

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ  
DAINF - DEPARTAMENTO ACADÊMICO DE INFORMÁTICA  
CURSO DE BACHARELADO EM SISTEMAS DE INFORMAÇÃO

RODRIGO GIÁCOMO MORONI DE SOUZA

**INTEGRAÇÃO DE DADOS PARA AUXÍLIO À TOMADA  
DE DECISÃO EM MATERNIDADES**

TRABALHO DE CONCLUSÃO DE CURSO

CURITIBA  
2020

RODRIGO GIÁCOMO MORONI DE SOUZA

**INTEGRAÇÃO DE DADOS PARA AUXÍLIO À TOMADA  
DE DECISÃO EM MATERNIDADES**

Proposta de Trabalho de Conclusão de Curso apresentado ao Curso de Bacharelado em Sistemas de Informação da Universidade Tecnológica Federal do Paraná, como requisito parcial para a obtenção do título de Bacharel.

Orientador: Prof<sup>a</sup> Dr<sup>a</sup> Rita Cristina Galarraga Bernardi  
DAINF - Departamento Acadêmico de Informática -UTFPR

CURITIBA  
2020

**RODRIGO GIÁCOMO MORONI DE SOUZA**

**INTEGRAÇÃO DE DADOS PARA AUXÍLIO À TOMADA DE DECISÃO EM  
MATERNIDADES**

Trabalho de Conclusão de Curso de Graduação  
apresentado como requisito para obtenção do título  
de Bacharel em Sistemas de Informação da  
Universidade Tecnológica Federal do Paraná  
(UTFPR).

Data de aprovação: 09/Outubro/2020

---

ADRIANA MARIA WAN STADNIK

Doutorado

Universidade Tecnológica Federal do Paraná (UTFPR)

---

SÍLVIA AMÉLIA BIM

Doutorado

Universidade Tecnológica Federal do Paraná (UTFPR)

---

RITA CRISTINA GALARRAGA BERARDI

Doutorado

Universidade Tecnológica Federal do Paraná (UTFPR)

**CURITIBA**

**2020**

## AGRADECIMENTOS

Agradeço a todas pessoas que contribuíram durante minha trajetória na graduação, familiares, colegas de curso e professores. Considerações especiais para Viviane Moroni e prof Rita Berardi.

*Art. 196. A saúde é direito de todos e dever do Estado, garantido mediante políticas sociais e econômicas que visem à redução do risco de doença e de outros agravos e ao acesso universal e igualitário às ações e serviços para sua promoção, proteção e recuperação. (Constituição da República Federativa do Brasil de 1988)*

## RESUMO

MORONI DE SOUZA, Rodrigo Giácomo. Integração de Dados para Auxílio à Tomada de Decisão em Maternidades. 2020. 69 f. Trabalho de Conclusão de Curso – Curso de Bacharelado em Sistemas de Informação, Universidade Tecnológica Federal do Paraná. Curitiba, 2020.

Na área da saúde, uma tomada de decisão deve se apoiar, essencialmente, em informações precisas e completas. Isto se deve ao caráter sensível quando à vida humana destas decisões. Por isso, muitos fatores, como uma boa contextualização sobre os pacientes, são fundamentais na gestão da informação de uma unidade de saúde. Neste sentido, este trabalho buscou desenvolver um modelo ontológico para uma unidade de saúde, com intuito de tornar sua base de dados semanticamente integrável a bases relacionadas com esta. Isto possibilitaria aprimorar recursos como recuperação de informação, obtenção de indicadores sobre a execução de protocolos, redução de problemas semânticos, entre outros. Alguns registros dos dados abertos de Curitiba foram utilizados como fonte de dados para esta ontologia. Paralelamente, outra ontologia foi elaborada para representar os dados abertos de Curitiba, para que fosse possível testar a integração entre ambas. Em seguida, algumas consultas foram elaboradas para avaliar seu funcionamento. Após os testes com a ontologia, alguns resultados positivos puderam ser notados, como melhorias na recuperação de informação e redução de problemas semânticos.

**Palavras-chave:** Ontologia. Saúde. Dados Abertos.

## ABSTRACT

MORONI DE SOUZA, Rodrigo Giácomo. Data Integration to Assist Decision Making in Maternity Hospitals. 2020. 69 f. Trabalho de Conclusão de Curso – Curso de Bacharelado em Sistemas de Informação, Universidade Tecnológica Federal do Paraná. Curitiba, 2020.

In the health area, decision-making must essentially be based on accurate and complete information. Due to the sensitive nature of human life on these decisions. Therefore, many factors, such as a good contextualization about patients, are fundamental in the management of information in a health unit. In this way, this work sought to develop an ontological model for a health unit, in order to make its database semantically integrable with related databases. This would enable to improve resources such as information retrieval, obtaining indicators on protocols execution, semantic problems reduction and others. Some records of open data from Curitiba were used as data source for this ontology. In parallel, another ontology was elaborated to represent the open data from Curitiba, so that it was possible to test the integration between both. Then, some queries were elaborated to evaluate its functioning. After testing the ontology, some positive results were noted, such as improvements in information retrieval and semantic problems reduction.

**Keywords:** Ontology. Health. Open Data.

## LISTA DE FIGURAS

Figura 1 – Topologia de rede descentralizada em um hospital. . . . .	15
Figura 2 – Porcentagem de artigos recuperados por base de dados bibliográficos. .	19
Figura 3 – Exemplo de setores hospitalares com bases de dados separadas. . . . .	20
Figura 4 – Exemplo de hospitais com bases heterogêneas. . . . .	22
Figura 5 – Exemplo de relação é-um em ontologias. . . . .	27
Figura 6 – Exemplo de relação parte-de em ontologias. . . . .	28
Figura 7 – Exemplo de ontologia gráfica. . . . .	29
Figura 8 – Exemplo de relacionamento em forma de triplas. . . . .	29
Figura 9 – Exemplo de RDF/RDF-S. . . . .	30
Figura 10 – Estimativa de crescimento do volume de dados digitais de 2010 a 2020.	31
Figura 11 – Camadas da <i>Web Semântica</i> . . . . .	32
Figura 12 – 5 Estrelas dos Dados Abertos Conectados. . . . .	34
Figura 13 – <i>Ontology Development 101</i> . . . . .	36
Figura 14 – Trajetória do FACR na unidade de saúde. . . . .	38
Figura 15 – Informações preenchidas em cada setor. . . . .	39
Figura 16 – Cenário com a Unidade de Saúde, separado por setores. . . . .	40
Figura 17 – Hierarquia das classes. . . . .	44
Figura 18 – Cardinalidades para a classe <i>RegistroHospitalar</i> . . . . .	46
Figura 19 – Consulta para QC4. . . . .	53
Figura 20 – Resultado da Consulta para QC4. . . . .	54
Figura 21 – Consulta para QC5. . . . .	54
Figura 22 – Resultado da Consulta para QC5. . . . .	54
Figura 23 – Consulta para QC3. . . . .	55
Figura 24 – Resultado da Consulta para QC3. . . . .	55
Figura 25 – Consulta para QC6. . . . .	56
Figura 26 – Resultado da Consulta para QC6. . . . .	56
Figura 27 – Consulta para QC1. . . . .	57
Figura 28 – Resultado da Consulta para QC1. . . . .	57
Figura 29 – Consulta para QC2. . . . .	57
Figura 30 – Resultado da Consulta para QC2. . . . .	57



## LISTA DE QUADROS

Quadro 1 – Questões de competência levantadas e setores do cenário envolvidos. .	41
Quadro 2 – Questões de competência relacionadas com necessidade de informação.	41
Quadro 3 – Propriedades de objeto. . . . .	45
Quadro 4 – Restrições de propriedades de objeto. . . . .	46
Quadro 5 – Restrições de tipo de dado e domínio. . . . .	47

## LISTA DE ABREVIATURAS E SIGLAS

ACR	Acolhimento com Classificação de Risco
CSV	<i>Comma-separated Values</i> , do inglês, valores separados por vírgula.
CID	Classificação Internacional de Doenças e Problemas Relacionados à Saúde
CNS	Cartão Nacional de Saúde
COREN	Conselho Regional de Enfermagem
CRM	Conselho Regional de Medicina
DPP	Data Provável do Parto
DUM	Data da Última Menstruação
FACR	Formulário de Acolhimento com Classificação de Risco
HC	Hospital de Clínicas
HL7	<i>Health Level 7</i> , do inglês, Camada de Saúde 7.
HTML	<i>Hypertext Marker Language</i> , do inglês, linguagem de marcação de hipertexto.
JSON	<i>JavaScript Object Notation</i> , do inglês, notação de objeto JavaScript.
OWL	<i>Ontology Web Language</i> , do inglês, linguagem de ontologia.
OMG	<i>Object Management Group</i> , do inglês, grupo de gerenciamento de objetos.
PDF	<i>Portable Document Format</i> , do inglês, formato de documento portátil.
RDF	<i>Resource Description Framework</i> , do inglês, estrutura de descrição de recursos.
RDF-S	<i>Resource Description Framework Schema</i> , do inglês, esquema para estrutura de descrição de recursos.
SPARQL	<i>Sparql Protocol and RDF Query Language</i> , do inglês, linguagem de consulta RDF.
SQL	<i>Structured Query Language</i> , do inglês, linguagem de consulta estruturada.

SWRL	<i>Semantic Web Rule Language</i> , do inglês, linguagem de regras semânticas.
UTFPR	Universidade Tecnológica Federal do Paraná
UFPR	Universidade Federal do Paraná
URI	<i>Universal Resource Identifier</i> , do inglês, identificador de recurso universal.
W3C	<i>World Wide Web Consortium</i> , do inglês, consórcio da rede mundial de computadores.
WWW	<i>World Wide Web</i> , do inglês, rede mundial de computadores.
WSMO	<i>Web Service Modeling Ontology</i> , do inglês, ontologia para modelagem de serviço <i>web</i> .
XLSX	<i>Microsoft Excel Open XML Spreadsheet</i> , do inglês, planilha <i>excel</i> em XML aberto.
XML	<i>Extensible Markup Language</i> , do inglês, linguagem de marcação extensível.

## SUMÁRIO

<b>1 – Introdução</b>	<b>14</b>
1.1 Contexto do Trabalho	16
1.2 Objetivo Geral	16
1.3 Objetivos Específicos	16
1.4 Organização do Documento	16
<b>2 – Trabalhos Correlatos</b>	<b>18</b>
2.1 Protocolo	18
2.2 Resultados	19
2.2.1 Ontologias em Sistemas de Informação em Hospitais	20
2.2.2 Ontologias para integração de dados entre unidades de saúde	21
2.2.3 Ontologias para integração com dados abertos	24
2.3 Considerações	25
<b>3 – Referencial Teórico</b>	<b>26</b>
3.1 Ontologia	26
3.2 Dados Abertos	31
3.2.1 Dados Abertos Conectados	33
<b>4 – Metodologia</b>	<b>35</b>
4.1 Modelagem da Ontologia - <i>Ontology Development 101</i>	35
4.2 Ferramentas Utilizadas	37
<b>5 – Resultados</b>	<b>38</b>
5.1 Modelagem da Ontologia	38
5.1.1 Determinação do Escopo	38
5.1.2 Considerações de Reuso	40
5.1.3 Enumeração dos Termos	41
5.1.4 Definição das Classes	43
5.1.5 Definição das Propriedades	43
5.1.6 Definição das Restrições	45
5.1.7 Criação das Instâncias	45
5.2 Validação da Ontologia	53
5.2.1 Primeiro Cenário	53
5.2.2 Segundo Cenário	55
5.2.3 Terceiro Cenário	56

6 – Considerações Finais . . . . .	58
Referências . . . . .	59
<b>Apêndices</b>	<b>63</b>
APÊNDICE A – <i>Script</i> para Filtrar Registros nos Dados Abertos de Curitiba . . . . .	64
A.1 Código-fonte em Python . . . . .	64
<b>Anexos</b>	<b>67</b>
ANEXO A – Formulário de Acolhimento com Classificação de Risco . . .	68

## 1 Introdução

Nas organizações contemporâneas, os Sistemas de Informação são componentes estratégicos, pois possibilitam análise de cenários, melhoram o apoio à tomada de decisão e contribuem para definição e implementação de novas estratégias organizacionais (SBC, 2018). No que tange a tomada de decisões, a carência de ferramentas auxiliares, bem como informações incompletas, desatualizadas ou ausentes podem influenciar, de forma negativa, a conduta de profissionais da saúde. Por isso, em unidades de saúde os Sistemas de Informação também se mostram fundamentais (MURPHY; REDDY, 2014).

A precisão na escolha e execução de procedimentos é de suma importância no campo da saúde, levando em conta seus objetivos e diretrizes. Segundo Costa e Stadnik (2019), a ausência de um Sistema de Informação adequado em um pronto atendimento permite o surgimento de inúmeros problemas, tais como:

- Dificuldade na obtenção históricos médicos de pacientes;
- Limitação na troca de informações entre diferentes setores;
- Ausência de mecanismos para controle de cumprimento do ACR - Acolhimento com Classificação de Risco.

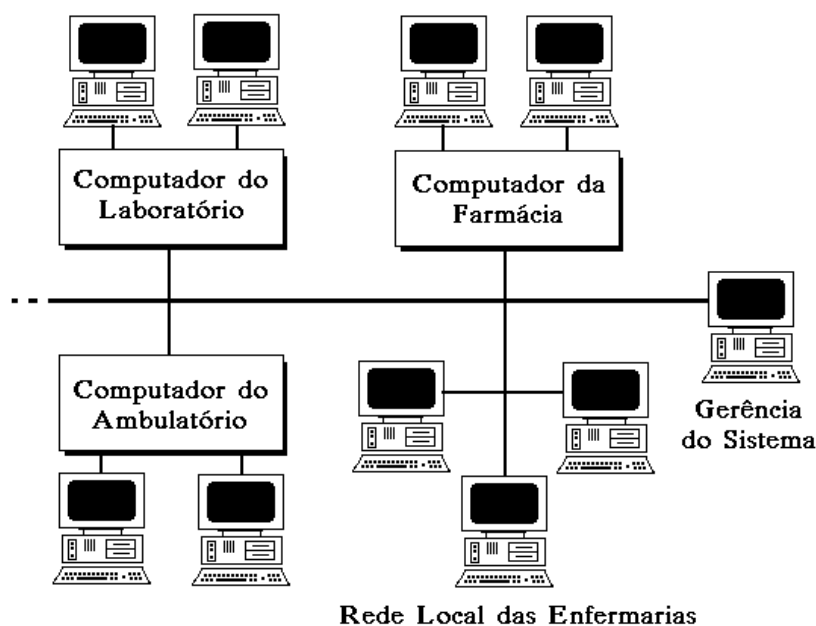
A não constatação da evolução do fator de risco inicialmente classificado no protocolo ACR é um erro comum decorrente dos fatores listados anteriormente, o que afeta a execução do próprio protocolo (BRASIL, 2009). Ademais, a ausência de um Sistema de Informação pode tornar o processo de gerenciamento muito mais custoso e impreciso. Com isso, pode haver uma complicação para gerência tanto da unidade como da rede de saúde, por não estarem estabelecidos os mecanismos adequados de apoio à tomada de decisão (COSTA; STADNIK, 2019).

A limitação na troca de informações entre diferentes setores, dentre outros fatores, está ligada à diferença de ordem semântica de cada um deles. Dessa maneira, na falta de mecanismos apropriados, surge uma limitação na comunicação intersetorial. Um exemplo disso é um Sistema de Informação Hospitalar moderno que, tipicamente, possui uma topologia descentralizada, portanto, uma separação semântica entre os setores, como ilustrado na Figura 1 (JOHANSTON, 1993).

Durante a listagem dos sintomas de um paciente, é possível perceber como a diferença semântica afeta a troca de informações entre os setores. Nessa etapa, o paciente irá (em diferentes momentos) relatar para a enfermagem e a equipe médica seus sintomas, que deverão ser registrados. É possível que no registro feito pela enfermagem, esse relato seja denominado apenas como relato do paciente, enquanto a equipe médica pode nomear como queixa.

Para resolução desse tipo de problema, umas das abordagens possíveis é uma camada lógica que atenua as diferenças semânticas das bases de dados, portanto, assegura

Figura 1 – Topologia de rede descentralizada em um hospital.



Fonte: Johanston (1993)

um meio em que todos setores interajam da mesma forma, podendo também ser um meio para troca de informações (RIANO et al., 2012). Uma hipótese nesse caso é um modelo ontológico que minimize a diferença semântica entre eles. A utilização de uma ontologia para este propósito é plausível, pois é um modelo que permite a representação formal de um domínio, compreensível por humanos e computadores. Além disso, é um modelo recomendado para aplicações que trabalham com bases de conhecimento (ISOTANI; BITTENCOURT, 2015).

Esse modelo possibilita, também, a integração de diferentes bases de dados. Isso poderia propiciar à unidade de saúde buscar informações, como históricos médicos de seus pacientes, em outras unidades de saúde, o que ajudaria a suprir um dos problemas levantados. Essa recuperação de informações poderia ocorrer de forma transparente para a unidade de saúde, de forma similar a uma consulta local, evitando empecilhos na troca de dados. Logo, o modelo pode agregar informações para serem empregadas no gerenciamento da unidade de saúde, possibilitando uma análise mais holística sobre a rede de saúde (PATRAO et al., 2013).

Uma das bases de dados hospitalares disponíveis, que poderia ter sua integração facilitada com o modelo proposto, são os dados abertos de Curitiba. A história dessa base de dados é fortemente vinculada com a Lei de Acesso à Informação, sancionada em 2011. Após esta entrar em vigor, diversos órgãos governamentais brasileiros passaram a disponibilizar, na rede mundial de computadores, documentos, dados e informações de domínio público para livre utilização, conforme prevê a legislação (BRASIL, 2011;

BRASIL, 2019). Consequentemente, a prefeitura de Curitiba passou a disponibilizar dados municipais de várias áreas, entre elas: segurança, saúde e transporte (CURITIBA, 2019).

## 1.1 Contexto do Trabalho

O desenvolvimento deste trabalho irá utilizar como referência um Pronto Atendimento Tocoginecologia, medicina especializada no sistema reprodutor feminino, de uma Maternidade de Alto Risco (HC/UFPR). Esta unidade de saúde também foi estudada em uma dissertação de mestrado do Programa de Pós-Graduação em Engenharia Biomédica da UTFPR (COSTA et al., 2020), que tinha por objetivo propor uma modelagem de um Sistema de Informação para gestão da informação na unidade. Com intuito de complementar os estudos da dissertação apresentada, foi elaborada uma proposta para o desenvolvimento de um modelo ontológico afim de tornar a base de dados desse Sistema de Informação semanticamente integráveis com outras bases, por exemplo, os dados abertos de Curitiba. Essa proposta caracteriza o objetivo principal deste trabalho. Desta forma, alguns dos resultados obtidos na dissertação serviram como base para o desenvolvimento deste.

