



## Guidance for Ithenticate Reports

This is a guide to authors about copied text, plagiarism and copyright infringement found in papers. It relates especially to the interpretation of Ithenticate output and other similarity-check reports. Ithenticate (<http://www.ithenticate.com/about>; related to Turnitin) is an online plagiarism detection software used by academic publishers.

First, plagiarism and copyright are reviewed, and some background to Ithenticate is given. Below, you will find categories of different types of copied text/plagiarism and action that authors should take.

### What is copied text?

Copied text means that part of a paper has been published somewhere before. If something is highlighted as copied text in a similarity report it doesn't necessarily mean there is a problem or that the authors are deliberately intending to mislead. It does mean that the text might need checking further, to be certain that the wording is acceptable.

### What is plagiarism?

Plagiarism means representing someone else's ideas as your own. This could mean literally using the same words, or using the same approach without giving proper acknowledgment. Ithenticate cannot detect all forms of plagiarism. It cannot detect:

- Plagiarism of general ideas or approaches where different words and terms are used.
- Any plagiarized work that isn't included in its database, including unpublished work, very recently published documents, and work that is not published online.
- Copying of figures, although copied text in the caption can give an indication.

It can detect:

- Direct copying of one source into another, ranging from short words or phrases up to whole paragraphs.
- Most cases where a paragraph was trivially modified by the authors.

Most importantly, Ithenticate cannot detect the authors' motives. Copied text might be due to inadvertently removing or forgetting to add a citation; alternatively, it could be that the authors have deliberately taken words directly from another author without attribution.

### **What is copyright infringement?**

Copyright infringement refers to illegal reuse of certain kinds of work, including writing. Note that it is not exactly the same as plagiarism, which is more concerned with ideas; copyright is concerned with how something is concretely expressed—the exact words used. All published text is covered by a license that determines when and where it can be reused. Papers published in subscription journals typically have a license that does not automatically allow reuse, whereas those in open access journals typically allow reuse with citation of the original paper.

Ithenticate can detect copyright infringements, but the license of the original needs to be checked to see whether the text was legally copied or not.

### **What is self-plagiarism?**

Copying from the authors' own previous work is referred to as self-plagiarism. In reality, it is not plagiarism since it does not copy the view of another author. It may breach copyright law if the author does not hold the copyright of the original piece. Self-plagiarism is considered unethical because it means that the new paper contains limited novelty. However, most publishers tolerate larger amounts of self-plagiarism than they would plagiarism of other authors.

The authors of a document are indicated for most Ithenticate sources, so self-plagiarism can be detected in many cases.

### **What does an Ithenticate report show?**

Ithenticate reports highlight parts of the text where a match with previously published work was found. Each work that has been copied from is numbered, and each match is assigned the number corresponding to that work. If multiple texts were copied from, Ithenticate usually shows only the top result, i.e., the one with the largest number of matching words.

An overall percentage of copied text is given. It is difficult to assign a 'safe' or 'unsafe' level to this value, since it depends on the type of copying found and sometimes where it is located in the document. A document with a single copied and uncited paragraph is problematic but can have a low matching score, but one containing many short phrases that appear in a variety of different other documents can be fine but have a much higher score.

# Issues found in Ithenticate reports

The following outlines the main kinds of problems found in Ithenticate reports and recommendations on how to solve them.

## **Rephrasing**

### *The problem*

The author has taken text from a paper and changed some of the words so that it doesn't look as much like the original. This is unethical, since it plagiarizes the ideas and expression of the original work.

### *The solution*

Check that the original text has been cited. In addition, either completely rewrite the paragraph in your own words, or change it back to the original and use quotation marks.

## **Short direct quotes**

### *The problem*

A phrase copied directly from another paper that is more than simply a standard expression. These are usually not a problem from a copyright point of view, as they count as 'fair use' (under US law, other jurisdictions have similar exemptions). However, it may count as plagiarism if it contains a significant idea that readers would interpret as coming from the author. It is probably ok if it is self-plagiarism, however publishers vary in policies towards self-plagiarism.

### *The solution*

Check that the original paper has been cited and consider putting the text in quotations marks or rewriting it.

## **Long direct quotes**

### *The problem*

Several consecutive sentences or paragraphs taken directly from another piece of work. Unless it is indicated as a quotation this is considered plagiarism, even if the other paper was cited. It is also likely to infringe copyright. Even where this is self-plagiarism, most publishers would not accept such a long passage repeated from previous work. In some experimental fields, it can be acceptable in the methods sections, provided that the original source of the method is clearly cited. In mathematics, statements of theorems can also fall into this category without being problematic.

### *The solution*

The section should be removed and rewritten in the authors' own words. Consider a shorter text summarizing the main points and referring readers to the other publication for full details. For

methods or theorems, make sure the original source is clearly cited and that readers would have the impression that the text is original.

### **Pick 'n mix**

#### *The problem*

Many short words and phrases (ranging from several words up to whole sentences) taken from different publications with almost no original text in between. To write this part of the paper, the authors copied and pasted from multiple other works. This issue is more commonly found in review papers and introduction sections. It is also a common approach from students who may not be aware that the practice is unethical. Even if they include the original citations for all quotations, this is considered plagiarism as the authors have not been able to present the ideas in their own words.

#### *The solution*

The authors should rewrite the section and reconsider their approach to writing. As well as listing previous results, they should offer more analysis and critique.



© MDPI. Published under a Creative Commons CC BY-NC v4.0 license.