



Search engines and software for manuscript editing

Yeonwook Kim

Guhmok Publishing Co., Seoul, Korea

Abstract

In recent years, manuscript editing has become extremely important for academic journals. Using appropriate software for manuscript editing results in improved work efficiency and increased accuracy; therefore, this article aimed to introduce search engines and software that can be used for manuscript editing. First, a variety of search engines and academic databases can be used to reduce errors and to create accurate references. Google, the world's leading search engine, provides users with information with the highest probability of accuracy, regardless of the reference language or the search term. If it is not possible to find certain information on Google, one can consult WorldCat, PubMed, Naver Academic, KoreaScience, Research Information Sharing Service, DBpia, Crossref, and Edifix. In particular, Naver Academic provides search results for some materials that cannot be found on Google. Second, PerfectIt facilitates the correction of errors that occur frequently in English-language documents. Finally, Grammarly is a helpful tool for checking and correcting grammar and spelling errors. As the academic publishing environment changes, the role and demands of manuscript editors are also changing. In a fast-paced environment, the software and search engines discussed herein are highly useful tools for manuscript editing.

Keywords

Language; PubMed; Publishing; Search engine; Software

Introduction

Many academic journals have their manuscripts edited by a manuscript editor (also known as a copy editor, technical editor, or managing editor) [1]. A manuscript editor is a person who edits manuscripts submitted by researchers for clarity, precision, readability, technical compliance, and structural accuracy [2]. In addition, a manuscript editor rewrites and revises the text, graphs, tables, and references according to the journal's publication guidelines. In recent years, it has become necessary to consider several new aspects of academic publications, such as authors' contributions and ORCID IDs, which have increased manuscript editors' workload.

Received: December 5, 2019
Accepted: January 2, 2020

Correspondence to Yeonwook Kim
yeonwook.kim@gmail.com

ORCID
Yeonwook Kim
<https://orcid.org/0000-0003-4631-9407>

This is an open access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Using appropriate software for these tasks can reduce the overall time required, while improving accuracy. Verifying the references section is the most time-consuming task, and the use of appropriate software can greatly increase the efficiency of this process. In this article, we review websites that can be helpful for finding references, and consider additional software that can be useful for manuscript editing.

How to Verify References

Ideally, the author of a manuscript should ensure the consistency of all the included references; however, in reality, inaccurate references are often included in accepted submissions. References that do not undergo verification may include a variety of errors. For example, when a reference is modified to

meet the formatting guidelines of the target journal, some of the information may be omitted or misspelled. In some cases, incorrect information may be reproduced when inaccurately written references are copied and pasted from other papers

Table 1. Essential characteristics of software used to search for references

Characteristics	Description
Amount of data	How much data (papers, books, etc.) does the database contain?
Data consistency	Is the software reliable and usable?
Price	Is the price affordable?
Required time	Can users save time using the software?
User interface	Is the software intuitive and easy to use?

Table 2. Software that can help with manuscript editing tasks

Software	Feature	URL	Payment	Whether it is included in Google search results ^{a)}
Google	Starting point for online searches Contains the largest amount of information	https://www.google.com	Free	Included
Google Scholar	Specialized search for academic information	https://scholar.google.com	Free	Separate search recommended
Google Books	Book contents with front matter	https://books.google.com	Free	Included
WorldCat	World's largest book database	https://www.worldcat.org	Free	Separate search recommended
PubMed	Well-managed biomedical database developed by the NLM	https://www.ncbi.nlm.nih.gov/pubmed	Free	Included
Naver Academic	Searches for some academic information not found in Google	https://academic.naver.com	Free	Separate search recommended
KoreaScience	Database that provides PDFs of all registered articles	http://www.koreascience.or.kr	Free	Included
RISS	Contains numerous Korean master's & PhD theses	https://www.riss.kr	Free	Included
DBpia	Commercial database that offers PDF preview	https://www.dbpia.co.kr	Free ^{b)}	Included
Crossref Link References	Searches the DOIs of a list of references	https://search.crossref.org/references	Free	NA
Crossref Simple Text Query	Searches the DOIs, PubMed IDs, and PMC IDs of a list of references	https://doi.crossref.org/simpleTextQuery	Free	NA
Edifix	Website that directly corrects references	https://edifix.com	Required	NA
PerfectIt	Checks for general errors in English-language documents Microsoft Office Word add-in	https://intelligentediting.com	Required	NA
Grammarly	Checks for spelling and grammar errors in English Microsoft Office Word add-in, Chrome extension, etc.	https://www.grammarly.com	Free ^{c)}	NA

NLM, US National Library of Medicine; NA, not applicable; PMC, PubMed Central.

