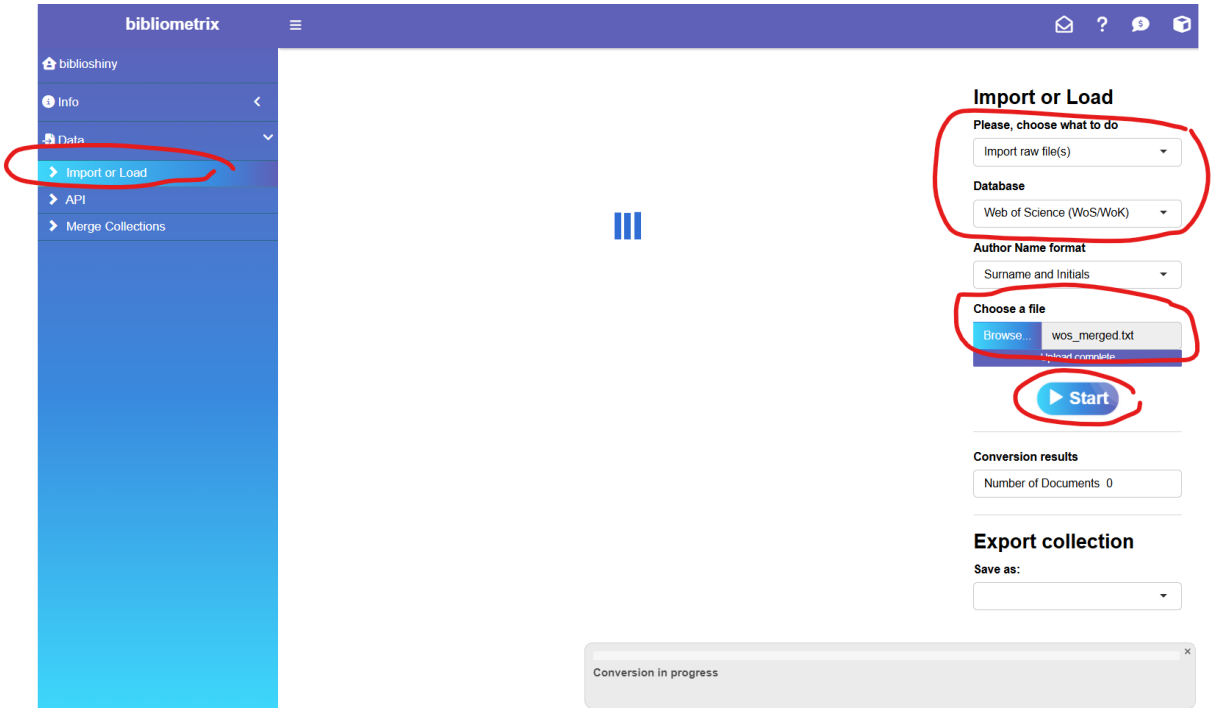
Vosviewer

- Used for generating Figure 5, 6, and 9.
- UI: Load data as shown in the figure below, click “Create”, then choose “Create a map based on bibliographic data”, then click “Next”.



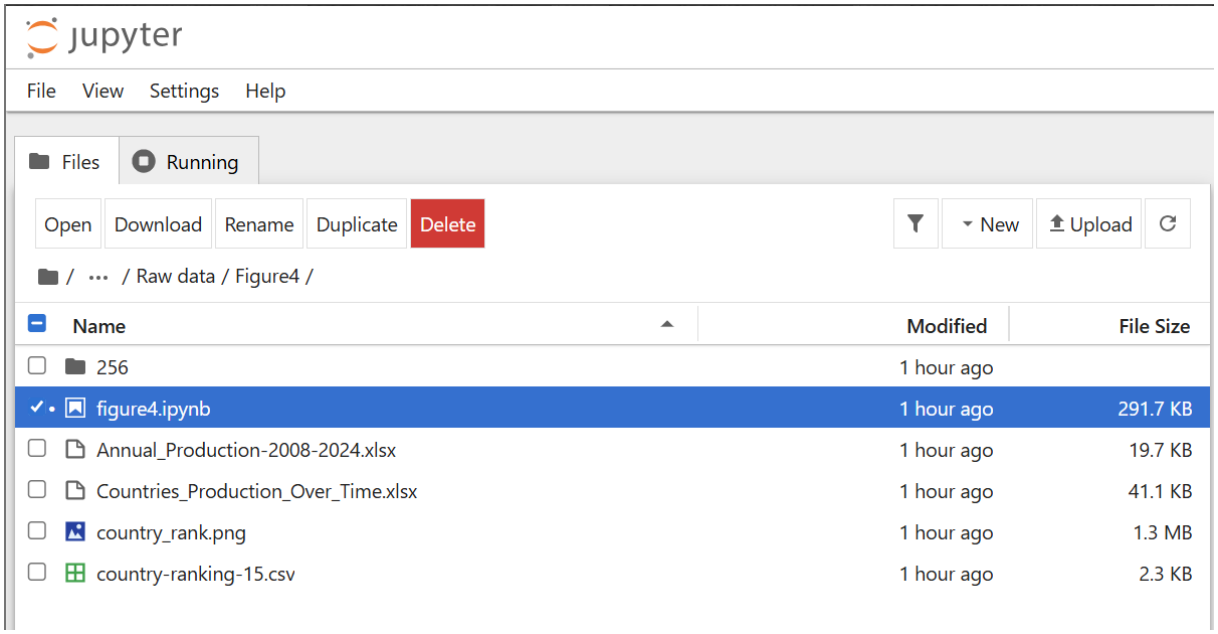
Biblioshiny

- Used for generating Figure 3, 7, 8, and 10.
- UI: as shown below, in "Import or Load," set "Import raw file(s)," set "Database" to "Web of Science (WoS/WoK)," and "Choose a file" to "wos_merged.txt."



Jupyter notebook

- Necessary for generating Figure 4.
- Run figure4.ipynb.



Data

Raw data:

- .\data and code\Raw data\wos
- .\data and code\Raw data\sotm

Ready to analyze data:

- .\data and code\wos_merged.txt
- .\data and code\sotm_merged.xlsx

Figures

- .\data and code\Raw data\

Figure 2

Step 1: load the data

In Biblioshiny software, in "Import or Load," set "Import raw file(s)," set "Database" to "Web of Science (WoS/WoK)," and "Choose a file" to "wos_merged.txt."

Step 2: run analysis

On the left panel, select "Overview" -> "Annual Scientific Production."

On the top panel, select "Table." Download the table: "Annual_Production-2008-2024.xlsx"

Step 3: generate the figure

Generate the line chart in Excel.

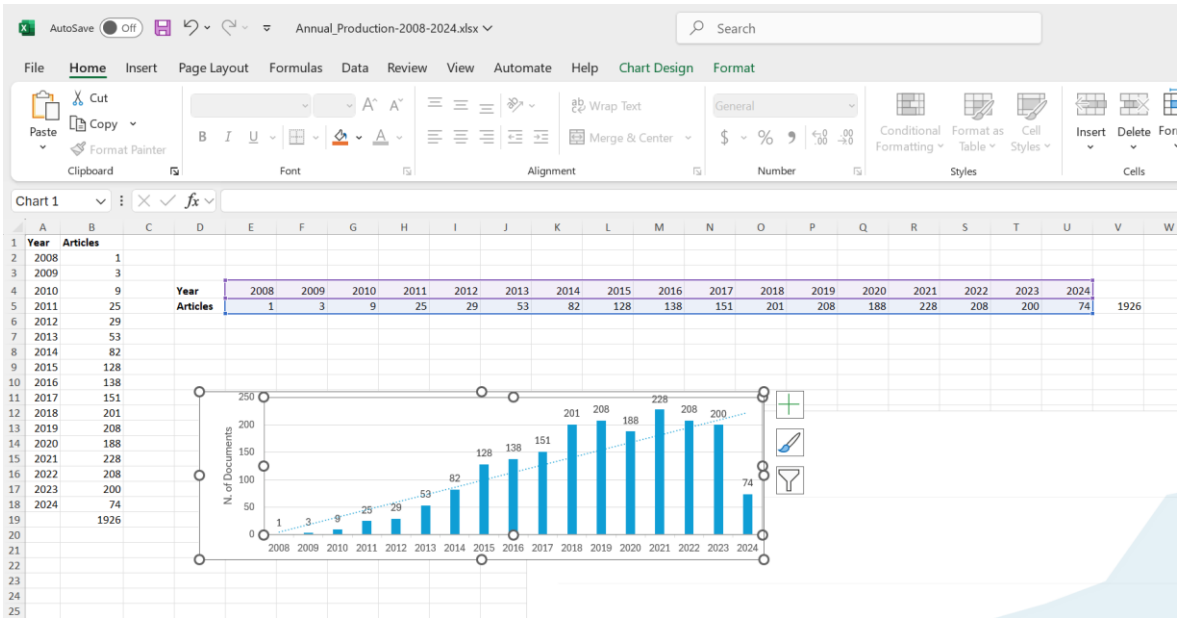


Figure 3

Figure 3 contains three sub-figures, as explained below.

