

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE BIBLIOTECONOMIA E COMUNICAÇÃO
DEPARTAMENTO DE CIÊNCIAS DA INFORMAÇÃO
GRADUAÇÃO EM BIBLIOTECONOMIA

Isadora Oliveira Turcatel

ONTOLOGIAS PARA A DESCOBERTA DE RECURSOS NA CIÊNCIA:
análise de VIVO-ISF

Porto Alegre
2014

Isadora Oliveira Turcatel

**ONTOLOGIAS PARA A DESCOBERTA DE RECURSOS NA CIÊNCIA:
análise de VIVO-ISF**

Trabalho de Conclusão de Curso apresentado como requisito parcial para a obtenção de grau de Bacharel em Biblioteconomia, pela Faculdade de Biblioteconomia e Comunicação, da Universidade Federal do Rio Grande do Sul.

Orientador: Dr. Rafael Port da Rocha

Porto Alegre
2014

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Carlos Alexandre Netto

Vice-Reitor: Rui Vicente Opperman

FACULDADE DE BIBLIOTECONOMIA E COMUNICAÇÃO

Diretora: Ana Maria Mielniczuk de Moura

Vice-Diretor: André Iribure Rodrigues

DEPARTAMENTO DE CIÊNCIAS DA INFORMAÇÃO

Chefe: Maria do Rocio Fontoura Teixeira

Chefe-substituto: Valdir José Morigi

COMISSÃO DE GRADUAÇÃO DO CURSO DE BIBLIOTECONOMIA

Coordenadora: Samile Andréa de Souza Vanz

Vice-Cordenadora: Glória Isabel Sattamini Ferreira

CIP - Catalogação na Publicação

Turcatel, Isadora Oliveira
Ontologias para a descoberta de recursos na
ciência: análise de VIVO-ISF / Isadora Oliveira
Turcatel. -- 2014.
84 f.

Orientador: Rafael Port da Rocha.

Trabalho de conclusão de curso (Graduação) --
Universidade Federal do Rio Grande do Sul, Faculdade
de Biblioteconomia e Comunicação, Curso de
Biblioteconomia, Porto Alegre, BR-RS, 2014.

1. Ontologias. 2. Web Semântica. 3. Linked Open
Data. 4. VIVO. 5. VIVO-ISF. I. Rocha, Rafael Port
da, orient. II. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da UFRGS com os
dados fornecidos pelo(a) autor(a).

Departamento de Ciências da Informação

Rua: Ramiro Barcelos, 2705

CEP: 90035-007 – Porto Alegre/RS

Telefone: (051) 3308-5143

E-mail: dci@ufrgs.br

Isadora Oliveira Turcatel

ONTOLOGIAS PARA A DESCOBERTA DE RECURSOS NA CIÊNCIA:
análise de VIVO-ISF

Trabalho de Conclusão de Curso apresentado como requisito parcial para a obtenção de grau de Bacharel em Biblioteconomia, pela Faculdade de Biblioteconomia e Comunicação, da Universidade Federal do Rio Grande do Sul.

Aprovado em: 05 de dezembro de 2014.

BANCA EXAMINADORA

Prof. Dr. Rafael Port da Rocha (orientador)
Universidade Federal do Rio Grande do Sul

Prof^a. Dra. Sônia Elisa Caregnato (examinadora)
Universidade Federal do Rio Grande do Sul

Prof^a. Me. Rita do Carmo Ferreira Laipelt (examinadora)

*Dedico este trabalho a todos aqueles
que ecoam em mim*

AGRADECIMENTOS

Então, chegou a hora....folha em branco, olhos marejados e dedos inseguros, mas vamos lá!

Pai, obrigada pelo dom da vida, pela dádiva da reencarnação. Obrigada por me amar incondicionalmente apesar de toda minha imperfeição.

Meu anjo protetor, obrigada por zelar meus passos e meu sono.

Aos amigos espirituais, agradeço a proteção e as boas energias.

Mãe, é impossível agradecer em poucas linhas tudo o que a sra fez por mim e continua fazendo. Obrigada por me dar o privilégio de ser sua filha. Meu amor eterno!

Pai, apesar de não estar fisicamente comigo aqui na Terra, sei que o sr. sempre me cuida, olha e protege. O que eu mais quero nessa vida é que o sr sinta orgulho de mim. Um dia vamos nos reencontrar! Meu segundo amor eterno!

Minha querida família, obrigada por tudo! Pelo carinho, amor, confiança e respeito! Irmã, vó, vô, nonna, nonno, tios, tias e primos.

Elo (negri), minha irmã de coração, amada minha! Juntas desde o primeiro semestre até o final! Não gosto nem de pensar como minha vida vai ser daqui pra frente...

Mirela, minha amiga iluminada, meu presente no final do curso! Nos conhecemos há pouco tempo, mas parece que faz uma eternidade! Muito obrigada por tudo!

Geise, outro presentinho! Picorrucha, obrigada por me ajudar a segurar a barra nesses últimos meses! Minha grande companheira de eventos!

Raquel, Rose, Leandro! Obrigada por fazerem parte da minha vida e por compartilharem tantos momentos bons comigo!

Josi, minha amiga do técnico que tem me acompanhado nesses últimos anos!

À minha super plateia que foi prestigiar minha defesa!

Aos meus professores da graduação: professora Glória, professora Martha e professora Rita. Obrigada por terem me escolhido como monitora nas disciplinas de vocês. Aos professores Valdir e Luciano pelas aulas perturbadoras.

Angélica e professora Elena, obrigada pelas oportunidades de aprendizado profissional e pessoal.

Agradeço às professoras Sônia e Rita por aceitarem ser minha banca!

Mãe UFRGS! Obrigada por me aceitar como sua filha e por ter me dado a chance de realizar meu maior sonho! E eu vou voltar, hein? Me aguarde!!

Meu professor, chefe, coautor, orientador e amigo Rafael! Eu seria muito ingrata se não te dedicasse pelo menos um parágrafo, no mínimo, de agradecimento. Se não fosse pela tua paciência quase divina, tua bondade, generosidade, respeito, confiança e humanidade, eu não teria chegado até aqui. Foram 4 anos de convivência que passaram rápido. Muitas vezes pensei em desistir de tudo, inclusive de ti (porque as brigas não foram poucas), mas aprendemos a conviver com nossas diferenças, deu certo e não me arrependo! Não tenho palavras pra agradecer todo teu zelo e empenho na elaboração desse trabalho, que tem muito de mim, mas que também tem muito de ti. Obrigada por entender minhas choradeiras e por sempre me receber na tua sala, por mais ocupado e/ou atrasado que tu pudesse estar. Acredito que nada na vida é por acaso e o teu nome é o melhor que tu poderia receber, pois foste meu anjo nesses anos de Fabico. Espero que minha fama de “menina do Rafael” não seja um ocorrente, mas um continuante!

A ontologia é, assim, a própria Filosofia e o conhecimento do Ser, isto é, das idéias, é a passagem das opiniões sobre as coisas sensíveis mutáveis rumo ao pensamento sobre as essências imutáveis. Passar do sensível ao inteligível – tarefa da Filosofia – é passar da aparência ao real, do Não-Ser ao Ser.

Marilena Chauí

RESUMO

Esta monografia apresenta a análise da ontologia (VIVO-ISF) do sistema VIVO. Explica que a Web Semântica e os Dados Ligados (Linked Data) trazem novas possibilidades para a comunicação e o intercâmbio de recursos da ciência. Cita projetos na ciência que envolvem a colaboração em um ambiente semântico com a finalidade de promover a interoperabilidade e a padronização dos recursos. Apresenta como justificativa a necessidade da análise, do questionamento e da reflexão no processo de construção e/ou adaptação de uma ontologia. Traz como problema de pesquisa a extensão e a adequação da ontologia VIVO-ISF para a descrição de recursos do domínio da ciência. Apresenta como objetivo geral a investigação da ontologia VIVO - ISF quanto a sua capacidade para descrição de recursos da ciência, em âmbito geral, no domínio da ciência. Apresenta o sistema VIVO como contexto do estudo. Aborda os temas de ontologias e Web Semântica no referencial teórico. Cita o estudo de caso como procedimento técnico para realização da pesquisa. Apresenta os resultados da pesquisa em quatro partes: escopo, fundamentação, evolução e reúso. Explica que a ontologia VIVO-ISF é do domínio da ciência e de abrangência geral e específica. Explica a importância do uso da fundamentação na remodelagem conceitual e na estrutura da ontologia. Explica que a estrutura conceitual proporcionada pela ontologia de fundamentação favorece a evolução. Explica que o reúso permite uma maior interoperabilidade semântica com outros ambientes que também utilizam as mesmas ontologias utilizadas por VIVO. Conclui que a ontologia VIVO-ISF pode ser utilizada na descrição de recursos do domínio da ciência, e é genérica o suficiente, em sua extensão, para ser utilizada em outros ambientes. Sugere a realização de novos estudos na área.

Palavras-chave: Ontologias. Web Semântica. Linked Open Data. VIVO. VIVO-ISF.

ABSTRACT

This paper presents the analysis of the ontology (VIVO-ISF) VIVO system. Explains that the Semantic Web and Linked Data (Linked Data) bring new possibilities for communication and the exchange of resources of science. Cites projects in science involving collaboration on a semantic environment in order to promote interoperability and standardization of resources. Presents as a justification of the need analysis, questioning and reflection in the construction process and/or adaptation of an ontology. Brings as research problem the extent and adequacy of VIVO-ISF ontology for describing domain resources of science. Presents the general objective of the research VIVO ontology - ISF as its ability to resource description of science in general framework in the field of science. Displays the VIVO system as the study context. Addresses the issues of ontologies and the Semantic Web in the theoretical framework. Cites the case study as a technical procedure for the research. Presents search results in four parts: scope, rationale, evolution and reuse. Explains that the VIVO-ISF ontology is the science and general and specific scope. Explains the importance of using the reasoning in the conceptual remodeling and ontology structure. Explains that the conceptual framework provided by the reasoning of ontology evolution favors. Explains that the reuse allows greater interoperability with other semantic environments that also use the same ontologies used by VIVO. Concludes that the VIVO-ISF ontology can be used in resource description of the field of science, and is generic enough in its length, for use in other environments. Suggests new studies in the area.

Keywords: Ontologies. Semantic Web. Linked Open Data. VIVO. VIVO-ISF.

LISTA DE ILUSTRAÇÕES

Figura 1 -	Perfil de pesquisador em VIVO	22
Figura 2 -	Mapa da Ciência	23
Figura 3 -	Rede de Coautor	24
Figura 4 -	Rede de Coinvestigador	25
Figura 5 -	Componentes básicos de uma ontologia	31
Figura 6 -	Uso de ontologias	32
Figura 7 -	BFO	36
Figura 8 -	FOAF	37
Figura 9 -	VCard	38
Figura 10 -	<i>Geopolitical Ontology</i>	39
Figura 11 -	Event	40
Figura 12 -	LSC	43
Figura 13 -	OBI	44
Figura 14 -	Exemplo de busca no <i>Google</i>	46
Figura 15 -	Exemplo de busca no <i>Factbites</i>	47
Figura 16 -	Tripla RDF	50
Figura 17 -	<i>Links</i> semânticos	53
Figura 18 -	Nuvem LOD em 2010	55
Figura 19 -	Nuvem LOD em 2014	56
Figura 20 -	Diagrama resumido da versão 1.6 da ontologia de VIVO	61
Figura 21 -	Conceitos em VIVO-ISF	63
Figura 22 -	VIVO & <i>Eagle-i</i>	68

LISTA DE QUADROS

Quadro 1 -	Categorias e subcategorias da <i>University of Florida</i>	21
Quadro 2 -	Os filósofos e a Ontologia	28
Quadro 3 -	Tipos de ontologias	33
Quadro 4 -	5 estrelas e LOD	50
Quadro 5 -	Linked Open Data na ciência	53
Quadro 6 -	Coleta dos dados	56
Quadro 7 -	Fundamentação em VIVO-ISF	62
Quadro 8 -	Principais modificações até a versão 1.5	65
Quadro 9 -	Ontologias em VIVO - ISF	68

LISTA DE SIGLAS

AIISO	Academic Institution Internal Structure Ontology
BFO	Basic Formal Ontology
BIBO	Bibliographic Ontology
Fabio	FRBR-aligned Bibliographic Ontology
FOAF	Friend of a Friend
LOD	Linked Open Data
LSC	Linked Science Core Vocabulary Specification
MIREOT	The Minimum Information to Reference an External Ontology Term
OBI	Ontology for Biomedical Investigations
OWL	Web Ontology Language
RDF	Resource Description Framework
SKOS	Simple Knowledge Organization System
UFRGS	Universidade Federal do Rio Grande do Sul
URI	Uniform Resource Identifier
VIVO-ISF	VIVO Integrated Semantic Framework

SUMÁRIO

1 INTRODUÇÃO	15
1.1 JUSTIFICATIVA/MOTIVAÇÃO.....	17
1.2 PROBLEMA DE PESQUISA	18
1.3 OBJETIVOS	18
1.3.1 Objetivo geral	18
1.3.2 Objetivos específicos	19
2 CONTEXTO DO ESTUDO: O SISTEMA VIVO	20
3 REFERENCIAL TEÓRICO	27
3.1 (O) ONTOLOGIA (S)	27
3.1.1 A Ontologia Segundo a Filosofia e a Inteligência Artificial	27
3.1.2 Componentes de uma Ontologia	30
3.1.3 Ontologias: Porque Usá-las e Tipos	32
3.1.4 Ontologias na Web Semântica	34
3.1.4.1 <i>Basic Formal Ontology (BFO)</i>	35
3.1.4.2 <i>Friend of a Friend (FOAF)</i>	37
3.1.4.3 <i>VCard</i>	38
3.1.4.4 <i>Geopolitical Ontology</i>	39
3.1.4.5 <i>Event</i>	40
3.1.4.6 <i>Simple Knowledge Organization System (SKOS)</i>	40
3.1.4.7 <i>Bibliographic Ontology (BIBO)</i>	41
3.1.4.8 <i>FRBR-aligned Bibliographic Ontology (Fabio)</i>	41
3.1.4.9 <i>Linked Science Core Vocabulary Specification (LSC)</i>	42
3.1.4.10 <i>Academic Institution Internal Structure Ontology (AIISO)</i>	43
3.1.4.11 <i>Ontology for Biomedical Investigations (OBI)</i>	44
3.2 WEB SEMÂNTICA	45
3.2.1 Agentes Inteligentes	49
3.2.2 Uniform Resource Identifier (URI)	49
3.2.3 Resource Description Framework (RDF) e Resource Description Framework Schema (RDF Schema)	50
3.2.4 Web Ontology Language (OWL)	51
3.2.5 Linked Open Data (LOD)	52
3.2.6 Linked Open Data na Ciência	55

4 METODOLOGIA	58
4.1 NATUREZA, OBJETIVO, ABORDAGEM E PROCEDIMENTO TÉCNICO DA PESQUISA	58
4.2 CORPUS DA PESQUISA	58
4.3 INSTRUMENTO DE COLETA DE DADOS.....	59
4.4 PROCEDIMENTOS DE COLETA DE DADOS.....	59
4.5 ANÁLISE DOS DADOS	60
5 RESULTADOS	61
5.1 ESCOPO DE VIVO-ISF	61
5.2 A FUNDAMENTAÇÃO EM VIVO-ISF	65
5.3 EVOLUÇÃO	67
5.4 REÚSO EM VIVO-ISF	69
6 CONCLUSÕES	72
REFERÊNCIAS.....	74
APÊNDICE – Formulário de coleta de dados.....	79
ANEXO A - Diagrama da versão 1.1 da ontologia de VIVO.....	80
ANEXO B - Diagrama da versão 1.2 da ontologia de VIVO.....	81
ANEXO C - Diagrama da versão 1.3 da ontologia de VIVO.....	82
ANEXO D- Diagrama da versão 1.4 da ontologia de VIVO.....	83

1 INTRODUÇÃO

Clico, logo existo! Uma paráfrase de Descartes que pode representar a sociedade a qual pertencemos. Sociedade da informação e/ou sociedade do conhecimento são denominações utilizadas para emoldurar as contínuas e rápidas transformações a partir do advento da tecnologia. A internet é um divisor de águas: as relações entre ser humano - ser humano e ser humano – máquina não são mais as mesmas. A “Era www” provocou e vem provocando mudanças no seu *modus operandi*, ou seja, a web se apresenta e se forma sob nós, teias, redes e artefatos que permitem ao usuário ser parte desse sistema.

A *web* surgiu com a finalidade de ser uma rede global para publicação e interligação de documentos. Hoje, evoluiu para um ambiente, no qual, seus usuários são colaboradores ativos na sua construção (Web 2.0). Apesar de ser intangível, intocável e estranhamente abstrata, a *web* está presente de tal forma que é difícil estabelecer uma linha tênue entre a vida real e a vida virtual. Estamos saindo da posição de passividade para assumirmos o papel de agentes na construção de uma realidade cibernética que está cada vez intrínseca a nós. Não somos mais meros consumidores de *software*, mas construtores de plataformas e ambientes que tiram proveito da inteligência coletiva.

