Metasearchers in the biomedical field
An update

RODOLFO RODRÍGUEZ-GÓMEZ • BOGOTÁ, D.C. (COLOMBIA)

DOI: https://doi.org/10.36104/amc.2023.2587

Abstract

The biomedical field has benefited greatly from the influx of information and communication technologies. Over the last few years, enormous strides have been made in organizing, structuring and seeking information, facilitating this task, as evidenced by the various databases and search engines. In this context, metasearchers are a powerful tool for optimizing the effectiveness of searches for scientific evidence and, therefore, represent an added value both in clinical practice as well as in the research setting. This article is aimed at healthcare workers. The purpose is to update the basic concepts about metasearchers, their generalities and the main characteristics of the most outstanding ones in the biomedical field. **(Acta Med Colomb 2022; 48. DOI: https://doi.org/10.36104/amc.2023.2587).**

Keywords: *search engine, database, internet, evidence-based medicine, medical informatics. (Source: DeCS).*

Dr. Rodolfo Rodríguez-Gómez: Médico epidemiólogo, Magíster en Salud Pública, Magíster en Bioética. Asesor Metodológico de Investigación en Salud. Bogotá, D.C. (Colombia).
Correspondencia: Rodolfo Rodríguez-Gómez. Bogotá, D.C. (Colombia)
E-Mail: fitopolux@hotmail.com
Received: 23/II/2022 Accepted: 12/X/2022

Introduction

Today, literature search skills are very important for healthcare students and professionals. In this context, mastery of the basic knowledge of metasearchers is an added value to better approach the body of evidence and, thus, optimize clinical practice. Within the subject of search engines, a special section belongs to metasearchers, which represent a higher level in the search for scientific evidence and, without a doubt, are an invaluable computer resource for extensive literature reviews or systematic reviews. This article is aimed at healthcare staff and intends to provide basic but updated information on metasearchers. We provide a brief description of what a metasearcher is, the first metasearchers and generalities about these data systems and their main characteristics. Details are provided on the important metasearchers in the biomedical field and, finally, their main advantages and disadvantages are mentioned.

What is a metasearcher?

A metasearcher is a data system that can carry out simultaneous searches in different search engines, but not directly in the web pages. Therefore, this type of searcher conducts *searches in searchers*. Metasearchers do not have their own databases, but they can consolidate, in a single list, the results they obtain from the search engines they consult. It is worth noting that metasearchers perform very well in retrieving document collections, which is why they are more effective than search engines (1). It is also relevant to consider that they eliminate duplicates (2) and optimize the efficacy of the searches by automating the search process simultaneously in several search engines, which decreases the workload for the user. In addition, results are presented in different ways, as some show them by categories, others classify them by relevance, and some separate them according to the search engine consulted (3).

Metasearchers basically use two search strategies; some use search engines and gather the information, while others like, for example, Metta, search for information in specific databases. There are federated metasearchers, such as SUMSearch and TRIP, which could be employed by users who do not know which database would be most appropriate for the research question. The metasearcher concept has a similar meaning to *federated searcher*; in fact, some texts use them as synonyms. The term *federated* may cause some confusion. Basically, it refers to a search which consults information in several sources or resources, which are generally independent systems like library catalogs, internet resources or a corporate intranet (4). Thus, a federated search could be defined as one which allows different web sources to be consulted simultaneously, in order to present grouped results to the user (5).

The first metasearcher came into existence in 1994 and was named MetaCrawler. This data system was derived from Eric Selber and Oren Etzioni's master´s thesis (6) at the University of Washington (7) and initially allowed access to search systems like WebCrawler, Lycos, Excite, Infoseek, AltaVista, Inktomi, Go To, LookSmart, Thunderstone and Yahoo (8). MetaCrawler organized the results in a standard format and showed results according to the relevance of the link and the type of domain. Over time, MetaCrawler became famous and was relaunched in 2017. Currently, its website is active with a minimalist interface which allows searches of the web, news, images and videos. This resource can be accessed through https://www.metacrawler.com/.

SavvySearch was launched in March 1995, developed by Daniel Drillinger as a class project at the University of Colorado (6). This data system learned to identify the appropriate search engines for certain questions and conducted parallel searches. SavvySearch was designed to maximize the probability of returning appropriate links and minimizing the use of computational resources. There were several search engines at that time, from general ones like AltaVista or WebCrawler, to more specific ones like FTPSearch. After entering a term or keyword, SavvySearch would run simultaneous searches in Excite, AltaVista, Infoseek, OpenText, Lycos, WebCrawler, HotBot and Deja News, retrieving the information and presenting it in a homogenous form (8). Then, SavvySearch would conduct traceability of the search engines to determine which were the most appropriate for the searches and allowed key terms or words to be included which could be combined with logical operators (9).

