**User guider of Searching Web of Science 1.0**

NOTE:

* This software may fail to act due to the casual or regular revision of WoS. Please indicate the error on <https://github.com/zhouyp97/PSC-PC-Top/discussions>. In this case, we will revise the software.
* The software will access the internet with the help of Microsoft Internet Explorer. If encountering an authority problem of Internet Explorer, such as script errors, please add related contents to the registry table of your operating system according to **2. Add contents to registry table**.
* This software searches WoS by simulating the manual searching operation without any special network techniques. Please permit its running if encountering any cautions from security software or other software.
* If the user needs to manually sign to the WoS, please do it through the WoS website of the software after running it. After loading, we should go to the initial page.
* We should set the language of the WoS website as “English”.
* The whole searching and recording process may need one to several weeks depending on the conditions of the user’s computer and network.
* The searching task should avoid the maintenance period of WoS.
* The various searching tasks should act independently. And therefore, the software should be restarted before a new searching task starts.
* The functions in “Manual search” should not be used by normal user which are not introduced in this guider.
* The data treating function (except the search function) can only treat the data not more than 10 publishers, and therefore, the data file for data treating should be divided into several files if it contains the data for more than 10 publishers.
* When the software runs in the silence mode by clicking a button in light brown color, user should avoid any manual operation of the computer by using a mouse, a keyboard and other manual input devices. The software window should be kept in the frontest position. The mouse cursor should be moved to the left-bottom blank space in the software window. In addition, user should close all other software that may pop up a new window. Any energy conservation mode including screen saver should be disabled.
* Normally, user can choose freely whether the software runs in the silence mode or not. Sometimes, if the normal mode exceeds the 6000 items limitation of WoS when searching an article group, user have to search the information about the article group in the silence mode.

Yangping Zhou ([zhouyp@mail.tsinghua.edu.cn](mailto:zhouyp@mail.tsinghua.edu.cn))

INET, Tsinghua University, 2021

**Contents**

[1. Operating system and hardware requirements 3](#_Toc58485150)

[2. Add contents to registry table 3](#_Toc58485151)

[3. List of journal-publisher 3](#_Toc58485152)

[3.1 Load scopus source list 3](#_Toc58485153)

[3.2 Load ESI master journal list 4](#_Toc58485154)

[3.3 Match journal with publisher 4](#_Toc58485155)

[3.4 Export matching results 4](#_Toc58485156)

[3.5 Load matching results 4](#_Toc58485157)

[NOTE: 4](#_Toc58485158)

[4. Year, type, database, number of journals in a group and running mode 5](#_Toc58485159)

[NOTE: 5](#_Toc58485160)

[5. Publisher citation 5](#_Toc58485161)

[5.1 Search 5](#_Toc58485162)

[5.2 Search (special) 6](#_Toc58485163)

[NOTE: 6](#_Toc58485164)

[6. Journal citation 7](#_Toc58485165)

[6.1 Search journal citation (M1) 7](#_Toc58485166)

[6.2 Search journal citation (M2) 7](#_Toc58485167)

[6.3 Combine 8](#_Toc58485168)

[6.4 Treat duplicate name 8](#_Toc58485169)

[6.5 Match pub 8](#_Toc58485170)

[NOTE: 8](#_Toc58485171)

[7. Category 9](#_Toc58485172)

[7.1 Search 9](#_Toc58485173)

[7.2 Search annual articles 9](#_Toc58485174)

[7.3 Pretreat 9](#_Toc58485175)

[7.4 Treat 10](#_Toc58485176)

[7.5 Load 10](#_Toc58485177)

[7.6 Combine 11](#_Toc58485178)

[NOTE: 11](#_Toc58485179)

[8. Organization 12](#_Toc58485180)

[8.1 Search (Simple) 12](#_Toc58485181)

[8.2 Get organization list 12](#_Toc58485182)

[8.3 Search 12](#_Toc58485183)

[8.4 Search (Special) 13](#_Toc58485184)

[8.5 Pretreat 13](#_Toc58485185)

[8.6 Treat 13](#_Toc58485186)

[8.7 Combine 13](#_Toc58485187)