Para o desenvolvimento do trabalho, devemos considerar a estrutura do nosso objeto de estudo. Neste sentido, a unidade de saúde sendo um pronto atendimento, trata situações urgentes que necessitam de uma intervenção médica imediata. A unidade possui três setores, são eles: recepção, enfermagem e equipe médica. A prioridade de atendimento é definida pelo ACR, protocolo que define um tempo máximo de espera pelo atendimento de acordo com critérios preestabelecidos (COSTA et al., 2020).

## 1.2 Objetivo Geral

O objetivo geral deste trabalho é desenvolver um modelo ontológico, capaz de representar a base de dados de um Pronto Atendimento Tocoginecologia de uma Maternidade de Alto Risco, afim de torná-la integrável semanticamente.

## 1.3 Objetivos Específicos

Os objetivos específicos deste trabalho consistem em: Reduzir a diferença semântica das bases de dados dos diferentes setores da unidade de saúde; Possibilitar a integração com outras bases de dados de maternidades de alto risco; Melhorar aspectos de qualidade da informação na obtenção de históricos médicos.

## 1.4 Organização do Documento

Este trabalho está organizado da seguinte forma: O capítulo 2 apresenta os trabalhos correlatos, com análise de trabalhos feitos na mesma área de conhecimento; No



capítulo 3, encontra-se o referencial teórico; No capítulo 4, é apresentada a metodologia aqui utilizada; No capítulo 5, são apresentados os resultados; Por fim, no capítulo 6, encontram-se as considerações finais.

## 2 Trabalhos Correlatos

Realizou-se uma busca por trabalhos que utilizassem ontologias aplicadas na área da Saúde, sem restringir uma área específica. A ideia era encontrar trabalhos que buscassem resolver problemas similares à proposta aqui apresentada, utilizando diversas ferramentas e abordando contextos variados. Para tal, foi feita uma revisão sistemática, que seguiu o protocolo apresentado por Kitchenham e Charters (2007). Segundo este protocolo, a revisão deve consistir em três etapas, sendo elas: Planejamento, Execução e Escrita.

O Planejamento serve para identificar as necessidades da revisão, formular questões de pesquisa e definir quais são os critérios de inclusão e exclusão de trabalhos científicos. Na Execução, é realizada a seleção dos trabalhos segundo os critérios definidos no Planejamento. A Escrita consiste em reportar as conclusões após a execução, respondendo, assim, às questões elaboradas anteriormente.

### 2.1 Protocolo

Para a revisão sistemática, foram selecionadas bases de dados bibliográficos nas áreas de Computação, Saúde e Ontologia. As bases selecionadas foram:

- *IEEE Xplore Digital Library*<sup>1</sup>;
- *Springer*<sup>2</sup>;
- *ACM Digital Library*<sup>3</sup>;
- *Science Direct*<sup>4</sup>;
- Ontobras<sup>5</sup>;
- BVS - Biblioteca Virtual em Saúde<sup>6</sup>.

Três questões elaboradas para pesquisa:

- Ontologia é utilizada na integração semântica de dados em sistemas de informação hospitalares?
- Ontologia é utilizada na integração de dados entre diferentes unidades de saúde?
- Ontologia é utilizada na integração semântica de dados locais de hospitais e dados abertos?

As palavras-chave escolhidas para pesquisa foram: Ontologia, Hospitais, Sistemas de Informação e Dados Abertos. Para execução das buscas, através do mecanismo de pesquisa de cada base de dados bibliográficos, as palavras-chave foram traduzidas para o inglês e arranjadas com os seguintes operadores lógicos: *Ontology AND Hospitals*; *Ontology*

---

<sup>1</sup><https://ieeexplore.ieee.org/Xplore/home.jsp>

<sup>2</sup><https://www.springer.com/br>

<sup>3</sup><https://dl.acm.org/>

<sup>4</sup><https://www.sciencedirect.com/>

<sup>5</sup><http://ceur-ws.org/>

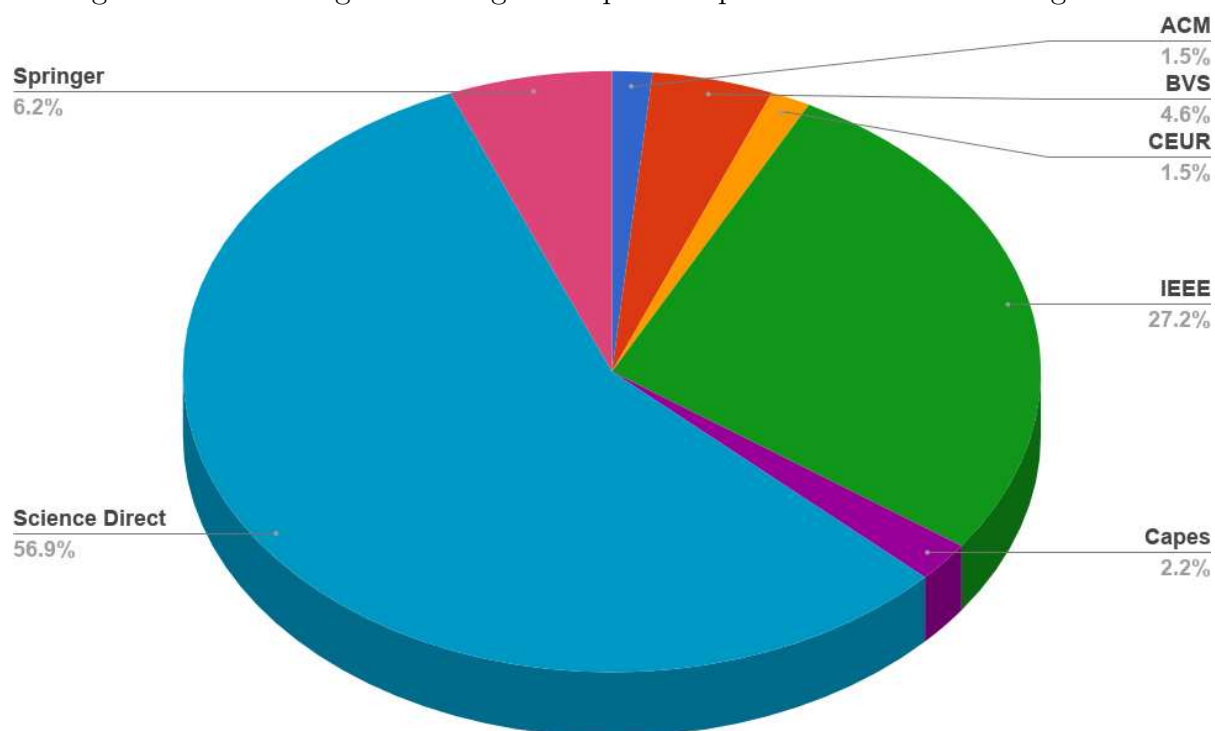
<sup>6</sup><http://brasil.bvs.br/>

*AND Information Systems; Ontology AND Hospitals AND Open Data*. Para seleção dos artigos, consideraram-se apenas publicações disponíveis integralmente *online* e nos idiomas inglês ou português, por questões de domínio linguístico. Além disso, foram considerados apenas artigos publicados entre 2013 e 2019. Os critérios de exclusão utilizados foram respectivamente: título não relacionado ao tema, resumo não relacionado ao tema e artigos redundantes. As buscas foram realizadas em maio de 2020.

## 2.2 Resultados

A partir da aplicação dos parâmetros de pesquisa nas bases, seguindo os critérios de inclusão estabelecidos, foram encontrados 2974 artigos. A Figura 2 representa a quantidade de artigos retornados por base de dados bibliográficos. Observa-se que a base com maior retorno de artigos foi a *Science Direct*. Para evitar o descarte pelo título, o artigo deveria apresentar relação com soluções para sistemas de informações com integração de dados na área da saúde através de ontologias. Após a leitura dos títulos dos trabalhos, foram descartadas 2901 publicações, sendo pré-selecionados 73 artigos.

Figura 2 – Porcentagem de artigos recuperados por base de dados bibliográficos.



Fonte: Autoria própria

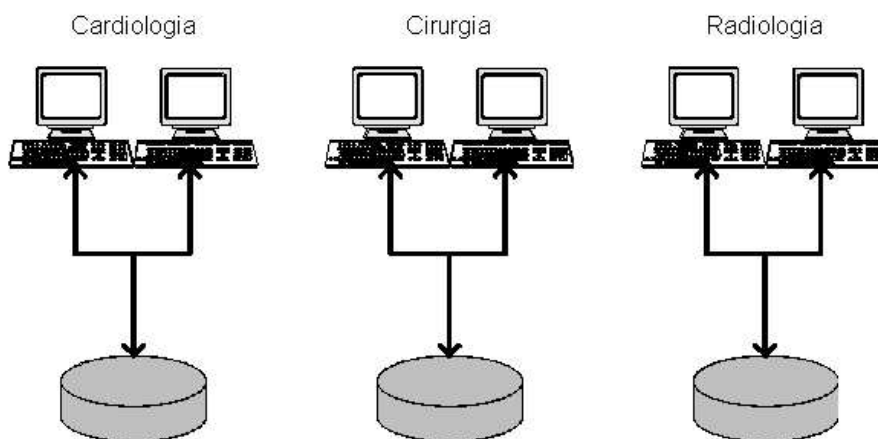
A partir da leitura do resumo dos artigos pré-selecionados, foram descartadas 56 publicações, selecionando, assim, 17 trabalhos que apresentaram relação direta com o tema buscado. Em seguida, aplicou-se o critério de exclusão de artigos redundantes, porém, não houveram descartes nesta etapa. Por fim, restaram 17 artigos para leitura íntegra.

Dos 17 artigos selecionados, apenas dois eram escritos em português, os demais escritos em inglês, apesar de dois destes serem publicações nacionais. Nas seções seguintes são descritos os 17 artigos resultantes e como eles ajudam a responder as questões de pesquisa a revisão sistemática propostas por este trabalho.

### 2.2.1 Ontologias em Sistemas de Informação em Hospitais

Após a leitura dos artigos, os trabalhos relacionados à primeira questão de pesquisa quanto ao uso de ontologias em Sistemas de Informação em Hospitais estão descritos nesta seção. Esta questão enfoca no cenário em que um mesmo hospital pode possuir diferentes Sistemas de Informação e, portanto, diferentes bases de dados (como ilustrado na Figura 3). Nesse tipo de cenário podem surgir ambiguidades para certos termos, por exemplo, queixa, que já foi discutido no capítulo anterior.

Figura 3 – Exemplo de setores hospitalares com bases de dados separadas.



Fonte: Porto (2006)

Nie et al. (2013) têm como proposta modelar uma ontologia para o HL7, um conjunto de normas para representação e transferência de dados hospitalares, que possui sintaxe em formato de texto puro. Primeiramente foi desenvolvida uma ontologia para representar os conceitos principais, presentes no HL7, e somente então foi feito um mapeamento semântico para o WSMO - *Web Service Modeling Ontology*. Por fim é apresentado um caso onde a ontologia é utilizada para representar a troca de mensagens, em HL7, entre Registros Médicos Eletrônicos e um Sistema de Informação Laboratorial, exemplificando seu comportamento em um caso real.

Keramaris e Danas (2016) relatam o caso de um hospital com inconsistências em sua base de dados, na qual foram encontradas falhas no mapeamento do sistema hospitalar local para o HL7. Os autores propõem o uso de ontologias como uma possível forma de solucionar o problema de troca direta de dados entre hospitais distintos. Além disso, propuseram a combinação de algumas metodologias para desenvolvimento de sistemas, como

o Modelo de Aceitação de Tecnologia e a Metodologia de Sistemas Flexíveis, para que essas falhas fossem evitadas no desenvolvimento de outros sistemas. Com isso, desenvolveram um *framework* com objetivo de combinar essas metodologias para tornar os Sistemas de Informações capazes de compartilhar conhecimentos específicos de um domínio, além de garantir interoperabilidade e qualidade de informação. Esse *framework* desenvolvido é chamado de “OntoDrive”.

Riaño et al. (2012) desenvolveram uma ontologia para apoio à tomada de decisão em uma clínica de pacientes com doenças crônicas. A hipótese era auxiliar os profissionais da saúde, visto que a estrutura das ontologias permite explicitar o relacionamento entre objetos o que proporciona informações mais precisas e estruturadas. Com objetivo de reduzir possíveis erros gerados por uma troca incorreta de informações, a ontologia desenvolvida pelos autores busca fornecer informações relevantes encontradas nos prontuários dos pacientes, pretendendo apresentar detalhes individuais e dados de diagnóstico importantes para os profissionais de saúde. Além disso, a ontologia é utilizada como base de conhecimento para outra ferramenta, que auxilia os profissionais de saúde a detectarem situações anormais, diagnósticos errados, etc.

Bettin, Teixeira e Berardi (2016) sugerem o desenvolvimento de uma ontologia para gerenciar informações operacionais em um hospital. A base para essa ontologia é o Sistema de Lousa Eletrônica, já presente no Hospital Universitário de Maringá e a ideia é de que essa ontologia possa melhorar a qualidade da informação apresentada neste sistema, visto que são informações essenciais para tomada de decisão dos profissionais de saúde.

Peral et al. (2018) propõem uma arquitetura para integrar informações de diferentes setores da área da Saúde. Os autores discutem o crescimento de dados na *Web* e como lidar com uma grande quantidade de dados localizadas em bases de dados heterogêneas. A arquitetura proposta utiliza ontologia para sua base de conhecimento, que irá permitir a integração de dados de diferentes fontes. Os autores ainda propõem a utilização de Mineração de Dados e Processamento de Linguagem Natural para tratamento dos dados obtidos.

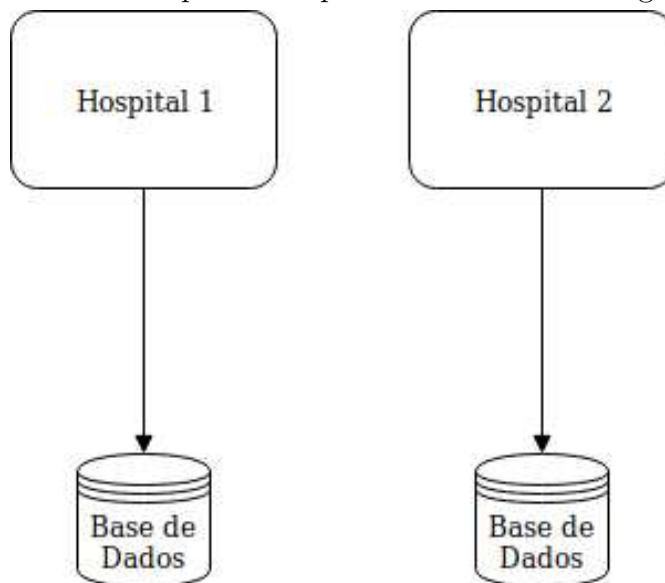
Nazarenko et al. (2015) sugerem o desenvolvimento de uma ontologia para automação de registros médicos. A modelagem dessa ontologia foi feita com base nos registros tecnológicos já existentes no hospital. A ontologia foi desenvolvida para o Centro Médico do Banco da Rússia e seu principal objetivo era criar um Registro Médico automatizado.

### 2.2.2 Ontologias para integração de dados entre unidades de saúde

Os trabalhos a seguir estão relacionados à segunda questão de pesquisa, que é uso de ontologias para integração de dados em que o cenário é entre unidades de saúde diferentes e, portanto, com bases de dados heterogêneas (como ilustrado na Figura 4).

Nesse tipo de cenário, podem surgir variações de vocabulário, isto é, a utilização de termos distintos para falar sobre o mesmo conceito. Um exemplo para isso é a hemorragia,

Figura 4 – Exemplo de hospitais com bases heterogêneas.



Fonte: Autoria própria

que pode ser entendida como perda de sangue. Em outro contexto hospitalar, esse conceito pode ser referido como sangramento, mesmo que, de uma forma geral, indiquem o mesmo problema.

Liu et al. (2014) buscam projetar uma ontologia para dados clínicos, essa ontologia pode auxiliar no monitoramento de *Clinical Pathway*, que são padrões de qualidade na saúde. Os dados necessários foram retirados de Registros Médicos Eletrônicos, tornando necessário que essa ontologia fosse capaz de representar os dados presentes nesses registros. Os registros estavam armazenados de forma isolada, dificultando o trabalho de integração de dados desses registros. Após o desenvolvimento dessa ontologia em OWL, foi desenvolvida uma ontologia baseada em regras (SWRL - *Semantic Web Rule Language*), com o objetivo de descrever a *Clinical Pathway*. Com essas novas ontologias, eles puderam monitorar o fluxo de trabalho em hospitais e integrar esses dados com as *Clinical Pathways*, através da SWRL.

Zhao e Zhang (2013) avaliam a possibilidade de converter registros médicos eletrônicos em uma ontologia e de que forma isso beneficiaria os profissionais da área. Esses registros são, basicamente, prontuários eletrônicos de cada paciente e a conversão é feita com processamento de linguagem natural. Após isso, é desenvolvido um SWRL baseado nos registros médicos eletrônicos.

Patrao et al. (2013) têm como proposta o desenvolvimento de um sistema de integração de base de dados, o que possibilita a integração de base de dados distintas e de diferentes hospitais. Para isso foi necessário o mapeamento desses dados para uma ontologia e com isso foi possível integrar bases de dados diferentes, o que abre possibilidade para integração de base de dados pública, por exemplo. Alguns testes foram realizados em

alguns setores do Hospital Oncológico A. C. Camargo Cancer Center na cidade de São Paulo, por exemplo os setores de patologia e de serviços clínicos, além disso foram feitos testes com tecnologias similares para fins de comparação com a ferramenta desenvolvida.

Kilintzis et al. (2019) projetaram uma ontologia para representar os termos presentes no HL7 e, dessa forma, possibilitar a integração de registros médicos que utilizem este padrão. Os autores defendem a utilização de ontologias para representação de registros médicos e outras aplicações na saúde. Como a ontologia é focada no padrão HL7, a reutilização para integrar dados de diferentes unidades de saúde é possível, se esse padrão for utilizado por elas. Os autores afirmam ainda que o uso da ontologia pode ser expandido para diversos domínios da área da Saúde.

