

6 Conclusão

Foi introduzida nesta dissertação a área de Mineração de Textos (MT), que nada mais é do que a descoberta de conhecimento novo e inovador a partir de grandes volumes de texto. MT é bastante interdisciplinar, o que levou-se a discutir sobre as áreas de conhecimento que a compõem, como a Linguística e a Inteligência Artificial. Alguns casos de sucesso também foram apresentados, reforçando a importância e a aplicabilidade do assunto tratado.

Em seguida, foi explorada uma metodologia para investigação textual proposta em outros trabalhos encontrados na literatura e composta pelas etapas de coleta, pré-processamento, indexação, mineração e análise, encadeadas na mesma ordem apresentadas. Para cada uma destas, foi apresentado o que atualmente é considerado “estado da arte” pela comunidade acadêmica, através de técnicas e algoritmos comprovados e reconhecidos como sendo de sucesso.

Partindo do princípio que é necessário determinar o que se pretende obter dos dados, isto é, qual objetivo final do processo, foram apresentadas as principais tarefas de MT, presentes na etapa de mineração.

Finalmente, foi apresentado o desenvolvimento de um sistema de mineração textual, o qual foi subdividido em módulos que representam cada uma das etapas da metodologia estudada, sendo possível a obtenção de resultados práticos obtidos a partir da execução de um estudo de caso.

Nosso estudo destacou a enorme importância dos dados não-estruturados para o mundo atual, e a necessidade de aperfeiçoamento constante de técnicas, algoritmos - *softwares* - e *hardware*, visto que a velocidade de processamento de dados textuais não consegue acompanhar o poder de produção do homem, haja vista o imensurável volume de textos circulantes na Internet.

Esta dissertação teve como motivação a tese de doutorado de Christian Nunes Aranha [29], e terá continuidade com as dissertações de mestrado de Roberto Miranda Gomes e Fábio de Azevedo Soares, dentre outras.