[8.8 Match org 13](#_Toc58485188)

[NOTE: 14](#_Toc58485189)

# 1. Operating system and hardware requirements

Microsoft Windows 7 or 10

Microsoft .NET Framework 4.5

Microsoft Internet Explorer: 11.0

RAM: 4.0GB

Hard disk: 2.0GB

Monitor: 1920×1200

# 2. Add contents to registry table

* 32-bit operating system:

HKEY\_LOCAL\_MACHINE\SOFTWARE\Microsoft\Internet Explorer\MAIN\FeatureControl\FEATURE\_BROWSER\_EMULATION

* 64-bit operating system:

\HKEY\_LOCAL\_MACHINE\SOFTWARE\WOW6432Node\Microsoft\Internet Explorer\Main\FeatureControl\FEATURE\_BROWSER\_EMULATION

Name: searchWoS-publisher.exe

Type: REG\_DWORD

Value: 2af9

# 3. List of journal-publisher

## 3.1 Load Scopus source list

Load the Scopus source list file. The original list file was download from <https://www.elsevier.com/zh-cn/solutions/scopus>. The software uses the version download on Feb. 16, 2019. Now, only the newest version is available on Scopus. The new version of the file can also be used and it may introduce some ignorable differences.

## 3.2 Load ESI master journal list

Load the ESI master journal list file. The excel file download from http://help.incites.clarivate.com/incitesLiveESI/ESIGroup/overviewESI/esiJournalsList.html can be directly used. At present, the file downloaded on Feb. 15, 2019 (http://help.incites.clarivate.com/incitesLiveESI/10678-TRS/version/9) is used by this software. The new version of the file can also be used and it may introduce some ignorable differences.

## 3.3 Match journal with publisher

Match the Scopus source list file with the ESI master journal list file. The process may take around 1-2 minutes depending on the speed of your computer.

## 3.4 Export matching results

Export the matching results to an indicated excel file.

## 3.5 Load matching results

Load the matching results file.

## NOTE:

* For getting the matching results, the action sequence should be **3.1**, **3.2**, **3.3** and **3.4**.
* The treated Scopus source list file, the ESI master journal list file and the matching results file in .xlsx format are provided.
* User can load the exported file of matching results to start different search tasks after restarting the software. In this case, steps **3.1**, **3.2**, **3.3** and **3.4** can be skipped.

# 4. Year, type, database, number of journals in a group and running mode

* User can set a year for searching, for each year (2000-present year), 2017 (presently used).
* User can check the “High” checkbox to search for high citation articles.
* User can check the “Open” to search open access articles.
* User can check or uncheck related databases, “SCIE”, “SSCI” and “AHCI”.
* User can input the textbox of “Cited articles” and click the “Set” button to set the number of articles in an article group.
* User can check the “Silence mode” checkbox to enable or disable the silence mode.

## NOTE:

* If neither the “High” checkbox nor the “Open” checkbox is checked, all articles are searched.
* When the Silence mode is disabled or enabled, the color of some buttons will be changed to indicate whether the related functions will run in silence mode or normal mode.
* Two functions, “Get journal list” and “Get organization list” usually run in silence mode not matter whether the “Silence mode” checkbox is checked.

# 5. Publisher citation

*It is recommended that the search results by the search function in this section should be stored into one file for related following automation by this software.*

## 5.1 Search

Search citation of articles of a publisher in a year by indicating the No. of the start publisher and the No. of the start article group of the publisher. The first textbox before the button is the No. of the start publisher. For example, 1 indicates Elsevier and 2 indicates Springer Nature. It should be bigger than 0. The second textbox is the No. of the start group. For example, 0 indicates the first article group of a publisher.

## 5.2 Search (special)

Search citation of publisher by indicating the No. of the start publisher and the No. of the start group of the publisher. The first textbox and the second text box are the same as those in **5.1**. Different from **5.1**, the software will only search the citation of the last 10000 articles of the top 100 journals and the top 101-200 journals (two special subsets) of a publisher if the second text is set as “0”. Accordingly, the software will only search the citation of the last 10000 articles of the top 101-200 journals and the top 201-300 journals (two special subsets) of a publisher if the second text is set as “1”. In this way, the journal number of a publisher should be more than 100 if the second text is set as “0”.

