

UNIVERSIDADE DE SÃO PAULO
ESCOLA DE ENGENHARIA DE SÃO CARLOS
FACULDADE DE MEDICINA DE RIBEIRÃO PRETO
INSTITUTO DE QUÍMICA DE SÃO CARLOS

FELIPE CARVALHO PELLISON

**Desenvolvimento e avaliação de um sistema interoperável para a gestão da
informação de pacientes com tuberculose**

São Carlos

2021

FELIPE CARVALHO PELLISON

**Desenvolvimento e avaliação de um sistema interoperável para a gestão da
informação de pacientes com tuberculose**

VERSÃO CORRIGIDA

Tese apresentada ao Programa de Pós-Graduação Interunidades em Bioengenharia da Escola de Engenharia de São Carlos – Faculdade de Medicina de Ribeirão Preto e Instituto de Química de São Carlos da Universidade de São Paulo, como requisito para a obtenção do Título de Doutor em Ciências.

Área de concentração: Bioengenharia

Orientador: Prof. Dr. Domingos Alves

Coorientador: Prof. Dr. Rui Pedro Charters
Lopes Rijo

São Carlos

2021

AUTORIZO A REPRODUÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica elaborada pela Biblioteca Prof. Dr. Sérgio Rodrigues Fontes da EESC/USP com os dados inseridos pelo autor.

P391a Pellison, Felipe Carvalho
Desenvolvimento e avaliação de um sistema interoperável para a gestão da informação de pacientes com tuberculose / Felipe Carvalho Pellison; orientador Domingos Alves; coorientador Rui Pedro Charters Lopes Rijo. São Carlos, 2021.

Tese (Doutorado) - Programa de Pós-Graduação Interunidades em Bioengenharia e Área de Concentração em Bioengenharia -- Escola de Engenharia de São Carlos; Faculdade de Medicina de Ribeirão Preto; Instituto de Química de São Carlos, da Universidade de São Paulo, 2021.

1. Interoperabilidade. 2. Sistemas de Informação em Saúde. 3. Web Semântica. I. Título.

Eduardo Graziosi Silva - CRB - 8/8907



FOLHA DE JULGAMENTO

Candidato(a): Felipe Carvalho Pellison

TÍTULO: "Desenvolvimento e avaliação de um sistema interoperável para a gestão da informação de pacientes com tuberculose"

Data da defesa: 14/04/2021

Comissão Julgadora	Assinatura	Resultado
Prof(a). Dr(a). Domingos Alves FMRP/USP -- Orientador		<u>Não votante</u>
Prof(a). Dr(a). Afrânio Lineu Kritski UFRJ		<u>Aprovado</u>
Prof(a). Dr(a). Valdes Roberto Bollela FMRP/USP		<u>Aprovado</u>
Prof(a). Dr(a). Ricardo João Cruz Correia Universidade do Porto		<u>Aprovado</u>
Prof(a). Dr(a). Antonio Ruffino Neto FMRP/USP		<u>Aprovado</u>
Prof(a). Dr(a). Mauro Niskier Sanchez UnB		<u>Aprovado</u>

DEDICATÓRIA

Aos meus pais.

AGRADECIMENTOS

À Diretoria da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo e à chefia da Seção Técnica de Informática da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo por permitirem a capacitação e formação de seus funcionários.

Ao meu orientador, Prof. Dr. Domingos Alves, e ao meu coorientador, Prof. Dr. Rui Pedro Charters Lopes Rijo, pela confiança e apoio.

A minha família, especialmente meus pais, pelo amor, suporte e cuidado.

À Márcia pelo amor, carinho e compreensão em todos os momentos.

Ao Vinícius, Filipe, Sidney, José Francisco, Ricardo, Leandro e à Vânia pela amizade. Especialmente à Nathalia pelo companheirismo durante o período de intercâmbio em Portugal.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) pelo financiamento do período de doutorado sanduíche no Instituto Politécnico de Leiria – Portugal (processo nº 88887.141211/2017-00).

RESUMO

PELLISON, F. C. **Desenvolvimento e avaliação de um sistema interoperável para a gestão da informação de pacientes com tuberculose**. 2021. 125 f. Tese – Programa de Pós-Graduação Interunidades em Bioengenharia da Escola de Engenharia de São Carlos, Faculdade de Medicina de Ribeirão Preto e Instituto de Química de São Carlos da Universidade de São Paulo, São Carlos, 2021.

O Sistema de Vigilância Epidemiológica de tuberculose (TB) necessita, não apenas da implementação de ações básicas de prevenção e assistência, mas também do fortalecimento da integração entre os diferentes serviços de saúde, programas e níveis de atenção existentes, cuja resolubilidade varia de acordo com recursos financeiros, técnicos, humanos e de infraestrutura dos serviços que compõem a rede de atenção local. Diante deste cenário, torna-se imprescindível o desenvolvimento de sistemas de informação e técnicas de interoperabilidade de sistemas que venham contribuir com as atividades de vigilância epidemiológica e com o acompanhamento dos pacientes com TB. Tais sistemas e técnicas devem permitir além do básico de se cadastrar, acompanhar e avaliar o paciente e seus respectivos contatos, armazenando desde as informações mais básicas até dados sobre o tratamento, exames e internações, como também permitir a integração dos dados com outros sistemas, isto é, informatizar um serviço prestado a comunidade, facilitando a coleta de dados e dificultando sua perda, e, dessa forma, possibilitar uma melhora na análise, visualização e manipulação dos dados. Para que isto seja possível este projeto pretende aprofundar e ampliar o conhecimento na área de informática em saúde pública a partir de técnicas de interoperabilidade entre sistemas baseados em web semântica. A interoperabilidade semântica é garantida por meio da criação de uma ontologia voltada ao mapeamento dos termos referentes à TB, em especial ao tratamento diretamente observado. Objetiva-se melhorar a recuperação e integração dos dados e tanto na gestão organizacional quanto na clínica dos casos de TB em Ribeirão Preto, por meio da interoperabilidade entre o e- SUS AB, SITETB, GAL, SISTB, os sistemas de informação de hospitais e o TBWEB.

Palavras-chave: Interoperabilidade. Sistemas de Informação em Saúde. Web Semântica

ABSTRACT

PELLISON, F. C. **Development and evaluation of an interoperable system for information management of tuberculosis patients** 2021. 125 f. Tese – Programa de Pós-Graduação Interunidades em Bioengenharia da Escola de Engenharia de São Carlos, Faculdade de Medicina de Ribeirão Preto e Instituto de Química de São Carlos da Universidade de São Paulo, São Carlos, 2021.

The Epidemiological Surveillance System requires, not only the implementation of basic prevention and care actions, but also the strengthening of the integration between the different health services, programs and levels of care, whose resolvability varies according to financial, technical, human resources and infrastructure of the services that make up the network. Given this scenario, it is imperative to develop information systems and systems interoperability techniques that contribute to epidemiological surveillance activities and patient follow-up. Such systems and techniques should allow the basic registration, monitoring and evaluation of the patient and his/her contacts, storing from the most basic information to data on treatment, examinations and hospitalizations, as well as allowing the integration of data with other systems, that is, computerize a service provided to the community, facilitating the collection of data and making it difficult to lose and thus enabling an improvement in the analysis, visualization and manipulation of the data. For this to be possible this project intends to deepen and broaden the knowledge in the area of public health informatics and interoperability techniques between systems based on semantic web. The semantic interoperability will be guaranteed by the creation of an ontology that maps the terms referred to tuberculosis, specially the directly observed treatment. Always trying to improve the quality of the data recovered and the integration between applications that hold both organizational and clinical data of tuberculosis in Ribeirão Preto from e-SUS AB, SITETB, GAL, SISTB, hospital information systems and TBWEB.

Keywords: Interoperability. Health Information Systems. Semantic Web

LISTA DE FIGURAS

Figura 1 - Ontologia para suportar o TDO no cenário brasileiro. -----	81
Figura 2 - Propriedades dos objetos relacionados ao TDO.-----	82
Figura 3 - Primeira parte das entidades relacionadas à TB mapeadas na BFO.-----	83
Figura 4 -Segunda parte das entidades relacionadas à TB mapeadas na BFO. -----	84
Figura 5 - Termos traduzidos a fim de permitir a troca de dados entre aplicações nacionais e internacionais -----	85
Figura 6 - Interface de seleção de mapeamento de versão da ontologia.-----	86
Figura 7 - Exemplo de resultado do mapeamento. -----	87
Figura 8 - Fluxo de informação para tratamento de TB após implementação.-----	88
Figura 9 - Captura de tela e parte do código semanticamente marcado do ambiente de testes do GAL. -----	90
Figura 10 - Captura de tela e parte do código semanticamente marcado do ambiente de testes do TBWeb.-----	90
Figura 11 - Captura de tela e parte do código semanticamente marcado do ambiente de testes do SINAN. -----	91
Figura 12 - Captura de tela e parte do código semanticamente marcado do ambiente de testes do SisTB.-----	91
Figura 13 - Resultados de uma query SPARQL simultânea no SisTB e TBWeb para o paciente com id = 10-----	92
Figura 14 - Exemplo de distribuição de casos de TB no Estado de São Paulo para o período de 2006 a 2014 -----	94
Figura 15 - Arquitetura de interoperabilidade proposta entre aplicações que utilizam WS e outros padrões -----	97

LISTA DE ABREVIATURAS E SIGLAS

AB	Atenção Básica
ALT	Alanina Aminotransferase
API	<i>Application Programming Interface</i>
BFO	<i>Basic Formal Ontology</i>
CID-10	Décima Versão da Classificação Internacional das Doenças
CNCT	Campanha Nacional Contra a Tuberculose
CPG	<i>Clinical Practice Guidelines</i>
DAML	<i>DARPA Agent Markup Language</i>
DL	Description Logics
DL	Description Logic
DSM-IV	<i>Diagnostic and Statistical Manual of Mental Disorder</i>
ECG	Eletrocardiograma
EHR	Registros eletrônicos de saúde
EORCA	<i>Even Oriented Representation for Collaborative Activities</i>
FNI	Ficha de Notificação Individual
GAL	Sistema de Gerência Laboratorial
GHO	<i>Global Health Observatory</i>
HAART	<i>Highly Active Antirretroviral Therapy</i>
HCLS IG	<i>Semantic Web Health Care and Life Sciences Interest Group</i>
IDH	Índice de Desenvolvimento Humano
JSTL	<i>Java Server Pages Standard Tag Library</i>
KDOM	Mapeador ontológico de dados de conhecimento
LD	<i>Linked Data</i>
MDG	Desenvolvimento de Metas relacionadas à Doença para o Milênio
MDR	Multi drogaresistente
MS	Ministério da Saúde
OBO	<i>Open Biomedical Ontologies</i>
OIL	<i>Ontology Inference Layer</i>
OMS	Organização Mundial de Saúde
OWL	<i>Web Ontology Language</i>
PACS	Programa de Agentes Comunitários em Saúde
PATO	<i>Phenotypic Quality Ontology</i>

PCR	Reação em Cadeia de Polimerase
PNIS	Política Nacional de Informação e Informática em Saúde
PSF	Programa de Saude da Família
RDF	Resource Description Framework
RDFS	Resource Description Framework Schema
RNTCP	Programa Nacional Revisado de Controle da Tuberculose
RNTCP	Programa Nacional de Controle e Gestão da Tuberculose da Índia
RR	Resistência à Rifampicina
SINAN	Sistema de Informação de Agravos de Notificação
SISREG	Sistema Nacional de Regulação
SNOMED	Nomenclatura Sistematizada de Medicina
SPARQL	<i>SPARQL Query Language</i>
SUS	Sistema Único de Saúde
TARV	Tratamento anti-retroviral
TB	Tuberculose
TB-DR	TB drogarresistente
TBEMR	TB extensivamente multirresistente
TB-MDR	TB multidroga resistente
TBWEB	Sistema de Notificação e Acompanhamento dos Casos de TB em SP
TDO	Tratamento Diretamente Observado
UMLS	<i>Unified Medical Language System</i>
URI	<i>Uniform Resource Identifier</i>
URL	<i>Uniform Resource Locator</i>
URN	<i>Uniform Resource Name</i>
WS	Web Semântica
WSA	<i>Web Service Architecture</i>
WSDL	<i>Web Service Description Language</i>
WSML	<i>Web Service Modeling Language</i>
WSMO	<i>Web Services Modeling Ontology</i>
WSS	Web services
XDR	Resistência Extensiva
XML	<i>Extensible Markup Language</i>

SUMÁRIO

1. INTRODUÇÃO	23
2. OBJETIVOS	27
3. REVISÃO BIBLIOGRÁFICA	28
3.1. TUBERCULOSE.....	28
3.1.1. Agente infeccioso e patogenia.....	28
3.1.2. Diagnóstico e tratamento.....	30
3.1.3. Epidemiologia.....	32
3.1.4. Medidas de controle da doença	34
3.1.5. Tratamento diretamente observado (TDO).....	38
3.1.6. Acompanhamento do tratamento	41
3.1.7. TB droga Resistente	41
3.1.8. Fatores de risco associados com abandono do tratamento	43
3.1.9. Retratamento.....	46
3.2. WEB SEMÂNTICA.....	46
3.2.1. Histórico e evolução	46
3.2.2. Unicode	50
3.2.3. Uniform Resource Identifier (URI)	50
3.2.4. Extensible Markup Language (XML)	51
3.2.5. RDF.....	53
3.2.6. Linked data.....	55
3.2.7. Armazenamento de RDF.....	56
3.2.8. SPARQL Protocol and RDF Query Language.....	57
3.2.9. Linguagem D2RQ.....	59
3.2.10. Aplicações da WS na área de saúde e medicina	60
3.3. ONTOLOGIAS	65
3.3.1. Componentes de uma Ontologia.....	67
3.3.2. Basic Formal Ontology (BFO).....	68
3.3.3. OWL	69
3.3.4. Mapeamentos de Ontologias	72
4. MÉTODOS	79
4.1. Cenário de Aplicação: Sistemas de Informação com dados de TB	79
4.2. Utilizando a BFO para construir a Ontologia da TB.....	80
5. RESULTADOS E DISCUSSÃO	83
5.1. Ontologia para apoiar a interoperabilidade no caso da TB	83
5.2. Mapeamento de ontologias via web	85

SUMÁRIO

5.3. A arquitetura de interoperabilidade para TB.....	87
5.4. Integrando dados de TB no estado de São Paulo: uma prova de conceito	93
5.5. Resultados Adicionais.....	95
6. CONCLUSÃO E TRABALHOS FUTUROS.....	98
7. REFERÊNCIAS.....	99
8. APÊNDICE A – Publicações.....	116

1. INTRODUÇÃO

No âmbito da saúde, de uma maneira geral, as informações e os dados coletados e sistematizados, resultantes das pesquisas experimentais, das pesquisas clínicas, das pesquisas epidemiológicas/operacionais e da prática profissional em saúde são a base para a construção do conhecimento nessa área que, por sua vez, deve subsidiar o estabelecimento de práticas e políticas adequadas, voltadas para a população e para o sistema de saúde vigente. Dessa forma, a capacidade de "executar" a gestão do conhecimento em saúde também depende de uma inteligência organizacional de saúde que, em primeira instância, deve promover a sinergia entre as atividades de ciência e tecnologia da informação reunindo mecanismos e instrumentos, bem como métodos e formas de organização, que possam incrementar a capacidade de promover as atividades em redes (BELLUZO, 2007).

O desafio a ser enfrentado é o da integração e articulação das informações em saúde, que facilmente podem subsidiar o conhecimento de ações e serviços nessa área, e de seu impacto na situação de saúde da população e nos indicadores epidemiológicos propostos pela Organização Mundial da Saúde (OMS). Para isso, o armazenamento computacional dos dados relacionados com um determinado evento (individual ou coletivo) em saúde é uma estratégia fundamental e deve objetivar a melhoria da qualidade e eficiência dos processos de trabalho, automaticamente alimentando a cadeia de informações e eliminando todos os demais instrumentos paralelos de coleta relativos a este evento.

É importante frisar que essa estratégia de integração tem de estar imbuída das necessidades e dificuldades encontradas no dia-a-dia da assistência em saúde na região, sem nunca esquecer que os métodos e produtos propostos estejam alinhados com as possibilidades das instituições parceiras. A partilha de informação e experiências entre grupos de investigação e/ou entre indivíduos singulares, permite a criação de equipes multidisciplinares para a resolução de problemas e torna efetivamente possível a colocação dos utilizadores dos Sistemas de Informação em Saúde - nomeadamente doentes, profissionais de saúde, profissionais de tecnologia da informação, pesquisadores e estudantes - no centro das soluções tecnológicas e organizacionais necessárias à melhoria dos serviços e produtos.

Diante desse cenário temos a tuberculose (TB) que representa um sério problema de saúde mundial e permanece como uma doença negligenciada em recente crescimento no Brasil (WHO, 2020). Estima-se que 95% dos casos de TB do mundo ocorram nos países em desenvolvimento (MALHÃO et al., 2010). No ano de 2019, o Brasil apresentou uma alta incidência de casos, ficando no grupo dos oito países responsáveis pela maior carga da doença

no mundo e, embora a TB seja tratável e curável com medicamentos de baixo custo e alta eficácia (WHO, 2020).

Em 2014, a OMS lança sua estratégia de extinção da TB (em inglês, *End TB Strategy*) visando reduzir drasticamente a incidência e as mortes em países com baixa renda. A estratégia é baseada em três pilares: prevenção e cuidados integrados centrados no paciente (pilar 1); sistemas que dêem suporte à políticas ousadas (pilar 2) ; estímulo à pesquisa e inovação (pilar 3). Dentro dos campos abordados no pilar 3, um componente chave para que se desenvolvam novas estratégias e políticas integradas é fomentar o desenvolvimento da *Digital Health* com o objetivo de unificar e extrair conhecimento de diversas fontes de dados heterogêneas em saúde.(WHO, 2014)

Assim, embora haja uma preocupação na elaboração de normas e documentos oficiais que possam servir como norteadores para que os profissionais de saúde se orientem na execução das ações, uma série de dificuldades é percebida em relação ao controle da TB, como o déficit na capacitação das equipes, a alta rotatividade de profissionais de saúde, a falta de conhecimento dos profissionais em relação à forma de transmissão da TB e métodos/procedimentos para a detecção de casos, a sobrecarga de trabalho, além da visão fragmentada da organização das ações nos serviços de saúde para o manejo clínico e seguimento adequado do caso em tratamento (SILVA, 2021).

O acompanhamento dos casos TB envolve o preenchimento de diversos instrumentos de registro padronizados pelo Ministério da Saúde (MS), tais como a Ficha de Notificação Individual (FNI), a Ficha de Registro do Tratamento Diretamente Observado e o Livro de Registro e Acompanhamento de Tratamento dos Casos de TB (BRASIL, 2006), além do preenchimento de informações no Sistema de Gerência Laboratorial (GAL), do prontuário clínico e do Sistema de Notificação e Acompanhamento dos Casos de TB do Estado de São Paulo (TBWEB) ou Sistema de Informação de Agravos de Notificação (SINAN) nos outros Estados. Nesse sentido, as diversas formas de registro podem gerar informações contraditórias e/ou a perda dos dados, não garantindo necessariamente a integralidade da assistência.

Deste modo, torna-se difícil gerenciar e organizar as informações dos pacientes e obter relatórios históricos para avaliar melhores estratégias durante o tratamento. Além disso, a deficiência na integração das informações entre os diferentes níveis de atenção por meio de um sistema compartilhado entre ambos dificulta o acompanhamento da enfermidade, e conseqüentemente na continuidade da assistência no que se refere às condições crônicas (ALMEIDA et al., 2010; Gansel et al 2019)

Além disso, tem sido observado em diferentes estudos a existência de uma estreita

relação entre os fatores ligados ao insucesso do tratamento da TB e o modo como as ações de controle são desempenhadas no âmbito da Atenção Básica (AB). Verifica-se que, apesar do acompanhamento do doente de TB nesses locais, viabilizar a identificação dos fatores que fragilizam a adesão terapêutica, questões relacionadas ao fluxo de atendimento, à referência e contra-referência na prestação da assistência, bem como as debilidades de integração dos níveis assistenciais e laboratoriais e também de ações e serviços prestados mostram-se relevantes para o controle da TB (SILVA, 2021).

Nesse sentido, frente ao impacto proporcionado pela morbimortalidade relacionado à TB, parte-se da premissa que a integração de ações e serviços efetiva é fundamental, bem como a identificação dos pontos de estrangulamento da atenção no controle da doença.

A atenção integral às pessoas acometidas pela TB, necessita, não apenas da implementação de ações básicas de prevenção e assistência, mas também do fortalecimento da integração entre os diferentes serviços de saúde, programas e níveis de atenção existentes no município, cuja resolubilidade varia de acordo com recursos financeiros, técnicos, humanos e de infraestrutura dos serviços que compõem a rede de atenção (BRASIL, 2006).

Conceitualmente, o sistema de informação em saúde pode ser entendido como um instrumento para adquirir, organizar e analisar dados necessários à definição de problemas e riscos para a saúde, avaliar a eficácia, eficiência e influência que os serviços prestados possam ter no estado de saúde da população, além de contribuir para a produção de conhecimento acerca da saúde e dos assuntos a ela ligados (BRANCO, 1996).

Visando ganhos de eficiência e qualidade permitidos pela tecnologia, o MS criou a Política Nacional de Informação e Informática em Saúde (PNIIS), que tem como propósito:

Promover o uso inovador, criativo e transformador da tecnologia da informação, para melhorar os processos de trabalho em saúde, resultando em um Sistema Nacional de Informação em Saúde articulado, que produza informações para os cidadãos, a gestão, a prática profissional, a geração de conhecimento e o controle social, garantindo ganhos de eficiência e qualidade mensuráveis através da ampliação de acesso, equidade, integralidade e humanização dos serviços e, assim, contribuindo para a melhoria da situação de saúde da população (BRASIL, 2004, p. 12).

Tendo em vista o propósito e as diretrizes do PNIIS e o controle da TB, é fundamental a criação de um instrumento que vise o gerenciamento dos casos, a fim de facilitar as atividades dos profissionais de saúde e gestores, além de garantir a qualidade dos dados, evitando informações incompletas e incorretas, de modo a ajudar no planejamento e tomada de decisões que possam melhorar o amparo ao paciente, bem como prevenir a doença.

Na DRS XIII, a atenção à TB lida com, pelo menos, sete principais sistemas de informação em saúde, o SISTB, o Hygia Web, o TBWeb, o SINAN, o GAL, o e-SUS AB e o SITETB. Em algumas das aplicações, os profissionais de saúde precisam reintroduzir as

mesmas informações e, quando são necessárias informações globais sobre um paciente, é preciso fazer tal trabalho manualmente sistema a sistema. A principal motivação para o presente trabalho, em vista do cenário atual, propõe-se o uso da web semântica (WS) para apoiar a interoperabilidade do Sistema Único de Saúde (SUS) brasileiro de apoio à eliminação da TB.

Na *Web* de hoje, a maior parte do conteúdo disponível é adequada para interpretação humana. No entanto, este conteúdo não é facilmente acessível por computadores e sistemas. A WS, definida por (BERNERS-LEE; HENDLER; LASSILA, 2001), pode ser especificada como uma extensão da web atual, com o objetivo de agregar lógica ao conteúdo para expressar o significado de uma informação, suas propriedades e as complexas relações existentes entre diferentes tipos de dados, para que seja possível interpretar o significado de um determinado dado sem se preocupar com sua forma de representação (ROBU; ROBU; THIRION, 2006). A ideia geral é criar uma maneira eficiente de representar dados na World Wide Web para construir um banco de dados conectado à nível global (LAUFER, 2015), através da marcação semântica de páginas *web* e a adaptação de bancos de dados relacionais existentes, baseados em ontologias.

2. OBJETIVOS

O objetivo principal desta tese foi desenvolver e implementar um ensaio de uma camada de interoperabilidade baseada em ferramentas previstas na WS para, em conjunto, servirem de base para um sistema informatizado no acompanhamento dos pacientes de TB na cidade de Ribeirão Preto, no âmbito da Vigilância Epidemiológica. Tais ensaios proveram possíveis meios de agregar e disponibilizar as informações de maneira acessível aos gestores e profissionais de saúde. Foram testados um conjunto de aplicativos e técnicas que poderão servir como base de conhecimento e para a ampliação de pesquisas que dêem suporte ao acompanhamento de pacientes com TB ao integrar dados de cadastro, acompanhamento e avaliação do paciente e seus respectivos contatos, especialmente dos dados referentes ao Tratamento Diretamente Observado, com outras bases de dados disponíveis como tratamento, exames laboratoriais e internações.

3. REVISÃO BIBLIOGRÁFICA

Nesta seção vamos rever os principais conceitos e trabalhos abordados relacionados à esta tese.

3.1. TUBERCULOSE

3.1.1. Agente infeccioso e patogenicidade

Mycobacterium tuberculosis é um bastonete aeróbio obrigatório, de parasitismo intracelular, que se desenvolve em tecidos com elevada concentração de oxigênio, e também de outros gases como hidrogênio, carbono, nitrogênio. (RICH AR, 1944; TARANTINO AB et al, 2008).

Pertencente à ordem dos Actinimycetales, família das *Mycobacteriaceae* e gênero *Mycobacterium*, caracterizando-se por germes móveis, álcool-ácido-resistentes, esporulados e aeróbios. Após serem corados, por meio de microscopia, apresentam-se isolados ou em pares, como bastonetes finos, levemente encurvados ou retos. O substrato mais comum para reprodução é o meio de cultura sólido Löwenstein-Jensen, a uma temperatura de 37 °C, em torno de 14 a 15h. O crescimento lento, em torno de 20-45h, apresenta colônias sob a forma eugônica, verrucosas e secas. É possível utilizar o teste da nitrato-redutase positivo pois é uma micobactéria que sintetiza niacina. Existe uma gama de micobactérias com potencial patogênico, a exemplo: *Mycobacterium africanum*, *Mycobacterium microti*, *Mycobacterium bovis*. Atenção especial seja dada ao *Mycobacterium bovis*, que entre os anos de 1896 e 1898 foi descoberto. É reconhecido pelas colônias lisas, disgnônicas, com crescimento demorado e altamente patogênico para diversas espécies de mamíferos, como ovinos, suínos, bovinos e, inclusive, humanos – onde sua patogenicidade é igual a *M. tuberculosis*. Várias crianças no final do século XX foram contaminadas ao ingerir leite destes animais infectados (HOPEWELL FC, 2005).

M. tuberculosis pode parasitar com uma atividade metabólica baixa, perdurando viável por longos períodos. É uma bactéria intracelular facultativa e infecta células do sistema mononuclear fagocitário. A alta concentração de ácido micólico na sua parede torna-o um bacilo álcool-ácido-resistente. Isto neutraliza a coloração do gram, mantendo fucsina e resistindo à lavagem por solução ácido-alcoólica. Nas lâminas realizadas com o método Ziehl-Neelsen, aparece corado em vermelho sob o fundo azul. (TARANTINO AB et al, 2008).

Partículas, contendo bacilos, pulverizadas por indivíduos com a doença no pulmão e absorvidas pela via inalatória são o meio mais comum de contágio da TB. As menores partículas, com aproximadamente 5,0 micrômetros, podem carregar de um a três bacilos. Estas dispersam-se rapidamente na atmosfera sendo facilmente inaladas, podendo progredir pela árvore traqueobrônquica (RILEY RL et al, 1961). No entanto a maioria destas gotículas é retida ali mesmo, sendo posteriormente deglutidas após chegar a orofaringe via tapete muco-ciliar. No entanto, as partículas que chegarem aos alvéolos terão uma chance de iniciar a atividade patogênica, formando o foco parenquimatoso, ou também chamado de cancro de inoculação, com multiplicação dos germes. Dali seguem para os gânglios mediastino-pulmonares, formando o Complexo de Ranke. Pela drenagem linfática, os bacilos chegam ao confluente venoso júbulo-subclávio homolateral (grande veia linfática à direita; ducto torácico à esquerda), ganhando a corrente sanguínea, disseminando-se livremente pelos diversos órgãos (semeadura precoce, ou bacilemia assintomática), enquanto a imunidade específica não se desenvolve, o que ocorre dentro de 3 a 8 semanas, então com a presença de sintomas gerais como febre, astenia e eventualmente eritema nodos na pele, coincidindo com a viragem tuberculínica, de reação negativa para positiva, e contenção da multiplicação bacilar. Pacientes podem evoluir para TB doença nos próximos meses, denominada TB primária (DANNENBERG AM Jr, 1989; MOREIRA JS, 2012).

A evolução dos focos pulmonares, ganglionares e em outros órgãos afetados dependem do potencial de imunidade celular desenvolvido. Se a imunidade for insuficiente, os sinais e sintomas evoluem, caracterizando a TB primária evolutiva, acometendo principalmente crianças em suas mais drásticas complicações, como por exemplo meningoencefalite e TB miliar.

No adulto, a TB pós-primária ocorre a partir da evolução dos focos que resultaram da infecção primária ou em novo contágio (reinfecção exógena). As lesões geralmente localizam-se nos segmentos apical e posterior dos lobos superiores e no segmento superior dos lobos inferiores. Começando com focos broncopneumônicos paucibacilares, com ou sem necrose caseosa, evoluindo para liquefação e eliminação do cáseo por meio da árvore brônquica. Surgem então cavidades ricamente habitadas por bacilos e conseqüente disseminação brônquica, proporcionando o surgimento de novos focos acinares no parênquima pulmonar, em geral evolutivos. Desta maneira, a TB progride com a destruição gradativa do pulmão, apresentando zonas recentes da doença, exsudativas, com ou sem cavidades, entremeadas por zonas de fibrose (CANETTI G, 1954; PICON PD et al, 1993).

3.1.2. Diagnóstico e tratamento

Acometendo em sua maioria crianças, a TB primária apresenta-se de forma insidiosa, com, sudorese noturna, irritação, inapetência e febre persistindo por mais de dez dias sem outra explicação aparente. Por meio do radiograma torácico pode-se notar a presença de adenopatia mediastino-pulmonar unilateral e a reação tuberculínica positiva deve estar também presente. É comum que o exame de escarro não apresente resultado positivo nestas condições, visto que apresenta uma coleta efetiva muito difícil e neste tipo de TB as lesões pulmonares praticamente não estão presentes. (MANUAL DE RECOMENDAÇÕES PARA O CONTROLE DA TUBERCULOSE NO BRASIL, 2011).

Na apresentação da TB pós-primária, os sintomas são de tosse persistente com expectoração purulenta ou não, com ou sem sangue, febre vespertina, sudorese noturna, anorexia e perda de peso. Por meio do radiograma torácico verificam-se lesões como focos de consolidação, ou até mesmo já escavadas, predominando em quadrantes póstero-superiores de um ou ambos os pulmões, frequentemente com focos pericavitários. Comumente a reação tuberculínica é positiva. O exame de cultura em meio de Löwenstein-Jensen do escarro e a pesquisa de BAAR (bacterioscopia – coloração Ziehl-Nielsen) mostram-se positivos em 60 a 80% dos casos. (MANUAL DE RECOMENDAÇÕES PARA O CONTROLE DA TUBERCULOSE NO BRASIL, 2011).

Outro exame comumente utilizado é o Teste Rápido Molecular para TB (TRM-TB), atualmente o teste Xpert MTB RIF, recomendado pela OMS desde 2010. Lançando mão da técnica de reação em cadeia de polimerase (PCR) em tempo real, amplifica-se os ácidos nucleicos a fim de tentar detectar o DNA do *M. tuberculosis* e também triar variantes resistentes à rifampicina. Apesar de ser comumente indicado para a detecção de TB pulmonar e laríngea em adultos e crianças, o TRM-TB também pode ser executado utilizando-se amostras de materiais obtidos de outras formas de TB sistêmica (pleural, pericárdica, ganglionar, óssea, meníngea, peritoneal).

