Information-seeking strategies

Juan D. Machin-Mastromatteo

Universidad Autónoma de Chihuahua, Mexico

jmachin@uach.mx

The final version was published as (please use the following reference when citing it):

Machin-Mastromatteo, J. D. (2025). Information-seeking strategies. In D. Baker & L. Ellis (Eds.), *Encyclopedia of libraries, librarianship, and information science* (Vol. 2, pp. 249-258). Elsevier. https://doi.org/10.1016/B978-0-323-95689-5.00196-6

Keypoints

- This entry distinguishes between scientific productivity, scientific and academic production, and scientific dissemination, emphasizing the importance of understanding these categories for effective information seeking.
- Classification of scientific journals: according to their publication models, journals can be divided into diamond/platinum, green, and gold open-access journals.
- The peer-review process is identified as a crucial element for a journal to be considered scientific, ensuring the credibility and reliability of the published research.
- A variety of document types are essential for academic research, including articles, books, theses, and conference proceedings.
- Digital repositories and databases provide full-text access to scientific and academic documents, while bibliometric indices are limited to metadata.
- Information-seeking behavior models: this entry enumerates several models and theories of information-seeking behavior, offering insights into how individuals search for and use information.
- Developing a search query implies selecting the words we are going to put in the search box of an information system and setting the relationships that these terms must have in the resulting documents, which is expressed with the aid of Boolean logic.
- Boolean logic represents relationships between entities in symbolic form. Hence, in information-seeking, Boolean operators (AND, OR, NOT) are used to build priory relationships that describe the contents of the documents we wish to retrieve.
- This entry proposes various strategies and tactics tailored to different academic tasks, emphasizing a nuanced approach to information retrieval.
- Boolean logic is the strategy used to refine searches for academic materials and navigate large data sets.
- The future of information-seeking will be impacted by artificial intelligence and natural language processing, as it might allow information systems to offer more intuitive search experiences in the short term.

Abstract

This entry presents an examination of information-seeking strategies within the academic and scientific realms. It delineates the characteristics of scientific and academic information, including distinctions between scientific productivity, production, academic production, and dissemination; while it also provides a categorization of the types of scientific documents, the usage of citation indices, retrieving full texts, and complementary information sources. After presenting a summary of some information-seeking behavior models, it details several useful information-seeking strategies to broaden or narrow search results in information systems, including Boolean logic.

Keywords: information-seeking strategies, scientific research, bibliometric indices, digital repositories, open access, Boolean logic, peer-reviewed journals, academic information, information retrieval, information behavior, search methodologies.

Introduction

Before tackling the specifics of information-seeking strategies, it may be useful to provide some characteristics of academic and scientific information. First, we could distinguish among different concepts that characterize the level of information (e.g., its specialization and intended audience), which would be the following: 1) scientific productivity: measurement of the behavior of knowledge institutions or their researchers with respect to their scientific culture and activity; such measuring could be achieved in various scientific information systems, but particularly in bibliometric indices (e.g., number of results after a given search, number of publications per author, institution, published in a given journal, the most cited authors and sources, among others); 2) scientific production or communication: it consists on the publication of the scientific results of researchers; all scientific information systems provide access to such production; 3) academic production: generation of other kinds of products derived from academic or scientific activities (curricula, guides, presentations, work books), such materials would generally be available online in repositories or in some websites, mainly for teaching purposes; 4) scientific dissemination: special types of scientific products that have been adapted and recoded in their language to cater to a wider audience, if possible, they will aim to be consumed by society in general (Tarango & Machin-Mastromatteo, 2017). Additionally, when seeking information for academic or scientific purposes, we could add a separate category of materials, integrated by reports, statistics, or even news, which can be considered complementary sources in most cases, while in others, they may be the main sources used if the topic is very new; which tends to happen more often in technology-related fields.

Types of scientific documents

Scientific journals and articles are the most widely appreciated information sources for scientific research. Journals are comprised of articles that contain current research data and information in a given scientific field. For a journal to be granted the level of scientific, its contents must be peer-reviewed before being published. Journals originated in scientific communities and societies, but universities, research centers, government agencies, and private publishers also publish them. They have a distinctive identity, expressed in their thematic scope, and have different specifications as to the type of content they accept, their length, format, and citation style. Journals are registered with the International Standard Serial Number (ISSN) and to ensure the permanence of publications, the use of the Digital Object Identifier (DOI) is incorporated.

It is possible to categorize journals in several ways. First, according to their publication and access models, we have: 1) diamond or platinum open access (OA) journals: they do not charge publication fees to authors and offer access to the full texts of their contents at no cost to their readers. They are generally published by non-profit organizations (such as universities or scientific societies); 2) Subscription journals that offer some OA possibilities: they allow OA author versions of articles to be archived in repositories. There are two main types: green OA journals allow self-archiving of the author's version, either immediately or after an embargo period (depending on its specific policy); and gold OA journals: their articles are published at no cost for readers, but the authors or their institutions must pay an article processing fee (APC) to the publisher; and 3) other commercial journals without OA possibilities. When seeking scientific articles, it may be useful to consider this classification, as it may not be possible to retrieve the full text of some articles because of the publication policies, and thus, the user will require an institutional or individual subscription to access them.

