

AQUISIÇÃO E REPRESENTAÇÃO DE CONHECIMENTO SOBRE CIRURGIA BARIÁTRICA – FASE 2

Vinícius Kenzo Fukace (PIBIC/CNPq/FA/UEM), Heloise Manica Paris Teixeira (Orientadora), Raqueline Ritter de Moura Penteadó (Coorientadora), Samuel Lincon Batista Bispo (Co-autor), e-mail: ra115672@uem.br

Universidade Estadual de Maringá / Centro de Tecnologia / Maringá, PR.

Área e sub-área do conhecimento: Ciência da Computação, Sistemas de Informação

Palavras-chave: Ontologia, Integração de Dados, Dados Abertos SUS

Resumo:

Existem diferentes técnicas computacionais utilizadas para analisar dados clínicos e de gestão sobre a cirurgia bariátrica, permitindo assim sua avaliação e tomada de decisão por especialistas e gestores em saúde. Para que sejam utilizados em um sistema computacional, o conhecimento sobre as cirurgias realizadas, medicamentos, dados clínicos e de gestão devem estar integrados e expressos de modo que sistemas computacionais possam processá-los. Neste sentido, o objetivo deste projeto foi aplicar a ontologia desenvolvida na primeira fase, por meio de um estudo de caso sobre execução de consultas e integração de dados abertos do Sistema Único de Saúde. Os resultados obtidos no estudo indicam que as ferramentas e técnicas estudadas podem ser utilizadas, porém, exigem conhecimento técnico, o que dificulta sua adoção por usuários leigos.

Introdução

Existem diversas técnicas computacionais utilizadas para analisar dados clínicos e de gestão sobre a cirurgia bariátrica que auxiliam sua avaliação. Para que sejam utilizados por um sistema computacional, o conhecimento sobre as cirurgias realizadas, medicamentos, dados clínicos e de gestão devem estar integrados e expressos de modo que sistemas computacionais possam processá-los. Uma abordagem para representar o conhecimento em linguagem que possa ser processada por máquinas são ontologias. Conforme Almeida e Bax (2003), uma ontologia define regras que regulam a combinação entre conceitos (organizados em uma taxonomia) e suas relações (i. é, o tipo de interação entre os conceitos) em um domínio do conhecimento e permite aos usuários formularem consultas a instâncias (elementos específicos, ou seja, os próprios dados) usando conceitos definidos pela ontologia.

A fase 1 da pesquisa (SILVA; TONON; TEIXEIRA, 2021) teve como objetivo principal extrair conhecimento na área de cirurgia bariátrica e estruturá-lo em uma ontologia no domínio. A fase 2 teve como objetivo principal aplicar a ontologia desenvolvida na fase 1 em um estudo na área de integração de dados.

O presente estudo explorou o modelo ontológico por meio de execução de consultas que requerem integração de diferentes bases de dados. Para isso, utilizou-se a ferramenta de edição de ontologias Protégé¹ com o plug-in Ontop², que agrega a funcionalidade de tradução de consultas SPARQL, realizadas utilizando uma ontologia para consultas em bases de dados relacionais.

Materiais e Métodos

Para a definição das consultas, primeiro foi realizado um estudo sobre as plataformas de dados abertos públicos para definir o escopo da aplicação na integração de base de dados. Em seguida realizou-se a seleção e tratamento dos dados abertos do SUS que serviram para povoar a ontologia. Posteriormente, foram implementadas regras de mapeamento entre as bases utilizando as ferramentas Ontop e Protégé.

Neste estudo utilizou-se a ontologia desenvolvida na fase 1 do projeto (SILVA; TONON; TEIXEIRA, 2021), que foi complementada no trabalho de Nunes (2021), criando a Ontology for Bariatric Surgery (OBaS). Elementos adicionais foram acrescentados à OBaS conforme necessário, a fim de adequar a ontologia às questões realizadas nesta pesquisa.

Em relação às fontes de dados, informações sobre cirurgia bariátrica foram obtidas nos sistemas de informação do Departamento de Informática do SUS (DATASUS³), em arquivos denominados ABO e RD. Arquivos ABO contêm dados sobre laudos de APAC (Autorização de Procedimento de Alta Complexidade) referentes a atendimentos ambulatoriais de acompanhamento a cirurgia bariátrica. Arquivos RD possuem dados a respeito de registros de AIHs (Autorizações de Internação Hospitalar) de todos os atendimentos provenientes de internações hospitalares financiados pelo SUS, incluindo aqueles relacionados a cirurgias bariátricas.