Entretanto, seu principal desafio é fazer com que máquinas (agentes de *softwares* inteligentes) possam compreender o significado de seus conteúdos, a fim de prover serviços inteligentes às pessoas. Dessa forma, surge a Web Semântica (Web 3.0 ou Web dos Dados), cuja “espinha dorsal” é a estruturação dos dados, a partir do uso de artefatos tecnológicos e ferramentas.

A *web* atual vem provocando mudanças em todos os aspectos de nossas vidas. As transformações ecoam e reverberam em sentidos técnicos, filosóficos e científicos, provocando uma reação em cadeia que culmina em uma nova sociedade e por que não em uma nova ciência?

Segundo Colon e Holmes (2012):

A formação de uma equipe de nível mundial requer de recursos ricos e variados para descoberta no campo da pesquisa, através dos quais seus cientistas membros possam ter uma consciência ampla e profunda do que está acontecendo em suas disciplinas. Estão encerrados os tempos de simplesmente manter-se atualizado com a literatura atual, ou mesmo antecipar-se sobre o que está vindo ao ser

editor de uma revista [...]. O problema com essa abordagem é que ela só pode fornecer um ponto de vista vantajoso a partir de uma pessoa de uma rede ou de um quadro de referência, e não favorece a descoberta de projetos para além do seu horizonte imediato (COLON; HOLMES, 2012, documento não paginado, tradução nossa *online*).

A Web Semântica e os Dados Ligados (*Linked Data*) trazem novas possibilidades para a comunicação e o intercâmbio de recursos da ciência. Enquanto Arquivos Abertos estão focados na publicação e na interoperabilidade de publicações científicas na web, a plataforma Web Semântica/*Linked Data* possibilita a **interoperabilidade semântica** de recursos da ciência que, atualmente, vão além dos bibliográficos, sendo: dados de pesquisa, pesquisadores, instituições, projetos, processos e eventos.

Observa-se, então, a necessidade de descrever conceitualmente não apenas publicações, mas outros recursos que fazem parte da ciência, a partir do uso de uma ferramenta poderosa que utiliza modelos de dados, linguagens, dados estruturados e modelagem conceitual: as ontologias. São as ontologias que tornam legíveis e compreensíveis os recursos da ciência para um computador, que por ser desprovido de inteligência, não compreende o real significado das coisas. Pode-se afirmar, portanto, que as ontologias são uma representação formal do mundo abstrato (as ontologias contém os conceitos utilizados para descrever os recursos), em uma linguagem codificada por máquina.

A representação de recursos da ciência na Web Semântica tem sido chamada de *Linked Science*, uma abordagem que: amplia as possibilidades de trocas no desenvolvimento da ciência, promove a descoberta de recursos, e aproxima pesquisas em um ambiente compartilhado.

Dessa forma, é necessário que sejam desenvolvidos ambientes que possam se adequar a essa nova ciência. Esses ambientes são chamados de *linked open data* (LOD), os quais, tem por objetivos, o compartilhamento de informações, a descoberta de recursos e a interoperabilidade semântica. Os ambientes, que já fazem parte do cenário mundial, abarcam as mais diversas áreas do conhecimento e descrevem diversos recursos, além dos conhecidos bibliográficos.

Existem projetos na ciência que envolvem a colaboração em um ambiente semântico com a finalidade de promover a interoperabilidade e a padronização dos recursos, lançando mão de ferramentas e serviços provenientes da Web 2.0 e da Web semântica. Esses projetos (como VIVO, *Eagle-i* e LODUM), ampliam a

descrição e interligação semântica de recursos da ciência envolvendo, além de documentos, pesquisadores (e seus perfis/currículos), instituições, unidades, laboratórios, projetos, auxílios (financiamentos), cursos, palestras, orientações, protocolos, dentre outros.

VIVO é um ambiente *linked open data* que tem por objetivos a descoberta e o compartilhamento de recursos da ciência sendo direcionado para cientistas das instituições participantes. Para que isso ocorra, o ambiente utiliza a ontologia VIVO – ISF, a qual, modela, organiza, padroniza e representa os principais conceitos da ciência. Dessa forma, a ontologia VIVO – ISF é o objeto de estudo dessa pesquisa, pois descreve genericamente recursos do domínio da ciência, reutiliza diversas ontologias conhecidas no cenário *linked data* e pode ser adaptada a outros ambientes.

Entretanto, antes de efetuar a escolha e a incorporação de uma ontologia em um sistema LOD devem ser analisados alguns fatores que são fundamentais para o sucesso e a garantia de uma ontologia. Esses fatores abarcam questões relacionadas à estrutura e à modelagem conceitual.

O trabalho inicia apresentando a justificativa, o problema de pesquisa e os objetivos pretendidos. O capítulo 2 apresenta o contexto do estudo. O capítulo 3 trata do referencial teórico que serviu de embasamento para esta pesquisa. O capítulo 4 descreve a metodologia utilizada para a concretização dos objetivos. O capítulo 5 apresenta a análise dos resultados obtidos. E o capítulo 6 trata das conclusões.

1.1 JUSTIFICATIVA/MOTIVAÇÃO

A *web* provocou, provoca e vem provocando transformações que ecoam nos mais diversos setores da sociedade. O usuário assume o papel de agente e de consumidor ativo em ambientes que promovem a descoberta e o compartilhamento de dados e de informações. Mas o que há de especial nessa **nova web**? As características e ferramentas oriundas da Web 2.0 e da Web Semântica permitem **novas** formas para publicação de dados (*linked open data*) e **novas** formas para descrição de **novos** recursos.

Projetos na ciência também fazem uso desses benefícios na publicação e na promoção da descoberta de dados científicos em ambientes *linked open data*. Para que o *modus operandi* do ambiente seja efetivado, é fundamental que exista uma ferramenta que padronize e que descreva os recursos pertencentes ao domínio da ciência: as ontologias. As ontologias alicerçam a estrutura desses ambientes, pois são artefatos tecnológicos que representam a realidade, permitindo que as informações sejam inteligíveis por humanos e por máquinas.

Entretanto, cada um desses ambientes possui uma realidade própria, ou seja, recursos e conceitos que representam determinado domínio (geral ou específico). Dessa forma, a ontologia a ser implementada deve contemplar as reais necessidades e o escopo do ambiente. Uma ontologia pode ser ideal para um dado ambiente, mas não para o outro. Recursos gerais podem não ser suficientes para um domínio específico e vice-versa, por isso a análise é fundamental.

Esta é a justificativa deste trabalho: **é preciso conhecer a Ontologia da ontologia** a fim de evitar futuras mudanças indesejadas em sua estrutura. O questionamento, a análise e a reflexão são imprescindíveis no processo de construção e/ou adaptação de uma ontologia.

1.2 PROBLEMA DE PESQUISA

A ontologia VIVO – ISF pode ser utilizada na descrição de recursos do domínio da ciência? É genérica o suficiente, em sua extensão, para ser utilizada em outros ambientes?

1.3 OBJETIVOS

Nesta subseção serão apresentados os objetivos desse trabalho.

1.3.1 Objetivo geral

Investigar a ontologia VIVO - ISF quanto a sua capacidade para descrição de recursos da ciência, em âmbito geral, no domínio da ciência.

1.3.2 Objetivos específicos

- a) Investigar o escopo da ontologia e sua adequação para descrição de recursos, em âmbito geral, no domínio da ciência;
- b) Investigar os fundamentos na elaboração da ontologia;
- c) Investigar o processo de evolução da ontologia;
- d) Investigar o reúso de ontologias visando a interoperabilidade semântica.

2 CONTEXTO DO ESTUDO: O SISTEMA VIVO

VIVO é uma rede interdisciplinar desenvolvida (2003) e implementada na *Cornell University* (2004), cujos objetivos são: apoiar a pesquisa, a navegação e a visualização, a fim de permitir a descoberta de pessoas, programas, instalações, financiamento, trabalhos acadêmicos e eventos. (VIVO, 2013).

Plataforma *linked open data* que conecta pesquisadores, a partir de um modelo de ontologia que interrelaciona entidades e propriedades, permitindo, dessa forma, a visualização de perfis de professores e/ou pesquisadores, a fim de promover a descoberta de recursos da ciência, para cientistas das instituições pertencentes à rede. De acordo com Conlon e *et al.* (2006):

Ao armazenar e expor dados como RDF e utilizando ontologias padrão, as informações em VIVO podem ser exibidas em uma página da web ou entregues a outros sistemas como RDF, permitindo que metadados sejam colhidos, agregados, e integrados na nuvem *Linked Open Data*. Dados reutilizados e reaproveitamento aumentam a precisão dos dados e a eficiência em toda a organização. (CONLON *et al.*, 2006, p. 429, tradução nossa *online*,).

Atualmente, as instituições participantes da rede são: *University of Florida, Cornell University, Indiana University, Washington University School of Medicine in St. Louis, The Scripps Research Institute, Weill Cornell Medical College, and Ponce Medical College*. Hoje, devido ao seu reconhecimento nacional e internacional, o projeto VIVO é coordenado pelo Consórcio *CTSACConnect* ao lado de outros projetos: *OHSU Ontology Development Group, Oregon Clinical & Translational Research Institute, Eagle-i e ShareCenter*.

Cada instituição pertencente à rede VIVO possui uma página própria que contém informações acerca da entidade, sendo *layout* e interface similares na sua apresentação. A página inicial difere de uma instituição para outra, entretanto, são sempre apresentadas quatro categorias principais: pessoas, organizações, pesquisa e eventos. O que varia é a especificidade das subcategorias, sendo algumas mais exaustivas do que outras.

Para facilitar a descoberta de informações, VIVO oferece uma busca facetada, onde os resultados obtidos incluem: pessoas, organizações, pesquisa, auxílios/financiamento e atividades. Categorias essas, presentes no sistema, o que

confirma a padronização dos recursos e que valida VIVO como objeto de estudo ideal para essa pesquisa.

O Quadro 1 traz, como exemplo, as categorias e subcategorias encontradas na página da *University of Florida*.

Quadro 1 - Categorias e subcategorias da *University of Florida*

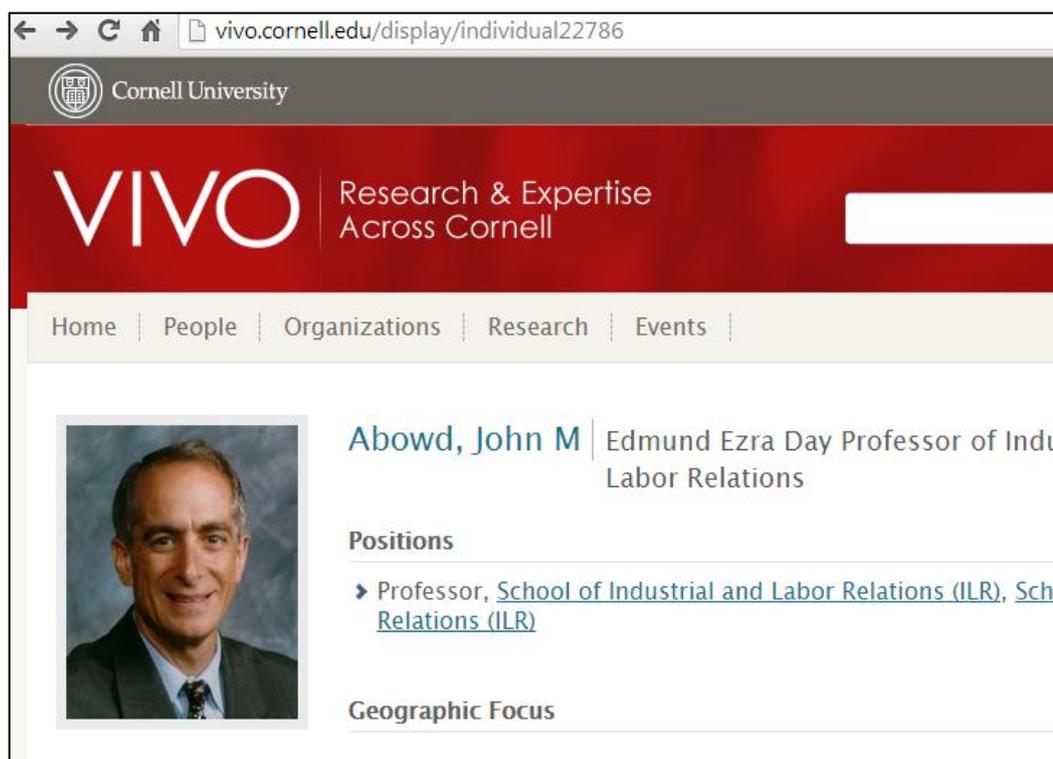
Categoria	Subcategorias
Pessoas	Faculdade Clínica, Consultor, Membro de Faculdade, Membro Emérito de Faculdade, Estudante Graduado, Bibliotecário, Não-Acadêmico, Não-Acadêmico de Faculdade, Pessoa, Pós-Doc, Professor Emérito.
Organizações	Departamento Acadêmico, Unidade Administrativa, Associação, Centro, Organização Clínica, Comitê, Companhia, Consórcio, Núcleo de Laboratório, Departamento, Divisão, Unidade de Extensão, Fundação, Organização Financiada, Agência do Governo, Grupo, Hospital, Instituto, Laboratório, Biblioteca, Museu, Organização Não-Governamental, Organização, Programa, Publicador, Organização de Pesquisa, Organização de Estudantes, Time, Universidade.
Pesquisa	Artigo Acadêmico, Artigo, Prêmio, Postagem de Blog, Livro, Estudo de Caso, Catálogo, Capítulo, Conceito, Artigo de Conferência, Base de Dados, Livro Editado, Arquivo Editorial, Equipamento, Jornal, Informativo, Patente, Anais de Eventos, Revisão, Séries, Software, Tese, Vídeo, <i>Webpage</i> , <i>Website</i> .
Auxílios	Auxílios.
Atividades	Acordos, Certificação, Licenciamento, Projeto, Projeto de Pesquisa, Serviço.
Cursos	Curso, Seção de Curso
Evento	Competição, Conferência, Série de Conferência, Evento, Exposição, Reunião, Apresentação, Série de Seminário, <i>Workshop</i> , Série de <i>Workshop</i> .
Localizações	Endereço, Edifício, Campus, Cidade, Continente, País, Localização Geográfica, Estado ou Província, Região Transnacional, Endereço Postal.

Fonte: autora (2014).

Cada pessoa possui um perfil que contém as seguintes informações: Posição, Foco Geográfico, *Website*, Afiliações, Pesquisa, Ensino, Publicações, Serviço, Identidade e Outros.

A Figura 1 apresenta o perfil do pesquisador John M. Abowd da *Cornell University*.

Figura 1 – Perfil de pesquisador em VIVO

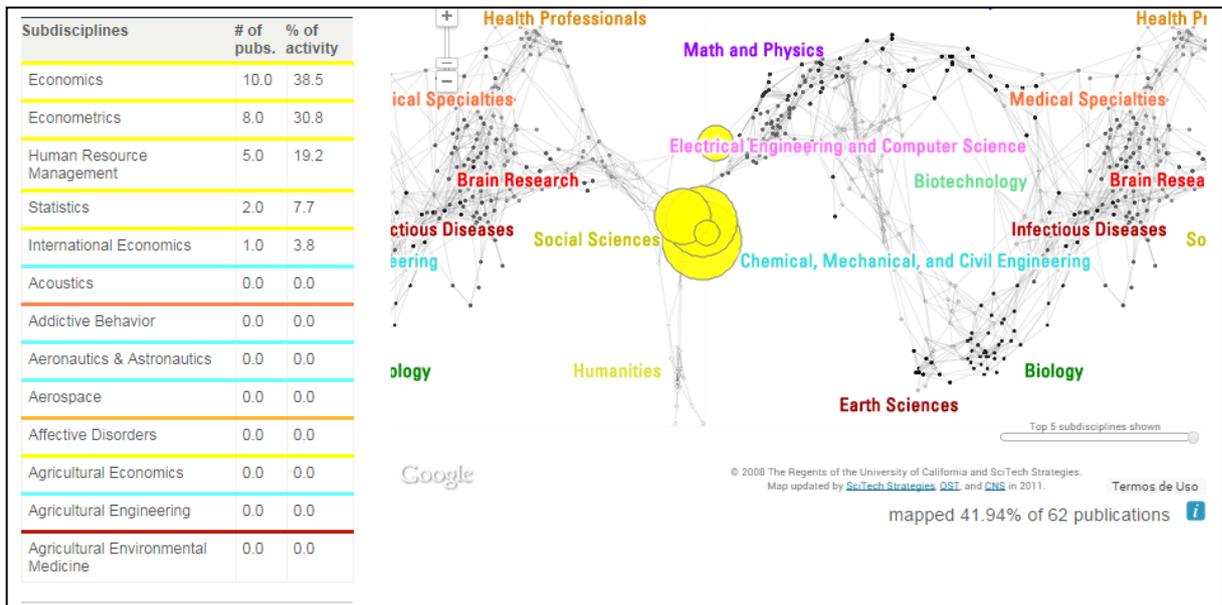


Fonte: Vivo (2014).