Roehrs et al. (2018) desenvolveram uma aplicação para integrar dados de registros médicos com diferentes padrões de representação. A arquitetura proposta envolve algumas camadas, desde a obtenção dos registros até o armazenamento deles. Após obtenção dos registros, estes são tratados por um componente que utiliza Processamento de Linguagem Natural. Feito esse processamento, o registro é convertido para ser utilizado na camada ontológica. Após todos tratamentos e conversões, os registros são armazenados em uma base de dados relacional e uma base de conhecimento (utilizando uma ontologia).

Azami, Malki e Tahon (2012) apresentam uma arquitetura para Sistemas de Informação Hospitalares, a ideia é que essa arquitetura proporcione um Sistema de Informação com interoperabilidade, sem que seja necessário modificações para a comunicação entre diferentes instituições de saúde. Essa arquitetura é organizada em camadas, com responsabilidades diferentes em cada uma delas. Uma ontologia foi utilizada para unificar conceitos da área da saúde, permitindo a troca de dados entre instituições diferentes.

Alkahtani et al. (2017) têm como proposta desenvolver um sistema multi-agente que permita a troca de dados entre diferentes hospitais. Para isso é proposta uma arquitetura que descreve o fluxo do sistema, onde os hospitais podem buscar informações presentes em outras bases de dados. Para implementação desse processo, foram utilizadas ontologias. Os testes, feitos sobre uma situação fictícia, mostraram que o sistema implementado conseguiu cumprir com a proposta, sendo capaz de estabelecer a comunicação entre diferentes bases de dados, este sistema foi chamado de *Statistics and Collaborative Knowledge Exchange*.

Rao, Makkithaya e Gupta (2014) propõem o desenvolvimento de uma ontologia de saúde pública, ou seja, que faça a integração de dados presente em diferentes hospitais da rede pública. Essa integração deve ocorrer de forma contínua e não pode ser afetada por diferenças nas tecnologias utilizadas. Com isso é possível ter uma visão mais abrangente sobre o setor, abrindo muitas possibilidades de uso. Para o desenvolvimento da ontologia, foi utilizada a base de dados presentes em um hospital específico da rede pública. Após o desenvolvimento da ontologia foi possível fazer a conversão de um banco de dados relacional para instâncias dessa ontologia, gerando resultados satisfatórios que abrem possibilidades para integração de dados na rede pública.

Mistry et al. (2013) discutem a necessidade de usar ontologias e sua aplicação para o domínio hospitalar. São apresentadas vantagens e desvantagens de utilizar ontologias, além de buscar como modelar uma para área da saúde, onde é utilizada a metodologia de desenvolvimento "Ontology Development 101", própria para ontologias. Após isso, é apresentada uma ontologia desenvolvida pelos autores, que foi testada em situações fictícias através da criação de instâncias para simular seu funcionamento.

Pereira, Wassermann e Salvador (2018) propõem um estudo de caso para avaliar a utilização de dois *frameworks* baseados em ontologias para melhorar a obtenção de dados hospitalares. Este trabalho discute como problemas de diferença semântica podem surgir mesmo na área da Saúde, que possui uma terminologia bem definida. Também existe uma discussão sobre as inúmeras bases de dados do Sistema Único de Saúde serem heterogêneas. O Ministério da Saúde possui cerca de 140 sistemas de informação para a área. Estes sistemas são independentes e desenvolvidos sob demandas diferentes, por isso surge a dificuldade para integração dos dados destes. O trabalho analisa os *frameworks* Ontop e *Linked Data Mashup* utilizando os dados abertos de São Paulo.

E por fim, Souza, Medeiros e Martins (2015) têm como proposta o desenvolvimento de uma aplicação ontológica para facilitar o acesso à informações sobre saúde, visto que boa parte dessas informações estão presentes na *Internet* de forma não-estruturada e desconexa, o que pode, em alguns casos, dificultar seu acesso. Apesar de ser uma modelagem nacional, os termos utilizados foram escritos na língua inglesa, com a reutilização de ontologias conhecidas internacionalmente. Os dados necessários foram extraídos manualmente de sites tanto da rede particular, como da rede pública de saúde no Brasil. Em seguida alguns exemplos de consultas foram apresentados, mostrando parcialmente o funcionamento da estrutura gerada.

### 2.2.3 Ontologias para integração com dados abertos

Com relação à terceira questão da pesquisa, não foram encontrados trabalhos que apresentassem, explicitamente, propostas de integração de dados locais de hospitais com dados abertos. Este fenômeno pode ocorrer por alguma característica da metodologia de busca empregada ou, até mesmo, pela novidade da temática escolhida. Outro fator que pode ter contribuído para esse fenômeno é o fato dos dados disponíveis ainda exigirem um trabalho manual para sua manipulação, devido ao seu formato e forma de distribuição. No entanto, a confirmação dessas hipóteses demandariam novos estudos.

Além disso, o uso das ontologias para integração com dados abertos não é uma tarefa trivial. Uma das dificuldades que pode ser encontrada durante essa tarefa é o formato de arquivo utilizado para divulgação dos dados abertos. Tomando como exemplo os dados disponíveis no portal brasileiro de dados abertos, onde os três formatos mais utilizados são: CSV, HTML e JSON (BRASIL, 2020). Para se ter uma ideia, no momento do levantamento, foram encontrados 14463 bases disponíveis nesses formatos, enquanto



que no formato RDF, que seria mais adequado para ontologias, foram encontrados apenas 24. Há, ainda, a diferença semântica entre bases de dados, que deve ser resolvida para possibilitar essa integração (PATRAO et al., 2013; RAO; MAKKITHAYA; GUPTA, 2014).

### 2.3 Considerações

Com a tecnologia existente na atualidade, os Sistemas de Informação em saúde integrados entre unidades mostraram-se implementáveis. Algumas ressalvas se devem ao fato de se tratar de uma área recente e ainda não existir uma grande quantidade de trabalhos realizados neste sentido. Os artigos selecionados apresentaram modelos e tecnologias necessários para o desenvolvimento desse tipo de aplicação, como arquitetura de Sistemas de Informação e modelos de ontologias para unidades de saúde. Além disso, mostraram como as ferramentas apresentadas podem solucionar problemas de integração entre diferentes bases e também melhorar o apoio ao processo de decisão.

Foram identificados vários desafios para a integração da base de dados de unidades de saúde com dados abertos. Entre eles, vale destacar os diferentes formatos de arquivo utilizados para publicação dos dados abertos e a diferença de ordem semântica entre as bases de dados (SOUZA; MEDEIROS; MARTINS, 2015). Embora nenhum artigo selecionado tenha levantado a questão sobre essas diferenças, vale ressaltar que os trabalhos que tratam da integração entre diferentes bases de dados também contribuem para discussão deste tópico.

Apesar das diversas bases de dados existentes não seguirem um padrão de representação, através de ferramentas como mineração de dados e processamento de linguagem natural, é possível realizar conversões de registros eletrônicos e não-eletrônicos para uma ontologia (ZHAO; ZHANG, 2013). Dessa forma, o problema do formato dos dados pode ser contornado com ferramentas preexistentes. Através das ontologias é possível estabelecer troca de dados entre diferentes unidades de saúde e ainda melhorar questões, como a qualidade da informação, podendo beneficiar o acesso aos dados abertos governamentais (SOUZA; MEDEIROS; MARTINS, 2015). Ademais, é possível estabelecer troca de dados locais com dados abertos, provendo uma visão mais holística sobre a unidade de saúde ou um determinado paciente, auxiliando à tomada de decisão para profissionais e gestores na área da saúde (RAO; MAKKITHAYA; GUPTA, 2014).

Por fim, de acordo com os resultados obtidos da busca aplicada, é possível inferir que a integração de dados tem sido amplamente utilizada em diferentes cenários da área da saúde, tanto em hospitais com uma base de dados heterogêneas quanto entre diferentes hospitais. Porém, a integração local com dados abertos, aplicado na gestão de saúde pública, ainda não foi um aspecto amplamente explorado. Apesar das ontologias apresentadas terem sido utilizadas em diversos contextos na área da saúde, não foram encontradas aplicações em maternidades de alto risco ou em prontos atendimentos com classificação de risco.

### 3 Referencial Teórico

Neste capítulo são apresentados os conceitos utilizados neste trabalho. Começando por Ontologia, onde são apresentadas algumas definições, boas práticas e usos. Após, são descritos os Dados Abertos e Dados Abertos Conectados, que fazem parte de uma discussão mais ampla sobre democratização da informação. Estes conceitos potencializam a abrangência da integração de dados.

#### 3.1 Ontologia

Ontologia é um conceito, com origem na Filosofia, que aborda a natureza do ser e a existência. Situa-se em um campo interdisciplinar com aplicação em diversas áreas como Ciência da Informação, Inteligência Artificial, etc. Na abordagem adotada pela Ciência da Computação, o termo está mais ligado com a contextualização de um domínio. Usualmente, uma ontologia é descrita como uma representação formal de um domínio tal que seja compreensível por humanos e computadores (MIZOGUCHI, 2004).

Segundo Gruber (1993), ontologia é uma especificação explícita de uma conceituação, isto é, uma visão simplificada e abstrata do mundo real, que desejamos representar conceitualmente para algum propósito. A representação formal de um domínio deve possuir, em sua declaração, quais são os objetos que compõem aquele universo, bem como suas relações. Ele também propõe critérios para o desenvolvimento de ontologias que tenham por objetivo o compartilhamento de conhecimento e interoperabilidade entre sistemas. Esses critérios são:

- Clareza: uma ontologia deve efetivamente comunicar o significado dos termos definidos. Essas definições devem ser claras e objetivas.
- Coerência: as inferências obtidas devem ser consistentes com as definições. Por isso, os axiomas definidos devem ser logicamente consistentes com o domínio que está sendo tratado.
- Extensibilidade: uma ontologia deve ser projetada de forma que seja possível o compartilhamento de um vocabulário. Isso significa que se for necessário utilizar parte do vocabulário desenvolvido para outro fim, não seja necessário revisar as definições já existentes.
- Mínimo de problemas com codificação: a conceituação deve atuar em um nível de conhecimento onde não exista ou seja minimizada a dependência de símbolos presentes em codificações alternativas, evitando problemas para os agentes consumidores dessa base de conhecimento.
- Mínimo de compromisso ontológico: as definições que devem ser feitas são as essenciais para a ontologia. Isto facilita a participação de terceiros sobre a base desenvolvida.

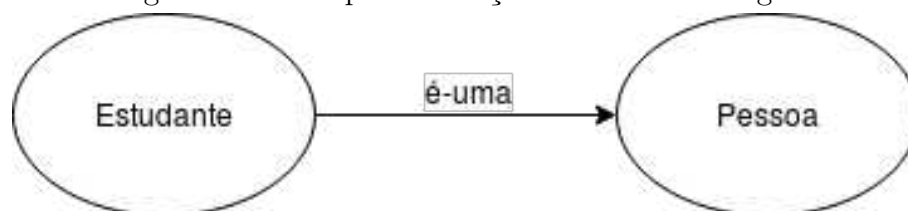
Para Guarino (1995), o termo "ontologia" pode ser lido de duas formas. Do ponto de vista filosófico, se refere a uma forma de descrever uma visão que temos do mundo, sem depender de uma linguagem específica. Para alguns campos da Ciência da Computação, ontologia consiste em um artefato que contém um vocabulário específico utilizado para descrever algo. Este vocabulário se refere a conceitos e suas relações, que, em uma ontologia, são organizados de forma hierárquica.

Guizzardi (2005) aborda a utilização de ontologias em quatro áreas da Ciência da Computação, são elas: Banco de Dados, Sistemas de Informação, Engenharia de *Software* e Inteligência Artificial. Embora sejam utilizadas com objetivos diferentes, as ontologias mostram-se úteis em diversas aplicações, principalmente quando existe a necessidade de reduzir a ambiguidade de termos, ter bases de conhecimento reutilizáveis, etc. As ontologias também podem ser utilizadas em outras áreas que não fazem parte da Ciência da Computação, entre as várias citadas por Guizzardi (2005), vale ressaltar Medicina, Direito e Química.

Para a Inteligência Artificial, uma ontologia pode ser compreendida como uma coleção de conceitos fundamentais e pela estrutura do conhecimento que será descrito (ISOTANI; BITTENCOURT, 2015). Estes podem ser entendidos da seguinte forma:

- Coleção de conceitos fundamentais: são parte do conhecimento básico sobre um domínio específico e pode ser representado pela forma de um vocabulário.
- Estrutura do conhecimento: o domínio deve ser descrito de acordo com as relações entre os conceitos especificados. O resultado é uma hierarquia na forma de classe/subclasse, esta relação é chamada é-um (do inglês *is-a*) e indica que uma classe é subclasse de outra (como pode ser visto na Figura 5), e um conjunto de relações não hierárquicas como o parte-de (do inglês *part-of*), que indica uma composição de duas partes (conforme mostrado na Figura 6). Além disso, as limitações semânticas do domínio, bem como em suas relações, devem estar descritas nesta estrutura.

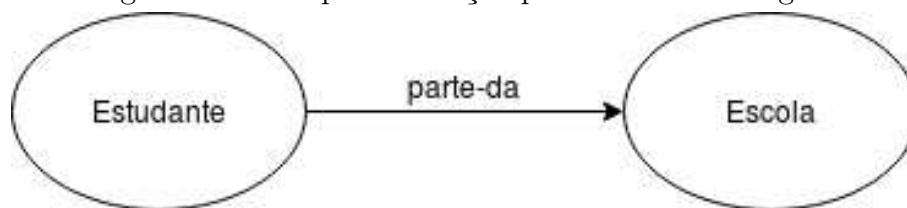
Figura 5 – Exemplo de relação é-um em ontologias.



Fonte: Autoria própria

Formalmente, uma ontologia é descrita da seguinte forma: Ontologia = [C, R, I, A]. Onde C é o conjunto das classes existentes no domínio, que representam conceitos presentes neste domínio. R representa as relações entre os conceitos presentes nesse domínio. I são as instâncias, ou seja, a representação de objetos específicos através da ontologia. A são os axiomas, ou seja, as restrições semânticas do domínio, que permitem também

Figura 6 – Exemplo de relação parte-de em ontologias.



Fonte: Autoria própria

a interpretação dos conceitos envolvidos (KIRYAKOV; OGNYSANOV; MANOV, 2005; GRUBER, 1993).

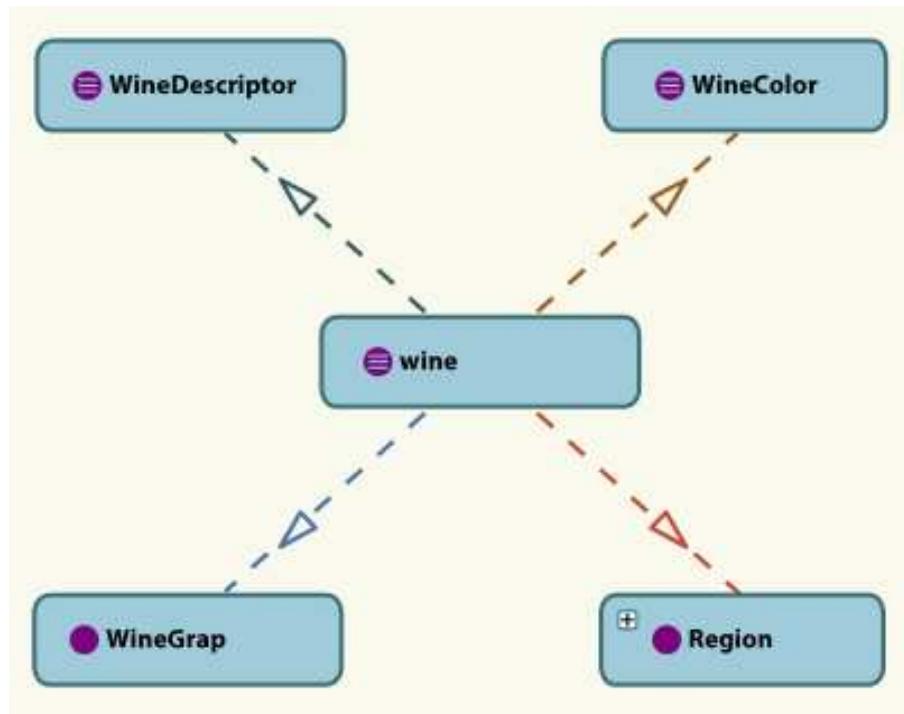
Embora a terminologia utilizada em uma ontologia seja relevante, Isotani e Bittencourt (2015) defendem que o desenvolvimento de uma ontologia deve focar na compreensão dos conceitos que fazem parte do domínio que desejamos descrever. Segundo Isotani e Bittencourt (2015), o vocabulário ou a terminologia utilizada é basicamente um rótulo para os conceitos existentes no mundo real. Uma ontologia representa e expressa o significado de conceitos e não de termos, embora também possa ser utilizada dessa forma.

A representação de uma ontologia pode ser feita de duas formas. A primeira forma é gráfica (como pode ser visto na Figura 7), utilizada para compreensão humana. Neste formato, usualmente são empregados grafos de forma que os vértices simbolizem as classes e as arestas as relações entre essas classes. A segunda forma é a representação formal, utilizada pelos computadores através de uma linguagem descritiva. Para a representação formal de uma ontologia, as linguagens descritivas mais utilizadas são: OWL e RDF/RDF-S (ISOTANI; BITTENCOURT, 2015).

A W3C, órgão que define diversos padrões na *Web*, teve como proposta inicial a especificação RDF/RDF-S. O RDF permite a criação do relacionamento entre dados em triplas da seguinte forma: sujeito, predicado, objeto (conforme mostrado na Figura 8). O sujeito é um recurso único, isto é, um elemento único na ontologia, identificado através de um URI. Na Figura 8 "Leonardo Da Vinci" e "The Mona Lisa" são exemplos de recursos. O objeto poder ser um literal ou um recurso na ontologia. Um literal é um dado puro, por exemplo, "14 July 1990" que representa a data de nascimento de Bob da Figura 8. O predicado é uma propriedade que viabiliza a relação entre os dois nós. Para descrição de uma ontologia utiliza-se o RDF-S, responsável pela definição de classes e propriedades, que resultam em *tags* no XML, como pode ser visto na Figura 9 (BERNERS-LEE, 2000).

