

## Plataforma para Especificação, Visualização e Recuperação de Informações em Base Prontuários Eletrônicos Semânticos

Bolsista Ricardo Antonio Zanetti (CTI) rzanetti@cti.gov.br

### Resumo

*Os avanços tecnológicos e computacionais também trazem novos desafios, na mesma proporção em que grandes quantidades de dados são gerados e armazenados. Este projeto consiste na revisão, pesquisa e desenvolvimento para o aprimoramento de prontuários eletrônicos, com foco inicial em oncologia, utilizando técnicas de web semântica e aprendizado de máquina. No presente momento, são apresentados os conceitos, estudos e revisões sistemáticas realizados, que servirão de base para o desenvolvimento de uma plataforma para registro e análise de informações padronizadas que auxiliarão profissionais de saúde e pesquisadores em diversas outras áreas de saúde.*

*Palavras-chave: Prontuário Eletrônico do Paciente, Recuperação de Informação, Visualização de Informação, Inteligência Artificial, Informática Médica.*

### 1. Introdução

A evolução da tecnologia da informação possibilitou o armazenamento de grandes volumes de dados em Prontuários Eletrônicos do Paciente (PEPs). Embora exista uma evolução recente dos padrões para modelagem de PEPs, bem como dos sistemas de organização do conhecimento (KOS – Knowledge Organization Systems), problemas relacionados à qualidade de dados, à recuperação de informações e à interoperabilidade semântica ainda persistem. Tais problemas se tornam ainda mais claros em PEPs de áreas especializadas, como a oncologia pediátrica. Portanto, se faz necessária a pesquisa por meios para prover descrições e anotações padrões de PEPs considerando tanto os aspectos sintáticos quanto os semânticos (Figueiredo et al., 2022).

Nesta linha, o presente projeto contribui inicialmente na investigação e desenvolvimento de modelo de PEP baseado no uso de descritores semânticos utilizando ontologias, tesouros e padrões. Esses PEPs devem ser capazes de prover modelos mais precisos, interoperáveis e que resultam em maior qualidade de conteúdo para posterior análise e visualização no contexto da oncologia pediátrica. Com isso, a solução irá apoiar a prática, pesquisa e gestão por profissionais de saúde.

Evidências iniciais do aprimoramento do reuso, interoperabilidade semântica e qualidade dos dados de PEPs, obtidas com a utilização do *framework* semântico BioFrame (Figueiredo, 2022), fazem parte do conjunto de ideias envolvidas na concepção do projeto.

O projeto também prevê a investigação do design de soluções interativas para coleta de dados em PEPs, por meio de atividades participativas com profissionais de saúde do CIB e preparação de ambiente computacional para aplicação de *deep learning* na base de dados da Fundação Oncocentro de São Paulo. Isso permitirá criar uma solução mais adequada à realidade, rotina dos profissionais e particularidades da especialidade. Um sistema de PEP adequado, que facilite a interação do profissional de saúde e o correto preenchimento, é fundamental para garantir a qualidade e confiabilidade dos dados armazenados nos PEPs.

## 2. Fundamentos Teóricos e Práticos

O prontuário do paciente constitui o conjunto das informações mais relevantes durante o período de assistência prestada ao mesmo, como relatos textuais, exames clínicos (imageamento, análise sanguínea e de tecidos, sinais vitais, etc) e observações gerais que permitem a comunicação entre membros da equipe multiprofissional (CFM, 2002). Ele é de vital importância tanto para o tratamento do próprio indivíduo quanto para a análise científica de casos (em caráter sigiloso quanto a informações pessoais) e levantamentos estatísticos para fins de pesquisa e políticas públicas de saúde (Cruz, 2011).

