

# **Construção de uma Ontologia para Auxiliar no Mapeamento de Laudos Médicos de Endoscopia Digestiva Alta da Região Anatômica Duodeno**

**Simone Aparecida Pinto Romero<sup>1</sup>, Luiz Henrique Dutra da Costa<sup>1</sup>, Hueli Diana Lee<sup>1,2</sup>, André Gustavo Maletzke<sup>1</sup>, Carlos Andrés Ferrero<sup>1</sup>, Cláudio Saddy Rodrigues Coy<sup>2</sup>, João José Fagundes<sup>2</sup>, Feng Chung Wu<sup>1,2</sup>**

<sup>1</sup>Laboratório de Bioinformática – LABI, UNIOESTE, Foz do Iguaçu, PR

<sup>2</sup> Serviço de Coloproctologia, DMAD, FCM, UNICAMP, Campinas, SP

## **Objetivos**

Laudos Médicos textuais (LMt) são utilizados para representar dados médicos. Este relato de resultado, quando processado e representado adequadamente, torna-se possível a aplicação de técnicas para a extração de conhecimento. Este trabalho tem como objetivo a construção de uma ontologia para auxiliar no mapeamento de LMt de Endoscopia Digestiva Alta (EDA) para Bases de Dados (BD).

## **Métodos/Procedimentos**

Para o mapeamento das informações contidas nos LMt para uma BD estruturada foi adaptado o método descrito em [1], o qual se divide em duas fases. Na Fase 1 são identificados os primeiros padrões textuais contidos nos LMt no qual, por meio de frases que se repetem, é constituído um Conjunto de Frases Únicas. Após, é realizada a geração de  $n$ -gramas e a definição de palavras e expressões que serão utilizadas para padronizar os termos do domínio. Essas palavras e expressões são armazenadas em um arquivo de padronização, sendo essa fase fortemente apoiada por especialistas da saúde. Em seguida, é construído um Dicionário (DI), no qual estão descritos de modo hierárquico, os termos citados nos LMt. Esses termos são utilizados para definir os atributos e os respectivos valores que representam locais, características e subcaracterísticas do domínio. Na Fase 2, o conteúdo dos LMt é padronizado e, baseado no DI, o Algoritmo de Busca e Preenchimento (ABP) é aplicado com a finalidade de mapeamento e preenchimento dos LMt e da BD, respectivamente. Contudo, o DI construído não é capaz de representar integralmente as relações no domínio. Desse modo, foi

construída uma ontologia para substituir o DI no processo de mapeamento e de preenchimento da BD. Assim, é apresentada a construção da ontologia baseado nos dados referentes à região duodenal de 609 LMt de EDA provenientes de uma Instituição de Saúde no período de 2000 a 2005.

## **Resultados**

Após a avaliação de especialistas de domínio médico, verificou-se que a ontologia construída permitiu melhor representação do conjunto de termos descritos nos LMt em relação ao DI, devido às características inerentes do conceito de ontologia [2], como a flexibilidade e a capacidade de representar de modo eficiente e completo as relações entre os termos por meio de representação lógica.

## **Conclusões**

Neste trabalho constatou-se, junto aos especialistas de domínio, que a ontologia construída representou adequadamente os termos dos LMt utilizados e relacionados ao mapeamento de LMt. Como trabalhos futuros destacam-se a expansão do domínio da ontologia e a adaptação do ABP para a aplicação do Método de Mapeamento Apoiado por Ontologias.

## **Referências Bibliográficas**

- [1] Honorato DF, Cherman EA, Lee HD, Monard MC and Wu FC. Construction of an Attribute-Value Representation for Semi-structured Medical Findings Knowledge Extraction. CLEI Electronic Journal, 2008; 11(2): 1-12.
- [2] Gruber TR. Toward Principles for the Design of Ontologies used for Knowledge Sharing. Formal Ontology in Conceptual Analysis and Knowledge Representation, 1993.