In the 2000s, other metasearchers besides MetaCrawler and SavvySearch were developed, like Dogpile, Megasearch and Pro Fusion. Most metasearchers worked in the same way but differed in how to present the results to the user. Some, like ProFusion and MetaCrawler, combined the results using data fusion techniques based on scores assigned to the documents and reported broken (unavailable) links, which lengthened the consultation times (10). From its beginning, the internet has grown constantly and continued to expand like the universe itself. This makes the amount of information immeasurable, and both the retrieval of all the information as well as the review of such a large number of results becomes a complex and cumbersome task. At the beginning of the 2000s, it was proposed that metasearchers use the cluster technique to better structure the search results. This concept refers to classifying or grouping data with homogenous topical characteristics (11), which reduces the volume of information and facilitates the construction of categories or taxonomies which are arranged hierarchically.

Metasearchers in the biomedical field

TRIP Database

This metasearcher has been online since 1997 and was created at the University of Wales (12). Its acronym "TRIP" stands for Turning Research Into Practice, and this metasearcher was created with medical practice in mind. The Trip Database conducts simultaneous searches in recognized data sources like the Database of Abstracts of Reviews of Effects (DARE) or the Cochrane Library. Currently, in the context of evidence-based medicine (EBM), Trip is one of the most well-known metasearchers, allowing search terms to be tracked in dozens of databases, although it can also access image and electronic book banks. Therefore, TRIP is considered to be an excellent tool not only for clinicians, but also for anyone conducting a systematic review.

According to the Trip Database website, its slogan is "Find evidence fast." In practice, it really is a speedy resource which not only retrieves textual clinical information, but also other content like videos, images, courses, pamphlets and news. Based on the user's needs, Trip Database offers a system of filters aimed at facilitating searches (for example, filtering by year of publication). According to the Trip website, the algorithms take three factors into account when deciding how to position the search results. First is the text score; for example, if a term is in the title, its score will be higher than if it is found in other parts of the document. The other factor is the publication's score, which considers the quality of the source; thus, resources like Cochrane represent a high-quality source. Finally, the date is considered, with more recent documents having a higher score.

Within the search options, Trip offers four main strategies. The first is to use the search box, which functions like any other searcher when keywords are entered. A second option is to use the PICO strategy, which refers to *patient, intervention, comparison* and *outcomes*. A third option is an advanced search, which is only available for PRO users. A final option is to search beginning with the most recent information; however, as with the advanced search, this is only available for PRO users. The last two options appear on the start-up screen along with a small label indicating this restriction. An interesting Trip tool is called "Evidence Maps," with which the results can be viewed as a graph, using circles to show different interventions for a given situation or disease. When you click on one of these circles, a cartesian graph is displayed in which each circle represents a certain article. When the cursor hovers over each circle, the name of the article appears; the size of the circle indicates the sample size, while the color indicates the risk of bias: light green represents a lower risk of bias. TRIP offers free access and may be found at https://www.tripdatabase.com/.

SUMSearch

This free metasearcher was initially called Medical SmartSearch. It began in October 1998 and was developed by Robert Badgett (14), an internist and professor at the University of Texas at San Antonio (USA) (14). Beginning in August 2010, SUMSearch became SUMSearch 2, which is faster, validates search strategies and provides bibliometric markers of the quality of the articles. Linked to the University of Kansas School of Medicine, SUMSearch 2 can carry out simultaneous searches of original studies, systematic reviews and clinical practice guidelines on PubMed. It also allows searches of high-impact medical journals in the National Library of Medicine, DARE, and the National Guidelines Clearinghouse (15). Currently, the SUMSearch 2 interface is simple and includes filters like age, adults, intervention, diagnosis, and humans, among others. It also allows the use of the Boolean operator AND, and its controlled language is MeSH. It can be found at http://sumsearch.org/.

Epistemonikos

Epistemonikos is a multilingual health field evidence metasearcher founded by Gabriel Rada and Daniel Pérez. It currently operates with the support of various institutions, most of which are non-profits, as well as support from Epistemonikos Foundation, with headquarters in Santiago de Chile. It uses color coding in the results and a filter system. Epistemonikos relies on a broad collaborative work between humans and robots to consult frequently updated databases like PubMed, EMBASE, LILACS, DARE, the Campbell Collaboration and CINAHL, among others. Epistemonikos groups articles according to three categories: 1) broad synthesis, in which it groups different types of articles in order to create systematic review syntheses; 2) syntheses of primary studies, mainly systematic reviews which meet certain criteria; and 3) primary studies. An additional category is structured summaries, which are article summaries for non-researchers. After conducting a search, Epistemonikos shows the results, grouped by category (broad syntheses, systematic reviews, structured summaries, primary studies), on the left side of the interface. It also presents a filter by year or personalized range and an automatic translation option. Epistemonikos can be found at https://www.epistemonikos.org/es/.