^{a)}Resources that are not included in Google search results need to be searched separately; ^{b)}The first page of the article's PDF is provided free of charge, and paid members can access the full PDF; ^{c)}Basic features are free, but they have paid premium functions.

without examining the original article. Therefore, manuscript editors should carefully check the reference information, with the assumption that the references could contain errors.

Usually, manuscript editors use search engines and academic databases to validate references. The critical characteristics of these search sites are listed in Table 1. However, no single resource is completely satisfactory, and if one has a thorough understanding of the characteristics of each website and can use them efficiently, it is possible to work more quickly when verifying references.

The basic process of verifying references is to obtain the original PDF file and to check the information provided by the author through a direct comparison. In some cases, however, a website might not exist for a certain journal, and in other cases, articles cannot be directly downloaded because they are not free. In such cases, the consistency of references needs to be verified using information from various academic databases and other sources.

If the information that must be verified is registered on a journal's website or in a database, verification can be easily accomplished. However, if the information cannot be found, there are two primary possibilities: either it is wrong, or it might not be available online. Outdated reports or books, as well as those that do not have an official English title but are translated into English in the references section, are hard to find online. In such cases, it might be helpful to search for a certain term contained within the reference, rather than typing in the full details of the title, author, and year. If searching for a journal article by title or author seems difficult, another possible technique is to check the volume, issue, and pages directly after navigating to the archive on the journal's website.

If a reference cannot be found at all, it is best to ask the author for confirmation. In some cases, if the author did not directly check the original article, he or she may ask for the reference to be deleted. Moreover, in some cases, the author may find a reference and send it directly. If part of a reference is missing, it is necessary to include the missing part after checking with the author to confirm the exact part of the reference to be included.

The software presented in Table 2 is freely available or provides a free trial. All the information is based on data gathered in October 2019, and the features and prices of these tools are likely to change with time.

Google

Google is the most popular search engine globally. Irrespective of the language of the reference or the information used for searching, Google provides the most accurate reference information. Most of the websites mentioned in the later sections of the article are also included in Google's search results;

therefore, Google is the most preferable search engine for all references. Users can usually find most of the information they need by clicking on the top two or three Google search results. Google is capable of providing search results that are more accurate than the search function provided by various databases.

Google uses a web crawler called "Googlebot" to find and classify data by independently surfing the web. The policies of some sites prevent this form of web crawling, and some sites are not optimized for crawling by bots; the information contained on such sites is therefore not available on Google. If certain information cannot be found on Google, then the other websites listed in Table 2 can be used.

Google Scholar

Google Scholar is the world's largest site for freely accessing scholarly information. Unlike Google, Google Scholar offers features optimized for scholarly literature. When Google Scholar identifies a document based on the content of the journal's source site, PubMed, and various databases, it collects the data and outputs it as a single result. When it is difficult to identify accurate bibliographic information in one place, additional content might be found on other sites. References can be displayed using the Modern Language Association (MLA) and American Psychological Association (APA) styles and can also be downloaded as RIS and BibTeX files, which can help find additional information.

Even if Google Scholar is unable to find a source site, it generates bibliographic information if many articles cite a specific article (Fig. 1). For example, even if an article that was published in 1940 does not have its original text on the Internet, the information of the cited article is analyzed and provided as an output. This may be valuable information about hard-to-find data, but in some cases, non-existent literature can be generated as if it were real. A search result that is not directly linked to the original data can be used to check whether the information that the author gave is accurate by comparing the result from Google Scholar with the information provided by the author. In short, the basic principle is to find the original source and then to write an appropriate reference on that basis.



Fig. 1. Google Scholar results without a reference link to the original site.

Google Books

Google Books is a service where users can browse and search the full text of books provided by publishers. Some books offer the full text with the publisher's permission, while others have partially-released contents. To make a reference to a book, the publisher's name, city, and year of publication are required. If a chapter is cited, the chapter name and author are also required. In such cases, one cannot create a correct reference without directly checking the front matter, which includes the copyright and table of contents. Google Books allows users to see the cover, copyright, and table of contents of a book, which enables manuscript editors to check the original source and to correct references as needed. When users search for a book on Google, they often find results for online bookstores—such as Amazon and Kyobo Books, which is mainly used in Korea—and publishers' homepages—such as Elsevier, Springer, and Wiley—in addition to Google Books. Some of these sites also offer free front matter for certain books, which manuscript editors can use to verify references.

WorldCat

WorldCat provides data for over 2 billion books from 10,000 libraries worldwide [3]. If Google fails to retrieve the information needed for a book, it is recommended to search WorldCat for the book's name, year, and other information. WorldCat aims to provide more accurate information about books than Google, and it does contain more information about books than Google.