Journals can also be categorized by their contents and status in the scientific community as 1) indexed journals, as they are included in bibliometric indices, i.e., Web of Science (Clarivate Analytics, 2024), or Scopus (Elsevier, 2024); 2) peer-reviewed journals: all indexed journals are refereed, but not all refereed journals are indexed. To be peer-reviewed, as stated above, is the basic condition for a journal to be considered scientific; and 3) scientific dissemination or popular science journals: intended for a wide audience, their contents are presented in a simple language that is easier to understand by non-scientists.

Regarding articles, there are comprehensive categorizations in the literature, such as the one available in the Publication Manual of the American Psychological Association (2020): 1) quantitative articles; 2) qualitative articles; 3) mixed-methods articles; 4) replication articles; 5) quantitative and qualitative meta-analyses; 6) literature review articles; 7) theoretical articles; 8) methodological articles; and 8) other types of articles.

However, we could also work with a simpler categorization in relation to their contents and purposes: 1) theoretical or literature review articles: these contain reviews of the specialized literature on theories, current research or methods, and philosophical contents, with scientific explanations and explorations; 2) systematic reviews: comprehensive reviews that are conducted by employing a specific methodology for the selection of sources and this is what differentiates them from the literature review articles, which are selective reviews and they do not follow a set methodology; 3) scientific articles: these include information supported by fieldwork and are usually presented in the IMR&D (introduction, methodology, results, and discussion) structure; 4) other minor genres: editorials, letters, columns, book reviews, essays, reports, case studies, working papers or protocols (Tarango & Machin-Mastromatteo, 2017. A more recent category worth mentioning is preprints, which are usually articles (although there are also preprints of chapters or books). However, these are shared by their authors before undergoing peer review, in order to provide their research results as quickly as possible to the scientific community. As these are not peer-reviewed, it must be kept in mind that they might contain imprecise or even erroneous information. When seeking information, it may also be relevant to consider the different types of articles, as some will be more useful than others, depending on our particular needs. Also, the IMR&D structure allows us to quickly find the content we are looking for at a given time.

According to Tarango & Machin-Mastromatteo (2017), scientific books are non-periodical publications that conform to treatises of a certain thematic coherence. They may also be the result of research, but often documentary, as they analyze, synthesize, and integrate a specific subject within the scientific, technological, humanistic, or different phenomena through a description, narration, or detailed exposition of its topics. They can be authored books or compiled or edited books and because of their length (much larger than articles), they allow for a more exhaustive and broad level of expression about an area of knowledge. The cited authors also define book chapters by providing the following characteristics: 1) they are the main subdivision of a monograph or book; 2) such fragmentation facilitates reading and optimizes the organization and integration of contents; 3) their length varies depending on its purposes; 4) they conform a way of dividing authored books. Edited books contain independent content and authorship units (each chapter may have its own authors). As stated before, other types of documents that may also be valid information sources depending on specific needs will include reports, statistics, or even news, as well as theses, patents, and conference proceedings. Information systems will have a variety of these, while some may specialize, for instance, in providing access to a few or even one kind of document.

Although this entry is centered on information seeking for academic and scientific purposes, there are many other purposes why we need to seek information, such as to solve a doubt, confirm a fact, look for data, or make decisions in our everyday lives. Regarding research purposes, it is important to locate the best research sources and academic work, since ideas do not come from a vacuum. For instance, I may come up with a research idea that I consider novel, but it turns out that if I had conducted an appropriate search before, I would have known that it was an already well-researched idea. But, if I become well informed about the topic and studies similar to the one I want to do, I will be able to come up with a more novel approach to it. Do not underestimate the search for and the use of good sources, as these are the basis for new research because every new study builds on what has already been published.

When searching for information, I should think about the following questions: what, why, where and how do I search. All these questions are interrelated. The *what* and *why* I search for information has to do with my research topic and the type of work I wish to do, or the kind of information need I require to address. Answering these first two questions will determine *where* I will look for information and in which specific place, usually a digital information system. *How* I search will depend on the system I have chosen because it is the type of system where I most likely will find the types of documents that will address my information needs.

The usage of citation indices

Suppose I need to examine the publication dynamics of a discipline or topic through a sample of the best scientific publications. In that case, the system of choice will be the citation indices, which include the most cited and well-ranked scientific publications. They are also very good information sources for starting a research process and determining the key documents that should be included in a literature review. The indexes include only the referential information of each document; that is, you cannot download full texts but only visualize the descriptive data or metadata of the publications and the trends resulting from their characteristics. On the commercial side, we can find Web of Science's indices (e.g., Science Citation Index Expand, Social Sciences Citation Index, Emerging Sources Citation Index, Arts & Humanities Citation Index, among others) and Scopus. However, these require institutional subscriptions to access them. There are alternatives with free versions, such as the ones offered by Dimensions (2024) or Wizdom.ai. However, we should note that these latter options are different than the former, which require journals to pass a rigorous evaluation before being added to a specific index. In the cases of Dimensions or Wizdom, they may offer larger datasets, but they do not evaluate sources. Hence, if I am interested in results exclusively from well-ranked journals, these latter options will not be adequate.

