



TRATAMENTO DE DADOS EM BIG DATA PARA PORTAL RELACIONADO A LÉXICO ESPECÍFICO

Autoria: Larissa Pavarini da Luz - Lucas Baltazar Cayres - Caio Beraldo -

Resumo: Modalidades de Tecnologia da Informação e Comunicação (TICs) possuem características relacionadas ao processamento, análise e recuperação no tratamento de dados, aspectos de visualização e design, cujos dados são observados nos quesitos variedade, velocidade e volume. Uma das vantagens da análise a disposição dessas categorias está na possibilidade de poder buscar informações com amplitude, garantindo assim, a credibilidade dos resultados e das interpretações. Projetos de Big Data de um modo geral organizam dados de diferentes formatos passíveis de serem trabalhados nos conceitos de Tecnologia da Informação. O objetivo deste trabalho é provar a pertinência quanto à usabilidade e acessibilidade de uma plataforma web que se constitui como objeto dos estudos voltados para a análise e desenvolvimento de Big Datas que concentram bases distintas às buscas de palavras, principalmente aquelas organizadas em léxicos específicos. Para comprovar a eficácia serão utilizados dados do projeto LHisPAR - Léxico Histórico do Paraná - cujo portal, em desenvolvimento, possibilitará a consulta de verbetes, suas abonações, os fac-símiles dos documentos originais de onde os verbetes foram extraídos e todas respectivas transcrições dos documentos. Com devem ser considerados todos os diferentes formatos de arquivos que pode comportar – inicialmente textos, imagens e, posteriormente, arquivos de áudio, vídeos e eventuais recursos gráficos - a diversidade e volume de dados alinha-se perfeitamente ao conceito de Big Data. A relevância da pesquisa relaciona-se ao modo como são propostos os levantamentos de requisitos e disseminação da informação para análise e retorno das solicitações feitas em um portal web – especificamente o do LHisPAR – e na agilidade que pode proporcionar ao usuário/pesquisador quando da consulta dos léxicos envolvidos.