VIVO permite a visualização de redes relacionadas ao Mapeamento da Ciência. Tais redes fornecem informações sobre Coautor, Coinvestigador e Mapa da Ciência, tanto de pesquisadores, como das instituições participantes.

As Figuras 2, 3 e 4 mostram o Mapa da Ciência, a Rede de Coautor e Rede de Coinvestigador do pesquisador supracitado.

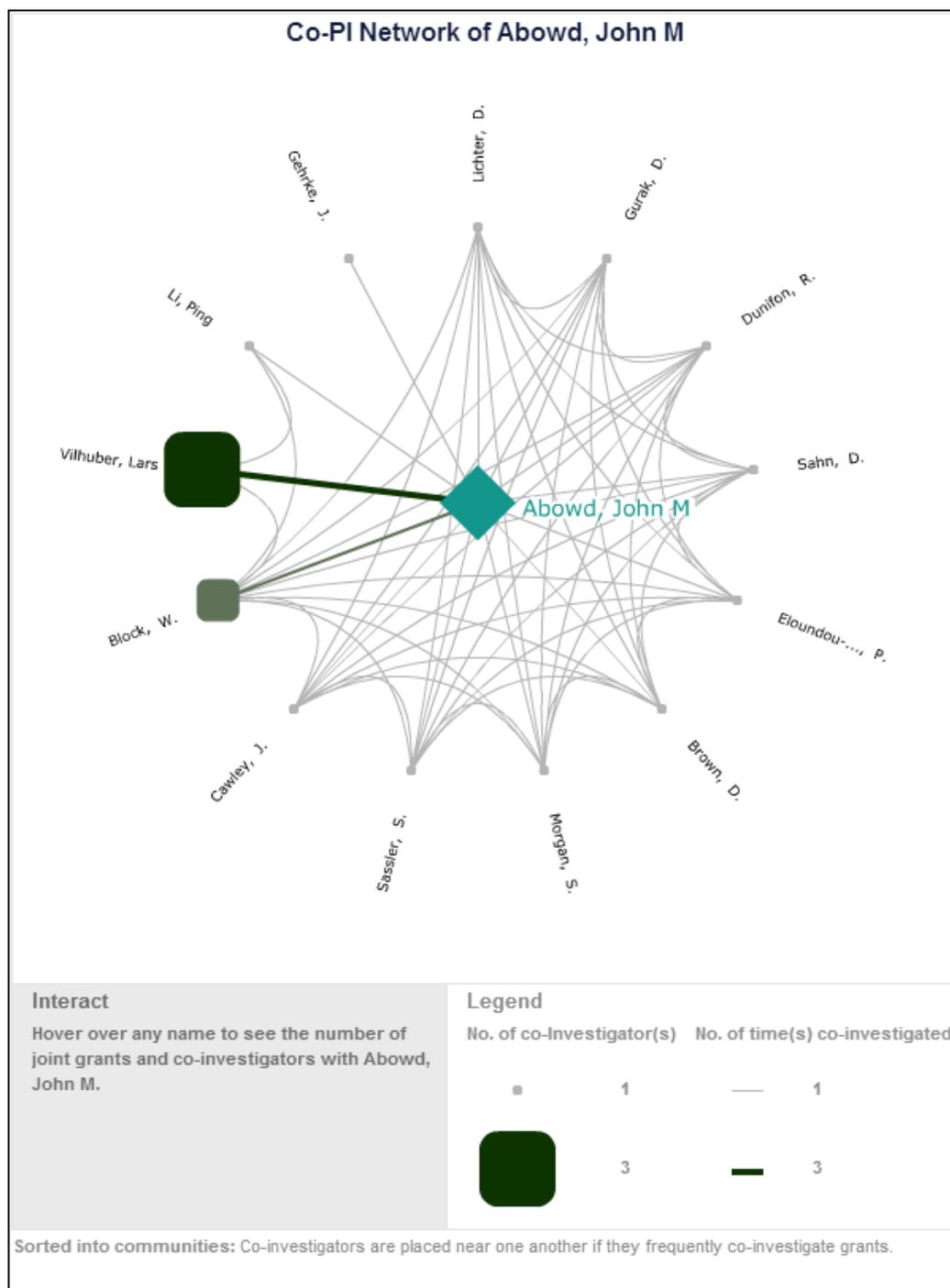
Figura 2 – Mapa da Ciência



Fonte: Vivo (2014).

No Mapa, são visualizadas as subdisciplinas com as quais o cientista está relacionado, a partir do número de publicações e de atividades referentes a essas. Nesse caso, o pesquisador possui uma ligação maior com as Ciências Sociais, mais especificamente, com a Economia.

Figura 4 – Rede de Coinvestigador



Fonte: Vivo (2014).

A Figura 4 apresenta os coinvestigadores do pesquisador, tendo, como parâmetro, o número de vezes em que ocorreu a coinvestigação. Nesse caso, o coinvestigador que mais pesquisou com Jonh M. Abowd também é Lars Vilhuber.

VIVO tem como palavras de ordem: **conecte, compartilhe, descubra**. Para que seu *modus operandi* seja efetivado há, por trás desse sistema, a ontologia VIVO

- ISF, a qual é responsável pela padronização e modelagem conceitual dos recursos e pela interoperabilidade semântica. Desde a criação do ambiente, a ontologia do sistema é atualizada continuamente, seja pela exclusão/inclusão de classes e/ou propriedades, seja pela remodelagem conceitual.

Atualmente, a Ontologia VIVO-ISF está na “versão” 1.6, a qual, é o objeto de estudo deste trabalho.

3 REFERENCIAL TEÓRICO

Esta seção apresenta o referencial teórico utilizado neste trabalho.

3.1 (O) ONTOLOGIA (S)

Nesta subseção serão abordadas as principais definições, os componentes, uso, tipos, porque usá-las e uso na Web Semântica.

3.1.1 A Ontologia Segundo a Filosofia e a Inteligência Artificial

A palavra ontologia gera dificuldades quanto ao seu significado, visto que é objeto de estudo da Filosofia, Inteligência Artificial, Engenharia do Conhecimento, Linguística e da Ciência da Informação. Entretanto, para este trabalho, serão abordadas as definições atribuídas pela Filosofia e pela Inteligência Artificial.

Para a Filosofia, o termo ontologia tem origem grega, em que, *ontos* significa ser e *logos* significa discurso. Portanto, ontologia é o estudo, a Ciência do Ser. De acordo com Lima-Marques (2006, p. 17) é: “[...] o estudo da existência de todos os tipos de entidades, abstratas ou concretas, que constituem o mundo. (LIMA-MARQUES, 2006, p.17).

Conforme Chauí (2000), apesar de a ontologia ter começado com Parmênides e Platão, é comum atribuir seu nascimento a Aristóteles. Isso se deve a três motivos:

- 1) Aristóteles não julga o mundo das coisas sensíveis, ou a Natureza, um mundo aparente e ilusório. Pelo contrário, é um mundo real e verdadeiro cuja essência é, justamente, a multiplicidade de seres e a mudança incessante.
- 2) Aristóteles considera que a essência verdadeira das coisas naturais e dos seres humanos e de suas ações não está no mundo inteligível, separado do mundo sensível, onde as coisas físicas ou naturais existem e onde vivemos. As essências, diz Aristóteles, estão nas próprias coisas, nos próprios homens, nas próprias ações e é tarefa da Filosofia conhecê-las ali mesmo onde existem e acontecem.
- 3) Ao se dedicar à Filosofia Primeira ou metafísica, a Filosofia descobre que há diferentes tipos ou modalidades de essências [...]. (CHAUÍ, 2000, p. 276 *online*).

Ontologia, Metafísica e Filosofia Primeira são sinônimos. Mas por que a palavra Metafísica se sobressaiu em relação à ontologia? Conforme aponta Chauí (2006, p. 267 *online*), a palavra Metafísica “[...] significa o estudo de alguma coisa que está acima e além das coisas físicas ou naturais e que é a condição da existência e do conhecimento delas.”. Assim sendo, o termo Metafísica consagrou-se como a Ciência do Ser.

Apenas a partir do século XX a ontologia passou a ser considerada como um estudo diferenciado do ser. (LIMA-MARQUES, 2006).

Existem, na própria Filosofia, diferentes visões para o conceito de ontologia, que variam conforme o pensamento de cada filósofo. O Quadro 2 apresenta algumas dessas visões.

Quadro 2- Os filósofos e a Ontologia

Filósofo	Ontologia
Aristóteles	Filosofia Primeira: o estudo do ser das coisas.
Christian Wolff Hegel	Filosofia Primeira ou ser enquanto ser (no sentido geral Aristotélico).
Abraham Calovius	Se refere ao próprio tema ou objeto.
Leibniz	Trata de um formalismo separado ou exclusivamente lógico.
Immanuel Kant	Leva em consideração a ideia de ser como um saber <i>a priori</i> , desvinculado da experiência e por ela utilizado.
Edmund Husserl	Estudo do ser diferenciado em entes dotados de essências próprias e irreduzíveis uns aos outros.
Heidegger e Maurice Merleau-Ponty	Empenha-se em buscar a essência ou sentido do ente físico ou natural, do ente psíquico, lógico, matemático, espacial, temporal, etc.
Stanislaw Lesniewski	É uma teoria e cálculo de classes e relações, diferenciando-se do cálculo proposicional e da álgebra de classes criando uma axiomática ontológica.
Quine	A ontologia de uma teoria, distinta das ideias que por ela se podem representar.
Gustav Bergmann	Quantificação do que existe é diferente da existência da Ontologia tradicional. Sugestão de um padrão ontológico formado por uma linguagem fictícia ideal.

Fonte: adaptado de Lima-Marques (2006).

Conforme pode ser observado no Quadro 2, o termo Ontologia carrega visões similares e ao mesmo tempo diferentes sobre seu significado. Entretanto, essa diversidade não se restringe apenas ao conceito, pois a grafia também pode interferir na interpretação da palavra. De acordo com Lima-Marques (2006) Ontologia refere-se à Ciência do Ser, enquanto que ontologias se referem ao uso, à finalidade. É, portanto, um conceito amplo, complexo e passível de modificações e adaptações, como acontece, por exemplo, com a Inteligência Artificial.

A Inteligência Artificial não se apropriou desse termo filosófico por acaso. O que é óbvio para nós humanos não é para uma máquina. Um computador não sabe diferenciar uma pessoa de um animal, uma tese de um artigo, um professor de um aluno, pois a máquina é desprovida de raciocínio, de cognição. Dessa forma, torna-se necessário buscar compreender o significado dos objetos, das “coisas” que representam o mundo das ideias e as relações semânticas existentes. Essa compreensão é fundamental quando se fala em ontologias.

As ontologias são ferramentas úteis tanto para a Inteligência Artificial quanto para a Web Semântica (subseção 3.2), pois oferecem sentido e estruturação aos dados, a partir da padronização de recursos que descrevem um determinado domínio do conhecimento.

Uma ontologia pode ser definida como “[...] uma especificação formal e explícita de uma conceitualização compartilhada.” (GRUBER, 1993, p. 1, *online*, tradução nossa). Para um melhor entendimento é necessário “desmembrar” essa definição, sendo que suas principais características são:

- a) Especificação: ser o mais específica possível;
- b) Formal: ser legível por máquina;
- c) Explícita: ser clara, objetiva;
- d) Conceitualização: representação do mundo das ideias;
- e) Compartilhada: compartilhada entre a comunidade envolvida.

Para a área de Inteligência Artificial, a ontologia é um artefato tecnológico desenvolvido em máquina e para máquinas, composto por uma estrutura controlada de termos, entidades, classes, subclasses, relações e propriedades, a partir de uma conceitualização compartilhada de uma determinada área do conhecimento.

Para Gruber (1993, p.5, *online*, tradução nossa) uma ontologia “[...] define os termos que serão utilizados na representação do conhecimento”. A partir dessa definição, observa-se que a ontologia também é objeto de estudo da área de Engenharia do Conhecimento, a qual, envolve a “extração” do conhecimento das pessoas envolvidas, a partir de técnicas utilizadas por especialistas.

3.1.2 Componentes de uma Ontologia

As ontologias podem variar sua estrutura, pois estão implícitas questões de especificidade, extensão e finalidade no seu processo de construção. Entretanto, de acordo com Almeida e Bax (2003), existem características e componentes básicos que compõe a maioria delas, sendo: classes, relacionamentos e instâncias.

Da mesma forma que ocorre em relação à definição do termo, não há um consenso acerca de todos os componentes que fazem parte de uma ontologia. Para essa pesquisa, serão apresentados os componentes básicos apontados por Ramalho (2010):

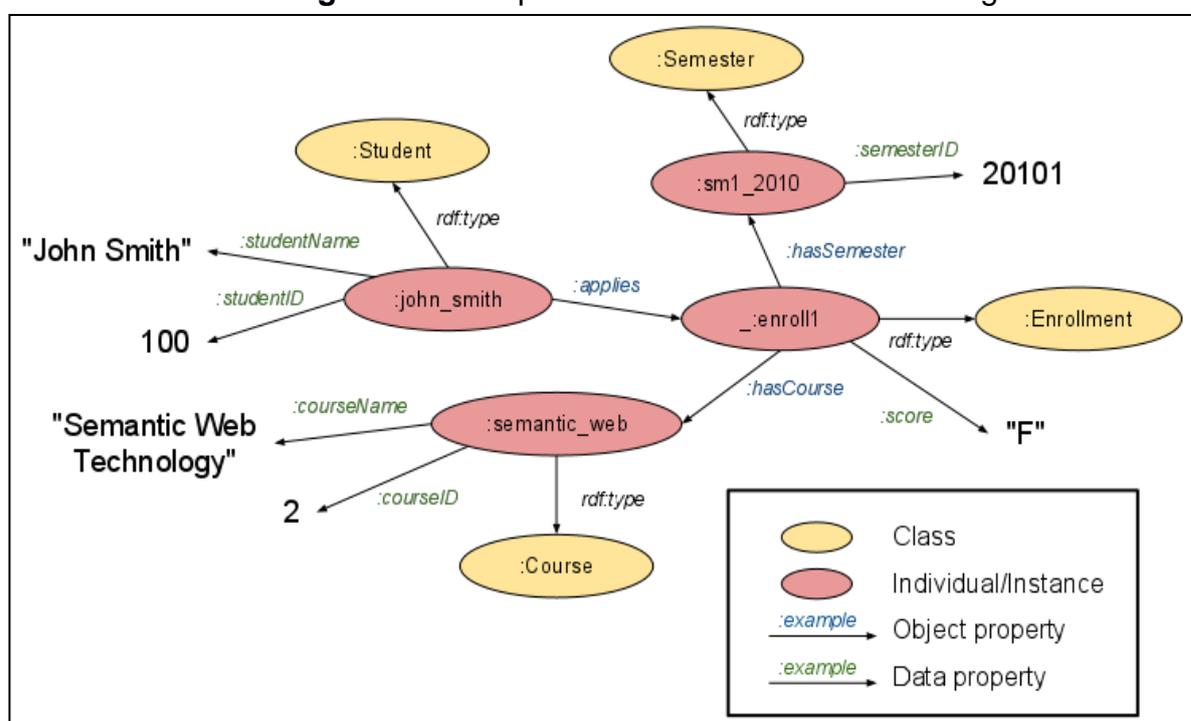
- a) **Classes e Subclasses:** As classes e subclasses de uma ontologia agrupam um conjunto de elementos, “coisas”, do “mundo real”, que são representadas e categorizadas de acordo com suas similaridades, levando-se em consideração um domínio concreto. Os elementos podem representar coisas físicas ou conceituais, desde objetos inanimados até teorias científicas ou correntes teóricas;
- b) **Propriedades Descritivas:** Descrevem as características, adjetivos e/ou qualidades das classes;
- c) **Propriedades Relacionais:** Trata-se dos relacionamentos entre classes pertencentes ou não a uma mesma hierarquia, descrevendo e rotulando os tipos de relações existentes no domínio representado;
- d) **Regras e Axiomas:** Enunciados lógicos que possibilitam impor condições como tipos de valores aceitos, descrevendo formalmente as regras da ontologia e possibilitando a realização de inferências automáticas a partir de informações que não necessariamente foram explicitadas no domínio, mas que podem estar implícitas na estrutura da ontologia;
- e) **Instâncias:** Indicam os valores das classes e subclasses, constituindo uma representação de objetos ou indivíduos pertencentes ao domínio modelado, de acordo com as características das classes, relacionamentos e restrições definidas;
- f) **Valores:** Atribuem valores concretos às propriedades descritivas, indicando os formatos e tipos de valores aceitos em cada classe. (RAMALHO, 2010, p. 38 *online*).

Berners-Lee, Hendler e Lassila (2001, tradução nossa) afirmam que uma ontologia é delimitada por uma taxonomia (define classes, subclasses e suas relações) e por um conjunto de regras de inferência.

A partir das definições acima, compreende-se que: as **classes e subclasses** representam os conceitos, as **propriedades** tratam das características, os **axiomas** são as regras e as **instâncias** indicam os valores das classes e subclasses.

A Figura 5 mostra alguns dos componentes básicos de uma ontologia e suas instâncias.

Figura 5 – Componentes básicos de uma ontologia



Fonte: Babbage (2010)

A ontologia exemplificada, anteriormente, representa a situação de um aluno matriculado no curso de “Tecnologia da Web Semântica”.

