

GROSSI JÚNIOR, José Alberto. Análise comparativa de ferramentas de extração de metadados em artigos científicos. 84 p. Dissertação (Mestrado em Ciência da Informação) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2015.

*São inúmeras as ferramentas para extração de metadados em artigos científicos, tendo cada uma sua particularidade, tecnologia e técnicas utilizadas. Porém, com a crescente produção científica e a grande variedade de editoras, eventos e congressos, um número cada vez maior de artigos permanece sem uma extração de metadados eficaz, o que dificulta a disseminação de conhecimento e principalmente a pesquisa eletrônica desses documentos. Este trabalho compara a capacidade de extração de metadados de algumas ferramentas pré-selecionadas - Cermine, CiteSeer, CrossRef e ParsCit - utilizando para isso um experimento empírico com um conjunto de artigos. Esse conjunto abrange diversas áreas do conhecimento, diversos eventos e formatos visuais diferentes. O experimento foi realizado em ambientes pré-configurados de acordo com a característica técnica de cada ferramenta, permitindo que todos os artigos tivessem seus metadados extraídos por cada uma delas e os resultados comparados individualmente. Desta forma, com base nos resultados apresentados, pôde-se identificar o comportamento de cada uma das ferramentas no tocante à sua capacidade de extração correta de metadados. Exceto pela ferramenta CrossRef, todas as demais obtiveram resultados acima de 60%, chegando a 86.83% da ferramenta Cermine. Além disso, foram evidenciadas as principais fragilidades dessas ferramentas; pontos onde seriam necessários ajustes; metadados para os quais se obtém um maior sucesso na extração. Ademais, é apresentado também um índice de confiabilidade, que permite estabelecer para cada ferramenta uma nota. Essa nota é calculada com base nos resultados obtidos na extração de metadados pela seleção de artigos realizada.*