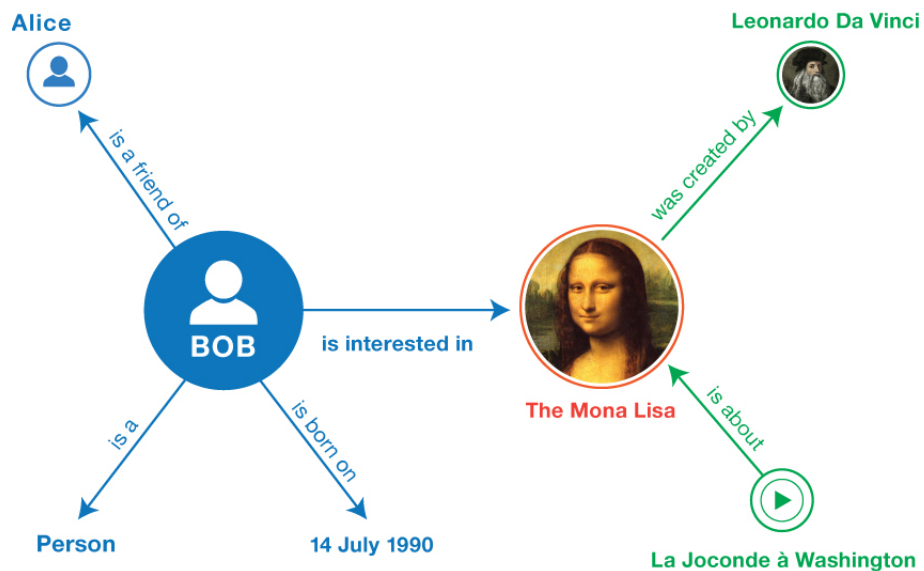
A OWL é uma linguagem, recomendada pela W3C, que foi projetada a fim de melhorar algumas limitações existentes no RDF, por isso existem algumas similaridades entre as duas linguagens. Em um arquivo OWL é possível definir as classes e propriedades de uma ontologia e, neste mesmo arquivo, criar as instâncias do seu domínio. Assim como no RDF, a OWL também é baseada no XML, utilizando *tags* para descrever sua estrutura (MIZOGUCHI, 2004).

Figura 7 – Exemplo de ontologia gráfica.



Fonte: Isotani e Bittencourt (2015)

Figura 8 – Exemplo de relacionamento em forma de triplas.



Fonte: W3C (2014)

Para realização de consultas sobre uma base de dados desenvolvida em RDF, é possível utilizar uma linguagem de consulta. O SPARQL é uma linguagem orientada a dados que foi desenvolvida pela W3C e permite a consulta sobre ontologias através da busca por padrões de triplas (W3C, 2008). A linguagem também possui algumas similaridades

Figura 9 – Exemplo de RDF/RDF-S.

```

<?xml version="1.0" encoding="utf-8" ?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:foaf="http://xmlns.com/foaf/0.1"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:ns0="http://dbpedia.org/ontology/"
  xmlns:schema="http://schema.org/"
  xmlns:geo="http://www.w3.org/2003/01/geo/wgs84_pos#">

  <foaf:Person rdf:about="http://dbpedia.org/resource/Bob_Marley">
    <rdfs:label xml:lang="en">Bob Marley</rdfs:label>
    <rdfs:label xml:lang="fr">Bob Marley</rdfs:label>
    <rdfs:seeAlso rdf:resource="http://dbpedia.org/resource/Rastafari"/>
    <ns0:birthPlace>
      <schema:Country rdf:about="http://dbpedia.org/resource/Jamaica">
        <rdfs:label xml:lang="en">Jamaica</rdfs:label>
        <rdfs:label xml:lang="it">Giamaica</rdfs:label>
        <geo:lat rdf:datatype="http://www.w3.org/2001/XMLSchema#float">17.9833</geo:lat>
        <geo:long rdf:datatype="http://www.w3.org/2001/XMLSchema#float">-76.8</geo:long>
        <foaf:homepage rdf:resource="http://jis.gov.jm"/>
      </schema:Country>
    </ns0:birthPlace>

  </foaf:Person>

</rdf:RDF>

```

Fonte: Addlesee (2018)

com o SQL, utilizado para consultas em banco de dados relacionais.

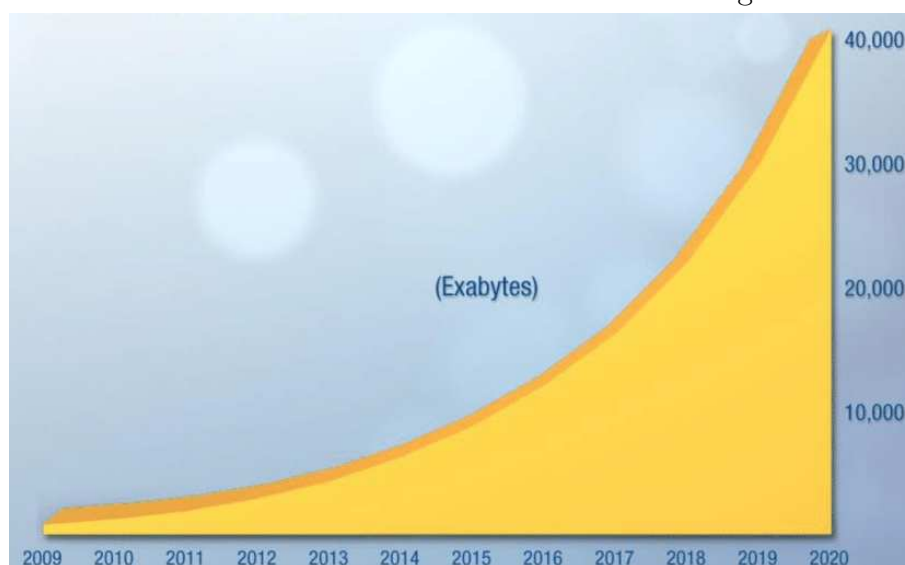
Uma ontologia tem por objetivo propiciar a representação abstrata de um determinado domínio de conhecimento. Este domínio pode ser compreendido como um conjunto de conceitos, relações e propriedades (GUIMARÃES, 2015). Os tipos de aplicações que podem ser feitas com esse modelo são diversos, mas é notável sua utilização em bases de conhecimentos que necessitam de melhorarias em questões semânticas. Dentre várias dessas bases, uma que vale destacar é a *Web*, que adotou as ontologias como modelo para melhorar sua semântica (GUIZZARDI, 2005).

A discussão sobre a necessidade de semântica na *Web* teve seu início no ano de 1994, durante a primeira conferência internacional da WWW - *World Wide Web*. Nessa conferência, o criador da WWW, o físico Timothy John Berners-Lee, mencionou a forma com que os documentos, até então, estavam estruturados na *Web* e como isso poderia ser melhorado. A organização de páginas através de *links* e nós dificulta a compreensão semântica dos conteúdos para as máquinas, mesmo que seja uma tarefa trivial para humanos. Isso propicia uma série de complicações para o futuro da *Web* e esse problema foi notado já em 1994, apenas 5 anos após sua criação (ISOTANI; BITTENCOURT, 2015; GUIZZARDI, 2005).

Essa questão começou a ganhar mais evidência com o crescimento da *Web*, que em meados de 2012 já havia ultrapassado um zettabyte de dados e possui uma estimativa

de crescimento ainda maior para os anos seguintes, conforme mostrado na Figura 10. A busca de informações em uma base com tal quantidade de dados é uma atividade inviável para seres humanos, sendo necessário que esta seja realizada por máquinas. Para que os computadores possam realizar essa tarefa de forma eficaz, as páginas *web* precisam ter semântica, que é possível através do uso de uma linguagem apropriada para tal. Todo esse processo, com mais detalhes e de forma mais abrangente, compõem a *Web Semântica*, modelo proposto por Berners-Lee em 2001 (ISOTANI; BITTENCOURT, 2015; BERNERS-LEE et al., 2001).

Figura 10 – Estimativa de crescimento do volume de dados digitais de 2010 a 2020.



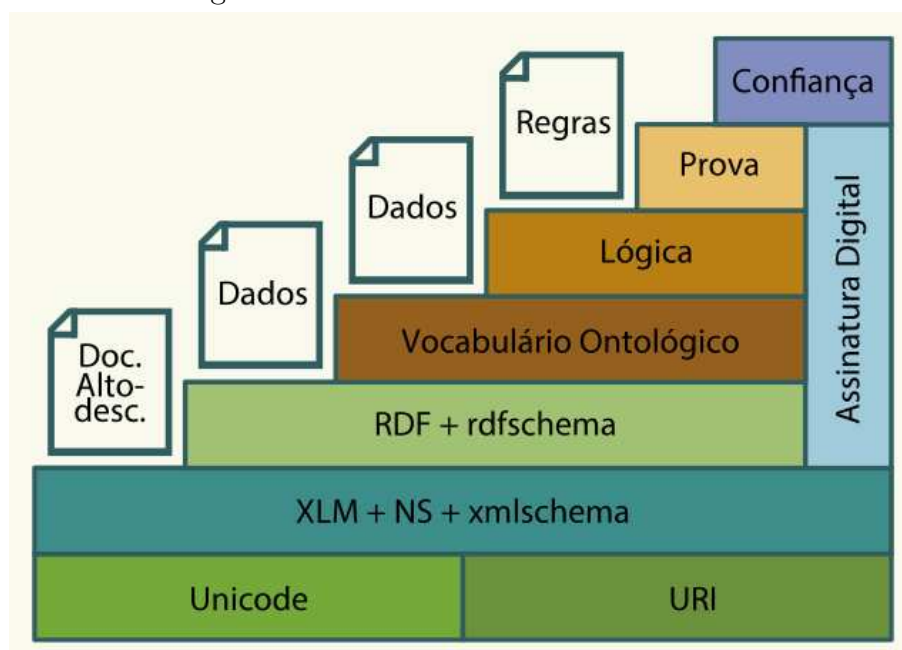
Fonte: EMC (2012)

Dessa forma, as ontologias fizeram parte do desenvolvimento da *Web Semântica*, dado que é um modelo de dados com características que suprem as necessidades do modelo proposto por Berners-Lee, como a capacidade de diminuir ambiguidades de termos existentes na linguagem natural e a possibilidade de compartilhamento de informação. Berners-Lee et al. (2001) propôs inicialmente um modelo em camadas para a *Web Semântica*, como pode ser visto na Figura 11, onde as ontologias compõem uma dessas camadas. Com isso, é possível perceber que as ontologias tornaram-se fundamentais para o desenvolvimento de ferramentas que lidam com grandes quantidades de dados, em questão de eficácia, além de casos onde é necessário lidar com bases de conhecimento (HORROCKS, 2008; GUIZZARDI, 2005).

### 3.2 Dados Abertos

Dados Abertos são dados cujo uso, reutilização e redistribuição são livres, ou seja, qualquer pessoa pode livremente acessar, modificar e compartilhar esses dados. Isso

Figura 11 – Camadas da Web Semântica.



Fonte: Isotani e Bittencourt (2015)

deve acontecer de forma que não existam mecanismos de restrição ao acesso dos dados, possibilitando o acesso por qualquer indivíduo, para qualquer finalidade (BRASIL, 2019).

A definição de Dados Abertos, de acordo com Foundation (2019), possui três regras essenciais, sendo elas:

- Disponibilidade e acesso: o custo para acessar os dados não deve ser superior ao seu custo de reprodução (impressão, acesso à *Internet*). De preferência, disponível para *download* na *Internet* e em formato conveniente.
- Reutilização e redistribuição: a licença sobre qual esses dados estão disponíveis deve permitir sua reutilização e redistribuição, alinhado à legislação vigente.
- Participação universal: o acesso sobre esses dados não pode ser restrito a um grupo de pessoas, não podendo haver quaisquer tipo de exclusão ou restrição. O acesso deve ser permitido para qualquer pessoa que tenha interesse.

Esse tipo de base de dados permite que as organizações trabalhem de maneira colaborativa, o que pode ser benéfico para todos que participam. Apesar disso, a adesão de organizações, principalmente as privadas, ainda é relativamente baixa na hora de compartilhar suas bases de dados (BLOOMBERG, 2017). Historicamente, o termo Dados Abertos surgiu em 1995 com a divulgação de dados ambientais e geofísicos por parte de um órgão estadunidense. Atualmente, os Dados Abertos ganharam notoriedade devido ao movimento de diversos grupos a favor da transparência, principalmente por parte dos governos. Isso estimulou o crescimento de movimentos do tipo na *Internet*, com resultados positivos, visto que diversos governos passaram a adotar medidas em prol da transparência



(CHIGNARD, 2013; ISOTANI; BITTENCOURT, 2015).

No Brasil, a disponibilidade de dados governamentais em formato aberto virou lei em 2011, notoriamente o Brasil foi um dos fundadores do *Open Government Partnership*, um grupo de colaboração internacional para promover políticas à favor da transparência, e por consequência, dos Dados Abertos. Atualmente esse grupo conta com 79 países (PARTNERSHIP, 2019).

Nesta perspectiva, a prefeitura de Curitiba passou a disponibilizar dados abertos em seu portal na *Internet*. Essa base faz parte de uma iniciativa para melhorar a transparência do órgão, atendendo algumas das exigências presentes na Lei de Acesso à Informação, já mencionada neste trabalho. A base dispõe de dados sobre diversos setores do governo municipal, sem qualquer restrição para seu acesso e com uma licença aberta, cumprindo todos os requisitos para ser classificado como dados abertos. Esses dados estão disponíveis no *website* do órgão em diferentes formatos de arquivo, sendo alguns deles: CSV, XLSX e PDF. É possível ainda solicitar bases de dados que, atualmente, não estejam disponíveis no portal, desde que esteja de acordo com a legislação vigente (CURITIBA, 2019).

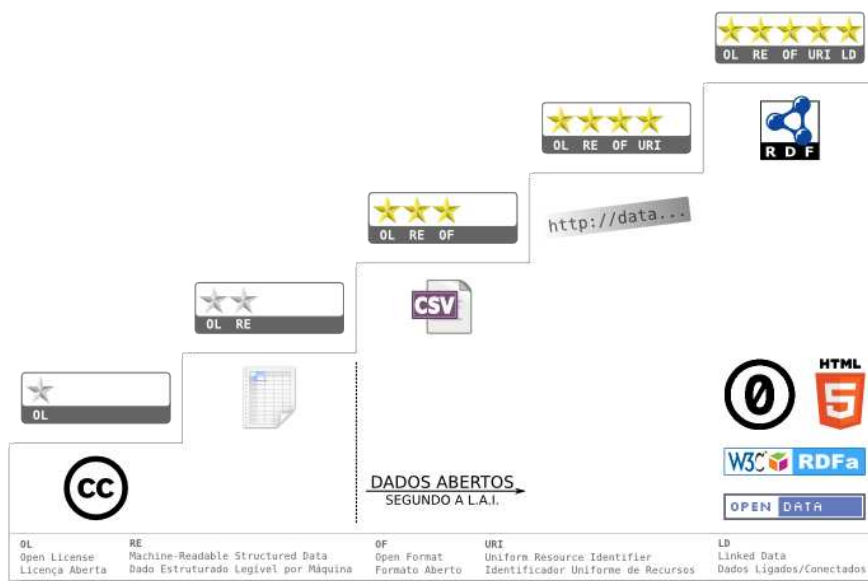
### 3.2.1 Dados Abertos Conectados

A discussão sobre a *Web Semântica*, citada anteriormente, ganhou força e o modelo proposto por Timothy John Berners-Lee começou a receber diversas contribuições para seu desenvolvimento. Um dos conceitos que passaram a integrar o desenvolvimento da *Web Semântica* foi o de Dados Conectados. Esse conceito (quem vêm do inglês *Linked Data*) faz referência a práticas para conectar um conjunto de dados estruturados na *Web*. Isso significa que agentes de *software* podem ler, de maneira automática, esse conjunto de dados, sem a necessidade de ferramentas adicionais para tal (ISOTANI; BITTENCOURT, 2015).

Uma base de dados abertas com essas características é definida como Dados Abertos Conectados e para que seja possível mensurar a conectividade dos dados abertos, existe uma métrica específica para tal. Esse é o chamado "Sistema de 5 Estrelas", proposto por Berners-Lee (2009), que pode ser visto na Figura 12. Nessa métrica (onde mais estrelas implica melhor conectividade), os dados abertos recebem de uma a cinco estrelas de acordo com algumas características dessas bases de dados. O uso de uma licença aberta garante apenas uma estrela, enquanto a utilização de RDF, URI e formato de arquivo aberto geram uma classificação de cinco estrelas (ONTOTEXT, 2010).

A discussão sobre a qualidade dos dados abertos é importante, pois garante que não ocorra um simples cumprimento da legislação ou uma divulgação desleixada de dados na *Web*. A qualidade destes assegura uma maior democratização, visto que a simples divulgação de bases de dados não garante a compreensão semântica destas. Nesse sentido, Dados Abertos Conectados é um importante pilar nessa área, uma vez que não

Figura 12 – 5 Estrelas dos Dados Abertos Conectados.



Fonte: Berners-Lee (2009)

apenas garantem aos agentes de *software* uma leitura automática de bases de dados, como permitem, por meio desses agentes, maior acessibilidade aos consumidores destas bases (ISOTANI; BITTENCOURT, 2015).

## 4 Metodologia

A metodologia escolhida para desenvolvimento da ontologia foi a *Ontology Development 101*. A escolha desta se deu pelo seu reconhecimento na área e também por sua simplicidade. Importante ressaltar que na primeira etapa de desenvolvimento da ontologia, estão inseridos alguns dos resultados da dissertação de mestrado citada anteriormente Costa et al. (2020), que servem como base e justificativa para o desenvolvimento da ontologia.