A adoção de PEPs em ambulatórios e hospitais agiliza os processos de registro, catalogação e análise de dados. Os itens constituintes de um PEP levam em consideração o embasamento científico e metodologias claras no registro de informações. Seus elementos podem ser orientados a tempo, problemas ou a fontes. Isso significa: dados apresentados em ordem cronológica, anotações para cada problema atribuído ao paciente com as informações subjetivas, objetivas, avaliações e plano (Babbott et al., 2014), bem como registros organizados de acordo com o método pelo qual a informação foi obtida (Häyrinen et al., 2008).

Há diversos sistemas e plataformas que implementam PEPs, tanto proprietários quanto livres (abertos). Neste projeto, a solução será baseada em padrões e iniciativas internacionais amplamente adotadas, como o openEHR (que engloba especificações, educação, modelos clínicos e *softwares* abertos) que busca soluções de interoperabilidade para saúde. Cada um de seus aspectos envolve a participação de profissionais clínicos e especialistas em informática da saúde, para a construção e uso de arquétipos a serem utilizados em programas nacionais e locais de saúde eletrônica, trazendo resultados técnicos de outros programas para o mundo real.

A tecnologia da Web Semântica é útil para sistemas baseados em arquétipos, tais como os baseados em openEHR. Essa tecnologia pode ajudar a encontrar o arquétipo apropriado para uma situação específica ou selecionar os arquétipos a serem usados na construção de um sistema específico. A classificação e verificação de consistência do conteúdo também podem ser realizadas de maneira mais eficiente e eficaz do que usando somente uma ADL (*Archetype Definition Language*), o que traria limitações de possibilidades de raciocínio e disponibilidade de ferramentas para gerenciar conteúdo. O uso de tecnologias semânticas, tais como técnicas de medição de similaridade semântica (disponíveis na comunidade da Web Semântica), permitem que um sistema específico possa receber arquétipos desconhecidos, que devem ser classificados na biblioteca de arquétipos específica. (Huang et al., 2017).

## 3. Objetivos do Projeto

A obtenção de PEPs semânticos mais avançados, que possibilitem prover desambiguação, estrutura e qualidade de dados (sejam eles textuais, numéricos ou de imagens de exames clínicos) para pesquisa por tratamentos médicos (incluindo preparação para análise futura por algoritmos de aprendizagem de máquina) é o principal objetivo deste projeto. Para tanto, a proposta prevê etapas a serem cumpridas na sequência:

1. Propor e descrever um modelo de PEP utilizando padrões abertos para domínio de saúde piloto (com foco primário em oncologia pediátrica).
2. Propor e descrever um modelo semântico para PEP, considerando o uso de ontologias, *thesauri* e vocabulários internacionais, de modo a propiciar desambiguação e precisão dos dados clínicos armazenados.

3. Elaborar modelo de interação para uso dos PEPs em ambiente clínico por profissionais de saúde.
4. *Design* de solução de recuperação e visualização de informações em PEPs semânticos.
5. Desenvolvimento de protótipos e validação empírica em domínio de saúde piloto (e.g., oncologia pediátrica).

De acordo com os tipos de informações disponíveis nas bases de dados de prontuários eletrônicos que serão estudados adiante, teremos diversas possibilidades de aplicações, tais como visualização de dados, rotulação de informações e classificação de tumores. Ou seja, a disponibilidade dos dados é o principal aspecto que determinará o desenvolvimento de aplicações e as técnicas de IA empregadas (aprendizado de máquina, *deep learning*, visão computacional e outras técnicas correlatas).

As possibilidades de aplicação em tecnologias e *web* semânticas e ciência de dados podem ser extrapoladas para além da oncologia pediátrica. Em teoria, quaisquer bases de dados estruturados com padrões bem definidos para PEPs seriam de utilidade para a área de saúde.

#### **4. Resultados Parciais**

Foi executada uma revisão integrativa composta de três revisões que identificaram e analisam a pesquisa e uso de padrões e KOS para especificação semântica de PEPs para a área de oncologia pediátrica. Os resultados obtidos foram compilados na elaboração de um artigo submetido à revista científica *Journal of Health Informatics* (Figueiredo et al. 2002).