## NOTE:

* Before starting **5**, please restart the software, load the matching results in **3** and confirm and/or change the year, type and database in **4**.
* After clicking the **5.1** or **5.2** button, user should indicate an excel file to save the searching results.
* The search results gotten by **5.1** and **5.2** will be saved in the different sheets of the indicated excel file by the reverse order of publisher, group and subset as following:



* The Group No. of the sheets obtained by **5.2** will begin with “S” to distinguish those sheets obtained by **5.1**.
* If the search process is interrupted or stopped for a long time (usually more than 1 minute) due to network problems, etc., user should firstly close the software. Then, user should delete the unfinished results for the last article group of related publishers in the excel file. For example, if the search process is interrupted when searching the second article group of the third publishers, user should delete all sheets in the excel files started with “3-1”. Finally, user can restart the software and the searching process by doing **3.5**, **4**, and **5**. When doing **5**, user should set the start publisher as “3” and start the group as “1” and choose the same excel file.

# 6. Journal citation

*It is recommended that the search results by the search function in this section should be stored into one file for related following automation by this software.*

## 6.1 Search journal citation (M1)

Search the citation information of the journals with Method 1. The first textbox before the button is the No. of the start journal. The searched journals are sequenced according to the numbers of their articles in the indicated year. The journal that has the most articles in the year will be the first journal. The journals contained more than 10000 articles in a year, whose citation reports are not available for WoS. And therefore, the journal will be treated as two journals each of which published 10000 articles in chronological or reverse order of a year. The second textbox is the article number of a sheet of the indicated excel file to save the search results. 50 is recommended.

## 6.2 Search journal citation (M2, usually is faster than M1)

Search the citation information of the journals with Method 2 in which a journal list in the indicated year will be used. The first textbox before the button is the number of published articles of the start journal in the indicated year. The second textbox is the alphabetical order (start from 0 by default) of the journal in the journal list with the same number of published articles. The third textbox is the article number of a sheet of the indicated excel file to save the search results. 50 is recommended. Method 2 does not search the journal having more than 10000 articles in a year. User should use Method 1 to search these journals.

For the journal citation in 2016, the new search should start by setting the first textbox as a big number which is bigger than the number of the articles in any journals, such as “10000”, and the second textbox should be “0”. In this way, the first journal is the first journal which has articles less than 10000 in the journal list.

To use this function, user should firstly indicate an excel file including a journal list from the databases of SCIE, SSCI and AHCI of WoS. The excel file can be obtained by clicking the “Get journal list” button before clicking the “Search” button here.

With Method 2, WoS may fail to search some journals whose title begins with “NEAR” or “OR” or consists of some special words such “AND NEAR”, which are listed as followings:

JOURNAL OF BALKAN AND NEAR EASTERN STUDIES, OR SPECTRUM, NEAR EASTERN ARCHAEOLOGY, and NEAR SURFACE GEOPHYSICS.

User has to search the citation of these journals with Method 1 or a manual way (recommended) by using their ISSNs, as 1944-8953, 0171-6468, 1094-2076, and 1569-4445. In this way, user have to manually delete these four journals from the journal list if using Method 2.

## 6.3 Combine

Combine citation contents in different sheets of the excel files consisting of the search results.