By using the indexes, I can check and analyze general publication trends after conducting a search, such as the number of papers published over time on the topic of my search, who are the authors and institutions that have published the most of these results, and the journals that have published most of them. This lets me know how broad and dynamic the area I searched in the indexes is. For instance, if I see that the publications over time conform to an exponential curve (or the number of publications increases each year), which will happen in most cases, then I can infer that it is a topic of interest and every year, there are more publications on it; conversely, if the curve (or the number of publications) decrease each year, then it is a topic that might not be of interest anymore. Indexes can also help reveal who the experts are by looking at the authors with the largest numbers of publications on the topic I searched; also, where they work (through the institutional affiliations), and where the research in this area is published (specific journals or books). The indexes provide this data both for a general area of knowledge, which is possible to check by using the browsing features, and on a very specific topic, which requires conducting a search.

Retrieving full texts

In order to retrieve full texts (i.e., articles, chapters, books, and even theses), then the information system we should use include aggregator portals such as those offered by EBSCO and ProQuest. There are also databases of commercial publishers, such as Elsevier's ScienceDirect, Springer Nature's SpringerLink, Taylor & Francis Online, Wiley Online Library, and Sage Journals, just to mention the so-called big five scientific publishers. However, other smaller publishers also offer their databases or online libraries, but these require institutional or individual subscriptions to them in order to be able to download full texts. Otherwise, we can still download some content from them, but these will be limited to those published in their gold OA journals, or those offered in OA for a limited time.

If we do not have access to these resources, we can use disciplinary, institutional, national, or international repositories; obviously, each of them will have a different coverage, and some will be larger than others. For instance, institutional repositories will be the smallest ones, as they will offer only the publications from people affiliated with a given institution that were archived. If there are no systematic archival policies and procedures in a given institution, then the repository will have fewer works, and it might not represent well the overall intellectual production of such an institution. Also, even nowadays, many institutions do not have an institutional repository, especially in developing countries. Although many repository management software solutions are open source, they require specific hardware, technical know-how, and a genuine institutional willingness to run a repository, as well as appropriate processes and policies to support it.

Some disciplines will have international repositories. For instance, library and information science counts with *e-LIS: ePrints in Library and Information Science*, and there are other larger international repositories that are multidisciplinary, such as *The Social Science Research Network* (SSRN), *Zenodo*, *Arxiv*, *COnnecting REpositories* (CORE), which is the largest international repository, as it aggregates records from many smaller repositories through the OAI-PMH standard. Another good resource for retrieving open-access journals and articles is *The Directory of Open Access Journals* (DOAJ), which supports searching by journal and also by topic. In the Ibero-American context, we have the regional repository portal *LA Referencia*, which harvests records from 12 national repositories (from various Latin American countries and also including Spain), and there are full-text databases and journal portals that also serve the function of repositories such as *The Online Scientific Electronic Library* (SciELO), *The Network of Scientific Journals of Latin America and the Caribbean, Spain and Portugal* (RedALyC) and *DIALNET*.

A very useful resource is *Google Scholar*, since it gathers practically all academic and scientific documents under the same search engine, regardless of whether a commercial or fully OA publisher published them. Indeed, Google Scholar incorporates the records and web locations of virtually all documents contained in all resources, whether commercial or open, as long as they follow the technical best practices and standards used in scientific publishing (which is the basic requirement so that Google Scholar indexes a given digital source). Although Google Scholar does not store any full texts, it is a very useful resource for retrieving them through its version feature, as it allows OA versions of documents to be located if available.

As mentioned before, in the section about the bibliometric indices, it is important to consider the size of the datasets in each information system. Indices might have smaller datasets because they only gather those sources that have approved their evaluation processes and are included due to their content selection criteria in the case of book publishers and book series. In turn, information systems that do not aim to evaluate the sources they index will have larger datasets. These considerations may be important for certain information needs, and depending on them, one will choose to work with either indexed sources or larger datasets. Aguillo (2023) offered some record count details into the most important systems; although these numbers constantly change, they allow illustrating the differences among their sizes; as such, in millions of records, Google Scholar has over 400, followed by Lens (252), OpenAlex (249), Semantic Scholar (211), Scilit (154), Dimensions (134), Web of Science (117), and Scopus (88). Additionally, as of 2024, CORE has 299 million records. Considering these numbers, if we conduct an exhaustive search in Google Scholar, it might not be necessary to replicate it in other systems, as Scholar will most likely include the records present in other systems. However, it has fewer and less sophisticated options for filtering and limiting results than other systems.

