**Lecture 7: Abstract and Citation Database Scopus**

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

**Vanita Khanchandani**

**Central Library**

**IIT Delhi**

**Week: 03**

**Lecture: 07**

Dear Learners, I welcome you all to the second week of the course on Science Communication: Research Productivity and Data Analytics using Open Source Software.

Here we will learn about data collection sources on research metrics. Today I will brief you on Scopus. This module will impart knowledge for conducting basic advanced search, how to export data to know the author and institutional outputs, to impart knowledge about editing the author profiles, and to understand the journal level metrics available in Scopus. We can reach Scopus through the URL www.scopus.com.

Scopus is a comprehensive, multi-disciplinary, abstracts and citation database with enriched data and links scholarly literature across a wide variety of disciplines. Scopus can help you understand who is researching what and where with greater accuracy providing invaluable data for strategic decision making. It can help you in conducting systematic literature reviews. Scopus has over 7000 publishers, 21.7 millions open access items, 17 millions author profiles, 94.8 thousands organization profiles, and 27.9 thousands active journals. To access personalized features, one should create an account in Scopus by clicking create account. See now I am signing in Scopus as I already have an account with Elsevier.

So, I am using the credentials and signing in Scopus. After registration, sign-in into Scopus to take advantage of a range of personalized features of the Scopus. You can create and save a list of documents, authors or sources, save your searches, set up alerts, customize the export feature of the preferred reference manager tool, manage any request you have made to Scopus, support team, manage your Elsevier account information and privacy options.

So, now we move to document search. Documents include journals, articles, book chapters, conference proceedings, articles in press, and data papers. We can follow the following steps to conduct a basic search. See we first enter the term. I am putting in a

one term Coffee Product. In this I am writing just a produc and then putting one asterisk and closing this term.

I will let you know about the details why I am putting the asterisk. This is a loose phrase searching term, and this will search coffee products, coffee product, and coffee productivity. All the terms containing coffee product, or related terms together. So, now I am searching. This is the search base, and see I have got 2966 documents.

On the right hand side we can save the search, set the alert. I am saving the search with this term Coffee Product. So, my search term has been saved and I can further refine the search using open access, year, subject area, and different refined features. One can show this result according to date, newest, oldest, cited by highest, cited by lowest, according to relevance of first author name A to Z, or Z to A, source title A to Z, and Z to A. We can download this result or even export the result. If we want we can export the results by different file types such as CSV, RIS, BibTex, plain text, and we can also export the result into following reference managers, Mendelay, Refworks, Zotero and more. So, one can export up to 20,000 documents in CSV format having citation information only. And for other information, all documents we can get other bibliographic information, abstracts and keywords, funding and this has less than 20000 results. So, I can get the full information and can export the results. We can email the results, view their references, view to whom they are cited by, save the list, and we can get the citation overview of the first 2000 results.

For analyzing the data, see I am now analyzing the data. So, these 2966 documents are analyzed, documents by year and then document by source, by author, affiliation, country or territory, subject area, type, funding and sponsor. You can email these graphs, you can export or take the printouts of these graphs and even change the year range from. This is from 1934 to 2024. You can change the range even. This is an overview of basic document search. You can even add the search field. One can add the search field like I am adding one search field climate change. So, now the search will be coffee product climate change. So, now I have got 279 documents in this area and I can analyze the results, export the results. I have already told you how to export, analyze the results. You can combine the queries if you want and you can save the search.

I have saved the search with the name Coffee Product. You can even set the alert and this will come to your mail every week or every month. You can see every day, every week, you can choose every month, and you can choose the dates also and set the alert. Whenever any new document comes to this search, you will get an alert in your email. Now we move to advanced documents search.

Advanced documents search allows you to create more complex queries using field codes, proximity operators, boolean operators, etc. In this example, I am interested in

finding articles on lung cancer. So, I am looking for field codes, TITLE-ABS-KEY (title-abstract-keywords). You can select the field code from here, TITLE-ABS-KEY and then write “Lung Cancer”. Another code which I want to add is AND NOT so that it can, the search will be on Lung Cancer only, not on other terms like Tumor.

