Library News
News and information from the Research Medical Library
THE UNIVERSITY OF TEXAS MD Anderson Cancer Center
Making Cancer History®

Interpreting iThenticate Reports

Although authors and journals regularly use iThenticate to review biomedical manuscripts for plagiarized text, the program itself cannot definitively determine whether plagiarism is present. Therefore, care must be taken—and some effort expended—when interpreting iThenticate reports.

iThenticate uses a proprietary algorithm to determine whether a submitted manuscript has text that matches sources in its vast content database, which includes billions of web pages and millions of published articles. iThenticate then generates a report showing the manuscript with matching text highlighted, the sources of matching text, and the proportion of matching text in the manuscript, also known as the similarity index.

Even experienced iThenticate users tend to focus on the similarity index, believing it to reflect the amount of problematic text in the manuscript: the higher the index, the thinking goes, the greater the likelihood of inappropriate reuse of previously published text. But this is not necessarily the case. The same similarity index can have vastly different meanings for different manuscripts; for example, a similarity index of 25% might indicate that 25% of the manuscript matches text in only one source, or it might indicate that 25% of the manuscript matches text in 25 different sources. Or it could indicate that multiple paragraphs in the manuscript were copied verbatim from published sources, or it could indicate that 25% of the manuscript is a list of references that also appear in bibliographies across multiple published articles.

Therefore, rather than focusing on the similarity index, users should focus their efforts on reviewing the annotated manuscript to identify matching text that is genuinely problematic.

Matching text that is not problematic but still contributes to the similarity index includes proper names, such as those of authors and their institutions on the title page; technical or discipline-specific terms or phrases; equations or formulas; boilerplate text, such as conflict-of-interest disclosures and funding acknowledgments; references in the bibliography; and quoted material that is properly cited. Sentences or phrases that are ubiquitous in scientific writing, such as "Breast cancer is the second most common cancer among women in the United States" and "*P* values of less than 0.05 were considered statistically significant," are also fine.

Text that is genuinely problematic (and warrants substantial revision) typically takes the form of a string of one or more sentences that exactly or closely repeats text in one or more of the sources identified by iThenticate. Problematic text can also take the form of uncommon phrases. Whether matching text is problematic may depend less on its length than its location; an entire paragraph in the Methods section might not be problematic—if it's one's own previously published text, say—but a single sentence in the Discussion might be.

To reduce the similarity index and facilitate the review of the annotated manuscript, one can adjust iThenticate's filters to exclude quotations and/or certain sections of the manuscript, including the bibliography. (Such sections must have an identifying heading to be excluded.) One can also adjust the filters to exclude individual text matches shorter than a certain number words and/or sources whose entire matched text is less than a certain number of words or percentage of the manuscript.

For more information about iThenticate, including instructions for requesting an account: https://mdanderson.libanswers.com/faq/42956

Creating Accessible Links

Hyperlinks serve as valuable tools for readers to quickly access information. When reviewing documents, emails, or webpages, many readers scan the content in search of pertinent details, often relying on hyperlinks. It is important to ensure that the link text is accessible to all readers. Link text, which is the clickable text associated with a URL, should be crafted with descriptive language, standardized when possible, and designed to stand out from the surrounding text.

The words used for the hyperlink should clearly indicate where the link will direct the reader without requiring additional context, and the link should not be the full URL. This will make it clear to readers where the link will take them at a glance. This practice is especially important for readers who use screen readers. Screen readers can navigate a page by skipping through links, allowing readers to skim for the information they need. When link text lacks context, such as using generic terms like "click here," the reader will not know the link's destination. Similarly, if the link text is the full URL, the screen reader will read it out, often letter by letter, causing delays and impairing understanding.
For example:

- Incorrect: Read library news here: https://mdandersonorg.sharepoint.com/teams/LibraryNews/SitePages/RML.aspx
- Incorrect: To read library news, click here.
- Correct: Read library news.