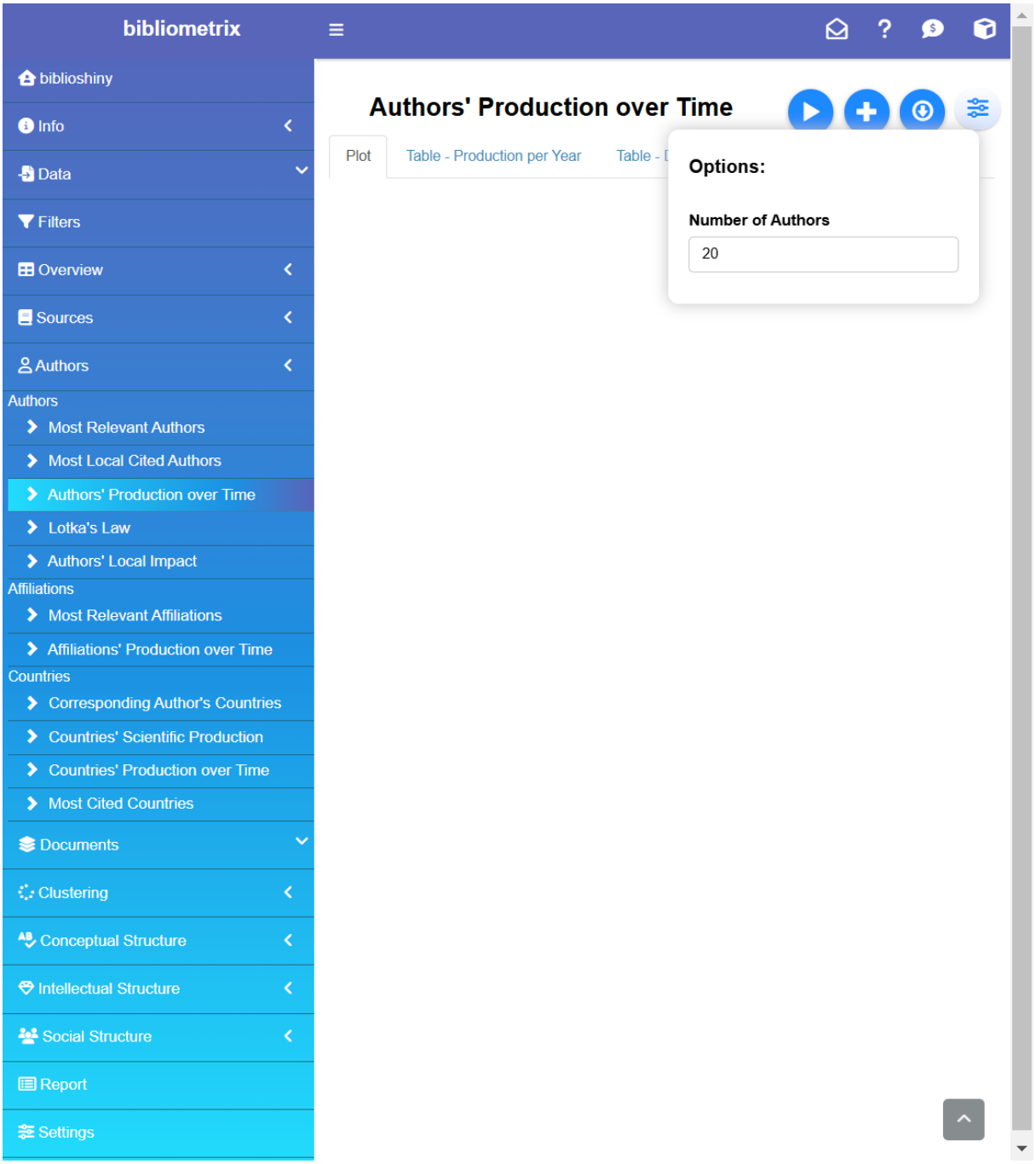
Subfigure 1:

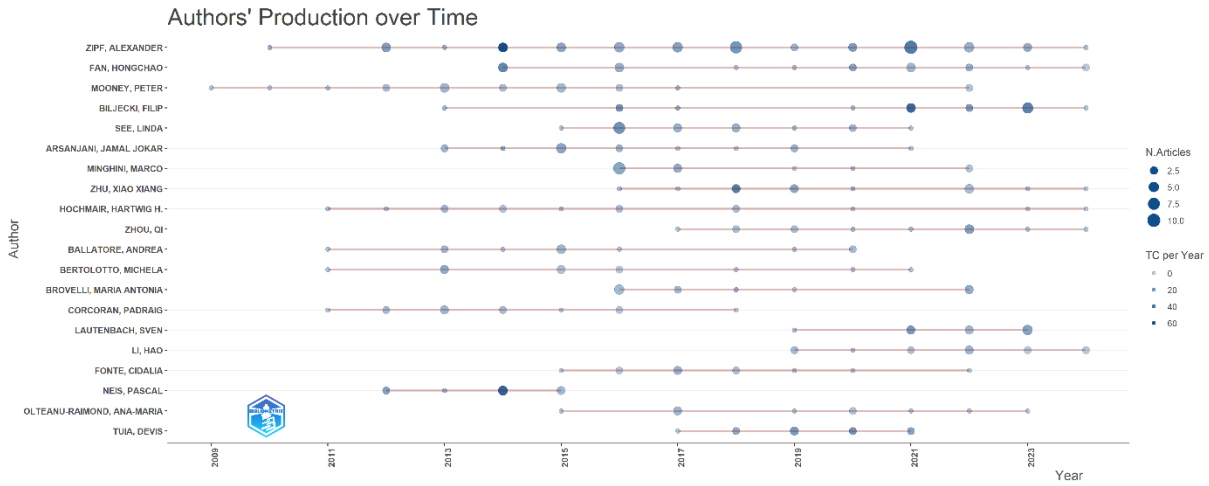
Step 1: load the data.

In Biblioshiny software, in "Import or Load," set "Import raw file(s)," set "Database" to "Web of Science (WoS/WoK)," and set "Author Name format" to "Fullname (if available)," then "Choose a file" to "wos_merged.txt."

Step 2: run analysis

On the left panel, select "Authors" -> "Authors' Productivity over Time," then set "Options: Number of Authors" to 20, and click "Run the Analysis."





Step 3: Download the tables of top authors.

On the top panel, select "Table - Documents," input the names of the top authors, and download their corresponding tables. Name the table using the format "lastname.xlsx."

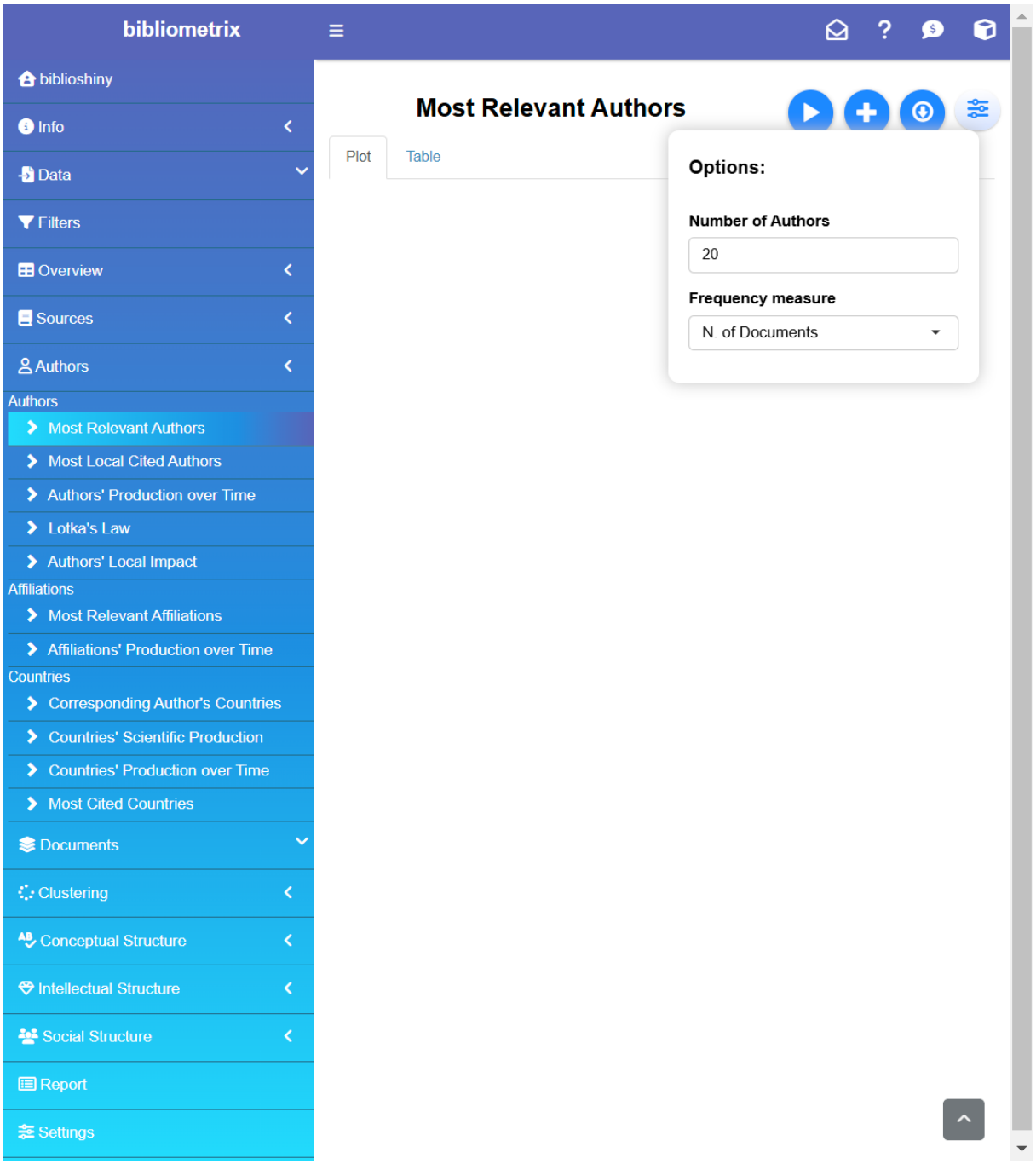
Subfigure 2:

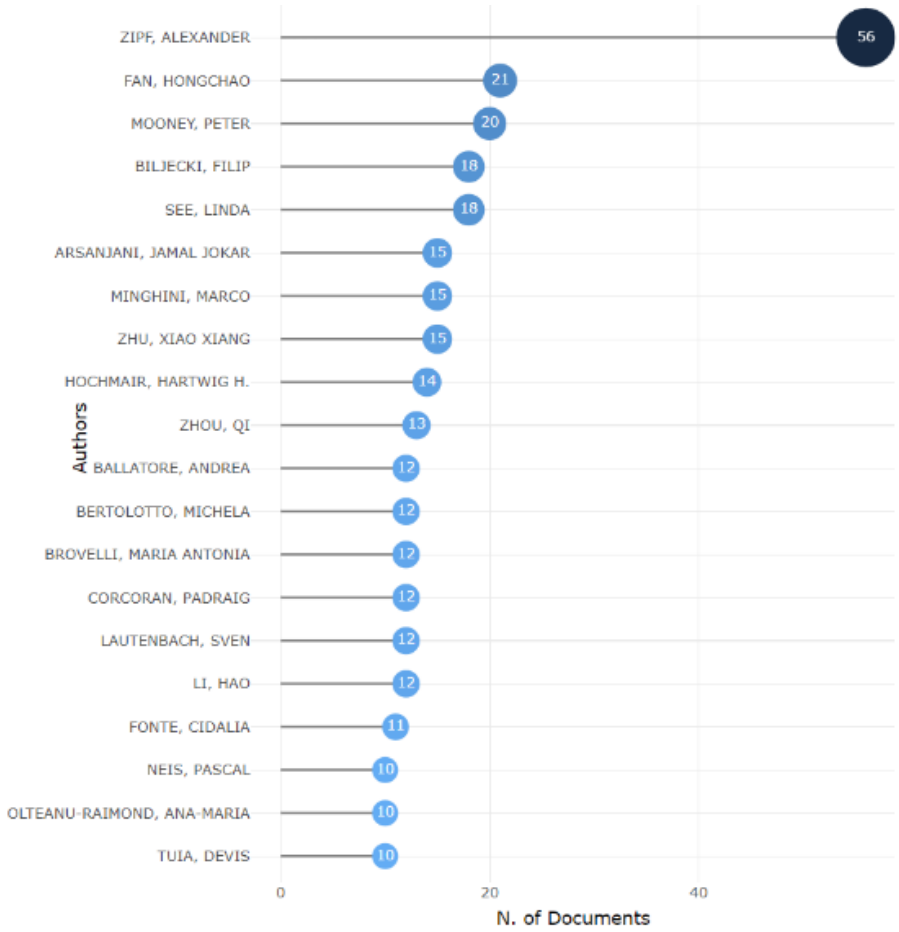
Step 1: load the data.

In Biblioshiny software, in "Import or Load," set "Import raw file(s)," set "Database" to "Web of Science (WoS/WoK)," and set "Author Name format" to "Fullname (if available)," then "Choose a file" to "wos_merged.txt."

Step 2: run analysis

On the left panel, select "Authors" -> "Most Relevant Authors," then set "Options: Number of Authors" to 20, and click "Run the Analysis."


bibliometrix
biblioshiny
Info
Data
Filters
Overview
Sources
Authors
Authors
Most Relevant Authors
Most Local Cited Authors
Authors' Production over Time
Lotka's Law
Authors' Local Impact
Affiliations
Most Relevant Affiliations
Affiliations' Production over Time
Countries
Corresponding Author's Countries
Countries' Scientific Production
Countries' Production over Time
Most Cited Countries
Documents
Clustering
Conceptual Structure
Intellectual Structure
Social Structure
Report
Settings
Most Relevant Authors
Plot
Table
Options:
Number of Authors
20
Frequency measure
N. of Documents



Subfigure 3:

Step 1: check affiliations for each author.

For each author, in the table "lastname.xlsx" generated in the previous step, we checked their affiliations corresponding to the publications in each year. Then, add the affiliations to the tables.

Step 2: make a colored table of all affiliations.

Draw a table, as shown below, add all affiliations, and then use different colors to fill the cells in the table.

Ruprecht Karl University of Heidelberg
Wuhan University (WHU)
Norwegian University of Science and Technology (NTNU)
Maynooth University
Environmental Protection Agency, Dublin
Delft University of Technology (TU Delft)
National University of Singapore (NUS)
International Institute for Applied Systems Analysis (IIASA)
Aalborg University Copenhagen (AAU CPH)
Polytechnic University of Milan (PoliMi)
Joint Research Centre (JRC)

Technical University of Munich (TUM)
German Aerospace Center (DLR)
University of Florida
China University of Geosciences (CUG)
International Research Center of Big Data for Sustainable Development Goals (CBAS)
University College Dublin (UCD)
University of California
University of London
Cardiff University
University of Coimbra / INESC
IGN / LASTIG
University of Zurich
Wageningen University

Combine the subfigures:

- Overlay a table in Subfigure 1 and color its cells according to Subfigure 3.
- Arrange the three figures in a row and adjust their size. Remove repeated information, e.g., author names in the middle figure.
- Adjust the color of the Subfigure 1 and 2.
- Save as one combined figure, as shown below.

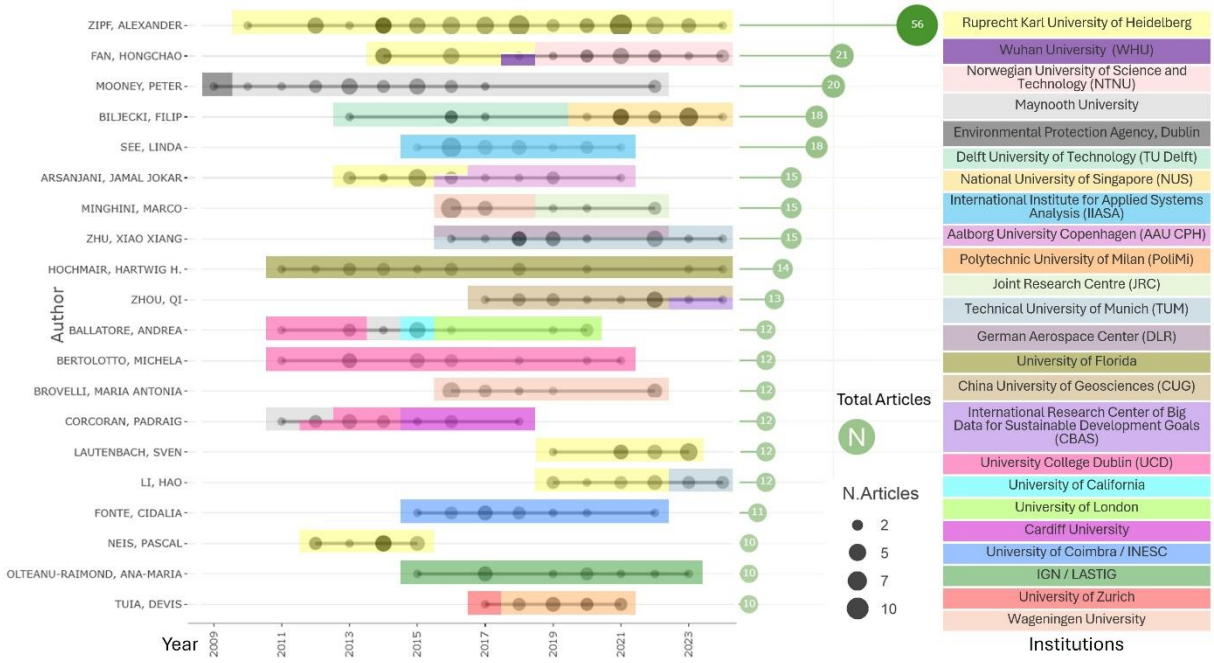


Figure 4

Code directory: .\data and code\Raw data\Figure4

Figure 2 contains two sub-figures, as explained below.