Após a confirmação do diagnóstico e detecção do forma de TB pelos diversos exames, os serviços de saúde devem proporcionar ao paciente, o mais breve possível, o tratamento adequado. O acolhimento do paciente por uma equipe capacitada e humanizada, tentando minimizar os entraves durante o tratamento e ao mesmo tempo respeitando a autonomia e dignidade de cada pessoa são fatores que contribuem para o sucesso (RUFFINO-NETTOA & VILLA TCS, 2006). A equipe de saúde tem diversos papéis educativos durante o tratamento do paciente: informando-o sobre a doença que o acomete, explicando detalhes sobre o tratamento, a importância da regularidade no uso da medicação e as consequências do

abandono ou até mesmo da interrupção do tratamento. É recomendado estabelecer uma relação de cooperativismo e confiança entre o agente de saúde, o paciente e seus familiares onde todos devem estar cientes do seu papel no sucesso do tratamento.

O tratamento da TB no Brasil foi revisado várias vezes com o tempo. Em 1979, foi estabelecido um programa de tratamento composto por três esquemas, segundo (MANUAL DE RECOMENDAÇÕES PARA O CONTROLE DA TUBERCULOSE NO BRASIL, 2011):

Esquema I (2RHZ/4RH): para novos casos;

Esquema I reforçado (2RHZE/4RHE): para retratamento; *Esquema II (2RHZ/7RH):* para forma meningocefálica;

Esquema III (3SZEet/9EEt): para casos que evoluem para falência do tratamento; Onde R = Rifampicina; H = Isoniazida; Z Pirazinamida; E Etambutol;

Posteriormente, em 2009, o Etambutol foi introduzido como quarto fármaco nos dois primeiros meses, também considerada a fase intensiva do tratamento, do Esquema I (2RHZE/4RH). Esta decisão foi baseada levando-se em consideração o aumento da resistência primária à isoniazida de 4,4% para 6,0% relatado no II Inquérito Nacional de Resistência aos Medicamentos anti-TB. A combinação destas quatro drogas em comprimidos com doses fixas combinadas (R150mg, H 75mg, Z 400mg e E 275mg) é preconizada pela OMS. Os esquemas I reforçado e III foram removidos do programa de tratamento. (MANUAL DE RECOMENDAÇÕES PARA O CONTROLE DA TUBERCULOSE NO BRASIL, 2011).

O Esquema básico (I) preconizado para adultos e adolescentes é 2RHZE/4RH e está indicado em: todos os casos novos adultos e adolescentes (> 10 anos), de todas as formas de TB pulmonar e extrapulmonar (exceto a forma meningoencefálica), infectados ou não por HIV; retratamento por recidiva (independentemente do tempo decorrido do primeiro episódio) ou por retorno após abandono com doença ativa em adultos e adolescentes (>10 anos), exceto a forma meningoencefálica.

Após resultado da cultura e teste de sensibilidade, o esquema medicamentoso dos pacientes será ajustado de acordo com a resistência apresentada aos fármacos.

A evolução de casos para a falência do tratamento devem ser cuidadosamente avaliados quanto ao histórico terapêutico, se houve efetiva adesão aos tratamentos anteriores e eventuais testes para comprovar resistência aos medicamentos. Estes casos, em especial, receberão o Esquema Padronizado para Multirresistência ou então Esquemas Especiais individualizados, de acordo com as resistências encontradas pelo teste de sensibilidade.

Para efeito de indicação de esquemas terapêuticos, consideramos (MANUAL DE

RECOMENDAÇÕES PARA O CONTROLE DA TUBERCULOSE NO BRASIL, 2011):

Caso novo de tratamento: pacientes que nunca se submeteram ao tratamento anti-TB ou o fizeram por até 30 dias;

Retratamento ou com tratamento anterior: pessoa já tratada para TB por mais de 30 dias que necessite de novo tratamento por recidiva após cura ou retorno após abandono;

Falência: persistência da positividade do escarro ao final do tratamento; ou aqueles que, no início do tratamento, são fortemente positivos (++ ou +++) e mantém essa situação até o quarto mês; ou aqueles com positividade inicial seguida de negatificação e nova positividade por dois meses consecutivos, a partir do quarto mês de tratamento;

Para alguns grupos de pacientes existem esquemas especiais de tratamento. Para os casos em que os valores da alanina aminotransferase (ALT) aumentam mais de três vezes o limite da normalidade com presença de sintomas ou valores da ALT aumentados mais que cinco vezes na ausência de sintomas; casos de presença de icterícia; casos de pacientes com cirrose hepática, recomenda-se o esquema alternativo para hepatotoxicidade. Existem também os esquemas que são individualizados de acordo com a combinação e classificação da resistência aos fármacos anti-TB.

3.1.3. Epidemiologia

Estima-se que um indivíduo com TB pulmonar ativa pode transmitir o *Mycobacterium tuberculosis* a outras 20 pessoas num período aproximado de 24 meses, enquanto não for efetuado o diagnóstico e iniciado o tratamento. Portanto, uma população de 100.000 pessoas com 50 casos de doentes bacilíferos pode produzir 1000 infectados anualmente, ou seja, 1,0% da população. Os indivíduos imunocompetentes que tenham tido contato com tais doentes, e infectados, apresentam uma chance de 5,0 a 10,0% de desenvolver TB ao longo da vida. A proporção de adoecimento por TB aumenta em pacientes imunossuprimidos, sendo maior nos indivíduos HIV-positivos a chance de desenvolvimento de TB chega a 8,0-10% ao ano.

Em 2019, em todo mundo, a estimativa de novos casos era de 10 milhões. Neste ano, 1,4 milhões de pessoas morreram devido à TB, sendo 208.000 coinfectados com HIV. Segundo a OMS, as regiões do Sudeste Asiático e África, juntas, são responsáveis por 69% dos casos. Índia, Indonésia, China, Filipinas, Paquistão, Nigéria, Bangladesh e África do Sul são responsáveis por dois terços do total de casos à nível global. (WHO, 2020)

No Brasil, a TB é a principal causa de mortalidade entre as pessoas infectadas pelo HIV, mesmo com *Highly Active Antiretroviral Therapy* (HAART) sendo usada por todas as pessoas infectadas pelo vírus desde 1997 (SECRETARIA DE VIGILÂNCIA EM SAÚDE,

2016).

Os indivíduos HIV-negativos e bacilíferos, sem tratamento para a TB, apresentam uma estimativa alta de óbito de 65,0%, a maioria num período de 2 anos, enquanto entre os não bacilíferos este número cai para 10,0 ou 15,0%. Pacientes que realizam tratamento irregular ou apresentam resistência a algum fármaco anti-TB, bem como os pacientes com coinfeção TB-HIV em tratamento apresentam proporções de mortalidade acima de 10%. Na ausência de infecção por HIV e seguindo regularmente o tratamento especificado, a proporção de óbitos em bacilíferos cai drasticamente para menos de 2,0% (WHO, 2008).

Conforme dados da testagem para HIV entre os casos novos de TB, no Brasil, havia 8,4% de pessoas com a co-infecção TB-HIV em 2019. A região Sul apresentou maior percentual de co-infectados (12,4%). As capitais que destacaram-se nos percentuais da co-infecção TB-HIV estão: Campo Grande (20.1%), Porto Alegre-RS (17,5%) e Maceió (15,0%)(SECRETARIA DE VIGILÂNCIA EM SAÚDE, 2020). No Brasil, em 2019, 76,2% dos casos novos de TB foram submetidos à testagem para HIV superando a meta de 70% preconizada pela Agenda Estratégica do MS (SECRETARIA DE VIGILÂNCIA EM SAÚDE, 2020). A testagem para HIV é importante, pois pessoas portadoras do vírus são mais propensas a desenvolver a TB ativa em comparação à população geral. É comum que o diagnóstico da infecção pelo HIV ocorra durante o curso da TB.

A proporção de cura de TB bacilífera no Brasil, em 2015, foi de 74,2%. Alguns estados alcançaram valores acima da média do país (Acre 89,5%; São Paulo 81,1%; Paraná 80,4%). No entanto, em outros estados essa proporção foi baixa: Ceará 65,1%; Rio Grande do Sul 65,9%; Mato Grosso do Sul 64,4%. A proporção de abandono do tratamento em casos de TB pulmonar bacilífera encontra-se acima de 5,0% em diversos estados, especialmente em Rondônia (16,2%), e no Rio Grande do Sul (16,1%) (SECRETARIA DE VIGILÂNCIA EM SAÚDE, 2016).

Conforme os boletins epidemiológicos, no Brasil, houve redução do coeficiente de incidência da TB, passando de 42,8/100 mil hab em 2001 para 33,5/100 mil hab em 2014, o que corresponde a uma redução média de 2,3% ao ano nesse período (SECRETARIA DE VIGILÂNCIA EM SAÚDE, 2016). Também ocorreu diminuição do número de casos diagnosticados, sendo, em 2001, 73.797 casos novos e, em 2014, 67.966. O coeficiente de mortalidade por TB também apresentou redução no período de 2004 a 2013, de 2,8/100 mil para 2,3/100 mil.

No entanto, o foco de preocupação atual do MS passou a ser a TB multidroga resistente (TB-MDR). Foram registrados 2.690 casos durante o período de 2000 a 2007, com

98,0% de formas pulmonares. 90% dos casos ocorreram nas regiões metropolitanas do Nordeste, Sudeste e Sul do Brasil. O grande número de óbitos, seguido pelo grande número de abandono reforça o quão crítico é o acompanhamento clínico desses casos e a grande dimensão dos desafios para o controle dos casos, sejam os que apresentam resistência medicamentosa ou mesmo os que utilizam o esquema básico (SECRETARIA DE VIGILÂNCIA EM SAÚDE, 2016).

3.1.4. Medidas de controle da doença

As primeiras movimentações para combater a TB no Brasil, curiosamente, não vieram do Estado nem de outro órgão público, mas sim diretamente da área da saúde e da sociedade, no final de século XIX (PICON PD et al, 1993). O Estado não tomava parte no controle da TB pois julgava que a doença não era uma ameaça para si nem para a economia (GONÇALVES H, 2002).

A Liga Brasileira Contra a Tuberculose e a Liga Paulista contra a Tuberculose, fundadas em 1899, foram as primeiras instituições criadas com o intuito de combater a doença. As principais ações consistiam em educar por meio de campanhas sanitárias métodos de tratamento e profilaxia, focando o atendimento à classe mais pobre, capacitando profissionais e implantando sanatórios (HIJJAR MA et al, 2007). Posteriormente estas entidades receberam apoio financeiro do estado e iniciaram a construção de casas populares, bem como fomentaram debates sobre alimentação adequada e trabalhos insalubres (GONÇALVES H, 2002).

Com a Reforma Carlos Chagas, em 1920, deu-se a criação do Departamento Nacional de Saúde Pública, instituindo-se a Inspeção de Profilaxia da TB, o primeiro órgão governamental de combate à TB (HIJJAR MA et al, 2007). Dentre os papéis desempenhados pela Inspeção, estavam: encontrar os doentes e educa-los para não disseminar a TB, bem como realizar ações de profilaxia, notificação, isolamento e desinfecção. As medidas terapêuticas adotadas eram isolamento, repouso, assepsia, boa alimentação e climatoterapia, ao recomendar a mudança do doente para as montanhas em busca de um ar mais puro. Dentre as cidades brasileiras que mais recebiam pacientes com TB para este tipo de tratamento, destacam-se Belo Horizonte e Campos do Jordão (GONÇALVES H, 2002).

Após a criação do Serviço Nacional de Tuberculose em 1941 e da Campanha Nacional Contra a Tuberculose (CNCT) em 1946, pode-se considerar que a TB passou a ser vista como uma questão de saúde pública. A CNCT almejava atuar em 66 municípios brasileiros que eram responsáveis por cerca de 80% dos óbitos por TB. Dentre as iniciativas, destacava-se a

criação de um modelo de sanatório popular, incrementar a vacinação BCG, promover o desenvolvimento científico e tecnológico e também capacitar pessoal para atuar diretamente nestes programa de atenção a saúde (HIJJAR MA et al, 2007). Ao constatarem a ocupação crítica dos serviços de saúde para o atendimento dos pacientes com TB, criou-se a política de visita a domicílio. Aproveitava-se a oportunidade da visita para verificar a estrutura econômica, social e familiar dos pacientes.

Durante as décadas de 50 e 60, o cenário da TB começou a mudar graças ao uso de antimicrobianos para seu tratamento, com queda da mortalidade, em nível mundial, principalmente em países de elevado nível sócio-econômico. (PICON et al., 1993).

No ano de 1965, já não era mais necessário o isolamento do paciente para o tratamento pois o mesmo era feito utilizando drogas anti-tuberculostáticas por 12 meses. Isto permitia que o doente convivesse na sociedade. A responsabilidade pelo controle e disseminação da doença, agora, recaía em grande parte sobre o paciente e não mais no meio social. Tal mudança de paradigma se deu pelo fato de ser permitido ao paciente começar a realizar o tratamento medicamentoso auto administrado. No entanto, tal política de autonomia acarretou no descontrole ao invés do controle dos casos pois constatou-se que ter a liberdade de manipular e dispor de sua terapêutica não necessariamente significava segui-la à risca (GONÇALVES H, 2002).

Na década de 60, foram realizados estudos na Índia, nos quais, ficou evidenciado que, com os novos medicamentos, a cura da TB ocorria na mesma frequência em pacientes internados em sanatórios ou atendidos em nível ambulatorial. Na década de 70, as orientações da OMS era promover o fechamento de sanatórios e criação de ambulatórios para atendimento dos pacientes com TB. No final da década de 70, era consenso que a TB estava controlada e não era mais necessário a participação da sociedade civil e da academia. As ações de controle passaram a ser coordenadas pelos Programas Nacionais de TB, em países de alta carga. No entanto, ao final da década de 80, ocorreram surtos de TB resistentes nos EUA e Europa, principalmente em pacientes HIV positivos com elevada proporção de mortalidade. Em paralelo, observou-se que em países africanos com elevada proporção de infecção de HIV na população geral, os indicadores epidemiológicos de TB sinalizavam que a TB estava aumentando.

Atenta ao considerável aumento dos casos de TB mundialmente, a OMS criou em 1993 o Plano Emergencial para o Controle da Tuberculose. Tal plano repercutiu na forma de administrar o tratamento da doença com a estratégia do Tratamento Diretamente Observado (TDO). Os objetivos eram ambiciosos para a época: Atingir 85% de cura no tratamento, 70%

de detecção de casos e até 5% no máximo de casos de abandono (WHO,1996).

No final da década de 90, ocorreu aumento TB resistente as drogas no leste europeu, na China e na Índia. Em seguida a OMS reconhece que a TB droga resistente passa a ser um desafio importante para o controle da TB em nível global (WHO, 2000).

No Brasil, o fornecimento de medicamentos tornou-se gratuito para o tratamento da TB em 1971. Em 1979, ocorreu a padronização do tratamento com o esquema de curta duração, previsto para 6 meses, e auto administrado. E no final da década de 70 e início da década de 80, foi enfatizada a recomendação da OMS, e iniciou-se o fechamento de sanatórios ou mudança de suas características para hospital geral e a priorização do atendimento ambulatorial no controle da TB (MOREIRA, 2010)

O Serviço Nacional de Tuberculose foi substituído pelo PNCT na década de 80. Com esta mudança o tratamento da doença tornava-se centralizado e dava subsídios operacionais e técnicos às instituições de saúde. Tal programa sugeria uma organização com supervisão permanente nas unidades de saúde, capacitação profissional e estabelecimento da forma auto-administrada como única modalidade de acompanhamento do tratamento do paciente. Na época, apesar das limitações, a remodelação das ações técnicas e operacionais tiveram um impacto positivo, resultando na diminuição de até 6% dos casos de TB por anos em todas as regiões (GONÇALVES, 2002).

Com a implementação do SUS na década de 90, houve também a descentralização do atendimento de pacientes de TB. Esta descentralização previa que as autoridades sanitárias federais seriam responsáveis pelo financiamento e os estados e municípios encarregariam-se da gestão dos recursos humanos e financeiros. No entanto, da forma como a descentralização ocorreu, vários municípios não tinham as capacidades gerenciais para desempenhar as tarefas acordadas, dada a falta de integração entre os diferentes níveis. O SUS havia sido implementado sem um planejamento racional para integrar os programas e a descentralização do serviço. Vários recursos gerenciais importantes que deveriam dar suporte aos municípios não foram providos, acarretando em uma perda da capacidade regulatória, sem diretrizes e incentivos. (KRITSKI, 2000).

No Brasil, a situação tornou-se mais difícil quando, também no final da década de 80 e início da década de 90, houve aumento de infecção por HIV associada a TB. O cenário não era favorável devido aos diversos fatores: coinfeção TB-HIV, dificuldades nos recursos humanos nos serviços de saúde, ausência de articulação entre os gestores de TB e HIV/Aids, piora das condições socioeconômicas da população devido à alta inflação.

A implantação do TDO no Brasil iniciou-se em 1998, em alguns municípios,

gerenciada pelos Programas de Controle a Tuberculose. No mesmo ano, em caráter nacional, o PNCT “definiu a TB como prioridade entre as políticas governamentais de saúde, estabelecendo diretrizes para as ações e fixando metas para o alcance de seus objetivos”. O plano tinha como pedra angular a atuação das equipes do Programa de Agentes Comunitários em Saúde (PACS) e Programa de Saúde da Família (PSF) para estimular a adesão terapêutica e diminuir o abandono do tratamento, criando ações de controle da TB, realizando o tratamento supervisionado, além de estimular que os vários municípios também aderissem aos Programas propostos pela Secretaria Estadual e pelo governo federal. Estas políticas de controle da TB foram implementadas com o objetivo de avançar com a implantação do TDO e aumentar paulatinamente a cobertura das Unidades Básicas de Saúde: de 3% em 1998 para 63,8% em 2005 (RUFFINO- NETTO A & VILLA TCS, 2006).

Em 2006 foi lançada a Estratégia “*Stop TB*”, no Dia Mundial da Tuberculose (24 de março), destinada a encontrar um Desenvolvimento de Metas relacionadas à Doença para o Milênio (MDG), bem como para estabelecer parcerias e metas para estruturar um Plano Global para 2006-2015 (WHO, 2015). As metas propostas foram: interromper e reverter a tendência do coeficiente de incidência da TB até 2015 (meta estabelecida nos Objetivos do Milênio); e reduzir em 50,0% a prevalência e a mortalidade por TB em 2015, em comparação a 1990 (meta do *Stop TB Partnership*). Nesta Estratégia *Stop TB*, a pesquisa foi incluída como parceira no controle da TB, pois não era considerada relevante desde o final da década de 70.

O final do ano de 2015 foi marcado pela transição do MDG, estabelecido em 2000, para a estrutura de desenvolvimento pós-2015. A meta futura da estratégia global pós-2015 visava um mundo livre da TB, com redução das mortes em 75% e em 50% na taxa de incidência, até 2015. Continuando com seus objetivos ambiciosos também estabelecia uma meta de redução em 95% das mortes e 90% na taxa de incidência para 2035 (WHO, 2015).

A Estratégia Global para Acabar com a Tuberculose (*Global End Tuberculosis Strategy*), proposto pela OMS em 2015 e referendado pela Assembleia Mundial de Saúde em 2018, visava reduzir as mortes e a incidência de TB em todos os países aos níveis observados em países de alta renda. Isso poderia ser alcançado por meio da redução da mortalidade, melhoria do diagnóstico precoce, tratamento mais eficaz, monitoramento de possível resistência micobacteriana e expansão do rastreamento de contato e controle de infecção. A estratégia é baseada em três pilares: atenção e prevenção integrada e centrada no paciente (Pilar 1); políticas ousadas e sistemas de apoio (Pilar 2); e intensificação da pesquisa e inovação (Pilar 3). (WHO, 2015; KRITSKI et al., 2016).

No Brasil em 2017, foi incorporada a Estratégia pelo Fim da TB, seguindo as

diretrizes da OMS. O combate à TB, coordenado pelo PNCT, passou a priorizar os três pilares com os seguintes objetivos:

Pilar 1: Diagnosticar precocemente todas as formas de TB, com oferta universal de cultura e teste de sensibilidade, incluindo o uso de testes rápidos; Tratar de forma adequada e oportuna todos os casos diagnosticados de TB visando a integralidade do cuidado; Intensificar as atividades colaborativas TB-HIV; Intensificar as ações de prevenção;

Pilar 2: Fomentar ações para garantir a realização das atividades de cuidado e prevenção da doença com recursos adequados (humanos, infraestrutura e financeiros); Fortalecer a articulação intra e intersetorial para garantia dos direitos humanos e cidadania nas ações de controle da doença; Fortalecer a participação da sociedade civil nas estratégias de enfrentamento da doença; Melhorar a qualidade dos sistemas informatizados de registro de casos para tomada de decisão mais oportuna;

Pilar 3: Estabelecer parcerias para afomentar a realização de pesquisa no país em temas de interesse para a saúde pública; Promover a incorporação de iniciativas inovadoras para aprimorar o controle da TB; (MINISTÉRIO DA SAÚDE, 2017)

Os dados dos indicadores prioritários para controle da TB objetivavam uma proporção de cura dos casos novos de TB pulmonar de 85%, aliada a uma proporção de abandono de tratamento abaixo de 5% e proporção de detecção acima de 70%, para que ocorresse redução da incidência em torno de 5,0% a 10,0% ao ano.

3.1.5. Tratamento diretamente observado (TDO)

Implantado em 1994, primeiramente na Índia, país que era líder na quantidade de casos de TB, o TDO, foi primeiramente denominado como Programa Nacional Revisado de Controle da Tuberculose (RNTCP). Dentre suas principais políticas estavam: comprometimento das autoridades governamentais, tentar alcançar o maior número de detecções passivas dos casos, confirmação dos casos por bacterioscopia, fornecimento de medicamentos adequados, monitoramento e avaliação constante do sistema.

Em 2001, foi dado o pontapé inicial no primeiro plano de ação para controle da TB. Previsto para o período de 2001 a 2005, tinha como objetivo o aumento da cobertura da estratégia TDO em 50%, atingindo cerca de 4 milhões de pessoas, e o aumento da detecção de casos novos (RUFFINO-NETTO A & VILLA TCS, 2006).

Durante a Conferência Ministerial sobre Tuberculose e Desenvolvimento Social, em 2003, foi firmada a “Declaração de Amsterdã”, cujo principal tema abordado era o desenvolvimento e aperfeiçoamento de soluções para o controle da TB em países em

desenvolvimento. Várias frentes foram definidas para nortear as ações: envolvimento da sociedade civil, a utilização de tecnologia acessível e eficaz para proporcionar um diagnóstico confiável e abarcar serviços de AB no atendimento aos pacientes com TB.

Ao final de 2003, 82 países haviam aplicado a estratégia no Mundo, o que correspondia a 77,0% da população mundial, envolvendo 3,7 milhões de todos os casos de TB, com uma taxa de detecção de 45,0%. Só na região das Américas, 78,0% da população possuía cobertura pela estratégia TDO, com 16 casos novos a cada 100 mil habitantes.

Um ano depois, em 2004, a quantidade de pessoas tratadas mundialmente com o TDO era de mais de 20 milhões de pessoas, com cerca de 16 milhões atingindo a cura. A taxa de incidência havia se mantido, no entanto a mortalidade havia caído. As exceções à estas tendências foram as regiões do extremo leste europeu e na região sub-saariana.

No Brasil, o MS implantou, em 1996, o Plano de Ação Emergencial para municípios considerados prioritários, com base nos seguintes critérios: população com mais de 50 mil habitantes, coeficiente de incidência acima da média brasileira (58,4 por 100 mil habitantes em 1995), óbitos por TB acima de 5,0% dos casos novos no ano, taxa de cura abaixo de 85,0%, taxa de abandono acima de 10,0%, índice sem informação do resultado do tratamento acima de 5,0% e elevada incidência de casos de AIDS. Inicialmente foram selecionados 206 municípios prioritários no País, distribuídos pelas regiões da seguinte maneira: Norte – 26 municípios, 76,1% dos casos da região; Nordeste – 81; 66,5%; Centro-Oeste – 27; 66,6%; Sudeste – 43; 55,7% (sendo 25 destes municípios localizados no Estado de São Paulo); Sul – 29 municípios e 51,0% dos casos da região (RUFFINO-NETTO A & VILLA TCS, 2006).

Devido à delicada colocação que o Brasil encontrava-se em 1998 - 10º lugar no ranking da OMS em número de casos de TB; levando em consideração a persistência do problema com alta proporção de abandono, baixo sucesso nas curas e detecção dos casos de TB, o Conselho Nacional de Saúde, por meio da Resolução nº284/98, foram lançadas dentro do Plano Nacional de Controle da TB as metas:

- Diagnosticar 90,0% dos casos esperados até 2001;
- Tratar com sucesso 85,0% dos casos diagnosticados;
- Reduzir em nove anos (até 2007) a incidência para pelo menos 50,0%;
- Reduzir a mortalidade em dois terços.

Segundo o plano, o modelo de atenção do SUS seria remodelado com o PSF e dos PACS. Estratégias de valorização do desempenho dos municípios em combater a TB e melhorarem a taxa de alta por cura foram criadas. Os municípios receberiam um incentivo financeiro em forma de bônus caso tivesse melhora no desempenho. Todas estas ações

fizeram parte da estratégia de Tratamento Supervisionado que foi implantada no Brasil (RUFFINO-NETTO A & VILLA TCS, 2006).

Em 2002, a estratégia TDO estava disponível no Brasil para 35,0% da população, apresentando aumento de 34,0% no ano seguinte e de 18,0% de aumento na proporção de detecção de casos novos, resultando na taxa global de detecção de 55,0% nas áreas com a cobertura da estratégia. A proporção de sucesso em 2002 foi 75,0%, com abandono em 18,0% (RUFFINO-NETTO A & VILLA TCS, 2006).

O TDO consiste na observação da ingestão dos medicamentos, preferencialmente todos os dias, realizado por profissionais de saúde, e nos finais de semana e feriados os medicamentos auto-administrados. Mas também são considerados em TDO aqueles doentes com 24 doses supervisionadas na primeira fase do tratamento e 48 doses supervisionadas na segunda fase. Sendo que cada unidade de saúde deve assegurar que todo o doente receba TDO na unidade de saúde mais próxima da sua residência e, naqueles casos que não possa comparecer a unidade, a medicação deve ser administrada no seu domicílio (MANUAL DE RECOMENDAÇÕES PARA O CONTROLE DA TUBERCULOSE NO BRASIL, 2011).

O emprego do TDO aproxima os profissionais do contexto social dos indivíduos e também auxilia a identificar grupos de risco para a não adesão ao tratamento e estabelece vínculos entre serviço de saúde-doente-família (RUFFINO-NETTO A & VILLA TCS, 2006). Além disso, proporciona maior probabilidade de curar a TB e diminuir o surgimento de resistência aos medicamentos.

O TDO tem como objetivos:

- Melhorar a atenção ao doente por meio do acolhimento humanizado;
- Possibilitar a adesão, garantindo a cura;
- Reduzir a taxa de abandono;
- Interromper a cadeia de transmissão da doença;
- Diminuir o surgimento de bacilos multirresistentes;
- Reduzir a mortalidade;
- Realizar uma educação em saúde mais efetiva, de forma individualizada voltada para orientar e co-responsabilizar o indivíduo, a família e a comunidade nas ações de saúde. A fim de organizar, supervisionar, capacitar, monitorar e definir ações de vigilância epidemiológica, foi então criado um sistema de informação específico, o SINAN-TB (RUFFINO-NETTO A & VILLA TCS, 2006).

3.1.6. Acompanhamento do tratamento

O acompanhamento do paciente com TB envolve a realização de baciloscopia de controle mensal. As baciloscopias do segundo, quarto e sexto mês são cruciais pois elas orientam o protocolo a ser seguido, contribuindo para a definição de alta por cura, falência do tratamento e também na identificação de TB droga-resistente. Os desfechos são concluídos da seguinte forma:

Alta por cura: quando o paciente tem pelo menos duas baciloscopias negativas, sendo uma na fase de acompanhamento e outra ao término do tratamento.

Falência do tratamento: é verificada quando há persistência da positividade do escarro ao término do tratamento; ou quando aqueles que, no início do tratamento, são positivos (+ + ou + + +) e mantêm a positividade até o quarto mês; ou quando aqueles com baciloscopia positiva inicial seguida de negativação e nova positividade por dois meses consecutivos, a partir do quarto mês de tratamento.

Abandono: quando o paciente fica mais de 30 dias, consecutivamente, sem tomar a medicação desde o dia da última consulta médica.

Casos de resistência aos medicamentos geralmente são verificados e estão vinculados ao cenário de abandono ou falência do tratamento.

3.1.7. TB droga Resistente

A resistência aos fármacos é dividida em:

Natural: é aquela que é inerente ao processo multiplicação do bacilo.

Primária: ocorre em pacientes que nunca tiveram TB e são contaminados por bacilos já resistentes.

Adquirida ou secundária: é a situação em que inicialmente os pacientes eram sensíveis às medicações mas, no decorrer do tratamento, tornaram-se resistentes.

As principais associações com os casos de resistência adquirida estão: tabagismo durante o tratamento, aumentando o risco de recidiva da TB em casos clínicos já tratados, pouca adesão e uso não contínuo da medicação por falta de medicação ou irregularidade e abandono. O sistema de informação responsável por registrar os pacientes que apresentam resistência aos fármacos é o SITE-TB. Segundo a OMS, os tipos de resistência são divididos em (WHO, 2015):

Monorresistência: resistência a somente um medicamento anti-tuberculostático de primeira linha.

Polirresistência: resistência à mais de um medicamento de primeira linha, que não seja

isoniazida e rifampicina simultaneamente.

Multi drogarr resistente (MDR): resistência a, pelo menos, isoniazida e rifampicina juntos.

Resistência Extensiva (XDR): resistência à alguma fluorquinolona e, pelo menos, um dos três medicamentos injetáveis de segunda linha (capreomicina, canamicina e amicacina), adicionado à multidrogarr resistência.

Resistência à Rifampicina (RR): resistência à rifampicina detectada usando métodos de fenótipos ou genótipos, com ou sem outras resistências associadas. Inclui qualquer RR, seja monorresistência, polirresistência, MDR ou XDR.

Em 2019, no mundo, ocorreram 3,36% de novos casos de TB drogarr resistente (TB-DR) e outros 17,8% que também apresentaram resistência mas, neste caso, eram pacientes em retratamento. Dos 465.000 casos de TB-DR ao redor do mundo, em 2019, 182.000 vieram a óbito por TB. (WHO, 2020)

Entre 2015 e 2020, foram diagnosticados 7.749 casos de TB-DR. 5.377 (69,5%) dos casos apresentaram resistência a rifampicina ou multirresistência, 1.731 (22,4%) dos casos apresentaram monorresistência, 542 (7,0%) polirresistência, e 92 (1,2%) resistência extensiva (SECRETARIA DE VIGILÂNCIA EM SAÚDE, 2021).

Infelizmente, para os casos de resistência bacteriana aos fármacos anti-TB, os tratamentos indicados são, muitas vezes, insatisfatórios se levarmos em consideração a duração, custo e efetividade. São tratamentos com uma janela de tempo grande, com posologia e formas de aplicação inconvenientes para o paciente, além de diversos efeitos colaterais. Neste sentido, incentiva-se o desenvolvimento de novas pesquisas nesta área visando o desenvolvimento de novos fármacos que possibilitem menor tempo de tratamento, facilitando a adesão ao tratamento.

Mudanças nas intervenções na área do cuidado em saúde relativo ao processo de tratamento da TB tornam-se necessárias para melhorar a qualidade de vida destes pacientes. Ressalta-se também a importância da equipe de cuidados paliativos para os casos terminais e graves. Equipe esta que presta suporte fundamental quando resta apenas a opção de oferta de conforto e dignidade no final de suas vidas. O aspecto psicológico também não deve passar despercebido pela equipe responsável pelos cuidados, visto que há uma grande discriminação dos pacientes acometidos pela TB, sendo amplamente estigmatizados pela sociedade, e as vezes até pela família, em razão do desconhecimento sobre TB (WHO, 2015).