A noteworthy feature about WorldCat is that it provides both the publisher's name and location. However, since the books' front matter is not provided, it is difficult to identify chapter titles and authors. Further, as WorldCat does not store information on all the editions published in various regions, book information that is not verifiable in WorldCat should not automatically be considered to be inaccurate. However, WorldCat is still useful for finding missing information and checking erroneous parts of references.

PubMed

PubMed is a biomedical abstract database maintained by the US National Library of Medicine. It verifies all the data carefully before abstracts are uploaded to the database, so the accuracy and reliability of the data are quite high compared to other sites. All information, including changed journal names, errata, volume and issue numbers, page numbers, and DOIs, is well maintained. As PubMed consistently provides reliable information, if Google determines that PubMed has relevant data, it ranks PubMed higher than the original site. Authors' names and journal abbreviations are accurately maintained, which makes it easier to distinguish between first and last names and to quickly identify journal abbreviations.

Naver Academic

Naver Academic is an academic database created by Naver, which directly contracts with several globally renowned publishing organizations, such as Crossref, PubMed, British Library, and the IEEE [4]. As of January 2019, it stores more than 200 million articles, and some resources that cannot be found on Google can be found by searching here.

KoreaScience

KoreaScience is a free full-text database provided by the Korea Institute of Science and Technology Information (KISTI) that is used as a DOI landing page by many journals in Korea. Original PDFs are freely available for all the articles that it contains. However, it can be a problem if a journal's name has been altered or modified, because the journal's name would then be changed retroactively. Therefore, although it is cumbersome to do so, it is important to check the original PDF for accurate information when verifying a reference.

Research Information Sharing Service

The Research Information Sharing Service (RISS) is an academic database provided by Korea Education & Research Information Service that allows researchers to search academic resources produced and possessed by Korean universities [5]. In particular, the RISS has numerous Korean master's and doctoral dissertations that are difficult to find in other databases.

DBpia

DBpia is a commercial database provided by Nurimedia for articles produced in Korea. The first page of the article's PDF is provided free of charge, and paid members can access the full PDF. As most articles provide the information needed for a reference on the first page of the PDF, this resource can be quite helpful when checking the references for Korean papers. Basic bibliographic information and export citation functions are provided based on the information written in Korean. As most Korean journals have references written in English, it is recommended to check the PDF directly to ensure the accuracy of the English-language wording. As a point of comparison, other Korean commercial academic databases (such as Koreanstudies Information Service System [KISS], e-article, and Korea Scholar) do not provide a free PDF preview function.

Crossref Link References

Crossref Link References is a DOI search service provided by Crossref. Although it allows several DOIs to be obtained quickly, errors in a reference cannot be identified based on these search results. However, searching for the DOI allows users to directly access the original site to find the necessary

information and edit the reference accordingly. If a DOI is assigned to an article, most of the references will be searched. However, if a DOI is not available, such as for books or reports, other search methods need to be used.

Crossref Simple Text Query

Crossref Simple Text Query is similar to Crossref Link References, as it is a resource for DOIs to be searched and printed. By selecting the 'Include PubMed IDs in results' option, search results for DOI, PubMed ID, and PubMed Central ID will be obtained. If most of the references are for journals that have DOIs, they can be quickly accessed using their DOIs.

Edifix

Edifix is a paid reference proofing website, launched by Inera, which developed the eXtyles program that helps to edit JATS XML in Microsoft Word [6]. Edifix finds, modifies, and provides notifications for the incorrect parts of reference lists using data registered in PubMed, Crossref, and other sources. For biomedical journals, which contain references to articles listed in PubMed, Edifix corrects references with a very high rate of accuracy (Fig. 2). The reference output style uses pre-designed, well-known styles (e.g., Vancouver, American Med-

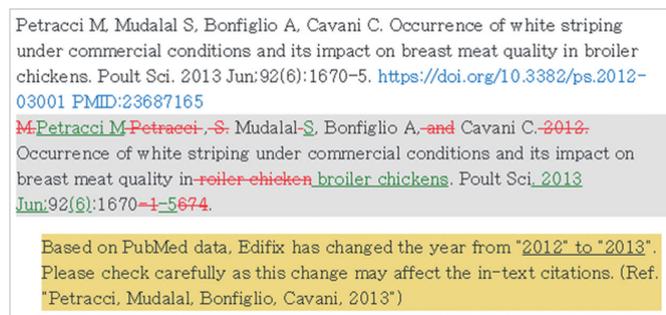


Fig. 2. Example of Edifix modifying a reference with errors to the Vancouver style.