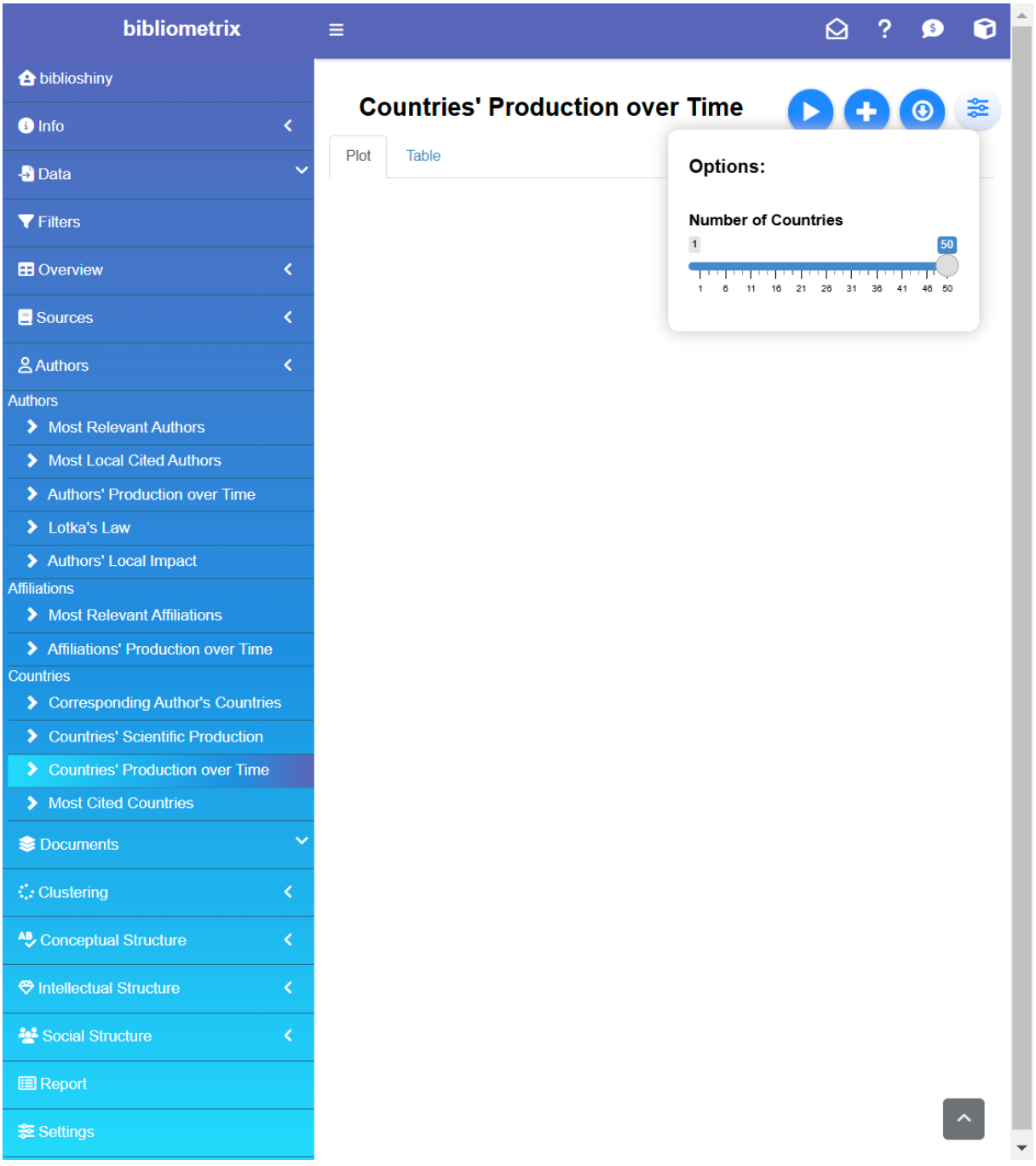
Subfigure 1:

Step 1: load the data.

In Biblioshiny software, in "Import or Load," set "Import raw file(s)," set "Database" to "Web of Science (WoS/WoK)," and "Choose a file" to "wos_merged.txt."

Step 2: run analysis

On the left panel, select "Authors" -> "Countries" -> "Countries' Production over Time," set "Options: Number of Countries" to 50, and click "Run the Analysis."



Step 3: Download the data

On the top panel, select "Table" and download: "Countries_Production_Over_Time.xlsx."

Step 4: rank the countries

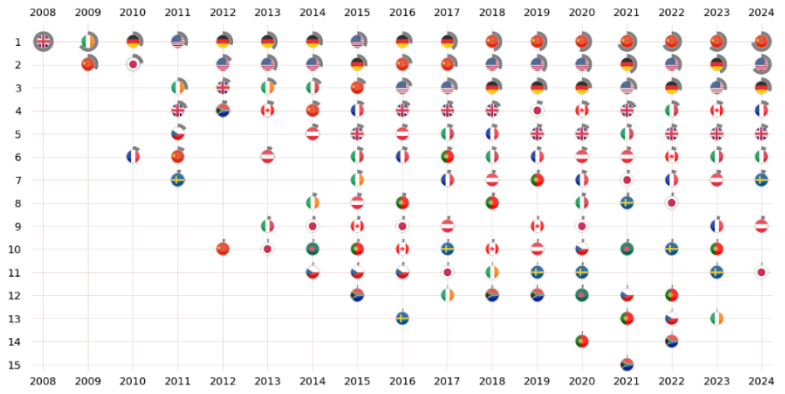
In Excel, rearrange the table. For each country per year, compute its production over the total number of papers in that year. Rank them, select the top 15 countries, and make the table "country-ranking-15.csv".

Step 5: Generate the figure. In jupyter notebook, run "figure4.ipynb".

```
        im, (x, y), frameon=False,
        xycoords="data",
    ))
```

In [8]:
```
def add_text(ax, value, max_value, y):
    trillions = round(value / 1e12, 1)
    ratio_to_max = round(100 * value / max_value, 1)

    text = "{}\n${:,}T ({}%)".format(
        row.country_name,
        trillions,
        ratio_to_max
    )

    ax.annotate(
        text, (1.03, y),
        fontsize=10,
        linespacing=1.7,
        va="center",
        xycoords=("axes fraction", "data")
    )
```

In [14]:
```
background_color = "#FFFFFF"
text_color = "#080520"
grid_color = "#E4C9C9"

set_seaborn_style(font_family, background_color, grid_color, text_color)

years = ["2008", "2009", "2010", "2011", "2012", "2013", "2014", "2015", "2016", "2017", "2018", "2019"

df = pd.read_csv("country-ranking-15.csv", index_col=None)
df, rank_columns = create_rankings(df, years)


fig, ax = plt.subplots(nrows=1, ncols=1, figsize=(25, 0.8*len(df)))
format_ticks(ax, years)

for i, row in df.iterrows():

    for j, rc in enumerate(rank_columns):

        add_pie_and_flag(
            ax, j, -row[rc],
            name=row.country_name,
            ratio=row[years[j]]
        )

plt.show()
# plt.savefig('country_rank.png', transparent=True, dpi=300)
```