Todos esses cuidados são importantes no tratamento e acolhimento desses pacientes para que se possa aumentar a aderência e eficácia da terapêutica instituída e, como

consequência, a redução das proporções de não aderência/abandono, propagação da TB-DR e surgimento de cepas com resistência extensiva.

3.1.8. Fatores de risco associados com abandono do tratamento

Em locais onde o programa de TB apresenta bons indicadores de sucesso no controle da TB, necessário focar em alguns grupos de pacientes possuem uma tendência a abandonar o tratamento. Podemos citar como exemplo a população carcerária, usuários de drogas, fumantes, alcoólatras, moradores de rua, portadores de HIV e portadores de síndromes e desordens mentais (CAMPANI STA et al, 2009). Tais grupos exigem uma dedicação especial e o acompanhamento da equipe do TDO (WHO, 2006). O acompanhamento constante do paciente realizado pela equipe auxilia a identificar sinais preditivos que podem indicar um possível abandono do tratamento ou interrupção. Idealmente o TDO deve acompanhar todos os pacientes no uso regular da medicação até o término do tratamento e a cura, diminuindo o risco de surgimento de isolados de *M.tuberculosis* com resistência aos fármacos. No entanto, dado os diversos cenários encontrados na área da saúde brasileira, onde há escassez de recursos humanos e financeiros, além das mais variadas diferenças culturais e socioeconômicas, em algumas situações, fica praticamente impossível exercer o que está estabelecido nos manuais de tratamento proposto pelo nível federal. Nestes cenário, a equipe de saúde local, passa a adotar medidas que identifiquem e priorizem o TDO para os grupos de alto risco para o abandono, afinal são estes pacientes que mais precisam do cuidado em saúde (BARREIRA D, 2012).

Baixo nível socio econômico e a dificuldade de acesso à informação e educação promovem o analfabetismo, baixa escolaridade e baixo nível de desenvolvimento humano, causando um cenário de desconhecimento dos sintomas, causas e tratamento da doença, bem como de sua importância no contexto social. As dificuldades econômicas de certos grupos de pacientes pode ser tão grande que estas pessoas não têm condições financeiras de custear o transporte até as unidades de atendimento para realizar suas consultas e exames (CRISTINA RFM et al, 2012; TESFAHUNEYGN G et al, 2015). A situação socioeconômica de uma população é fator influenciador direto de vulnerabilidade no processo saúde-doença, acarretando no abandono e persistência do ciclo de contaminação (MUTURE BN et al, 2011; HARGREAVES JR et al, 2011).

O preconceito e a não aceitação da doença no círculo de convívio do doente também acarreta em negligência por parte do paciente, bem como no abandono do tratamento. A faixa etária acometida pela TB entre 15 e 54 anos é constituída de indivíduos economicamente

ativos, de maioria do sexo masculino, que não possuem hábitos alimentares regulares e apresentam dificuldade para cumprir seus compromissos médicos e para manter uma regularidade na administração da medicação.

Pode-se avaliar o nível socioeconômico utilizando diversas variáveis (analfabetismo, escolaridade, renda). No entanto o mais conhecido é o Índice de Desenvolvimento Humano (IDH). Seu cálculo baseia-se em um relatório do Programa das Nações Unidas em três pilares: expectativa de vida ao nascer (saúde), média de anos de estudo (educação/escolaridade) e renda. O Brasil tem IDH = 0,754.

Um fator de risco e um mau prognóstico para o desenvolvimento da TB é quando o paciente está infectado por HIV (LANNON LH et al, 2008). Uma maior mortalidade é observada nos pacientes soropositivos em comparação com os pacientes soronegativos para HIV. Em alguns estudos em países do Mediterrâneo, notou-se uma grande diferença nos desfechos dos tratamentos de pacientes HIV-positivos e HIV-negativos, onde o sucesso foi de 53% e 82%, respectivamente. A chance de pacientes HIV-positivos evoluírem para o óbito no decorrer do tratamento para TB foi três vezes maior que os demais, com proporções de 11% contra 3,4%, respectivamente. Nas Américas, a proporção é de 19% dos pacientes com coinfeção HIV-TB que evoluíram para a óbito no decorrer do tratamento e somente 5% dos HIV-negativos (GLOBAL TUBERCULOSIS REPORT, 2014).

No Brasil, em 2015, 68,9% dos pacientes com TB realizaram testagem para HIV e revelaram que 9,7% dos casos apresentavam co-infecção TB-HIV. Na região Sul (Paraná, Santa Catarina e Rio Grande do Sul) observou-se o maior percentual de co-infectados (17,3%). As capitais Porto Alegre (RS), Curitiba (PR) e Florianópolis (SC) destacaram-se com as maiores proporções – 25,2%, 21,5% e 21,2%, respectivamente. Também estão entre os municípios que realizam testagem em mais de 80% de todos os casos de TB (SECRETARIA DE VIGILÂNCIA EM SAÚDE, 2016).

A mortalidade no grupo de co-infectados HIV-TB pode ser reduzida ao se realizar a testagem e o diagnóstico precoce do HIV, possibilitando ações rápidas para o início do tratamento anti-retroviral (TARV). Considerando que a TB é a doença que mais leva pacientes com HIV a óbito, esta intervenção rápida e no momento correto é importantíssima para os indivíduos destas populações terem uma sobrevida e uma chance maior de sucesso. No Brasil, em 2016, somente 39,7% dos novos casos com coinfeção HIV-TB utilizavam a terapia antirretroviral. A coinfeção HIV-TB deve receber uma atenção em especial no tratamento da TB não somente devido aos altos índices de mortalidade, mas também pela maior propensão ao desenvolvimento de efeitos colaterais como hepatotoxicidade e insuficiência renal.

Outros fatores comuns de risco são o uso de drogas ilícitas e lícitas pelos pacientes. O alcoolismo, em associação ao desenvolvimento do quadro de TB é considerado um fator de mau prognóstico por dois motivos: clinicamente devido às intercorrências e complicações causadas pelos efeitos colaterais como hepatotoxicidade mas, principalmente, pela alta proporção de abandono verificada nesta população. No estudo de ALBUQUERQUE MFM et al (2001), o abandono ocorreu com maior frequência entre os dependentes de álcool, e o de BERGEL ES & GOUVEIA N (2005), apresentou uma probabilidade quase quatro vezes maior de abandono do tratamento quando os pacientes eram dependentes de álcool. Recomenda-se, portanto, que o alcoolismo associado à TB receba atenção pela equipe de saúde de uma forma especial, pois o comportamento desse grupo apresenta uma grande tendência ao abandono do tratamento (SILVA CCAV et al, 2013).

A associação do etilismo com a TB é amplamente estudada e conhecida, mas outras drogas como tabagismo e drogas ilícitas também estão associadas com abandono do tratamento. Em 2008, STORY A et al realizaram em Londres um estudo para identificar a associação do uso de crack e a infecção por *Mycobacterium tuberculosis*. Neste estudo evidenciou-se que os usuários de crack apresentaram-se com quadro clínico mais grave, desnutridos, com má aderência ao tratamento e, conseqüentemente, com maiores índices de falência, e aumento da resistência bacteriana. Como o crack causa lesões pulmonares levando à alteração da função do macrófago alveolar e prejuízo da produção de citocinas, o “pulmão do crack” fica mais suscetível às doenças infecciosas, como a TB. Além disso, essa população apresenta-se, em geral, com a forma bacilífera da doença; são tossidores crônicos, apresentando lesões infiltrativas e edema pulmonar, hemoptise e dispnéia (STORY A et al, 2008). Esses fatores relacionados ao acometimento pulmonar, associados com a má aderência, e também com o fato de os usuários de drogas conviverem em ambientes confinados e em situação de rua, expõem sua fragilidade, exigindo uma atenção maior. A realização de TDO nesse grupo de usuários de drogas ilícitas deve ser utilizada como estratégia para redução do abandono, da falência ao tratamento e, conseqüentemente, do surgimento de cepas resistentes nessa população de risco.

Por fim, o tabagismo está relacionado com a progressão da doença, gravidade da doença e aumento do risco de recidiva pós tratamento, bem como o aumento da mortalidade (ANGELES AJM et al, 2016).

Exposição ao tabaco causa alterações imunológicas nos pulmões, atuando nos macrófagos alveolares, reduzindo a produção de TNF-alfa, IFN- γ , alteração do movimento mucociliar da célula epitelial, demora para negatificação do escarro, levando ao aumento do

período de transmissibilidade.

Em estudo realizado no município de São Paulo concluiu-se que os tabagistas, os etilistas e os usuários de drogas ilícitas abandonaram o tratamento da TB com maior frequência que os demais ($p=0,009$, $p=0,008$, $p=0,002$, respectivamente) (RIBEIRO SA et al, 2000). Portanto, pacientes tabagistas também devem ter uma atenção especial, tanto por se exporem a um fator de risco para abandono, como também pela sua relação com maior recidiva e associação com maior mortalidade (HARGREAVES JR et al, 2011).

3.1.9. Retratamento

Em 2015, ocorreram no Brasil 12337 casos de retratamento de TB, representando 16,3 do total de casos. O estado com a maior proporção é o RS, com 24,5% acompanhado da capital Porto Alegre apresentando 35,1% de casos de retratamento, ficando em primeiro lugar entre as capitais com a maior proporção (SECRETARIA DE VIGILÂNCIA EM SAÚDE, 2016).

Culturas de escarro foram feitas em apenas 30,7% dos casos, com destaque especial para Palmas como o melhor índice entre as capitais com 83,3% e Cuiabá com o pior (5,6%). 73,5% dos casos de retratamento foram testados para HIV, apresentando uma proporção de coinfeção TB- HIV quase duas vezes maior neste cenário (18,1%) se comparados aos novos casos (9,7%). Nos estados do Rio Grande do Sul e Santa Catarina observam-se os maiores percentuais de coinfeção sendo 35,3% e 35,4%, respectivamente. Não surpreende que suas respectivas capitais sigam este padrão apresentando uma proporção de 43,6% em Porto Alegre e 55,1% em Florianópolis. (SECRETARIA DE VIGILÂNCIA EM SAÚDE, 2016).

Os dados supracitados demonstram uma tendência para casos de retratamento, seja por abandono ou recidiva, onde deve-se dispender um cuidado maior e um acompanhamento mais próximo da equipe de saúde. Isto se deve ao fato de que, em ambos os casos, a proporção de cura é menor e a proporção taxa de abandono é muito maior do que os pacientes que estão em seu primeiro tratamento da TB.

Demais estudos realizados em nível global e nacional reforçam esta hipótese (ALBUQUERQUE MFM et al, 2001; KIGOZI G et al, 2017; OLIVEIRA HB et al, 2000).

3.2. WEB SEMÂNTICA

3.2.1. Histórico e evolução

A maior parte da informação contida na *web* foi pensada para ser lida e compreendida por

humanos. Programas computacionais ou páginas na web tem a habilidade de interpretar apenas alguns atributos das próprias páginas em si ou de seus conteúdos (LIMA; CARVALHO, 2004; ISOTANI; BITTENCOURT, 2015). No entanto, há uma limitação na compreensão e processamento do significado dos conteúdos hospedados na *web* (PINOCHET, 2011). Para solucionar esta deficiência, o paradigma da WS foi introduzido (LOPES, 2004). O objetivo da WS é nutrir um ambiente onde os agentes de *software* (programas) podem executar tarefas sofisticadas, em nome dos usuários, usando informações com um significado bem estabelecido de forma que este processo possa ser feito de forma automática, resultando num esforço colaborativo entre humano e computadores (VARGAS-VERA et al., 2002; BRATSAS et al., 2010; GIANNOPOULOU et al., 2014; ISOTANI; BITTENCOURT, 2015).

No paradigma da WS, serviços online e agentes de *software*, assistidos pelo uso de ontologias (que descrevem o domínio de aplicação, determinam seu vocabulário e descrevem os relacionamentos entre os seus elementos), são capazes de interagir de forma autônoma para satisfazer as necessidades dos usuários (SOUZA; ALVARENGA, 2004). Serviços são descritos pela especificação de serviços. A especificação do serviço não está limitada às suas funções ou características; há também uma preocupação na definição de propriedades e requisitos não-funcionais (LIMA; CARVALHO, 2004).

A WS utiliza duas tecnologias que se complementam: XML e o RDF. O XML permite a construção de uma estrutura para conteúdos e documentos, mas não provê uma maneira clara de especificar o significado representado pela estrutura (BERNERS-LEE; HENDLER; LASSILA, 2001; LIMA; CARVALHO 2004; DEVEDZIC, 2006; ISOTANI et al., 2009; TILAHUN, 2012; YU, 2014). O RDF compensa esta deficiência, provendo significado para o conteúdo e seus documentos usando uma arquitetura de triplas para descrever os dados (BERNERS-LEE; HENDLER; LASSILA, 2001).

A WS é desafiada no seu âmbito de descrever dados e regras usadas em inferências quando um significado é definido por dois ou mais termos diferentes (ISOTANI, BITTENCOURT, 2015). Este percalço pode ser resolvido incluindo uma coleção de ontologias (também conhecido como vocabulários estruturados) que definem os relacionamentos entre os termos e conceitos (classes) (FREITAS, 2003; LIMA; CARVALHO, 2004; CANTARELLI, 2005; ISLAM, 2010; MODESTO, 2010). Ontologias melhoram a comunicação e identificação dos serviços que estão disponíveis para satisfazer uma determinada requisição de serviço.

A ontologia para a *web* é composta por uma “taxonomia e um conjunto de regras de inferência”. A taxonomia provê a definição das classes dos objetos e os seus relacionamentos.

O conjunto de regras provê os mecanismos de processamento necessários para manipular os termos e extrair significados (YAGUE et al., 2003; BASSIN, 2004; KIFER et al., 2005; HORROCKS et al., 2005; DUBOST; HERMAN, 2008).

Atualmente ontologias são comumente expressas na *Web Ontology Language* (OWL) que é fundamentada na *Description Logics* (DL). OWL permite inferir relacionamentos entre conceitos e suas definições. É uma evolução de seus predecessores DARL e OIL.

Outras abordagens também foram desenvolvidas para prover alternativas para habilitar a descrição semântica de ontologias (VRANDECIC et al., 2003; BERNERS-LEE, 2005; BERNERS-LEE, 2006; RAMALHO; VIDOTTI; FUJITA, 2007; MODESTO, 2010). A W3C desenvolveu o *Web Service Architecture* (WSA) *framework* para prover definições padrões para os modelos de web services, a o relacionamento entre diferentes componentes (BERNERS-LEE, 2005; BERNERS-LEE, 2006). *Web services* (WSS) são definidos como “um sistema de *software* desenvolvido para suportar a interoperabilidade máquina-máquina através de uma rede”.

WSS são caracterizados por duas entidades principais: os provedores e os clientes dos serviços. Estas entidades trocam mensagens através do uso da *Web Service Description Language* (WSDL), que define as especificações funcionais dos serviços (por exemplo, seus tipos de dados, formatos, protocolos) (DÜRST, 2001; ISAAC; SUMMERS, 2009; NOWACK, 2009; DIMITROV, 2010; HERMAN, 2015; KIFER; BOLEY, 2013; ABIDA; BIRBECK, 2015; PRUD’HOMMEAUX; CAROTHERS, 2014; ISOTANI, BITTENCOURT, 2015). A semântica em WSS refere-se ao seu comportamento esperado durante as interações com o mesmo. Representa um contrato ou acordo entre os participantes.

Serviços são tipicamente identificados pelo URI, e se apoiam em tecnologias para a troca de mensagem e protocolos de transporte (HTTP ou SOAP). Uma URI pode ser usada para identificar recursos físicos, como celulares, tvs, etc. A linguagem RDF pode descrever atributos destes recursos, e a forma com que eles podem operar em conjunto como se fossem agentes de *software*.

A descrição dos serviços incluem informação sobre a interface dos serviços e os metadados que podem ser usados durante o processo de descoberta do serviço. Todos estes devem, idealmente, serem escritos em uma linguagem que seja compreendida por máquinas.

O termo perfil refere-se aos “documentos trocados entre dispositivos que descrevem as capacidades dos mesmos”. O uso de ontologias em perfis permitem busca e descoberta automática de funcionalidades necessárias ao usuário que estão bem descritas e disponíveis em dispositivos para eventual utilização de seus serviços.

Os agentes são tangíveis “fragmentos de *software* ou *hardware* que enviam e recebem mensagens”, enquanto serviços são pacotes de funcionalidades que são ofertados para serem usados através de requisições. Agentes desempenham um papel chave no sucesso da WS pois eles são os responsáveis por coletar conteúdo, processá-lo e distribuí-lo; eles podem até permitir que agentes que não foram previamente desenhados para atuarem em conjunto a colaborarem com a transferência de dados sempre que solicitado pelo significado semântico dos dados.

WSS e WS se complementam, evoluindo para o que é conhecido com *semantic web services*. *Semantic web services* são entidades autossuficientes e auto descritas que semanticamente ofertam suas operações e descrições e auxiliam na descoberta, composição e união de serviços. *Semantic web services* são suportados por ontologias e usam linguagens que podem ser interpretadas por máquinas e geralmente funcionam em ambientes heterogêneos e abertos.

A descoberta de serviços é comumente realizada por agentes que buscam por descrições de serviços compatíveis com um critério funcional e semântico desejado, especificado no momento da requisição feita pelo usuário. Há dois importantes aspectos da descoberta de serviços: a arquitetura de descoberta de serviços, determinando onde os serviços podem ser divulgados e disponibilizados, e o mecanismo de compatibilidade de serviço, determinando como serviços podem ser comparados com os requisitos das solicitações dos usuários até se encontrar um que se adeque com a maior perfeição possível. Tipicamente, apenas características funcionais de um serviço são utilizadas na comparação de serviços, ignorando-se as propriedades não funcionais. Propriedades não funcionais adicionam critérios de compatibilidade extra, estreitando o número de web services com funcionalidades equivalentes. Várias abordagens a fim de facilitar a especificação semântica de web services foram propostas, incluído as seguintes:

O programa DARPA propôs o *DARPA Agent Markup Language* (DAML) como uma extensão do XML e RDF. A *Ontology Inference Layer* (OIL) adiciona uma ontologia sob a camada da WS. DAML e OIL foram sucedidas pela DAML+OIL, uma linguagem que incluía funcionalidades de ambos antecessores, e eventualmente evoluiu na OWL.

A organização da DAML propôs o uso da OWL como a “linguagem de representação escolhida para a proposta OWL-S”. A ontologia de serviços, proposta pela *OWL Service Coalition*, provê informação acerca: (a) o perfil de serviço que representa cada serviço e descreve suas funcionalidades e características; (b) o modelo de serviço que descreve como o serviço funciona; (c) a base do serviço que provê detalhes como o serviço pode ser acessado

(DÜRST, 2001; ISAAC; SUMMERS, 2009; NOWACK, 2009; DIMITROV, 2010; HERMAN, 2015; KIFER; BOLEY, 2013; ABIDA; BIRBECK, 2015; PRUD'HOMMEAUX; CAROTHERS, 2014; ISOTANI, BITTENCOURT, 2015).

A *Web Services Modeling Ontology* (WSMO) permite a especificação de serviços e provê suporte para identificar incoerências na descrição dos serviços. Ela utiliza a *Web Service Modeling Language* (WSML).

Num esforço para incorporar significado semântico nas descrições de WSS, a W3C estendeu sua linguagem WSDL na WSDL-S. WSDL-S é um padrão em formato XML, e é usado para descrever uma rede de serviços como um conjunto de endpoints baseando-se na sintaxe dos serviços. A linguagem WSDL-S refina as descrições da WSDL adicionando significado semântico e incorporando à ela alguns conceitos da OWL, WSMO, por exemplo (HERMAN, 2015).

3.2.2. Unicode

O Unicode é um modelo de codificação que provê um número identificador único para cada caractere, independente da plataforma, aplicativo ou linguagem. Age, portanto, como um padrão da indústria da computação para codificação, representação e manipulação de texto de forma padronizada entre os aplicativos de escrita.

O padrão Unicode foi adotado pelos líderes da indústria da computação, como Apple, HP, IBM, Microsoft, Oracle, SAP, Sun, Sybase além de ser usado por diversos padrões modernos como XML e Wireless Markup Language.

Ao incorporar o Unicode nos dados e aplicações, tem-se uma grande vantagem em relação ao uso de conjuntos de caracteres legados específicos de cada fabricante. Unicode permite que uma única aplicação seja alvo de múltiplas plataformas, linguagens e países sem que sofra um processo de reengenharia. Permite que os dados sejam transportados através de diversos sistemas sem que sejam corrompidos. Portanto, é a escolha ideal para a representação e codificação dos caracteres na WS.

3.2.3. Uniform Resource Identifier (URI)

A WS é construída sob os princípios e tecnologias da Web atual, reutilizando a indexação global e o sistema de nomes para os dados contidos nos documentos da WS. Indexação global e sistema de nomes significa que cada informação pode ser acessada por browsers padrões, bem como por aplicações semanticamente sensíveis, de maneira unívoca através de um identificador único. A WS utiliza o URI para este propósito.

Uma URI é uma *string* de caracteres formatados, também chamados de identificadores web, que definem uma fonte da Internet. É usado para identificar (dar nome) fontes físicas ou abstratas pela internet. Uma URI pode ser classificada como um localizador (URL), um nome (URN), ou ambos.

O URI mais comum é o Uniform Resource Locator (URL) que identifica o endereço de um domínio na Internet. Outro tipo comum de URI é o Uniform Resource Name (URN).

O termo URL refere-se a um subconjunto de URIs que identificam recursos via representação de seus mecanismos de acesso primários ao invés de identifica-los pelo nome ou por outro atributo. Diferente das URLs, as URIs não se limitam a identificar recursos que necessariamente possuem uma localização na rede.

O termo URN refere-se ao subconjunto de URI que são necessárias para permanecerem globalmente únicas e persistentes mesmo quando o recurso deixa de existir ou fica indisponível. As URIs, na WS, provê um jeito simples de identificar um recurso na Internet, bem como seus relacionamentos com outros recursos. O RDF, usa referências baseadas no URI para dar nome a sujeitos, objetos e predicados.

3.2.4. Extensible Markup Language (XML)

XML é uma linguagem de marcação que permite ao usuário definir suas próprias tags de marcação que podem ser usadas para criar seus próprios padrões de documentos. Um documento XML contém informação estruturada através de texto plano e tags de marcação. As tags de marcação permitem que máquinas interpretem os diferentes blocos de conteúdo (chamados elementos). Documentos XML carregam dados de uma maneira auto descritiva, melhorando a interoperabilidade entre as aplicações envolvidas com a troca de informação. É aceito como um padrão para a troca de dados estruturada na Web e pode ser usado na WS para prover a interoperabilidade sintática.

Em HTML, existe um conjunto de tags que são padrões e suas respectivas interpretações são definidas diretamente pela especificação da HTML. Enquanto isso, o XML permite que o usuário defina sua própria linguagem de marcação (tags e estrutura na qual elas podem aparecer), adaptando-as às suas especificações em particular.

O XML, por si só, não agrega nenhum significado à informação que possa ser processada por máquinas. O significado embarcado nas tags XML é definido pelos usuários humanos que estão trocando os dados. Somente com um acordo prévio entre as partes acerca dos termos usados nas tags, chamado de ontologia, é possível que máquinas processem tais documentos sem uma intervenção humana.

Em um ambiente distribuído como a Web, é possível que diferentes pessoas utilizem os mesmos termos como tags em suas próprias linguagens de marcação, porém com significados totalmente distintos. Por exemplo, a tag <telefone> pode representar um número de celular de uma pessoa em determinada linguagem de marcação, enquanto em outra pode significar o modelo do referido celular. Portanto, mecanismos adicionais são necessários para evitar tais confusões, relacionadas a interpretação das tags, quando os dados de diferentes documentos XML são integrados. Este processo de desambiguação pode ser realizado com a ajuda de XML Namespaces. Um Namespace identifica a parte da web que age como qualificador para um conjunto específico de tags. Um Namespace é definido para uma linguagem de marcação XML através da criação de uma URI para ele. Esta URI geralmente refere-se à página web que descreve a linguagem de marcação (nomes das tags e seus significados). Por qualificar nomes para as tags com as URIs de seus respectivos Namespaces, qualquer um pode criar suas próprias tags e pode diferenciá-las de tags que possuem caracteres idênticos e foram criadas por outros.

O quadro abaixo exemplifica o uso de um XML Namespace.

```
<usp:aluno xmlns:usp="http://usp.br/termos/">
  <usp:curso> Doutorado em Bioengenharia </usp:curso>
  <usp:duracao> 3 anos </usp:duracao>
</usp:aluno>
```

Quadro de código 1 - Exemplo de marcação XML Fonte: própria

No Quadro de Código 1, o atributo `xmlns:usp="http://usp.br/termos/"` declara o Namespace. Ele mapeia o prefixo “usp” para o Namespace URI “http://usp.br/termos/”. O conteúdo XML que queira aderir ao significado das tags descritas nesta URI pode também usar como tags seus qualificadores de nomes (também chamados de QName) como, por exemplo, “usp:curso. Um QName contém um prefixo que identifica o Namespace, seguido de um sinal de dois pontos, e então o nome local para uma tag XML. Ao preencher todos estes requisitos para qualificar os nomes, garante-se que o significado considerado para aquela tag em particular, seja o definido especificamente no seu respectivo Namespace URI e não haja confusão com tags similares criadas por outros usuários. Duas tags que são escritas da mesma

forma (possuem no mesmo nome) só são consideradas a mesma caso elas tenham o mesmo Namespace URI, evitando portanto o conflito de nomes durante as integrações (RDF PRIMER, n.d.).

3.2.5. RDF

O modelo de dados da WS é chamado de Resource Description Framework (RDF). RDF pode ser conceitualmente representado como um modelo baseado em grafo com arcos (predicados) representando as relações entre as entidades (nós) (BASSIN, 2004; LEE et al., 2010; LOPES; OLIVEIRA, 2011; TIMM; RENLY; FARKASH, 2011; ABANDA; TAH; KEIVANI, 2013; GIANNOPOULOU et al., 2014; ISOTANI; BITTENCOURT, 2015). O nó de origem é denominado sujeito e o nó destino é chamado de objeto (DUMONTIER, 2010; LEE et al., 2010; LOPES; OLIVEIRA, 2011; TILAHUN, 2012; ABANDA TAH; KEIVANI, 2013).

Os nós, na WS, podem ser URIs, nós vazios ou um valor literal (KATAYAMA et al., 2013). No entanto, apenas objetos podem conter valores literais: o sujeito e o predicado devem ser URIs (ISOTANI; BITTENCOURT, 2015). URIs são endereços únicos semelhantes aos endereços de páginas web, com a diferença de URIs da WS podem se referir a qualquer coisa que exista no universo do domínio.

O formato de dados do RDF representa a relação de predicado entre o sujeito e objeto como uma tupla constituída por três membros, chamada de tripla (ISLAM, 2010; MACHADO et al., 2013; ISOTANI, 2009; VAN WOENSEL et al., 2014). Os três membros da tupla são sujeito, predicado e objeto. Por exemplo, a seguinte tripla representa que o Brasil tem como cidade capital Brasília: (*Brasil, capital, Brasília*).

A seguinte tripla, da base de dados da DBpedia, representa que o Brasil tem como cidade capital Brasília:

(<http://dbpedia.org/resource/Brazil>,
<http://dbpedia.org/ontology/capital>,
<http://dbpedia.org/resource/Brasilia>)

URIs podem serem abreviadas usando abreviação por namespaces para a primeira parte da URI, dois pontos e, então, a parte única da URI. O uso de namespaces torna a leitura de RDFs muito mais simples. A base de dados da DBpedia usa *dbr:* e *dbo:* como abreviações de namespace para recursos e objetos, respectivamente. A tripla Brasil tem-capital pode ser escrita em uma forma de tupla assim:

dbr:Brasil dbo:capital dbr:Brasilia

Valores literais são usados para exibir rótulos ou valores de um estado. Rótulos para recursos são comumente expressados usando o predicado *rdfs:label* do namespace *rdfs:*. Por exemplo o rótulo para o recurso Brasil é o valor literal “*Brasil*”. Isto é expressado com o literal entre aspas duplas, conforme demonstrado na tripla:

dbr:Brasil rdfs:label “Brasil”

Na prática, é comum nas URIs de sujeitos ter uma tripla com um predicado *rdfs:label* e esta tripla terá necessariamente o valor referente a uma string literal como objeto (IANELLA et al., 2013; KATAYAMA et al., 2013). Esta string será representada na interface do usuário como uma representação escrita da entidade. O exemplo a seguir, mais completo, demonstra a representação para um browser suportar a WS e mostrar a afirmação de Brasil tem-capital Brasília como a string “*Brasil capital Brasília*”. Cada uma das URIs da primeira tripla torna-se o sujeito de uma nova tripla com o predicado *rdfs:label* e o literal que rotula a respectiva URI como objeto.

dbr:Brasil dbo:capital dbr:Brasilia

dbr:Brasil rdfs:label “Brasil”

dbr:capital rdfs:label “capital”

dbr:Brasilia rdfs:label “Brasilia”

Quando as triplas tem a mesma URI para o sujeito, então estas triplas tratam da mesma coisa. Triplas com a mesma URI para o predicado descrevem a mesma coisa sobre o sujeito. Triplas com URIs compartilhadas do mesmo objeto têm a mesma coisa em comum.

Não há restrições que impeçam uma tripla de ter em comum com outras triplas diferentes combinações de seus membros. É comum duas triplas terem o mesmo sujeito e predicado quando se quer demonstrar que são membros de uma lista desordenada ou então para especificar alternativas. Nos exemplos a seguir, tanto Fernando Henrique quanto Luís Inácio foram líderes do Brasil além de especificar que Brasil é chamado Brasil em Português e Brazil em inglês.

dbr:Brasil dbo:leader dbr: Fernando Henrique

dbr:Brasil dbo:leader dbr: Luís Inácio

e

dbr:Brasil dbo:label “Brasil”@pt-br

dbr:Brasil dbo:label “Brazil”@en

Uma coleção de triplas expressa um grafo RDF. Um conjunto de nós em um grafo RDF são todos os sujeitos e objetos em uma coleção de triplas. O conjunto de arcos em um grafo RDF são todos os predicados contidos nas triplas que representam o grafo. Um grafo

RDF é serializado em triplas, e estas são transmitidas, tipicamente usando HTTP/s, através da web em um documento RDF.

Se um documento RDF é focado em apenas um único sujeito, então tal documento RDF pode não incluir a informação de *rdfs:label* para todos os predicados e objetos que estão contidos no documento RDF. Por exemplo, o documento RDF retornado para Brasília na DBpedia inclui uma tripla *rdfs:label* para Brasília mas não *rdfs:labels* para nenhuma das outras entidades referenciadas no mesmo documento RDF. Sem realizar o cache, encontrar os *rdfs:label* para cada entidade requer outra transação HTTP/S para retornar o documento RDF para cada entidade. Comparado a uma ao acesso a uma máquina ou rede local, transações HTTP/S através da Internet são relativamente mais lentas –eventualmente lentas demais para uma interface de usuário. Dado que buscar na Internet por um único item é comparativamente lento e há vários documentos RDF para serem retornados, resolver os *rdfs:labels* em documentos RDF é temporalmente custoso. Já que a consulta de *rdfs:labels* é uma operação custosa e comum, alguns browsers para navegação na WS fazem o cache de dados de *rdfs:label*.