Complementary information sources

In academic and scientific works, it might be pertinent also to include complementary sources such as news, statistics, public policies, and reports from governmental or non-governmental organizations. However, it must be stressed that the keyword is *complementary*, as these sources are best to formulate a research problem, establishing its context and justifying its importance, as an academic paper must not be filled with news or other complementary documents. *Most* of the bibliography, especially in a scientific article, should include other scientific articles and sources recently published because that is where the latest research is published. In order to locate complementary sources, regular Google may suffice, which also supports many of the search strategies, as detailed below.

Information seeking

When seeking information as a library and information scientist and especially when providing advice or training to users to develop their information-seeking skills, it might be useful to have some background about information-seeking behavior. This is a large topic in library and information science. However, some insights can be derived from the following summary of a non-exhaustive selection of the information-seeking behavior models that have been proposed over the years, which are integral to understanding and enhancing processes of information retrieval:

- Static Model: describes users' activities during the information-seeking process. It is largely founded on statements and theoretical diagrams, attempting to describe user activities or decipher the causes and consequences of those activities (Allam et al., 2019).
- Wilson's Model: a pioneering framework and probably the most influential model for in understanding information-seeking behavior. Initially proposed in 1981, it includes concepts related to the information user, the use of information, the exchange of information, and the informal transfer of information between individuals and emphasizes that physiological, affective, and cognitive needs trigger information seeking rather than a mere 'information need' (Agarwal, 2022). It also emphasizes the concepts of success or failure in information seeking, which link to processes of evaluating and reformulating information (Agarwal, 2022; Krishnamurthy et al., 2022).
- Ellis's Model: it focuses on the various stages and tactics associated with the information-seeking process, such as initiation, exploration, and extraction stages, among others (Lopatovska & Sessions, 2016).
- Kuhlthau's Information Search Model focuses on the process of information-seeking with six distinct stages that concentrate on the cognitive and emotional states throughout the process (Oza & Patel, 2021). These include "task initiation, topic selection, prefocus exploration, focus formulation, information collection, and search closure" (Savolainen, 2015, p. 181). During these stages, information seekers carry out the following tasks: recognize, identify, investigate, formulate, gather, and complete (Allam et al., 2019).
- Belkin's Episodic and Anomalous Models: the first model emphasizes interactions with information, while the latter is focused on the realization of a gap in existing knowledge by the researcher (Oza & Patel, 2021).
- Information Foraging Theory: developed by Card, Chi, and Pirolli, represents the user as a forager who needs to adapt his strategies to optimize the intake of valuable information (Oza & Patel, 2021).
- Chatman's Life in the Round: contends that unless an initial problem arises, there is no point in initiating the information-seeking process (Oza & Patel, 2021).
- Dervin's Sense-Making Model: it poses that the information seeker makes sense of uncertain situations and interprets the information used for information-related decisions (Oza & Patel, 2021).

- Principle of Least Effort: it states that people prioritize the most convenient path to acceptable information (Oza & Patel, 2021).

Information-seeking strategies

In the specialized literature, it is possible to find many information-seeking strategies. For instance, the information triangle categorizes strategies into three types: 1) formal system strategies, related to the use of formal resources like databases and online scholarly sources; 2) informal resource strategies, which pertain to browsing and citation tracing, typically referencing print sources; and 3) interactive human strategies, which refer to consulting with knowledgeable humans, such as librarians or colleagues, directly or through electronic means (Savolainen, 2016). We will center mostly on the first type.

A mapping of information-seeking tactics across different academic tasks includes tactics like identifying and selecting resources, searching for specific facts or keywords, tracking sources, and verifying the accuracy of information (Lopatovska & Sessions, 2016). Savolainen (2016) summarizes some of the strategies as: 1) footnote chasing, citation searching and journal run; 2) many still use diverse information-seeking strategies beyond surfing the web, combining formal system strategies with resources like OPACs, and with informal resource strategies and human strategies; 3) planned-situational models emphasize information retrieval tactics to identify diverse ways in which information seeking occurs as a combinatorial process spanning methods, modes, and resources. Moreover, in an academic search context, there may be four different information-seeking scenarios: “fact verification, knowledge enhancement, knowledge acquisition, and knowledge discovery” (Hoeber et al., 2019, p. 231).

Some advice that can be provided when searching for a topic could include: 1) selecting a topic of interest; 2) focus or narrow the topic to retrieve a good and manageable number of results; 3) if the topic is too general, a search will yield a large number of results, but if it is too specific, then results will be fewer; 4) read some of the descriptive fields of the results (title, abstract, keywords) or even some full texts if you need to narrow and reduce results, as reading will help to better describe the topic and identify main concepts, theories, methods, models and previous related research, which will be useful to develop a list keywords for improving the search query; 5) keep in mind that document retrieval is based on using the right search terms; 6) it is recommended to use specific text strings, not generic words; 7) when unfamiliar with the topics, refer to thesauri or the keywords used in the documents that have been already retrieved and that represent are what I am really looking for.

