

Modelo Não Relacional Para a Manipulação de Linguagem de Sinais no Cassandra

Uriel Cairê Balan Calvi; Luciana Brasil Rebelo dos Santos; Lucas Venezian Povoia

uriel.caire@gmail.com

O objetivo deste projeto é desenvolver um modelo de dados não relacional capaz de manipular linguagem de sinais, com foco na LIBRAS. O modelo deve atender às particularidades do Cassandra, banco de dados NoSQL escolhido para o armazenamento destes dados. Os sinais serão capturados por meio do sensor Kinect com apoio de um software de captura dos sinais da LIBRAS que foi desenvolvido em uma Iniciação Científica. Os dados coletados são relacionados a um autor e ao significado dos sinais realizados. Em função do tipo e do grande volume de dados, este projeto busca estruturar os dados coletados (i.e., imagens de cor e profundidade, coordenadas dos pontos de articulação, significados dos sinais) e gerados sobre uma plataforma com alta escalabilidade. Para esse fim, serão empregadas técnicas de desnormalização e será feita a utilização de bancos de dados baseados na abordagem Bigtable (CHANG et al., 2008). Com o banco previamente selecionado, foi então realizada a primeira modelagem das informações. Por se tratar de NoSQL, buscou-se aplicar a desnormalização. A partir deste primeiro modelo, foram construídos, em Java, os 'scripts' para realizar a persistência dos dados. Para realizar a comunicação entre o Cassandra e o Java, se fez necessária a utilização do Kundera (IMPETUS, 2016) - uma mapeador de objetos com interface JPA, capaz de utilizar as 'annotations' do JAVA para lidar com bancos não relacionais. Depois de gravar, foi dado início à recuperação dos dados. Em vista da necessidade de pesquisas por texto, e do fato do Cassandra em sua versão 2.2.x não suportar tais pesquisas, a 'search engine' Elasticsearch foi agregada ao projeto. Assim, enquanto o Cassandra armazena os dados, o Elasticsearch indexa os atributos necessários para futuras consultas. Durante o estudo da integração do Cassandra e do Elasticsearch, notou-se que o modelo desenvolvido no início não estava suficientemente desnormalizado, afetando o desempenho da aplicação. Em vista disso, os dados foram novamente modelados. Para realizar uma desnormalização mais profunda, foram levantadas as 'queries' que seriam utilizadas neste processo. Com base no que foi levantado, os dados foram remodelados e as redundâncias necessárias foram aplicadas. De acordo com os testes de persistência e busca, o modelo dos dados

desenvolvido ao fim do projeto atende às tecnologias e às necessidades deste projeto. O levantamento das 'queries' foi essencial para a desnormalização correta destes dados.

LIBRAS. Inclusão social. Tecnologia assistiva. Inteligência artificial