In [256]:
```
df[["country_name"] + rank_columns]
```

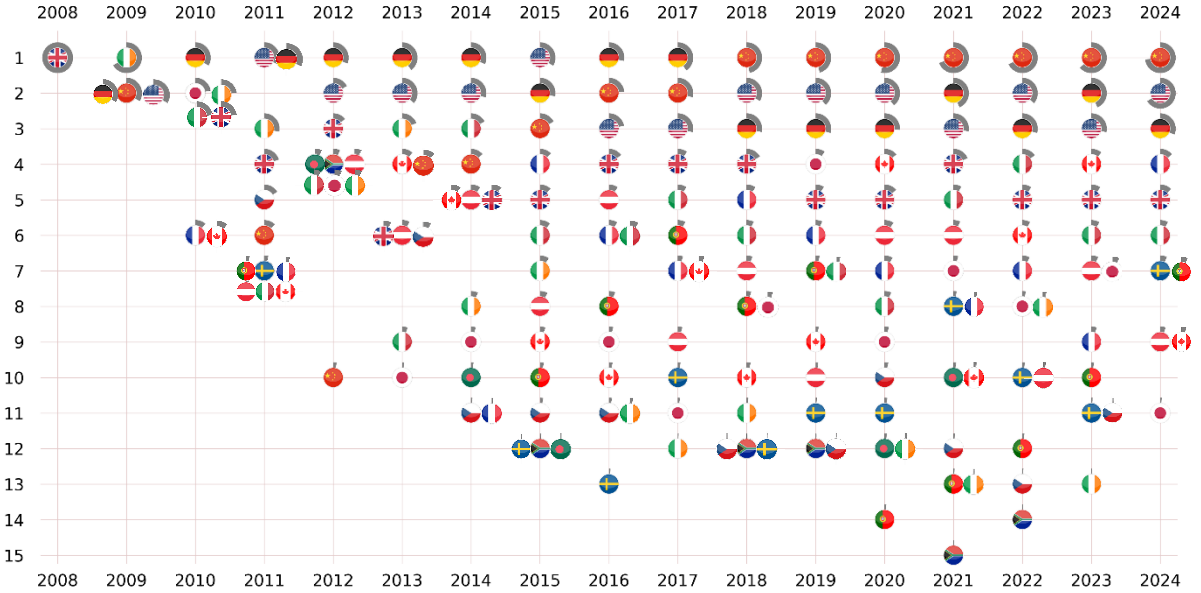
Out[256]:

	country_name	rank_0	rank_1	rank_2	rank_3	rank_4	rank_5	rank_6	rank_7	rank_8	rank_9	rank_10	rank_11	rank
0	United Kingdom	1.0	NaN	2.0	4.0	3.0	6.0	5.0	5.0	4.0	4.0	4.0	5.0	
1	Ireland	NaN	1.0	2.0	3.0	4.0	3.0	8.0	7.0	11.0	12.0	11.0	12.0	1
2	Germany	NaN	2.0	1.0	1.0	1.0	1.0	1.0	2.0	1.0	1.0	3.0	3.0	
3	USA	NaN	2.0	NaN	1.0	2.0	2.0	2.0	1.0	3.0	3.0	2.0	2.0	
4	China	NaN	2.0	NaN	6.0	10.0	4.0	4.0	3.0	2.0	2.0	1.0	1.0	
5	Italy	NaN	NaN	2.0	7.0	4.0	9.0	3.0	6.0	6.0	5.0	6.0	7.0	
6	Japan	NaN	NaN	2.0	NaN	4.0	10.0	9.0	15.0	9.0	11.0	8.0	4.0	
7	Canada	NaN	NaN	6.0	7.0	NaN	4.0	5.0	9.0	10.0	7.0	10.0	9.0	
8	France	NaN	NaN	6.0	7.0	NaN	11.0	11.0	4.0	6.0	7.0	5.0	6.0	
9	Czech Republic	NaN	NaN	NaN	5.0	NaN	6.0	11.0	11.0	11.0	13.0	12.0	12.0	1

In [ ]:

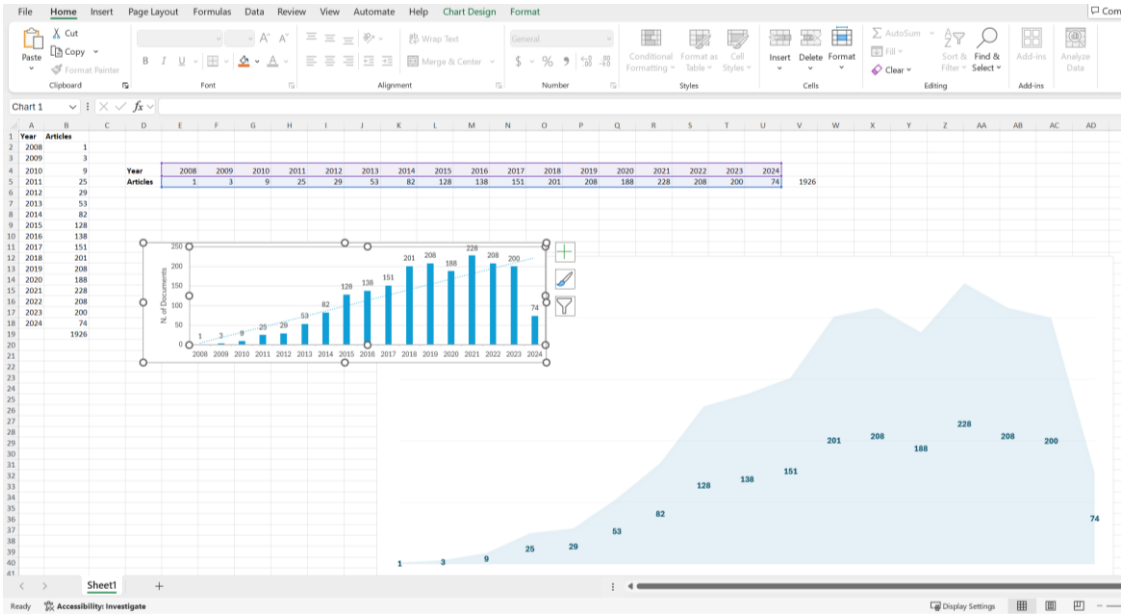
Step 6: Adjust the figure

Countries with the same outputs overlay each other. Adjust the table for all countries and combine the generated figures with the figure below.



Subfigure 2:

Adjust Figure 2 in “Annual_Production-2008-2024.xlsx”.



Combine the subfigures:

Finally, add Subfigure 2 to the background of Subfigure 1. Adjust the location of numbers in Subfigure 2. Add labels on the x- and y-axis.

Figure 5

Data directory: .\data and code\Raw data\Figure5

Step 1: load the data.

In Vosviewer software, set "choose the type of data" to "Create a map based on bibliographic data," "Choose data source" to "Read data from bibliographic database files," select "wos_merge_switchC1C3.txt" to load the data.

Step 2: set the methods and parameters:

- "type of analysis" → Co-authorship
- "Unit of analysis" → Organizations
- "counting method" → Full counting
- "choose thresholds" → Minimum number of documents of an organization: 10

Step 3: Create the map (Explained in Section 4.3.2)

Deselect: Arizona State University, Helmholtz Association, Purdue University System, State University System of Florida, Swiss federal institutes of technology domain, University of California System, University of Colorado System, University of Minnesota System, University System of Ohio, and then click "Finish".

- Set parameters according to the screenshot below.

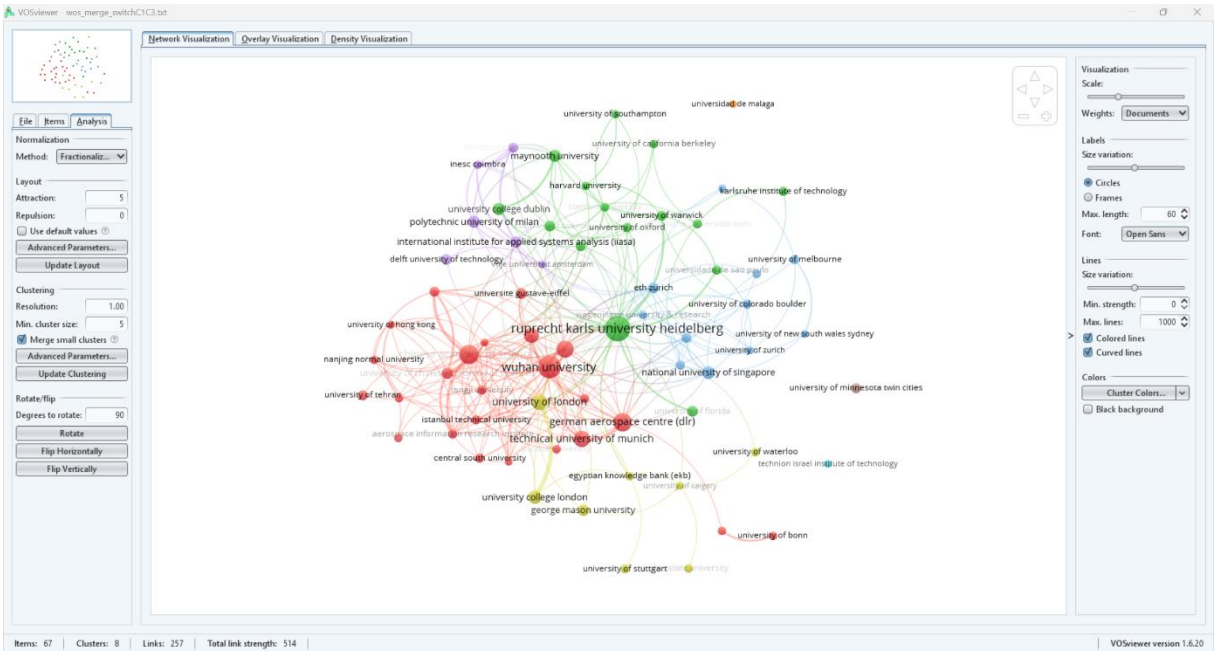


Figure 6

Step 1: load the data.