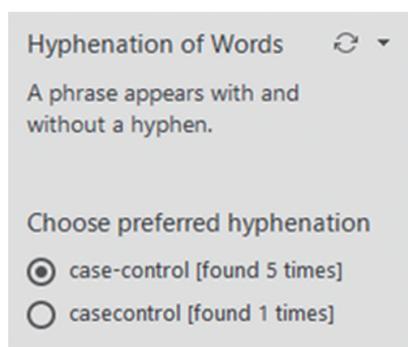


Fig. 3. Example showing the use of PerfectIt to check the consistency of hyphenation.

ical Association [AMA], and APA style), which the users cannot easily customize—to do so, they would need to contact Edifix directly. In addition, it is difficult to compare references for non-journal references, such as books and reports. However, even in such cases, the order of components within the reference is corrected to follow the selected output style. Edifix charges per reference, with rates ranging from 0.09 to 0.16 US dollars depending on the payment plan. Initially, it offers a free trial for 100 cases.

Helpful Software for Manuscript Editing

PerfectIt

PerfectIt is a Microsoft Office Word add-in program that helps to fix common errors in English-language documents [7]. This program can be used for all documents produced in English, not just journal articles. The following tasks can be completed with a few clicks, similar to the Word 'find and replace' process (Fig. 3): 1) identifying abbreviations that have not been defined (or that appear before their definition); 2) ensuring hyphenation consistency (e.g., 'fine tuning' vs. 'fine-tuning'); 3) checking capitalization consistency (e.g., 'Government' vs. 'government'); and 4) ensuring consistency of italics (for foreign words such as '*in vivo*' vs. 'in vivo'). It is offered as a 14-day free trial and charges 70 US dollars per year per user.

Grammarly

Grammarly is a program that finds and corrects errors in English grammar and spelling [8]. It is available for all platforms, including Chrome, Word, Gmail, and Outlook. It suggests corrections by checking for incorrectly spelled words, tense mismatches, inappropriate use of articles, and singular/plural mismatches (Fig. 4). By default, all these features are

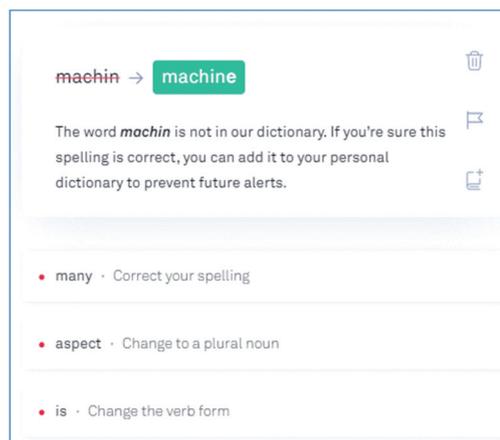


Fig. 4. Example showing the use of Grammarly to check for English-language grammar and spelling errors.

available free of charge, with a premium version of 29.95 US dollars per month. The premium version provides additional features, such as readability, vocabulary enhancement suggestions, and genre-specific writing style checks.

Conclusion

Manuscript editing involves thoroughly improving the quality of an article. The scope of work varies, and the best approaches to produce optimal results may also vary depending on the academic field, academic journal, and the editor's preferences. Therefore, it is difficult to establish a standardized method that everyone can follow. However, to promote the development of the manuscript editing industry and overall quality improvement, it is essential to standardize manuscript editing-related tasks and it is highly important to make efforts to share related skills and knowledge. In addition, as the academic publishing environment changes, the role and demands of manuscript editors are continuing to change, making it increasingly necessary for manuscript editors to participate in training and workshops to remain up to date.

Conflict of Interest

No potential conflict of interest relevant to this article was reported.

References

1. Yi HJ, Chang JH, Seo YJ. Analysis of the results of the first implementation of the Korea Manuscript Editors Certification. *Sci Ed (Seoul)* 2017;4:34-8. <https://doi.org/10.6087/kcse.87>
2. Iverson C. "Copy editor" vs "manuscript editor" vs ...: venturing onto the minefield of titles. *Sci Ed (Oakbrook Terrace)* 2004;27:39-41.
3. WorldCat [Internet]. Dublin, OH: Online Computer Library Center; 2019 [cited 2019 Nov 2]. Available from: <https://www.worldcat.org>
4. Naver Academic. Data provider [Internet]. Seongnam: Naver; 2019 [cited 2019 Nov 2]. Available from: <https://academic.naver.com/cp.naver>
5. KERIS. Guideline for using RISS [Internet]. Daegu: Korea Education and Research Information Service; 2018 [cited 2019 Nov 2]. Available from: <https://www.riss.kr/AboutRissGuide.do>
6. Edifix [Internet]. Belmont: Inera; 2019 [cited 2019 Nov 2]. Available from: <https://edifix.com>
7. PerfectIt [Internet]. London: Intelligent Editing; 2019 [cited 2019 Nov 2]. Available from: <https://intelligentediting.com>
8. Grammarly [Internet]. San Francisco, CA: Grammarly; 2019 [cited 2019 Nov 2]. Available from: <https://www.grammarly.com>