3.2.6. Linked data

A expressão *Linked Data* (LD) refere-se a um conjunto de boas práticas e uma coleção de tecnologias que compõe a WS, como RDF, OWL, SPARQL, dentre outros, que permite aos usuários publicar, compartilhar e conectar dados na www ou em intranets (BERNERS-LEE, 2006). Um dos aspectos principais da LD é que ao contrário de outros métodos de armazenamento de dados, o relacionamento entre os dados são tão importantes quanto os dados em si tornando-se a cola que permite que bases de dados distintas sejam conectadas (BIZER et al., 2009; TILAHUN, 2012; ISOTANI; BITTENCOURT, 2015; SPERONI et al., 2017). As principais recomendações para LD, definidas por Tim Berners-Lee:

- Usar URIs como nome para coisas.
- Usar HTTP URIs para que os usuários possam buscar estes nomes
- Quando alguém buscar por uma URI, prover informações úteis usando padrões
- Incluir links para outras URIs para que os usuários possam descobrir novas coisas

Fica claro que quando os dados são apresentados desta forma e tornam-se buscáveis, eles também podem ser publicados e ligados à qualquer outra fonte de dados na web, ou então mencionados em uma intranet corporativa, para que a informação seja compartilhada dentro

da organização (MACHADO et al., 2013; ISOTANI; BITENCOURT, 2015; PEINL, 2016; SPERONI et al., 2017). Outro benefício da LD é que quando os padrões, como por exemplo o RDF, são aplicados, os dados ganham a possibilidade de serem lidos por máquinas e serem explicitamente definidos (BERNERS-LEE, 2006; BEEK et al., 2014; ISOTANI; BITENCOURT, 2015; SPERONI et al., 2017). Esta auto-descrição dos dados, completamente separada da forma como é apresentada e formatada permite aos autores não serem restritos a um único vocabulário já que as conexões e os relacionamentos permitirão que a própria aplicação ache uma definição apropriada.

A capacidade de tornar os dados legíveis para as máquinas, o uso do protocolo HTTP como mecanismo de recuperação de informação, combinado com a flexibilidade e a facilidade de se gerar conexões em um grafo RDF permite que robos naveguem por grandes conjuntos de dados interligados e descubram novas fontes de dados e suas respectivas ligações (GIANNOPOULOU et al., 2014; ISOTANI; BITENCOURT, 2015).

O RDF foi desenvolvido como um meio de se descrever metadados sobre recursos através da representação de propriedades e relacionamentos que os conectam. É basicamente o modelo de dados da LD e teve sua serialização original baseada em XML. Os relacionamentos modelados por ele podem ser visualizados ou imaginados em termos de grafos, onde os nós pertencem aos recursos na web e as propriedades e relacionamentos são as arestas ou conexões.

Resource Description Framework Schema (RDFS) é um meio para definir um vocabulário usado dentro de um RDF através do uso de classes e propriedades *rdfs:type* (VARGAS-VERA, 2002; CORCHO; CIMA; GÓMEZ-LÓPEZ, 2006; MACHADO et al., 2013; TILAHUN et al., 2014; YU, 2014; ISOTANI; BITENCOURT, 2015). Uma classe RDF é um agrupamento de recursos RDF que são um tipo particular do *rdfs:class*. Classes ou recursos podem ser futuramente agrupados através da propriedade *rdfs:subClassOf*.

3.2.7. Armazenamento de RDF

Similar a um banco de dados relacional convencional, no sentido de armazenar dados para uso futuro, há o repositório RDF, também conhecido como repositório de triplas ou armazenamento de RDF. A maior diferença está no formato dos dados, que são armazenados e recuperados na forma de triplas, que podem ser combinadas e interconectadas de acordo com o padrão RDF (OBRST; JANSSEN; CEUSTERS, 2010; ALLEMANG; HENDLER, 2011; YU, 2014; SHARMA et al., 2017). Dessa forma, um repositório RDF também é considerado um banco de dados em forma de grafo, visto que cada registro armazenado representa um

conjunto de triplas RDF interligadas entre si de maneira lógica (ROBINSON et al, 2015).

Assim como banco de dados convencionais, é possível incluir um mecanismo de consulta de triplas RDF no repositório, como o *SPARQL Query Language* (SPARQL) através de *endpoints* (SERNADELA et al., 2017). Além disso, os repositórios RDF também possuem uma variedade de recursos de desempenho, que incluem o volume de dados a serem armazenados, a velocidade com que os dados podem ser acessados ou atualizados e a variedade de linguagens de consulta suportados pelos mecanismos de busca (ALLEMANG; HENDLER, 2011).

Existem diversas ferramentas para armazenamento e manipulação de triplas RDF, que, por meio de *Application Programming Interfaces* (APIs), funcionam como *middlewares* entre a

aplicação da WS e o repositório de triplas e podem atuar como uma extensão de banco de dados existentes, repositórios exclusivos (nativos) para dados RDF, como o OWLIM e AllegroGraph®, bem como repositórios híbridos que suportam armazenamento relacional e exclusivo, como o OpenLink Virtuoso (ISOTANI; BITTENCOURT, 2015; PEINL, 2016).

3.2.8. SPARQL Protocol and RDF Query Language

SPARQL é uma linguagem de consulta que recupera dados de um repositório RDF, recomendado pela W3C (ALLEMANG; HENDLER, 2011; LOPES; OLIVEIRA, 2011; ABANDA; TAH; KEIVANI, 2013; YU, 2014).

SPARQL pode ser usado para expressar consultas em uma ou mais fontes de dados, armazenados nativamente como RDF ou acessados como RDF via middleware. O SPARQL também contém recursos para consultar e retornar diferentes grafos RDF juntamente com suas conjunções, disjunções e restrições (PRUD'HOMMEAUX; SEABORNE, 2008).

SPARQL 1.1 foi desenvolvido pelo W3C SPARQL-WG (ALLEMANG; HENDLER, 2011; LOPES; OLIVEIRA, 2011; ABANDA; TAH; KEIVANI, 2013; YU, 2014). Esta nova versão é uma extensão do SPARQL 1 para adicionar elementos que não podiam ser incluídos anteriormente (DUMONTIER et al., 2010; OBRST; JANSSEN; CEUSTERS, 2010; TILAHUN, 2012; KATAYAMA et al., 2013; JAFARPOUR et al., 2016). Isto estende o documento de consulta SPARQL 1.0 com subconsultas e operadores de agregação e adiciona vários novos documentos com novas funcionalidades: Atualização SPARQL 1.1, Regimes de Entailment, Descrição de Serviço, Consulta federada, protocolo, caminhos de propriedade e formatos de resultado de consulta JSON, CSV e TSV:

SPARQL 1.1 Query: Os novos recursos incluem funções de agregação que permitem

operações como contagem, numérica mín. / máx. / média, operando sobre colunas de resultados. As subconsultas também são especificadas, fornecendo uma maneira de aninhar os resultados de uma consulta em outra consulta. Por exemplo, um o usuário pode identificar as 10 últimas postagens de blog criadas em um blog da web, com um único nome de autor para cada blog post. O recurso de negação, implicitamente presente no SPARQL 1.0 (por meio da filtragem de variáveis) também está presente na consulta com o filtro NOT EXISTS e o MINUS como palavras-chaves. NOT EXISTS trata do teste de negação, ou seja, se existe um padrão nos dados, dadas as ligações já determinadas pelo padrão de consulta. O filtro NOT EXISTS remove correspondências com base na avaliação de dois padrões. Expressões: projetar expressões representa a capacidade de consultas SPARQL SELECT para projetar qualquer expressão SPARQL, em vez de apenas variáveis. Uma expressão projetada pode ser uma variável, um URI constante, um literal constante ou um expressão arbitrária (incluindo chamadas de função) em variáveis e constantes.

SPARQL 1.1 Update: Esta extensão SPARQL fornece recursos para permitir a inserção de novas triplas RDF em um conjunto de dados RDF exposto por um terminal SPARQL. Isso fornece a capacidade de inserir triplas em um gráfico RDF, excluir triplas de um gráfico RDF, carregar um gráfico RDF, limpar um gráfico RDF, criando um novo gráfico RDF, tudo no triplestore especificado pelo endpoint. Também é possível retornar um grafo RDF do armazenamento de grafo e copiar, mover ou adicionar o conteúdo de um grafo RDF de um triplestore para outro. Também permite que um grupo de operações de atualização possam ser realizadas com uma única ação.

SPARQL 1.1 Protocol descreve o protocolo SPARQL, como um meio de transmitir consultas SPARQL e atualizações de clientes para agentes de consulta. O protocolo SPARQL é descrito de duas maneiras: primeiro, como uma interface abstrata independente de qualquer realização, implementação ou ligação concreta a outro protocolo; segundo, como uma ligação HTTP dessa interface.

SPARQL 1.1 Service Description fornece um método para descobrir e um vocabulário para descrever Serviços SPARQL disponibilizados por meio do protocolo SPARQL 1.1 para RDF. Essas descrições fornecem um mecanismo pelo qual um cliente ou usuário final pode descobrir informações sobre o serviço SPARQ, com suporte e detalhes sobre o conjunto de dados disponível.

SPARQL 1.1 Entailment Regimes: fornece um regime básico de vinculação para SPARQL. Isto permite que terminais SPARQL pudessem ser capazes de suportar formas de vinculação, como RDFS. Usuários, ao consultar o endpoint remoto, obterão em seus

resultados retornados aqueles que também refletem as correspondências de RDFS.

SPARQL 1.1 triplestore em HTTP Protocol descreve o protocolo para atualizar e buscar um grafo RDF de conteúdo de um triplestore sobre HTTP no estilo REST.

SPARQL 1.1 Federated Query define a sintaxe e semântica de SPARQL 1.1 Federated Query para executar consultas distribuídas em diferentes terminais SPARQL. O serviço palavra-chave estende SPARQL 1.1 para suportar consultas que mesclam dados distribuídos pela web.

SPARQL 1.1 Property Paths: os caminhos de propriedade correspondentes às consultas SPARQL são definidos. Um caminho de propriedade é uma rota possível através de um grafo entre dois nós. Um caso trivial é um caminho de propriedade de comprimento, que é um padrão de tripla.

3.2.9. Linguagem D2RQ

A linguagem D2RQ faz o mapeamento de uma base de dados relacional para vocabulários RDF e ontologias. O arquivo de mapeamento D2RQ permite a conversão de uma base de dados relacional em um grafo RDF.

Termos importantes a saber:

ClassMap: D2RQ ClassMap expressa uma classe uma coleção de classes idênticas de um vocabulário RDF ou ontologia OWL.

PropertyBridges: define as propriedades do recurso ClassMap RDF, associando uma coluna da base de dados com uma propriedade RDF.

Database: no D2RQ, Database expressa a conexão JDBC para bases de dados nativas e remotas.

dataStorage: no D2RQ, dataStorage faz referencia a uma Database.

Class: class significa RDFS ou classe OWL.

uriPattern: D2RQ usa uriPattern para reconhecer instancias do ClassMap.

classDefinitionLabel: define uma classe de rótulos. Age como um rdfs:label para uma classe associada.

belongsToClassMap: mostra que uma propriedade ponte refere-se a um ClassMap. É requisito para todas as PropertyBridges.

property: no D2RQ, property indica a propriedade do RDF que associa o ClassMap com o objeto ou literal.

pattern: especifica propriedades com um valor literal.

propertyDefinitionLabel: representa um rótulo como rdfs:label para a definição da

propriedade correspondente.

column: especifica a coluna na base de dados que armazena valores literais.

datatype: define os tipos de dados literais do RDF.

3.2.10. Aplicações da WS na área de saúde e medicina

O desenvolvimento da infraestrutura computacional e dos sistemas de informação em saúde geram grandes quantidades de dados em saúde. Junto com o crescimento exponencial dos dados, revelam-se também outros desafios de como gerir e integrar estes dados. Talvez o maior desafio neste sentido seja o da interoperabilidade entre os sistemas de informação em saúde. Isto se deve ao fato de ser um cenário heterogêneo, onde há diversas fontes de informação, diferentes estruturas e diferentes formatos. É complexo o trabalho de recuperar os dados, organizá-los e permitir uma navegação simplificada e amigável ao usuário, quando as fontes são as mais diversas possíveis. O tratamento destes dados, extração de indicadores e estatísticas tem potencial para contribuir no âmbito administrativo da saúde, auxiliando a tomada de decisão para definir políticas e melhores práticas. Como consequência, tem-se um melhor resultado na otimização da distribuição de recursos humanos e financeiros para atender especificamente as necessidades encontradas. (FREITAS; SCHULZ, 2009; LYTRAS; SAKKOPOULOS; PABLOS, 2009; DUMONTIER, 2010; LOPES; OLIVEIRA, 2011; KATAYAMA et al., 2013; YU, 2014; ZENUNI et al., 2015).

Na área da saúde, as ontologias têm recebido muita atenção de diferentes áreas, da representação do conhecimento às aplicações preditivas. Essas áreas incluem, mas não se limitam a, conhecimento representação, registros eletrônicos de saúde (EHR), padronização, diretrizes clínicas e suporte à decisão, terminologias padrão, análise qualitativa e preditiva.

Um exemplo do sucesso da ontologia em representações de conhecimento foi relatado no artigo “*Even Oriented Representation for Collaborative Activities (EORCA)*” ao implementar um método que abordava a observação e a representação de uma coleção atividades durante o manejo dos pacientes que poderiam ser reutilizadas pelos membros da equipe a fim de se prepararem para o credenciamento (L PELLEGRIN ET AL., 2007). A ontologia era usado para construir uma representação de conhecimento de diretrizes padrões para observação de tarefas pela equipe da UTI. No estudo na área de cardiologia, a ontologia baseada em UMLS apoiou os procedimentos de cardiologia para suporte cognitivo na tomada de decisão médica permitindo que diferentes partes interessadas e grupos de saúde compartilhassem a gestão do conhecimento e comunicação (BIOLCHINI, 2002). A ontologia também permitiu que o conhecimento do domínio fosse independente da tecnologia, para que

pudesse ser reutilizada de forma eficaz em várias plataformas. Um exemplo está no artigo “Sistema de Apoio à Decisão Clínica para Ponto de Assistência” utilizando um modelo de ontologia e regras de inferência em aprendizado de máquina, provando que quando o domínio de conhecimento foi alterado, apenas o resultado da inferência foi alterado sem qualquer trabalho no sistema de *software* (FARION ET AL., 2009). Além disso, Doyle e sua equipe implementaram uma base de conhecimento ontológico para vigilância em saúde pública que poderia ser referenciada por várias aplicações. Devido à sua independência tecnológica, o estudo concluiu que essa ontologia foi benéfica para a troca de informações em EHRs (DOYLE, MA, GROSECLOSE, & HOPKINS, 2005).

Formalizar um conceito é a primeira e mais importante etapa para estabelecer uma base de conhecimento para qualquer tipo de software na área da saúde. Uma das principais lutas dos EHRs é a falta de padronização. Sem padronização, o conceito de compartilhamento de dados entre os EHRs não seriam viáveis. O objetivo da aplicação da Ontologia em EHRs é fornecer uma representação sistemática de vários conhecimentos médicos usados para diferentes tipos de raciocínio nas atividades de saúde (A. JOVIC, M. PRCELA, & D. GAMBERGER, 2007). Por exemplo, um sistema usando Ontologia para aprimorar a funcionalidade semântica dos EHRs é apresentado no XOntoRank, a ferramenta para pesquisa ontológica de EHRs usada para resolver o problema de extração de dados com reconhecimento de ontologia para banco de dados EHR baseado em XML (FARFAN, HRISTIDIS, RANGANATHAN, BURKE, 2008). Com definições ontológicas na Nomenclatura Sistematizada de Medicina (SNOMED), o sistema poderia executar uma pesquisa semântica nos documentos XML. Não havia necessidade de uma perfeita correspondência entre palavras-chave em uma consulta e palavras em documentos devido aos termos semânticos mapeados na ontologia. Primeiro, o autor listou as palavras-chaves potencialmente problemáticas. Em seguida, ele desenvolveu os conceitos semânticos para essas palavras-chaves. Então, o grau de associação entre conceitos ontológicos e palavras-chave foram atribuídos. Por último, o estudo calculou e avaliou os algoritmos para responder a consultas de palavras-chave com reconhecimento de ontologia em EHRs. Como um resultado, o mecanismo de pesquisa XOntoRank encontrou mais correspondências. Ele encontrou a correspondência exata entre palavras-chaves e os conceitos correspondentes com uma melhor taxa de precisão e taxa de recall quando comparado ao algoritmo controle. O estudo concluiu que usar mapeamentos conceituais pode ser mais eficiente do que o mapeamento terminológico.

Peleg também desenvolveu outra ferramenta de mapeamento baseada em ontologia em

EHRs apresentada no ‘Mapeamento Computacional de Diretrizes clínicas para registros médicos eletrônicos: Mapeador ontológico de dados de conhecimento (KDOM)’ artigo (PELEG M., 2008). O KDOM era uma estrutura de mapeamento de ontologia que conectava as diretrizes interpretáveis por computador aos EHRs para criar compartilhamento e reutilização entre várias instituições. Com a ontologia, foi encontrada a solução para um problema incompatibilidade, separando o conhecimento do domínio médico do conhecimento operacional, tornando explícitas as suposições de domínio e desenvolvendo uma base para a ontologia. As classes foram divididas por propriedades em dois grupos: propriedades conceituais descrevendo o conhecimento abstrato e propriedades de campos referentes às tabelas do EHR. Os slots na ontologia de mapeamento foram usados para se referir aos campos de destino no bancos de dados e para especificar como o valor recuperado foi correspondido com um valor constante. Como resultado, a estrutura da ontologia mapeou com sucesso as diretrizes para o EHR e preencheu a lacuna entre as abstrações em um CIG e EHR, que poderia ser configurada posteriormente independentemente.

Um dos usos mais conhecidos da ontologia na área da saúde é para sistemas de terminologia, como UMLS (*Unified Medical Language System*), SNOMED, e LOINC (PISANELLI, GANGEMI, & STEVE, 1998). Esses sistemas oferecem padrões de comunicação, documentação e classificação de vocabulários médicos / de saúde (Cole, 2004). No entanto, mesmo esses sistemas de terminologia são todos baseados em um *framework* estruturado; os conceitos entre esses sistemas não são totalmente consistentes nem compatíveis (B. BOLBEL, 2006). Além disso, a ontologia tem desempenhado um papel essencial na vinculação do conhecimento de estudo de um domínio a um sistema de terminologia padrão. Por exemplo, em o "Facilitando a representação das diretrizes de avaliação pré-operatória usando SNOMED Artigo CT ", os autores investigaram se o SNOMED CT cobriria os termos usados no pré-operatório assim como o artigo de diretrizes de avaliação (L. AHMADIAN, 2009). Eles descobriram que 71% dos diretrizes foram mapeadas com SNOMED, enquanto 69% de 39 não foram totalmente cobertos ou violaram pelo menos um dos formatos SNOMED CT. Os autores propuseram que ontologia pudesse ser uma solução em potencial para formalizar as diretrizes no SNOMED CT. Outra ilustração é o “Relatório de Surto de Doenças em Direção à Filtragem Baseada em Funções”, que usa ontologia para analisar classificações conceituais de doenças infecciosas que não foram apresentadas em nenhum sistema de terminologia (S. DOAN, 2009). Também no “CardioOP data Class”, o estudo relatou que nenhum dos sistemas terminológicos; ou seja, ICD 10, SNOMED, UMLS nem MESH forneceram granularidade suficiente de conteúdo ou

completude de domínio para metadados em dados de multimídia no domínio Cardio (FRIED, KLAS, & WESTERMANN, 2003). No estudo de Elkin, a Ontologia foi usada para construir a terminologia estruturada para sistemas automatizados fornecendo classificação em notas clínicas (ELKIN ET AL., 2005). Assim, esses domínios de terminologia existentes podem integrar definições e terminologias existentes em diferentes níveis de saúde (D. PAPPA, 2006), conforme apresentado no “Na Estrutura de Categorias para um Sistema Terminológico em Sistema tradicional na medicina tradicional, sintomas, sinais e sua combinação Artigo de padrões ”usando ontologia para desenvolver uma terminologia padrão para diferentes áreas e países (PARK ET AL., 2009). Outro exemplo de uso de ontologia no mapeamento terminologias em todos os idiomas foram apresentadas no “Associação da comunicação com pacientes indígenas - Comunicação cultural amigável com diretrizes de cuidados clínicos para diabéticos tipo 2 Mellitus” artigo propondo um sistema padrão para as instalações locais, a fim de reduzir quaisquer preconceitos de equipes médicas locais, possivelmente influenciados por culturas e ambientes (FORBES, SIDHU E SINGH, 2011). A ontologia foi uma solução que oferece um modelo de mapeamento entre as diretrizes locais de taxonomia clínica australiana e a versão de inglês aborígine, resolvendo as questões de ambigüidade individual e má interpretação encontradas em literatura educacional aborígine com diretrizes para orientação baseada em pesquisa. Também, Nardarni e a equipe apresentaram em seu artigo “*Migrating Existing Clinical Content from ICD- 9 para SNOMED* ” mostrando como a ontologia poderia ser usada para mapear os conceitos entre dois sistemas diferentes (PRAKASH M NADKARNI, 2010). Da mesma forma, Saunders usou ontologia para construir o vocabulário para o *Anatomical Therapeutic Chemical* e Termo Crítico para Medicamentos Australianos e seus dados de segurança, aprimorando os métodos de regra de associação para explorar mais dados de doenças raras. No outra amostra, no estudo em “Critérios de Adequação ACR: Tradução para a Prática e pesquisa ”, o estudo foi desenhado para converter o conteúdo atual baseado em texto em uma base de dados relacional (SISTROM, 2005). Com a implementação da ontologia, o estudo representou com sucesso uma estrutura formal de um banco de dados relacional que armazena o versão principal das diretrizes, o que permitiu que os critérios fossem distribuídos rapidamente.

Diante desse cenário, a WS vem sendo utilizada na tentativa de superar tais desafios. Dessa forma encontram-se diversas iniciativas, pesquisas, projetos e grupos na área.

VALLE et al. (2005) em uma iniciativa semelhante a (LOPES, 2015), defende a adoção do paradigma da web semantica com tecnologias amplamente utilizadas no mercado, como Extensive Markup Language e web services, para alcançar a interoperabilidade dos

sistemas de informação e registro em saúde. Essa abordagem é defendida pelos autores, pois promove a combinação de interoperabilidade sintática e semântica entre aplicativos.

Hitzler e Janowicz (2010) enfatizam o crescente uso da WS. Tal crescimento é possível graças à aplicação do paradigma da WS não estar vinculado a um tipo específico de conhecimento ou área. A WS vem apoiar ambientes onde interdisciplinaridade e heterogeneidade estão implícitas no cotidiano. Nesse sentido, justifica-se o aumento no número de convenções e pessoas interessadas no assunto, sejam elas provenientes do meio acadêmico ou comercial.

Ogundele et al., (2016) desenvolveram uma ontologia para representar, consolidar e estruturar o conhecimento sobre os fatores que influenciam o comportamento de adesão ao tratamento em pacientes com TB. O repositório criado por eles poderia ser usado para encontrar alguns potenciais fatores que afetam a adesão ao tratamento da TB em comunidades semelhantes que foram utilizados no estudo, gerando seus índices de risco e também ajudando no acompanhamento dos pacientes e seus acompanhamentos.

Gonçalves et al. (2011) afirma em seu trabalho que a heterogeneidade entre conceitualizações deve ser resolvida antes de lidar com a heterogeneidade sintática a nível do termo, de modo que a interoperabilidade semântica possa ser conduzida de maneira efetiva. Ele fez uma prova ao conduzir um experimento que forneceu evidências de que a ontologia do eletrocardiograma (ECG) pode ser efetivamente usada para dar suporte ao projeto de outras versões interoperáveis que se referem ao ECG, como o HL7 aECG. Ele também afirma, por seus resultados, que sua metodologia também pode ser aplicada em outros domínios.

Kumar (2014) discute seu trabalho sobre alguns vocabulários, assim como eles podem ser usados para alcançar a interoperabilidade entre os aplicativos. No entanto, também discute as questões relacionadas à privacidade e interoperabilidade de aplicativos que usam a WS.

ZENUNI et al. (2015) revisaram todos os vocabulários de referência publicados que estão de acordo com os princípios e práticas da WS e concluíram que o mapeamento de formatos proprietários em ontologias é uma tarefa complexa e intensa. Eles também mencionam a manutenção de ontologias como um dos pontos delicados e enfatiza a importância de criar uma prova de conceito para aumentar a confiança neste tipo de abordagem para alcançar a interoperabilidade.

McMurray et al., (2015) desenvolveu um modelo conceitual de intercâmbio eletrônico clínico regional entre saúde. A ontologia permitiu a visualização do modelo e também instâncias pelo modelo computacional e foi utilizada para validar um subconjunto dos dados coletados utilizando um banco de dados diferente, embora ainda permaneça um banco de dados

relacionado com os interesses da pesquisa.

Jiang et al (2013) descrevem em seu trabalho o desenvolvimento de uma ferramenta para solidificar as definições recomendadas pela Classificação Internacional de Doenças na versão 11. A classificação gerada pelo seu trabalho através de um modelo de governança que envolve consenso de especialistas, validação e suporte colaborativo e distribuído permite que esses parâmetros sejam otimizados. Essas classificações são então comparadas com outros valores encontrados no Sistema Unificado de Linguagem Médica e na Nomenclatura Sistematizada de Medicina - Termos Clínicos. Por fim, a compilação dos resultados é submetida a especialistas para avaliação e atribuição efetiva do grau de usabilidade.

Sharma et al. (2017) desenvolveram e avaliaram um sistema de gerenciamento de metadados do estudo do genoma do câncer. Este sistema baseado em ES serviria como base e ponto crítico para apoiar a modelagem de informação.

O aumento significativo no uso da WS como solução na área da saúde reforça uma tendência de consolidação em diversos cenários práticos. Neste sentido diversas formas de aplicação contribuem para um amplo espectro de validação e comprometimento na entrega de valor semântico aos usuários dos sistemas.

3.3. ONTOLOGIAS

O termo ontologia origina-se da filosofia que o estabelece como uma descrição sistemática do ser, descrevendo a natureza de objetos. Na Ciência da Computação, ontologias são usadas para conceitualizar domínios de conhecimento. Algumas definições para a ontologia são propostas nas áreas de sistemas de informação e conhecimento, seguem:

a) Ontologia é uma especificação explícita de uma conceitualização compartilhada (GRUBER, 1993).

b) Ontologia é uma especificação (representação) formal (explícita) de uma conceitualização compartilhada por uma comunidade de pessoas acerca de um domínio em particular. (STUDER, 1998)

c) Uma ontologia é um modelo de um domínio em particular, construído com um propósito em particular (ANTONIOU, 2008)

d) Ontologia é uma especificação formal de uma conceitualização compartilhada. Por (BORST,1997) e como citado em (GHAWI, 2011)

e) Ontologia é uma tupla $O = (C, R, I, T)$ onde C é um conjunto finito de símbolos conceituais, R é um conjunto finito de símbolos relacionais, I é um conjunto finito de símbolos de instâncias, e T é um conjunto de teorias e regras que satisfazem as relações dentro

de um domínio. (LIE, et all. 2006) Ontologia é a especificação de uma conceitualização que é criada para ser reusada entre múltiplas aplicações e implementações. Uma especificação de uma conceitualização é a descrição formal e escrita de um conjunto de conceitos e relações dentro de um domínio de interesse. (OWG-MGED3, n.d.)

f) Ontologia é um modelo de dados que representa um conjunto de conceitos dentro de um domínio e as relações entre estes conceitos é usada para inferências sobre os objetos daquele domínio. (NUNO, 2003)

Os diferentes aspectos de uma ontologia destacados por estas definições são:

Conceitualização significa que a ontologia representa um conceito de modelo que é derivado de uma abstração de algum conceito relacionado ao fenômeno no mundo real.

Explícita significa que ontologias devem ser definidas precisamente e de forma não ambígua. Ontologias expressam um conceito precisamente de forma que possa ser interpretada e usada por sistemas computacionais (máquina).

Formal refere-se ao fato das ontologias serem reconhecidas e poderem ser interpretadas, não necessariamente entendidas, e usadas por máquinas. Em tempo, máquinas podem fazer inferências sobre os conhecimentos contidos em uma ontologia para deduzir novos conhecimentos.

Inferência refere-se à capacidade de abstrair significados não claramente explícitos chegando a novas conclusões. Por exemplo: podemos definir o significado ou conceito semântico de relacionamento “É um ancestral de” como sendo uma relação transitiva dentro da ontologia. Portanto, se uma ontologia contém o fato de que X é ancestral de Y e Y é ancestral de Z, portanto é possível que máquinas concluam que X é ancestral de Z.

Compartilhada refere-se ao fato de ontologias serem desenvolvidas em um consenso. O conhecimento contido na ontologia é coletivo e acordado entre os usuários.

Reuso trata-se do fato de ser viável operacionalmente. Usuários podem reutilizar o vocabulário de uma ontologia já existente para descreverem seus próprios domínios ao invés de criarem suas próprias.

Se todos estes aspectos forem verificados ao se desenvolver uma ontologia, ela então será viável como solução para problemas de interoperabilidade entre fontes de dados heterogêneas e contemplarão completamente tudo que está previsto na WS.

Em termos gerais, ontologia é um consenso para um vocabulário, criado com a ajuda de elementos estruturantes bem estabelecidos a fim de construir conhecimento de alto nível, para especificar o significado semântico de sistemas de terminologias de uma maneira definida e não ambígua. Tais elementos estruturantes são utilizados com o intuito de construir

conhecimento estruturado e formalizado (SMITH, et al., 2004)

Para um domínio em particular, a ontologia representa uma linguagem rica que permite um controle sobre os tipos de recursos e suas propriedades. Comparada às taxonomias, ontologias melhoram o significado semântico ao prover relacionamentos ricos entre os termos do vocabulário (SMITH, et al., 2004).

Um ponto chave para as ontologias é que através de terminologias formalmente especificadas, com significados extraídos direto do mundo real e frutos de um consenso, entrelaçam o entendimento humano e das máquinas. Isto facilita o reuso e compartilhamento de ontologias entre humanos e também entre máquinas. A representação explícita do significado semântico sobre informações, programas e outros recursos disponíveis permite que a web baseada em conhecimento seja apresentada com uma melhor qualidade das informações, servindo melhor seus usuários.

A próxima sessão descreve os componentes de uma ontologia: instâncias, conceitos (classes), propriedades e axiomas.

3.3.1. Componentes de uma Ontologia

Instâncias (Objetos) também conhecidas por Indivíduos ou Objetos, são os componentes básicos de uma ontologia. É tudo que existe e pode ser diferenciado dos outros, dizendo-se algo sobre. Pode incluir objetos concretos como pessoal, livros, mesas e também indivíduos abstratos como doença ou especialidade médica. Por exemplo, um livro específico “Introdução à WS” é um objeto.

Conceitos (Classes) são grupos abstratos, conjuntos ou coleções de objetos. Eles podem conter indivíduos, outras classes ou a combinação de ambos. Um conceito é um conjunto de indivíduos ou objetos dentro de um domínio. Conceitos são tipicamente organizados em taxonomias (hierarquia de classes) usando-se relações de generalização ou especialização para implementar o conceito de herança.

Se C e C' são dois conceitos tal que se algum objeto $x \in C$ implica também que $x \in C'$, portanto C é um subconceito de C' ou C' é um superconceito de C .

Propriedade do objeto também é conhecido como relações ou relacionamentos dado que ela liga um conceito a outro, representando um tipo de associação entre ambos. O domínio e *range* são ambos propriedades do conceito. Pode-se denotar por $R(C1,C2)$, que estabelece para a instância de relação $r \in R$ que o domínio é $C1$ e *range* é $C2$. Isto é, r liga dois conceitos, $c1 \in C1$ e $c2 \in C2$.

Propriedade do dado também é conhecida por atributos. Atributos são características e

parâmetros que objetos podem possuir e compartilhar. Objetos em uma ontologia podem ser descritos ao designar atributos à eles. Atributos geralmente são diferenciados de relações pois seu *range* é um tipo de dado, como por exemplo uma *string* ou um inteiro, enquanto o range de uma relação é um conceito.

Axiomas são afirmações que de uma maneira lógica, juntas, formam a teoria geral sob o domínio em consideração. Axiomas são usados para moderar as informações contidas numa ontologia, para garantir a consistência e integridade ao conhecimento representado nela, bem como para deduzir novas informações.