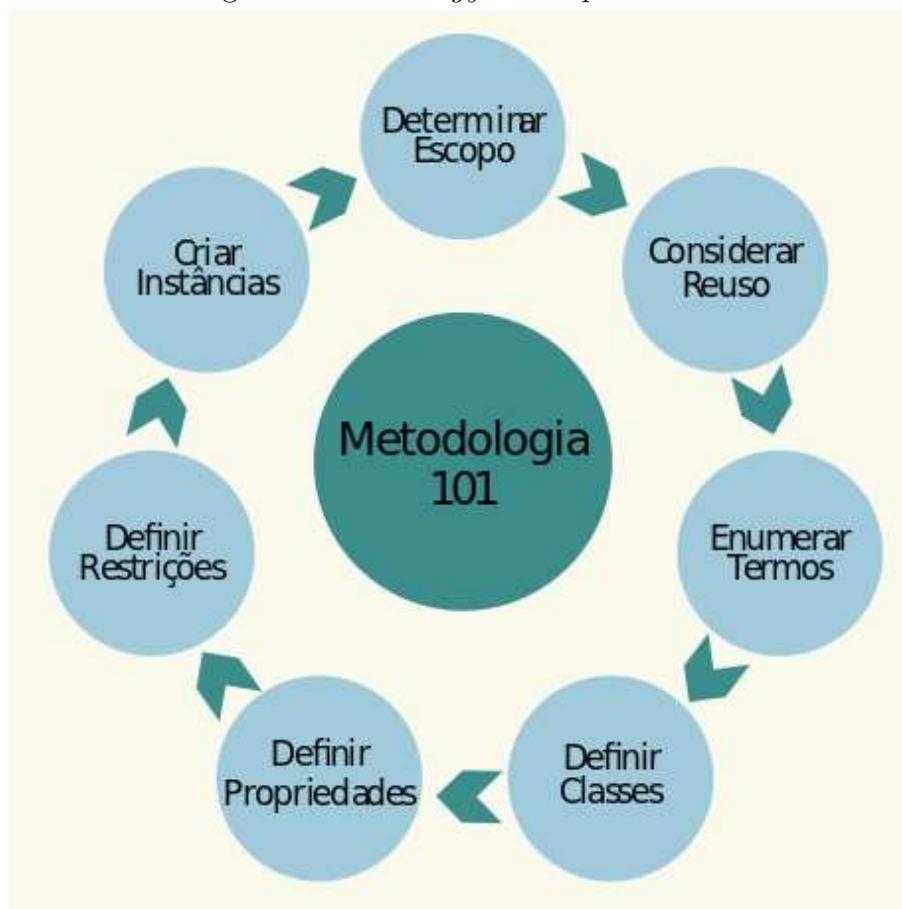
### 4.1 Modelagem da Ontologia - *Ontology Development 101*

A *Ontology Development 101* foi proposta por Noy e McGuinness (2004) e seu desenvolvimento é composto por sete passos, conforme mostrado na Figura 13. De acordo com as autoras, o primeiro passo dessa metodologia é determinar o domínio e o escopo da ontologia. Determinar esse escopo envolve descrever qual domínio a ontologia irá cobrir, buscando entender qual objetivo para construção dela. Também faz parte desse primeiro passo pensar sobre o tipo de informação que a ontologia irá fornecer, bem como quem irá utilizá-la. As autoras apresentam, também, uma outra forma de determinar o escopo da ontologia, que é através de questões de competência, isto é, questões que a ontologia deve ser capaz de responder quando pronta. Essa estratégia também serve para testar a ontologia posteriormente.

O segundo passo consiste em buscar ontologias já desenvolvidas e disponíveis para trabalhar conceitos específicos, por exemplo, os dados vitais de uma paciente, que é medido pela equipe de enfermagem. É possível que já exista uma ou mais ontologias que tenham trabalhado com esse conceito. Desta forma, seria possível reaproveitar parte do que já foi desenvolvido sobre.

O terceiro passo é enumerar os termos que serão utilizados na ontologia. Os termos dizem a respeito de conceitos existentes no domínio e nesta etapa eles não são separados em classes ou propriedades. Nesta fase, o mais importante é conseguir listar os termos que fazem parte do domínio, sem se preocupar com as relações ou restrições que eles representam.

O quarto passo é definir quais serão as classes da ontologia. Nessa etapa também é definida a hierarquia das classes. Para isso, as autoras apresentam três estratégias diferentes. A *top-down*, que consiste em definir conceitos mais genéricos e abstratos e só então os mais específicos, por exemplo, definir uma classe Vinho e depois definir Vinho Branco, Vinho Rosé, etc. A *bottom-up*, que é o caminho contrário da estratégia anterior, onde primeiro são definidos conceitos mais específicos e depois suas abstrações. A última estratégia combina as duas anteriores, iniciando com alguns poucos conceitos abstratos.

Figura 13 – *Ontology Development 101*.

Fonte: Isotani e Bittencourt (2015)

O quinto passo é definir as propriedades das classes criadas. Essas propriedades podem ser encontradas nos termos criados no terceiro passo, visto que nem todos termos resultam em classes. Um exemplo de propriedade pode ser cor de um vinho, onde vinho é uma classe definida no passo anterior.

O sexto passo é definir as restrições para as classes e suas propriedades. Tipos (texto, numérico, etc), valores permitidos e cardinalidade são exemplos de restrições. Basicamente são axiomas, que tornam a ontologia logicamente consistente para o domínio tratado. Por fim, o último passo é criar instâncias na ontologia. Também funciona como teste da ontologia desenvolvida, possibilitando ao desenvolvedor verificar se os requisitos levantados foram atendidos.

Para alimentar a ontologia desenvolvida, foram utilizados alguns registros presentes nos dados abertos de Curitiba, atualizados em abril de 2020 com espectro temporal de 3 meses. Esses registros foram filtrados para que fossem utilizados apenas atendimentos da área de Ginecologia. Por se tratar de um escopo muito específico, foi necessário a utilização de dados sintéticos mesclados com os dados reais obtidos anteriormente. Isso porque a base utilizada não contempla todos os campos de dados necessários para alimentar a ontologia.

Para testar a ontologia, foram elaboradas consultas em SPARQL. Isto possibilitou avaliar a ontologia quanto aos requisitos levantados.

## 4.2 Ferramentas Utilizadas

Para filtrar registros nos dados abertos de Curitiba, foi desenvolvido um programa utilizando a linguagem Python<sup>1</sup> (versão 3.5.3), com a biblioteca Pandas<sup>2</sup> (versão 0.25.1). O código-fonte resultante pode ser encontrado no Apêndice A. Para desenvolvimento da ontologia, foi utilizado o programa Protégé<sup>3</sup> (versão 5.5.0) com *plugin*<sup>4</sup> para execução de *queries* em SPARQL.

---

<sup>1</sup><https://www.python.org/downloads/release/python-353/>

<sup>2</sup><https://pandas.pydata.org/pandas-docs/stable/whatsnew/v0.25.1.html>

<sup>3</sup><https://protege.stanford.edu/>

<sup>4</sup><https://github.com/protegeproject/autoupdate/tree/master/snap-sparql-query>

## 5 Resultados

### 5.1 Modelagem da Ontologia

A modelagem da ontologia iniciou-se de acordo com os passos da metodologia 101. A execução dos passos está descrita nas seções a seguir. O arquivo OWL da ontologia pode ser encontrado no *link*: [https://github.com/RMoroni/tcc2/blob/master/maternidade\\_pa.owl](https://github.com/RMoroni/tcc2/blob/master/maternidade_pa.owl).

#### 5.1.1 Determinação do Escopo

O escopo da ontologia envolve as áreas de atuação da unidade de saúde estudada, são elas: Maternidade, Maternidade de Alto Risco e Tocoginecologia. Para elaboração das questões de competência, foi utilizado o levantamento das necessidades de informação feito por Costa et al. (2020) no Pronto Atendimento Tocoginecologia na Maternidade de Alto Risco do HC da UFPR. Este levantamento foi feito através de diversos passos, entre eles, vale ressaltar, uma série de entrevistas estruturadas com os profissionais de saúde envolvidos e uma avaliação dos atributos relativos à qualidade da informação, seguindo a metodologia proposta por Sordi (2015). A avaliação da qualidade da informação foi aplicada sobre o FACR - Formulário de Acolhimento com Classificação de Risco.

O FACR, que é um arquivo impresso conforme o modelo mostrado no Anexo A, é utilizado para obter e armazenar informações sobre a paciente. No processo de acolhimento, é possível perceber que a troca de informações entre os setores também está vinculada ao FACR. Na Figura 14 está ilustrada a trajetória do FACR na unidade de saúde, isto é, por quais setores ele transita durante o processo de acolhimento da paciente na unidade de saúde.

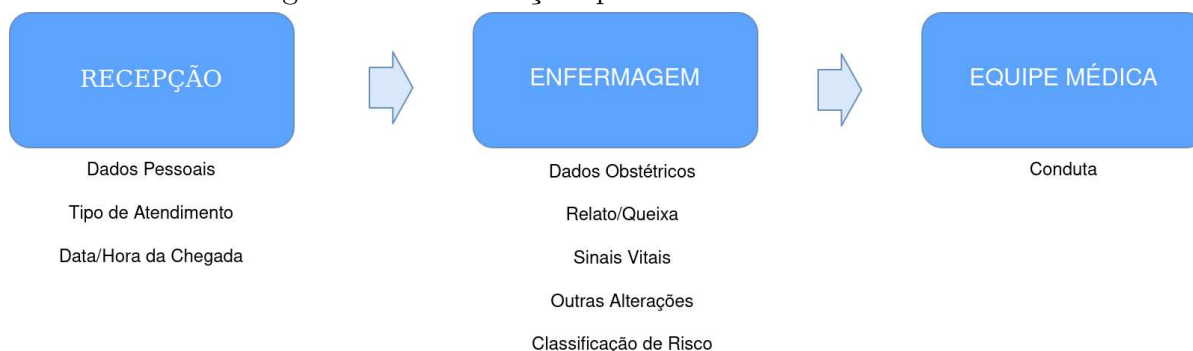
Figura 14 – Trajetória do FACR na unidade de saúde.



Fonte: Adaptado do trabalho de Costa et al. (2020)

Percebe-se que a trajetória do FACR entre os setores é um elemento crucial para o fluxo de informações. Nessas etapas ocorre o preenchimento do FACR, cada setor possui responsabilidades diferentes quanto ao preenchimento do mesmo. No entanto, em vários momentos o FACR trafega entre os setores apenas como uma fonte de informações sobre a paciente, servindo como base para decisões. Os grupos de informações inseridos por cada setor podem ser vistos na Figura 15, as etapas em que o FACR é apenas consultado, sem inserção de novas informações, foram omitidas.

Figura 15 – Informações preenchidas em cada setor.



Fonte: Autoria própria

Existem ainda outras ações presentes nesse processo, como a geração de uma carta de encaminhamento (quando necessário) e algumas trocas de informações que ocorrem verbalmente. No entanto, é notória a fragilidade desse processo, não pela ausência de um sistema computacional, mas pelas características do processo em si. O fato de todo processo ocorrer por meios manuais, não garante uma comunicação plena entre os setores, permitindo o envio de informações equivocadas ou incompletas.

A recuperação de informação é uma outra etapa com problemas evidentes. Caso a paciente já tenha sido atendida naquela unidade de saúde, é possível que a equipe médica solicite seu histórico (armazenado no setor de arquivamento). Fica sob responsabilidade da recepção realizar essa busca nos arquivos da unidade de saúde. Normalmente essas fichas são ordenadas pelo nome da paciente, facilitando uma busca posterior. Entretanto, não existem garantias de que as fichas de fato estejam ordenadas no momento na busca, além disso, existe a possibilidade de extravio dessas informações. Outros agravantes são a falta de garantia de unicidade das fichas para uma mesma paciente e a utilização de diferentes vocabulários em cada setor, resultando em uma diferença semântica entre eles.

Ainda tratando-se da recuperação de informação, outra limitação notada é a impossibilidade de obter registros ou informações de outras bases de dados. Com isso, os profissionais de saúde têm dificuldade em acessar informações das pacientes oriundas de outras unidades de saúde. Além disso, uma integração com outras unidades ou com uma base de dados abertos local poderia agregar mais para a obtenção de indicadores da rede de saúde. Esse tipo de informação é essencial para um gestor da saúde, pois ajuda a entender onde existe demanda de investimento, entre outras ações possíveis.

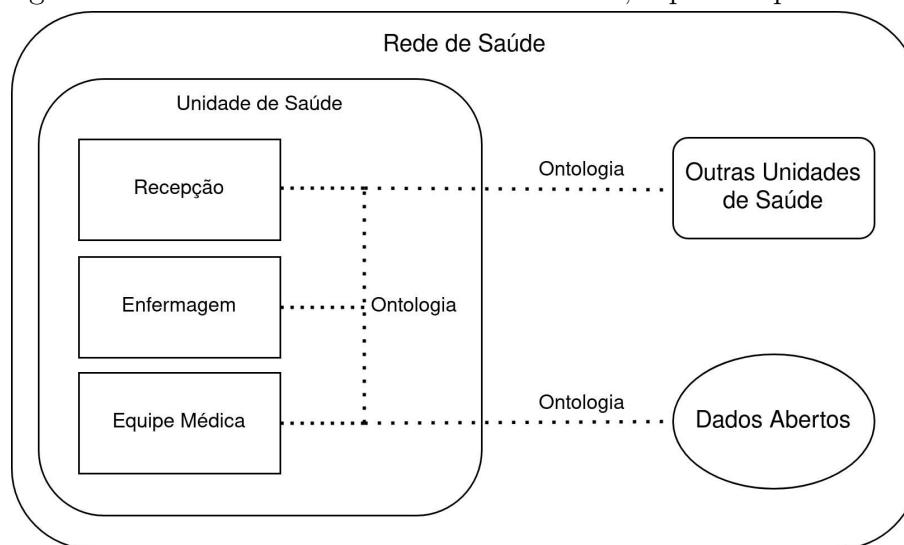
Para atender o escopo deste trabalho, foram selecionadas algumas das necessidades de informação levantadas por Costa et al. (2020). As necessidades selecionadas têm relação com os objetivos específicos deste trabalho e estão listadas a seguir:

- Informações completas e precisas sobre o histórico médico da paciente;
- Informações agregadas sobre atendimentos realizados;
- Indicadores sobre a execução do ACR;

- Informações sobre a paciente proveniente de outras unidades de saúde.

A partir das necessidades de informação, procedeu-se com a elaboração das questões de competência para a ontologia proposta, através das quais as necessidades de informação serão atendidas. Essas questões de competência foram baseadas nas necessidades de informação, de acordo com o cenário proposto, ilustrado na Figura 16. A ontologia deve minimizar as diferenças semânticas entre os setores da unidade de saúde e, também, ser integrável com outras bases de dados da saúde, para assim auxiliar no levantamento de indicadores de gerenciamento.

Figura 16 – Cenário com a Unidade de Saúde, separado por setores.



Fonte: Autoria própria

As questões de competência foram listadas no Quadro 1 de acordo com os setores envolvidos, segundo o cenário apresentado na Figura 16. Dessa forma, é mostrado o trajeto da informação desejada para cada questão.

O Quadro 2 mostra quais necessidades de informação estão relacionadas com cada questão de competência levantada.

### 5.1.2 Considerações de Reuso

Para finalidade de reutilização, inicialmente foram analisadas ontologias utilizadas em trabalhos na área da saúde, assim como algumas oriundas dos trabalhos selecionados para o Capítulo 2 (Trabalhos Correlatos). Algumas ontologias eram recorrentes em diversos trabalhos, o que contribuiu para sua seleção. Por fim, as ontologias selecionadas para reutilização foram: *Ontology for General Medical Science*<sup>1</sup>, *Phenotype And Trait Ontology*<sup>2</sup>,

<sup>1</sup><https://bioportal.bioontology.org/ontologies/OGMS>

<sup>2</sup><http://obofoundry.org/ontology/pato.html>



Quadro 1 – Questões de competência levantadas e setores do cenário envolvidos.

<b>Questão de Competência</b>	<b>Setores Envolvidos</b>
QC1: Quais remédios são utilizados pela paciente?	Enfermagem e Médico
QC2: Qual situação da gravidez da paciente?	Enfermagem e Médico
QC3: Quais sintomas apresentados nas últimas consultas?	Outras Unidades de Saúde
QC4: Quais sintomas estão associados com determinados diagnósticos?	Unidade de Saúde e Dados Abertos
QC5: A classificação de risco está adequada (de acordo com o manual) para os sintomas e sinais vitais apresentados?	Unidade de Saúde e Dados Abertos
QC6: Quais diagnósticos já foram feitos para uma paciente?	Outras Unidades de Saúde

Fonte: Autoria própria

Quadro 2 – Questões de competência relacionadas com necessidade de informação.

<b>Necessidade de Informação</b>	<b>Questão de Competência</b>
Informações completas e precisas sobre o histórico médico da paciente	QC1 e QC2
Informações agregadas sobre os atendimentos realizados	QC4
Indicadores sobre a execução do ACR	QC5
Informações sobre a paciente proveniente de outras unidades de saúde	QC3 e QC6

Fonte: Autoria própria

*Human Disease Ontology*<sup>3</sup>, *Drug Ontology*<sup>4</sup> e o *Friend of a Friend*<sup>5</sup>.

### 5.1.3 Enumeração dos Termos

Os primeiros termos elaborados para a ontologia foram: registro hospitalar e paciente. Isso porque cada FACR é um registro novo e nele está contido uma paciente. Pensando nesses dois termos, os demais são derivados deles, pois, primeiro foram elaborados os termos mais abstratos e, em seguida, os mais específicos. Todos os termos levantados, para registro hospitalar e paciente, estão listados a seguir:

- Dados do Registro:
  - Data e Hora do Acolhimento;
  - Data e Hora do Atendimento;

<sup>3</sup><http://www.obofoundry.org/ontology/doid.html>

<sup>4</sup><https://bioportal.bioontology.org/ontologies/DRON>

<sup>5</sup><http://www.foaf-project.org/>

- Data e Hora da Chegada na Unidade;
- Data e Hora do Recolhimento;
- Tipo de Atendimento;
- Registro do COREN - Conselho Regional de Enfermagem;
- Registro do CRM - Conselho Regional de Medicina.
- Dados Pessoais:
  - Nome da Paciente;
  - Idade;
  - Sexo;
  - CNS - Cartão Nacional de Saúde.
- Alergia:
  - Nome da alergia
- Comorbidade:
  - Descrição da comorbidade
- Medicamento em uso:
  - Nome do medicamento
- Dados Obstétricos:
  - Quantidade de Gestações;
  - Quantidade de Partos Normais;
  - Quantidade de Cesarianas;
  - Quantidade de Abortos;
  - Idade Gestacional;
  - DUM - Data da Última Menstruação;
  - DPP - Data Provável do Parto.
- Sintomas:
  - Queixa;
  - Intensidade da Dor;
  - Descrição;
  - Movimento Fetal;
  - Sangramento Vaginal;
  - Contrações Uterinas;
  - Perda de Líquidos.
- Sinais Vitais:
  - Pressão Arterial;
  - Pulso;
  - Frequência Respiratória;
  - Temperatura;
  - Saturação.
- Classificação de Risco:

- Categoria da Classificação.
- Conduta:
  - Conduta Definida.
- Diagnóstico:
  - Descrição do diagnóstico.

#### 5.1.4 Definição das Classes

De forma semelhante à enumeração dos termos, as primeiras classes elaboradas foram Registro Hospitalar e Paciente. As duas classes têm uma relação, onde a classe Registro Hospitalar tem uma Paciente. Na ontologia *Friend of a Friend*, uma das selecionadas, temos a classe *Person* ou Pessoa. Entende-se que paciente é um pessoa, logo a classe Paciente foi definida como subclasse de Pessoa. No FACR, médico e enfermeiro possuem apenas uma referência quanto ao número dos respectivos conselhos (CRM e COREM). Porém, entende-se que os dois conceitos também podem ser tratados como subclasses de Pessoa.