If you link to the same page more than once, use the same text to describe it every time. Conversely, don't use that same text to link to a different page. If link text is too similar, it can confuse your readers.
For example:

- Incorrect: Read library news or subscribe to library news.
- Correct Read library news or subscribe to library news.

The standard way to denote a hyperlink is to change the text color and underline it. It's important to do both, as some readers may have difficulty differentiating colors and need the underline to easily see the link. To test if your hyperlinks are accessible, scan your document reading only the links. Make sure you can find the links easily and can tell where the links will go without any additional information.

References
Harvard University. (n.d.) *Write helpful links*. Digital Accessibility. Retrieved December 14, 2023.
New York State Education Department. (n.d.) *Create Accessible Hypertext Links*. Retrieved December 14, 2023.
Research Medical Library. (2023, October 23). *Creating Accessible Educational Materials: Linking*. Research Guides. Retrieved December 14, 2023.
WebAIM. (n.d.) *Links and Hypertext*. Retrieved December 14, 2023.

Research Fronts 2023: Annual Report

Clarivate and the Chinese Academy of Sciences (CAS) released Research Fronts 2023. The report unveiled the latest progress and the evolving direction of scientific fields by identifying the significant research specialties in sciences and social sciences.

Research Fronts are groups of highly cited core papers within a specialized topic, identified by tracking the world's most influential scientific and scholarly literature. When such a group of highly cited papers attains a certain level of

activity and coherence, a Research Front is formed, with these highly cited papers serving as the front's foundational "core."

This year's report identifies 128 Research Fronts based on co-citation analysis, comprising 110 "hot" Research Fronts and 18 "emerging" Research Fronts. Hot Research Fronts track active areas while emerging Research Fronts signify rapidly developing areas in scientific research. The selection of "key" hot or emerging Research Fronts involves both professional judgment and quantitative analysis. This report provides a valuable resource for administrators, policymakers, and others involved in monitoring, supporting, and advancing research amidst finite resources.

In the 2023 report, the research Fronts are classified into 11 broad areas in the sciences and social sciences.

In **Clinical Medicine**, the Top 10 Research Fronts focus on subfields such as immunotherapy and gene therapy for genetic diseases. Tumor immunotherapy and targeted therapy continue to hold a prominent position in the Research Fronts ranking. The two key research fronts identified by Clarivate are:

- new gene therapies, such as CRISPR/Cas9 gene editing and shRNA targeting BCL11A for sickle cell disease and β-thalassemia, and
- KRAS(G12C) Inhibitors and tumor-targeted therapy.

In **Biological Sciences,** the Top 10 Research Fronts include protein structure prediction using artificial intelligence (AI), and pan-cancer analysis of whole genomes. Clarivate identified two key research fronts as:

- spatial transcriptomics technology and
- analysis of structural variations in the human genome using the third generation long-read sequencing technology.

The CAS report also identifies the top countries and institutions producing core and citing papers for the key hot Research Fronts. For those interested, the full report Research Fronts 2023 is available for download.

Are There Any Times We Should Use "Anytime

If you've ever been confused about whether to use *any time* (two words) or *anytime* (one word), you're not alone. Even experienced writers sometimes have to pause and think about which of these two options is correct in a given sentence.

The two words *any time* form a noun phrase in which the adjective *any* modifies the noun *time*. In contrast, *anytime* is an adverb meaning "at any time,"[1] and it always modifies a verb or verb phrase.

To decide whether to use *anytime*, simply ask yourself whether it could be replaced by its definition, "at any time." If the answer is yes, then *anytime* is correct.[2]

Examples:
We do not have any time to spare. (*Any* modifies *time*, and the sentence would not make sense if the words "any time" were replaced by "at any time.")
I'm not busy today, so call me anytime. (The adverb *anytime* is correct because it could be replaced by "at any time.")