3.3.2. Basic Formal Ontology (BFO)

BFO é classificada como uma *upper ontology*, isto é, uma ontologia que é domínio independente e projetada para servir como um *framework* para formalizar qualquer domínio. BFO é guiada pelo realismo, a proposição filosófica na qual “a realidade e seus constituintes existem independentemente de nossas representações (linguísticas, conceituais, teóricas e culturais) deles” (GRENON, 2003c). A BFO serviu como base para algumas ontologias de domínio colaborativas, como por exemplo: *Open Biomedical Ontologies (OBO) Foundry* (SMITH et al., 2004) e também a *Phenotypic Quality Ontology (PATO)*.

Além do realismo, a arquitetura da BFO pressupõe também perspectivalismo, falibilismo e adequatismo. A proposição de perspectivalismo admite que existem múltiplas representações da realidade que são igualmente boas e podem “capturar diferentes e importantes aspectos de uma mesma realidade” (SPEAR, 2007). Isto inclui a visão de se separar elementos do mundo real em diferentes níveis de granularidade. Falibilismo concede a possibilidade de que teorias científicas podem, na verdade, estarem incorretas, dado que a realidade “existe independentemente de nossa maneira de conceitualizá-la” (SPEAR, 2007). Finalmente, o adequatismo é o princípio que uma ontologia que represente um domínio deve permanecer fiel aos elementos fundamentais do domínio, ao invés de se limitar somente aos aspectos básicos da realidade (SPEAR, 2007). Por exemplo, uma ontologia de um fenômeno biológico não deve somente incluir componentes como células e órgãos, mas também populações (SPEAR, 2007). Uma ontologia deve descrever o domínio em um nível de granularidade suficientemente e minimamente necessária para que seja compreensível.

BFO divide suas entidades em continuantes e ocorrentes. O que diferencia ambos os tipos de entidades são suas relações com o tempo. Continuantes são entidades que são possuem uma linearidade e completude em relação ao tempo e não possuem partes temporais. Isto é resultado da implicação direta da noção de que continuantes estão completamente

presentes em qualquer instante do tempo no qual existem. Por exemplo, uma célula é um continuante. Ocorrentes são entidades que estão ligadas à fatores temporais ou “desdobramentos”. Isto é resultado da implicação direta da noção de que ocorrentes tem o tempo como parte deles. Por exemplo, o processo de divisão celular é um ocorrente. Regiões temporais são ocorrentes cujas partes consistem somente de espaços de tempo.

Na BFO, continuantes têm dois subtipos: continuantes independentes e continuantes dependentes. Continuantes dependentes, como as qualidades, são inerentes à continuantes independentes. Isto significa que eles existem somente enquanto características de um continuante independente específico. Por exemplo, a forma de uma célula é inerente à uma célula específica.

Continuante é também um subtipo de Região Espacial. Uma região espacial é alguma parte do espaço-tempo que pode não existir a qualquer momento ou que não pode existir a partir de um instante de tempo (BITTNER; DONNELLY, 2004). BFO considera o espaço-tempo como uma entidade absoluta de quatro dimensões. Um continuante dependente ou independente ocupa ou está localizado em uma região espacial por algum tempo.

Na BFO, ocorrente podem ter os subtipos: Entidade Processual e Região Espaço-Temporal. Para entidades processuais o foco geralmente se dá no processo, por exemplo a evolução de um quadro de doença ou a vida de uma abelha. Uma região espaço-temporal é uma parte do espaço-tempo. Uma entidade processual ocupa uma região do espaço-tempo.

Ocorrentes também possuem o subtipo região temporal. Regiões temporais são consideradas como um “tipo neutro de tempo” (SPEAR, 2007) e indexa ambas entidades continuantes que existam em regiões espaciais e entidades processuais que existam em regiões espaço-temporais. As expressões “tipo neutro de tempo” e “indexa” são termos admitidamente vagos e não claros bem como a natureza do tempo não é esclarecida completamente pela literatura da BFO.

3.3.3. OWL

A *Web Ontology Language* (OWL) é a linguagem recomendada pela W3C para representar e descrever formalmente ontologias (MCGUINNESS; HARMELEN, 2004; ARGÜELLO; EL-HASIA; LEES, 2006; TIMM; RENLY; FARKASH, 2011; YU, 2014; LIANG et al., 2015). Sua primeira versão foi desenvolvida como uma extensão e melhoria para a DAML+OIL e para os vocabulários RDF/RDFS, que são representados com sintaxe XML (VARGAS-VERA et al., 2002; ISOTANI; BITTENCOURT, 2015). Comparado aos seus predecessores, OWL assume alguns pré-requisitos como verdadeiros, como

independência de classes, cardinalidade, equivalência de classes, simetria e transitividade de propriedades, que ajudam a organizar e categorizar a informação em classes e indivíduos (MCGUINNESS; HARMELEN, 2004; LIMA; CARVALHO, 2005; VAN WOENSEL et al., 2014; ISOTANI; BITTENCOURT, 2015).

Uma típica ontologia OWL é estruturalmente composta pelas seguintes seções:

- O Cabeçalho declara as informações que identificam a ontologia, como o identificador da ontologia, a versão e outras anotações associadas a ontologia. Referências a outras ontologias (que permitem reuso) também são declaradas neste campo.
- Axiomas de classe define as classes, que provêm uma abstração para o agrupamento de recursos com características similares. Classes podem ser descritas nos termos de um identificador de classes, listagem exaustiva dos indivíduos, restrição de propriedades, intersecção, união e complemento de classes.
- Axiomas de propriedades definem as características das propriedades, através do uso de construtores padronizados tais como declarações de propriedade, domínio, intervalo, propriedade inversa, propriedade funcional, equivalência, simetria e transitividade.
- Afirmações de fatos sobre indivíduos declaram fatos acerca da identidade de um indivíduo (URI), fatos acerca de quais classes eles pertencem e os valores das propriedades de um indivíduo.

A linguagem OWL versão 1 provê três linguagens derivadas ou também chamadas de perfis, desenvolvidas para cenários especiais:

- OWL FULL é definido como a união da DAML+OIL com RDF. Tem grande poder expressivo e pode, inclusive, considerar classes como sendo indivíduos. No entanto, ontologias escritas neste perfil não tem garantia de que poderão ser usadas em processos computacionais.
- OWL-DL inclui todos os padrões de construção da OWL, com algumas restrições que garantem que possam ser usadas em processos computacionais. Esta sublinguagem é semanticamente equivalente a *Description Logic* (DL), contribuindo para que algoritmos possam tirar vantagem da formalização de propriedades pela DL a fim de suportar otimizações e verificações. Qualquer ontologia válida no formado OWL-DL também é uma ontologia válida OWL FULL.

- OWL Lite é um conjunto simplificado da OWL-DL que provê suporte instantâneo para a classificação em uma hierarquia de classes e para cardinalidade binária. É sugerida como um caminho para a migração a partir de dicionários taxonômicos.

A OWL versão 2 é uma extensão e revisão da OWL que adiciona novos recursos que facilitam a engenharia dos conhecimentos (LIMA; CARVALHO, 2005; KATAYAMA et al., 2013; YU, 2014; ISOTANI; BITTENCOURT, 2015).. Há a retro compatibilidade com a OWL 1, embora as ontologias escritas em OWL 2, assim como na sua predecessora, provê os mesmos elementos estruturantes, como classes, propriedades, indivíduos, valores de dados e anotações. Os recursos na OWL 2 são armazenados como documentos web e identificados por uma IRI (internationalized resource identifier), que é uma extensão Unicode das tradicionais URLs. OWL 2 oferece as mesmas funcionalidades além de adicionar nova expressividade a OWL, tais como:

- Chaves (indivíduos podem ser unicamente identificados por uma propriedade);
- Cadeia de propriedades (composição de propriedades);
- Novos tipos e intervalos de dados;
- Restrições de cardinalidade qualificadas;
- Propriedades assimétricas, reflexivas e desmembradas;
- Capacidade de anotação melhorada;

Comparada a OWL 1, a OWL 2 é mais uma especificação estrutural e conceitual, de tal forma que não é mais limitada a somente uma sintaxe. A W3C, inclusive, sugere cinco tipos de sintaxes para as ontologias especificadas na OWL 2:

- RDF/XML representa uma ontologia como grafos RDF, similares a OWL 1. É a sintaxe primária e padrão para a troca de dados, necessariamente implementada por todas as ferramentas OWL 2, tornando-se uma das mais indicadas para interoperabilidade.
- OWL/XML representa uma serialização direta da OWL para XML, portanto torna o processo de serialização do XML natural (sendo facilmente processada por ferramentas prontas para XML).
- Sintaxe Funcional representa diretamente a especificação estrutural, portanto torna fácil a visualização estrutural da ontologia.
- Sintaxe de Manchester é gramaticalmente próxima da description logic, portanto é fácil de ser usada com a logica semântica da OWL.
- Turtle também realiza o mapeamento para grafos RDF e é uma forma

compacta de representar triplas.

Semelhante à OWL 1, a OWL 2 também tem sublinguagens e perfis com diferentes características e significados para se usar de acordo com necessidades específicas:

- Semântica baseada em OWL 2 RDF atribui significado diretamente aos grafos RDF, portanto é inteiramente compatível com a semântica RDF bem como é menos restrita que a versão OWL 2. “OWL 2 FULL” é o nome informal usado para se referir aos grafos RDF gerados para as ontologias OWL 2.
- Semântica Direta OWL 2 é a atribuição computacional padrão da OWL 2 que é compatível com o modelo teórico de descrição de lógica semântica.
- OWL 2 EL é o perfil da OWL 2 com semântica direta que permite algoritmos de complexidade polinomiais realizarem inferências de vários tipos. É recomendada em grandes ontologias onde a expressividade tem baixa prioridade comparada a performance computacional.
- OWL 2 QL é um perfil da OWL 2 com semântica direta que permite que uma intersecção de buscas sejam respondidas em um espaço logarítmico usando padrões tecnológicos de base de dados relacionais. É especialmente útil quando os dados necessitam ser acessados por queries em SQL, por exemplo.
- OWL RL é um perfil da OWL 2 com semântica direta que permite implementações de inferência de complexidade polinomial em base de dados baseadas em regras. É especialmente útil quando manipula-se diretamente triplas RDF.

O teorema da correspondência da semântica baseada em RDF e semântica direta afirma que as inferências tiradas com semântica direta ainda serão válidas se a ontologia for mapeada e interpretada com a semântica baseada em RDF. Além disso, qualquer ontologia OWL 2 EL, QL ou RL é, obviamente, também uma ontologia OWL 2 que pode ser interpretada usando a semântica direta (DL).

3.3.4. Mapeamentos de Ontologias

A WS apresenta novas oportunidades que permitem a modelagem, compartilhamento e extração de conhecimento de dados disponíveis na web. Estas possibilidades tornam-se possíveis através da representação formal de domínios do conhecimento através de ontologias. Atualmente, as ontologias são vastamente usadas na WS para prover significado à um domínio em particular. Estas ontologias de domínios representam conceitos de uma maneira específica e independente, e geralmente são incompatíveis. Enquanto o número de

sistemas que dependem de ontologias cresce, verifica-se a necessidade de fundir ontologias para uma representação mais generalizada.

Em vários domínios do conhecimento, há algumas sobreposições entre ontologias que diferem apenas no nível de abstração e no método de representação dos mesmos conceitos (ZAHRA, et al., 2007). Isto ocorre devido às diferentes percepções do domínio baseadas em ideologias, bagagem cultural e até mesmo no uso de diferentes linguagens de representação.

A web possui a característica intrínseca de ser distribuída, bem como as ontologias. A característica distribuída do desenvolvimento de ontologias levou à um grande número de diferentes ontologias cobrindo ou sobrepondo o mesmo domínio. Conforme mais e mais pessoas envolvem-se, várias ontologias distintas são criadas. A interoperabilidade entre as diferentes ontologias torna-se essencial para ganhar força dentro da WS.

A fim de que dois atores, neste caso máquinas ou pedaços de *software*, compreendam-se, é preciso que eles usem a mesma representação formal (ontologia) para a conceptualização compartilhada.

Infelizmente, não é fácil fazer com que todos os envolvidos concordem em usar uma mesma ontologia para um determinado domínio quando há diferentes ontologias para tal. É aí que os problemas aparecem e os atores não entendem-se. O uso de ontologias fundacionais, a fim de prover uma definição comum de termos fundamentais, pode mitigar este problema até certo ponto. No entanto, não existem tais ontologias fundamentais para a maioria dos domínios das diferentes áreas. Portanto a solução é prover um sistema de mapeamento e fusão de diferentes ontologias. Atualmente, o processo de mapeamento de ontologias é na sua maior parte manual e portanto consome grande tempo e esforço.

Comumente, mapeamento e fusão são atividades consideradas similares, mas são bem diferenciadas segundo:

Mapeamento de ontologias acha correspondências entre duas entidades que cobrem um mesmo domínio e que estão relacionadas semanticamente.

Fusão de ontologias é o processo onde duas ontologias são semanticamente relacionadas à nível conceitual e que as instâncias ontológicas são transformadas em uma nova ontologia alvo de acordo com as respectivas relações semânticas.

O processo genérico de mapeamento de ontologias pode ser definido como: dado ontologias A e B que abordam o mesmo domínio, mapear uma ontologia na outra significa que para cada conceito (nó) na ontologia A, encontra-se um conceito (nó) correspondente, que possui um significado igual ou semelhante, na ontologia B e vice-versa. Esta relação de correspondência pode ser de um para um, um para vários, vários para um ou vários para vários.

Algumas das relações genéricas que podem ligar conceitos da ontologia fonte para a ontologia alvo são: igualdade, disjunção, sub ou supra conceito e sobreposição. O resultado de um processo de mapeamento é um conjunto de regras de mapeamento e suas correspondências entre as entidades que pertencem às diferentes ontologias. Estas regras de mapeamento conectam conceitos na ontologia A a conceitos na Ontologia B, com o mesmo grau de confiança. Em adição às duas ontologias, o processo de mapeamento de ontologias usa alguns tipos de recursos externos e parâmetros que são definidos pelo usuário, podendo usar também mapeamentos previamente encontrados ou aceitos.

Dado duas ontologias, a correspondência, também conhecida como elemento de mapeamento por (SHVAIKO, 2005), é uma tupla $EM = (id, e1, e2, n, r)$ onde id é um identificador única para a correspondência; $e1$ e $e2$ são entidades (conceitos) da primeira e da segunda ontologia, respectivamente; n é a medida de confiança (geralmente estipulada entre 0 e 1) correspondente entre $e1$ e $e2$; e r é a relação estabelecida entre $e1$ e $e2$, que pode ser, por exemplo igual, geralmente maior, geralmente menor, disjunção ou sobreposição.

A correspondência $(id, e1, e2, n, r)$ assegura que a relação r estabelecida entre as entidades $e1$ e $e2$ das ontologias tenham a o grau de confiança n . Esta tarefa de mapeamento, geralmente é feita manualmente, e demanda um grande esforço dos especialistas dos respectivos domínios. Métodos automáticos ou semiautomáticos devem ser desenvolvidos para reduzir tal esforço. Isto torna o mapeamento de ontologias um grande desafio para atingir a interoperabilidade na construção da WS.

3.3.4.1. Demandas

O progresso nas tecnologias de informação e comunicação tem disponibilizado uma grande quantidade de dados. O número de diferentes fontes de informação têm crescido significativamente e, portanto, o problema de gerenciar a heterogeneidade entre eles também tem aumentado junto. Este problema também aplica-se às ontologias usadas na WS.

A WS, RDF e OWL não são por si só a resposta cega para que se atinja uma interoperabilidade transparente entre as partes. Algumas dificuldades precisam ser superadas para que tal nível de transparência seja atingido (NOY, 2005). Existem vários problemas que devem ser cuidados no processo de engenharia da ontologia de um determinado domínio, sendo um deles o mapeamento de ontologias (TJOA, et al., 2005).

O mapeamento de ontologias é um pré-requisito para prover a interoperabilidade entre agentes na WS. Quando for oportuno, agentes precisam compartilhar e reusar ontologias correta e consistentemente em detrimento da criação de suas próprias. Caso não seja possível, surge a

necessidade de se criar correspondências entre os termos das diferentes ontologias usadas. É preciso que eles alinhem suas ontologias para permitir que as informações sejam trocadas tanto à nível sintático quanto semântico.

3.3.4.2. Problemas e desafios

Falta de um framework padrão

Não existe um framework padrão que suporte as atividades de mapeamento de ontologias. A performance das ferramentas existentes precisam ser investigadas conforme (KOTIS, 2008). No entanto, algum esforço neste sentido têm sido feito por pesquisadores. Por exemplo, um meta modelo para o mapeamento de ontologias baseado em teoria de relações de conjuntos é descrito em (LIE, et. al., 2006) enquanto em (LIPING, et al., 2006) descreve-se a modelagem de um framework de mapeamento de ontologias.

Semiautomático versus automático

A maior parte das ferramentas de mapeamento de ontologias disponíveis atualmente são semiautomáticas (KOTIS, 2008). As técnicas atuais requerem um nível de intervenção e validação humana. No entanto, mesmo para ontologias com tamanho moderado, é um trabalho que consome muito esforço dos engenheiros. É esperado que as ferramentas de mapeamento de ontologias permitam que o usuário realize suas atividades em diferentes sessões divididas em diversos períodos. Um cenário melhor ainda seria que as ferramentas permitissem múltiplos usuários trabalharem em colaboração (ou paralelamente), unindo esforços de todos.

Precisão

O requisito para precisão e recuperação varia de aplicação para aplicação. Algumas aplicações requerem uma recuperação completa (aplicações de segurança) enquanto outras requerem uma precisão completa (prescrição de medicamentos). Eventualmente, em determinados cenários, é interessante abrir mão de ambas para que se ganhe velocidade. Bons exemplos destes cenários são interações de curta duração entre dispositivos mobile ou em redes P2P.

Complexidade computacional

A ferramenta que oferece precisão e recuperação ótima tende a ser computacionalmente tão complexa que não pode ser usada com ontologias usadas em aplicações do mundo real.

Velocidade

O processo de mapeamento de ontologia deve ser rápido o suficiente, mesmo que ao custo de outros parâmetros, em aplicações de tempo real que são dependentes do tempo (ERIG, 2004).

Aspectos culturais

Outro importante problema é integrar as noções culturais ao processo de mapeamento de ontologias. Pessoas diferentes pensam de forma diferente acerca de um determinado conceito, e isto está enraizado nas suas respectivas bagagens cultural e histórica (TJOA, et. al., 2005). Outros problemas similares que podem surgir é quando uma mesma palavra é usada para significados diferentes (homônimos) ou diferentes palavras são usadas para o mesmo significado (sinônimos), mesmo que entre pessoas de uma mesma cultura. Não obstante, um determinado conceito pode ser modelado em diferentes níveis de granularidade (SAMPSON, 2005).

Escalabilidade

A maioria das ferramentas de mapeamento de ontologias desenvolvidas atualmente lidam com ontologias leves e não são capazes de lidar com ontologias grandes e complexas, como por exemplo as ontologias de genes e médicas (KOTIS, 2008). O método de mapeamento de ontologias deve ser escalável.

Interface de Usuário

É muito complexo, se não impossível, automatizar completamente o processo de mapeamento de ontologias devido à algumas incompatibilidades sintáticas e semânticas. A maioria das ferramentas requerem intervenção manual para a confirmação ou rejeição do pareamentos encontrados pela ferramenta. Esta limitação impõe a necessidade de que exista uma interface de usuário amigável a fim de proporcionar uma interação amigável e de maneira produtiva com o aplicativo. Assume-se também que a ferramenta disponha de explicações acerca dos pareamentos encontrados para auxiliar os usuários na tomada de decisão. A boa visualização dos resultados é essencial.

Necessidade de experiência e conhecimento no domínio

Enquanto se executa o mapeamento de ontologias para um domínio específico, o uso de tesouros externos e a experiência prévia com o assunto mostram-se benéficas ao processo. No entanto, como apontando em (EUZENAT, et., al., 2007) e citado em (KOTIS, 2008), esta realidade não é aplicável a ontologias complexas.

Considerações sobre o contexto

Foi observado por alguns pesquisadores que a ferramenta de mapeamento de ontologias pode melhorar o processo de mapeamento caso o contexto seja conhecido deles. A linguagem usada por humanos é fundamentalmente ambígua e o significado dos termos e conceitos dependem do contexto no quais são utilizados. Este significado implícito deve ser levado em consideração pela ferramenta de mapeamento de ontologia e isto só é possível caso

o contexto da informação seja, de alguma forma, disponibilizado.

Confiança e segurança

A WS deve ser capaz de prover confiança e segurança em suas camadas superiores (TJOA, et. al., 2005) e, portanto, as ferramentas futuras de mapeamento de ontologias devem ser capazes de considerar tais aspectos ao parear diferentes ontologias. O processo de mapeamento de ontologias deve ser capaz de garantir segurança nas afirmações e mapeamentos gerados.

Consideração de tempo de vida

Processos de mapeamento de ontologias devem estar cientes sobre o ciclo de vida das ontologias antes de mapeá-las. Os resultados gerados de uma ontologia desatualizada não podem ser usados de maneira significativa (TJOA, et. al., 2005). Inclusive, para se produzir informações com significados consistentes, deve-se levar em consideração os fusos nas quais as ontologias foram geradas.

3.3.4.3. Aplicações do mapeamento de ontologias Integração de informações

Aplicações de integração de informações necessitam integrar informações de várias fontes seja para satisfazer temporariamente a busca de um usuário ou permanentemente para armazenar as informações integradas. Por exemplo, quando um usuário busca pelo título de algum livro em conjunto de algum outro critério como, por exemplo, o preço, é requisitado que seja integradas ontologias de todas as lojas virtuais. Seguem-se alguns exemplos de casos especiais e seus requisitos para integração de informação:

Integração de catálogos de produtos: por exemplo, em aplicações B2B, um site que provê um Market place junto de aplicações de comparação de preços entre os produtos de diferentes marcas e fabricantes, necessita integrar os catálogos de produtos de todos os fornecedores de uma maneira única a fim de atender as necessidades de busca dos usuários que estão navegando no site.

Integração de aplicativos federais: pareamento de processos necessários para a construção de Data Warehouse e para prover a interoperabilidade em bases de dados de sistemas federativos.

Integração de dados: Quando uma organização incorpora outra empresa, como parte do processo, surge a necessidade de fundir suas informações de negócio que estão armazenadas em diversos formatos em suas respectivas bases de dados.

Sistemas com múltiplos agentes

Em sistemas com múltiplos agentes, mecanismos de alto nível para troca de mensagens

são padronizados até um certo nível. No entanto, o real conteúdo das expressões variam de agente para agente ou, pelo menos, de grupo de agentes para grupo de agentes. Quando dois agentes realmente independentes e autônomos desejam comunicarem-se, é preciso que eles pareiem e alinhem suas ontologias antes de trocarem mensagens.

Sistemas peer-to-peer

Sistemas peer-to-peer referem-se a aplicações distribuída capazes de trocar informações sob demanda, satisfazendo as necessidades de compartilhamento de informações de um determinado grupo de usuários que estão interessados em um determinado tipo de informação, como por exemplo arquivos de áudio ou vídeo. Cada peer pode atuar como um servidor para prover informação ou pode atuar como cliente para consumir informação. Enquanto um peer em um determinado sistema é autônomo, eles podem ter de estabelecer um acordo prévio sobre os vocabulários pre-definidos para descrever as informações de interesse. Ou então, será necessário que haja um sistema intermediário que possa realizar o mapeamento dos metadados para que se possa atingir o propósito do compartilhamento de informação com significado.

Integração de serviços de WS

Web Services provêm comunicação entre aplicações. Para que se possa usar um web service, a aplicação que irá consumi-lo precisa localizá-lo dentre um repositório centralizado de regras referenciais e documentações para que possa atender suas necessidades. A nova era da WS suporta web services descritos em termos de ontologias. Estes serviços podem ser descobertos e integrados dinamicamente, conforme a demanda, para satisfazer as necessidades dos usuários. Portanto, tais aplicações necessitam do mapeamento de ontologias para buscar todos os web services para sua aplicação cliente que satisfaçam, além da interoperabilidade sintática e funcional, a interoperabilidade semântica.

4. MÉTODOS

Para apreciar o conteúdo deste estudo, fazemos uma descrição mais detalhada dos métodos utilizados onde aproveitamos para discutir de maneira sucinta os procedimentos que foram utilizados para a sua realização.

A pesquisa possui essencialmente três temas principais, como descrito detalhadamente na revisão: WS, ontologias e interoperabilidade de sistemas de informação em saúde.

Trata-se de uma pesquisa aplicada, com contribuições práticas envolvendo a aplicação de seus resultados diretamente na solução de problemas encontrados em cenários reais de recuperação de dados de TB com significado relevante aos gestores e profissionais de saúde através do uso de ferramentas da WS.

4.1. Cenário de Aplicação: Sistemas de Informação com dados de TB

No cenário brasileiro, a Portaria do MS 2073/2011 regula o uso de normas de interoperabilidade no âmbito do SUS e do setor de saúde suplementar (BRASIL, 2015) para garantir a interoperabilidade funcional e semântica dos sistemas de informação em saúde.

Em relação especificamente à TB na DRS XIII, o acompanhamento dos casos de TB envolve o preenchimento de diversos instrumentos de registro padronizados pelo MS, como o FNI, o Registro de TDO e o Registro de Tratamento e Acompanhamento. Levantamento de casos de TB, além do prontuário eletrônico e do TBWEB ou SINAN em outros estados. Outros sistemas também estão envolvidos, a saber, e-SUS AB, SISTB (YUKIE et al., 2007), e sistemas de informação de hospitais - que são, respectivamente, sistemas de informação de saúde governamentais, estaduais e regionais:

- SISTB: Armazena e centraliza informações sobre o paciente, tratamento, exames e internações no município de Ribeirão Preto;
- HygiaWeb - Um software de gestão de saúde pública da cidade de Ribeirão Preto que está presente e conecta vários níveis de atenção à saúde regional;
- TBWEB - Software criado pelo governo do estado de São Paulo que tem como objetivo principal a vigilância epidemiológica do estado;
- e-SUS AB - Software de gerenciamento da Atenção Básica de Saúde em âmbito nacional;
- GAL: um acrônimo para Gerenciador de Ambiente Laboratorial, permite o gerenciamento de rotinas e o acompanhamento das etapas para a realização dos exames, contendo dados que podem estar associados aos registros de exames armazenados no SISTB.

- SITETB: Plataforma que apoia a notificação de toda prescrição de tratamento que não envolva os medicamentos comuns (rifampicina, isoniazida, pirazinamida e etambutol). Geralmente usado para TB resistente a medicamentos.

Levando-se em consideração o cenário brasileiro de sistemas de informação em saúde para a TB, o trabalho exposto nesta tese originalmente tem como inspiração o trabalho de ABHISHEK; M.P., 2016 que, por sua vez, descreve a criação de uma ontologia para auxiliar a tomada de decisão dos gestores do Programa Nacional de Controle e Gestão da Tuberculose da Índia (RNTCP). A base utilizada desta ontologia foi a BFO, caracterizada por ser uma meta-ontologia que permite hierarquizar e corrigir a divisão e classificação das entidades do domínio que se deseja representar.

Com a ontologia criada para o cenário brasileiro, ao utilizá-la junto das ferramentas da WS citadas em 3.2 desta tese, cria-se a base para atingir a interoperabilidade semântica e funcional entre as aplicações brasileiras que dispõem de dados em TB.

Esta abordagem, conforme também citado por ABHISHEK; M.P. 2016, permite um mapeamento entre outras ontologias que fazem uso da mesma meta-ontologia para sua construção, levando em consideração os desafios expostos em 3.3.4.2 desta tese. Também estaria contemplado pela adoção desta abordagem a consistência semântica para a eventual aplicação de algoritmos de extração de conhecimento baseados na Linguagem de Regras da WS (SWRL).

4.2. Utilizando a BFO para construir a Ontologia da TB

Como já comentado anteriormente, a pedra angular de uma solução WS é sua ontologia subjacente. Assim, o primeiro passo foi o desenvolvimento de uma ontologia considerando: a) os conceitos clínicos da TB; b) os diferentes sistemas existentes, FNI, TBWEB, SINAN, SISTB, HygiaWeb, e-SUS AB e GAL; c) e os conceitos relacionados à TB e ao TDO. O TDO é a estratégia internacionalmente recomendada para o controle da TB pela OMS, que tem sido reconhecida como uma estratégia altamente eficiente e econômica. Como mencionado anteriormente, esta é a estratégia adotada pelo governo brasileiro.

As entidades da ontologia foram alocadas sob as definições da BFO, que é uma ontologia de alto nível desenvolvida inicialmente para uso em domínios científicos como a biomedicina. O BFO vê a realidade em relação a uma divisão de nível superior de todas as entidades particulares (indivíduos) nas duas categorias separadas de continuantes e ocorrente. As entidades continuantes incluem objetos, atributos e locais e são contrastadas com entidades ocorrentes, que incluem processos e regiões temporais. Os processos acontecem no tempo e,

portanto, têm partes temporais. Continuantes, ao contrário, existem na totalidade em qualquer momento em que existam. Por se tratar de uma ontologia de alto nível e com um nível de representação realista, permite mapear várias entidades, processos e suas respectivas funções e características dentro de um espaço temporal, destacando-se sobre outras ontologias que somente representam visões instantâneas dessas situações. Dado o interesse em mapear os termos das áreas de medicina e administrativa relacionados à TB, essa ontologia de nível superior é uma excelente alternativa.

A Figura 1 apresenta a ontologia resultante do TDO para as políticas de saúde brasileiras e a Figura 2 representa as propriedades dos objetos relacionados ao TDO.

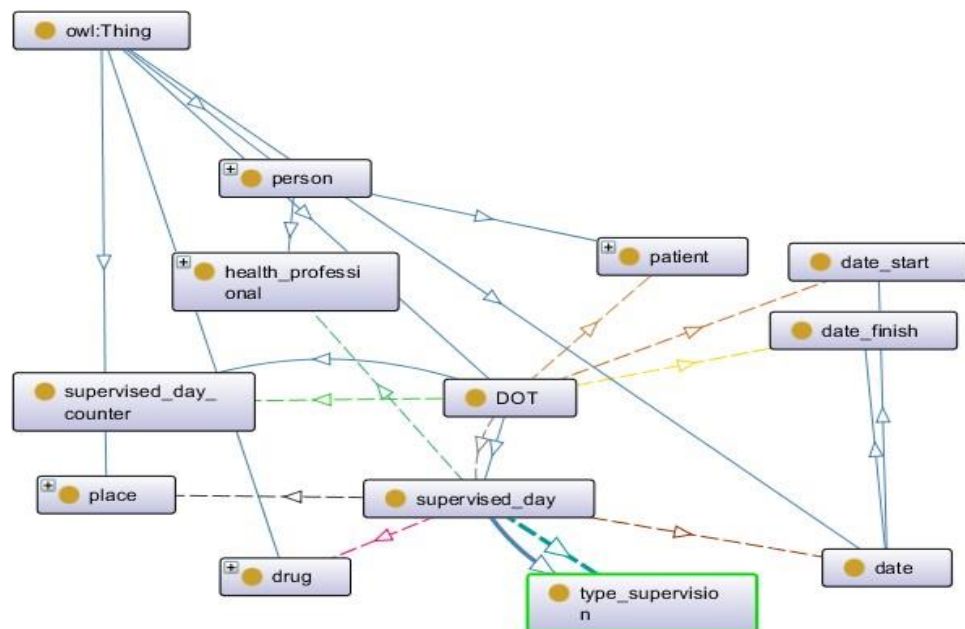


Figura 1 - Ontologia para suportar o TDO no cenário brasileiro (PELLISON, et al., 2020).

