

Guide de bonnes pratiques concernant iThenticate

L'analyse des résultats de similitude du logiciel iThenticate nécessite une certaine prudence. Il n'y a pas de nombre magique à valeur quasi-juridique de similitude car il n'y a pas deux documents complètement identiques, ne serait-ce que dans leurs objectifs (ex : thèse vs article scientifique vs rapport).

Néanmoins, un score de similitude trop élevé peut indiquer un usage trop important de citations directes non signalées ou de sources secondaires. Règle générale, un score de 15 à 20 % pourrait être considéré comme une bonne limite, selon le type de document. Si le résultat est plus élevé, il est important d'analyser les détails du rapport de similitude, et de tenir compte du contexte ET des configurations du logiciel.

Le rapport de similarité signale la répartition du pourcentage global entre les différentes sources. Un score de similitude relativement élevé composé de petits pourcentages (ex : 1%) provenant de nombreuses sources ne serait pas très inquiétant. À l'opposé, un score plus faible (ex : 15%) mais ne provenant que d'une ou deux sources serait nettement moins acceptable. Il faut donc examiner le rapport de similitude en détails plutôt que de se contenter de la note globale.

Règle générale encore une fois, une source de plus de 2% nécessite probablement une attention particulière, surtout pour un article à soumettre.

Il faut prendre en considération l'objectif du texte analysé. Ainsi, même si la méthodologie est la même qu'un article précédemment publié par le même auteur et ce, de façon tout à fait logique et éthique (ex : même méthodologie mais article présentant d'autres résultats), l'éditeur de la revue qui recevra le deuxième article n'aura pas nécessairement la même finesse de jugement et risque fort de rejeter immédiatement le manuscrit sans possibilité d'appel. Par prudence, il vaut donc mieux récrire avec une certaine souplesse cette partie du texte.

Il est possible d'ajuster la configuration du logiciel

- En excluant la bibliographie (recommandé)
- En excluant certaines phrases ou séquences de mots précises qui pourraient augmenter indûment le résultat (ex : longs noms d'organismes, etc.)
- En excluant les citations directes dans le texte, visibles soit par une indentation dans le texte, soit par des guillemets (recommandé)
- Ou en spécifiant un seuil minimal de mots contigus dans une même phrase à ne pas considérer dans l'analyse. A utiliser avec prudence, iThenticate ayant déjà un tel seuil par défaut calibré par une longue expérience.

D'autres options de configuration sont disponibles, dont ne pas considérer le résumé ou la section Methods. A utiliser avec une grande prudence entre autres pour les raisons évoquées ci-haut.

Quelques pistes pour en savoir plus :

<https://www.ithenticate.com/plagiarism-detection-blog/bid/63534/CrossCheck-Plagiarism-Screening-Understanding-the-Similarity-Score#.WJuaXYVOK71>

https://www.icevirtuallibrary.com/pb-assets/for%20authors/CrossCheck_interpretingscores-1638444252.pdf

<https://media.springernature.com/full/springer-cms/rest/v1/content/16185088/data/v2>