As classes representam os conceitos. Nesse caso, são: estudante, semestre, curso e matrícula. As instâncias são as responsáveis pela individualização das classes, sendo: John Smith, Web Semântica, matrícula 1, semestre 2010/1. As propriedades de objeto especificam as relações entre as instâncias, sendo: John Smith “possui” uma matrícula; matrícula 1 “referente” ao curso de Web Semântica; matrícula “referente” ao semestre 2010/1. E as propriedades dos dados “nomeiam” as instâncias, sendo: John Smith o nome do estudante; Tecnologia da Web

Semântica o nome do curso; 100 o número identificador de Jonh Smith; 2 o número identificador do curso; 20101 o número identificador do semestre e F a “nota” do aluno.

A partir da análise do exemplo anterior fica evidente que uma ontologia deve ser: específica, formal e explícita. Apesar de parecer complexa, a ontologia deve ser simples.

3.1.3 Ontologias: Porque Usá-las e Tipos

Ontologia é um assunto bastante em voga, seja pelo seu caráter atual, seja pela sua importância para as mais diversas áreas do conhecimento. Entretanto, apesar da notoriedade e da pertinência, os motivos para a construção de uma ontologia são por, algumas vezes, desconhecidos ou mal compreendidos. Sabe-se de sua relevância para a *web* atual, os diversos conceitos do termo, os componentes, os tipos e os métodos de construção.

Mas por que as ontologias são construídas, então? Bibliotecários, cientistas da computação, engenheiros do conhecimento, linguistas e filósofos podem responder conforme sua *expertise* na área, mas no mundo das ideias, as concepções são similares. As respostas de cada especialista podem variar na estrutura, mas são substancialmente as mesmas.

A Figura 6 traz os principais usos e vantagens para a construção das ontologias.

Figura 6 – Uso de ontologias



Fonte: autora (2014).

Estes são alguns dos motivos pelos quais uma ontologia é construída:

- a) **Descrição:** capacidade para descrever os mais variados recursos (bibliográficos, pessoas, organizações, dentre outros);
- b) **Modelagem:** modelagem conceitual de um domínio de conhecimento;
- c) **Padronização:** permite a padronização de recursos;
- d) **Interoperabilidade:** promove a interoperabilidade entre sistemas de informação;
- e) **Semântica:** promove significado às informações e aos dados;
- f) **Relações:** estabelece relações entre classes, subclasses e propriedades;
- g) **Compartilhamento:** facilita o compartilhamento de informações entre pessoas, comunidades, redes e instituições.

Existem diversos tipos de ontologias, com diferentes finalidades. A escolha é um processo que demanda planejamento, não apenas nas etapas, mas também na execução e na manutenção. Não há um padrão, pois tudo dependerá do objetivo a ser alcançado.

Para Almeida e Bax (2003) as ontologias podem ser classificadas de acordo com a função, o grau de formalismo, a aplicação, a estrutura e o conteúdo. Para este trabalho, será apresentada a tipologia de acordo com Guarino (1998, documento não paginado, tradução nossa, *online*), conforme o Quadro 3.

Quadro 3 – Tipos de ontologias

Tipos de ontologias	Definição
Ontologia de nível superior	Descrevem conceitos muito gerais como espaço, tempo, matéria, objeto, evento, ação, etc., que são independentes de um problema ou domínio particular.
Ontologia de domínio	Expressam conceituações de domínios particulares, descrevendo o vocabulário relacionado a um domínio genérico tal como Medicina ou Direito.
Ontologia de tarefa	Expressam conceituações sobre a resolução de problemas, independentemente do domínio em que ocorram, isto é, descrevem o vocabulário relacionado a uma atividade ou tarefa genérica, tal como, diagnosticar ou vender.
Ontologia de aplicação	Descrevem conceitos dependentes de um determinado domínio e tarefa particular. Estes conceitos frequentemente correspondem aos papéis desempenhados por entidades do domínio ao

executar uma determinada atividade.

Fonte: adaptado de Guarino (1998).

Uma ontologia representa formalmente um domínio de conhecimento, mas para que a representação do mundo abstrato seja a mais fidedigna possível em uma linguagem formal, é fundamental que existam questionamentos e uma espécie de “passo a passo”. Conforme Gómez-Pérez e Benjamins (1993) devem ser feitas as seguintes perguntas:

- a) O que é uma ontologia?
- b) Quais princípios devem ser seguidos na construção de uma ontologia?
- c) Quais são os componentes de uma ontologia?
- d) Quais tipos de ontologias existem?
- e) Como as ontologias são organizadas?
- f) Quais métodos podem ser usados para a construção da minha ontologia?
- g) Quais técnicas são apropriadas?
- h) Como um software pode ser utilizado na construção e no processo de desenvolvimento de ontologias?
- i) Quais são as ontologias mais conhecidas?
- j) Quais são os usos de ontologias?
- k) Quais são os princípios para selecionar a melhor ontologia para a minha finalidade? (GÓMEZ-PÉREZ; BENJAMINS, 1993, p.1 *online* tradução nossa).

As respostas para essas questões podem facilitar o planejamento da construção, pois uma ontologia varia conforme o escopo, o domínio do conhecimento, a comunidade envolvida e a finalidade. Ainda conforme Gómez-Pérez e Benjamins (1993) se as ontologias são de pequena escala, algumas etapas podem ser puladas, mas, se for uma ontologia de larga escala, devem estar inclusas garantias de correção e completude.

3.1.4 Ontologias na Web Semântica

Existem diversas ontologias que são utilizadas na Web Semântica para a descrição dos mais variados recursos. Para este trabalho serão abordadas as que descrevem: entidades, agentes, regiões geopolíticas, eventos, conceitos, recursos bibliográficos, acadêmicos, de pesquisa e de investigação clínica.

3.1.4.1 Basic Formal Ontology

A ontologia *Basic Formal Ontology* (BFO) foi desenvolvida em 2002 por Barry Smith e Pierre Grenon.

Conforme Arp e Smith (2008, p.1, tradução nossa, *online*):

BFO é uma ontologia de nível superior desenvolvida para apoiar a integração dos dados obtidos através da pesquisa científica. É deliberadamente concebida para ser muito pequena, de modo que deve ser capaz de representar de forma consistente as categorias de nível superior comum para ontologias de domínio desenvolvidas por cientistas em domínios diferentes e em diferentes níveis de granularidade. (ARP; SMITH, 2008, p. 1, tradução nossa, *online*).

A base da ontologia de fundamentação BFO é *Entity* (Entidade). Uma entidade pode ser *Continuant* (Continuante) ou *Occurrent* (Ocorrente). Nesse trabalho serão apresentadas apenas as categorias principais.

Continuante é o que persiste ao longo do tempo, o que existe por si mesmo como por exemplo: pessoas, qualidades e objetos. **Ocorrente** é um evento que acontece em um determinado tempo com o envolvimento de continuantes como, por exemplo: aula, congressos e reuniões. (ARP; SMITH, 2008).

Pertencem à entidade Continuante: *Independent Continuant* (Continuante Independente), *Generically Dependent Continuant* (Continuante Genericamente Dependente) e *Specifically Dependent Continuant* (Continuante Especificamente Dependente). (ARP; SMITH, 2011). Sendo:

- a) **Continuante Independente** são os objetos que servem como portadores de continuantes dependentes. As pessoas são continuantes independentes, pois elas existem por si mesmas, ou seja, elas perduram;
- b) **Continuante Genericamente Dependente** pode existir em diversos portadores, não sendo específicos a apenas um. Itens documentais e bases de dados são exemplos de continuantes genericamente dependentes;
- c) **Continuante Especificamente Dependente** não pode ser migrado, pois é específico para um dado objeto, como por exemplo: qualidades, funções e papel. “A” e “B” são pessoas (continuantes independentes) que possuem

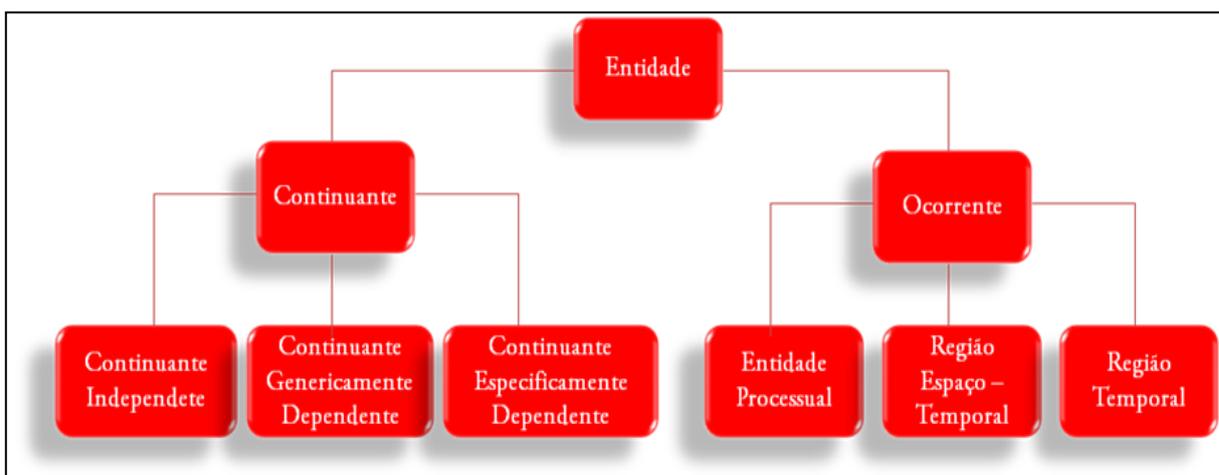
formações diferentes, caracterizando os papéis de aluna e de professor (continuintes especificamente dependentes), respectivamente. “A” e “B” não dependem dos papéis, mas os papéis de Aluna e de Professor dependem de “A” e “B”.

Pertencem à categoria Ocorrente as classes: *Processual Entity* (Entidade Processual), *Spatiotemporal Region* (Região Espaço-temporal) e *Temporal Region* (Região Temporal). (ARP; SMITH, 2011). Sendo:

- a) **Entidade Processual** são os processos que ocorrem em um determinado espaço de tempo tendo, portanto, início, meio e fim. São exemplos: eventos e ações;
- b) **Região Espaço-Temporal** inclui a localização temporal de um objeto em um local, como por exemplo: um organismo vive em lugar específico por um tempo determinado;
- c) **Região Temporal** representa períodos de tempo. São exemplos, o tempo de uma aula; o tempo que uma pessoa permanece hospedada em um hotel, em uma determinada cidade.

A Figura 7 apresenta de maneira sucinta a estrutura da BFO.

Figura 7 – BFO



Fonte: autora (2014).

BFO é uma ontologia amplamente utilizada por ontologias da área biomédica.

3.1.4.2 Friend of a Friend (FOAF)

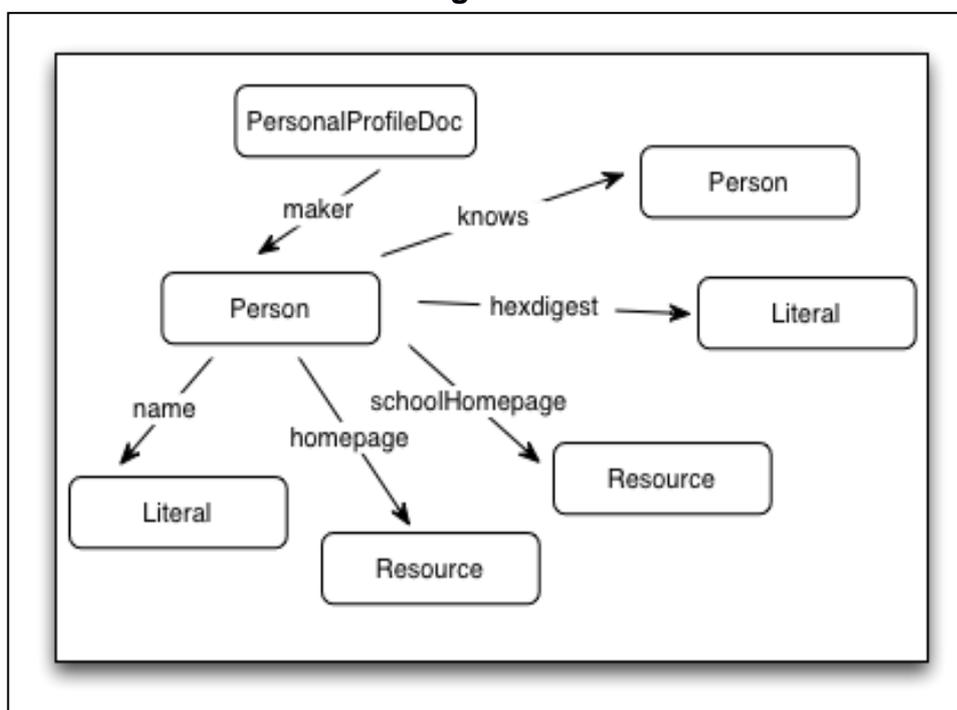
FOAF é um projeto criado por *Dan Brickley e Libby Miller* em 2000. Pode-se afirmar que este projeto é uma ontologia, pois sua função é descrever pessoas na *web*, sendo

uma tecnologia simples que faz com que seja mais fácil compartilhar e utilizar as informações sobre as pessoas e suas atividades (fotos, calendários, blogs), para transferir informações entre sites, e para automaticamente estender, fundir e ser reutilizada online. (FOAF, [200-?] *online* tradução nossa).

FOAF é uma ontologia bastante útil, pois gerencia comunidades, expressa identidade e indica autoria (DING et al. [200-?] tradução nossa *online*).

A Figura 8 mostra as relações e os atributos de uma pessoa a partir do uso da ontologia FOAF.

Figura 8 - FOAF



Fonte: Weitzner et al. (2006).

Atualmente, existem 16 classes, 32 propriedades de objetos e 26 propriedades de dados.

3.1.4.3 VCard

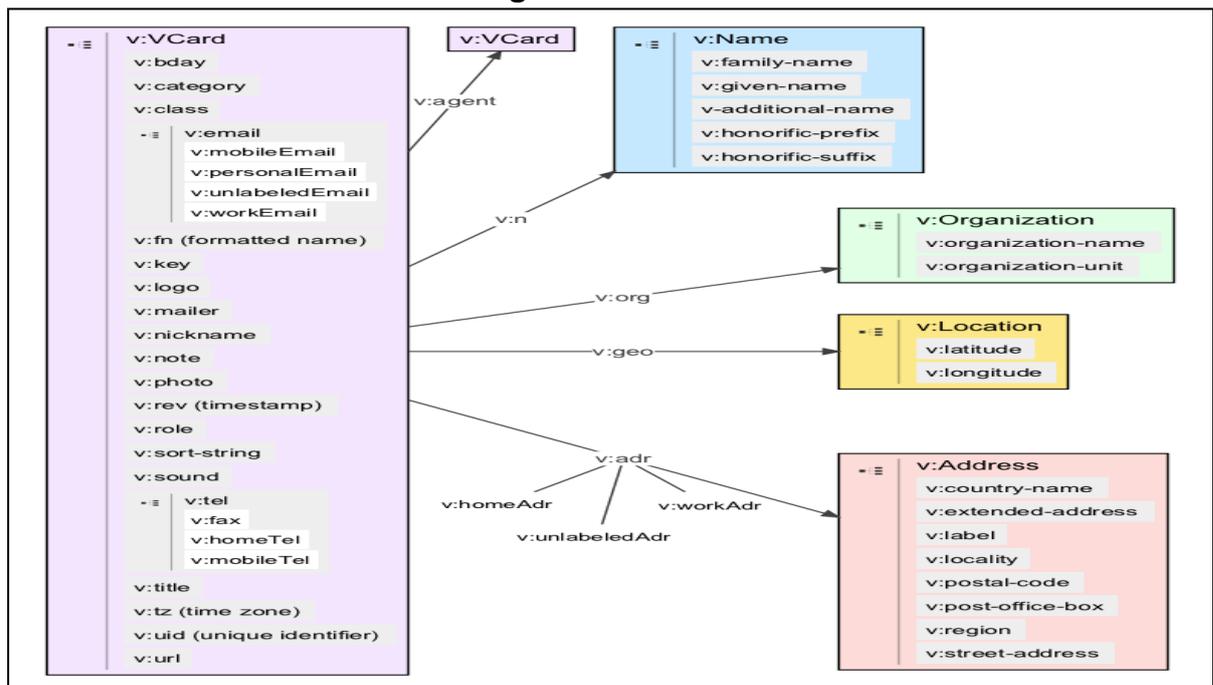
VCard é uma ontologia cujo objetivo é descrever pessoas e organizações.

Há semelhanças entre as ontologias FOAF e Vcard, pois ambas tratam da descrição de pessoas. Entretanto, de acordo Ianelle e McKinney (2014, documento não paginado, *online*, tradução nossa):

[...] a Ontologia VCard também enfocou a descrição de pessoas e organizações, incluindo informações sobre a localização e os grupos de tais entidades. A ontologia FOAF concentra-se mais sobre as relações entre as pessoas, agentes, coisas e entidades sociais da web [...]. (IANELLA; MCKINNEY, 2014, documento não paginado, tradução nossa *online*).