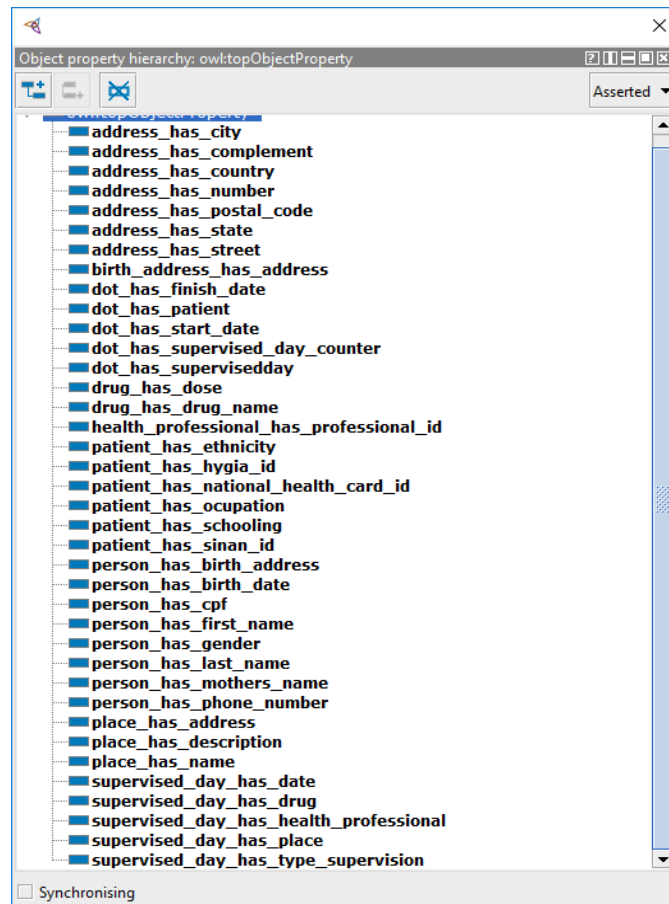


Figura 2 - Propriedades dos objetos relacionados ao TDO (PELLISON, et al., 2020).

O trabalho de (ABHISHEK; M.P., 2016) demonstrou a criação de uma ontologia para o cenário indiano da TB, servindo de base para a criação da ontologia apresentada neste trabalho. Quando utilizamos a mesma estrutura de classificação para o desenvolvimento da ontologia que reflete o cenário brasileiro, o mapeamento ontologia-ontologia será facilitado. Isso porque o BFO já classifica esses termos, tornando trivial o estágio de abstração de significados e apenas sendo necessário relacionar os termos com significados semelhantes.

5. RESULTADOS E DISCUSSÃO

5.1. Ontologia para apoiar a interoperabilidade no caso da TB

Como já especificado na seção de Métodos, foi dada a continuidade à construção da ontologia com base na BFO. Ao eliminar a etapa de abstração de significado, eliminamos também uma carga de subjetividade inerente que possa levar a erros no mapeamento de uma ontologia para outra. As figuras 3 e 4 são trechos dos conceitos que foram mapeados no BFO.

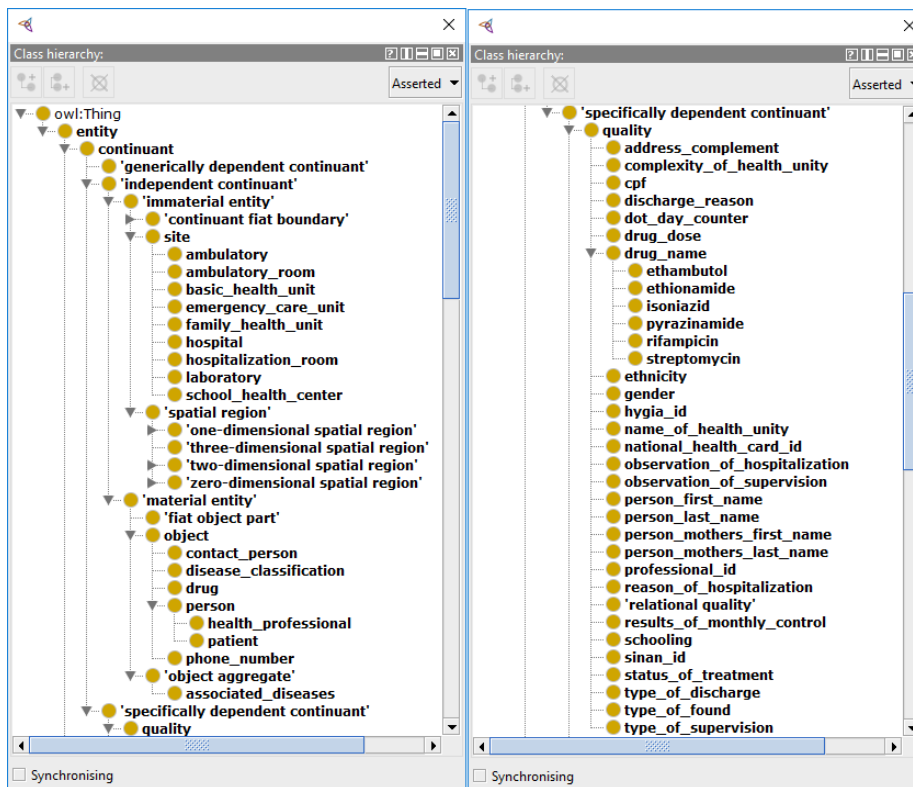


Figura 3 - Primeira parte das entidades relacionadas à TB mapeadas na BFO (PELLISON, et al., 2020).

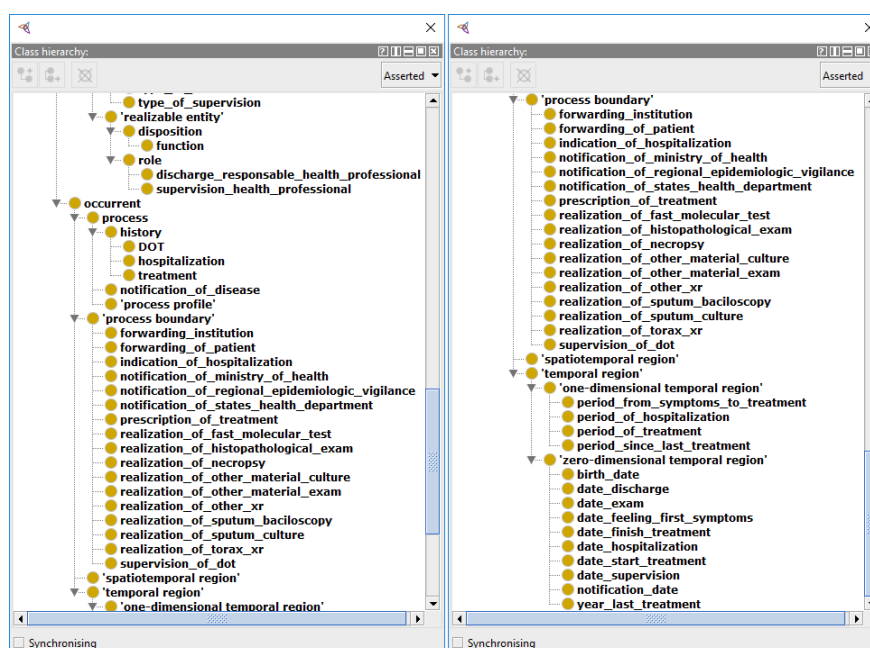


Figura 4 -Segunda parte das entidades relacionadas à TB mapeadas na BFO (PELLISON, et al., 2020).

O mapeamento de termos na estrutura do BFO apresentou uma dificuldade considerável, dada a complexidade filosófica envolvida nas definições dos mesmos segundo o grau de abrangência que escolhemos tomar. É claro que a granularidade das ações especificadas na construção da ontologia pode crescer até o nível desejado. No entanto, para este trabalho, a formalização resultante mostrada nas Figuras 3 e 4 foi o resultado de um consenso entre os especialistas colaboradores do trabalho, e julgado suficiente para apoiar o desenvolvimento do trabalho e permitir a interoperabilidade entre os sistemas propostos. O mapeamento das entidades apoiados pela BFO e validados por colaboradores reforça a possibilidade de interoperabilidade com iniciativas como a apresentada por (ABHISHEK, 2016).

Os principais conceitos que permitiram a interoperabilidade entre sistemas e a recuperação dos dados TDO estão descritos na Figura 5. Nesta figura, estão descritos especialmente os dados pessoais dos pacientes (endereço, data de nascimento, cpf, nome da mãe) e também informações relacionadas à saúde (cartão nacional de saúde, hygia id, sinan id). As características de transitividade dos dados recuperados, juntamente com a garantia de interoperabilidade semântica fornecida pela ontologia, permitem a fusão dos dados dos pacientes dos aplicativos participantes. Isso aumenta a relevância das informações disponíveis para o profissional de saúde que precisa tomar decisões estratégicas e clínicas.

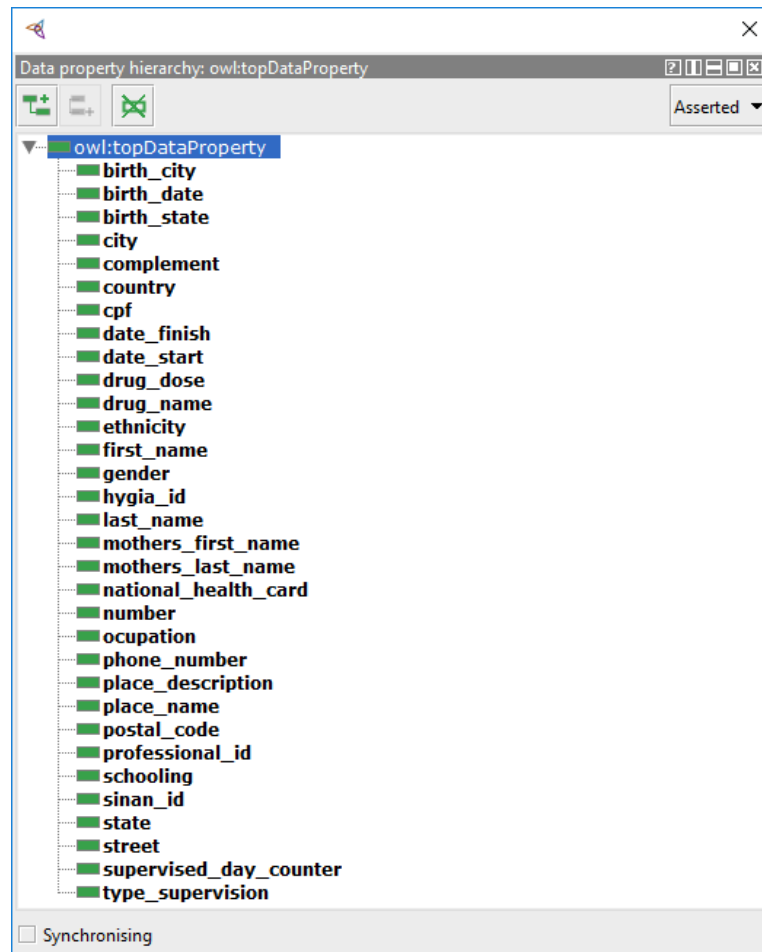


Figura 5 - Termos traduzidos a fim de permitir a troca de dados entre aplicações nacionais e internacionais (PELLISON, et al., 2020).

5.2. Mapeamento de ontologias via web

Segundo as necessidades apresentadas em 4.2 houve a necessidade de se desenvolver uma aplicação para que pudesse ser feita a remarcação semântica dos dados baseado em pareamentos de ontologias previamente cadastrados. Dada a facilidade de publicação e os conhecimentos adquiridos durante o desenvolvimento do projeto, realizou-se o desenvolvimento de uma interface onde o usuário seleciona qual a ontologia de origem em que os dados estão marcados e para qual ontologia ele deseja marcar os dados.

Ontology Mapping Service

Source Ontology
 Person - version 1

Target Ontology
 Person - version 2

What is your data format?
 CSV HTML

⚠ The CSV must have a well-defined structure:

- The first line is the header that must contain the properties that you want to map
- Make sure to use ";" or "," as column delimiter

Your CSV should look like this

```
property-1; property-2; property-3; property-n
value-1; value-2; value-3; value-n
value-1; value-2; value-3; value-n
```

Example using properties from ontology <http://schema.org/Person>

```
givenName; additionalName; familyName; gender; age
Barack; Hussein; Obama; Male; 56
Donald; John; Trump; Male; 71
```

We will try to map all properties found in the header from the source ontology to the target ontology

Now, inform how you want to send your data
 Paste CSV text Upload CSV File

Paste your CSV here

Go!

Figura 6 - Interface de seleção de mapeamento de versão da ontologia. Fonte: própria

O usuário pode inserir os dados no campo de texto destinado à esta finalidade (*paste your CSV here*) ou então realizar o upload de um arquivo CSV conforme exemplo especificado na interface da Figura 6 e clicar no botão “Go” para que seja realizada a remarcação baseada no pareamento entre as duas ontologias selecionadas. Temos um exemplo de resultado na Figura 7 onde as tags correspondentes à “*givenName*” foram mapeadas para “*firstName*” e as tags *additional name* mapeadas para “*middleName*”.



Figura 7 - Exemplo de resultado do mapeamento. Fonte: própria.

Esta ferramenta torna-se útil quando o mapeamento entre uma ontologia e outra já está pré-estabelecido e deseja-se importar dados juntamente com seu significado semântico. É possível, também, por meio da ferramenta manter um histórico das versões de mapeamento entre as ontologias. Isto é importante pois a ontologia, como representação de um domínio, pode sofrer alterações de acordo com o tempo.

5.3. A arquitetura de interoperabilidade para TB

Na Figura 8, com base em nosso trabalho (PELLISON et al., 2020), descrevemos a implementação da camada de interoperabilidade entre TBWEB, SINAN, SISTB, HygiaWeb, e-SUS AB, SITETB e GAL, que permitirá o intercâmbio de informações ocorra de forma transparente ao do utilizador. Em nosso trabalho anterior, consideramos apenas um subconjunto desses sistemas. No presente, a camada de interoperabilidade foi desenvolvida usando o paradigma da WS e os padrões preconizados pelo W3C. Este paradigma permitiu a extração de conhecimento e conteúdo destes sistemas de forma otimizada para máquinas por meio de um serviço web, abrindo um leque de possibilidades para geração de informações úteis, conforme especificado por Tim Berners-Lee, James Hendler e Ora Lassila para tomadas de decisão.

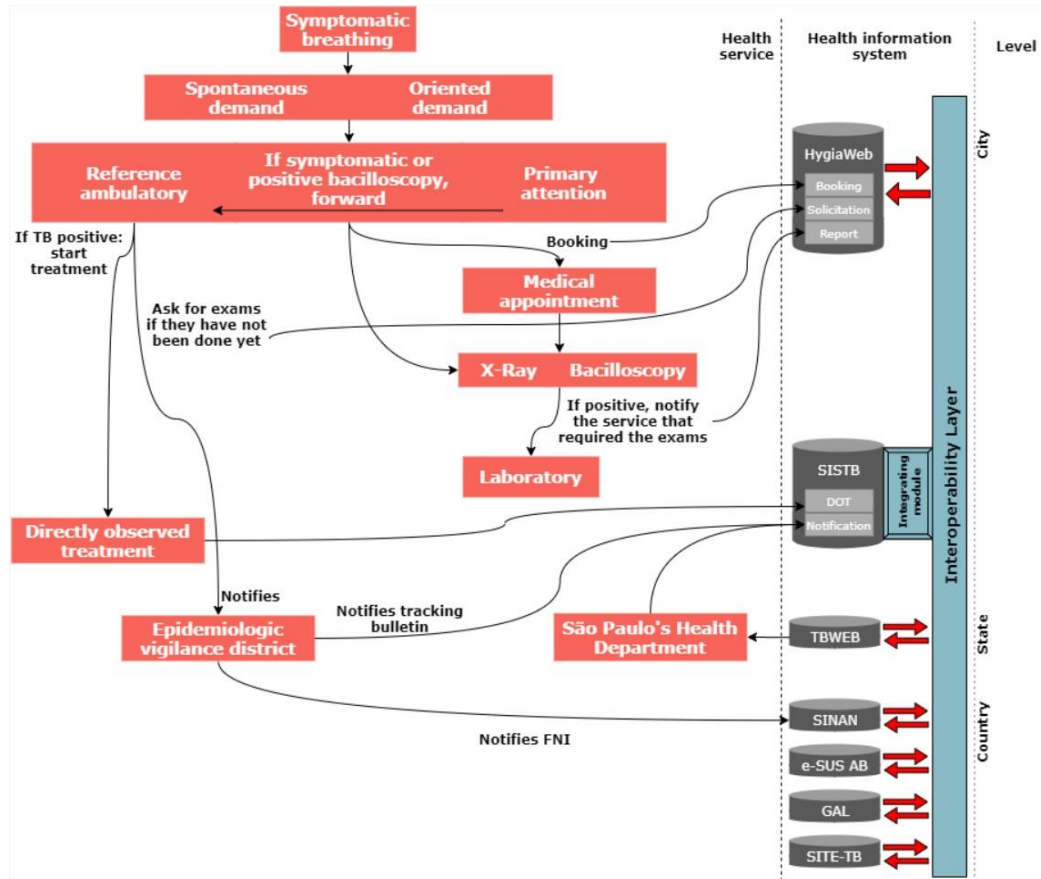


Figura 8 - Fluxo de informação para tratamento de TB após implementação (PELLISON et al, 2020).

A busca pelo serviço de saúde ocorre quando o paciente apresenta algum sintoma relacionado ao sistema respiratório. Essa demanda pode ser espontânea ou orientada. A demanda espontânea se refere quando o paciente procura o médico que atua na atenção primária em saúde (clínico geral/médico de família) ou especialista em pneumologia que atua nos níveis secundário ou terciário de saúde. A demanda orientada é quando o clínico geral encaminha o paciente para o especialista. Se os sintomas corroboram a suspeita de TB, são solicitados exame radiológico e de escarro e agendada consulta médica. Com os resultados dos testes em mãos, o paciente é encaminhado para um ambulatório de TB, onde o tratamento será iniciado. O tratamento é diretamente observado e todos os dados de acompanhamento são armazenados no SisTB. Para que a camada de interoperabilidade seja alcançada, é necessário que todos os sistemas sejam marcados semanticamente usando a ontologia criada através de *Microdata*.

Marcação de página foi feita por meio da inserção de tags HTML usando o formato *Microdata*. Tags referem-se diretamente aos termos usados na ontologia criada. Esse formato

de marcação foi escolhido porque permite que os mecanismos de pesquisa extraiam facilmente o conhecimento dos campos marcados com as tags, pois a linguagem HTML é uma base comum para aplicativos da Web. A grande vantagem dessa extração de conhecimento é que a interoperabilidade semântica já está implicitamente inserida no contexto da página, já que a ontologia fornece toda estrutura lógica à qual os dados marcados estão vinculados, evitando um trabalho adicional de atribuir significado aos dados retornados.

Em cada um desses sistemas, também deve haver um serviço de *endpoint* SPARQL ativo, permitindo que as consultas SPARQL também sejam executadas em informações armazenadas em bancos de dados legados. Esse *endpoint* é fornecido pelo *software* D2R. Isso é fundamental para extrair dados que foram armazenados antes de adotar o paradigma semântico da Web para habilitar a camada de interoperabilidade.

A extração de informações diretamente das páginas marcadas é realizada por meio da API Any23, que extrai diretamente os objetos ou literais e suas tags (correspondentes à ontologia criada). Com os dados extraídos, é possível realizar vários tipos de consultas SPARQL e incorporar tais informações em seu banco de dados local para quaisquer consultas.

No Quadro de Código 2, podemos ver o exemplo de uma consulta SPARQL, onde todas as propriedades de todos os pacientes são retornadas. Esta é uma consulta muito simples sendo usada como base para a incorporação de dados extraídos de outros sistemas marcados com a respectiva ontologia.

```
"SELECT ?propriedade ?valor \n" +
"WHERE { \n" +
"  ?paciente ?propriedade ?valor .\n" +
"  FILTER(?valor != \"Null\") .\n" +
"  FILTER(isLiteral(?valor)) .\n" +
"}";
```

Quadro de Código 2 - Exemplo de uma query SPARQL utilizada. Fonte: própria.

É importante observar que essa consulta pode ser executada nos dados de todos os sistemas marcados. Isso garante que os dados retornados tenham o mesmo significado, uma vez que foram marcados com a mesma ontologia e, portanto, interoperáveis entre os sistemas.

Essa abordagem permite que os dados marcados com essa mesma ontologia sejam buscados em qualquer sistema (via *endpoint* HTTP ou SPARQL). Os dados retornados são então tratados e incorporados ao sistema, se desejado, ou podem ser usados para consultas e estatísticas ad hoc, permitindo uma rápida tomada de decisão pelos profissionais de saúde.

Nesse sentido, tanto a interoperabilidade semântica quanto a interoperabilidade funcional atuam em função do ideal de integralidade e resposta rápida buscado pelos sistemas de informação em saúde.

The screenshot shows the GAL interface with a patient record for 'PACIENTE TESTE DEZ'. The interface includes a header with the GAL logo, a navigation bar with buttons like 'Consultar Exames', and several data sections: 'Requisitante', 'Dados do paciente', 'Dados de Residência', and 'Dados clínicos gerais'. To the right of the interface, a portion of the semantic markup code is visible, showing HTML tags with semantic attributes like 'itemprop' and 'itemtype'.

Figura 9 - Captura de tela e parte do código semanticamente marcado do ambiente de testes do GAL. Fonte: própria.

The screenshot shows the TBWeb interface with a patient record for 'PACIENTE TESTE DEZ'. The interface includes a header with the TBWeb logo, a navigation bar, and several data sections: 'Dados de identificação', 'Dados de nascimento', 'Dados de residência', 'Dados clínicos gerais', and 'Dados de exames'. To the right of the interface, a portion of the semantic markup code is visible, showing HTML tags with semantic attributes like 'itemprop' and 'itemtype'.

Figura 10 - Captura de tela e parte do código semanticamente marcado do ambiente de testes do TBWeb. Fonte: própria.

SINAN - Notificação

Identificação
Número de Notificação: 10

Dados Gerais

COD(CIO10)	Agravo/Doença Tuberculose	Data de Notificação 2017-07-23	UF SP	Município de Notificação Ribeirão Preto
Nome do paciente PACIENTE TESTE DEZ			Data de Nascimento 2000-05-04	Idade aproximada 17
Sexo Masculino		Gestante Não se aplica	Raça/Cor Branca	Escolaridade Ensino médio completo

Dados de Residência

CEM (CENSA) Nacional de Saúde 130456	Nome de mãe MÃE TESTE DEZ			
UF SP	Município de Residência Ribeirão Preto	Código (IBGE) Logradouro (Rua/Av/Elts) Rua Lala	Distrito II	Código
Nome Lala	Complemento		CEP 11111-111	
Telefone 23232323	Zone Urbana		País Brasil	

Dados Complementares do Caso

Número do prontuário 10	Tipo de Entrada Caso Novo			
Populações Especiais Não	Beneficiário de programa de transferência de renda do governo Não			
Forma Pulmonar + Extrapulmonar	Se Extrapulmonar Não			
Bactioscopia de escarro (diagnóstico) Positiva	Radiografia do Tórax Suspeita	HIV Em andamento		
Terapia Antimicrobiana Durante o Tratamento para a TB Não	Histopatologia Sugestivo de TB	Teste de sensibilidade Em andamento		
Cultura Positivo	Teste Molecular Rápido TB (TMR-TB) Não realizado			
Data de Início do Tratamento Atual 2017-07-23	Total de Contatos Identificados 2			
Município/Unidade de Saúde Ribeirão Preto/Unidade A	Cod. Unidade de Saúde 12			
Nome Prof. Lala	Função Auxiliar			

Figura 11 - Captura de tela e parte do código semanticamente marcado do ambiente de testes do SINAN. Fonte: própria.

Dados do Paciente

Nome paciente: NOME TESTE 10
CPF: 00000000000
Cartão Nacional de Saúde: 80000105
Número do Prontuário: 6629270
Número Hygia: 105
Data de Nascimento: 30/04/2000
Idade: 17
Sexo: Masculino
Gestante: Não
Nome da Mãe: MÃE TESTE 10
Etnia: Pardo
Naturalidade: Ribeirão Preto
Escolaridade: 0 ano(s) concluído(s)
Tipo de Ocupação: Sem ocupação
Outro tipo de ocupação:
Observações:
Número SINAN: 10

Contato

Endereço: Rua 105
Cidade: RIBEIRÃO PRETO
UF: SP
Telefone: (16) 1051-0510

Figura 12 - Captura de tela e parte do código semanticamente marcado do ambiente de testes do SisTB. Fonte: própria.

Pode-se notar por meio das capturas de tela, apresentadas nas Figuras 9 a 12, que todos

os sistemas de informação em saúde possuem dados comuns que identificam o paciente (CPF, Cartão Nacional de Saúde, Data de Nascimento). Ou seja, com a marcação semântica deles, é possível retornar os dados referentes a um determinado paciente e agregá-los em um único sistema ou importar apenas dados suficientes para a tomada de decisão, como é mostrado no exemplo da Figura 13. A intersecção desses registros permite um registro completo da situação de saúde atual do paciente, corroborando com a integralidade na assistência à saúde do paciente. A recuperação dessas informações a partir de vários sistemas de informação em saúde permite que os profissionais de saúde tomem suas decisões de maneira mais sólida, pois contém detalhes que não estariam presentes se estivessem usando apenas um dos sistemas.

Do ponto de vista gerencial, o avanço na agregação de informações para a realização de estudos demográficos também é aprimorado. O desenvolvimento de políticas públicas de saúde mais especializadas e com melhor eficácia também é facilitado, uma vez que a disponibilidade de informações sobre a maioria dos pacientes é aumentada.

Resultado da SPARQL Query SisTb		Resultado da SPARQL Query TBWeb	
propriedade	valor	propriedade	valor
<http://sis1b-dev.ddns.net/ontology/Paciente/nroProntuario>	"6629276"	<http://sis1b-dev.ddns.net/ontology/Exame/nome>	"RX Outro"
<http://sis1b-dev.ddns.net/ontology/Paciente/nroSinamo>	"18"	<http://sis1b-dev.ddns.net/ontology/Exame/status>	"Normal"
<http://sis1b-dev.ddns.net/ontology/Paciente/idade>	"17"	<http://sis1b-dev.ddns.net/ontology/Exame/nome>	"Cultura do Escarro"
<http://sis1b-dev.ddns.net/ontology/Paciente/dataNascimento>	"19/04/2000"	<http://sis1b-dev.ddns.net/ontology/Exame/status>	"Em andamento"
<http://sis1b-dev.ddns.net/ontology/Paciente/gestante>	"Não"	<http://sis1b-dev.ddns.net/ontology/Exame/nome>	"Baciloscopia de escarro"
<http://sis1b-dev.ddns.net/ontology/Paciente/nomePaciente>	"NOME TESTE 10"	<http://sis1b-dev.ddns.net/ontology/Exame/status>	"Em andamento"
<http://sis1b-dev.ddns.net/ontology/Paciente/cartaSus>	"80000185"	<http://sis1b-dev.ddns.net/ontology/Medicamento/nome>	"Amoxicilina (A)"
<http://sis1b-dev.ddns.net/ontology/Paciente/nomeMae>	"MÃE TESTE 10"	<http://sis1b-dev.ddns.net/ontology/Exame/status>	"Outros"
<http://sis1b-dev.ddns.net/ontology/Paciente/genero>	"Masculino"	<http://sis1b-dev.ddns.net/ontology/Medicamento/nome>	"Osixacina (OF1)"
<http://sis1b-dev.ddns.net/ontology/Paciente/escolaridade>	"0 ano(s) concluido(s)"	<http://sis1b-dev.ddns.net/ontology/Exame/nome>	"Outras Drogas"
<http://sis1b-dev.ddns.net/ontology/Paciente/etnia>	"Pardo"	<http://sis1b-dev.ddns.net/ontology/Exame/status>	"Necropsia"
<http://schema.org/Person/telephone>	"(16) 1851-0510"	<http://sis1b-dev.ddns.net/ontology/Medicamento/nome>	"Não sugestivo TB"
<http://schema.org/Person/address>	"Endereço: Rua 185Cidade: RIBEIRAO PRETOUF: SP"	<http://sis1b-dev.ddns.net/ontology/Medicamento/nome>	"Pirazinamida (Z)"
		<http://sis1b-dev.ddns.net/ontology/Medicamento/nome>	"Sin"
		<http://sis1b-dev.ddns.net/ontology/Medicamento/nome>	"Etambutol (E)"
		<http://sis1b-dev.ddns.net/ontology/Medicamento/nome>	"Sin"
		<http://sis1b-dev.ddns.net/ontology/Medicamento/nome>	"Isoniazida (H)"
		<http://sis1b-dev.ddns.net/ontology/Medicamento/nome>	"Sin"
		<http://sis1b-dev.ddns.net/ontology/Medicamento/nome>	"Clofazimina (Clo)"
		<http://sis1b-dev.ddns.net/ontology/Exame/nome>	"Baciloscopia Outro Material"
		<http://sis1b-dev.ddns.net/ontology/Exame/status>	"Em andamento"
		<http://sis1b-dev.ddns.net/ontology/Exame/status>	"DX de Tórax"
		<http://sis1b-dev.ddns.net/ontology/Exame/nome>	"Normal"
		<http://sis1b-dev.ddns.net/ontology/Exame/status>	"Cultura do Escarro"
		<http://sis1b-dev.ddns.net/ontology/Exame/status>	"Bovetivo"
		<http://sis1b-dev.ddns.net/ontology/Paciente/nroSinamo>	"18"
		<http://sis1b-dev.ddns.net/ontology/Paciente/genero>	"Masculino"
		<http://sis1b-dev.ddns.net/ontology/Paciente/dataNascimento>	"19/04/2000"
		<http://sis1b-dev.ddns.net/ontology/Paciente/nomePaciente>	"NOME TESTE 10"
		<http://sis1b-dev.ddns.net/ontology/Paciente/ocupacao>	"Estudante"
		<http://sis1b-dev.ddns.net/ontology/Paciente/nomeMae>	"Outro"
		<http://sis1b-dev.ddns.net/ontology/Paciente/escolaridade>	"De 0 a 11 anos"
		<http://sis1b-dev.ddns.net/ontology/Paciente/idade>	"17"
		<http://sis1b-dev.ddns.net/ontology/Paciente/cartaSus>	"18"
		<http://sis1b-dev.ddns.net/ontology/Paciente/cpf>	"18"
		<http://sis1b-dev.ddns.net/ontology/Paciente/rg>	"18"
		<http://sis1b-dev.ddns.net/ontology/Medicamento/nome>	"Terizidona (T)"
		<http://sis1b-dev.ddns.net/ontology/Medicamento/nome>	"Estreptomicina (S)"
		<http://sis1b-dev.ddns.net/ontology/Medicamento/nome>	"Etionamida (E)"
		<http://purl.org/dc/terms/title>	"BMB"
		<http://sis1b-dev.ddns.net/ontology/Medicamento/nome>	"Fluamplicina (R)"
		<http://sis1b-dev.ddns.net/ontology/Medicamento/nome>	"Sin"
		<http://sis1b-dev.ddns.net/ontology/Exame/nome>	"Anti-HIV"
		<http://sis1b-dev.ddns.net/ontology/Exame/status>	"Em andamento"
		<http://sis1b-dev.ddns.net/ontology/Exame/nome>	"Histopatológico"
		<http://sis1b-dev.ddns.net/ontology/Exame/status>	"Não sugestivo TB"

Figura 13 - Resultados de uma query SPARQL simultânea no SisTB e TBWeb para o paciente com id = 10 (PELLISON et al, 2020).

