Science Communication: Research Productivity and Data Analytics using Open Source Software

Vanita Khanchandani

Central Library

IIT Delhi

Week: 03

Lecture: 08 Abstract and Citation Database: Web of Science

Dear Learners, Welcome to the course Science Communication, Research Productivity and Data Analytics using Open Source Software.

So, earlier we learned about Scopus. Now we will learn about Web of Science, an Abstracting and Citation database, and helping in data extracting and collection. Web of Science database can be searched from www.webofscience.com, and its database has Core Collection Science: Citation Index Expanded, Social Science Citation Index, Arts and Humanities Citation Index, Emerging Sources Citation Index, Books Citation Index, Conference Proceedings Citation Index, Index Chemicals, and Current Chemical Reactions. It has other special collections also like specialty hosted collections CABI, CAB Abstracts and Global Health, Food Science and Technology Abstracts, INSPEC, MEDLINE. Patents Data and Preprint collections like Preprint Citation Index, Derwent Innovations Index, Data Citation Index. Speciality collections like Zoological Records, Current Contents Connect, Biological Abstracts, Biosis Citation Index and Biosis Preview. In regional collections it has SCIELO Citation Index, KCI Korean Journal Database, Chinese Science Citation Database, ARCI Arabic Citation Index. It has more than 34000 journals across the platform, 21900 total journals in Core Collection, 2.2 billion cited references, 196 billion records, 20 billion records with funding data, 109 million patents, 14.5 million datasets and data studies. It has back files up to 1900, and more than 3 lakhs conference proceedings, and 137000 books. So, this is the Web of Science interface. It is available in nine languages.

Okay. So, one can register in the Web of Science for assessing personalized features. After registering you can sign in. So, now moving towards the demo of this. Now I am logging through my account.

So, you can see that this interface is available in nine languages. And on the left hand side, the menu is there. And after creating the account one can see the profile. So, moving back to presentation again. One can do the document search as well as researcher search. So, first of all we will do the document search.

So, in document search various fields are there like topic, title, author, publication title, year and so on. So, I am taking one example in all fields like "water consumption". You can take any of the keywords. So, one can see that we have got 108,853 results. On the left hand side we have various filters, like we can select on review articles, open access, early access articles and enriched cited references, open publisher-invited reviews.

We can filter by year, by document type, by researcher profile, and so on. And if you see one document you can see that in one document abstract is there, document type, keywords, author information is there, funding agency information is there, ISSN, current publisher, research area, Web of Science categories, and journal citation indicator is there. So, you can add these results to the marked list. All records on the page or records from this and maximum records you can add is 50,000. See I am adding this to create a new list on water consumption or water, and creating a list.

So, a list has been created and you can see on the left hand side that you have a list on water consumption. So, moving to export option, see there are various options available to export and if we want to export the results in Excel, so you can export all the records available on this page, or maximum records you can export is 1000 at a time. So, I want to export 1000 then, after 1000 I can take again 1000 results, and so on. And suppose I need a full record then I can select the full record in record content and then export the results. It will take 1 or 2 minutes.

See in downloads we have the result in Excel file. We have the full record or bibliographical details of 1000 data. You can see here that we have all the 1000 records exported in the Excel file. Okay. So, we can again search for different key terms. Suppose we want to search "water consumption", and we can add a row and want to search about "water treatment". We can add date range also. Suppose not all years but the publication date of the last 5 years say and we can do the search. So, we have got 8,468 results, and you can see water consumption and water treatment is there for the 5 years from 2020 to 2024. So, you can export the data, you can put this data to a mark list and you can even view the history for the current session like water consumption and water treatment. Okay. From history you can even combine the search. If you want, you can combine the search queries. Now we are searching for water pollution for the last 5 years.

This much results we have got and in history we can combine the term water consumption and treatment along with water pollution. You can combine these two search terms to combine searches and go to advanced search. Terms 3 and 2 can be

combined and added to the query. And you can combine the query using Boolean operators AND, OR. Now you can add. #4 has been added to the query and the result from this combined search is 1236. And you can see some of the articles are on pollution, some are on water treatment and some are on water consumption. See first article is water consumption and pollution. Okay, and second is on water treatment and so on. So, we have till now learned how to conduct a basic search and how to combine two searches. How data can be exported and maximum of 1000 results can be exported at a time. How rows can be added, how to combine search results and now we move to Boolean operators.

So, Boolean operators are of three kinds OR, AND, and NOT. For OR means one term must appear like water consumption OR treatment. AND means both terms will appear water consumption AND treatment. NOT one term will be excluded. Water consumption NOT treatment means only results pertaining to water consumption will be available, treatment will be left. So, for this you can see in PPT results for all three search terms are there. Now we move to a demo. Water consumption OR treatment. So, this much result we have. So, in this water consumption AND treatment is there. In title, in abstract. Okay, I have selected all the fields. So, even in publications. So, instead of all fields you can select topic, title, author. Like in the title, now you can see that water consumption or treatment will be available. See here water consumption is there. And in another article. For all water treatment there is either water treatment OR consumption.