A Figura 9 exemplifica o uso de VCard. Nota-se que são descritas informações sobre contatos, data de nascimento e endereço. Essa ontologia é uma espécie de cartão virtual de visitas.

Figura 9 - VCard



Fonte: Vivo (2013).

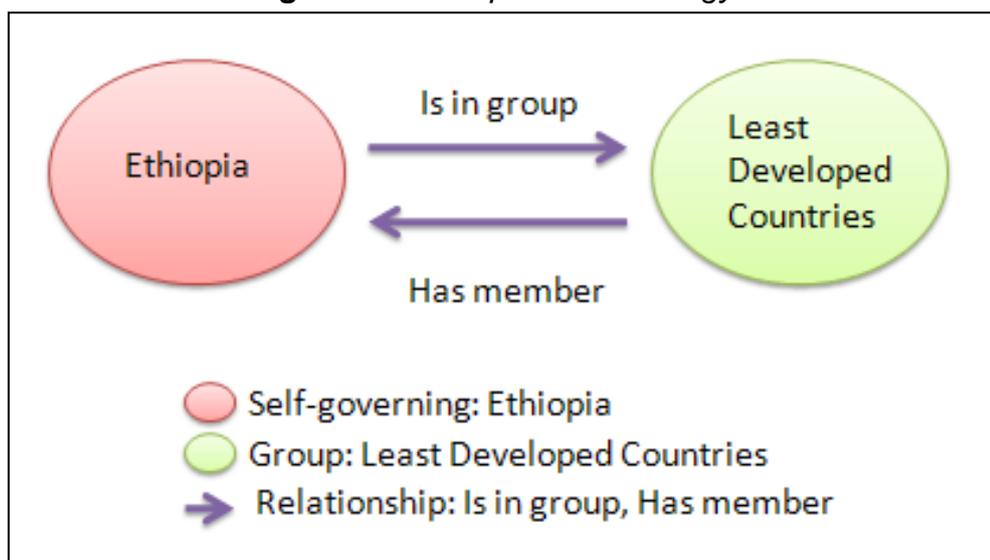
Atualmente, a ontologia é composta por 62 classes, 51 propriedades de objetos e 33 propriedades de dados.

3.1.4.4 Geopolitical Ontology

Esta ontologia foi lançada em 2008 pela *Food and Agriculture Organization of the United Nations* (FAO) como um mecanismo para promover interoperabilidade e relações semânticas entre países e entidades geográficas. (IGLESIAS-SUCASAS; KIM; VIOLLIER, [200-?], p.2, tradução nossa, *online*). *Geopolitical Ontology* foi desenvolvida para descrever áreas geopolíticas.

A Figura 10 exemplifica uma relação entre um país e o grupo ao qual pertence.

Figura 10 – Geopolitical Ontology



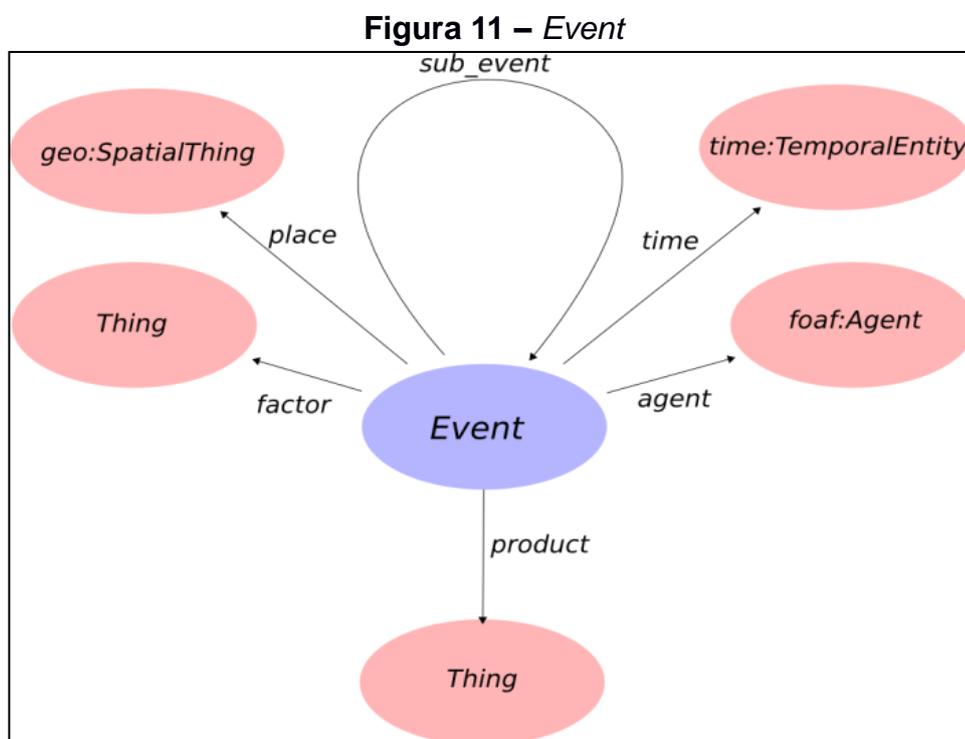
Fonte: Wikipedia (2013).

Para Iglesias-Sucasas, Kim e Viollier ([200-?], p. 2, tradução nossa, *online*) a Ontologia Geopolítica FAO foi desenvolvida para: “[...] melhorar a gestão da informação, a fim de facilitar compartilhamento de dados e interoperabilidade, e entregar as informações mais atualizadas relacionadas à geopolítica.”.

3.1.4.5 Event

A ontologia *Event* foi desenvolvida no *Centre for Digital Music* na Universidade de Londres no ano de 2004. Sua finalidade é descrever recursos relacionados a eventos, tais como: conversas em uma conferência, descrição de um concerto, festivais, dentre outros. (RAIMOND; ABDALLAH; 2007, p.2, tradução nossa, *online*).

A Figura 11 abaixo exemplifica a estrutura de *Event* e o uso de outras ontologias.



Fonte: Event (2007).

Duas características desta ontologia merecem ser destacadas: a facilidade e a simplicidade no uso, já que são poucas as classes que a compõem.

3.1.4.6 Simple Knowledge Organization System (SKOS)

SKOS é uma ontologia que utiliza um modelo de dados comum cujo objetivo é o compartilhamento e *linkagem* de sistemas de organização do conhecimento, via Web Semântica.

Atualmente, a ontologia é composta por 4 classes e 28 propriedades.

Para Peroni e Shotton (2011, documento não paginado, tradução nossa, *online*):

The Simple Knowledge Organization System (SKOS) é um modelo RDFS para apoiar o uso de sistemas de organização do conhecimento (KOS), tais como dicionários, esquemas de classificação, listas de cabeçalhos de assuntos e taxonomias no âmbito da Web Semântica. (PERONI; SHOTTON, 2011, documento não paginado, tradução nossa *online*).

SKOS não é uma linguagem para a modelagem de ontologias, mas um modelo de dados.

3.1.4.7 Bibliographic Ontology (BIBO)

BIBO tem por finalidade descrever recursos bibliográficos na Web Semântica. Porém, essa ontologia também pode ser utilizada de outras formas. Conforme *The Bibliographic Ontology* (2014, documento não paginado, *online*): “Esta ontologia pode ser usada como uma ontologia para citação, para classificação de documentos, ou simplesmente como um meio para descrever qualquer tipo de documento em RDF, como por exemplo: artigo, livro, patente, manual, tese, dentre outros.

BIBO é um vocabulário de termos específicos para metadados bibliográficos relacionados a documentos legais e a alguns tipos de eventos. (PERONI; SHOTTON, 2011, documento não paginado, tradução nossa, *online*).

Atualmente, a ontologia possui 69 classes, 52 propriedades de objetos e 54 propriedades de dados.

3.1.4.8 FRBR-aligned Bibliographic Ontology (Fabio)

FaBIO é uma ontologia alinhada com o *Functional Requirements for Bibliographic Records* (FRBR).

Para Peroni e Shotton (2011, documento não paginado, tradução nossa, *online*) FaBIO é uma ontologia para:

[...] a gravação e publicação na Web Semântica de descrições de entidades que são publicadas ou potencialmente publicáveis, e que contêm ou são referidos por referências bibliográficas, ou entidades usadas para definir tais referências bibliográficas. Entidades FaBIO

são publicações essencialmente textuais, como livros, revistas, jornais e periódicos, e itens de seu conteúdo, como poemas, textos de conferências e editoriais. No entanto, elas também incluem blogs, páginas web, bases de dados, algoritmos de computador, protocolos experimentais, especificações formais e vocabulários, registros legais, documentos governamentais, relatórios técnicos e comerciais e publicações similares, e também antologias, catálogos e coleções similares. (PERONI; SHOTTON; 2011, documento não paginado, tradução nossa *online*).

As classes de FaBIO estão organizadas de acordo com o esquema FRBR, sendo: Obra, Expressão, Manifestação e Item. Atualmente, a ontologia possui 237 classes e 27 propriedades de objeto.

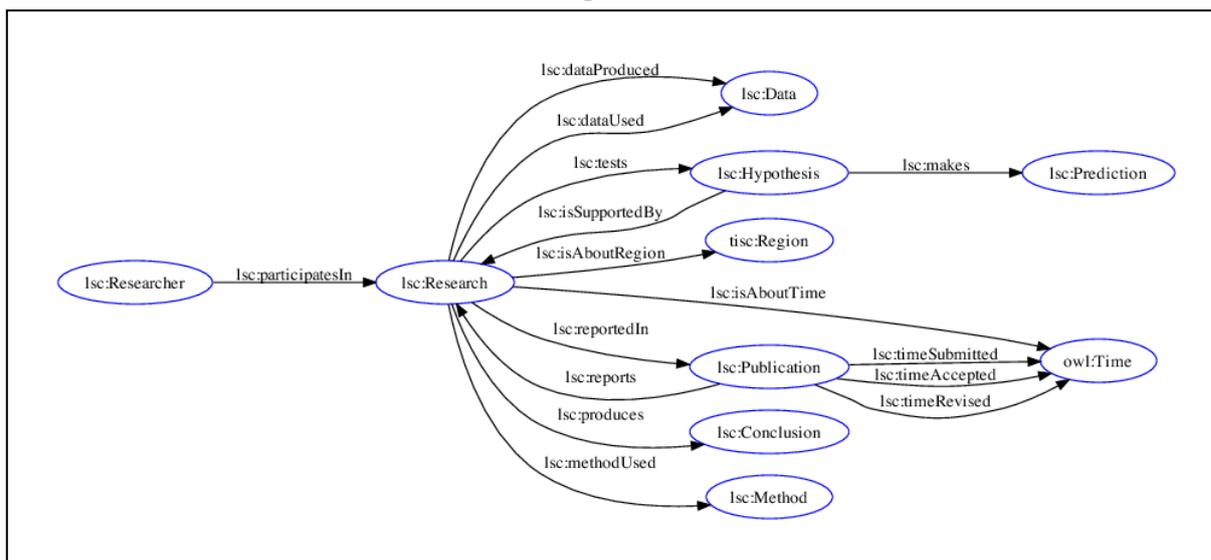
3.1.4.9 *Linked Science Core Vocabulary Specification (LSC)*

Ontologia que descreve recursos ligados à pesquisa. De acordo com Baglatzi, Kauppinen e Kebler (2011, documento não paginado, tradução nossa, *online*):

LSC, *Linked Science Core Vocabulary Specification*, é um vocabulário leve que proporciona condições para permitir que os editores e pesquisadores relacionem as coisas da ciência em tempo, espaço e temas. Mais precisamente, LSC é projetado para descrever recursos científicos, incluindo elementos de pesquisa, seu contexto, e sua interligação. (BAGLATZI; KAUPPINEN; KEBLER, documento não paginado, tradução nossa, *online*).

A Figura 12 apresenta a estrutura da ontologia LSC.

Figura 12 - LSC



Fonte: Baglatzi, Kauppinen e Kebler (2011).

Conforme consta na Figura 12, são representadas as relações entre o pesquisador e a pesquisa (etapas, hipóteses, método, publicação, dados, região e conclusões).

Atualmente, a ontologia possui 7 classes e 17 propriedades.

3.1.4.10 Academic Institution Internal Structure Ontology (AIISO)

A ontologia AIISO descreve recursos ligados à área acadêmica, mais especificamente, a estrutura interna da instituição.

Conforme Styles e Shabir (2008, documento não paginado, tradução nossa, *online*):

A Academic Institution Internal Structure Ontology (AIISO) fornece classes e propriedades para descrever a estrutura organizacional interna de uma instituição acadêmica. AIISO é projetado para trabalhar em parceria com Esquema de Participação [...], FOAF [...] e Papéis - AIISO [...] para descrever os papéis que as pessoas desempenham dentro de uma instituição. (STYLES; SHABIR, 2008, documento não paginado, tradução nossa, *online*).

Atualmente, a ontologia é composta por 15 classes e 10 propriedades.

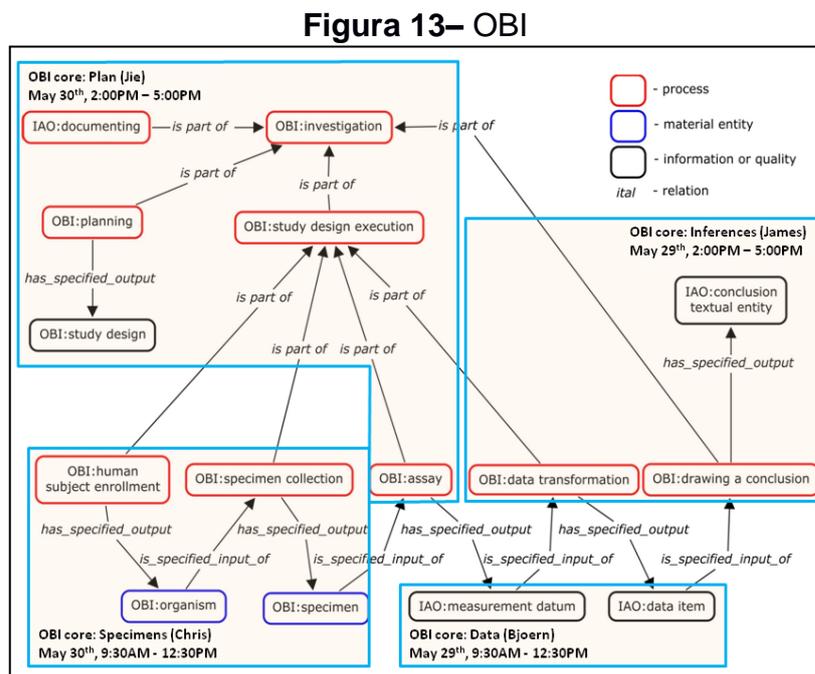
3.1.4.11 Ontology for Biomedical Investigations (OBI)

A ontologia faz parte de um Consórcio com o mesmo nome, o qual tem por objetivo, colaborar na construção de uma ontologia na área biomédica. A ontologia está começando a utilizar a *Basic Formal Ontology* (BFO).

OBI tem por objetivo

o desenvolvimento de uma ontologia integrada para a descrição de investigações biológicas e clínicas. Isto inclui um conjunto de termos "universais", que são aplicáveis em vários domínios biológicos e tecnológicos, e os termos específicos do domínio que são relevantes apenas para um determinado domínio. Esta ontologia irá apoiar a anotação constante de investigações biomédicas, independentemente do campo específico de estudo. (THE ONTOLOGY..., 2014, tradução nossa *online*).

A Figura 13 mostra um exemplo do uso da ontologia OBI.



Fonte: Philly (2013).

Como pode ser observado na Figura 13, as propriedades representam as relações entre os processos, as entidades materiais e as qualidades.

O escopo da ontologia abrange materiais biológicos, instrumentos, conteúdo informacional, desenho e execução de uma investigação e transformação de dados.

As ontologias são ferramentas bastante úteis, pois descrevem os mais variados recursos, representam os conceitos em uma linguagem formal (inteligível por máquinas e por humanos) e promovem a interoperabilidade semântica.

3.2 WEB SEMÂNTICA

A segunda geração da *web*, apesar de oferecer novas ferramentas, serviços e muitas vantagens, foi implementada de forma descentralizada, além de ter crescido de maneira exponencial, ou seja, essa *web* se apresenta como um enorme repositório de documentos (SOUZA; ALVARENGA, 2004). Conseqüentemente, as pesquisas realizadas em motores de buscas atuais são, em grande parte, insatisfatórias, pois a recuperação da informação é realizada através do uso de palavras-chave, o que dificulta a precisão nas consultas.

Para Souza e Alvarenga (2004) a limitação dos motores de busca suscitou a necessidade de se criar:

Uma padronização de tecnologias, de linguagens e de metadados descritivos, de forma que todos os usuários da *Web* obedçam a determinadas regras comuns e compartilhadas sobre como armazenar dados e descrever a informação armazenada e que esta possa ser 'consumida' por outros usuários humanos ou não, de maneira automática e não ambígua." (SOUZA; ALVARENGA, 2004, p, 134).