Pelo acima exposto, foi realizada a construção de uma ontologia que sustenta o cenário brasileiro da TB, com ênfase no TDO, similar ao descrito por (ABHISHEK, 2016), para cenário indiano. No entanto, no escopo deste trabalho, a ontologia foi aplicada em um ambiente de interoperabilidade, demonstrando através de um experimento de integração a viabilidade da solução. Tal ontologia foi construída usando o meta-ontologia BFO para a classificação de termos. Essa formalização permitirá o mapeamento de entidades ontológicas brasileiras para a TB para outras ontologias que também utilizaram o BFO como modelo e

possuem entidades similares, facilitando a interoperabilidade semântica. Com a ontologia construída, uma arquitetura foi desenvolvida para permitir a interoperabilidade funcional entre aplicativos e sistemas que possuem dados de saúde relacionados à TB. A arquitetura foi implementada em um ambiente de teste, no qual versões de desenvolvimento de tais aplicativos são instaladas. As aplicações foram marcadas via microdata com os termos da ontologia criada. Essa marcação permite a extração de conteúdo de vários sistemas de informação em saúde a partir de uma única consulta SPARQL nos respectivos endpoints instalados. Vale ressaltar que foram implementadas soluções que permitem consultas SPARQL em bancos de dados relacionais e triplestores, permitindo a manutenção de bancos de dados legados. Um teste conceitual foi realizado, e os dados fictícios de um determinado paciente de teste foram retornados de todos os *softwares* que tinham o mesmo registro a respeito do paciente. Os dados retornados mantiveram seus significados e a interoperabilidade semântica e funcional foi alcançada.

5.4. Integrando dados de TB no estado de São Paulo: uma prova de conceito

Numa fase seguinte, utilizando os recursos da WS como *queries* SPARQL e RDF, pretendemos demonstrar com uma prova de conceito o potencial de uma possível integração de dados de TB no Estado de São Paulo. Por meio de uma única *query* federada, procuramos recuperar simultaneamente dados obtidos do TBWeb, um sistema governamental de nível estadual, e também dos dados contidos no SisTB, um sistema de informação em saúde usado principalmente em Ribeirão Preto. Ao realizar esta *query*, foi possível combinar estes dados de ambas as fontes com um valor semântico agregado. Quando ocorreram divergências entre os trechos de dados recuperados, os mesmos foram marcados e os usuários podem tomar ações acerca deles.

Dentre as ferramentas tecnológicas usadas, foi escolhido para o desenvolvimento do *frontend* HTML5, Java Script e Java Server Pages Standard Tag Library (JSTL). Foram também aplicados componentes do Framework Bootstrap para tornar o site responsivo e por fim o Bing Maps API foi escolhido para entregar a funcionalidade de mapa para o georeferenciamento.

Para o *backend*, foi usada a linguagem Java, Enterprise Edition Versão 8, aproveitando-se de *frameworks* disponíveis para lidar com aplicações semânticas, como o Apache Jena.

A arquitetura de interoperabilidade do SisTB inclui um *endpoint* SPARQL, fornecido pelo D2RQ Server. O aplicativo D2RQ permite que bases relacionais sejam acessadas virtualmente como RDF através de *queries* SPARQL. Além do que o endpoint SPARQL

permite que sejam feitas *queries* federadas em bases distintas, funcionalidade esta que é fundamental para a integração semântica dos dados.

Para este ensaio, uma amostra anonimizada do TBWeb com 7277 casos de TB no estado de São Paulo, do ano de 2006 a 2014, foi usada em conjunto com toda a base de dados do SisTB. Um identificador comum único (SINAN ID) foi escolhido para fazer o pareamento dos pacientes entre ambas as bases de dados. Os dados demográficos foram usados para localizar os casos no mapa e comparar valores entre os conjuntos de dados, como por exemplo: latitude, longitude, idade, sexo, cidade de notificação, unidade da federação (estado), escolaridade e etnia.

Para localizar os dados entre os sistemas uma ontologia simples baseada na Person ontology (<http://schema.org/Person>) foi usada, designada especificamente para identificar dados e para finalidades de testes. A ontologia foi usada por meio do D2RQ Server, que mapeia o esquema de base de dados relacionais e seu conteúdo. No entanto, o servidor retorna os dados como um grafo RDF virtual, obtido através de uma *query* SPARQL.

Aplicando-se esta ferramenta o usuário pode obter um *heatmap* de acordo com o número de casos notificados em uma determinada região, conforme a Figura 14.

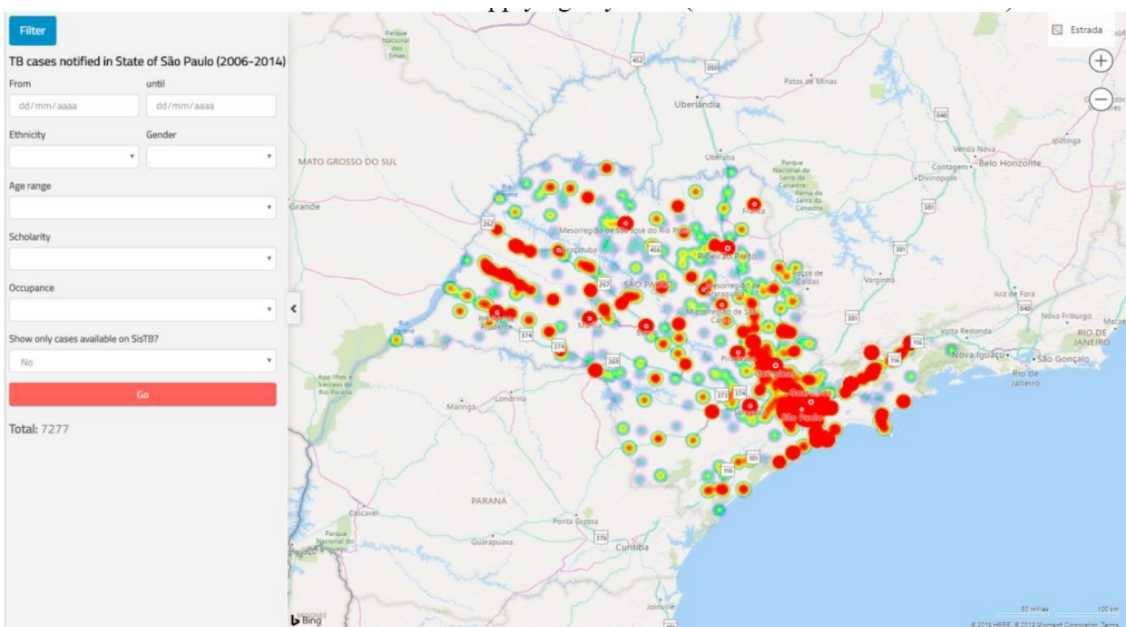


Figura 14 - Exemplo de distribuição de casos de TB no Estado de São Paulo para o período de 2006 a 2014 (PELLISON, et al., 2019).

Ao identificar divergências entre o pareamento de informações armazenadas nas bases de dados, a qualidade dos dados pode ser melhorada por meio de uma intervenção manual nos

dados ou então através de um método automatizado. Este cenário também permite que, num ambiente integrado, não haja o retrabalho de inserção dos mesmos dados em diferentes sistemas pelos profissionais de saúde. Os dados consolidados entre os sistemas de informação podem levar à valiosas bases de dados para o desenvolvimento de aplicativos de suporte a decisão.

Por meio da WS foi possível demonstrar que é possível recuperar e assimilar dados de fontes heterogêneas agregando significado semântico através das ontologias; realizar queries nestes dados utilizando o protocolo SPARQL; usar grafos RDF que facilitam a modelagem e a troca de dados semânticos.

5.5. Resultados Adicionais

A construção da ontologia baseada no BFO se beneficia do fato de ser facilmente relacionada às outras ontologias construídas no mesmo framework. Isso permite que diferentes ontologias sejam relacionadas com muito mais facilidade, porque suas entidades são organizadas sistematicamente e hierarquicamente, de acordo com seu significado semântico defendido pela mesma meta-ontologia. Essa relação inter-ontológica tem grande potencial, pois permite que dados marcados sejam compartilhados entre as instituições, levando seus respectivos valores semânticos, ampliando seu potencial de pesquisa multicêntrica. Atualmente, existem iniciativas para mapear termos e construir ontologias relacionadas ao tratamento da TB em outros cenários, conforme citado em (ABHISHEK; M.P., 2016). O mapeamento dos processos do TDO no âmbito brasileiro foi realizado com base na mesma meta-ontologia e devido à esta relação inter-ontológica, a possibilidade de transposição dos dados entre os sistemas marcados é factível e pode ser interessante para ambas as partes, permitindo assim a interoperabilidade semântica entre dois ou mais países distintos.

Durante o desenvolvimento da arquitetura foram encontrados alguns desafios que demandaram atenção e cuidado e tornam-se alvo de trabalhos futuros. Todas as ontologias podem ser modificadas, seja pelas mudanças na definição do domínio, seja por correções na sua construção, ou por causa dos ajustes considerados necessários para o bom funcionamento da ontologia pelo grupo que as mantém. No entanto, essas alterações afetam diretamente os sistemas marcados com ela, e é necessária uma atenção especial para os dados marcados pela ontologia. É necessário controlar as versões da ontologia para permitir a reconstrução da vida de uma ontologia e traçar o significado dos dados extraídos na versão em que foi marcada. Portanto, é necessário encontrar soluções para ajustes, da maneira mais fácil e intuitiva possível, que permita a remarcação dos dados de maneira simples e com impactos reduzidos nos sistemas. Alguns trabalhos já foram desenvolvidos nessa área, com foco na previsão de padrões

de mudanças na ontologia, por exemplo (JAVED; ABGAZ; PAHL, 2013).

Em conjunto com o controle de versão, também é necessário prestar atenção em como marcar esses dados, já que essa tarefa exige um esforço considerável em sistemas que usam ontologias com muitas entidades e termos. Tal marcação automática possui um pequeno número de trabalhos realizados e documentados, talvez justificado pela necessidade de marcadores especializados no domínio para os quais as respectivas ontologias foram desenvolvidas. Ainda é possível enfatizar algumas iniciativas que são propostas para ajudar os desenvolvedores das arquiteturas a realizar a marcação de maneira pelo menos semi-automática. Entre eles, podemos mencionar UCCA-App (ABEND; YERUSHLAMI; RAPPOPORT, 2013), MnM (VARGAS-VERA et al., 2002) e SemTag and Seeker (DILL et al., 2003). A redução do esforço de marcação, seja antes ou após a publicação de uma nova versão de uma dada ontologia, deve ser um foco do estudo de trabalhos futuros, para proporcionar maior facilidade na manutenção dos sistemas que utilizam este paradigma de WS.

Com esses problemas resolvidos, o próximo passo seria tornar essa abordagem interoperável com outros padrões, como os perfis HL7, OpenEHR e IHE.

No caso do HL7, talvez uma abordagem viável seria criar um middleware capaz de traduzir dados extraídos por meio de marcação semântica em mensagens HL7 (sejam elas V2.x ou 3.x). No Reino Unido, é descrito por (PLASTIRAS; O'SULLIVAN, 2017) o desenvolvimento de um middleware semelhante ao proposto para interoperar dados de registros pessoais de saúde em registros eletrônicos de saúde. Nossa proposição de middleware, entretanto, receberia como entrada os dados extraídos de aplicações semanticamente marcadas, construindo as mensagens e encaminhando-as para os destinatários - e também fazendo a via inversa. Neste cenário, um esforço deve ser feito para mapear as entidades das ontologias específicas de domínio nos campos pré-especificados, recomendados pelo padrão HL7. Tal esforço seria semelhante ao mapeamento de quais entidades compõem um determinado arquétipo do OpenEHR consolidado. Ou seja, para interoperar sistemas que usam a WS com sistemas que usam o padrão OpenEHR, é necessário garantir que os arquétipos sejam totalmente representados por entidades de uma ontologia. O recíproco também é verdadeiro, já que o mapeamento de arquétipos em entidades de uma ontologia também é extremamente necessário. Tais processos são exemplificados na Figura 15.

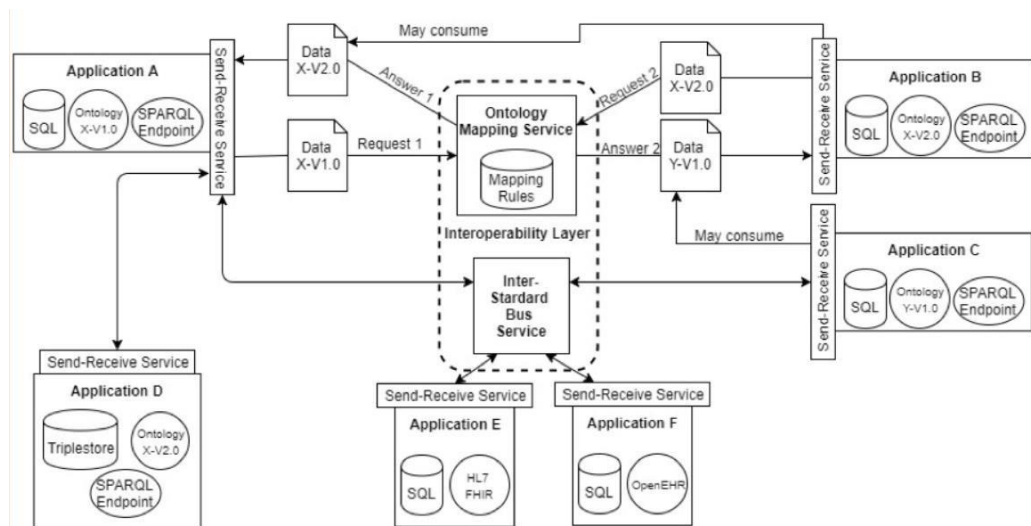


Figura 15 - Arquitetura de interoperabilidade proposta entre aplicações que utilizam WS e outros padrões (PELLISON, et al., 2020).

A flexibilidade na construção das ontologias e o alto nível de abstração que eles possuem tornam o processo de mapeamento de arquétipos relativamente trivial quando já se tem um arquétipo aceito para a fundação responsável pelo gerenciamento do repositório padrão.

Ainda é importante ressaltar que a abordagem da WS pode ser incorporada em vários perfis IHE para que a integração possa ocorrer em diversas áreas, bastando ter uma ontologia que suporte todas elas e seus respectivos processos. No entanto, o uso da WS tem a vantagem de ter uma governança de escopo de domínio muito mais dinâmica do que outros padrões de interoperabilidade, como o HL7 e o OpenEHR anteriormente citados. Tais dinâmicas favorecem a evolução dos termos e adaptação às novas tendências, algo que é inerentemente recorrente na saúde, onde técnicas, novos procedimentos são incorporados e novos protocolos clínicos evoluem a cada dia.

6. CONCLUSÃO E TRABALHOS FUTUROS

Os trabalhos citados nas seções 5.1, 5.2, 5.3, 5.4 e 5.5 demonstram a viabilidade de uma estrutura de interoperabilidade semântica e funcional para o cenário da TB. Seus resultados são diretamente aplicáveis ao SisTB do município de Ribeirão Preto, e servem de referencial teórico-prático para outros trabalhos dos grupos de pesquisa do Laboratório de Inteligência em Saúde da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo.

O desenvolvimento de uma arquitetura de interoperabilidade baseada em WS torna o processo dinâmico e com uma liberdade de governança considerável, permitindo aos gestores que sejam feitas adequações de maneira ágil. Fator que é fundamental, visto as características de mutabilidade dos diversos termos e fluxos na área da saúde.

Uma limitação do nosso trabalho é o mapeamento das entidades que dizem respeito à TBMR, à TB extensivamente multirresistente (TBEMR), à aplicação da notificação nesses casos (SITETB) e às comorbidades. Em trabalhos futuros incluirá o mapa desses fluxos de trabalho e outras aplicações de suporte, como o Sistema Nacional de Regulação (SISREG) e as aplicações demográficas do Instituto Brasileiro de Geografia e Estatística.

Apesar da facilidade de usar bancos de dados legados, como trabalho futuro, há uma necessidade de melhorar os serviços que facilitariam a implementação desta solução na prática diária. A marcação automática de dados pode ser uma área de interesse de estudo, auxiliando no esforço necessário para atribuir significado semântico aos dados. Outros exemplos viáveis no curto prazo seriam a implementação de um serviço de roteamento e barramento de mensagens inter-padrões. Isto permitiria a interoperabilidade entre aplicativos que usam paradigmas e padrões diferentes da WS, como o HL7 FHIR e o OpenEHR, completando o arcabouço tecnológico necessário para o cumprimento dos requisitos de integralidade e multidisciplinaridade do atendimento em saúde.

7. REFERÊNCIAS

AAMIR, M. J.; MANSOOR, A. **Testing Web Application from usability perspective**. In: Proceedings of the 3rd International Conference on Computer, Control & Communication. Karachi, Pakistan, 2013.

ABANDA, F. H.; TAH, J. H. M.; KEIVANI, R. **Trends in built environment semantic Web applications: Where are we today? Expert Systems with Applications**, v. 40, n. 14, p. 5563- 5577, 2013.

ABEND, O.; YERUSHLAMI, S.; RAPPOPORT, A. **UCCAApp: Web-application for Syntactic and Semantic Phrase-based Annotation**. p. 2–7, 2013.

ABHISHEK, K.; M.P., S. **An Ontology based Decision support for Tuberculosis Management and Control in India**. International Journal of Engineering and Technology, v. 8, n. 6, p. 2860– 2877, 2016.

ABIDA, B.; BIRBECK, M. **RDFa Core 1.1 - Third Edition**, 2015. Disponível em: <<https://www.w3.org/TR/rdfa-core/>>.

ALBUQUERQUE, M.F.M. LEITÃO, C.C.S. CAMPELO, A.R.I.; SOUZA, V.W.S. **Fatores prognósticos para o desfecho do tratamento da tuberculose pulmonar em Recife**, Pernambuco, Brasil. Rev Panam Salud Pública. 2001; 9:368-7.

ALLEMANG, D.; HENDLER, J. **Semantic web for the working ontologist: effective modeling in RDFS and OWL**. Amsterdam: Morgan Kaufmann, 2011. 384 p.

ALMEIDA P.F.; GIOVANELLA L.; MENDONÇA M.H.; ESCOREL S. **Challenges for healthcare coordination: strategies for integrating levels of care in large cities** . Cadernos de Saude Publica. 2010

ALMEIDA, C. P. F. et al. **Challenges for healthcare coordination: strategies for integrating levels of care in large cities**. Fev, v. 26, n. 2, p. 286–298, 2010.

ANGELES, A.J.M; RODRIGO, T. ALTET, M.N.; JIMÉNEZ-RUIZ, C.A.; CASALS, M.; PENAS,A.; MIR,I.;REINA,S.S.; RIESCO-MIRANDA, J.A; CAYLÁ,J.A. **Factors associated with smoking among tuberculosis patients in Spain**. BMC InfectDis. 2016; 16:486.

ANTONIOU, G.; HARMELEN, F. **A Semantic Web Primer** (Vol. 2e). Massachusetts London, England: The MIT Press Cambridge. 2008

ARGÜELLO, M.; EL-HASIA, A.; LEES, M. **A Semantic Web Portal to construction knowledge exchange**. WIT Transactions on Information and Communication Technologies, v. 37, p. 417-427, 2006.

BARREIRA, D. **Situação epidemiológica e política de controle da tuberculose no Brasil. Brasília, DF, 2012**. Documento apresentado no V Encontro Nacional de Tuberculose e ao II Fórum da Parceria Brasileira, Brasília, DF, 2012.

BASSIN, T. B. **Explorer's Guide to the Semantic Web**. Shelter Island: Manning, 2004. 305p.

BEEK, W. et al. **A web observatory for the machine processability of structured data on the web**. In: Proceedings of the 2014 ACM conference on Web science. Bloomington, Indiana, USA, 2014.

BELLEAU, F. et al. **Bio2RDF: Towards a mashup to build bioinformatics knowledge systems**. Journal of Biomedical Informatics, v. 41, n. 5, p. 706–716, 2008.

BELLUZZO, R. C. B. **Inteligência, informação e conhecimento em corporações**. Brasília, 2006. ISBN: 85-7652-063-x.

BERGEL, F.S.; GOUVEIA, N. **Retornos frequentes como nova estratégia para adesão ao tratamento de tuberculose**. Rev Saúde Pública. 2005; 39:898-905.

BERNERS-LEE, T. **Design Issues**. [S.l.]: W3C, 2006. Disponível em: <<https://www.w3.org/DesignIssues/>>.

BERNERS-LEE, T. **Semantic Web Concepts**, 2005. Disponível em: <<http://www.w3.org/2005/Talks/0517-boit-tbl>>.

BERNERS-LEE, T. **The Semantic Web Layer Cake**. In: XML-2000 Conference, Washington, United States of America, 2000.

BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. **The Semantic Web. A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities**.

Scientific American, v. 284, n. 5, p. 34–43, 2001.

BITTNER, T. AND DONNELLY, M. **The Mereology of Stages and Persistent Entities**. In Proceedings of the 16th European Conference on Artificial Intelligence, pages 283–287, Valencia. IOS Press. 2004.

BIZER, C.; HEATH, T.; BERNERS-LEE, T. **Linked Data - The Story So Far**. International Journal on Semantic Web and Information Systems, v. 5, n. 3, p. 1-22, 2009.

BORST, W. N.; **Construction of Engineering Ontologies for Knowledge Sharing and Reuse**. Enschede, The Netherlands: University of Twente. 1997

BRANCO, M. A. F. **Sistemas de informação em saúde no nível local**. Cadernos de Saúde Pública, v. 12, n. 2, p. 267–270, 1996.

BRASIL. Ministério da Saúde Secretaria Executiva Departamento de Informação e Informática do SUS. **Política Nacional de Informação e Informática em Saúde**. v. 1.0, 2004.

BRATSAS, C. et al. **Using Social Semantic Web Technologies in Public Health: A Prototype Epidemiological Semantic Wiki**. In: Proceedings of the XII Mediterranean Conference on Medical and Biological Engineering and Computing. Chalkidiki, Greece, 2010.

BURANARACH, M. et al. **A Semantic Web Framework to Support Knowledge Management in Chronic Disease Healthcare**. Metadata and Semantic Research, v. 1, p. 164-170,2009.

CAMPANI, S.T.A; MOREIRA, J.S.; TIETBOHEL, C.N. **Fatores preditores para o abandono do tratamento da tuberculose pulmonar (esquema de primeira linha) em Porto Alegre (RS)**. Dissertação de Mestrado do Programa de Pós-graduação em Ciências Pneumológicas da Universidade Federal do Rio grande do Sul. Porto Alegre, 2009.

CANETTI, G. **Primo-infection et reinfection dans La tuberculosis**. Paris, Flammanrion. 1954; C.11; p.143-156.

CANTARELLI, A. G. **D2R editor: habilitando a publicação automática de anotações semânticas de sites dinâmicos**. 2005. Dissertação (Mestrado) - Programa de Pós-graduação em Ciência da Computação, Universidade Federal de Santa Catarina, Florianópolis.

CAROTHERS, G. **RDF 1.1 N-Quads: A line-based syntax for RDF datasets**, 2014. Disponível em: <<http://www.w3.org/TR/n-quads/>>.

CORCHO, O.; CIMA, A. L.; GÓMEZ-LÓPEZ, A. **A platform for the development of semantic web portals**. In: Proceedings of the 6th international conference on Web engineering. Palo Alto, Estados Unidos, 2006.

CORLOSQUET, S. et al. **Produce and Consume Linked Data with Drupal**. In: 8th International Semantic Web Conference. Washington, United States of America, 2009.

CRISTINA, R.F.M.; OLIVEIRA, S.P.; MARCON, S.S. **Fatores associados ao abandono do tratamento de tuberculose no estado do Paraná**. Acta Paul Enferm. 2012; 25: 108-14.

DANNENBERG, A.M.Jr. **Immune mechanisms in the pathogenesis of pulmonary tuberculosis**. RvInfec Dis. 1989; 11(Suppl 2): S369-578.

DAS, S. et al. **Building biomedical web communities using a semantically aware content management system**. Briefings in Bioinformatics, v. 10, n. 2, p. 129-138, 2009.

DEVEDZIC, V. **Semantic Web and Education**. Berlin: Springer, 2006. 354 p.

DILL, S. et al. **SemTag and Seeker: Bootstrapping the semantic web via automated semantic annotation**. Proceedings of the 12th international conference on World Wide Web, v. 1, n. Section 2, p. 178–186, 2003.

DIMITROV, M. **Introduction to Semantic Technologies, Ontologies and the Semantic Web, 2010**. Disponível em: <https://pt.slideshare.net/marin_dimitrov/introduction-to-th>.

DING, Y. et al. **Semantic Web Portal: A Platform for Better Browsing and Visualizing Semantic Data**. In: International Conference on Active Media Technology. Toronto, Canada, 2010.

DING, Y.; KLEIN, K. **Model-Driven Application-Level Encryption for the Privacy of E-Health Data**. In: International Conference on Availability, Reliability and Security, Krakow, Poland, 2010.

DUBOST, K.; HERMAN, I. **State of the Semantic Web**. In: INTAP Semantic Web Conference, Tokyo, Japan, 2008

DUMONTIER, M. **Building an effective Semantic Web for Health Care and the Life Sciences**. *Semantic Web*, v. 1, n. 1,2, p. 131-135, 2010.

DÜRST, M. J. **Internationalized Resource Identifiers: From Specification to Testing**. In: 19th International Unicode Conference. San Jose, United States of America, 2001.

EUZENAT, J.; ISAAC, A.; MEILICKE, C. S.; STUCKENSCHMIDT, H.; SVAB, O.; SVATEK, V. ET AL. **First results of the Ontology Alignment Evaluation Initiative 2007**. Second International Workshop on Ontology Mapping. IEEE Xplore. 2007

FREITAS, F. **Ontologias e Web Semântica**. In: Anais do XXIII Congresso da Sociedade Brasileira de Computação, Campinas, Brasil, 2003.

FREITAS, F.; SCHULZ, S. **Ontologias, Web semântica e saúde**. *Revista Eletrônica de Comunicação, Informação & Inovação em Saúde*, v. 3, n. 1, p. 4-7, 2009.

GANDON, F.; SCHREIBER, G. **RDF 1.1 XML Syntax, 2014**. Disponível em: <<http://www.w3.org/TR/rdf-syntax-grammar/>>.

GANGWAR, M.; YADAV, R. S.; MISHRA, R. B. **Semantic Web Services for Medical Health Planning**. In: IEEE 1st International Conference on Recent Advances in Information Technology. Dhanbad, India, 2012.

GANSEL X.; MARY M.; VAN BELKUM A. **Semantic data interoperability, digital medicine, and e-health in infectious disease management: a review**. *Eur J Clin Microbiol Infect Dis*. 2019 doi: 10.1007/s10096-019-03501-6.

GIANNOPOULOU, E. et al. **A Semantic Web Approach in the Implementation of a Linked Data Portal Using a CMS**. In: Proceedings of the Tenth International Conference on

GONÇALVES, B.; GUIZZARDI, G.; PEREIRA FILHO, J. G. **Using an ECG reference ontology for semantic interoperability of ECG data**. *Journal of Biomedical Informatics*, v. 44, n. 1, p. 126–136, 2011.

GONÇALVES, H. **Peste branca: um estudo antropológico sobre a tuberculose**. Edição 1. Porto Alegre: Editora da UFRGS, 2002.

GRENON, P.; **Spatio-temporality in Basic Formal Ontology: SNAP and SPAN**, Upper-

Level Ontology, and Framework of Formalization (Part I), IFOMIS Technical Report Series, 2003.

GRUBER, T. R. **A Translation Approach to Portable Ontology Specifications**. Knowledge Acquisition, v. 5, n. 2, p. 199-220, 1993.

HADZIC, M.; CHEN, M.; DILLON, T. S. **Towards Mental Health Ontology**. In: IEEE International Conference on Bioinformatics and Biomedicine. Philadelphia, USA, 2008.

HARGREAVES, J.R.; BOCCIA, D.; EVANS, C.A.; ADATO, M.; PETTICREW, M.; PORTER, J.D.H. **The Social Determinants of Tuberculosis: From Evidence to Action**. Amer J Public Health.2011; 101(4): 654–662.

HAVLIK, D. **Building Environmental Semantic Web Applications with Drupal**. In: **9th International Symposium on Environmental Software Systems**. Brno, Czech Republic, 2011.

HERMAN, I. et al. **RDFa 1.1 Primer – 3rd edition**. W3C, 2015. Disponível em: <<http://www.w3.org/TR/2015/NOTE-rdfa-primer-20150317/>>.

HIJJAR, M.A.; GERHARDT, G.M.; TEIXEIRA, G.M.; PROCÓPIO, M.J. **Retrospecto do controle da tuberculose no Brasil**. Rev. Saúde Pública. 2007; 41(Supl.1):50-58.

HITZLER, P.; JANOWICZ, K. **Editorial: Semantic Web - Interoperability, Usability, Applicability**. Semantic Web, v. 1, n. 1–2, p. 1–2, 2010.

HOPEWELL, F.C. **Tuberculosis and other mycobacterial diseases**. In: Mason RJ et al. Textbook of Respiratory Medicine. 4th Ed, Philadelphia. Saunders 2005, v.1, c.33, pp.979-1043.

HORROCKS, I. et al. **Semantic web architecture: Stack or two towers?** In: Proceedings of the 3rd International Workshop on Principles and Practice of Semantic Web Reasoning. Dagstuhl Castle, Germany, 2005.

influencing treatment default among tuberculosis patients in a high burden province of South Africa. Int J InfectDiseases. 2017; 54: 95–102.

ISAAC, A.; SUMMERS, E. **SKOS Simple Knowledge Organization System Primer**, 2009. Disponível em: <<https://www.w3.org/TR/skos-primer/>>.

ISLAM, N. et al. **Semantic Web: Choosing the right methodologies, tools and standards.** In: Proceedings of 2010 International Conference on Information and Emerging Technologies. Karachi, Pakistan, 2010.

ISOTANI, S.; BITTENCOURT, I. I. **Dados abertos conectados.** São Paulo: Novatec, 2015. 177p.

ISOTANI, S. et al. **Estado da Arte em Web Semântica e Web 2.0: Potencialidades e Tendências da Nova Geração de Ambientes de Ensino na Internet.** Revista Brasileira de Informática na Educação, v. 17, p. 30-42, 2009.

J. C.; ALVES, D.; **Development and evaluation of an interoperable system based on the semantic web to enhance the management of patients' tuberculosis data.** Procedia Computer Science, Volume 121, 2017. <https://doi.org/10.1016/j.procs.2017.11.102>.

JAFARPOUR, B. et al. **Exploiting Semantic Web Technologies to Develop OWL-Based Clinical Practice Guideline Execution Engines.** IEEE Journal of Biomedical and Health Informatics, v. 20, n. 1, p. 388-398, 2016.

JAVED, M.; ABGAZ, Y. M.; PAHL, C. **Ontology Change Management and Identification of Change Patterns.** Journal on Data Semantics, v. 2, n. 2-3, p. 119-143, 2013.

JIANG, G.; SOLBRIG, H. R.; CHUTE, C. G. **Using Semantic Web technology to support icd- 11 textual definitions authoring.** Journal of Biomedical Semantics, v. 4, n. 1, p. 1-9, 2013.

KATAYAMA, T. et al. **The 3rd DBCLS BioHackathon: improving life science data integration with Semantic Web technologies.** Journal of Biomedical Semantics, v. 4, n. 1, p. 6, 2013.

KIFER, M. et al. **A realistic architecture for the semantic web.** In: Proceedings of the Rules and Rule Markup Languages for the Semantic Web, Galway, Ireland, 2005.

KIFER, M.; BOLEY, H. **RIF Overview (Second Edition),** 2013. Disponível em: [<https://www.w3.org/TR/rif-overview/>](https://www.w3.org/TR/rif-overview/).

KIGOZI, G.; HEUNIS, C.; CHIKOBVU, P.; BOTHA, S.; VAN RENSBRUG, D. **Factors**

KOTIS, K.; LANZENBERGER, M. **Ontology Mapping: Current status, dilemmas and future challenges**. International Conference on Complex, Intelligent and software Intensive Systems. IEEE Xplore. 2008

KRITSKI, A. L., RUFFINO-NETTO, A.; **Health sector reform in Brazil: impact on tuberculosis control**. The international journal of tuberculosis and lung disease, 2020.

KRITSKI, Afranio et al. **Brazilian Response to Global End TB Strategy : The National Tuberculosis Research Agenda**. Rev. Soc. Bras. Med. Trop., Uberaba , v. 49, n. 1, p. 135-145, Feb. 2016 .