## 6.4 Treat duplicate name

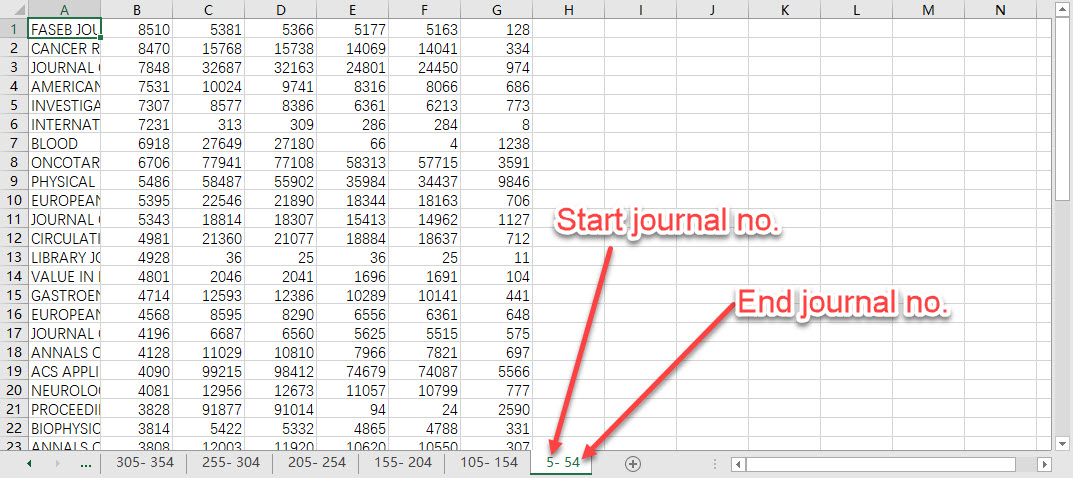
Remove the duplicate journal names in the excel file.

## 6.5 Match pub

Match the journal with the publishers.

## NOTE:

* Before starting **6**, please restart the software, load the matching results in **3** and confirm and/or change the year, type and database in **4**.
* To finish the process of the journal citation, user should finish **6.1 and/or 6.2, 6.3, 6.4** and 6.**5** sequentially.
* The search results gotten by **6.1 or 6.2** will be saved in the different sheets of the indicated excel file by the reverse order of journal No. as following:



* The journal No. of the sheets obtained by **6.2** will begin with M2 to distinguish those sheets obtained by **6.1**.
* If the search process is interrupted or stopped for a long time (usually more than 1 minute) due to network problems, etc., user should firstly close the software and all background processes name as Microsoft Excel. Then, user should delete the unfinished results for related journals in the excel file. For example, if the search process is interrupted when searching the 205-254 journals, user should delete the sheet named “205-254” in the excel files. Finally, user can restart the software and the searching process by doing **3.5**, **4**, and **6**. When doing **6.1**, user should set the start journal as “205” and choose the same excel file. When doing **6.2**, user should set article number of the related journal and its alphabetical order (start from 0).
* Before doing **6.3**, user needs to check the excel file to delete the redundant sheet such as “sheet1”, etc.
* The task for searching journal citation is the most time-consuming work. In addition, the task using Method 1 will become slower when searching the journals with less articles. It is recommended that user start the search with Method 1 and then shift to Method 2 when the searching becomes slow.

# 7. Category

*It is recommended that the search results by the search function in this section should be stored into different files according to the year for related following automation by this software. For example, if we searched WoSC profiles in 2017, 2018 and 2019, we should store the related information in three files for 2017, 2018 and 2019 respectively.*

## 7.1 Search

Search the category information of a publisher in a year (2000-present year). The first textbox before the button is the No. of the start publisher. The second textbox is the No. of the beginning article group. User should indicate an excel file saving the search results.

## 7.2 Search annual articles

Search the category information of all articles in a year (2000-present year). User should indicate an excel file saving the search results.

## 7.3 Pretreat

Pretreat the indicated excel file saving search results of the category information in a year. If the excel file includes the information of the subset with 10000 articles (usually from **5.1/5.2**), the obtained excel files will include two files: one is for the group articles named as the name of the indicated file (excluding the suffix) + “-hhmmss” + the suffix and the other is for the cited subset articles named as the name of the indicated file (excluding the suffix) + “-hhmmss-sub” + the suffix. If it only includes the information of the article group (usually from **7.1**), the obtained excel files is only for group articles named the indicated file (excluding the suffix) + “-hhmmss” + the suffix.