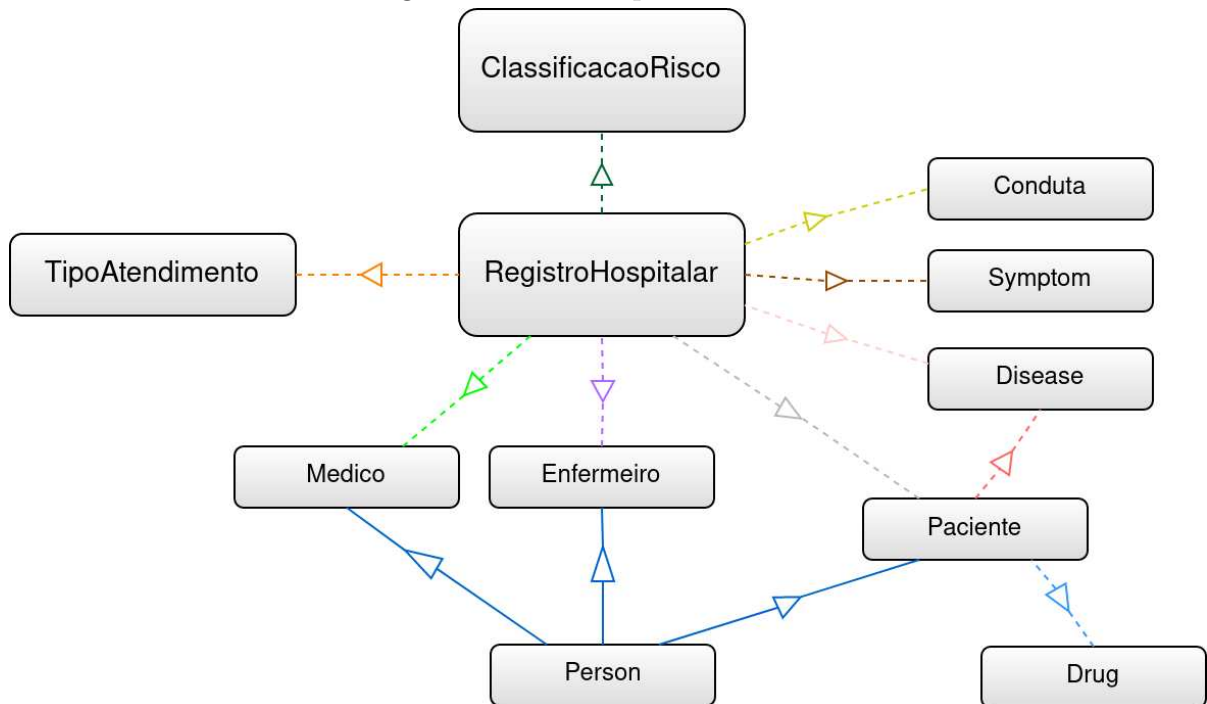
Alguns dos termos levantados anteriormente foram definidos como classes nesta etapa. A classe referente ao sintoma possui definição em duas das ontologias selecionadas, com o termo *Symptom*. Foi selecionada a classe definida na ontologia *Human Disease Ontology*. Comorbidade é definida com o termo *Disease* na ontologia *Human Disease Ontology*. Esta classe também abrange alergias (algumas doenças são alergias) e diagnóstico, logo, a classe contempla três termos levantados (Comorbidade, Alergia e Diagnóstico). Medicamento em uso (remédio) está definido como *Drug* na ontologia *Drug Ontology*. Não foram encontradas, nas ontologias, classes para os conceitos de Classificação de Risco, Tipo de Atendimento e Conduta, logo, estes foram definidos pelo autor. A hierarquia das classes pode ser vista na Figura 17. As linhas contínuas representam relações de subclasse, enquanto as tracejadas as propriedades de objeto, que serão apresentadas na seção a seguir.

#### 5.1.5 Definição das Propriedades

A maior parte das propriedades de dados foram definidas a partir dos termos mais específicos, levantados na terceira etapa. Também foi inserido como propriedade o CID - Classificação Internacional de Doenças para padronizar a comorbidade, sintomas, diagnóstico e conduta (quando for um procedimento médico). Algumas delas já existiam nas ontologias importadas e foram aproveitadas, como *age* e *name*, o conjunto final de propriedades está representado na lista abaixo:

- *data\_hora\_acolhimento*;
- *data\_hora\_atendimento*;
- *data\_hora\_recepcao*;
- *data\_hora\_recolhimento*;
- *tipo\_atendimento*;

Figura 17 – Hierarquia das classes.



Fonte: Autoria própria

- *COREN*;
- *CRM*;
- *name (foaf:name)*;
- *age (foaf:age)*;
- *gender (foaf:gender)*;
- *cartao\_nacional\_saude*
- *quantidade\_gestacoes*;
- *quantidade\_parto\_normal*;
- *quantidade\_cesariana*;
- *quantidade\_aborto*;
- *idade\_gestacional*;
- *DUM*;
- *DPP*;
- *queixa*;
- *intensidade\_dor*;
- *descrição\_sintoma*;
- *movimento\_fetal*;
- *perda\_liquido*;
- *pressao\_arterial*;
- *pulso*;

- *frequencia\_respiratoria*;
- *temperatura*;
- *saturacao*;
- *classificacao\_risco*;
- *CID*
- *conduta*.

As propriedades de objeto foram definidas manualmente, exceto "has\_symptom" que já estava definida na ontologia *Human Disease Ontology*. As propriedades, bem como os respectivos domínios e abrangências, estão listadas no Quadro 3.

Quadro 3 – Propriedades de objeto.

<b>Propriedade de Objeto</b>	<b>Domínio</b>	<b>Abrangência</b>
<i>has_symptom</i> ( <i>doid:has_symptom</i> )	<i>RegistroHospitalar</i>	<i>Symptom</i>
<i>temClassificacaoRisco</i>	<i>RegistroHospitalar</i>	<i>ClassificacaoRisco</i>
<i>temComorbidade</i>	<i>Paciente</i>	<i>Disease</i>
<i>temConduta</i>	<i>RegistroHospitalar</i>	<i>Conduta</i>
<i>temDiagnostico</i>	<i>RegistroHospitalar</i>	<i>Disease</i>
<i>temEnfermeiro</i>	<i>RegistroHospitalar</i>	<i>Enfermeiro</i>
<i>temMedicamento</i>	<i>Paciente</i>	<i>Drug</i>
<i>temMedico</i>	<i>RegistroHospitalar</i>	<i>Medico</i>
<i>temPaciente</i>	<i>RegistroHospitalar</i>	<i>Paciente</i>
<i>temTipoAtendimento</i>	<i>RegistroHospitalar</i>	<i>TipoAtendimento</i>

Fonte: Autoria própria

#### 5.1.6 Definição das Restrições

As restrições para as propriedades de objetos estão listadas no Quadro 4. Apenas a restrição funcional foi utilizada nas propriedades, por isso as demais foram omitidas.

As restrições definidas para as propriedades de dados estão listadas no Quadro 5, onde para cada propriedade é atribuído um tipo de dado e domínio.

Foi definido também a cardinalidade para as relações que envolvem a classe *RegistroHospitalar*, conforme pode ser visto na Figura 18.

#### 5.1.7 Criação das Instâncias

Primeiro foram criadas as instâncias que deveriam existir antes da ontologia ser populada. Começando pela classe *ClassificacaoRisco*, onde foram criadas as instâncias para cada classificação possível (cinco). Em seguida, a classe *TipoAtendimento*, em que foram instanciados os tipos de atendimento para obstetrícia e ginecologia, ambos presentes no FACR. Após isso, foram elaborados cenários sintéticos que permitissem uma possível avaliação da ontologia.

Quadro 4 – Restrições de propriedades de objeto.

Propriedade de Objeto	Funcional
<i>has_symptom (doid:has_symptom)</i>	Não
<i>temClassificacaoRisco</i>	Não
<i>temComorbidade</i>	Não
<i>temConduta</i>	Sim
<i>temDiagnostico</i>	Sim
<i>temEnfermeiro</i>	Sim
<i>temMedicamento</i>	Não
<i>temMedico</i>	Sim
<i>temPaciente</i>	Sim
<i>temTipoAtendimento</i>	Sim

Fonte: Autoria própria

Figura 18 – Cardinalidades para a classe *RegistroHospitalar*.

Fonte: Autoria própria

O primeiro cenário aborda a obtenção de indicadores e outras informações a partir da integração da base da unidade de saúde com os dados abertos de Curitiba. Para tal, uma outra ontologia foi desenvolvida para representar os dados abertos. Com isso, foi possível testar como ocorreria uma integração desse tipo caso os dados abertos tivessem uma base desenvolvida em RDF ou OWL. Espera-se que a consulta seja capaz de responder as questões de competência QC4 e QC5, que atendem, respectivamente, as seguintes necessidades de informação: "Informações agregadas sobre atendimentos realizados" e "Indicadores sobre a execução do ACR". Para a QC5, a ideia é avaliar se através de uma consulta na ontologia é possível identificar erros na execução do ACR. Essa verificação é baseada no manual definido por Brasil (2017). Os dados dos registros utilizados nesta etapa estão listados a seguir (as pacientes sem registro obstétrico tiveram

Quadro 5 – Restrições de tipo de dado e domínio.

<b>Propriedade de Dado</b>	<b>Tipo de Dado</b>	<b>Domínio</b>
<i>data_hora_acolhimento</i>	<i>dateTimeStamp</i>	<i>RegistroHospitalar</i>
<i>data_hora_atendimento</i>	<i>dateTimeStamp</i>	<i>RegistroHospitalar</i>
<i>data_hora_chegada</i>	<i>dateTimeStamp</i>	<i>RegistroHospitalar</i>
<i>data_hora_recolhimento</i>	<i>dateTimeStamp</i>	<i>RegistroHospitalar</i>
<i>tipo_atendimento</i>	<i>String</i>	<i>TipoAtendimento</i>
<i>COREN</i>	<i>String</i>	<i>Enfermeiro</i>
<i>CRM</i>	<i>String</i>	<i>Medico</i>
<i>name (foaf:name)</i>	<i>Literal</i>	<i>Thing</i>
<i>age (foaf:age)</i>	<i>Literal</i>	<i>Person</i>
<i>gender (foaf:gender)</i>	<i>Literal</i>	<i>Person</i>
<i>cartao_nacional_saude</i>	<i>String</i>	<i>Paciente</i>
<i>quantidade_gestacoes</i>	<i>Short</i>	<i>Paciente</i>
<i>quantidade_parto_normal</i>	<i>Short</i>	<i>Paciente</i>
<i>quantidade_cesariana</i>	<i>Short</i>	<i>Paciente</i>
<i>quantidade_aborto</i>	<i>Short</i>	<i>Paciente</i>
<i>idade_gestacional</i>	<i>Short</i>	<i>Paciente</i>
<i>dum</i>	<i>String</i>	<i>Paciente</i>
<i>dpp</i>	<i>String</i>	<i>Paciente</i>
<i>queixa</i>	<i>String</i>	<i>RegistroHospitalar</i>
<i>intensidade_dor</i>	<i>Short</i>	<i>RegistroHospitalar</i>
<i>descricao_sintoma</i>	<i>String</i>	<i>Symptom</i>
<i>movimento_fetal</i>	<i>Boolean</i>	<i>RegistroHospitalar</i>
<i>perda_liquido</i>	<i>Boolean</i>	<i>RegistroHospitalar</i>
<i>pressao_artorial</i>	<i>String</i>	<i>RegistroHospitalar</i>
<i>pulso</i>	<i>String</i>	<i>RegistroHospitalar</i>
<i>frequencia_respiratoria</i>	<i>String</i>	<i>RegistroHospitalar</i>
<i>saturacao</i>	<i>String</i>	<i>RegistroHospitalar</i>
<i>classificacao_risco</i>	<i>String</i>	<i>ClassificacaoRisco</i>
<i>temperatura</i>	<i>Short</i>	<i>RegistroHospitalar</i>
<i>conduta</i>	<i>String</i>	<i>Conduta</i>
<i>CID</i>	<i>String</i>	<i>Symptom, Disease, Conduta</i>

Fonte: Autoria própria

os dados obstétricos omitidos):

- Registro 1489:
  - Identificador da Paciente: 2270257;
  - Identificador do Enfermeiro: 55589;
  - Identificador do Médico: 3509;
  - Tipo de Atendimento: Ginecologia;
  - CID do Sintoma: D62, N939 e N92.6;

- Classificação de Risco: Amarelo;
- CID do Diagnóstico: D25.1.
- Registro 1502:
  - Identificador da Paciente: 1126909;
  - Identificador do Enfermeiro: 4471;
  - Identificador do Médico: 3210;
  - Tipo de Atendimento: Obstetrícia;
  - CID do Sintoma: N92.6 e N939;
  - Classificação de Risco: Amarelo;
  - CID do Diagnóstico: D25.1.
- Registro 1511:
  - Identificador da Paciente: 1592807;
  - Identificador do Enfermeiro: 93381;
  - Identificador do Médico: 91621;
  - Tipo de Atendimento: Ginecologia;
  - CID do Sintoma: N92.6;
  - Classificação de Risco: Amarelo;
  - CID do Diagnóstico: D25.1.
- Registro 1520:
  - Identificador da Paciente: 1125776;
  - Identificador do Enfermeiro: 4471;
  - Identificador do Médico: 3210;
  - Tipo de Atendimento: Ginecologia;
  - CID do Sintoma: N939 e N92.6;
  - Classificação de Risco: Amarelo;
  - CID do Diagnóstico: D25.1.
- Registro 17487:
  - Identificador da Paciente: 8333072;
  - Identificador do Enfermeiro: 37994;
  - Identificador do Médico: 12539;
  - Tipo de Atendimento: Obstetrícia;
  - CID do Sintoma: N939 e O62.4;
  - Classificação de Risco: Laranja;
  - Intensidade da Dor: 7;
  - Temperatura: 40.
- Registro 3070:
  - Identificador da Paciente: 6823578;
  - Identificador do Enfermeiro: 4471;
  - Identificador do Médico: 3210;



- Tipo de Atendimento: Obstetrícia;
- CID do Sintoma: N939 e O62.4;
- Classificação de Risco: Laranja;
- Intensidade da Dor: 7;
- Temperatura: 40.
- Registro 985:
  - Identificador da Paciente: 1125776;
  - Identificador do Enfermeiro: 93381;
  - Identificador do Médico: 91621;
  - Tipo de Atendimento: Obstetrícia;
  - CID do Sintoma: N939 e O62.4;
  - Classificação de Risco: Amarelo;
  - Intensidade da Dor: 8;
  - Temperatura: 40.
- Paciente 8333072:
  - Idade: 33;
  - Gênero: Feminino;
  - Quantidade parto normal: 1;
  - Quantidade cesariana: 0;
  - Idade Gestacional: 29;
  - DUM: 19/01/2020;
  - DPP: 26/10/2020;
  - Quantidade aborto: 0;
  - Quantidade gestações: 1.
- Paciente 6823578:
  - Idade: 32;
  - Gênero: Feminino;
  - Quantidade parto normal: 0;
  - Quantidade cesariana: 0;
  - Idade Gestacional: 22;
  - DUM: 12/03/2020;
  - DPP: 19/12/2020;
  - Quantidade aborto: 0;
  - Quantidade gestações: 0.
- Paciente 1125776:
  - Idade: 46;
  - Gênero: Feminino;
  - Quantidade parto normal: 0;
  - Quantidade cesariana: 0;

- Idade Gestacional: 23;
- DUM: 15/03/2020;
- DPP: 22/12/2020;
- Quantidade aborto: 0;
- Quantidade gestações: 0.
- Paciente 2270257:
  - Idade: 24;
  - Gênero: Feminino.
- Paciente 1126909:
  - Idade: 50;
  - Gênero: Feminino;
  - Quantidade parto normal: 0;
  - Quantidade cesariana: 1;
  - Idade Gestacional: 30;
  - DUM: 05/01/2020;
  - DPP: 12/10/2020;
  - Quantidade aborto: 0;
  - Quantidade gestações: 1.
- Paciente 1592807:
  - Idade: 25;
  - Gênero: Feminino.

O segundo cenário trata informações sobre uma paciente proveniente de outras unidades de saúde, neste caso também foi utilizada a ontologia desenvolvida para os dados abertos de Curitiba. O objetivo deste cenário é testar a possibilidade de integração dos dados locais com outras bases de dados de maternidade. Este cenário deve ajudar a responder as questões de competência QC3 e QC6, que atendem a seguinte necessidade de informação: "Informações sobre a paciente proveniente de outras unidades de saúde". Os dados dos registros utilizados nesta etapa estão listados a seguir (a paciente está sem registro obstétrico, por isso os dados obstétricos foram omitidos):

- Registro 768669:
  - Identificador da Paciente: 2447303;
  - Identificador do Enfermeiro: 95164;
  - Identificador do Médico: 42179;
  - Tipo de Atendimento: Ginecologia;
  - CID do Sintoma: N390;
  - Classificação de Risco: Verde;
  - CID do Diagnóstico: D251.
- Registro 769628:
  - Identificador da Paciente: 2447303;

- Identificador do Enfermeiro: 95164;
- Identificador do Médico: 13724;
- Tipo de Atendimento: Ginecologia;
- CID do Sintoma: A09;
- Classificação de Risco: Azul.
- Registro 773807:
  - Identificador da Paciente: 2447303;
  - Identificador do Enfermeiro: 51154;
  - Identificador do Médico: 42179;
  - Tipo de Atendimento: Ginecologia;
  - CID do Sintoma: N939;
  - Classificação de Risco: Amarelo;
  - CID do Diagnóstico: O20.0.
- Registro 769141:
  - Identificador da Paciente: 2447303;
  - Identificador do Enfermeiro: 55589;
  - Identificador do Médico: 3210;
  - Tipo de Atendimento: Ginecologia;
  - CID do Sintoma: N939;
  - Classificação de Risco: Amarelo;
  - CID do Diagnóstico: B373.
- Registro 769141:
  - Identificador da Paciente: 2447303;
  - Identificador do Enfermeiro: 61007;
  - Identificador do Médico: 3509;
  - Tipo de Atendimento: Ginecologia;
  - CID do Sintoma: N939;
  - Classificação de Risco: Amarelo;
  - CID do Diagnóstico: O20.0.
- Paciente 2447303:
  - Idade: 45;
  - Gênero: Feminino.

O terceiro cenário é focado na troca de informações entre os setores da unidade de saúde, possibilitando avaliar a redução da diferença semântica entre eles. Para as instâncias deste cenário foram usados alguns registros dos dados abertos de Curitiba, com a utilização de dados sintéticos para os campos não contemplados pela base. O objetivo deste cenário é mostrar que, com a ontologia, a dificuldade em buscar informações de outros setores é significativamente reduzida. Os campos disponíveis nos dados abertos de Curitiba contemplam as seguintes propriedades de dados da ontologia para a maternidade:

- *data\_hora\_acolhimento*;
- *data\_hora\_atendimento*;
- *tipo\_atendimento*;
- *age* (*foaf:age*);
- *gender* (*foaf:gender*);
- *CID* (sintoma);
- *CID* (diagnóstico ou procedimento).