Essa padronização tecnológica é uma das principais características da Web Semântica, que pode ser entendida como uma extensão da *web* atual, ao invés de uma separação da mesma. (BERNERS-LEE; HENDLER; LASSILA, 2001, tradução nossa).

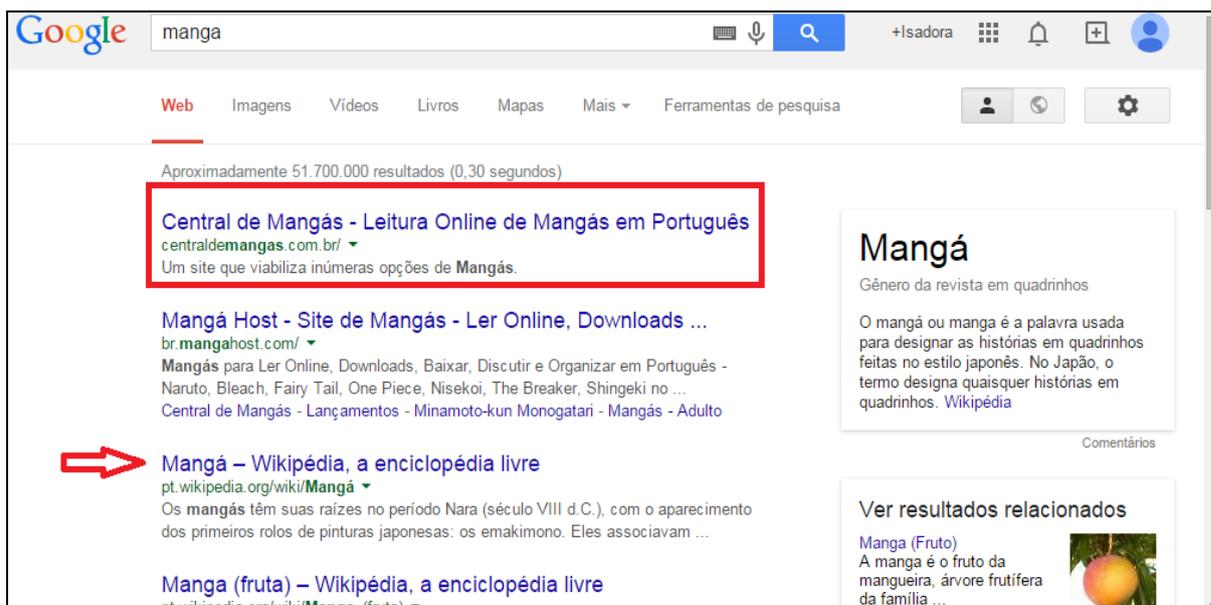
Vive-se em uma "era hipertextual" onde tudo e todos estão ligados e interconectados através de *links*. A linearidade passa a ser coadjuvante e a palavra **ligação** passa a ter o papel principal. Para Palmer (2001, documento não paginado, tradução nossa) a Web Semântica é:

[...] uma malha de informação ligada de maneira que seja facilmente processável por máquinas, em uma escala global e pode ser pensada como sendo uma forma eficiente de representar dados na web ou como uma base de dados globalmente vinculada. (PALMER, 2001, documento não paginado, tradução nossa *online*).

A Web Semântica vai ao encontro de uma das principais preocupações e deveres de um profissional da informação: a recuperação da informação. Assim sendo, a atuação do profissional vai além dos tradicionais catálogos *online* de bibliotecas, pois a informação na *web* também precisa ser tratada, organizada e disseminada.

Alguns motores de busca atuais utilizam os chamados *web crawlers* (rastreadores *web*), que nada mais são do que indexadores automáticos. A indexação das palavras é feita a partir da ocorrência, ou seja, esses rastreadores buscam determinada palavra, não importando o local da página em que se encontre, nem o contexto. A Figura 14 traz um exemplo de uma busca feita no *Google*.

Figura 14 – Exemplo de busca no *Google*



Fonte: Google (2014).

A Figura 14 mostra uma busca realizada no motor de busca *Google* com a palavra “manga”. Como pode ser observado, surgiram dois tipos de resultados: em primeiro lugar, o mangá (história em quadrinho em estilo japonês) e depois a fruta manga. Por ser desprovido de semântica, o buscador não consegue diferenciar a manga fruta da manga da camisa, nem do mangá.

Porém, a Web Semântica não é mais uma utopia, mas uma realidade cada vez mais presente, seja em bases de dados, seja em ambientes semânticos que promovem a descoberta e o compartilhamento. Inclusive, já existem alguns motores

de busca que utilizam semântica em sua estrutura, tais como: *LexxeBeta*¹, *SenseBot*², *Watson*³, *Exalead*⁴ e *Factbites*⁵.

A Figura 15 mostra um exemplo de uma pesquisa feita com o *Factbites*, pois trata-se de um motor de busca semântico atual.

Figura 15 - Exemplo de busca no *Factbites*

The screenshot shows the Factbites website interface. At the top left is the logo 'factbites' with the tagline 'Where results make sense'. To the right of the logo are navigation links: 'About us', 'Why use us?', 'Reviews', 'PR', and 'Contact us'. Below the logo is a search bar containing the word 'Mango' and a 'Find »' button. The main content area is titled 'Topic: Mango'. It features a list of search results, each starting with a green square icon containing a white letter 'i'. The first result is titled 'Mango' and contains three bullet points: 'Most Indian mangos are monoembryonic...', 'Non-fibrous mangos may be cut in half to the stone...', and 'Mango kernel decoction and powder...'. The second result is titled 'MANGO Fruit Facts' and contains three bullet points: 'Mangos may be pruned to control size...', 'Mango trees are very sensitive to root loss...', and 'Mangos ripen in June...'. The third result is titled 'Kate's Global Kitchen' and contains two bullet points: 'In Thailand, green mango slices are dipped...' and 'Ripe mangoes are soft...'. On the right side of the page, there is a 'Related Topics' section with a list of items: 'Alphonso (mango)', 'Fruit', 'Durian', 'Rambutan', 'Passion fruit', 'Jackfruit', 'Strawberry', 'Dragonfruit', 'Guava', 'Sapodilla', 'Kiwi Fruit', 'Avacado', and 'Mangosteen'. Each item in the 'Related Topics' list is preceded by a green square icon with a white letter 'i'.

Fonte: Factbites (2014).

A busca foi feita com o mesmo termo “manga” (nesse caso, em inglês, “mango”) e percebem-se algumas diferenças nos resultados dos dois motores de busca. Em *Factbites*, os resultados são organizados em categorias e são oferecidos alguns tópicos relacionados ao termo. A indexação também é feita pelas palavras-chave, entretanto, o diferencial do motor de busca está em outra forma de indexação: pela área do tópico. Dessa forma, são incluídos também os resultados que não mencionam o termo.

Existem alguns princípios que norteiam essa nova web e que merecem ser

¹<http://www.lexxe.com/index.html>

²<http://www.sensebot.net/>

³<http://watson.kmi.open.ac.uk/WatsonWUI/index.html>

⁴<http://www.exalead.com/search/web/>

⁵<http://www.factbites.com/>

destacados. De acordo com Miler e Koivunen⁶ (2002, p. 27-33 *apud* FEITOSA, 2006, p. 69):

1. Tudo pode ser representado por meio de URIs - é possível referenciar pessoas, lugares e coisas do mundo físico na *web* semântica, com a utilização de diversos identificadores unívocos;
2. Recursos e links podem ter tipos - a *web* atual consiste de *links* e recursos destinados à interpretação humana e, ainda não é amplamente difundido o uso de metadados para explicitar seu significado, bem como, suas relações com outros documentos, de modo que máquinas possam realizar inferências sobre o seu significado; tais relações podem ser, por exemplo: *depende de*, *é uma versão de*, *possui assunto*, *é autor de*, entre outras;
3. Informação parcial é tolerada- a *web* atual sacrifica a integridade dos *links*, em função de sua escalabilidade; de tal modo, páginas podem estabelecer ligações com outras sem a necessidade de que haja uma reciprocidade nessa ligação; de maneira similar, na *web* semântica, pode-se declarar qualquer coisa sobre qualquer coisa, sem que haja necessidade da criação de ligações entre os recursos declarados;
4. A verdade absoluta não é necessária - nem toda informação na *web* atual é verdadeira, e a *web* semântica não prevê mudanças nessa orientação;
5. A evolução é suportada - a *web* semântica utiliza convenções descritivas que podem expandir-se como a compreensão humana se expande; tais convenções permitem a combinação do trabalho independente e compartilhado entre diversas comunidades, mesmo quando estas utilizam vocabulários diferentes;
6. O projeto é minimalista - um dos objetivos para garantir o desenvolvimento da *web* semântica é padronizar apenas o que for necessário, favorecendo a implementação de aplicações flexíveis. (2002, p. 27-33 *apud* FEITOSA, 2006, p. 69).

A Web Semântica também é conhecida como Web dos Dados ou Web 3.0. Mas quais são os seus objetivos? Vários, mas os principais são: dar **significado** e **estruturação** aos dados; **organizar** seu conteúdo a partir do uso de **agentes inteligentes**; criar **conexões** para os usuários; **relacionar** bases de dados/ motores de busca e documentos; fazer com que **máquinas compreendam informações** até então compreensíveis apenas por seres humanos; propiciar um **ambiente** em que máquinas e pessoas possam trabalhar em conjunto.

Para que o *modus operandi* da Web Semântica seja efetivado, são utilizados alguns padrões e tecnologias, como: Agentes Inteligentes, *Uniform Resource*

⁶MILLER, E.; KOIVUNEN, M. WRC Semantic Web Activity. In: HYVONEN, Eero (editor). *Semantic Web Kick-Off in Finland*. Vision, technologies, research and applications. Helsinki, Finland: HIIT Publications. 2002. Disponível em: <<http://www.cs.helsinki.fi/u/eahyvone/stes/semanticwe/kickoff/index.html>>.

Identifier (URI), *Resource Description Framework* (RDF) e *Web Ontology Language* (OWL), além das ontologias.

3.2.1 Agentes Inteligentes

Os agentes inteligentes têm o intuito de organizar o conteúdo na Web Semântica, a partir de suas três funções: “Perceber dinamicamente o ambiente; tomar ações que afetem as condições do ambiente; entender as percepções, resolver problemas, construir inferências e determinar ações”. (HAYES-ROTH⁷, 1995 *apud* CRISPIM *et al.*, 2002, documento não paginado, *online*).

De acordo com Gonçalves e Carrapatoso (2009, p. 85 *online*): “Os agentes inteligentes tratam da recuperação da informação relevante aos utilizadores que dela necessitam e assistem-nos na tomada de decisão”.

Basicamente, um agente inteligente observa as ligações (percurso) que o usuário faz na *web*, e a partir disso, deduz o que interessa ao usuário. Os agentes inteligentes são fundamentais para uma melhor experiência do usuário, pois fazem com que a *web* fique personalizada de acordo com as preferências e os interesses de cada “navegador”.

3.2.2 Uniform Resource Identifier (URI)

Uniform Resource Identifier é um identificador persistente de objetos/recursos na *web*. Para Berners-Lee (1998, documento não paginado *online* tradução nossa) um URI é: “[...] um conjunto compacto de caracteres utilizado para a identificação de um recurso abstrato ou físico. Um URI fornece ao recurso um significado simples e extensível.”

Portanto, um URI possui um papel importante na Web Semântica, pois é a partir dele que os recursos serão identificados na *web*, assim como, um URI pode fornecer informações sobre outro. Segundo Feitosa (2006) é possível atribuir um URI para qualquer coisa, e qualquer coisa que possua um URI pode ser considerada como pertencente à *web*.

⁷HAYES-ROTH, B. An architecture for adaptive intelligent systems. **Artificial Intelligence**, v. 72, n. 1, p. 329-365, 1995.

3.2.3 Resource Description Framework (RDF) e Resource Description Framework Schema (RDF Schema)

Resource Description Framework (RDF) pode ser definido como um modelo padrão, ou uma estrutura, cujo objetivo é o intercâmbio de dados e informações na *web*.

RDF estende a estrutura de ligação da *web* para usar URIs para nomear a relação entre as coisas. Usando esse modelo simples, ele permite que dados estruturados e semi-estruturados possam ser misturados, expostos e compartilhados entre diferentes aplicações. (W3C, 2014, *online*, tradução nossa).

Os conceitos básicos de RDF, conforme Feitosa (2006) são: modelo de dados gráfico e vocabulário baseado em URI. O modelo gráfico é formado por **sujeito**, **predicado e objeto**, que juntos formam uma tripla. Para Rocha (2013, documento não paginado *online*):

O sujeito é o recurso da *web* que está sendo descrito, e deve ser identificado através da sua URI. O predicado indica uma propriedade, isto é, um aspecto, uma característica, atributo ou relação atribuída ao recurso pela sentença. O objeto é o valor dado ao recurso pela propriedade. (ROCHA, 2013, documento não paginado, *online*).

A Figura 16 abaixo mostra um exemplo de uma tripla RDF.

Figura 16 – Tripla RDF

Sujeito	Predicado	Objeto
www.sciam.com/article.cfm?id=the-semantic-web	dc:date	2001
www.sciam.com/article.cfm?id=the-semantic-web	rdf:type	bibo:AcademicArticle
www.sciam.com/article.cfm?id=the-semantic-web	dc:creator	www.w3.org/People/Berners-Lee/
www.w3.org/People/Berners-Lee/	rdf:type	foaf:Person
www.w3.org/People/Berners-Lee/	foaf:name	Tim Berners-Lee

Fonte: Rocha (2013).

Conforme Feitosa (2006, p. 111) RDF Schema: “[...] foi concebido para fornecer rigor sintático e semântico a um documento e funciona como uma extensão para o RDF, fornecendo a arquitetura para a descrição de classes e propriedades específicas [...]”.

RDF Schema é uma linguagem simples para representação de ontologias e é utilizada para definir e representar vocabulários específicos de um determinado domínio em RDF. A linguagem oferece uma estrutura para descrever propriedades,

classes e hierarquia de propriedades e de classes, além de restrições de propriedades.

3.2.4 Web Ontology Language (OWL)

Existem diversas linguagens para a construção de ontologias, entretanto, para este trabalho será abordada apenas a linguagem OWL, pois é a utilizada pelo sistema VIVO.

Web Ontology Language é uma linguagem para construção de ontologias que permite descrever não apenas aspectos semânticos, como também os relacionamentos existentes entre os conceitos de um determinado domínio, de uma forma mais abrangente. OWL oferece um vocabulário adicional com semântica formal, como por exemplo: relações entre classes, cardinalidade e características de propriedade.

Para Feitosa (2006, p.115):

Do mesmo modo que em RDF, a linguagem OWL utiliza estruturas de classes para a organização dos recursos. Assim, cada classe pode fornecer um mecanismo de abstração para o agrupamento de recursos e está associada a um conjunto de objetos individuais, a que se denomina *extensão* da classe. Cada objeto individual em uma classe é denominado instância de tal classe e possui um significado intensional. Devido a essa concepção, todos os recursos em um documento OWL são descritos como classes, subclasses ou propriedades de classes e subclasses. Isso garante, também, à linguagem, a possibilidade de representar significações com base em elementos que atuam como funções em algoritmos e que permitem a realização de inferências por agentes inteligentes. (FEITOSA, 2006, p. 115).

A linguagem OWL tem mais facilidades para expressar significado e semântica se comparada às linguagens XML, RDF e RDF Schema. Dessa forma, OWL supera as demais na capacidade de representar conteúdos na *web* interpretáveis por máquina. (W3C, 2004, *online*).

3.2.5 Linked Open Data (LOD)

Linked Data e *Linked Open Data* são termos usados como sinônimos. O primeiro significa dados ligados e o segundo, dados abertos ligados. Ambos transitam no universo da Web Semântica, por entre ontologias e linguagens.

Entretanto, para este trabalho, será adotado o termo *Linked Open Data*, por uma questão de preferência.

De acordo com Rocha (2012, p.282, *online*), *Linked Data* é

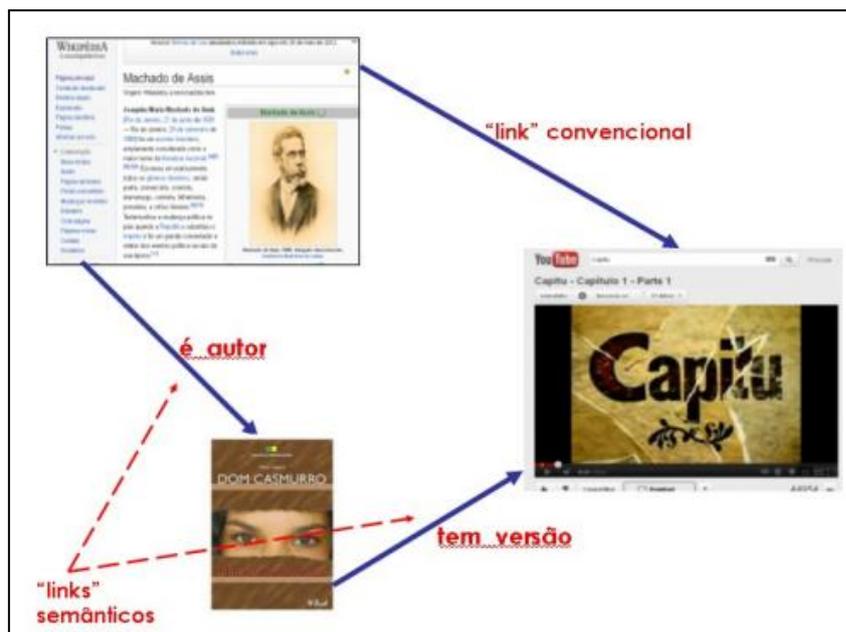
[...] o termo usado para descrever recomendações de melhores práticas para expor, compartilhar e conectar pedaços de dados e conhecimento na Web Semântica, usando *Uniform Resource Identifier* (URI) e RDF. Com o objetivo de permitir a ligação entre dados publicados por diversas organizações, *Linked Data* estabelece como princípio usar URIs para identificar unicamente as entidades representadas nesses conjuntos de dados. (ROCHA, 2012, p. 282, *online*).