KUMAR, V. K. **Semantic Web Approach Towards Interoperability and Privacy Issues**. v. 5, n. 3, p. 13–17, 2014.

LACHTIM, F. A; MOURA, A. M. C.; CAVALCANTI, M. C. **Ontology matching for dynamic publication in semantic portals**. Journal of the Brazilian Computer Society, v. 15, n. 1, p. 27- 43, 2009.

LANNOY, L.H.; CORTEZ-ESCALANTE, J.J.; EVANGELISTA, M.S.; ROMERO, G.A. **Tuberculosis incidence and risk factors among patients living with HIV/AIDS in public health service institutions in Brasilia, Federal District**. Revista da Sociedade Brasileira de Medicina Tropical. Nov-dez 2008; 41(6):549-555.

LARA, R et al. **An Evaluation of Semantic Web Portals**. In: Applied Computing International Conference 2004. Lisbon, Portugal, 2004.

LAUFER, C. **Guia Web Semantica**. p. 133, 2015.

LAUSEN, H. et al. **Semantic web portals: state-of-the-art survey**. Journal of Knowledge Management, v. 9, n. 5, p. 45-49, 2005.

LEE, M. et al. **Semantic association-based search and visualization method on the semantic web portal**. International Journal of Computer Networks & Communications, v. 2, n. 1, p. 140- 152, 2010.

LEI, Y.; LOPEZ, V.; MOTTA, E. **An Infrastructure for Building Semantic Web Portals**. In: International Workshop on Web Information Systems Modeling (WISM 2006). Luxemburgo, Luxemburgo, 2006.

LIANG, C.; SUN, J.; TAO, C. **Semantic Web Ontology and Data Integration: a Case Study in Aiding Psychiatric Drug Repurposing**. Studies in Health Technology and Informatics, v. 216, p. 1051, 2015.

LIE, X.-J.; WANG, Y.-L.; & WANG, J. **Towards A Semi- Automatic Ontology Mapping – An Approach Using Instance Based Learning and Logic Relation Mining**. Proceedings of the Fifth Mexican International Conference on AI. IEEE Xplore. 2006

LIMA, J. C.; CARVALHO, C. L. **Uma visão da Web Semântica**. Goiânia: Instituto de Informática - Universidade Federal de Goiás, 2004. 16 p. Relatório técnico, INF_001/94.

LIPING, Z.; GUANGYAO, L.; SHAJING; YONGQUM, L. **Design of Ontology Mapping Framework**. CIMCA-IAWTI'06. IEEE Xplore. 2006

LOPES, I. L. **Novos paradigmas para avaliação da qualidade da informação em saúde recuperada da Web**. Ciência da Informação, v. 33, n. 1, p. 81-90, 2004.

LOPES, P.; OLIVEIRA, J. L. **A semantic web application framework for health systems interoperability**. Proceedings of the first international workshop on Managing interoperability and complexity in health systems - MIXHS '11, p. 87, 2011.

LOPES, P.; OLIVEIRA, J. S. **A Semantic Web Application Framework for Health Systems Interoperability**. In: Proceedings of the first international workshop on Managing interoperability and complexity in health systems. Glasgow, Scotland, 2011.

LYTRAS, M. D.; SAKKOPOULOS, E.; PABLOS P. A. **Semantic Web and Knowledge Management for the health domain: state of the art and challenges for the Seventh Framework Programme (FP7) of the European Union (2007-2013)**. International Journal of Technology Management, v. 47, n. 1/2/3, 2009.

MACHADO, C. M et al. **The semantic web in translational medicine: current applications and future directions**. Briefings in Bioinformatics, v. 16, n. 1, p. 89-103, 2013.

MÄKELÄ, E. et al. **OntoViews - A Tool for Creating Semantic Web Portals**. In: **3rd International Semantic Web Conference**. Hiroshima, Japan, 2004.

MALHÃO, T. A. et al. **Avaliação da completitude do Sistema de Informação de Agravos de Notificação da Tuberculose, Brasil, 2001-2006** Epidemiologia e Serviços de Saúde scielo , , 2010.

MCMURRAY, J. et al. **Ontological modeling of electronic health information exchange.** Journal of Biomedical Informatics, v. 56, p. 169–178, 2015.

MENDES, A. DE M.; FENSTERSEIFER, L. M. **Tuberculose: porque os pacientes abandonam o tratamento?** Boletim de Pneumologia Sanitária scielo , , 2004.

MINISTÉRIO DA SAÚDE. DEPARTAMENTO DE VIGILÂNCIA DAS DOENÇAS TRANSMISSÍVEIS. **"Brasil Livre da Tuberculose : Plano Nacional pelo Fim da Tuberculose como Problema de Saúde Pública"** – Brasília, 2017.

MINISTÉRIO DA SAÚDE. SECRETARIA DE VIGILÂNCIA EM SAÚDE. DEPARTAMENTO DE VIGILÂNCIA EPIDEMIOLÓGICA. **Manual de recomendações para o controle da tuberculose no Brasil.** Brasília, DF, 2011.

MODESTO, F. **A pirâmide da web semântica para bibliotecários não egípcios, 2010.** Disponível em: <http://www.ofaj.com.br/colunas_conteudo.php?cod=571>.

MOHAMMED, O.; BENLAMRI, R. **Developing a semantic web model for medical differential diagnosis recommendation.** J Med Syst., v. 38, n. 10, p. 79, 2014.

MOREIRA, Adriana da Silva Rezende. **O Processo de Criação e Implantação do Programa de Controle de Tuberculose Hospitalar - Hospital Universitário Clementino Fraga Filho/RJ. Rio de Janeiro, 2010.** Dissertação (Mestrado em Ciências da Saúde) – Faculdade de Medicina, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2010.

MOREIRA, J.S. **Patogenia da tuberculose.** In Corrêa da Silva LC et al. Pneumologia, Princípios e Prática. Porto Alegre: Artmed, 2012. c. 39,p 314-17.

NOWACK, B. **The Semantic Web Technology Stack (not a piece of cake...), 2009.** Disponível em: <<http://linkeddatadeveloper.com/Projects/Linked-Data/media/fig11.2.png>>.

NOY, N. **p42-noy[1].pdf.** Retrieved April 2019, from www.acmqueue.com. 2005

NUNO, S., & JOAO, R.. **MAFRA – An Ontology Mapping FRAMework for the Semantic Web.** Proceedings of the 6th International Conference on Business information Systems. UCCS, Colorado Springs, CO. 2003

OBRST, L.; JANSSEN, T.; CEUSTERS, W. **Ontologies and Semantic Technologies for Intelligence**. Amsterdam: IOS Press, 2010, 236p.

OGUNDELE, O. A. et al. **An ontology for factors affecting tuberculosis treatment adherence behavior in sub-Saharan Africa**. Patient Preference and Adherence, v. 10, p. 669–681, 2016.

OGUYA, F. **Factors associated with default from treatment among tuberculosis patients in nairobi province, Kenya: A case control study**. BMC Public Health. Sept 2011; 11:696.

OLIVEIRA, H.B.; FILHO, D.C.M. **Abandono de tratamento e recidiva da tuberculose: aspectos de episódios prévios, Campinas, SP, Brasil, 1993-1994**. Rev Saúde Pública. 2000; 34(5):437- 43.

PEINL, R. **Semantic Web: State of the Art and Adoption in Corporations**. Künstliche Intelligenz, v. 30, n. 2, p. 131-138, 2016.

PELLISON, F. C.; RIJO, R. P. C. L.; LIMA, V. C; CREPALDI, N. Y.; BERNARDI, F. A.; GALLIEZ, R. M.; KRITSKI, A.; ABHISHEK, K.; ALVES, D.; **Data Integration in the Brazilian Public Health System for Tuberculosis: Use of the Semantic Web to Establish Interoperability**. JMIR MEDICAL INFORMATICS, v. 8, p. e17176, 2020.

PELLISON, F.C.; LIMA, V.C; RIJO, R.P.C.L.; ALVEZ, D; **Integrating Tuberculosis data in State of São Paulo over Semantic Web: a proof of concept**. Procedia Computer Science, Volume 164, 2019. <https://doi.org/10.1016/j.procs.2019.12.236>.

PICON, P.D.; RIZZON, C.F.; OTT, P.W. **Tuberculose: epidemiologia, diagnóstico e tratamento em clínica e saúde pública**. Rio de Janeiro: Medsi, 1993, 543-59.

PINOCHET, L. H. C. **Tendências de Tecnologia de Informação na Gestão em Saúde. O Mundo da Saúde**, v. 35, n. 4, p. 382-394, 2011.

PLASTIRAS, P.; O’SULLIVAN, D. M. **Combining Ontologies and Open Standards to Derive a Middle Layer Information Model for Interoperability of Personal and Electronic Health Records**. Journal of Medical Systems, v. 41, n. 12, p. 1–15, 2017.

PRUD’HOMMEAUX, E.; CAROTHERS, G. **RDF 1.1 Turtle: Terse RDF Triple Language**. W3C, 2014. Disponível em: <<http://www.w3.org/TR/turtle/>>.

PRUD'HOMMEAUX, E.; SEABORNE, A. **SPARQL Query Language for RDF**. W3C, 2008. Disponível em: <<https://www.w3.org/TR/rdf-sparql-query/>>.

RAMALHO, R. A. S.; VIDOTTI, S. A. B. G.; FUJITA, M. S. L. **Web semântica: uma investigação sob o olhar da Ciência da Informação**. DataGramaZero - Revista de Ciência da Informação, v. 8, n. 6, p. 1-18, 2007.

RICH, A.R. **The Pathogenesis of Tuberculosis**. Springfield. CC Thomas, San Martin 1944.

RICO, M. et al. **A Tool Suite to Enable Web Designers, Web Application Developers and End- users to Handle Semantic Data**. International Journal on Semantic Web and Information Systems, v. 6, n. 3, p. 38-60, 2010.

RILEY RL, MILLS CC, O'GRADY F, SULTAN LU, WITTSTADT F, SHIVPURI DN. **Infectiousness of air from a tuberculous ward Ultraviolet Irradiation of Infected Air: Comparative Infectiousness of Different Patients**. AmRevRespirDis. 1961, 85(4):511-525.

RIMNÁC, M. **OntoViews - Semantic web portal using remarks as RDF data**. In: **Proceedings of the 12th International Conference on Information Integration and Web-based Applications & Services**. Paris, France, 2010.

ROBINSON, I.; WEBBER, J.; EIFREIM, E. **Graph Databases, New Opportunities for Connected Data**. Sebastopol: O'Reilly Media, 2015. 238p.

ROBU, I.; ROBU, V.; THIRION, B. **An introduction to the Semantic Web for health sciences librarians**. Journal of the Medical Library Association : JMLA, v. 94, n. 2, p. 198–205, 2006.

ROVAN, L. **Realizing Semantic Web Portal Using Available Semantic Web Technologies and Tools**. In: 7th International Semantic Web Conference (ISWC). Karlsruhe, Germany, 2008.

RUFFINO-NETTO, A.; VILLA, T.C.S. **Tuberculose: implantação do DOTS em algumas regiões do Brasil histórico e peculiaridades regionais**. Instituto do Milênio REDE TB. Ribeirão Preto, 2006.

SAMPSON, J. **Ontology Alignment in Agent Systems: Current and future challenges**. Proceeding of the 2005 International Conference on Computational Intelligence for Modeling, Control and Automation. IEEE Xplore. 2005

SCHENKMAN, S. **Estudo caso-controle de indicadores de abandono em doentes com tuberculose.** J Pneumol [online]. 2000; vol.26, n.6, pp.291-296.

SCHWEIGER, D.; TRAJANOSKI Z.; PABINGER, S. **SPARQLGraph: a web-based platform for graphically querying biological Semantic Web databases.** BMC Bioinformatics, v. 15, p. 279, 2014.

SECRETARIA DE VIGILÂNCIA EM SAÚDE; MINISTÉRIO DA SAÚDE. **Boletim Epidemiológico: Detectar, tratar e curar: desafios e estratégias brasileiras frente à tuberculose.** Volume 46 N° 9; 2015.

SECRETARIA DE VIGILÂNCIA EM SAÚDE; MINISTÉRIO DA SAÚDE. **Boletim Epidemiológico de Tuberculose.** 1ª Edição, 2020.

SECRETARIA DE VIGILÂNCIA EM SAÚDE; MINISTÉRIO DA SAÚDE. **Boletim Epidemiológico de Tuberculose.** 1ª Edição, 2021.

SECRETARIA MUNICIPAL DE SAÚDE DE PORTO ALEGRE. **Boletim epidemiológico, No 58.** Programa Municipal de Controle da Tuberculose. Porto Alegre, 2015.

SERNADELA, P.; GONZÁLEZ-CASTRO, L.; OLIVEIRA, J. S. **SCALEUS: Semantic Web Services Integration for Biomedical Applications.** Journal of Medical Systems, v. 41, p. 54, 2017.

SHABAN-NEJAD, A. et al. **PopHR: a knowledge-based platform to support integration, analysis, and visualization of population health data.** Annals of the New York Academy of Sciences, v. 1387, n. 1, p. 44-53, 2017.

SHARMA, D. K. et al. **Building a semantic web-based metadata repository for facilitating detailed clinical modeling in cancer genome studies.** Journal of Biomedical Semantics, v. 8, p. 19, 2017.

SHVAIKO, P.; EUZENAT, J. **Tutorial on Schema and Ontology Matching.** ESWC'05. 2005.

SILVA, C.C.A.V.; ANDRADE, M.S.; CARDOSO, M.D. **Fatores associados ao abandono do tratamento de tuberculose em indivíduos acompanhados em unidades de saúde de referência na cidade do Recife, Estado de Pernambuco, Brasil, entre 2005 e 2010.**

Epidemiol. Serv. Saúde. Brasília. 2013; 22(1):77-85.

SILVA, S. Y. B. E; MEDEIROS, E. R. DE; SILVA, S. B. DA; ANDRADE, R. P. DA S.; BERALDO, A. A.; PINTO, ÉRIKA S. G. **Facilities and difficulties in implementation of the tuberculosis control program in Primary Health Care**. ABCS Health Sciences, v. 46, p. e021204, 15 Jan. 2021.

SMITH, M.; WELTY, C.; MCGUINNESS, D. **OWL Web Ontology Language Guide**. Recommendation . 2004

SONG, T. M.; PARK, H. A.; JIN D. L. **Development of Health Information Search Engine Based on Metadata and Ontology**. Healthcare Informatics Research, v. 20, n. 2, p. 88-98, 2014.

SOUZA, R. R.; ALVARENGA, L. **A Web Semântica e suas contribuições para a ciência da informação**. Ciência da Informação, v. 33, n. 1, p. 132-141, 2004.

SPEAR, A. **Ontology for the Twenty First Century: An Introduction with Recommendations**. Technical report, University at Buffalo. 2007

SPERONI, R. M. et al. **Usando dados ligados na representação de indicadores da inovação regional**. Navus - revista de gestão e tecnologia, v. 7, n. 3, p. 95-103, 2017.

STOLLBERG, M. et al. **Toward Semantic Web Portals**. In: Proceedings of the WWW2004 Workshop on Application Design, Development and Implementation Issues in the Semantic Web. New York, USA, 2004.

STORY, A.; OTHAMLEY, G.; HAYWARD, A. **Crack Cocaine and Infectious Tuberculosis**. *Emerging Infectious Diseases*, 2008; 14(9):1466-69.

STUDER, R. BENJAMINS, AND D. FENSEL. **Knowledge engineering: Principles and methods**. Data & Knowledge Engineering, 25(1-2):161-198, 1998

TARANTINO, A.B. et al. **Doenças Pulmonares. Capítulo 17: Tuberculose. Sexta edição**. Editora Guanabara, 2008. p. 266-328.

TESFAHUNEYGN, G.; MEDHIN, G.; LEGESSE, M. **Adherence to Anti-tuberculosis treatment and treatment outcomes among tuberculosis patients in Alamata District, northeast Ethiopia**. BMC Res Notes. Setembro, 2015; 8:503.

TILAHUN, B. **Design and Development of a Linked Open Data-Based Health Information Representation and Visualization System: Potentials and Preliminary Evaluation.** JMIR Medical Informatics, v. 2, n. 2, p. e31, 2014.

TILAHUN, B. **Potential of Linked Open Data in Health Information Representation on the Semantic Web.** In: Poster at International SWAT4LS Workshop - Semantic Web Applications and Tools for Life Sciences. Paris, France, 2012.

TIMM, J.; RENLY, S.; FARKASH, A. **Large Scale Healthcare Data Integration and Analysis using the Semantic Web.** Studies in Health Technology and Informatics, v. 169, p. 729-733, 2011.

TJOA, A. M.; ANDJOMSHOAA, A.; SHAYEGANFAR, F.; WAGNER, R. **Semantic Web Challenges and New Requirements.** Proceeding of the 16th International Workshop on Database and Expert Systems Applications (DEXA 05). 2005

VALLE, E. DELLA et al. **The Need for Semantic Web Service in the eHealth.** First International Workshop on Mediation in Semantic Web Services (MEDIATE 2005), p. 35–50, 2005.

VAN WOENSEL W. et al. **A Comparison of Mobile Rule Engines for Reasoning on Semantic Web Based Health Data.** In: IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technologies. Warsaw, Poland, 2014.

VANDERVALK, B.; MCCARTHY, L.; WILKINSON, M. **CardioSHARE: Web Services for the Semantic Web.** Semantic Web Challenge, v. 16, 2008.

VARGAS-VERA, M. et al. **MnM: A Tool for Automatic Support on Semantic Mark-up.** In: Proceedings of the 13th International Conference on Knowledge Engineering and Knowledge Management. Siguenza, Spain, 2002.

VARGAS-VERA, M. et al. **MnM: Ontology driven semi-automatic and automatic support for semantic markup.** Knowledge Engineering and Knowledge Management. Methods, Models, and Tools: Ontologies and The Semantic Web, v. 2473, p. 213–221, 2002.

WORLD HEALTH ORGANIZATION (WHO). **Anti-tuberculosis Drug Resistance in the World. Report no. 2: Prevalence and Trends.** Geneva, 2000.

WORLD HEALTH ORGANIZATION (WHO). **Global tuberculosis report 2020**. Geneva, 2020.

WORLD HEALTH ORGANIZATION (WHO). **Guidelines for surveillance of drug resistance in tuberculosis**. 2015.

WORLD HEALTH ORGANIZATION (WHO). **TB, a Global Emergency**, 1996.

WORLD HEALTH ORGANIZATION (WHO). **The end TB strategy**. 2014.

WORLD HEALTH ORGANIZATION (WHO). **The Stop TB Strategy: Building on and enhancing DOTS to meet the TB-related Millennium Development Goals**. WHO/HTM/TB/ 2006.

WORLD HEALTH ORGANIZATION (WHO). **World Health Statistics**. Geneva: World Health Organization, 2013.

WORLD HEALTH ORGANIZATION; **Implementing the WHO Stop TB Strategy: A Handbook for National Tuberculosis Control Programmes**. Geneva, 2008.

YAGUE, M. I. et al. **Applying the semantic Web layers to access control**. In: Proceedings of the 14th International Workshop on Database and Expert Systems Applications. Prague, Czech Republic, 2003.

YANG, Z. et al. **Development and validation of an instrument to measure user perceived service quality of information presenting Web portals**. Information & Management, v. 42, n. 4, p. 575- 588, 2005.

YOSHIURA, V. T. **Desenvolvimento de um modelo de observatório de saúde baseado na web semântica: o caso da rede de atenção psicossocial**. 2020. 251f. Tese (Doutorado) - Programa de Pós-Graduação em Saúde Pública, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2020.

YU, L. **A Developer's Guide to the Semantic Web. Segunda edição**. Berlim: Springer, 2014. 629p.

YUKIE, N. et al. **DESENVOLVIMENTO E IMPLANTAÇÃO DE UM SISTEMA PARA GESTÃO DE PACIENTES DE TUBERCULOSE**. n. January 2014, p. 1–8, 2007.

ZAHRA, E., NASSER, Y., & FARHAD, O. (2007). **A Vector Based Method of Ontology Mapping. Third International Conference on Semantics, Knowledge and Grid.** IEEE Computer Society, IEEE Xplore.

ZAVERI, A. et al. **Quality assessment for Linked Data: A Survey.** Semantic web, v. 7, p. 63- 93, 2016.

ZAVERI, A. et al. **ReDD-Observatory: Using the Web of Data for Evaluating the Research Disease Disparity for Emerging Regions.** In: Proc. of the IEEE/WIC/ACM International Conference on Web Intelligence. Lyon, France, 2011.

ZENUNI, X.; et al. **State of the Art of Semantic Web for Healthcare.** Procedia - Social and Behavioral Sciences, v. 195, p. 1990-1998, 2015.

ZOLHAVARIEH, S.; PARRY D.; BAI Q. **Issues Associated with the Use of Semantic Web Technology in Knowledge Acquisition for Clinical Decision Support Systems: Systematic Review of the Literature.** JMIR Medical Informatics, v. 5, n. 3, p. e18, 2017.

8. APÊNDICE A – PUBLICAÇÕES

Durante o período do doutorado, foram feitas dez publicações, três delas como primeiro autor. As publicações estão direta ou indiretamente ligadas à esta tese e serão apresentadas a seguir, ordenadas cronologicamente e com seus respectivos *abstracts*:

LIMA, V.C.; RIJO, R. P. C. L.; PELLISON, F. C.; LIMA, R. R.; CORREIA, R. J. C.; GIURIATI, H. T.; ALVES, D. From guidelines to decision-making: using mobile applications and semantic web in the practical case of guides to support patients. *Procedia Computer Science*, Volume 121, 2017. <https://doi.org/10.1016/j.procs.2017.11.104>.

Abstract

The use of guidelines, procedures and recommendations is very common and also a simple way to structure, organize and retrieve relevant information that, in turn, may impact on patient care. However, it is necessary to expedite the retrieval of stored information. By simplifying the access to the desired data, it is possible to make the process practical, fast and intuitive for the health professional. The main goal of this article is to present a development approach for a project that proposes decision-making tools, based on guidelines, to assist health professionals in the care process of patients with specific diseases (or a combination of them). These tools will be developed for use on mobile devices and as a web-based system, which will transform a passive and declarative manual into an interactive tool that will provide not only data and information, but also knowledge through semantic web paradigm. Further, we will show an example of usability and feasibility of the proposed application through a functional prototype developed for Android platform.

PELLISON, F.C.; RIJO, R.P.C.L.; LIMA, V.C; LIMA, R. R.; MARTINHO, R.; CORREIA, R. J. C.; ALVES, D.; Development and evaluation of an interoperable system based on the semantic web to enhance the management of patients' tuberculosis data. *Procedia Computer Science*, Volume 121, 2017. <https://doi.org/10.1016/j.procs.2017.11.102>.

Abstract

Tuberculosis (TB), despite all the efforts and progress made in the last decade, continues to be very relevant in Brazil as well as in other Portuguese-speaking countries, so there is the need to join efforts for increasing the effectiveness of the fight against TB. Therefore, the epidemiological surveillance system of TB requires, not only the implementation of basic prevention and care actions, but also the strengthening of the integration between the different health services, programs and levels of care, whose resolvability varies according to financial, technical and human resources, as well as service infrastructure that comprise the network of attention. The deepening and broadening of data management techniques must constantly be carried out to achieve, at a higher level, integration and integrity of the ideal and desirable health care. Then, the goal of this article is to describe the research methods used to develop an information system using systems interoperability techniques based on the Semantic Web paradigm that will contribute to the activities of epidemiological surveillance and the follow-up of TB patients.

LIMA, V.C; ALVES, D.; **PELLISON, F.C.**; YOSHIURA, V. T.; CREPALDI, N. Y.; RIJO, R.P.C.L.; Establishment of Access Levels for Health Sensitive Data Exchange through Semantic Web, *Procedia Computer Science*, Volume 138, 2018. <https://doi.org/10.1016/j.procs.2018.10.027>.

Abstract

Data exchange in health information systems must be carefully planned and needs to be protected from unauthorized access due to sensibility of stored content. Security aspects like authentication, authorization and encryption must be considered in this context. The main goal of this article is to present the implementation of security mechanisms to a semantic API that allows data extraction from a regional health information system designed to create notifications and to follow patients diagnosed with Tuberculosis. Data semantically tagged will be mapped individually to several access levels. It will be showed how external systems can connect, authenticate and retrieve only authorized data that are classified in the scope of its maximum access level.

YOSHIURA, V. T.; YAMADA, D. B.; PELLISON, F. C.; LIMA, I. B.; DAMIAN, I. P. M.; RIJO, R. P. C.; AZEVEDO-MARQUES, J. M.; ALVES, D. Towards a health observatory conceptual model based on the semantic web. *Procedia computer science*, v. 138, p. 131-136, 2018. <https://doi.org/10.1016/j.procs.2018.10.019> .

Abstract

Introduction. Health Observatories have gained global popularity and have been established worldwide since the 1970s. With the advent of the Semantic Web, machines can process, reuse, integrate and understand the meaning (semantics) of the information and knowledge on the World Wide Web to perform complex tasks. Objective. To propose a health observatory conceptual model based on Semantic Web in order to assist in the design, development and implementation processes of a Health Observatory. Methods. The proposed model was based on a conceptual analysis that include Semantic Web technologies narrative review, multi-layer software architecture and an integrative review of Health Observatories. Results. The proposed Health Observatory conceptual model consists of a chain of several related components based on information technology multi-layer architecture, Semantic Web technologies, Health Observatories stakeholders and key concepts. Conclusion. The proposed model can provide opportunities for the development and implementation of new observatories or for the adequacy of existing Health Observatories.

CREPALDI, N. Y.; LIMA, V. C.; BERNARDI, F. A; SANTOS, L. R. A.; YAMAGUTI, V. H.; PELLISON, F. C.; SANCHES, T. L. M.; MIYOSHI, N. S. B.; RUFFINO-NETTO, A.; RIJO, R. P. C. L.; ALVES, D.; SISTB: an ecosystem for monitoring TB. *Procedia Computer Science*, Volume 164, 2019. <https://doi.org/10.1016/j.procs.2019.12.224>.

Abstract

Tuberculosis (TB) is a curable disease, although it still kills thousands of people every year. There are several information systems and applications developed to assist the TB treatment, but many of these systems are isolated and are not able to exchange relevant information. In this way, a set of interoperable applications and decision support modules have been developed. A socio-technical approach was applied to incorporate key stakeholders in all development stages so that final tools fit the health service reality. A set of systems and applications, called the SISTB Ecosystem, has been designed to assist in the treatment of TB and to improve the routine of health professionals.

LIMA, V.; PELLISON, F.; BERNARDI, F.; CARVALHO, I.; RIJO, R.; ALVES, D.; Proposal of an integrated decision support system for Tuberculosis based on Semantic Web, *Procedia Computer Science*, Volume 164, 2019. <https://doi.org/10.1016/j.procs.2019.12.219>.

Abstract

Epidemiological surveillance of Tuberculosis (TB) requires a strong integration of different health services, programs and levels of care. The deepening and broadening of data management techniques must be constantly carried out to increase the integrality of healthcare. Otherwise, knowledge extraction and clinical and administrative decision-making processes are significantly hampered, directly affecting the management and quality of health services. Thus, this work aims to establish a computerized decision support system capable of collecting, integrating and sharing TB health data in Brazilian Unified Public Health System. Also, it will allow the monitoring of infected patients and the visualization of consolidated information of regular TB and its resistant variants for health professionals and managers. The data will be made available from heterogeneous, disconnected and unstructured sources by combining traditional web services, Semantic Web resources and security algorithms. A solid knowledge base applied to epidemiological surveillance, health information governance and clinical support will be enabled to integrate the multiple areas of TB patients care, as well as to support the creation of more accurate operational and diagnostics models.

PELLISON, F.C.; LIMA, V.C; RIJO, R.P.C.L.; ALVEZ, D; Integrating Tuberculosis data in State of São Paulo over Semantic Web: a proof of concept. *Procedia Computer Science*, Volume 164, 2019. <https://doi.org/10.1016/j.procs.2019.12.236>.

Abstract

Although tuberculosis is a curable disease and, in most cases, with low cost drugs, its mortality still is a global concern. This facts turns our attention to management issues and the difficulties related to retrieving data of interest that are powdered on many applications. This work presents a semantic web approach to achieve functional and semantic interoperability between two applications in State of São Paulo that contain tuberculosis data. By combining a theoretical-practical development, the geolocalization tool created is a proof of concept that could help managers to take strategic decisions and develop better health policies by showing the distribution of tuberculosis cases across the state. This work stands out the importance of working in solutions that could improve the quality of data in health field and daily activities of health professionals.

BERNARDI, F. A. ; LIMA, V. C. ; **PELLISON, F. C.** ; MARQUES, P. M. A. ; RIJO, R. P. C. L. ; GALLIEZ, R. ; KRITSKI, A. ; ALVES, D. . Blockchain Based Network for Tuberculosis: A Data Sharing Initiative in Brazil. STUDIES IN HEALTH TECHNOLOGY AND INFORMATICS, v. 262, p. 264-267, 2019.

Abstract

Data sharing, information exchange, knowledge acquisition and health intelligence are the basis of an efficient and effective evidence-based decision-making tool. A decentralized blockchain architecture is a flexible solution that can be adapted to institutional and managerial culture of organizations and services. Blockchain can play a fundamental role in enabling data sharing within a network and, to achieve that, this work defines the high-level resources necessary to apply this technology to Tuberculosis related issues. Thus, relying in open-source tools and in a collaborative development approach, we present a proposal of a blockchain based network, the TB Network, to underpin an initiative of sharing of Tuberculosis scientific, operational and epidemiologic data between several stakeholders across Brazilian cities.

LIMA, V. C. ; BERNARDI, F. A. ; **PELLISON, F. C.** ; RIJO, R. P. C. L. ; KRITSKI, A. ; GALLIEZ, R. ; ALVES, D. . Decision Support Systems for Tuberculosis: Protocol for a Scoping Review. *STUDIES IN HEALTH TECHNOLOGY AND INFORMATICS*, v. 262, p. 101-104, 2019.

Abstract

Tuberculosis (TB) represents a global challenge in terms of prevention, care and control. Decision support systems (DSS) can supply the necessary knowledge basis to underpin investigators, policy makers and health personnel actions and to provide crucial elements that can help reducing TB burden. Thus, the objectives of this work are to present the protocol to be followed for carrying out a scoping review to identify topics where DSSs are used, to define appropriate categories and to clarify main outcomes and research gaps. As part of the protocol, five electronic bibliographic databases will be searched for articles from 2006 to 2019 and two investigators will independently screen each work using the study inclusion criteria. Data extraction will be performed, and findings will be reported. The results will be used to provide a broad understanding of how DSSs for TB are being used.

PELLISON, F. C.; RIJO, R. P. C. L.; LIMA, V. C; CREPALDI, N. Y.; BERNARDI, F. A.; GALLIEZ, R. M.; KRITSKI, A.; ABHISHEK, K.; ALVES, D.; Data Integration in the Brazilian Public Health System for Tuberculosis: Use of the Semantic Web to Establish Interoperability. *JMIR MEDICAL INFORMATICS*, v. 8, p. e17176, 2020.

Abstract

Background: Interoperability of health information systems is a challenge due to the heterogeneity of existing systems at both the technological and semantic levels of their data. The lack of existing data about interoperability disrupts intra-unit and inter-unit medical operations as well as creates challenges in conducting studies on existing data. The goal is to exchange data while providing the same meaning for data from different sources. Objective: To find ways to solve this challenge, this research paper proposes an interoperability solution for the tuberculosis treatment and follow-up scenario in Brazil using Semantic Web technology supported by an ontology. Methods: The entities of the ontology were allocated under the definitions of Basic Formal Ontology. Brazilian tuberculosis applications were tagged with entities from the resulting ontology. Results: An interoperability layer was developed to retrieve data with the same meaning and in a structured way enabling semantic and functional interoperability. Conclusions: Health professionals could use the data gathered from several data sources to enhance the effectiveness of their actions and decisions, as shown in a practical use case to integrate tuberculosis data in the State of São Paulo.