De acordo com os resultados das três revisões realizadas, no que concerne aos aspectos-chave na descrição semântica, a maioria dos trabalhos revisados não destacaram o padrão internacional de PEP utilizado. Isso aponta que, apesar dos esforços para a adoção desses padrões na literatura, ainda existem projetos que não adotam ou não dão atenção ao uso de padrões. É importante ressaltar, no entanto, que há carência de trabalhos que mostrem uma visão mais abrangente envolvendo o uso de padrões internacionais, diferentes KOS e tecnologias recentes na área de interoperabilidade semântica. Sendo este um ponto chave em uma agenda de pesquisa futura na área.

Quando analisados trabalhos mais focados em KOS sobre oncologia pediátrica, foi possível observar um maior uso de padrões internacionais. Os tipos de câncer infantil, citados no trabalho, reforçam a importância das características relacionadas à genética e à hereditariedade ou são relacionados ao sangue e ao sistema imunológico.

Do ponto de vista tecnológico, três trabalhos abordam linguagens e ferramentas ligadas ao aprendizado de máquina para predição e identificação de informações relevantes. Entretanto, destaca-se que estes estudos não abordam a ligação destes KOS com PEPs para oncologia pediátrica, bem como a integração com outros KOS em PEPs interoperáveis.

Com base nos resultados, foi identificada a necessidade de novas pesquisas e aprimoramento na especificação de PEPs que integrem KOS com padrões internacionais. Assim, como próximos passos desta pesquisa, é proposto a pesquisa por processo e modelo para integração entre padrões para especificação de PEPs com os KOS mais utilizados para oncologia pediátrica.

## Referências

**BABBOTT, S., MANWELL, L. B., BROWN, R., MONTAGUE, E., WILLIAMS, E., SCHWARTZ, M., HESS, E., & LINZER, M.** *Electronic medical records and physician stress in primary care: Results from the MEMO Study*. Journal of the American Medical Informatics Association, 21(E2), 100–106, 2014. <https://doi.org/10.1136/amiajnl-2013-001875>

**CFM, CONSELHO FEDERAL DE MEDICINA.** *Resolução CFM nº 1.638/2002*, 2002. Disponível em: <https://sistemas.cfm.org.br/normas/visualizar/resolucoes/BR/2002/1638> (acesso Julho 2020).

**CRUZ, J. A. S.** *Prontuário Eletrônico de Pacientes (PEP): Políticas e Requisitos Necessários à Implantação no Hospital Universitário de Santa Maria (HUSM)*. I Seminário de Tecnologia da Informação em Saúde, 2011. <http://repositorio.ufsm.br/handle/1/10976>

**FIGUEIREDO, E. B.** *BioFrame: um framework semântico para prontuários eletrônicos em Oncologia*. Dissertação de Mestrado. UNIFACCAMP - Centro Universitário Campo Limpo Paulista, 2022. <https://www.cc.faccamp.br/Dissertacoes/ElaineBarbosaFigueiredo.pdf>

**FIGUEIREDO, E. B., BONACIN, R., ROSA, F., ZANETTI, R. A.** *Semantics in EHR for pediatric oncology: an integrative review*. Journal of Health Informatics. 2022. (Manuscrito submetido em avaliação)

**HÄYRINEN, K., SARANTO, K., & NYKÄNEN, P.** *Definition, structure, content, use and impacts of electronic health records: A review of the research literature*. International Journal of Medical Informatics, 77(5), 291–304, 2008. <https://doi.org/10.1016/j.ijmedinf.2007.09.001>

**HUANG, E. W., ZHANG, R., LIU, J., WANG, S., LIU, B., ZHOU, X., ZHAI, C., LI, B., ZHANG, R., & LIN, H.** *Hemnet: Integration of electronic medical records with molecular interaction networks and domain knowledge for survival analysis*. ACM-BCB 2017 - Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics, 378–387, 2017. <https://doi.org/10.1145/3107411.3107422>