LOD utiliza a Web Semântica como plataforma para a publicação de dados ligados a partir de *links* (URI). Cabe ressaltar que tais *links* não se conectam apenas a outros, mas também a recursos. O diferencial dos *links* está na semântica, pois expressam as relações entre os recursos a partir do uso das ontologias e de linguagens, como RDF.

Em *Linked Open Data*, entidades são representadas através de URIs. Uma pessoa, um objeto ou um artigo, é referenciado por um *link*, o qual, está relacionado a outros. Dessa forma, surge uma rede de dados, em que existe uma conexão semântica de ligações, muito mais estruturada do que uma simples relação hipertextual.

A Figura 17 traz um exemplo que mostra a diferença entre um *link* convencional e um *link* semântico.

Figura 17 – Links semânticos



Fonte: Marcondes (2012).

Como pode ser observado na Figura 17 os *links* semânticos proporcionam uma melhor experiência do usuário no momento da recuperação da informação.

Assim como na Web Semântica, *Linked Open Data* possui alguns princípios que norteiam a publicação dos dados. Segundo Berners-Lee, são:

- 1) Usar *Uniform Resource Identifier* para identificar os dados;
- 2) Utilizar o protocolo *Hypertext Transfer Protocol* (http) para facilitar a localização dos dados;
- 3) Fornecer informações úteis e utilizar padrões como o *Resource Description Framework* (RDF);
- 4) Incluir *links* para outras URIs para que os usuários possam descobrir mais informações relacionadas. (BERNERS-LEE, 2006, documento não paginado, tradução nossa *online*).

Os benefícios de LOD estendem-se aos consumidores e aos publicadores de dados.

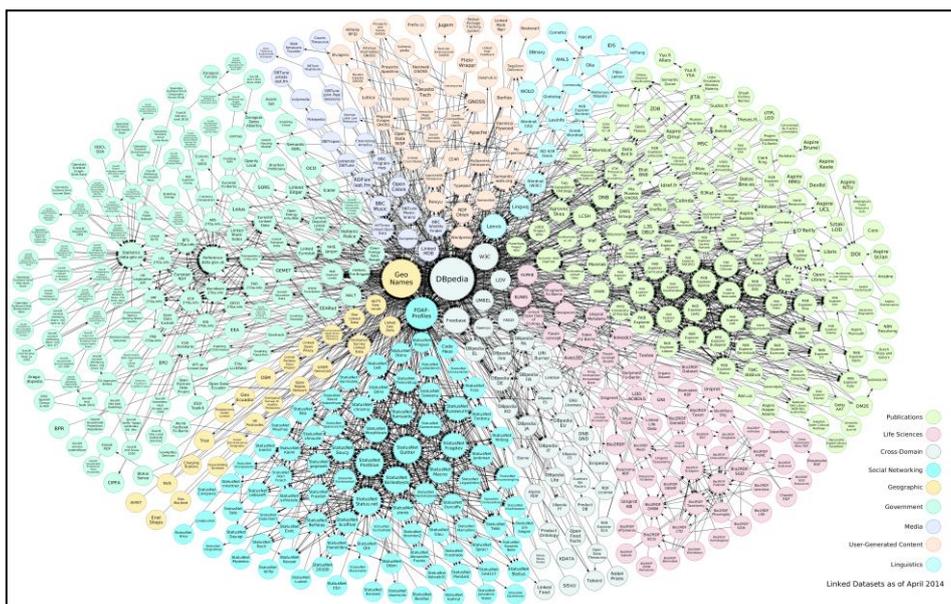
O Quadro 4 apresenta alguns dos benefícios proporcionados por LOD para consumidores e publicadores.

Quadro 4 – 5 estrelas e LOD

Estrelas	Consumidor	Publicador
	<ol style="list-style-type: none"> 1) Você pode vê-los. 2) Você pode imprimi-los. <p>Você pode armazená-los localmente (no seu disco ou em um USB).</p> <ol style="list-style-type: none"> 3) Você pode inserir os dados manualmente em outro sistema. 	É fácil publicar.
	<ol style="list-style-type: none"> 1) Você pode processá-los diretamente com um software proprietário para: agregá-los, realizar cálculos, visualizá-los, etc. 2) Você pode exportá-los em outro formato (estruturado). 	É fácil publicar.
	Você não tem que pagar por um formato sobre o qual uma única entidade mantém controle exclusivo.	É fácil publicar.
	<ol style="list-style-type: none"> 1) Você pode ligá-los de um lugar a outro, na web ou localmente. 2) Você pode marcá-los. 3) Você pode reutilizar partes dos dados. 	<ol style="list-style-type: none"> 1) Você terá que investir algum tempo para reordenar seus dados. 2) Você precisará atribuir URIs para itens de dados e pensar sobre como representá-los. 3) Você tem o controle sobre os itens de dados e pode otimizar seu acesso.
	<ol style="list-style-type: none"> 1) Você pode descobrir novos dados de interesse enquanto consome outras informações. 2) Você tem acesso ao esquema de dados. 	<ol style="list-style-type: none"> 1) Você terá de investir recursos para vincular os seus dados com outros na web. 2) Você faz os seus dados detectáveis. 3) Você aumenta o valor dos seus dados.

Fonte: adaptado de Bauer e Kaltenböck (2012).

Figura 19 - Nuvem LOD em 2014



Fonte: Schmachtenberg (2014).

Linked Open Data possibilita a **interoperabilidade semântica** de recursos da ciência que, atualmente, vão além dos bibliográficos, como: dados de pesquisa, pesquisadores, instituições, projetos, processos e eventos. Através da interoperabilidade os pesquisadores podem ter acesso ao que está sendo publicado na própria fonte e trocar informações entre seus pares.

O Quadro 5 apresenta alguns projetos que publicam seus dados em LOD e suas finalidades.

Quadro 5 - Linked Open Data na ciência

Projeto	Finalidade
<i>Europeana Linked Data Pilot</i>	Publicação dos dados da <i>Europeana</i> , um provedor de serviços que colhe metadados de 1500 instituições culturais na Europa.
<i>Semantic Web Dog Food</i>	Artigos, conferências, organizações e pessoas da área da Web Semântica.
<i>Faceted DBLP</i>	Publicação em <i>Linked Data</i> da base de dados bibliográfica DBLP da Ciência da Computação.
<i>Archaeology Data Service</i>	Publicação da base de dados <i>Archaeology Data Service</i> (ADS), que envolve recursos de pesquisa, aprendizado e ensino, preservando e disseminando dados de arqueologia.
<i>JISC Open Citations</i>	Publicação da base de dados <i>Open Citations</i> , uma base de dados de citações de literatura biomédica, colhidas a partir das referências de artigos abertos

	da PubMedCentral-UK.
LODUM	<i>Linked Open Data University of Muenster</i> . Ativos de ciência produzidos pela Universidade de <i>Muenster</i> .

Fonte: adaptado de Rocha (2013).

Linked Open Data surge como uma nova forma para a publicação dos dados, mesclando características e ferramentas da Web 2.0 e da Web Semântica: inteligência coletiva, *wikis*, compartilhamento, dados estruturados, *links* semânticos e padronização tecnológica.

4 METODOLOGIA

Nesta seção, serão abordados os princípios metodológicos utilizados para a realização dessa pesquisa.

4.1 NATUREZA, OBJETIVO, ABORDAGEM E PROCEDIMENTO TÉCNICO DA PESQUISA

A natureza e o objetivo da pesquisa foram, respectivamente, básica e exploratória. Básica porque não haverá, *a priori*, uma aplicação prevista dos resultados. Quanto ao objetivo, a pesquisa foi de cunho exploratório, pois pretendeu-se explicitar e conhecer a ontologia de VIVO. Para Gil (2009, p. 41) a pesquisa exploratória tem como objetivo:

[...] proporcionar maior familiaridade com o problema, com vistas a torná-lo mais explícito ou a constituir hipóteses. Pode-se dizer que estas pesquisas têm como objetivo principal o aprimoramento de ideias ou a descoberta de intuições. (GIL, 2009, p. 41).

Abordagem qualitativa porque foi efetuada uma análise global de como a ontologia do sistema VIVO “enxerga” a ciência, sem a pretensão de quantificar ou tabelar os dados coletados.

O procedimento técnico adotado foi o Estudo de Caso. Conforme Gil (2009, p.54) Estudo de Caso: “[...] consiste no estudo profundo e exaustivo de um ou poucos objetos, de maneira que permita seu amplo e detalhado conhecimento [...]”.

Sabe-se que este procedimento limita o estudo, já que, a pesquisa é direcionada a um único objeto de investigação. Entretanto, esse é o procedimento mais adequado para a finalidade do presente trabalho, visto que, o foco foi a análise estrutural e conceitual da ontologia de VIVO.

4.2 CORPUS DA PESQUISA

A escolha do *corpus* foi realizada por etapas. Inicialmente, foi realizado um levantamento no motor de busca *Google Scholar*, nas bases de dados LISA e BRAPCI e no repositório E-Lis com as expressões de busca “*linked open data*” and

“*project*” e “*dados abertos ligados*” and “*projeto*”. Por tratar-se de um tema atual e em constante atualização foram utilizados os seguintes filtros, tanto para a pesquisa bibliográfica, quanto para a construção do referencial teórico.

- a) **temporalidade:** de 2000 a 2014;
- b) **idiomas:** inglês e português (preferencialmente o primeiro);
- c) **fontes de informação:** teses, dissertações, artigos de periódicos, livros, anais de eventos, relatórios técnicos e *sítes* de projetos relacionados ao tema da pesquisa.

Primeiramente, foi selecionado o projeto VIVO. A escolha justifica-se pelo fato de VIVO ser uma rede formada por cientistas, de diversas áreas do conhecimento, interessados em compartilhar recursos e descobertas da ciência; e por ser um ambiente conhecido e utilizado por universidades internacionais de prestígio.

A partir da escolha do projeto VIVO foi selecionada a sua ontologia (VIVO-ISF) para ser o *corpus* dessa pesquisa, a fim de verificar sua adequação a outros ambientes que visam descrever recursos do domínio da ciência.

4.3 INSTRUMENTO DE COLETA DE DADOS

Os dados foram coletados a partir do uso do formulário de coleta de dados (APÊNDICE). O instrumento serviu como um roteiro, ou seja, as questões não tiveram o propósito de serem respondidas formalmente pelo pesquisador (como em um questionário), mas com o intuito de guiar a investigação. As onze perguntas que compõem esse formulário estão divididas em quatro categorias: escopo, fundamentação, evolução e reuso.

4.4 PROCEDIMENTOS DE COLETA DE DADOS

O Quadro 6 estabelece as relações entre os objetivos específicos e as questões do formulário utilizadas para a coleta dos dados.

Quadro 6- Coleta dos dados

Objetivos específicos	Questões
a) Investigar o escopo da ontologia e sua adequação de recursos, em âmbito geral, no domínio da ciência.	1, 2 e 3
b) Investigar os fundamentos na elaboração da ontologia.	4, 5 e 6
c) Investigar o processo de evolução da ontologia.	7 e 8
d) Investigar o reúso de ontologias visando a interoperabilidade semântica.	9, 10, e 11

Fonte: autora (2014).

4.5 ANÁLISE DOS DADOS

A análise da ontologia VIVO-ISF foi realizada em quatro etapas. A primeira etapa investigou o escopo da ontologia VIVO-SF visando identificar as intenções de uso, a abrangência e quem a utiliza. A segunda etapa investigou o uso da ontologia de fundamentação BFO e suas principais modificações na estrutura da ontologia de VIVO. A terceira etapa investigou a evolução da ontologia e as principais mudanças ocorridas em cada versão. A quarta etapa investigou o reúso de ontologias conhecidas no, cenário *linked open data*, em VIVO-ISF.

5 RESULTADOS

Esta seção apresenta os resultados obtidos durante a realização da pesquisa.

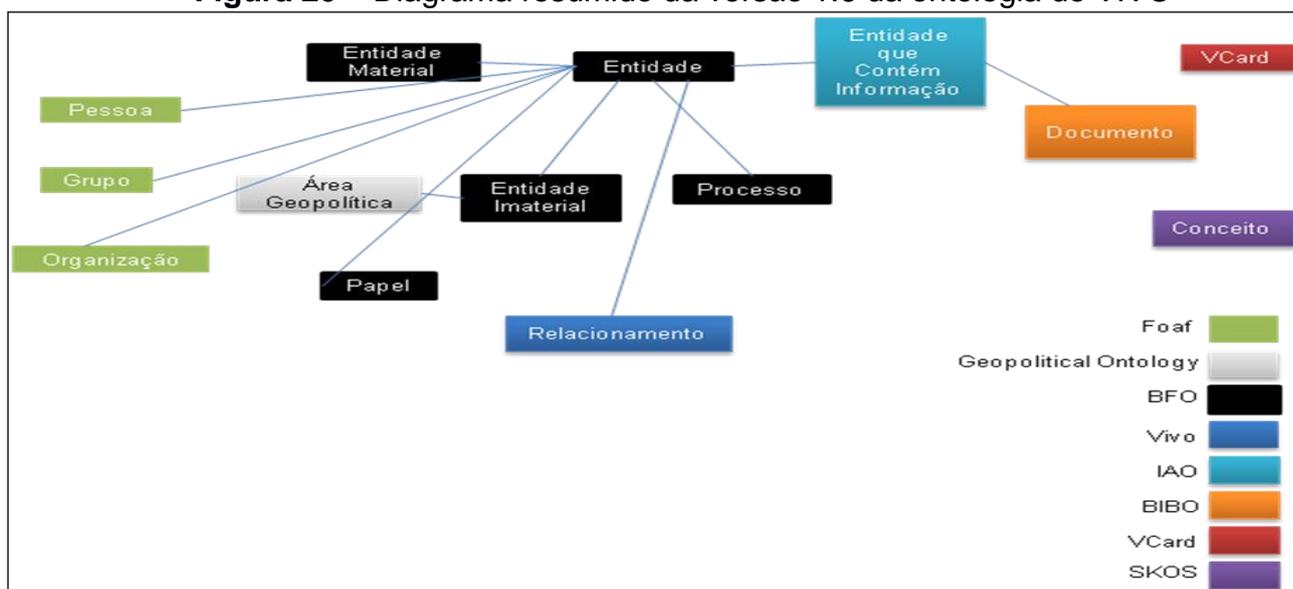
5.1 ESCOPO DE VIVO-ISF

Esta subseção investiga o escopo da ontologia VIVO – ISF (*Integrated Semantic Framework*) visando identificar a sua adequação para descrição de recursos do domínio da ciência. A investigação baseia-se na análise de seu possível uso em um sistema para descrever recursos de qualquer área da ciência.

Para analisar o escopo é fundamental compreender de que forma VIVO enxerga a ciência. A ontologia de VIVO foi criada com o intuito de modelar perfis de pesquisadores e descrever recursos de pesquisa clínica. A modelagem é feita estruturalmente e de forma individual. Na estrutura são utilizadas ontologias de cunho mais geral (pessoas, organizações, documentos). Na forma individual são utilizadas ontologias que descrevem recursos de: pesquisa, ensino, serviços e *expertise*. Na descrição de recursos de pesquisa clínica são utilizadas ontologias da área biomédica.

A Figura 20 traz o diagrama resumido da versão 1.6 da ontologia de VIVO e suas ontologias reutilizadas.

Figura 20 – Diagrama resumido da versão 1.6 da ontologia de VIVO



Fonte: autora (2014).

A ontologia VIVO-ISF tem como base a classe Entidade, a qual, é responsável por relacionar e integrar as demais classes. Os recursos estão representados conforme: Agentes, Área Geopolítica, Relacionamento, Papel, Processo, Entidades Material e Imaterial, Entidade que Contém Informação e Documento. VCard e Conceito complementam a estrutura integrada da ontologia. O primeiro representa as características dos agentes (endereço, identificação) e o segundo, os conceitos dos recursos (definição de Papel, Relacionamento, dentre outros).

Conforme observado na seção 2 (contexto do estudo), VIVO tem por objetivos a modelagem de perfis de pesquisadores e a descrição de recursos de pesquisa clínica. Dessa forma, a ontologia está integrada de acordo com esses objetivos, haja visto que uma das funções de uma ontologia é a representação formal dos conceitos em um artefato tecnológico.