Good topics are intriguing and encourage working with them, they lead to new or novel knowledge, allow supporting or elaborating theories and solve problems, generate research questions, and conducting research on them must be feasible (Gómez, 2009). When developing a search query, which implies selecting the words we are going to put in the search box of an information system and the relationships among these terms (expressed with the aid of Boolean logic, see below), we can ask ourselves the following questions: Why did you choose this topic? What interests you? Do you have an opinion on the subject? These allow for choosing the first search terms. Who has produced information on the subject? Who is affected by it? Do you know any institutions involved? These allow determining which kind of information system, document types, resources, and complementary sources of information I might need to seek. What questions can we ask about the topic? Are there any important debates? Are there different trends? What is the scope (local, national, international)? Is it a current or historical issue? Do you want to compare the issue at different times? These allow for narrowing results by publication date and discipline.

Boolean logic

In mathematics and mathematical logic, Boolean algebra or Boolean logic represents a distinct subdivision of algebra. Its divergence from basic algebra is noted in two primary aspects. Firstly, in Boolean algebra, variable values are confined to true and false truth values, often symbolized by 1 and 0, unlike in basic algebra, where variables assume numerical values. Secondly, it operates with logical operators such as conjunction (and) symbolized by ∧, disjunction (or) by ∨, and negation (not) by ¬, as opposed to the arithmetic operators like addition, multiplication, subtraction, and division used in basic algebra. Thus, Boolean algebra provides a structured method for articulating logical operations, paralleling how basic algebra articulates operations with numbers. George Boole first introduced the concept of Boolean algebra in his pioneering work, The Mathematical Analysis of Logic, in 1847, and was further elaborated in An Investigation of the Laws of Thought in 1854.

Boolean logic has different applications in many fields, such as mathematics, engineering, and library and information science. The latter is what interests us now. In information retrieval and digital libraries, Boolean logic is fundamental for constructing precise search queries or strategies to *pre-filter* and retrieve more specific documents from databases or search engines. Catalogers may be somewhat familiar with building these types of *a priory* relationships and using artificial languages to describe documents when assigning subject headings to them in a *pre-coordinated* manner. This is also applied to information-seeking. Boolean logic represents, in symbolic form, relationships between entities. Boolean operators must usually be written in capital letters.

When used to build search strategies, we use Boolean logic to establish *a priori* relationships among the words we use to build our search strategy. Deciding which words to use and defining the relationships that these words should present in the documents I need will determine the results we will get. Hence, Boolean logic is the cornerstone of information search and there are three basic Boolean operators (AND, OR, NOT). The Boolean operators and strategies we can use to prepare better and more precise search queries in almost any information system are the following:

- Use of quotation marks (“”): When searching for concepts integrated by two or more words (e.g., educational strategies, acetylsalicylic acid, or customer relationship management), we should put all these words within quotation marks to prevent the information system from searching for any other combination of such words. For instance, if I include the term “educational strategies” in my query, the system will offer as results the documents that contain such a specific combination of these terms and in this specific order, thus avoiding others such as educational technologies or military strategies. Otherwise, if we put the words educational strategies one after the other in a search engine without quotation marks, the information system will search at least for three sets of things: 1) documents with the word strategies; 2) documents with the word educational; and 3) and documents with the two words.
- Truncation (often expressed with an asterisk (*) at the end of the word): truncation allows searching for various words with the same root in common. Those are words that begin the same. For example, if we search for *educa**, the system will bring as results the documents

that have the words education, educational, and educative, without needing to type all these terms in the search box.

- Boolean operator AND: this operator is used for searching for two or more terms that must appear in the results, regardless of how far they appear from one another in the documents. We use it to tell the system to search for several terms that must appear in the results.
- Boolean operator OR: this operator will yield a larger number of results, because we are *telling the system* that we are interested in several terms, but not necessarily all of them must be together in the results, meaning that the system will retrieve all possible combinations of the terms specified in the search. This is useful when including optional terms or synonyms in our query. This operator is widely used for compiling results for a systematic review, as all possible synonyms must be incorporated in the search query. For example, LitCovid was set to monitor the publications on COVID-19, and the search string with which this resource has been updated is: ("coronavirus" OR "ncov" OR "cov" OR "2019-nCoV" OR "COVID-19" OR "SARS-CoV-2") (Chen et al., 2024). Here, it is possible to see that many different synonyms are incorporated in case some researchers did not use the most common one (COVID-19).
- Boolean operator NOT: when searching, we might have varying levels of clarity about what we are interested in finding. However, sometimes it seems that we are clearer about what we are not interested in. If this is the case, and we can identify some key terms that represent what we wish to avoid, we use the NOT operator, as it allows excluding results containing such words that we are not interested in. Please note that including the NOT operator in some systems should be expressed as AND NOT.
- Proximity operators (near, W/n, Pre/n) are more sophisticated and available in some systems, such as the bibliometric indices, which are used to define how close two terms should appear to one another.