In Vosviewer software, set "choose type of data" to "Create a map based on bibliographic data", "Choose data source" to "Read data from bibliographic database files", select “wos_merged.txt” to load the data.

Step 2: set the methods and parameters:

- “Type of analysis” → co-authorship
- “Unit of analysis” → Authors
- “Counting method” →Full counting
- “Choose thresholds” → minimum number of documents of an author: 5

Step 3: create map, click “Finish”.

- Set parameters according to the screenshot below.

Figure 7

Step 1: load the data.

In Biblioshiny software, in "Import or Load," set "Import raw file(s)," set "Database" to "Web of Science (WoS/WoK)," and "Choose a file" to "wos_merged.txt."

Step 2: run analysis

On the left panel, select "Authors" -> "Countries" -> "Corresponding Author's Countries," set "Options: Number of Countries" to 20, and click "Run the Analysis."

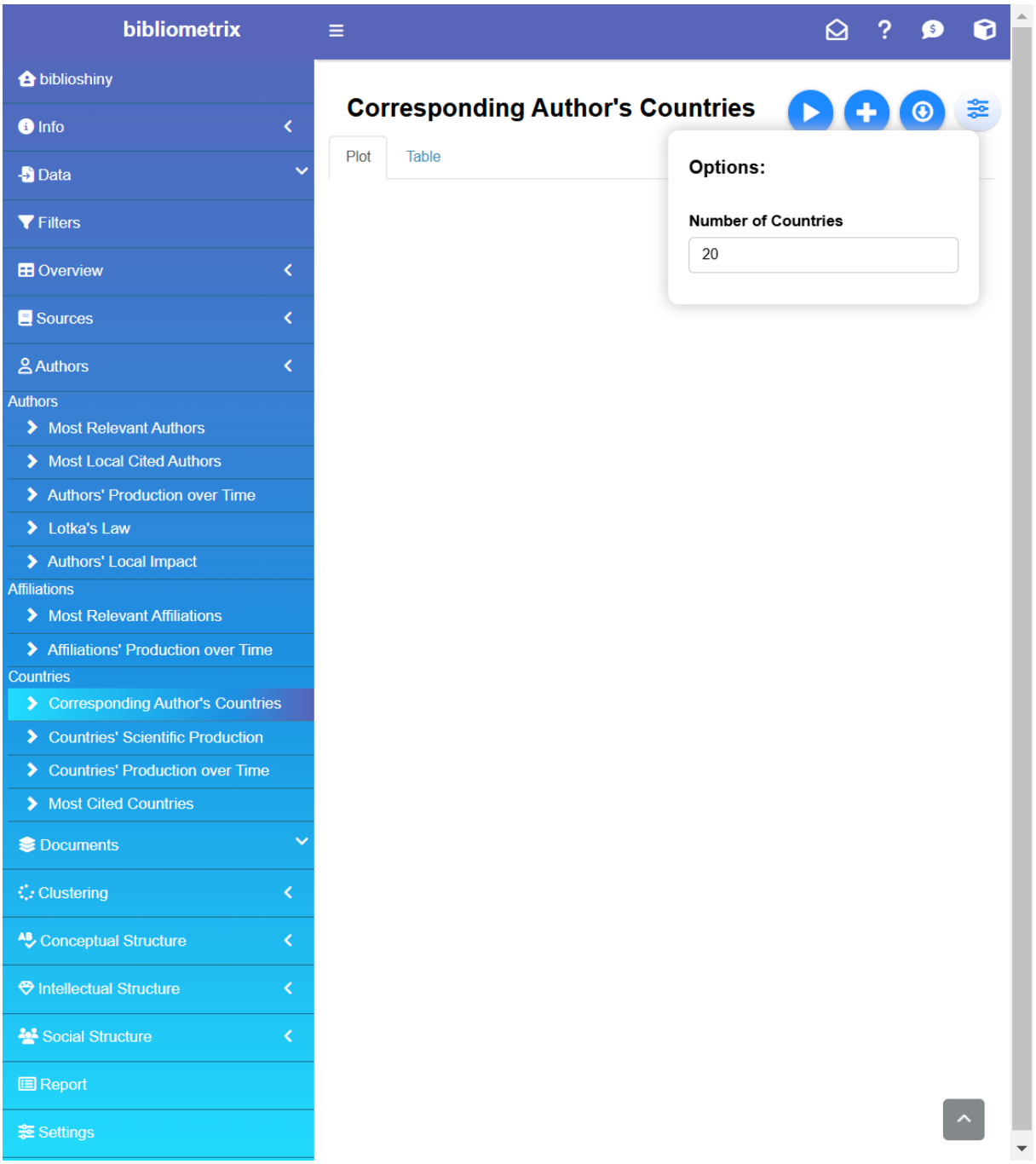

bibliometrix
biblioshiny
Info
Data
Filters
Overview
Sources
Authors
Authors
Most Relevant Authors
Most Local Cited Authors
Authors' Production over Time
Lotka's Law
Authors' Local Impact
Affiliations
Most Relevant Affiliations
Affiliations' Production over Time
Countries
Corresponding Author's Countries
Countries' Scientific Production
Countries' Production over Time
Most Cited Countries
Documents
Clustering
Conceptual Structure
Intellectual Structure
Social Structure
Report
Settings
Corresponding Author's Countries
Plot
Table
Options:
Number of Countries
20

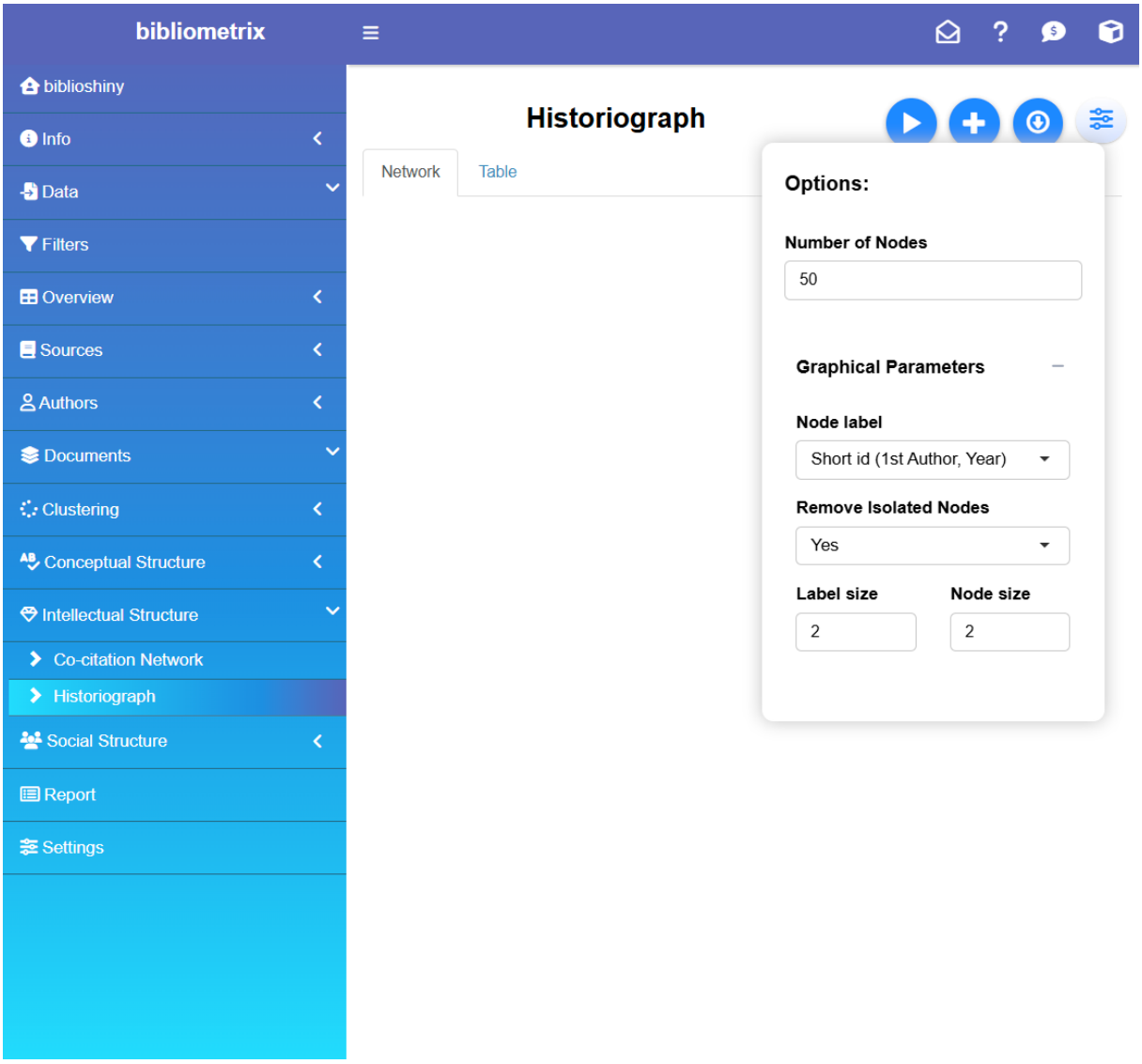
Figure 8

Step 1: load the data.