So, I am writing this and then I am putting AND NOT Tumor. To refine more, I will go to subject areas now and as Cancer is a part of health science and I am taking this term, I am searching in Medicine. Okay, now I am searching for this term. So, I have found 163,281 documents. You can search using subject codes also. If you have knowledge of subject codes from this option, we can get help from the Support Center, and you can search for the subject codes. So, you will get All Science Journal Classification Codes. You have got all the Journal classification codes. So, this can also be used for advanced search like you can add subject terms. Now our advanced search feature is over and I will move to other basic searching tips which I want to share with you people.

You can conduct a search using a loose phrase. You can use double quotation marks like earlier I have told you and this will search documents like I have given examples of climate change coffee product, that will appear together in not a loose phrase search or separate words. For example, in the TITLE-ABS-KEY, I will search for coffee product without quotes and this will search for the document where coffee and product appear together or separately. For Boolean operators, there are three Boolean operators OR, AND, AND NOT. OR will search where one term must appear, e.g., Coffee OR product. AND will search both the terms together e.g., Coffee AND Product. AND NOT will exclude one term, e.g., Coffee AND NOT Product. Now in this term the Product will not appear. It will search only for Coffee. There are two types of wild cards. Question mark will represent a single character. For example, WOM?N will retrieve both Woman and Women. WOMAN and WOMEN both.

And for representing any number of characters, even zero. One can use an asterisk. For example, COMPUT* will return Computer, Computers, Computerized, or Computerization. I have already told you about the example Coffee. Coffee Produc* will find Coffee Product, Coffee Products, and Coffee Productivity to cater to in a search.

For searching exact phrases one can use braces, or curly brackets for example, {Oyster Toadfish}. If you search without braces, Oyster Toadfish will retrieve 184 documents. And if you put curly brackets, {Oyster Toadfish}, then it will retrieve 170 documents. I will show you the example shortly, and for proximity operators we have to use W/n means within, PRE/N means precedes. Within will don't have any order, precede will have an order while searching the document. For example, if we are in title search, we search TITLE(COVID PRE/2 Virus). This means COVID precedes Virus within two words. After conducting this search we will retrieve documents in this search COVID-19 Virus is there every time. While within COVID can be found within a distance of two

words from Virus. Earlier, We should have an order that means one term should be preceded with the other term while in this that can be found within. So, no order is required. TITLE(COVID W/2 virus) will retrieve 1000 documents, and it is like COVID-19 Virus, Virus during the COVID-19.

I have already told you about All Science Journal Classification Codes and the example. Now I will show the demo of pre, within and exact phrase searching. So, I am searching for TITLE(COVID Pre/2 Virus). So, 472 searched documents are there. And you can see every time COVID-19 virus means COVID is preceded by just two terms. Now I will show you instead of Pre, I am writing TITLE(COVID W/2 Virus). So, this has retrieved 1004 documents, and COVID-19 virus, and updated COVID-19 vaccines. So, the Virus has two words and updated. Virus during COVID-19. So, this is Within and earlier we have Pre. Now I will show you the example of a curly bracket which means exact search also. In basic search you can just put the curly brackets. So, 170 documents are there and now I am removing this curly bracket. 184 documents. So, this search has more recall than precision, and when we put the curly bracket it has more precision than recall. Both our basic and advanced search is over and now I am moving to author search. So, Scopus author profiles allow researchers to demonstrate their reputation and impact of their outputs, and attract other researchers for networking and collaboration. Author profiles allow potential employers and funding agencies to evaluate your work quickly and easily in one place, inform hiring and promotion decisions, and decide which researchers or teams of researchers will receive research grants. To search an author, select the author type from this search using drop down and enter the last names.

