

CROSSCHECK

identificação de plágio por similaridade

Eixo temático: Periódicos Científicos

Modalidade: Pôster

1 INTRODUÇÃO

A produção científica universal é ampla e em constante crescimento anual. Identificar o conteúdo ideal para a publicação e aceite de um artigo pelos editores, textos redigidos pelos pesquisadores, orientandos e alunos, o conteúdo textual de uma tese e demais documentos textuais é importante para as instituições e revistas que não podem errar e publicar textos similares, e indiretamente assumir o plágio dos textos, mostrando erratas e retratações. A questão do plágio, e os aspectos éticos na comunicação científica sempre foram conhecidos pelos editores, autores, pesquisadores, docentes e discentes, de difícil identificação. Todos estes atores da comunicação científica são possíveis usuários e necessitam de sistemas de identificação de plágio para auxiliá-los em comparar similaridade de conteúdo textual a ser publicado. Um autor que pretende identificar se outros estão citando seus textos, esse mapeamento é muito importante para trabalhar em suas linhas de pesquisa e comparar sua produção científica. Sistemas de identificação de similaridade são essenciais no meio acadêmico.

[...] comprehensive databases and gives you a very good indication of whether the paper contains significant sections of duplicated text. A quick glance at the Similarity Report will often be all that is needed to confirm a manuscript only contains small matches comprised of frequently used terms or phrases, or at worst, poorly cited content that can be corrected. If, however, the Similarity Report identifies one or more matches that are quite large, or lots of smaller matches even with the bibliography excluded, the Content Tracking report should be your next port of call (MEEDINGS, 2011, p. 2).

A similaridade é estudada na cientometria e infometria, como objetos as palavras em bases de dados e com variáveis de recuperação e principalmente medir a relevância, por percentuais e números de palavras e frases idênticas ou com pequenas alterações (BORNMANN et al., 2008).

CrossCheck (2012) allows member organizations to run a text-similarity check of a submitted manuscript against the corpus of scholarly and professional literature (and optionally the wider internet) and returns a list of documents with any matching or similar text highlighted for further editorial review (MOED et al., 2012).

Os editores de periódicos científicos também são responsáveis pelo conteúdo textual enviado pelos autores, pela participação direta no aceite e publicação e indireta no conteúdo textual, ao qual, deve precaver-se sempre, assim não podem publicar conteúdos e dados de outros autores sem a citação ou prévia autorização, e quando identificarem plágio, rejeitar os artigos. O *CrossCheck* é um serviço destinado aos associados do *CrossRef*¹, em parceria com

a *iThenticate*^{TM2} e utiliza para o mapeamento e avaliação de conteúdos publicados na Internet e suas bases de dados, com objetivo de prevenir e identificar plágio em texto e a integridade de seu conteúdo.

2 MÉTODOS

Este trabalho tem por finalidade mostrar o *CrossCheck* e suas principais características como sistema de identificação de plágio, através de breve revisão bibliográfica, mostrando os principais objetivos. Serão apresentadas figuras e imagens com mapas de identificação de plágio do sistema (Pôster), com resultados de similaridade por relevância, por quantidade de palavras, resultados de busca com estatísticas, formas de acompanhar as correções solicitadas aos autores e índices de relevância de plágio em conteúdo textual, através de modelo de análise multidimensional de dados utilizado pelo sistema. Mostra também qual o funcionamento básico do sistema, seu acesso e custos para instituições e editores. Foram realizados testes do autor com artigos e textos enviados por editores científicos.