References
1. Merriam-Webster.com. Anytime. Accessed December 4, 2023. https://www.merriam-webster.com/dictionary/anytime
2. Dictionary.com. What's the difference between "anytime" vs. "any time"? Accessed December 4, 2023. https://www.dictionary.com/e/anytime-vs-any-time

BioCyc: a Pathway/Genome Database Collection

The Research Medical Library now has access to the Pathway/Genome Database Collection from Biocyc. Access is available while on campus or connected remotely through VPN, Velo Cloud, or the library's off-campus login. According to the Biocyc website, "BioCyc is a collection of 20,043 Pathway/Genome Databases (PGDBs) for model eukaryotes and for thousands of microbes." Please contact the Research Medical Library with any questions.

Must Read Books of 2023

What has our library staff been reading this year? Look below to see everyone's favorite books of 2023! Be sure to search our free eBooks and audiobooks on Libby to discover more.

It's Always Been Ours by Jessica Willson. I found this book incredibly enlightening and plan to make it a yearly read! – Saryah Leyton, Librarian

Lessons in Chemistry by Bonnie Garmus. Women deserve to have their stories told and believed. *Lessons in Chemistry* might be fiction, but https://www.pictureascientist.com/ shows that women in STEM still suffer gender inequality, discrimination, and harassment. – Lesli Moore, Library Information Analyst

Finding Me by Viola Davis. This is especially good on audiobook because she reads it! – Laurissa Gann, Associate Director

Newcomer by Keigo Higashino. A murder mystery with a unique way of telling the story—each section is told from the point of view of the people connected to the victim as they're interviewed, starting with those who only knew her

tangentially and building up to her family and friends. – Kelsey Hensler, Senior Librarian

A Future We Can Love by Susan Bauer-Wu. Inspired by a conversation between the Dalai Lama and teen activist Greta Thunberg. Subtitle: "How we can reverse the climate crisis with the power of our hearts and minds." Fresh perspective, with varied voices, very well written, hopeful. – Sunita Patterson, Senior Scientific Editor

A Tale for the Time Being by Ruth Ozeki. The author interweaves two fascinating stories, one narrated by Nao, a 16-year-old girl living in Japan, and the other about Ruth, a 50ish novelist living in the US Pacific Northwest. The two main characters are linked by a Hello Kitty lunchbox with a Japanese-language diary inside that washes up on a beach near Ruth's home after the 2011 tsunami. I don't often enjoy novels with elements of the supernatural, but this book was wonderful. – Stephanie Deming, Senior Scientific Editor

The Power by Naomi Alderman. This is a fascinating fictional alt history/future exploring what would happen if women were stronger than men. *The Tenant of Wildfell Hall* by Anne Bronte. Anne Bronte is the most obscure of the Bronte sisters but, in my opinion, underrated. – Erica Goodoff, Senior Scientific Editor

All the Pieces Matter: The Inside Story of The Wire by Jonathan Abrams. I can never get enough of the TV show *The Wire*! The interviews were done before the deaths of Michael K. Williams and Lance Reddick, so their voices are included in the book. The show also featured writing contributions from novelists George Pelecanos, Dennis Lehane, and Richard Price. – Travis Holder, Librarian

All You Need Is Ears by George Martin. He talks about a lot more than just working with The Beatles, including his plan to build a "floating studio" on a ship. – Donald Norwood, Scientific Editor

I found myself reading short story collections this year. – Amy Ninetto, Senior Scientific Editor

- *Stories of Your Life and Others* and *Exhalation* by Ted Chiang
- *How Long 'til Black Future Month* by N.K. Jemisin
- *The Assassination of Margaret Thatcher* by Hilary Mantel

Inciting Joy by Ross Gay and *Babel, or the Necessity of Violence* by R.F. Kuang. They pair well because the first one's a collection of essays (wide-ranging, ebullient and brilliant, yet approachable) by a writing teacher, and the second one's a dark academia fantasy about the magic and social/political power of language translation. – Sarah Bronson, Scientific Editor

Support your local library and independent bookstores!

- Brazos Bookstore
- Kindred Stories
- Blue Willow Bookshop
- Murder By The Book