In Biblioshiny software, in "Import or Load," set "Import raw file(s)," set "Database" to "Web of Science (WoS/WoK)," and "Choose a file" to "wos_merged.txt."

Step 2: run analysis

On the left panel, select "Intellectual Structure" -> "Historiography," then set "Options: Number of Nodes" to 50, and click "Run the Analysis."



Step 3: Label the key papers

The five key research interests' starting paper(s) are manually highlighted in the figure.

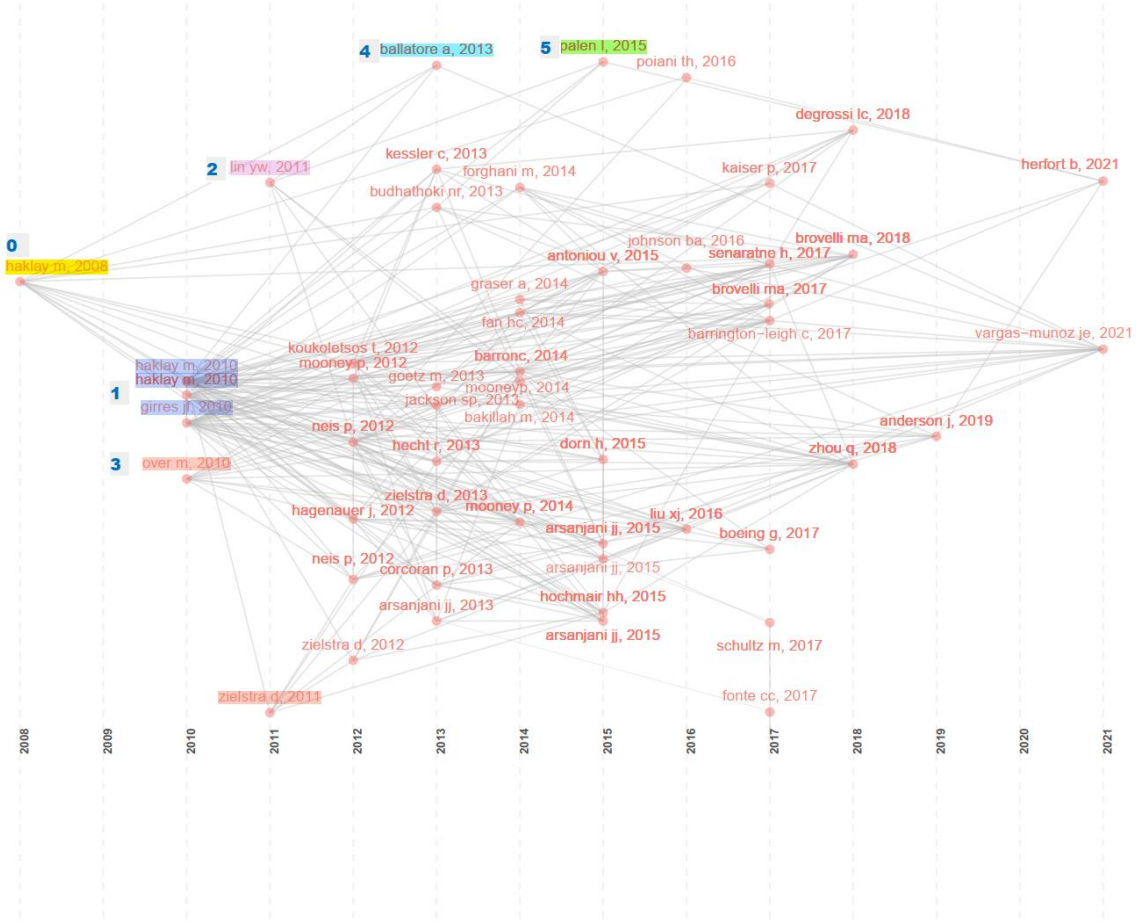


Figure 9

Step 1: load the data.

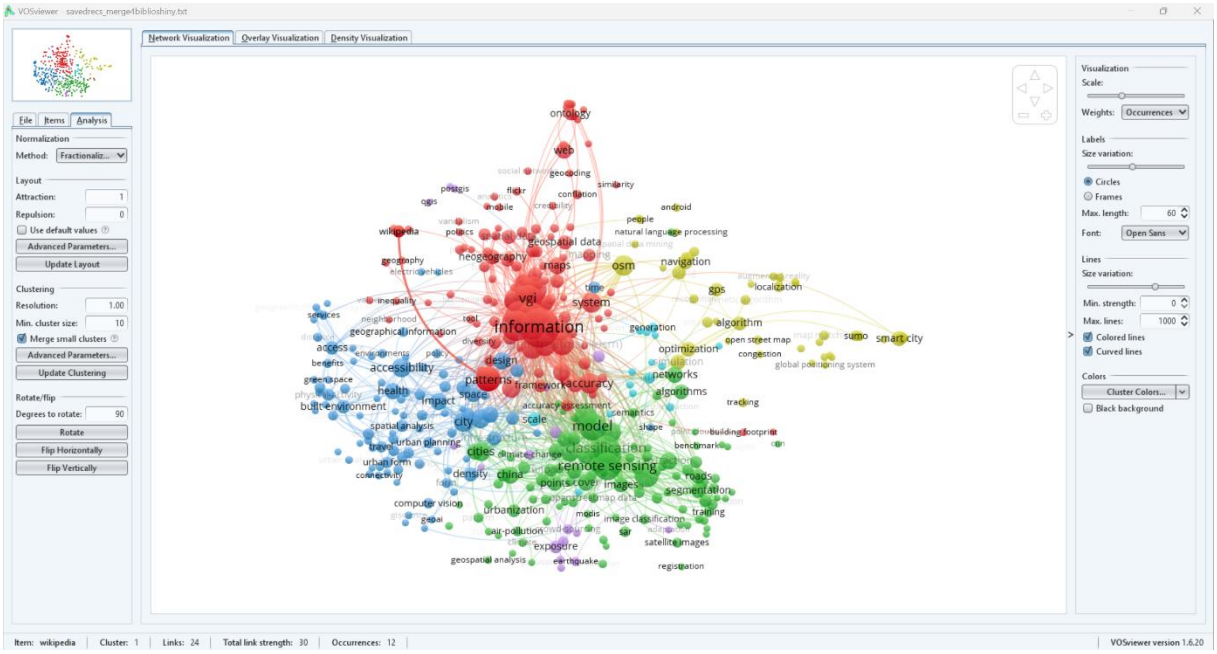
In Vosviewer software, set "choose type of data" to "Create a map based on bibliographic data", "Choose data source" to "Read data from bibliographic database files", select "wos_merged.txt" to load the data.

Step 2: set the methods and parameters:

- “Type of analysis” → Co-occurrence
- “Unit of analysis” → All Keywords
- “Counting method” →Full counting
- “Choose thresholds” → Minimum number of occurrences of a keyword: 5

Step 3: create map, deselect the top 2 keywords, “openstreetmap" and “volunteered geographic information", then click “Finish”.

- Set parameters according to the screenshot below.



- For Figure 9 (b), change to “Overlay Visualization”.

Figure 10

Figure 10 contains two WoS sub-figures and two SotM subfigures.

WoS figures:

Data directory: .\data and code\Raw data\Figure10\wos

Step 1: load the data.