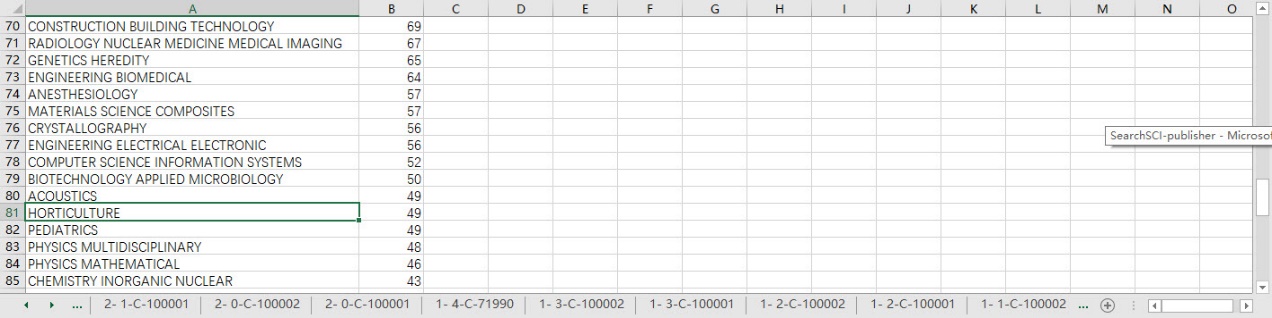
## 7.4 Treat

Treat the excel file saving search results of the category information of a publisher in a year. There are two kinds of category files:

* The category file, containing the numbers of the cited subset article in each category. This original excel file comes from **5.1/5.2**. User can obtain it through **7.3**.
* The category file, containing the number of the group article in each category. This original excel file may come from **5.1/5.2** or **7.1**. User can obtain it through **7.3**.

The contents of the above two types of files before treating should be as following which only includes the sheets related to category information:

*File for the subset article come from* ***5.1/5.2***



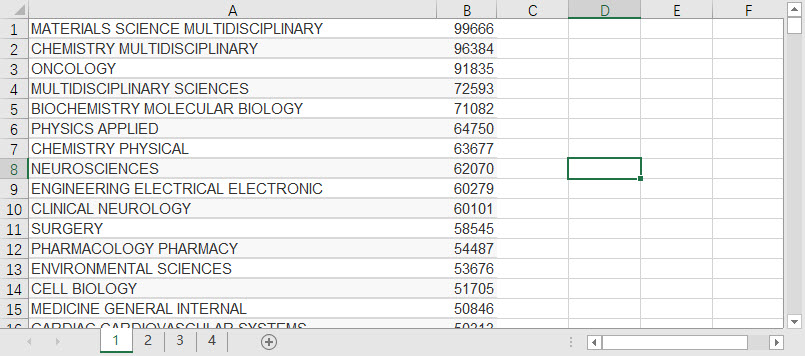
*File for the group article come from* ***5.1/5.2*** *or* ***7.1***



## 7.5 Load

Load the category files obtained by **7.2 and 7.4**. The value of the textbox before the button indicated the type of the category file:

* 0: The summary category file, containing the article numbers of each category in different years. These files are obtained by **7.2**. The contents of each year should be in a sheet with a year sequence. If it is not, user should manually sequence it. For example, if we treat the data in 2017-2020, the first one should be that of 2017 and the last one should be that of 2020. The contents of the file should be:



* 1: The category file for the cited articles. It is obtained after the treatment by **7.4** originally from **5.1/5.2**, containing the numbers of the cited article belonging to each category in an indicated year, such as 2017.
* 2017, 2018, 2019 and 2020 (if the year is set as 2017 in **4**): The category files for the articles for the related publishers in related years. They are obtained after the treatment by **7.4**. The category file in the same year as the cited articles is obtained by **5.1/5.2**. Other category files are obtained by **7.1**. These files contain the numbers of the articles belonging to each category which are published by the related publishers in a year, such as 2017, 2018, 2019 and 2020 if the cited articles were published in 2017.

## 7.6 Combine

Combine the contents obtained by **7.5** to a new indicated excel file for PSCR calculation.