After using a first search query, it is useful to examine some of the results to identify keywords that might be useful to incorporate in our query with any of the Boolean operators. The AND operator must be used to narrow the search if the word is very important. Conversely, if we need to broaden the search or include an important synonym (as well as include both plural and singular forms, if pertinent), we use the OR operator. Finally, if the word refers to something we are not interested in, then it must be used with the NOT operator. An example of a search query combining all three main Boolean operators would be the following:

> "young patients" AND (COVID-19 OR coronavirus) AND (cardiovascular AND treatment AND immunity) OR test OR vaccine NOT elderly

Almost any information system allows building search queries using at least the three main Boolean operators, even Google and Google Scholar. However, in these cases, please note that Google represents the AND with a blank space between words, the OR is supported as is, while the NOT must be represented by a minus or hyphen sign (-). The following is the same search query as the example above but adapted for either Google or Google Scholar (note that ANDs are replaced by blank spaces and a hyphen sign replaces the NOT):

"young patients" (COVID-19 cardiovascular treatment immunity coronavirus) OR test OR vaccine -elderly

Additionally, in Google and Google Scholar, it is possible to use the following commands in a search box: 1) filetype will allow searching for PDF files, spreadsheets, slides or documents, among others, by using this command, together with the file type extension (e.g., filetype:pdf); and 2) site will allow narrowing the search to a given type of web domain, such as .com, .gov, .edu., or .org (e.g., site:.edu). Google has other features, such as limiting results by language, country, and last update.

Search by field

Building stronger search queries will require working with the different metadata fields of the documents available in an information system. The main ones are the title, abstract, and keywords. In some systems, we might find them separate, but they are usually combined in one option that searches for any string in these three fields at the same time. Searching in these fields is a good way to narrow or get smaller numbers of results. Conversely, some systems will allow us to search within the full text of the documents, but this will be useful if we get smaller numbers of results, as it might drastically increase them. Other systems will allow searching by other fields, and we can obviously search by different fields at the same time and combine several Boolean operators in the same search query. Bibliometric indices will not allow full-text searching as they do not store them, but they offer the possibility of searching in many different fields. For instance, Elsevier's (2024) Scopus allows very sophisticated queries, as it is possible to search by the following fields: all fields; article title, abstract, keywords; authors; first author; source title (e.g., journal title or book title); article title; abstract; keywords; affiliation; affiliation name; affiliation city; affiliation country; funding information; funding sponsor; funding acronym; funding number; language; ISSN; CODEN; DOI; references; conference; article title, abstract, keywords, authors; chemical name; CAS number; and ORCID. The advanced search feature may require users to input abbreviations instead of the full label of the field in order to have the system search within a specific field (e.g., TI = title). An example of a search query in Scopus may be the following:

(TITLE-ABS-KEY (literacy) OR TITLE-ABS-KEY (literacies) AND TITLE-ABS-KEY (information) OR

TITLE-ABS-KEY (librar*) AND AFFILCOUNTRY (Mexico))

The above search intends to comprehensively find results about various literacies (information literacy, digital literacy, media literacy, new literacies, among many others) within the realms of either information in general or libraries (note the truncation in librar*, which will search for various terms at the same time (library, libraries, librarians, librarianship), and the results must have been published by authors within a Mexican institution.

Working with search results

When you search, you will probably get thousands or hundreds of thousands of results, so it is important to focus or narrow the topic in some way, either by geographic characteristics, by publication date, discipline, sub-discipline or by adding more words to our search query or strategy. In addition, any good information system will provide options for either broadening or for filtering and narrowing results.

Broadening results.

If the number of results is too small, there are several things that may be occurring: 1) the search query has errors; 2) the topic is too new; and 3) the search query is too specific and should be broadened. For this, it is useful to find out what are the most accepted technical terms (and relevant synonyms, if any) that describe the topic where are searching for and incorporate them into our search query. Supposing that we get a single document as the result of our search, it is possible to do two things: 1) check the technical words that are contained in this document's title, abstract, and keywords. Then, use these terms to improve the search query, which might provide more results; and 2) the sole document that was retrieved probably includes a reference section; we can check these references as some might be useful. Additionally, if the information system includes the references of a given document, as well as the list of other works that have cited it (such as Dimensions and the bibliographic indices), we can discover other documents that might be useful by checking both lists.

Other strategies to broaden results may include: 1) checking if the topic is too specific; 2) including or combining several elements that may have parallels or influence each other; 3) if the topic is too new, then incorporate complementary sources such as reports, statistics, news, websites, and newspapers; 4) perhaps the information system chosen does not have a large enough dataset, or is not specialized in the disciplines under which the topic is studied, then we can incorporate other databases, open access journals, and repositories, as well as Google and Google Scholar; 5) check the words used in the query, perhaps they are not the most commonly used by researchers, or they may not be the formal or specialized terms (colloquial terms should usually be avoided), we can also try incorporating synonyms.

Filtering and narrowing results.