3 RESULTADOS E DISCUSSÃO

Identificar o plágio em conteúdos textuais através de sistema de varredura semântica com análise multidimensional de conteúdo publicado na Internet é essencial para instituições e editores. Identificar ainda se um conteúdo já foi publicado ou está sendo avaliado em outro periódico em parte ou no todo faz parte das atividades gerenciais de um editor e também de orientadores e pesquisadores vinculados a programas de pós-graduação. No meio acadêmico a identificação de conteúdo já publicado é essencial para a originalidade e o ineditismo, juntamente com os métodos e resultados alcançados nas pesquisas acadêmicas. O *CrossCheck* é um sistema desenvolvido cooperativamente, com acordos de fidedignidade e segurança dos dados de convênio entre os principais editores científicos do mundo e associadas ao *CrossRef*. No Brasil ainda é incipiente sua utilização, principalmente devido ao desconhecimento do sistema, mas os editores necessitam de suporte. Existe a utilização de softwares por alguns pesquisadores e editores, mas, o conteúdo a ser enviado para sistemas de identificação de plágio requer que estes sistemas também garantam a segurança e originalidade, e que não será enviado para bancos de dados externos. Os sistemas de identificação de plágio tem de apresentar critérios de garantia do conteúdo a eles postado. A identificação de conteúdos similares de testes realizados com o sistema foram considerados muito importantes pelos editores.

CONSIDERAÇÕES FINAIS

O *CrossCheck* é utilizado pela editora *Elsevier* e *Nature*, entre outras das principais do mundo, tem como parceiros grandes editores como a *Ebsco*, *Thomson Reuters*, por exemplo, que também utilizam este sistema. Em números atuais é utilizado por 300 editoras, com 70.000 títulos de periódicos, 40 milhões de itens indexados e 30.000 documentos checados mensalmente. Existem métodos para que o conteúdo postado no sistema tenha segurança e garantias legais, assim, as instituições e editores terão confiabilidade do conteúdo enviado e checado no sistema. O serviço é oferecido um custo baixo, uma pequena anuidade e checagem de documentos ao custo de U\$0,75 (setenta e cinco centavos de dólar americano). Em 2010, um artigo publicado na *Nature*, mostra que com o *CrossCheck* uma revista chinesa

identificou que 31% dos artigos submetidos a ela eram plágio, e a utilização do sistema foi essencial para esta identificação (ZHANG, 2010). O sistema é uma ferramenta útil em inúmeros processos nas instituições, como: checagem de teses, dissertações, projetos, monografias, e editores científicos no conteúdo dos artigos que foram submetidos para suas revistas, assim, o editores de periódicos devem utilizá-lo principalmente para verificação dos artigos recém submetidos às revistas.

REFERÊNCIAS

- BORNMANN, L.; MUTZ, R.; NEUHAUS, C.; HANS-DIETER, D. Citation counts for research evaluation: standards of good practice for analyzing bibliometric data and presenting and interpreting results. **Ethics in Science and Environmental Politics**, Oldendorf, v. 8, no. 1, p. 93-102, 2008. doi: [10.3354/esep00084](https://doi.org/10.3354/esep00084)
- CROSSCHECK. Disponível em: < <http://www.crossref.org/crosscheck/index.html> >. Acesso em: 20 mar. 2012.
- CROSSCHECK fees. Disponível em: < http://www.crossref.org/crosscheck_fees.html >. Acesso em: 20 mar. 2012.
- CROSSREF. Disponível em: < <http://www.crossref.org> >. Acesso em: 20 mar. 2012.
- ITHENTICATE. Disponível em: < <http://www.ithenticate.com> >. Acesso em: 20 mar. 2012.
- MEDDINGS, K. CrossCheck plagiarism screening: what's the magic number? **EON Editorial Office News**, Aug. 2011. Disponível em: < <http://www.ismte.org/File/Aug11CrossCheck.pdf> >. Acesso em: 21 abr. 2012.
- MOED, H. F. et al. Citation-based metrics are appropriate tools in journal assessment provided that they are accurate and used in a informed way. **Scientometrics**, Amsterdam, v. 90, 2012. doi: [10.1007/s11192-012-0679-8](https://doi.org/10.1007/s11192-012-0679-8) ahead of print.
- PENG, C. Focus on quality, not just quantity. **Nature**, v. 475, p. 267, 21 July 2011. doi: [10.1038/475267a](https://doi.org/10.1038/475267a)
- ZHANG, Y. Chinese journal finds 31% of submissions plagiarized. **Nature**, v. 467, 9 Sept. 2010. doi: [10.1038/467153d](https://doi.org/10.1038/467153d)

¹ *CrossRef*, associação de editores científicos com milhares de editoras associadas e com 52 milhões de registros bibliográficos.

² *iThenticate*, líder mundial em identificação e prevenção de plágio.