PediaClic

PediaClic is a recent metasearcher for child and youth health resources. It consists of 10 specific searchers and a general searcher which consults various sources and retrieves information which it classifies in categories such as clinical practice guidelines, evidence-based summaries, clinical questions, health blogs, online pediatric texts, and information for families, among others. This resource is mainly aimed at healthcare professionals in Latin America and Spain who are involved in child and adolescent health care, but also has content for families seeking information on these topics. Several scientific associations are associated with PediaClic, such as Asociación Española de Pediatría de Atención Primaria (AEPap) [Spanish Association of Primary Care Pediatrics], or Confederación Nacional de Pediatría de México (CONAPEME) [Mexican National Confederation of Pediatrics], as well as blogs like Biblioteca Médica Virtual [Virtual Medical Library] and Biblioteca Virtual de Ciencias de la Salud [Virtual Healthcare Sciences Library]. Searches on PediaClic are simple, as it uses the autofill tool; it also allows the use of Boolean operators AND and OR, and the use of special commands like *allintext* or *allintitle* to find a term in any part of the contents or title, respectively. PediaClic's specific searchers conduct additional searches in multiple sites including the Federación Española de Fibrosis Quística [Spanish Cystic Fibrosis Federation] or family websites like ABC Pediatría and multiple blogs. PediaClic can be found at http://www.pediaclic.org/.

Exploraevidencia

This is a free access metasearcher created by Antonio José Morales at Escuela Andaluza de Salud Pública. Exploraevidencia has three important links: *tengo prisa* (I am in a hurry), *buscar evidencias* (search for evidence), and *información y salud* (information and health). The first, *tengo prisa,* provides access to a page with direct access to various abstract, clinical practice guideline, review, and database search resources, as well as other metasearchers like TRIPDatabase and Epistemonikos. The *información y salud* link provides different patient information links such as Cochrane Net Consumers, Cancer.gov, Escuela de Pacientes [Patient School] and Family Doctor, among others. It can be found at http://www.easp.es/exploraevidencia/.

ACCESSSS

ACCESSSS is a computer resource which conducts simultaneous searches in several evidence-based data services. The search results in this resource generate content that is organized hierarchically, with the body of evidence in all pyramid levels evaluated according to its scientific merit and clinical relevance. This hierarchy is based on what is known as the pyramid 5.0, covering, from the base, primary studies, systematic reviews, systematic guidelines, abstracts and, at the top, systems. On the website home page, ACCESSSS offers a view of the most read articles in all disciplines over the last 30 days, each scored with a star system, with a maximum score of seven stars. An important point is that you must register to conduct searches in this resource. ACCESSSS can be found at https://www.accessss.org/.

METTA

Metta is a metasearcher designed to retrieve biomedical literature and is conceived for those who conduct systematic reviews. Metta is a federated searcher which connects to five databases: PubMed, CINAHL, EMBASE, PsycINFO and Cochrane Central Register of Controlled Trials (16). This metasearcher is not available to the general public, but offers a different approach which can save time for those who conduct systematic reviews, meta-analyses or literature searches on a specific topic (16).

Advantages and disadvantages of metasearchers

One of the advantages of these computer-based systems is that they provide a general idea of the information available on a specific topic. In addition, they are very fast, and they increase the chances of finding relevant information. One element on which the developers have worked and which the users enjoy is the clean and friendly interface several of them have, which also affects the users' experience with a given information system. This deserves special mention, since, after evaluating different metasearchers, we have concluded that several do not have the best performance, as the developers' main concern has been, precisely, the interface (17). Another metasearcher advantage to highlight is the grouping or clustering of results, which is useful for people who do not know much about a specific area and are

not familiar with the keywords (18). On the other hand, the disadvantages of metasearchers include having fewer search options than some of the search engines and the user being limited to the way in which a given metasearcher performs and configures the searches (19).