If there are too many results, or they do not seem to be precisely what I am looking for, then the topic must be further specified, and hence we should improve the search query to narrow down its results; sometimes fewer results may mean that they are more relevant to the specific topic. We can ask ourselves the following questions to improve the search query: what do I already know about the topic? What is it that I do not know and wish to know? Are there various theoretical traditions? In the social sciences and humanities, we can identify if there are various schools of thought and limit results to one or a few of them. Are there any particular aspects of the topic I wish to know about? Consider only one or a few elements or problems related to the topic. Is there any specific period of time I wish to cover? Some research fields will prefer to work almost exclusively with sources published in the past five or ten years. Is there any particular population or sample of interest? Including some demographic variables (age, sex, educational or economic level, profession, geographic location) in the search query might help. Additionally, the following criteria can be used to filter and narrow the number of results of a given search query, particularly if such number is too high:

- Using filters: most information systems will include filters to narrow the search and hence reduce the number of results. For example, in any results page, Scopus allows filtering by: publication date; author name; subject area; document type; source title; keyword; affiliation; country; source type; language; publication stage; funding sponsor; and open access.
- Evaluate the sources you find. Among the most important evaluation criteria is authorship, i.e., that you can clearly identify who is the author of the document, who is responsible for the information, does he/she has academic credentials, and whether he/she is affiliated

with a serious or reputable institution. An evaluation of sources will center on the following criteria: authorship, obsolescence, reliability, depth, objectivity and balance, truthfulness, specialization, usefulness to answer the specific information need, relevance, use, and citation of sources. Additionally, online content evaluation may incorporate assessing: style, design and writing, alteration of data to attract attention (*clickbait*), domain and URL, title, social activity (comments, likes), and the about page.

- The currency or obsolescence of the sources. Depending on the discipline, limiting working with sources published during the past three, five, or ten years might be appropriate. This time period may be shorter in the case of technology, medicine, and natural sciences. Not all sources in a bibliography must be current, but it is recommended that most of them are.
- The reliability criterion concerns the authorship and the institution or organization that published a given document. Dubious websites are fairly easy to detect, and we should deter users from employing them as information sources. It could also be good advice to avoid using Wikipedia as the sole source of information for academic work. However, it might be a good starting point for both its entries and the references section, , and it might be worthwhile to check the community discussions behind an entry to understand its development, as well as for analyzing the disagreements and agreements about its contents.
- The depth and specialization of the sources will depend on the type of academic or scientific work we wish to do and the individual's own academic level. An elementary school paper is not the same as a Ph.D. thesis; the sources employed should match the level of the resulting work.
- Further evaluation of sources: these may include assessing a given source's objectivity, balance, and truthfulness if the paper cites sources, if its language is not biased, and if it presents diverse perspectives.
- When using websites as information sources, pay attention to their design and writing style and check the social activity of the page (e.g., are there favorable comments? Are they critical? Do they present arguments that make you doubt the integrity of the website's contents?). If the website has an 'about' page, this could allow for assessing its authority.
- Ordering results: some systems allow ordering results by author, title, publication date, or relevance. The choice will depend on the user, but some care should be taken when ordering by relevance, as there will be some algorithmic implications behind how a specific system sorts a group of results under this criterion.

Future perspectives

Although the use of these strategies will remain important in the near future, the information industry has been moving toward expanding and improving the use of natural language search queries. Such has been the case with Google since its inception and we are witnessing that artificial intelligence (AI) solutions are based on natural language queries. Regarding the retrieval of scientific information, some interesting developments have been made by Scite.ai (2024) and Consensus (2024), which are AI systems that allow asking questions and do not provide a specific answer. However, their results pages consist of published documents that may contain the answer. These systems highlight extracts that may contain the answer and mark those documents that are systematic reviews or are highly cited. As of 2024, more features are being developed and implemented in these systems. The future of information-seeking behavior can move in the following directions, which are relevant for information services and studies:

- There is the need to develop more dynamic and user-engaging approaches to information-seeking in a rapidly changing environment (Allam et al., 2019)
- The complexities of search tasks increase as we progress from students to professionals and as the nature of the task changes from a fact verification task to a knowledge enhancement one. This calls for improvement in search tools to cater to different needs (Hoeber et al., 2019).
- There is a need for a deeper understanding of how information is processed and used after it is acquired. Such insights can help various groups ranging from students, educators, librarians, and system designers enhance their practices and tools (Lopatovska & Sessions, 2016).
- Information theorists may arrive at a unified model of information-seeking behavior, which could incorporate different viewpoints, concepts, and terminology, providing a holistic understanding of user behavior (Agarwal, 2022).
- Authors highlight some opportunities for improving, evaluating, and modeling information-seeking behaviors: 1) Assessing the effectiveness and applicability of various models of information-seeking (Allam et al., 2019; Berget et al., 2021; Agarwal, 2022); 2) exploring and enhancing information-seeking behaviors among specific communities, such as researchers, graduate students, and users with impairments (Hoeber et al., 2019; Berget et al., 2021); 3) Analyzing the role of assistive technologies in information-seeking (Berget et al., 2021); 4) Focusing on individual capabilities and limitations that might affect information searching (Berget et al., 2021); and 5) Integrating existing models and frameworks for a more comprehensive understanding of information-seeking behavior (Agarwal, 2022).