Utilizando o TabWin⁴, ferramenta desenvolvida pelo Datasus, foi feita a conversão dos arquivos (disponibilizados no formato .dbc) para sql. Em seguida, foram importados no MySQL⁵, um sistema de gerenciamento de banco de dados (SGBD) relacional suportado pelo Ontop, e bases de dados nomeadas ABO e RD foram criadas e populadas utilizando os respectivos arquivos. Finalmente, a conexão entre o Ontop e o MySQL foi estabelecida utilizando a ferramenta Protégé.

Os arquivos escolhidos para este trabalho contêm dados referentes ao estado do Paraná no período do mês de Março de 2020, este período foi escolhido para evitar possíveis inconsistências nos dados devido ao retardo de notificação.

Com base no conteúdo das fontes de dados, duas questões foram elaboradas para explorar o processo de integração das bases ABO e RD. Questão 1: “As comorbidades dos pacientes implicam no custo da cirurgia bariátrica?”, e Questão 2: “A prática de atividades físicas e alimentação saudável pelos pacientes implica no custo da cirurgia bariátrica?”.

¹ <https://protege.stanford.edu/>

² <https://ontop-vkg.org/>

³ <https://datasus.saude.gov.br/>

⁴ <https://datasus.saude.gov.br/transferencia-de-arquivos/>

⁵ <https://www.mysql.com/>

No sistema Ontop, foram implementados onze mapeamentos entre as bases de dados e a ontologia, a fim de responder às questões previamente definidas. A criação de um mapeamento necessita da definição de uma fonte, que é uma consulta SQL na base de dados, e de um alvo, que define a tripla RDF que liga a fonte e a ontologia.

Resultados e Discussão

As consultas 1 e 2 (Figuras 1 e 2) foram elaboradas e executadas na linguagem SPARQL por meio da interface do Protégé com o Ontop.

```
SPARQL query editor:
SPARQL Query
PREFIX : <https://github.com/glaubernunes/ontology-for-bariatric-surgery/raw/main/ontology-for-bariatric-surgery.owl#>
SELECT ?internacao ?pessoa ?comorbidade ?valor
WHERE {
  ?internacao :temPessoa ?pessoa .
  ?internacao :custo ?valor .
  OPTIONAL {?pessoa :OBAS_0000172 ?comorbidade}
}
```

Figura 1 – Definição da consulta SPARQL relativa à primeira questão utilizando a interface Protégé.

```
SPARQL query editor:
SPARQL Query
PREFIX : <https://github.com/glaubernunes/ontology-for-bariatric-surgery/raw/main/ontology-for-bariatric-surgery.owl#>
SELECT ?internacao ?pessoa ?habito ?valor
WHERE {
  ?internacao :temPessoa ?pessoa .
  ?internacao :custo ?valor .
  OPTIONAL {?pessoa :OBAS_0000173 ?habito}
}
```

Figura 2 – Definição da consulta SPARQL relativa à segunda questão utilizando a interface Protégé.

A estrutura das consultas ilustradas nas Figuras 1 e 2 diferem no conteúdo presente no campo “OPTIONAL”. Ambas buscam registros de AIH que estejam ligados ao CNS de uma pessoa e que tenham um custo definido. Se a pessoa tem comorbidades, elas aparecem nos resultados da primeira consulta, e se ela tem hábitos de alimentação ou atividade física, estes aparecem na segunda. O Ontop realiza a tradução das consultas SPARQL em consultas SQL, que são aplicadas nas bases de dados relacionais do MySQL. Em seguida, a consulta SQL é processada e o retorno é enviado ao Ontop e apresentado ao usuário. No total, foram obtidos 28 resultados na consulta 1 e 44 na consulta 2. Uma análise mais detalhada das ferramentas utilizadas e resultados obtidos foi apresentada em (Bispo et al., 2021). A Figura 3 ilustra parte do resultado da consulta 2 na interface do Protégé, que retorna o número da AIH no campo “internacao”; o CNS, em “pessoa”; hábitos do paciente (se houver); e o valor da internação.