Okay, so some places you are seeing that consumption is there but that is not water consumption rate that is energy consumption. So, for AND. Now I am putting in as an AND operator. In the title search again. So, 55 results are there of water treatment AND water consumption. Okay. Sometimes energy consumption is also there. But the last operator is NOT. In the results in all results water consumption is there excluding the term treatment.

So, now I'm moving to proximity search. In Scopus P and W. Were there. While in the Web of Science NEAR. And the same are used as proximity operators. NEAR can be used NEAR/X where X is the number of the terms. If you don't put any terms then it automatically takes fifteen words. So, one can search for documents. Having fifteen words after that. Okay. If you put the number, then it will just search the sentence in which it has to come after that word.

Like for example Salmon NEAR Virus. If you are not putting any number then it will search salmon NEAR/15 virus means virus will become after fifteen words. Okay. And you can put NEAR double quotes also for searching a title from a source such as journal, book, proceeding, etc. So, now we move to an example. So, there are 583 articles. And the Virus is just two terms. Covid NEAR/2 virus. Sometimes, Covid and virus have a space. In some words, there is a colon parenthesis like this. Okay. AND is there. And, if

you don't put any number then see. It has 1749 results. And, NEAR can be used for searching documents from a source. Okay. So, these are all titles. So, NEAR can be used to search for a title from a source like titles from journals, books, etc. If you put NEAR double quotes. And same can be used for searching addresses NEAR can be used with double quotes and same also for searching addresses. So, coming back to one more example. So, in the address bar we can search McGill University in Canada. So, you can see that McGill University. In the country, Canada is there. And in affiliation, McGill University is there. Okay. So, the same is used for searching this. Yeah one thing I want to add is that in Boolean search there is an order of precedence. Okay, if we put the bracket then. This order or operator precedence can be overridden. So, the order operator precedence is as follows NEAR followed by the same, then NOT, AND, OR.

But if we put the brackets then this operator precedence can be overridden. Influenza OR flu AND Avian. So, it will search for avian influenza, OR avian flu. So, if we put the bracket then the operator precedence will be overridden. So, now we come to wildcards searching using wildcards. There are three types of wildcards in the Web of Science. First is asterisk (*), which represents any group of characters or including no character. Question mark (?), which represents a single character. Dollar sign ($) represents zero or one character.

Okay. So, now we come to demo once again like we search. Colour has two spellings. One is American and the other is British. So, the colour is British and American is the color. So, here we put an * for the character. So, it is searching for what colour and color. You can even search for Women by putting dollar sign also. So, it will search both A and E. And another search is for an Organization. You can put the question mark also. So, it will search Z and S both. And for the author field one can search using a question mark. So, it will search Barthold, Bartholdi. All these kinds of results. You see why I. All kinds of results are there. You can see if you put the question mark in the author's surname then I and Y all will come.

Now we move to Cited Reference search. An important search in the Web of Science after document search. So, how can we do the cited reference search? I'm telling you that. So, for cited reference search you have cited author, cited work, cited year. So, you can write a cited reference whatever you want to choose. You can see KK Pant's work in cited references, for chemical engineering for 2020-24. And you can see how much is citation these citing article citations they have received means the first article has been cited by 30 articles. You can do a cited reference search and in the last researcher search there is. So, you can put the researcher name, author identifier, organization. So, I am giving this writing the same name as Pant. You can also add some variants. So, this is. You can even do the combined search. You can view the combined records. If you want to combine, you can combine all his name variants for doing a cumulative search. You can even

submit the correction and you can merge the records. See his h-index is 51, total number of publications is 321.

In the Web of Science the last search is Structure Search. If you have a subscription to Index Chemicals and Current Chemical Reactions, then that will be useful for searching compounds and reaction data. And one needs to sign in through the credentials. And you can search for chemical compounds and reactions that match a structure query that you create using an elemental drawing tool from Dotmatics Knowledge Solutions. You can also search for data associated with compounds and reactions by entering your search term in compound and reaction text fields. So, if you have a subscription for Index Chemicals you can have a structure search also. So, before ending this I want to revise how you can export the data in the Web of Science.

So, in documents, you want to search all the fields you are not selecting any parameter from this and you are just putting a simple term like water pollution. And you are searching for the term and from this. From one left 81,423 records you want to export the results. So, you can click at export and then you can see there are various formats like EndNote Online, EndNote Desktop, add to my researcher profile, plain text, RefWorks, RIS, BibTex, Excel. So, for Excel I am selecting only 50 records. From 1 to 50 records, and in record content I am selecting the full record and exporting the results. So, this is the export file which I have got. So, for 50 results this has full records information like author, author full name, article title, source title, language, and so on. If you have any queries you can mail us in the Discussion Forum. So, now we end the Web of Science. Thank you.