Discussion

Since 1994, when MetaCrawler (the first metasearcher) appeared, these computer systems have evolved significantly and have in one way or another facilitated the users' work. It is clear that not all metasearchers report the same information, due to the algorithms they use. Therefore, since each metasearcher operates in its own way, different metasearchers may report different results. Thus, we recommend using various metasearchers for complex searches (for example, for systematic reviews), which may result in retrieving different information. Clearly, metasearchers have advantages and disadvantages and, like other information resources, are also subject to evaluation, a process which considers different aspects such as functionality, the interface, and the presentation of results. Over the last few years, the architecture of this type of information programs has improved ostensibly and, together with the sophistication of the algorithms, has optimized both the searches and the presentation of results, which ultimately provides great benefits to the final user.

Conclusions

Over the last few years, the contribution of Information and Communication Technologies (ICT) to the biomedical field has been momentous and has radically changed the way in which scientific knowledge is created, shared and incorporated. In the biomedical field, metasearchers are very helpful informatics tools for getting to the next level in any type of search and are highly recommended for retrieving information in the context of clinical practice and EBM. Despite certain differences in the search algorithms and interface encountered by the users, the metasearchers are an important literature search tool, especially when a large quantity of information is consulted, as is the case with reviews of the literature or systematic reviews. In summary, it all points to the fact that it is imperative for healthcare professionals to optimize their knowledge of scientific literature searching, in this case with metasearchers, since, today, these skills are a necessity rather than a luxury.

References

1. **Meng W, Yu C.** Advanced Metasearch Engine Technology. Morgan & Claypool; 2011.
2. **Amaya MA.** Evaluación de herramientas de recuperación de información electrónica. *Revista Venezolana de Información, Tecnología y Conocimient*o. 2017;**14(1)**:29-50.
3. **Oakes EH, Kia M**. Social Science Resources in the Electronic Age: World history. United States of America: Greenwood Publishing Group; 2004.
4. **Melucci M, Baeza-Yates R.** Advanced Topics in Information Retrieval. New York (USA): Springer Science & Business Media; 2011.
5. **de Smet E, Dhamdhere S.** E-Discovery Tools and Applications in Modern Libraries. United States of America: IGI Global; 2016.
6. **Kent A, Williams JG.** Encyclopedia of Microcomputers: Volume 28 (Supplement 7). United States of America: Marcel Dekker, Inc; 2002.
7. **Sharma V.** Text Book of Bioinformatic. Rastogi Publications; 2008.
8. **He J.** Internet Resources for Engineers. A practical handbook foe engineers and students. Australia: Reed International Books Australia Ltd; 1998.
9. **Howe A, Dreilinger D.** SavvySearch: a meta-search engine that learns wich search engines to query. *AI Magazine*. 1997;**18(2)**:19-25. doi: https://doi.org/10.1609/aimag.v18i2.1290
10. **Peters C**. Research and Advanced Technology for Digital Libraries: First European Conference, ECDL '97 Pisa, Italy, September 1-3, 1997 Proceedings. Vol. Volume 1. Germany: Springer; 1997.
11. **Adina L.** World Wide Web Metasearch Clustering Algorithm. *Revista Informática Economica*. 2008;**2(46):**5-11.
12. **Ferragud C, Vidal A, Bertomeu JR, Lucas R**. Documentación en Ciencias de la Salud. *Valencia: Nau Libres*; 2020.
13. **Crom V.** SUMSearch. *J Med Libr Asso*c. 2007;**95(4):**471-3. doi: 10.3163/1536-5050.95.4.471
14. **Ferneini EM, Goupil MT.** Evidence-Based Oral Surgery: A Clinical Guide for the General Dental Practitioner. United States of America: Springer International Publishing; 2019.
15. **Pencak M, Staffileno BA, Foreman MD.** Research for Advanced Practice Nurses: From Evidence to Practice. Third Edition. New York (USA): Springer Publishing Company, LLC; 2018.
16. **Smalheiser NR, Lin C, Jia L, Jiang Y, Cohen A, Yu C, et al.** Design and implementation of Metta, a metasearch engine for biomedical literature retrieval intended for systematic reviewers. *Health Inf Sci Syst*. 2014;**2(1).** doi: 10.1186/2047-2501-2-1
17. **Sastre-Suaréz S, Pastor-Ramon E.** Evaluación de metabuscadores gratuitos especializados en ciencias de la salud. El profesional de la información. 2011;**20(6):**6339-644. https://doi.org/10.3145/epi.2011.nov.06
18. **Sugumaran V.** Methodological Advancements in Intelligent Information Technologies: Evolutionary Trends: Evolutionary Trends. United States of America: IGI Blobal; 2010.
19. **Devanandan KV, Lakshmi R.** ICT embedded Education and Research: A Hand Guide for Teachers and Researchers. *Notion Press*; 2018.