For example I am putting Pant in the last name and for the first name I am writing Kamal Kishore. If you want to add affiliation you can also enter the affiliation. I am writing the Indian Institute of Technology Delhi. Or, one can search even using ORCID or some keywords. So, I am clicking this Kamal Kishore Pant or Pant, Kamal Kishore has 326 documents, 56 h-index, affiliation Indian Institute of Technology Delhi. This is the result of Pant, Kamal Kishore. One can set the alert or even document alert, or author citation alert. Whenever any new document becomes a part of this, or even the author receives the citation you will get the alert through your email. You can even save it to the list, and this will reflect in myScopus under the saved list. I am writing, KK Pant and creating one list. So, as you can see in myScopus, I have one author list KK Pant. Okay. Moving back again you can see the Scopus ID of professor Kamal Kishore Pant. His ORCID even. You can view his H-index, and h-graph also. You can see his 56 documents have received 56 citations. So, his h-index is 56. This is the slope where both the intercepts are meeting. So, you can change this by excluding his self-citations, or by excluding citations from books and update the graph. you can even change the date range. You can export the graph print or email the graphs by source, by type, by year, by subject, and professor Pant has received 11572 citations as of now having 150 co-authors. The important thing which

I want to share is that one can edit the author profile, if you find that there is some name variant or some institution change. One can edit the profile through this edit profile. in this you will get a dialog box Is this your own profile. Yes, then continue and you can change by reviewing the request. From this you can remove the articles. you can submit the request. or even you can remove the article from the profile. Or one can even add the article from the profile. if any of the documents are missing you can add even. If you move to this help box, one can get the web form for correcting the author profile. this is the web form you are getting. From help and support, here you will decide your role. you are a librarian, author, editor, publisher, user. and after deciding your role you can click the reasons. why you want to update the profile, if you have any missing document affiliation correction, or the profile correction, citation correction whatever. you can just choose one or the reason. and you can upload the file regarding which you want to update and send your request. this request will come to your Scopus profile. You can see from this request dashboard. if you have ever requested any request you will get details here. The important thing is that you can set the document alert for each of the documents. and if you click any one of his articles you can set the citation alert for each document individually even. see here we have author metrics which will tell you this is a new feature recently inserted in Scopus. So, you can see how much collaboration he has. He has international collaboration 25.4% and his Field with Citation Impact is 1.78. Field with citation impact (FWCI) is the ratio of citations received relative to the expected world average of the subject field, publication type, and publication year. here it means 1.78. So, it means he is globally receiving 78% more citations than expected. So, now I am moving to research discovery and this is a new inserted search in the Scopus. Research discovery allows you to connect with the researchers around the globe. you can search here by writing a simple keyword term on the topic you want to search. For example, I am just writing about Lung Cancer and searching. So, here you can refine the search through the year, through the country, organization, etc. So, I am not refining the search now. Just see I am getting these are the researchers who are working on the topic lung cancer around the world and I am interested in looking at the profile of He Lie from China. I am clicking on the preview profile and see. he is from the Chinese Academy of Medical Sciences and Peking Union Medical College, China, experience having 22 years plus. and you can see if I want to collaborate with him, I can connect him with you through his profile details. and moving to the last search is affiliation or organization search. I am interested in viewing documents on how Indian Institute of Technology Delhi is working. This is the affiliation ID of the Indian Institute of Technology. This is the Scopus ID. So, you are getting documents for the whole institution 54,500. While documents for affiliation 54,477. In Scopus there is a parent-child relationship, the whole institution is a parent, while affiliation is a child. So, obviously being a parent the whole institution will receive more documents, while affiliation being a child has less number of documents. IIT Delhi has 11,307 authors. You can save the author lists for the 2,000

items only. You can set the alert, you can export the subject area, and document by subject area. you can even move to another tab affiliation hierarchy. here see Indian Institute of Technology has one child or affiliation. The Centre for Biomedical Engineering has 769 documents. This is collaborating affiliations telling about the collaborations. IIT Delhi has 924 documents co-authored with AIIMS New Delhi, and then with University of Delhi. document by source. IIT Delhi authors are contributing more in *Lecture Notes in Computer Science*, including its sub series on *Lecture Notes in Artificial Intelligence*, and *Lecture Notes in Bioinformatics*. 580 documents are from that source followed by Journal of Applied Polymer Science and so on. It has patents as roughly estimated at 10,825. For documents for the whole institution you can click here. if you select you will view citation overview and through this citation overview this will let you know the h-index of IIT Delhi. Although the results are more than 20,000, if you want to select the documents in one go then you have to write your email. but for more than 20,000 documents they will mail you.