## NOTE:

* Before starting **7**, please restart the software, load the matching results in **3** and confirm and/or change the year, type and database in **4**.
* If the search process (**7.1**) is interrupted or stopped for a long time (usually more than 1 minute) due to network problems, etc., user should firstly close the software and all background processes name as “Microsoft Excel” and “Excel”. Then, user should delete the unfinished results for the category information of the related publisher in the excel file. For example, if the search process is interrupted when searching the third publishers, user should delete all sheets in the excel files started with “3-”. Finally, user can restart the software and the searching process by doing **3.5** and **7.1**. When doing **7.1**, user should set the start publisher as “3” and start the group as “0” and choose the same excel file.
* Steps, **7.5** and **7.6**, should be done sequentially. When donging **7.5**, all files should be treated including the summary category file, the category file after the treatment by **7.4** containing the numbers of the cited subset article belonging to each category published by related publishers in a year (such as 2017), and the category files after the treatment by **7.4** containing the article number of each category published by related publishers in different years (such as 2016, 2017, 2018 and 2019).
* Before **7.3**, **7.4** and **7.5**, the excel files should be checked to delete the abundant sheets named “sheet1”, etc.

# 8. Organization

*It is recommended that the search results by the search function* ***8.1*** *should be stored into different files according to the years for following automation by this software. In addition, results by functions* ***8.3*** *and* ***8.4*** *should be stored in one file while different years for different files. For example, if we searched author organization profiles in 2017, 2018 and 2019, we should store the related information in three files for 2017, 2018 and 2019 respectively.*

## 8.1 Search (Simple)

Search the organization information of the articles of a publisher in a year (2000-present year). The first textbox before the button is the No. of the start publisher. The second textbox is the No. of the start article group. User should indicate an excel file saving the search results.

## 8.2 Get organization list

Get the organization list that publish articles in the related databases (SCIE, SSCI and AHCI) in an indicated year.

## 8.3 Search

Search the organization information of the cited articles and the citing articles of a publisher in an indicated year (2000-present year). The first textbox before the button is the No. of the start publisher. The second textbox is the No. of the start article group. User should indicate an excel file saving the search results.

## 8.4 Search (Special)

Search the organization information of the cited articles (in special subsets) and the citing articles of a publisher in a year (2000-present year). The first textbox before the button is the No. of the start publisher. The second textbox is the No. of the start article group. User should indicate an excel file saving the search results.

## 8.5 Pretreat

Pretreat the indicated excel file saving search results of the category information in a year.

* If the indicated excel file comes from **8.1**, the obtained excel files is only for group articles named as the indicated file (excluding the suffix) + “-hhmmss” + the suffix.
* If the indicated excel file comes from **8.3/8.4**, the obtained excel files will include two files: one is for the cited articles named as the name of the indicated file (excluding the suffix) + “-hhmmss” + the suffix and the other is for the citing articles named as the name of the indicated file (excluding the suffix) + “-hhmmss-C” + the suffix.
* If the excel file including the information of the subset with 10000 articles (usually from **8.3/8.4**), the above two excel files will be changed as: one is for group articles named as the name of the indicated file (excluding the suffix) + “-hhmmss” + the suffix; the second is for the subset articles named as the name of the indicated file (excluding the suffix) + “-hhmmss-sub” + the suffix and the third is for the related citing articles named as the name of the indicated file (excluding the suffix) + “-hhmmss-C” + the suffix.

## 8.6 Treat

Treat the excel file saving search results of the organization information of articles in a year. The original organization files come from **8.5** after pretreating.

## 8.7 Combine

Combine the organization files for cited article and citing article obtained by **8.6** originally from **8.3/8.4**.

## 8.8 Match org

Match the organization file with the total organization list.

## NOTE:

* Before starting 8, please restart the software, load the matching results in **3** and confirm and/or change the year, type and database in **4**.
* If the search process (**8.1, 8.3** or **8.4**) is interrupted or stopped for a long time (usually more than 1 minute) due to network problems, etc., user should firstly close the software and all background processes name as “Microsoft Excel” and “Excel”. Then, user should delete the unfinished results for the category information of the related publisher in the excel file. For example, if the search process is interrupted when searching the third publishers, user should delete all sheets in the excel files started with “3-”. Finally, user can restart the software and the searching process by doing **3.5** and **8.1, 8.3 or 8.4**. When doing **8.1, 8.3** or **8.4**, user should set the start publisher as “3” and start the group as “0” and choose the same excel file.
* Before **8.5-8.8**, the excel files should be checked to delete the abundant sheets named as “sheet1”, etc.