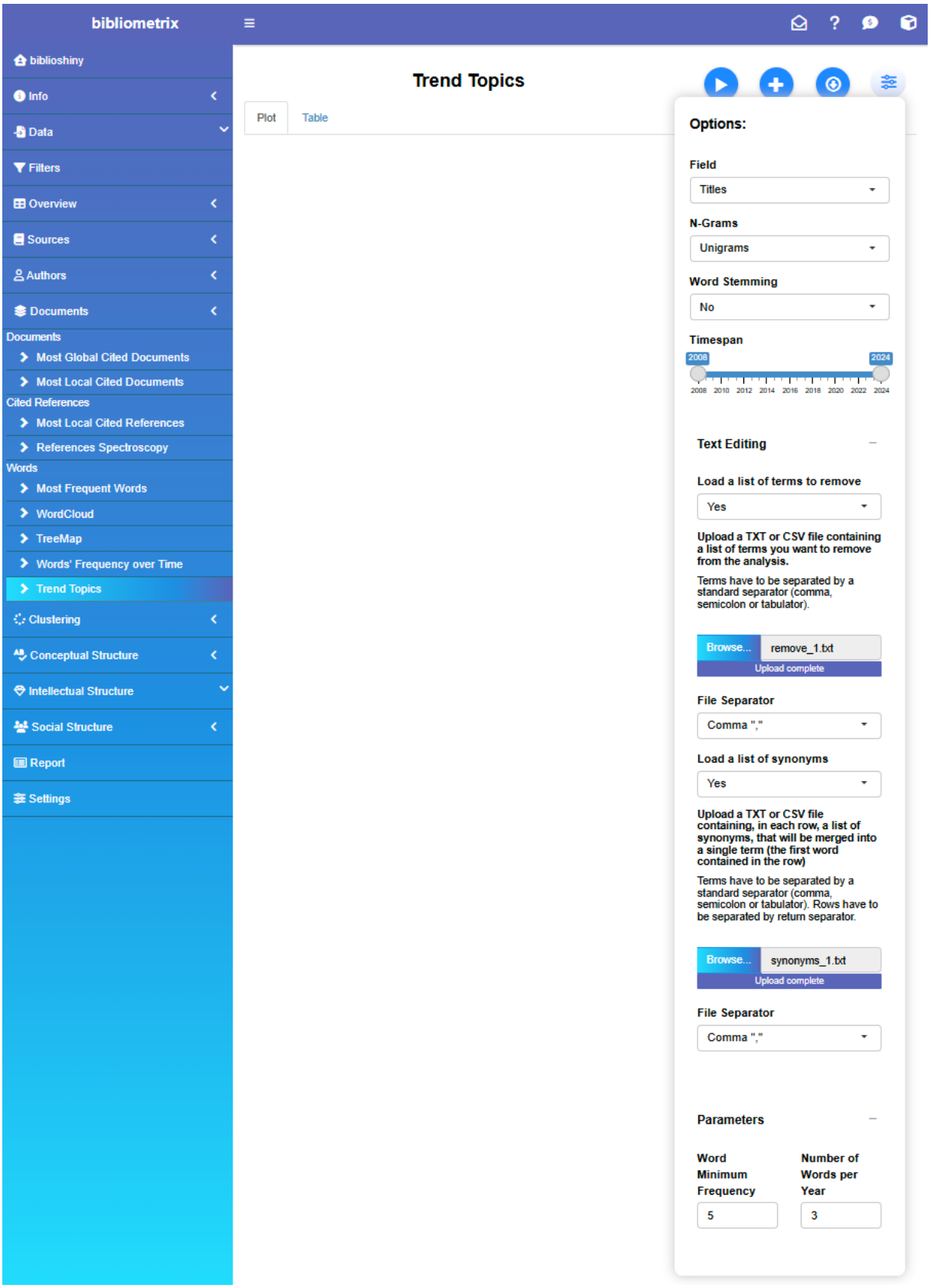
In Biblioshiny software, in "Import or Load," set "Import raw file(s)," set "Database" to "Web of Science (WoS/WoK)," and "Choose a file" to "wos_merged.txt."

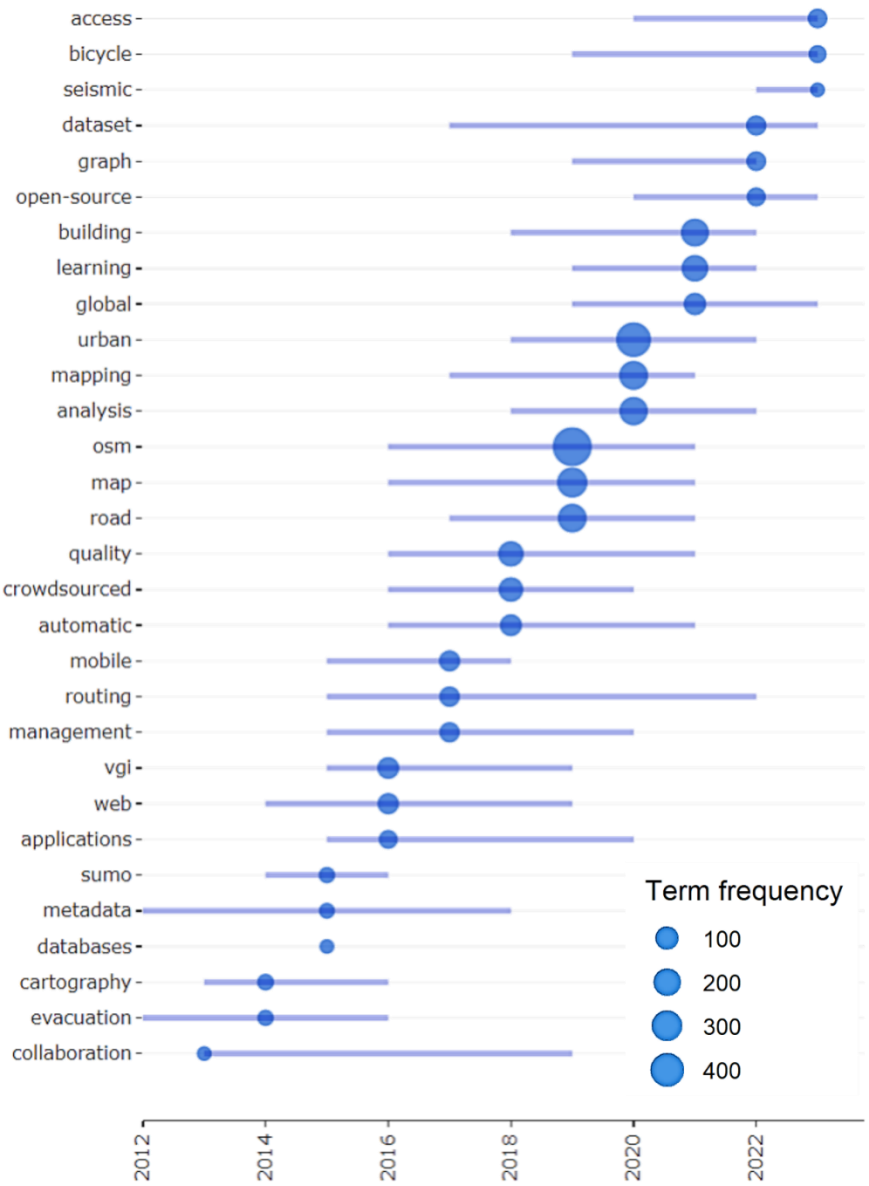
Step 2: set parameters

- For Unigrams:
  - On the left panel, select "Documents" -> "Words" -> "Trend Topics," and set "Options: Filed" to "Titles" and "N-Gram" to "Unigrams."
  - In "Text Editing," at "load a list of terms to remove," load "remove_1.txt"; at "load a list of synonyms," load "synonyms_1.txt".
  - In "Parameters," set "Word Minimum Frequency" to 5 and "Number of Words per Year" to 3.
- For Bigrams
  - On the left panel, select "Documents" -> "Words" -> "Trend Topics," and set "Options: Filed" to "Titles" and "N-Gram" to "Bigrams."
  - In "Text Editing," at "load a list of terms to remove," load "remove_bi.txt"; at "load a list of synonyms," load "synonyms_bi.txt."
  - In "Parameters," set "Word Minimum Frequency" to 3 and "Number of Words per Year" to 3.

Step 3: run analysis, click "Run the Analysis."

See the screenshot below for getting the Unigram of WoS. Follow the steps to get other figures.





SotM figures:

Data directory: .\data and code\Raw data\Figure10\sotm\

Step 0: prepare data for analysis:

- Open "sotm_merged.xlsx," and copy-paste the columns "TI," "AU," and "PY" to the corresponding columns in the template "temp_Bibliometrix.xlsx."
- Name it “SOTM_Bibliometrix.xlsx”.

Step 1: load the data.

In Biblioshiny software, in “Import or Load”, set “Load bibliometrix file(s)”, “Choose a file” to “SOTM_Bibliometrix.xlsx”.

Step 2: set parameters

- For Unigrams:
    - On the left panel, select "Documents" -> "Words" -> "Trend Topics,"” and set "Options: Filed" to "Titles,"” "and N-Gram" to "Unigrams."”

    - In "Text Editing," at "load a list of terms to remove," load "remove_1.txt"; at "load a list of synonyms," load "synonyms_1.txt".
    - In "Parameters", set "Word Minimum Frequency" to 5 and "Number of Words per Year" to 3.
- For Bigrams
    - On the left panel, select "Documents" -> "Words" -> "Trend Topics," and set "Options: Filed" to "Titles," and "N-Gram" to "Bigrams."
    - In "Text Editing," at "load a list of terms to remove," load "remove_bi.txt"; at "load a list of synonyms," load "synonyms_bi.txt."
    - In "Parameters", set "Word Minimum Frequency" to 2 and "Number of Words per Year" to 3.

Step 3: run analysis, click “Run the Analysis”.

Combine the subfigures:

Finally, combine the four subfigures and adjust the colors.