See you have almost 45,806 sources listed in the Scopus, and you can search and refine them according to the open access journals, and according to the Quartiles. There are four Quartiles in Scopus. The first Quartile having top 25 journals in it. then the second Quartile after 25 is 26 to 50. then 51 to 75. Then 76 to 100 in the fourth Quartile. You can search in it through subject area, title, publisher, or ISSN. This is the All Science Journal Classification Codes scheme. you can search from here also. Suppose I want to search on Oncology. Medical oncology is there. I want to search for this. So, the top journal is *Ca-A Cancer Journal for Clinicians*. and I am searching for that. and you can see in this one can have the Scopus coverage, publisher, ISSN, subject area, type, and on the right side you can see CiteScore of this journal, SJR value, SNIP value. I will tell you the details of it. Now come to the CiteScore. The CiteScore for the for 2022 will be calculated by the citations of four years divided by the total number of documents published during those four years means 2019, 2020, 2021, 2022. the citation received in the four years divided by the documents published during those four years. you will get the CiteScore while for CiteScore Tracker this is for the current. Every month it keeps on changing and you can see CiteScore tracker for 2023 and at the bottom you have CiteScore rank under CiteScore. CiteScore rank for Medicine > Oncology the rank is one and percentile is one. This means the *Cancer Journal for Clinicians* is the top one percent journal. Now come to SJR, another journal metrics available in Scopus after CiteScore. So, five years ago SJR Journal Rank is a major part of the prestige of scholarly journals as just the Scores are computed using network analysis of citations received by the journals. Methodology accounts for a number of citations, the source of citations, the citations from high prestige journals, being more than those from journals with low prestige. SNIP is a Source Normalized Journal Impact Assessment, is calculated as the number of citations given to the present year to the publication in the past three preceding years, divided by the total number of publications in the past three years. Simply SNIP is the number of citations

given to the present year, for example say 2022 to the publications in the past three years meaning 2021, 2020 and 2019. You will divide by 19 by the total number of publications in the past three years. A journal with an SNIP of one has the median number of citations for the journal in that field. The impact of a single citation is given to a higher value in the subject areas while citations that are likely and lower value in the subject area where citations are more likely. So, now we come to compare sources. Suppose I want to publish a review article on Lung Cancer in a journal, and I want to compare the sources on oncology. So, first I can compare up to 10 journals in this. In the source titles I will. You can compare this with *Cancer Cell* and another journal. So, you will see the graph. This is for the *Cancer Journal for Clinicians*. Then *Nature Reviews Cancer,* and last is *Cancer Cell*. So, how are they doing? This is the CiteScore of publications by year. You can check SJR value, SNIP, citations by year, documents by year, percentage not cited by year, percentage review articles. Obviously if you see this graph before sending your review article, you will see obviously the *Cancer Journal for Clinicians* will be the best suitable one. So, after CiteScore you can go to CiteScore rank and trend. You can see how a journal is performing. Obviously the *Cancer Journal for Clinicians* is on the top in the Oncology category. You can see the content coverage. Okay from here you can check the citations overview for 2024 of this journal. Exclude the citations from books. How this journal is performing you can see. For this Score you need to export data from the Scopus, so that you can analyze using open source software. So, exporting the data with the search string for Lung Cancer NOT Tumor under Medicine, document type articles, limit to language, and to the journals with the year range 2020 to 2024, and now I am limiting this. So, after refining the term I have got 19,441 documents. Now see how I am exporting. I am exporting the CSV file for both things. So, if I want to get the whole information then even I can save as preference. If I save this as preference, then for all the searches the same kind of export will be done. So, now it is exporting to a CSV file. So, the export has been completed. We can see how the results look like. So, here you have the documents. You can check the number of documents exported in Excel. This is for the Scopus. For your queries you can write to me at vanita@library.iitd.ac.in, or you can write to us on Discussions Forum. Thank you.