Apesar de não mostrar o COREN, CRM ou CNS do atendimento, os dados abertos trazem um código único para identificação do médico, enfermeiro e paciente.

Esse cenário envolve as questões de competência QC1 e QC2, que atendem a seguinte necessidade de informação: "Informações completas e precisas sobre o histórico médico da paciente". Os dados dos registros utilizados nesta etapa estão listados a seguir:

- Registro 564016:
  - Identificador da Paciente: 2270257;
  - Identificador do Enfermeiro: 78583;
  - Identificador do Médico: 28857;
  - Tipo de Atendimento: Obstetrícia;
  - CID do Sintoma: R509;
  - Classificação de Risco: Amarelo.
- Registro 908707:
  - Identificador da Paciente: 1592807;
  - Identificador do Enfermeiro: 55589;
  - Identificador do Médico: 3209;
  - Tipo de Atendimento: Obstetrícia;
  - CID do Sintoma: N939;
  - Classificação de Risco: Amarelo.
- Paciente 2270257:
  - Idade: 24;
  - Gênero: Feminino;
  - Medicamentos em uso: Carbonato de lítio e Cloridrato de Paroxetina;
  - Quantidade parto normal: 0;
  - Quantidade cesariana: 0;
  - Idade gestacional: 15;
  - DUM: 23/02/2020;
  - DPP: 30/11/2020;
  - Quantidade aborto: 0;
  - Quantidade gestações: 0.
- Paciente 1592807:
  - Idade: 25;

- Gênero: Feminino;
- Quantidade parto normal: 0;
- Quantidade cesariana: 0;
- Idade gestacional: 16;
- DUM: 10/01/2020;
- DPP: 17/10/2020;
- Quantidade aborto: 0;
- Quantidade gestações: 0.

## 5.2 Validação da Ontologia

Nesta seção foram elaboradas as consultas em SPARQL para testar os cenários listados anteriormente. Para execução das consultas, foram utilizados os seguintes prefixos:

- "PREFIX owl: <http://www.w3.org/2002/07/owl#>";
- "PREFIX doid: <http://purl.obolibrary.org/obo/doid#>";
- "PREFIX pat\_hc: <http://www.semanticweb.org/rodrigo/maternidade\_pa#>";
- "PREFIX dac: <http://www.semanticweb.org/rodrigo/d\_a\_curitiba#>".

O prefixo *pat\_hc* é referente à ontologia desenvolvida para a maternidade. Já o *dac* é referente aos dados abertos de Curitiba. O prefixo *doid* é de uma ontologia importada, enquanto o *owl* é sobre a linguagem de descrição OWL.

### 5.2.1 Primeiro Cenário

A seguir estão listadas as consultas elaboradas para o primeiro cenário. Este cenário tem relação com o objetivo específico: "Possibilitar a integração com outras bases de dados de maternidades de alto risco", visto que as consultas são sobre duas bases de dados distintas. A Figura 19 mostra a consulta elaborada para responder a questão de competência: "Quais sintomas estão associados com determinados diagnósticos?". Os sintomas devem estar ordenados pela quantidade, dessa maneira os sintomas mais recorrentes, para um determinado diagnóstico, devem aparecer primeiro no resultado da consulta.

Figura 19 – Consulta para QC4.

```
SELECT ?symptom_cid (COUNT (?symptom_cid) as ?qtd) WHERE { { ?temDiag owl:equivalentProperty pat_hc:temDiagnostico
, ?registro_h ?temDiag pat_hc:disease_cid_d25.1 , ?hasSymp owl:equivalentProperty doid:has_symptom , ?registro_h
?hasSymp ?symptom , ?symptom pat_hc:cid ?symptom_cid , } UNION { ?temDiag owl:equivalentProperty
dac:temDiagnosticoDa , ?registro_h ?temDiag dac:diagnostico_cid_d25.1 , ?temSintoma owl:equivalentProperty
dac:temSintoma , ?registro_h ?temSintoma ?sintoma , ?sintoma dac:CID ?symptom_cid , } } GROUP BY ?symptom_cid ORDER
BY DESC(?qtd)
```

Fonte: Autoria própria

O resultado da execução da consulta anterior pode ser visto na Figura 20. Na figura é possível ver quais os sintomas mais comuns para o diagnóstico escolhido, neste

caso leiomioma intramural do útero (CID: D251). Esse tipo de informação pode contribuir, por exemplo, para definição de um diagnóstico.

Figura 20 – Resultado da Consulta para QC4.

<b>?symptom_cid</b>	<b>?qtd</b>
N92.6	4
N939	3
D62	1

Fonte: Autoria própria

A Figura 21 mostra a consulta elaborada para a questão de competência: "A classificação de risco está adequada (de acordo com o manual) para os sintomas e sinais vitais apresentados?".

Figura 21 – Consulta para QC5.

```
SELECT ?registro_h ?cr_cor WHERE { { ?temClassRisc owl:equivalentProperty pat_hc:temClassificacaoRisco . ?registro_h ?temClassRisc ?class_risc . ?class_risc pat_hc:ClassificacaoRisco ?cr_cor . ?temPaciente owl:equivalentProperty pat_hc:temPaciente . ?registro_h ?temPaciente ?paciente . ?hasSymp owl:equivalentProperty doi:has_sympptom . ?registro_h ?hasSymp pat_hc:symptom_cid_n939 . ?registro_h ?hasSymp pat_hc:symptom_cid_o62.4 . ?paciente pat_hc:idade_gestacional ?idade_gest . ?registro_h pat_hc:intensidade_dor ?int_dor . ?registro_h pat_hc:temperatura ?temp . FILTER (?idade_gest > 21) . FILTER (?int_dor > 6) . FILTER (?temp > 39) . } UNION { ?temClassRisc owl:equivalentProperty dac:temCR . ?registro_h ?temClassRisc ?class_risc . ?class_risc dac:cr_cor ?cr_cor . ?temPaciente owl:equivalentProperty dac:temPaciente . ?registro_h ?temPaciente ?paciente . ?temSintoma owl:equivalentProperty dac:temSintoma . ?registro_h ?temSintoma dac:sintoma_cid_n939 . ?registro_h ?temSintoma dac:sintoma_cid_o62.4 . ?paciente dac:idade_gest ?idade_gest . ?registro_h dac:intensidade_da_dor ?int_dor . ?registro_h dac:temp ?temp . FILTER (?idade_gest > 21) . FILTER (?int_dor > 6) . FILTER (?temp > 39) . } }
```

Fonte: Autoria própria

O resultado da execução da consulta pode ser visto na Figura 22. Nela é possível observar qual classificação de risco atribuída para cada registro. Dessa forma, é simples detectar uma classificação incorreta, como o registro que recebeu uma classificação amarela. O correto, segundo Brasil (2017), seria uma classificação vermelha, devido a vários fatores do atendimento como intensidade da dor, sintomas e temperatura da paciente.

Figura 22 – Resultado da Consulta para QC5.

<b>?registro_h</b>	<b>?cr_cor</b>
pat_hc:registro_985	Amarelo
pat_hc:registro_17487	Laranja
pat_hc:registro_3070	Laranja

Fonte: Autoria própria

### 5.2.2 Segundo Cenário

As duas consultas a seguir foram elaboradas para o segundo cenário. Ambas consultas têm relação com dois objetivos específicos: "Possibilitar a integração com outras bases de dados de maternidades de alto risco" e "Melhorar aspectos de qualidade da informação na obtenção de históricos médicos". Isso porque ambas consultas são elaboradas para duas bases de dados distintas, além de agregarem mais informações aos históricos das pacientes.

A Figura 23 mostra a consulta elaborada para responder a questão de competência: "Quais sintomas apresentados nas últimas consultas?". Nesta consulta são buscados os sintomas para apenas uma paciente, porém em prefixos diferentes, ou seja, ontologias diferentes. Neste caso, alguns sintomas vêm da ontologia feita para a maternidade e outros vêm da ontologia para os dados abertos de Curitiba.

Figura 23 – Consulta para QC3.

```
SELECT ?cid_sintoma WHERE { { ?hasSymp owl:equivalentProperty
doid:has_sympptom . ?registro_h ?hasSymp ?sympptom . ?sympptom pat_hc:cid
?cid_sintoma . ?temPaciente owl:equivalentProperty pat_hc:temPaciente .
?registro_h ?temPaciente pat_hc:paciente_2447303 . ?registro_h
pat_hc:data_hora_acolhimento ?dh_acolhimento . } UNION { ?temSintoma
owl:equivalentProperty dac:temSintoma . ?registro_h ?temSintoma ?sintoma .
?sintoma dac:CID ?cid_sintoma . ?temPaciente owl:equivalentProperty
dac:temPaciente . ?registro_h ?temPaciente dac:paciente_2447303 . ?registro_h
dac:data_hora_acolhimento ?dh_acolhimento . } } ORDER BY DESC(?dh_acolhimento)
```

Fonte: Autoria própria

O resultado da execução da consulta anterior pode ser visto na Figura 24. Nela estão listados alguns sintomas, todos referenciados através do CID. Esse tipo de informação pode ajudar ao setor médico saber, por exemplo, se os sintomas apresentados naquele atendimento são recorrentes.

Figura 24 – Resultado da Consulta para QC3.

<b>?cid_sintoma</b>
N939^^xsd:string
N939^^xsd:string
R300^^xsd:string
A09^^xsd:string
N390^^xsd:string
N939^^xsd:string

Fonte: Autoria própria

A Figura 25 mostra a consulta elaborada para responder a questão de competência: "Quais diagnósticos já foram feitos para uma paciente?". Esses diagnósticos são oriundos das duas ontologias elaboradas, mas a consulta se refere a apenas uma paciente.

Figura 25 – Consulta para QC6.

```
SELECT ?cid_diag WHERE { { ?temDiag owl:equivalentProperty pat_hc:temDiagnostico
. ?registro_h ?temDiag ?diagnostico . ?diagnostico pat_hc:cid ?cid_diag . ?temPacient
e owl:equivalentProperty pat_hc:temPaciente . ?registro_h ?temPaciente
pat_hc:paciente_2447303 . } UNION { ?temDiag owl:equivalentProperty
dac:temDiagnostico . ?registro_h ?temDiag ?diagnostico . ?diagnostico dac:CID
?cid_diag . ?temPaciente owl:equivalentProperty dac:temPaciente . ?registro_h
?temPaciente dac:paciente_2447303 . } }
```

Fonte: Autoria própria

O resultado da execução da consulta anterior pode ser visto na Figura 26. Nela é possível visualizar os diagnósticos feitos para aquela paciente. Isso pode contribuir durante a definição de um diagnóstico, levando em consideração os que já foram realizados.

Figura 26 – Resultado da Consulta para QC6.

<b>?cid_diag</b>
O20.0^^xsd:string
B373^^xsd:string
D251^^xsd:string
O20.0^^xsd:string

Fonte: Autoria própria

### 5.2.3 Terceiro Cenário

As duas consultas a seguir foram elaboradas para o terceiro cenário. Por se tratarem de informações que transitam entre diferentes setores da unidade de saúde, estas consultas têm relação com o objetivo específico: "Reduzir a diferença semântica das bases de dados dos diferentes setores da unidade de saúde". Como as consultas agregam mais informações ao histórico da paciente, também se relacionam com o objetivo específico: "Melhorar aspectos de qualidade da informação na obtenção de históricos médicos". A Figura 27 mostra a consulta elaborada para responder a questão de competência: "Quais remédios são utilizados pela paciente?".

O resultado da execução da consulta anterior pode ser visto na Figura 28, onde foram retornados dois medicamentos em uso pela paciente. Dessa forma, a equipe médica poderia verificar a utilização de medicamentos contraindicados para gestantes e pedir a suspensão imediata do tratamento.



Figura 27 – Consulta para QC1.

```
SELECT DISTINCT ?drug_name WHERE { ?temMed owl:equivalentProperty
pat_hc:temMedicamento , pat_hc:paciente_2270257 ?temMed ?drug . ?drug
foaf:name ?drug_name . }
```

Fonte: Autoria própria

Figura 28 – Resultado da Consulta para QC1.

<b>?drug_name</b>
Carbonato de Lítio^^rdfs:Literal
Cloridrato de Paroxetina^^rdfs:Literal

Fonte: Autoria própria

A Figura 29 mostra a consulta elaborada para responder a questão de competência: "Qual situação da gravidez da paciente?". As informações selecionadas na consulta foram baseadas no FACR.

Figura 29 – Consulta para QC2.

```
SELECT ?qtd_gest ?qtd_part ?qtd_cesa ?qtd_abo ?idade_gest ?dum ?dpp
WHERE { ?paciente owl:sameAs pat_hc:paciente_1592807 . ?paciente
pat_hc:idade_gestacional ?idade_gest . ?paciente pat_hc:dpp ?dpp .
?paciente pat_hc:dum ?dum . ?paciente pat_hc:quantidade_gestacoes
?qtd_gest . ?paciente pat_hc:quantidade_parto_normal ?qtd_part .
?paciente pat_hc:quantidade_cesariana ?qtd_cesa . ?paciente
pat_hc:quantidade_aborto ?qtd_abo . }
```

Fonte: Autoria própria

O resultado da execução da consulta anterior pode ser visto na Figura 30. Nela, são mostradas algumas informações sobre a gestação como a idade gestacional, data provável do parto, etc.

Figura 30 – Resultado da Consulta para QC2.

<b>?qtd_gest</b>	<b>?qtd_part</b>	<b>?qtd_cesa</b>	<b>?qtd_abo</b>	<b>?idade_g...</b>	<b>?dum</b>	<b>?dpp</b>
0	0	0	0	16	10/01/2020	10/09/2020

Fonte: Autoria própria

## 6 Considerações Finais

A tomada de decisão na área da saúde é um fator determinante na prestação de auxílio a pacientes, mas para que esta seja eficaz é importante que existam mecanismos para gerenciamento de informações. O gerenciamento por meios manuais aumenta as chances de problemas relacionados à qualidade da informação, além de limitar a capacidade de recuperação dessas informações. A proposta apresentada buscou não apenas inserir um sistema computacional em uma unidade de saúde, mas, também pensar em uma camada com mecanismos que contribuíssem com o processo de decisão.

A obtenção de indicadores sobre a execução do protocolo de classificação de risco, por meio de consultas sobre a ontologia, mostrou um potencial promissor na gestão da informação. A construção de indicadores integráveis pode ajudar os gestores da saúde a verificarem a qualidade dos atendimentos nas unidades e, se necessário, buscarem melhorias. A proposta da abordagem também foi capaz de mostrar possíveis associações entre informações relevantes dos pacientes, como sintomas e diagnósticos, além de meios simples para recuperação de informações locais ou oriundas de outras unidade de saúde.

A utilização dos dados abertos de Curitiba, através de uma ontologia, demonstrou potencial avanço na integração de dados. Porém, existem, ainda, muitas limitações quanto a utilização desse tipo de base de dados, por exemplo, problemas de inconsistência e a quantidade limitada de campos disponíveis. Nos dados abertos de Curitiba, por exemplo, por uma questão de privacidade, não é possível descobrir quem de fato é o paciente em um determinado atendimento. Esse tipo de informação só poderia ser obtida caso o órgão responsável concedesse. Além do mais, frequentemente são utilizados formatos de arquivos que não facilitam a integração de dados, por exemplo, CSV e XLSX. Sendo necessário a realização de conversões entre bases de dados para que seja possível uma integração entre elas.

Por fim, os resultados mostram que a ontologia conseguiu melhorar aspectos de recuperação da informação, problemas semânticos, etc. Dito isto, é possível concluir que o objetivo geral do trabalho foi alcançado, visto que foi elaborada uma ontologia para a unidade de saúde proposta e esta é capaz de tornar sua base de dados integrável. Porém, para utilização em um ambiente real mais testes seriam necessários com dados inteiramente reais. Por isso, trabalhos futuros nessa área poderiam explorar outras camadas para uma aplicação semelhante, buscar outras possíveis fontes para alimentar a ontologia, aprofundar na integração semântica com os dados abertos, entre outros aspectos não abordados neste trabalho.

## Referências

- ADDLESEE, A. **Understanding Linked Data Formats**. 2018. <<https://medium.com/wallscope/understanding-linked-data-formats-rdf-xml-vs-turtle-vs-n-triples-eb931dbe9827>>. Acessado em 10/05/2020. Citado na página 30.
- ALKAHTANI, N. H. et al. A semantic multi-agent system to exchange information between hospitals. **Procedia Computer Science**, Elsevier, v. 109, p. 704–709, 2017. Citado na página 23.
- AZAMI, I. E.; MALKI, M. O. C.; TAHON, C. Integrating hospital information systems in healthcare institutions: a mediation architecture. **Journal of medical systems**, Springer, v. 36, n. 5, p. 3123–3134, 2012. Citado na página 23.
- BERNERS-LEE, T. et al. The semantic web. **Scientific american**, New York, NY, USA:, v. 284, n. 5, p. 28–37, 2001. Citado na página 31.
- BERNERS-LEE, T. J. **Design Issues for the World Wide Web**. 2000. <<https://www.w3.org/DesignIssues/>>. Acessado em 17/05/2019. Citado na página 28.
- BERNERS-LEE, T. J. **Linked Data**. 2009. <<https://www.w3.org/DesignIssues/LinkedData.html>>. Acessado em 06/03/2020. Citado 2 vezes nas páginas 33 e 34.
- BETTIN, G. C. d. S.; TEIXEIRA, H. M. P.; BERARDI, R. C. Modelagem de uma ontologia de contexto no domínio de urgência e emergência médica hospitalar. In: **ONTOBRAS**. [S.l.: s.n.], 2016. Citado na página 21.
- BLOOMBERG, T. A. **Stefaan Verhulst: Unlocking Private Data for Public Good**. 2017. <<https://www.techatbloomberg.com/blog/stefaan-verhulst-unlocking-private-data-public-good/>>. Acessado em 30/05/2019. Citado na página 32.
- BRASIL. **Acolhimento e Classificação de Risco nos Serviços de Urgência**. 2009. <[http://bvsmms.saude.gov.br/bvs/publicacoes/acolhimento\\_classificacao\\_risco\\_servico\\_urgencia.pdf](http://bvsmms.saude.gov.br/bvs/publicacoes/acolhimento_classificacao_risco_servico_urgencia.pdf)>. Acessado em 20/05/2019. Citado na página 14.
- BRASIL. **LEI Nº 12.527, DE 18 DE NOVEMBRO DE 2011**. 2011. <[http://www.planalto.gov.br/ccivil\\_03/\\_ato2011-2014/2011/lei/112527.htm](http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/112527.htm)>. Acessado em 10/05/2019. Citado 2 vezes nas páginas 15 e 16.
- BRASIL. **Manual de Acolhimento e Classificação de Risco em Obstetrícia**. 2017. <[https://bvsmms.saude.gov.br/bvs/publicacoes/manual\\_acolhimento\\_classificacao\\_risco\\_obstetricia\\_2017.pdf](https://bvsmms.saude.gov.br/bvs/publicacoes/manual_acolhimento_classificacao_risco_obstetricia_2017.pdf)>. Acessado em 10/06/2020. Citado 2 vezes nas páginas 46 e 54.
- BRASIL. **O que são dados abertos?** 2019. <<http://dados.gov.br/pagina/dados-abertos>>. Acessado em 30/05/2019. Citado 3 vezes nas páginas 15, 16 e 32.
- BRASIL. **Portal Brasileiro de Dados Abertos**. 2020. <<http://www.dados.gov.br>>. Acessado em 10/03/2020. Citado na página 24.

