



BEST PRACTICES FOR COMPARING PLAGIARISM DETECTION SERVICES





You have choices for plagiarism checking, so you'll probably want to compare them, a process that is not as simple as it appears. Here are a few things to consider when comparing plagiarism detection services.

1. All plagiarism services assign a score to each paper that is submitted and scanned.

However, each service has different methods for calculating scores, different comparative databases and different technical approaches, so the scoring reports among the services are not necessarily comparable. Scores are not absolute indicators of plagiarism but rather numeric suggestions for further investigation.

This means that you can't compare the 19% match you got on one service with the 32% score that was returned from another service. Both scores suggest that some text in the submitted paper matches other writing, but this difference in scoring doesn't suggest that one scan is more accurate or rigorous than the other. It simply means that the methods for comparing and calculating the score are different.

2. A perfectly original paper will not always get a score of 0% and a published paper will not always get a score of 100%.

Most comparisons of plagiarism detection services involve submitting papers with a known level of copied or plagiarized text. This might range from a completely original work with no copied text (expecting a score of 0%) to the entirety of a published paper (expecting a score of 100%). However, these exact scores are rarely achieved by any plagiarism detection service.

Plagiarism scanning allows for some variability when scoring papers. This may produce some false positive results (scores that suggest plagiarism when none has occurred). On the other hand, if only verbatim matches were identified, many more false negatives would happen (cases of plagiarism that are missed). So a perfectly original paper will not always get a score of 0% and a published paper will not always get a score of 100%.

If the software was rigid and didn't allow for variability in scoring, it would not be effective in detecting the many forms that plagiarism can take such as disguised plagiarism, patch-writing, shake and paste plagiarism, etc. By expanding the rules for matching and detecting more than verbatim matches, plagiarism-checking software increases the likelihood of identifying the misappropriated or improperly cited work of others while minimizing false positives.

This means that no plagiarism detection service is fully authoritative and that subjective analysis is always required. Keep in mind that the goal of plagiarism-detection software isn't to verify plagiarism, but to identify potential cases of plagiarism so the work in question can be inspected in more detail.



3. To simulate common scenarios, consider submitting original work that includes one or more passages that are copied from other sources.

Plagiarism scores are an estimated percentage of matching text relative to the overall size of the paper. Therefore, papers with more copied content should produce higher scores than those with less copied content.

4. Inspect the scoring reports carefully, and look for matches with the original publication, not secondary citations.

The more important or popular a licensed article is, the more likely it is that someone has already plagiarized it, so when you look for matches to the original source, you'll find matches to other examples of plagiarism, not the original. These secondary references are likely to produce a matching score. Inspect the scoring reports carefully, and look for matches with the original publications, not secondary citations.

5. No automated system can verify plagiarism. It requires human, subjective evaluation.

Instructors should calibrate their threshold for each plagiarism system they use or compare. Experience suggests that for any given plagiarism service, each instructor will establish an approximate level of scoring below which they do not investigate, and above which they do investigate. For example, as a rule of thumb, any paper with a score of 15% or higher in VeriCite probably warrants inspection by the instructor or teacher.

The point is that no automated system can verify plagiarism. It requires human, subjective evaluation. Plagiarism detection tools streamline this process by scoring each paper then providing the instructor with the tools to inspect the sources for suspect papers.

6. Sanctions, perhaps, should be a last resort. Engaging students in a discussion about writing with sources should always come first.

How exact should plagiarism detection be? It depends on your needs. If you are an instructor in a course, your focus is probably on teaching your students to write with sources. You want your students to author responsibly, to understand and respect the intellectual property of others and, of course, to avoid the sanctions of the campus judicial policies on academic integrity. The presence of a plagiarism-detection service in the assignment-submission process serves as an incentive, regardless of the process.

If most forms of plagiarism are flagged for inspection by the instructor, these goals will be realized. As in almost anything, achieving 100% certainty is quite difficult and expensive, if possible at all. Debora Weber-Wulff, a thought leader in plagiarism detection, has said [1] "...it is not important in an education setting to find all of the plagiarism in a paper. It is sufficient to find enough for a sanction to be necessary." Sanctions are perhaps a last resort, but engaging the student in a discussion about writing with sources comes first.

[1] Weber-Wulff, et al., 2013. Plagiarism Detection Software Test 2013. http://plagiat.htwberlin.de/?attachment_id=2916