Execution time: 95ms. Solution mappings returned: 44.

internacao	pessoa	habito	
<https://github.com/glaubernunes/ontology-for-bariatric-surgery...>	<https://github.com/glaubernunes/ontology-for-bariatric-surger...>	<https://github.com/glaubernunes/ontology-for-bariatric-surger...>	"6388.40" ^{^^}
<https://github.com/glaubernunes/ontology-for-bariatric-surger...>	<https://github.com/glaubernunes/ontology-for-bariatric-surger...>	<https://github.com/glaubernunes/ontology-for-bariatric-surger...>	"5927.70" ^{^^}
<https://github.com/glaubernunes/ontology-for-bariatric-surger...>	<https://github.com/glaubernunes/ontology-for-bariatric-surger...>	<https://github.com/glaubernunes/ontology-for-bariatric-surger...>	"6018.09" ^{^^}
<https://github.com/glaubernunes/ontology-for-bariatric-surger...>	<https://github.com/glaubernunes/ontology-for-bariatric-surger...>	<https://github.com/glaubernunes/ontology-for-bariatric-surger...>	"5941.80" ^{^^}
<https://github.com/glaubernunes/ontology-for-bariatric-surger...>	<https://github.com/glaubernunes/ontology-for-bariatric-surger...>	<https://github.com/glaubernunes/ontology-for-bariatric-surger...>	"5919.02" ^{^^}
<https://github.com/glaubernunes/ontology-for-bariatric-surger...>	<https://github.com/glaubernunes/ontology-for-bariatric-surger...>	<https://github.com/glaubernunes/ontology-for-bariatric-surger...>	"5948.15" ^{^^}
<https://github.com/glaubernunes/ontology-for-bariatric-surger...>	<https://github.com/glaubernunes/ontology-for-bariatric-surger...>	<https://github.com/glaubernunes/ontology-for-bariatric-surger...>	"6207.74" ^{^^}

Figura 3 – Parte do resultado da consulta 2 na interface do Protégé.

Em relação à integração das bases de dados, o CNS, hábitos e comorbidades do paciente pertencem à base ABO, o valor da internação à base RD, e o número de AIH está presente em ambas as bases, servindo como chave para a junção das duas. Os campos internação, pessoa e hábito aparentam retornar resultados repetidos, porém, o trecho acima somente ilustra o prefixo dos resultados, devido a restrições de espaço na imagem. Ao observar os números de CNS, é possível perceber que foram criptografados, de modo a proteger a identidade dos pacientes.

Conclusões

O presente estudo descreveu as etapas do processo de integração e execução de consultas em dados abertos sobre cirurgia bariátrica com o uso de uma ontologia no domínio. Conforme exposto, é possível a execução das consultas com as ferramentas estudadas, porém, se faz necessário conhecimento técnico sobre elas bem como sobre as bases de dados abertos e ontologias, inviabilizando o uso direto por usuários leigos. Como trabalho futuro, sugere-se um estudo com a execução de consultas mais complexas, observando-se seu desempenho com a utilização das ferramentas estudadas e com outros sistemas de bancos de dados.

Agradecimentos

Ao programa PIBIC/CNPq-FA-UEM pela concessão de bolsa de Iniciação Científica, e às Professoras Orientadoras.

Referências

ALMEIDA, M. B.; BAX, M. P. Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. **Ciência da Informação**, v. 32, n. 3, 2003.

NUNES, G. M. O.; BERARDI, R. C. G. **Multionto: método de construção de ontologia considerando heterogeneidade de fontes e tipos de conhecimentos- um estudo de caso sobre cirurgia bariátrica**. 2021. Dissertação de Mestrado. Universidade Tecnológica Federal do Paraná.

SILVA, J. E. C. de; TONON, M. M.; TEIXEIRA, H. M. P. Aquisição e Representação do Conhecimento Científico sobre Cirurgia Bariátrica - Fase 1. In: **30º Encontro Anual de Iniciação Científica (EAIC)**, Maringá: Universidade Estadual de Maringá, 2021.

BISPO, S. L.; FUKACE, V. K.; MAZUR, A. H.; PENTEADO, R. R. M.; TEIXEIRA, H. M. P. Integração de Dados Abertos em Saúde com o modelo OBDA: Um Estudo de Caso na Área de Cirurgia Bariátrica. In: **Anais do 22º Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS 2022)**, 2022.