Conclusion

Information-seeking implies some trial and error; hence, arriving at the ideal search query may require several tries until finding the most appropriate combination of words that yields the most useful results for the specific information need. It is also important to stress that we should not be satisfied with only seeing the first page of the search results.

Navigating the landscape of information-seeking strategies requires an understanding of both traditional and emerging models and a grasp of the practical applications of Boolean logic, filtering, and evaluation techniques. Librarians, researchers, and information seekers must adapt to a rapidly evolving information ecosystem, where the proliferation of digital platforms and the increasing sophistication of search engines demand a refined set of skills. The future of information-seeking behavior appears to be intertwined with advancements in AI, suggesting a shift towards more intuitive and conversational search queries. However, the core principles of diligent, strategic information searching—characterized by critical evaluation of sources, effective use of search operators, and adaptability to changing information landscapes—remain indispensable. As we move forward, it is essential to continue refining our understanding of these dynamics, ensuring that information professionals and researchers are equipped to navigate the complexities of the digital age effectively.

References

Agarwal, N. K. (2022). Integrating models and integrated models: Towards a unified model of information seeking behaviour. *Information Research, 27*(1). https://doi.org/10.47989/irpaper922

Aguillo, I. F. [@isidroaguillo]. (11 april 2023). *Algunas de las principales fuentes bibliométricas y su cobertura (millones de registros) actual (datos propios)* [Tweet]. Twitter (X). https://twitter.com/isidroaguillo/status/1645701614781464576

Allam, H., Bliemel, M., Nassiri, N., Toze, S., Peet, L. M., & Banerjee, R. (2019). A review of models of information seeking behavior. In *2019 Sixth HCT Information Technology Trends* (pp. 147–153). IEEE. https://doi.org/10.1109/ITT48889.2019.9075095

American Psychological Association. (2020). *Publication manual of the American Psychological Association: The official guide to APA style* (7th ed.). American Psychological Association.

Berget, G., MacFarlane, A., & Pharo, N. (2021). Modelling the information seeking and searching behaviour of users with impairments: Are existing models applicable? *Journal of Documentation, 77*(2), 381-400. https://doi.org/10.1108/JD-04-2020-0049

Chen, Q., Allot, A. and Lu, Z. (2024). LitCovid. https://www.ncbi.nlm.nih.gov/research/coronavirus

Clarivate Analytics. (2024). Web of Science. https://www.webofknowledge.com

Consensus. (2024). Consensus: AI search engine for research. https://consensus.app/search

Dimensions. (2024). https://app.dimensions.ai/discover/publication

Elsevier. (2024). Scopus. https://www.scopus.com

Gómez, M. (2009). Introducción a la metodología de la investigación científica (2a. ed.). Córdoba: Editorial Brujas

Hoeber, O., Patel, D., & Storie, D. (2019). A study of academic search scenarios and information seeking behaviour. In *Proceedings of the 2019 Conference on Human Information Interaction and Retrieval* (pp. 231-235). ACM. https://doi.org/10.1145/3295750.3298943

Krishnamurthy, M., Asundi, A. Y., & Subhash, R. B. (2022). Hidden Concepts of Library and Information Science in Information Seeking Behavior Models. *International Information & Library Review, 54*(3), 266-273. https://doi.org/10.1080/10572317.2021.2022389

Lopatovska, I., & Sessions, D. (2016). Understanding academic reading in the context of information-seeking. *Library Review, 65*(8–9), 502–518. https://doi.org/10.1108/LR-03-2016-0026

Oza, N. D., & Patel, P. Y. (2021). An Analytical study of Information seeking behaviour models. *Library Philosophy and Practice, 2021*, 6246. https://digitalcommons.unl.edu/libphilprac/6246

Savolainen, R. (2015). The interplay of affective and cognitive factors in information seeking and use: Comparing Kuhlthau's and Nahl's models. *Journal of Documentation, 71*(1), 175-197. https://doi.org/10.1108/JD-10-2013-0134

Savolainen, R. (2016). Information seeking and searching strategies as plans and patterns of action: A conceptual analysis. Journal of Documentation, 72(6), 1154-1180. http://doi.org/10.1108/JD-03-2016-0033

Scite. (2024). https://scite.ai

Tarango, J., y Machin-Mastromatteo, J.D. (2017). *The role of information professionals in the knowledge economy: Skills, profile and a model for supporting scientific production and communication*. Oxford: Elsevier. http://www.sciencedirect.com/science/book/9780128112229

See also

20003. Obsolescence

20009. Books

20010. Serials

20023. Scholarly Communication, an overview

20024. Peer Review

20045. Sources of Information

20046. Repositories

20047. Open Access

20051. Discovery Systems

50029. Current Trends in Information Behavior Research: Expanding Beyond Search, Seeking, Finding and Behavior