

MARTINS, Agnaldo Lopes. O uso do sintagma nominal na recuperação de documentos: proposta de um mecanismo automático para a classificação temática de textos digitais. 192f. Tese (Doutorado em Ciência da Informação) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2014.

*Esta tese objetivou avaliar o uso do sintagma nominal como fonte de dados para um sistema automático de classificação de documentos textuais armazenados no formato digital. Foram utilizadas diversas ferramentas tecnológicas que transformaram artigos científicos em uma lista de sintagmas nominais que foram utilizados para treinamento de um sistema classificador baseado em treinamento supervisionado. Dentre as ferramentas utilizadas o software Palavras foi o responsável pela identificação e remoção dos sintagmas nominais dos corporas utilizados. Para treinamento da máquina classificadora foi utilizado o aplicativo SVMLight. A metodologia foi desenvolvida em duas etapas; na primeira foi realizado um teste qualitativo na comparação entre os documentos do corpus; e na segunda etapa foi realizado o treinamento utilizando SVM com um número maior de documentos. Ao final, vários testes foram realizados sendo possível demonstrar que a metodologia proposta foi capaz de classificar documentos com alta precisão.*