- CHIGNARD, S. A brief history of open data. **Paris Tech Review**, v. 29, 2013. Citado na página 33.
- COSTA, G. R. d. et al. **Gestão da informação apoiada na gestão por processos em maternidade de alto-risco**. Dissertação (Mestrado) — Universidade Tecnológica Federal do Paraná, 2020. Citado 4 vezes nas páginas 16, 35, 38 e 39.
- COSTA, G. R. da; STADNIK, A. M. W. Health information systems applied to maternity hospitals: A systematic review. **XXVI Brazilian Congress on Biomedical Engineering**, v. 1, 2019. Citado na página 14.
- CURITIBA, P. M. **Dados Abertos Curitiba**. 2019. <<http://www.curitiba.pr.gov.br/dadosabertos/>>. Acessado em 10/05/2019. Citado 2 vezes nas páginas 16 e 33.
- EMC. **The Digital Universe in 202: Big Data**. 2012. <<http://www.emc.com/leadership/digital-universe/index.html>>. Acessado em 30/05/2019. Citado na página 31.
- FOUNDATION, O. K. **What is open?** 2019. <<https://okfn.org/opendata/>>. Acessado em 30/05/2019. Citado na página 32.
- GRUBER, T. R. **Toward Principles for the Design of Ontologies Used for Knowledge Sharing**. 1993. <<https://tomgruber.org/writing/onto-design.pdf>>. Acessado em 04/05/2020. Citado 2 vezes nas páginas 26 e 28.
- GUARINO, N. Formal ontology, conceptual analysis and knowledge representation. **International journal of human-computer studies**, Elsevier, v. 43, n. 5-6, p. 625–640, 1995. Citado na página 27.
- GUIMARÃES, R. C. M. Nomeação de elementos ontológicos para criação de ontologias: uma proposta metodológica. 2015. Citado na página 30.
- GUIZZARDI, G. Ontological foundations for structural conceptual models. 2005. Citado 3 vezes nas páginas 27, 30 e 31.
- HORROCKS, I. Ontologies and the semantic web. **Commun. ACM**, v. 51, n. 12, p. 58–67, 2008. Citado na página 31.
- ISOTANI, S.; BITTENCOURT, I. **Dados Abertos Conectados**. [S.l.]: Editora Novatec, 2015. Citado 10 vezes nas páginas 15, 27, 28, 29, 30, 31, 32, 33, 34 e 36.
- JOHANSTON, H. **Sistemas de Informação Hospitalar: Presente e Futuro**. 1993. <<http://www.informaticamedica.org.br/informed/halley.htm>>. Acessado em 29/05/2019. Citado 2 vezes nas páginas 14 e 15.
- KERAMARIS, V. A.; DANAS, K. "ontodrive" a multi-methodological ontology driven framework for systems analysis of health informatics. In: IEEE. **2016 IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom)**. [S.l.], 2016. p. 1–5. Citado na página 20.
- KILINTZIS, V. et al. A sustainable hl7 fhir based ontology for phr data. In: IEEE. **2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)**. [S.l.], 2019. p. 5700–5703. Citado na página 23.

- KIRYAKOV, A.; OGNANOV, D.; MANOV, D. Owlim—a pragmatic semantic repository for owl. In: SPRINGER. **International Conference on Web Information Systems Engineering**. [S.l.], 2005. p. 182–192. Citado na página 28.
- KITCHENHAM, B.; CHARTERS, S. **Guidelines for Performing Systematic Literature Reviews in Software Engineering**. 2007. <[https://www.elsevier.com/\\_\\_\\_data/promis\\_misc/525444systematicreviewsguide.pdf](https://www.elsevier.com/___data/promis_misc/525444systematicreviewsguide.pdf)>. Acessado em 30/05/2019. Citado na página 18.
- LIU, J. et al. An ontology-based real-time monitoring approach to clinical pathway. In: IEEE. **2014 7th International Conference on Biomedical Engineering and Informatics**. [S.l.], 2014. p. 756–761. Citado na página 22.
- MISTRY, M. et al. Ontologies: Need, usage and attainment of health care system. In: IEEE. **2013 International Conference on Intelligent Systems and Signal Processing (ISSP)**. [S.l.], 2013. p. 381–386. Citado na página 24.
- MIZOGUCHI, R. Part 3: Advanced course of ontological engineering. **New Generation Computing**, Springer, v. 22, n. 2, p. 193–220, 2004. Citado 2 vezes nas páginas 26 e 28.
- MURPHY, A. R.; REDDY, M. C. Identification and management of information problems by emergency department staff. In: AMERICAN MEDICAL INFORMATICS ASSOCIATION. **AMIA Annual Symposium Proceedings**. [S.l.], 2014. v. 2014, p. 1845. Citado na página 14.
- NAZARENKO, G. et al. Development of the ontology of patient management technological records for modeling of clinical workflows in a general hospital. **Scientific and Technical Information Processing**, Springer, v. 42, n. 6, p. 455–462, 2015. Citado na página 21.
- NIE, H. et al. From healthcare messaging standard to semantic web service description: generating wsmo annotation from hl7 with mapping-based approach. In: IEEE. **2013 IEEE International Conference on Services Computing**. [S.l.], 2013. p. 470–477. Citado na página 20.
- NOY, N. F.; MCGUINNESS, D. L. **Ontology development 101: A guide to creating your first ontology**. 2004. <[https://protege.stanford.edu/publications/ontology\\_development/ontology101.pdf](https://protege.stanford.edu/publications/ontology_development/ontology101.pdf)>. Acessado em 11/06/2019. Citado na página 35.
- ONTOTEXT. **What are Linked Data and Linked Open Data?** 2010. <<https://www.ontotext.com/knowledgehub/fundamentals/linked-data-linked-open-data/>>. Acessado em 07/03/2020. Citado na página 33.
- PARTNERSHIP, O. G. **About | Open Government Partnership**. 2019. <<https://www.opengovpartnership.org/about/>>. Acessado em 30/05/2019. Citado na página 33.
- PATRAO, D. F. et al. Ontocloud—a clinical information ontology based data integration system. In: CITESEER. **ONTOBRAS**. [S.l.], 2013. p. 118–129. Citado 3 vezes nas páginas 15, 22 e 25.
- PERAL, J. et al. An ontology-oriented architecture for dealing with heterogeneous data applied to telemedicine systems. **IEEE Access**, IEEE, v. 6, p. 41118–41138, 2018. Citado na página 21.

- PEREIRA, D. L. N. C.; WASSERMANN, R.; SALVADOR, L. Integraç ao semântica das bases de dados do município de sao paulo: Um estudo de caso com anomalias congênitas. 2018. Citado na página 24.
- PORTO, F. U. do. **Sistemas de Informação em Saúde**. 2006. <[http://im.med.up.pt/si\\_saude/si\\_saude.html](http://im.med.up.pt/si_saude/si_saude.html)>. Acessado em 11/06/2019. Citado na página 20.
- RAO, R. R.; MAKKITHAYA, K.; GUPTA, N. Ontology based semantic representation for public health data integration. In: IEEE. **2014 International Conference on Contemporary Computing and Informatics (IC3I)**. [S.l.], 2014. p. 357–362. Citado 2 vezes nas páginas 23 e 25.
- RIAÑO, D. et al. An ontology-based personalization of health-care knowledge to support clinical decisions for chronically ill patients. **Journal of biomedical informatics**, Elsevier, v. 45, n. 3, p. 429–446, 2012. Citado 2 vezes nas páginas 15 e 21.
- ROEHRS, A. et al. Toward a model for personal health record interoperability. **IEEE journal of biomedical and health informatics**, IEEE, v. 23, n. 2, p. 867–873, 2018. Citado na página 23.
- SBC. **Comissão Especial de Sistemas de Informação**. 2018. <<http://www2.sbc.org.br/ce-si/>>. Acessado em 12/05/2019. Citado na página 14.
- SORDI, J. O. **Administração da Informação - Fundamentos e Práticas Para Uma Nova Gestão do Conhecimento**. 2. ed. [S.l.]: Saraiva, 2015. ISBN 8502634801. Citado na página 38.
- SOUZA, A. C. R.; MEDEIROS, A. P.; MARTINS, C. B. An application ontology to support the access to data of medical doctors and health facilities in brazilian municipalities. In: **Ontobras**. [S.l.: s.n.], 2015. Citado 2 vezes nas páginas 24 e 25.
- W3C. **SPARQL Query Language for RDF**. 2008. <<https://www.w3.org/TR/rdf-sparql-query/>>. Acessado em 26/05/2019. Citado na página 29.
- W3C. **RDF 1.1 Primer**. 2014. <<https://www.w3.org/TR/rdf11-primer/>>. Acessado em 20/05/2019. Citado na página 29.
- ZHAO, C.; ZHANG, L. Research of information presentation for electronic medical record based on ontology. In: IEEE. **2013 6th International Conference on Information Management, Innovation Management and Industrial Engineering**. [S.l.], 2013. v. 3, p. 489–492. Citado 2 vezes nas páginas 22 e 25.

## Apêndices

## APÊNDICE A – *Script* para Filtrar Registros nos Dados Abertos de Curitiba

### A.1 Código-fonte em Python

```
1 import pandas as pd
2 import numpy as np
3 import matplotlib
4 import matplotlib.pyplot as plt
5 import sys
6
7 def carrega_base_dados(caminho='/home/rodrigo/Documents/pmc-
  dados-abertos-saúde/'):
8     sys.stdout.write('Carregando os dados...')
9     sys.stdout.flush()
10
11     # Abre o arquivo CSV no caminho especificado, especí
12     # ficado 'latin-1' para caracteres especiais, separação
13     # por ';' ao invés de ','
14     enfermagem = pd.read_csv(caminho + '2020-04-06_Sistema_E-
15     Saude_Enfermagem_-_Base_de_Dados.csv', low_memory=
16     False, encoding='latin-1', sep=';')
17     medicos = pd.read_csv(caminho + '2020-04-06_Sistema_E-
18     Saude_Medicos_-_Base_de_Dados.csv', low_memory=False,
19     encoding='latin-1', sep=';')
20
21     base_dados = [enfermagem, medicos]
22
23     sys.stdout.write('finalizado!\n')
24     sys.stdout.flush()
25
26     return base_dados
27
28 def aplica_filtros(base_dados=None, codigo_cbo='225250'):
29     if base_dados is None:
30         return None
31
32     sys.stdout.write('Aplicando os filtros...')
```



```
27     sys.stdout.flush()
28
29     # Separa as bases
30     enfermagem = base_dados[0]
31     medicos = base_dados[1]
32
33     # Define quais colunas serão utilizadas
34     colunas_enfermagem = ['Data do Atendimento', 'Data de
35     Nascimento', 'Sexo', 'Tipo de Unidade',
36     'Código da Unidade', 'Descrição da Unidade', 'Código do
37     Procedimento', 'Descrição do Procedimento',
38     'Código do CID', 'Descrição do CID', 'cod_usuario', '
39     cod_profissional']
40     colunas_medicos = ['Data do Atendimento', 'Sexo', 'Código
41     do Procedimento', 'Descrição do Procedimento',
42     'Código do CBO', 'Código do CID', 'Descrição do CID', '
43     cod_usuario', 'cod_profissional']
44
45     # Aplica o filtro para as clunas
46     enfermagem = enfermagem[colunas_enfermagem]
47     medicos = medicos[colunas_medicos]
48
49     # Aplica o filtro de tipo profissional (225250 ->
50     Ginecologista) e sexo, busca os usuários que estão na
51     base de enfermagem
52     medicos = medicos[medicos['Sexo'].map(lambda x: x == 'F')
53     ]
54     medicos = medicos[medicos['Código do CBO'].map(lambda x:
55     x == codigo_cbo)]
56     enfermagem = enfermagem[enfermagem['cod_usuario'].isin(
57     medicos['cod_usuario'].tolist())]
58
59     sys.stdout.write('finalizado!\n')
60     sys.stdout.flush()
61
62     return [enfermagem, medicos]
63
64 def salva_base_dados(caminho='/home/rodrigo/Documents/pmc-
65     dados-abertos-saúde/', base_dados=None):
```

```
55     if base_dados is None:
56         return
57
58     sys.stdout.write('Salvando base...')
59     sys.stdout.flush()
60
61     # Salva a base no caminho especificado, 'latin-1' para
        preservar os caracteres especiais
62     base_dados[0].to_csv(caminho + 'EnfermagemFiltrado.csv')
63     base_dados[1].to_csv(caminho + 'MedicosFiltrado.csv')
64
65     sys.stdout.write('finalizado!\n')
66     sys.stdout.flush()
67
68 if __name__ == "__main__":
69     base_dados = carrega_base_dados()
70     base_ginecologia = aplica_filtros(base_dados)
71     salva_base_dados(base_dados=base_ginecologia)
72     sys.exit(0)
```

Anexos

## ANEXO A – Formulário de Acolhimento com Classificação de Risco



**UNIDADE MATERNO INFANTIL-UMI- PRONTO ATENDIMENTO EM GINECOLOGIA E OBSTETRÍCIA  
FICHA DE ACOLHIMENTO E CLASSIFICAÇÃO DE RISCO**

**Atenção, dados obrigatórios para preenchimento\***

\* Horário de chegada na recepção: \_\_\_\_\_

\* Horário do acolhimento e classificação de risco: \_\_\_\_\_

 GINECOLOGIA

\* Horário do recolhimento se necessário: \_\_\_\_\_

\* Horário do atendimento médico: \_\_\_\_\_

 OBSTETRÍCIA**\*DADOS DO ACOLHIMENTO**

Nome: \_\_\_\_\_ Idade: \_\_\_\_\_

Registro: \_\_\_\_\_ Data: \_\_\_\_/\_\_\_\_/2018.

**\*DADOS OBSTÉTRICOS**

Histórico obstétrico: Gestação: \_\_\_\_\_ Parto Normal \_\_\_\_\_ Cesariana : \_\_\_\_\_ Aborto(s): \_\_\_\_\_

Idade gestacional: \_\_\_\_\_ D.U.M: \_\_\_\_\_ D.P.P: \_\_\_\_\_

**\*NÚMERO E RELATO DA QUEIXA PRINCIPAL**


Comorbidade \_\_\_\_\_

Medicamento de uso: \_\_\_\_\_

Alergias: \_\_\_\_\_

**\*Atenção, descrever queixa, início e outras comorbidades.**

1. Desmaio/ Mal estar 2. Dor abdominal/Lombar/Contrações uterinas 3. Dor de cabeça/ Tontura. 4. Falta de ar/Sintomas respiratórias 5. Febre/ Sinais de infecção 6. Náuseas e Vômitos  
7. Perda de Líquido vaginal/ Secreções 8. Perda de sangue via vaginal 9. Queixas urinárias 10. Parada/ Redução dos movimentos fetais 11. Relato de convulsão 12. Outras queixas.

**\*SINAIS VITAIS**

PA: \_\_\_\_\_ mmHg Pulso: \_\_\_\_\_ FR: \_\_\_\_\_ Temperatura: \_\_\_\_\_ °C Saturação: \_\_\_\_\_ (%)

											Em uma escala de 0 a 10, como você classifica a sua dor, considerando como 0 nenhuma?
<b>Azul</b>	<b>Verde</b>			<b>Amarelo</b>			<b>Laranja</b>				
<b>0</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>	
<b>Sem Dor</b>	<b>Dor Leve</b>			<b>Dor Moderada</b>			<b>Dor Intensa</b>				_____

**\*OUTRAS ALTERAÇÕES**

Movimento fetal: ( ) Sim ( ) Não \_\_\_\_\_

Sangramento vaginal: ( ) Sim ( ) Não \_\_\_\_\_

Contrações uterinas presentes: ( ) Sim ( ) Não \_\_\_\_\_

Perda de Líquidos: ( ) Sim ( ) Não \_\_\_\_\_

**\*CLASSIFICAÇÃO E CONDUTA**
 Vermelho     Laranja     Amarelo     Verde     Azul

Fluxo adotado \_\_\_\_\_ Descritor Chave: \_\_\_\_\_

Enfermeiro/COREN \_\_\_\_\_ Médico/CRM \_\_\_\_\_

O Conselho Federal de Medicina, por meio da Resolução CFM n° 2.077 de 16 de setembro de 2014, dispõe:  
Art. 3° Todo paciente que tiver acesso ao Serviço Hospitalar de Urgência e Emergência, deverá, obrigatoriamente, ser atendido por um médico, não podendo, sob nenhuma justificativa, ser dispensado ou encaminhado a outra unidade de saúde por outro profissional que não o médico.



<b>Campo exclusivo do paciente</b>	<b>TERMO DE CONSENTIMENTO</b>
	Eu, _____ estou ciente, que de acordo com a classificação de risco realizada pelo enfermeiro, através das queixas relatadas, mensuração de sinais vitais e observação clínica, ficam caracterizadas como atendimento não prioritário ou a possibilidade de encaminhar para outro serviço de referência depois da consulta médica. Diante disso, foi esclarecido que fica a minha escolha aguardar até 240 minutos, sujeita a reavaliação.

<b>Campo exclusivo do médico</b>	<b>TERMO DE CONDUTA</b>
	Eu, _____, médico responsável pelo atendimento dessa usuária, nas especialidades de ginecologia e obstetrícia, tenho com suspeita diagnóstica de _____, através do relato descrito e por isso, realizo o encaminhamento para _____ onde poderá ser assistida conforme sua necessidade. Abaixo, minhas considerações;  _____ _____ _____ _____ _____ _____ _____ _____  _____  Médico/ Registro profissional