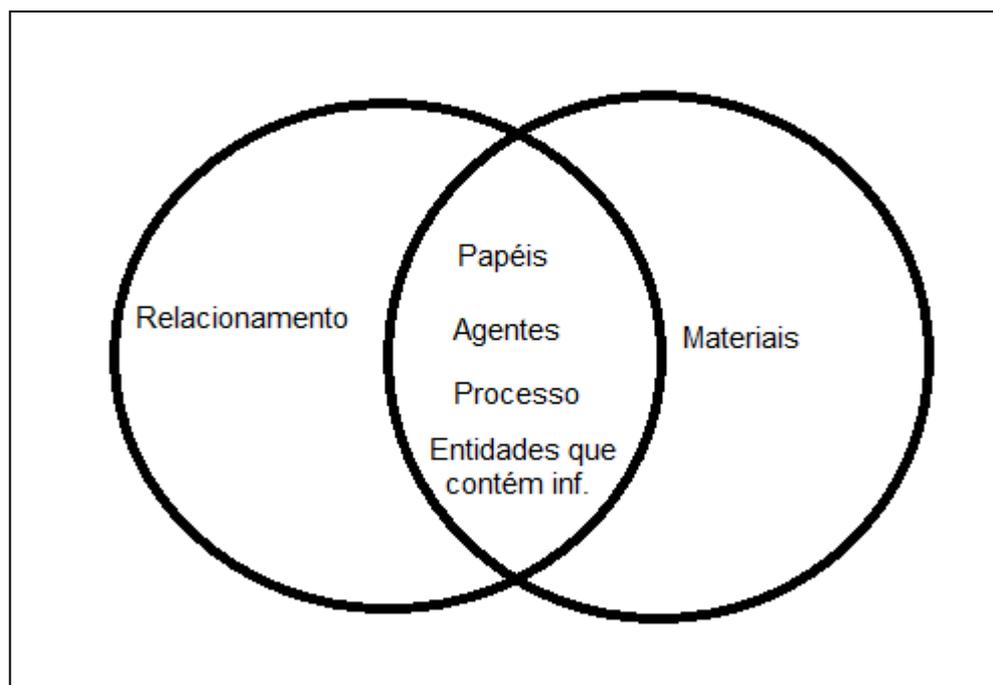
Agentes são as pessoas, grupos e organizações pertencentes à uma universidade ou a um centro de pesquisa. Esses agentes possuem **papéis** de professor, de aluno, de administrador, de pesquisador, dentre outros. Entretanto, para que esses papéis sejam efetivados é necessário que haja um **relacionamento** entre os agentes, como ocorre em uma relação de orientação entre um orientando e um orientador. Os relacionamentos são **processos** que ocorrem entre agentes e agentes e agentes e instituições, como: editoria, orientação e autoria. Os agentes e seus papéis, os relacionamentos e os processos ocorrem em uma determinada **área geopolítica**, isto é, em determinado local e em um determinado tempo. A área geopolítica é uma **entidade imaterial**, assim como o *site* de uma universidade e uma região espaço-temporal. Assim como há a entidade imaterial, há a **entidade material**, que são os reagentes e o maquinário utilizado na realização de uma pesquisa. São produtores e produtos de atividades de ensino e pesquisa, a **entidade que contém informação**, sendo as bases de dados e os dados de pesquisa que produzem os documentos utilizados e gerados durante esse processo.

Assim sendo, é possível também realizar uma análise sob o viés acadêmico. Sabe-se que os pilares de uma universidade são o ensino, a pesquisa e a extensão. A partir dessa tríade podem ser investigados os conceitos que pertencem ao fazer universitário.

A extensão é representada apenas pela subclasse *Unit Extension* (Unidade de Extensão), o que seria equivalente a uma Pró-Reitoria de Extensão. O ensino e a pesquisa possuem conceitos em comum e os que são específicos a cada um deles.

A Figura 21 mostra os principais conceitos que representam o ensino e a pesquisa.

Figura 21 - Conceitos em VIVO-ISF



Fonte: autora (2014).

Os conceitos são representados na ontologia VIVO-ISF por: Agentes, Papéis, Relacionamento, Processo, Entidades que contém Informação e Materiais.

Agentes: Os agentes são responsáveis, simultaneamente, pelo ensino e pela pesquisa. São exemplos de agentes: pessoas, grupos e organizações que pertencem a uma instituição. Cabe ressaltar que as subclasses Professor Emérito, Bibliotecário e Estudante são consideradas tipos de pessoas, ao invés de papéis.

Papéis: Esse conceito está diretamente ligado ao relacionamento, pois em um papel de orientador está implícita uma relação de orientação entre o orientador e seu orientando. A classe representa os papéis que os agentes assumem em atividades de ensino e pesquisa, como por exemplo: Educador, Orientador, Editor, Pesquisador e Estudante;

Relacionamento: Essa classe está relacionada ao ensino, descrevendo as relações entre os papéis dos agentes envolvidos. A relação de orientação só ocorre

se houver a participação do orientador (professor) e do orientando (aluno). São exemplos de relações: autoria, orientação, contrato e editoria;

Processo: Esse conceito representa recursos que possuem uma temporalidade (tempo de duração), ou seja, que tenham início, meio e fim. O processo ocorre tanto no ensino, quanto na pesquisa. São exemplos de processo: eventos, projetos, processo educacional e processo de orientação.

Entidades que contém informação: Esse conceito abrange a produção científica e acadêmica dos pesquisadores. As entidades estão presentes no ensino e na pesquisa, pois são fontes de informação e produção intelectual ao mesmo tempo. São exemplos: recursos bibliográficos, bases de dados e dados de pesquisa;

Materiais: O conceito de materiais abrange os recursos necessários para o desenvolvimento da pesquisa clínica, mais especificamente, da área da Biomedicina. Essa classe está diretamente relacionada à pesquisa. São exemplos de materiais: espécie biológica, organismo, instrumento, reagente e equipamento.

A ontologia de VIVO é utilizada pelas instituições pertencentes ao sistema: *University of Florida, Cornell University, Indiana University, Washington University School of Medicine in St. Louis, The Scripps Research Institute, Weill Cornell Medical College, and Ponce Medical College*. E é objeto de estudo de grupos de pesquisa na Tailândia e na Austrália. No grupo da Tailândia, partes da ontologia estão sendo estudadas no protótipo de desenvolvimento de uma rede de pesquisadores chamada *Linked OpenScholar*⁸. No grupo da Austrália, a ontologia está sendo estudada para ser utilizada na Iniciativa *Australian Nacional Data Service (ANDS)*⁹.

A partir das análises, verifica-se que a ontologia é do domínio da ciência, tendo abrangência **geral** e **específica**. É de abrangência geral, pois descreve diversos tipos de recursos, sendo assim, aplicáveis a qualquer área da ciência. Específica, pois também descreve alguns tipos de recursos específicos ligados à área da Biomedicina.

Não são contempladas na ontologia a descrição do método da pesquisa e recursos acadêmicos. Dessa forma, existem duas ontologias que poderiam

⁸NONTHARKN, C., WUWONGSE, V. Linked OpenScholar: a Researcher Network Using Linked Open Data. In: ICADL, 14., Taipei. **Proceedings...** Taipei: [s.n.], 2012, p. 325-328. Disponível em: <<http://arnetminer.org/publication/linked-openscholar-a-researcher-network-using-linked-open-data-3555554.html;jsessionid=8FE9645DEDFBF774336CD75DF9CF6883.tt>>. Acesso em: 08 nov. 2014.

⁹WOLSKI, M. et al. Developing the Discovery Layer in the University Research e-Infrastructure. In: KGCM, 5., Orlando. **Proceedings...** Orlando: [S.n.], 2011, documento não paginado. Disponível em <http://www.iiis.org/CDs2011/CD2011SCI/KGCM_2011/Abstract.asp?myurl=GA919ET.pdf, 2011.>. Acesso em: 07 nov. 2014.

incorporadas (via reutilização) ao sistema VIVO para preencher essa lacuna. São elas: *Linked Science Core Vocabulary Specification* e *Academic Institution Internal Structure Ontology* (AIISO). A primeira descreve recursos ligados à pesquisa (método, hipótese, resultado) e a segunda, recursos acadêmicos (curso, departamento, divisão, instituto, módulo, programa, dentre outros).

Pode-se concluir que a ontologia VIVO-ISF está adequada para a descrição de recursos científicos, pois abrange conceitos que podem ser usados para qualquer área da ciência.

5.2 A FUNDAMENTAÇÃO EM VIVO-ISF

Esta subseção investiga o uso da ontologia de fundamentação (BFO)¹⁰ na remodelagem conceitual e estrutural da ontologia VIVO-ISF. A fundamentação favorece o desenvolvimento de uma ontologia para a descrição de recursos da ciência, pois propicia um entendimento filosófico acerca das entidades que fazem parte da ciência refletindo, por exemplo, sobre aspectos da natureza dos recursos. Essa reflexão filosófica intrínseca à ontologia de fundamentação nos faz voltar aos pensamentos dos filósofos da antiguidade que discutiam a Ontologia: a essência das coisas.

Aparentemente, parece que o uso da fundamentação pouco interfere no processo de construção. Entretanto, não se pode esquecer que uma ontologia deve ser legível, formal e explícita e para isso é necessário que haja um entendimento comum sobre um determinado conceito. Uma ontologia pode ser comparada a um quebra-cabeça, ou seja, todas as peças devem ser encaixadas umas nas outras sem que nenhuma peça fique solta.

A fundamentação só começou ser a ser utilizada a partir da “versão” 1.6, (como será mostrado na seção 5.3).

VIVO-ISF incorporou apenas algumas partes de BFO, selecionando as categorias mais adequadas para a descrição dos recursos que são contemplados pela ontologia.

O Quadro 7 mostra o uso de BFO em VIVO-ISF.

¹⁰ Abordada no Referencial Teórico deste trabalho.

Quadro 7 – Fundamentação em VIVO-ISF

Ontologia de fundamentação	Uso
Continuante Independente	Descrição de agentes, como: <i>Person</i> (Pessoas), <i>Group</i> (Grupo), <i>Organization</i> (Organização) e <i>Manufacturer</i> (Fabricante). Os três primeiros são descritos pela ontologia FOAF, já o recurso fabricante é descrito pela ontologia IAO.
Continuante Genericamente Dependente	Abrange a classe <i>Information Content Entity</i> (Entidade que contém Informação). Essa classe descreve <i>Collection</i> (Coleção), <i>Contact</i> (Contato), <i>Database</i> (Base de Dados), <i>Document</i> (Documento) e <i>Service</i> (Serviço). São reutilizadas as ontologias BIBO, Vivo, IAO, ARG, Vcard, OBO, C4o, OBI, ERO e UO.
Continuante Especificamente Dependente	Descreve as classes <i>Relationship</i> (Relacionamento), <i>Quality</i> (Qualidade) e <i>Realizable Entity</i> (Entidade Realizável). A primeira é descrita pela ontologia Vivo e a as demais pela ontologia BFO. Também são reutilizadas as ontologias ARG e ERO.
Processo	Abrange as classes <i>Advising Process</i> (Processo de Orientação), <i>Attending Process</i> (Processo de Atendimento), <i>Educational Process</i> (Processo Educacional), <i>Event</i> (Evento), <i>Event Series</i> (Séries de Evento), <i>Organizing Process</i> (Processo de Organização), <i>Phase</i> (Fase), <i>Planned Process</i> (Processo Planejado), <i>Presenting Process</i> (Processo de Apresentação) e <i>Project</i> (Projeto).

Fonte: autora (2014).

Os recursos de pesquisa clínica ganharam mais destaque após a implementação de BFO. A partir da definição de *Entity* (Entidade), a classe *Material Entity* (Entidade Material) passou a ser designada para descrever: organismo, espécie biológica, instrumento, reagente, equipamento e *software*, ou seja, recursos pertencentes à pesquisa clínica.

O uso da fundamentação trouxe importantes benefícios à ontologia como um todo. Classes que estavam soltas e “penduradas” foram realocadas de acordo com a

natureza dos recursos. Além disso, VIVO-ISF foi remodelada e reestruturada semanticamente, o que a qualifica como ontologia para descrição de recursos do domínio da ciência.

5.3 EVOLUÇÃO

Esta subseção investiga a evolução da ontologia VIVO a partir das mudanças ocorridas em suas versões, analisando também como tais modificações podem impactar sistemas que utilizam essa ontologia.

Desde a primeira versão (2010), a ontologia de VIVO é atualizada constantemente, o que comprova que o sistema evolui conforme as constantes exigências que ocorrem na descrição de recursos do domínio da ciência. No total, a ontologia possui seis mudanças nas “versões”, sendo: 1.1, 1.2, 1.3, 1.4, 1.5 e 1.6 (os diagramas estão nos Anexos A, B, C e D deste trabalho). Serão apresentadas, brevemente, as modificações que ocorreram nas “versões” 1.1, 1.3 e 1.4, pois foram as mais significativas para a evolução da ontologia.

O Quadro 8 sintetiza as modificações mais relevantes em 1.1, 1.3 e 1.4.

Quadro 8 – Principais modificações até a versão 1.5

Versão	Modificações	Ontologias
1.1	Classes pertencentes à ontologia: <i>Agent</i> , <i>Autorship</i> , <i>Position</i> , <i>Role</i> , <i>Event</i> , <i>Process</i> , <i>Informational Resources</i> , <i>Relationship</i> e <i>Agreement</i> .	VIVO, BIBO, <i>Event</i> e FOAF
1.3	Classe <i>Relationship</i> removida. Com essa alteração, a então subclasse de <i>Relationship</i> , <i>Advising Relationship</i> tornou-se uma classe “solta” na ontologia. Classes relacionadas à área geográfica e a pesquisa clínica são adicionadas.	VIVO, BIBO, <i>Event</i> , FOAF, <i>Geopolitical</i> e <i>Eagle-i</i>
1.4	Classe <i>Process</i> removida e devido a essa alteração, a então subclasse, <i>Project</i> transformou-se em classe.	VIVO, BIBO, <i>Event</i> , FOAF, <i>Geopolitical</i> , <i>Eagle-i</i> , SKOS

Fonte: autora (2014).

Até a “versão” 1.4 as modificações foram pequenas, entretanto, ocorreram exclusões de classes, o que pode ser prejudicial para quem está utilizando a ontologia em seu próprio sistema. Cabe ressaltar o aumento do número e da

especificidade das ontologias reutilizadas conforme a evolução de VIVO-ISF. Em 1.1, quatro ontologias de cunho mais geral; em 1.3, seis ontologias (sendo quatro de cunho geral e duas específicas); e em 1.4 sete ontologias (sendo cinco de cunho mais geral e duas específicas).

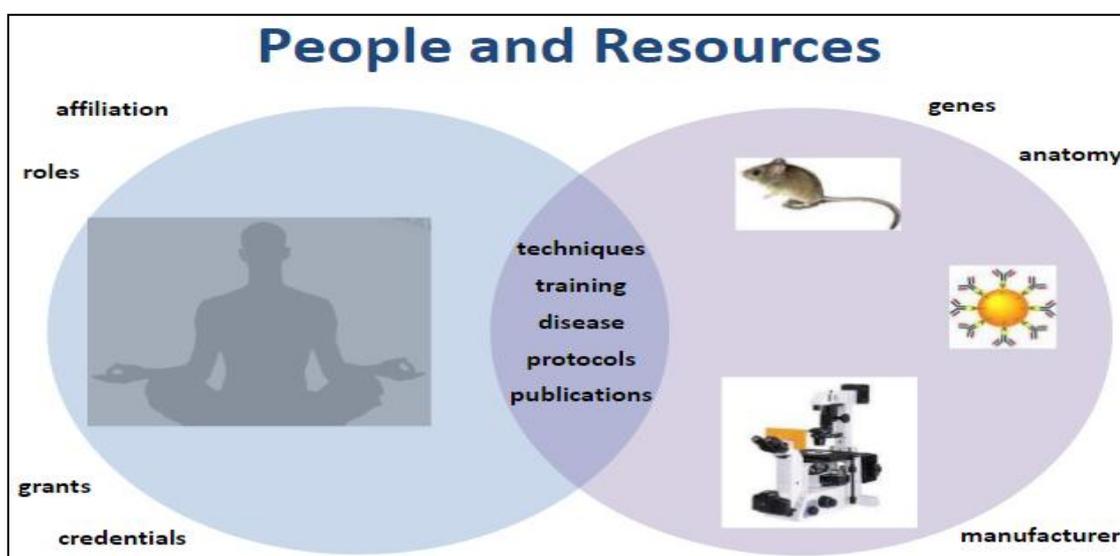
Entretanto, a grande transformação aconteceu na “versão” 1.6 (objeto de análise dessa pesquisa) com a aproximação da ontologia do projeto *Eagle-i*, o qual, é voltado para a área da Biomedicina. Essa fusão culminou no desenvolvimento da ontologia VIVO – ISF, que como o nome já diz, possui uma estrutura semântica integrada. A ontologia obteve um ganho conceitual importante devido ao uso da ontologia de Fundamentação BFO, a qual já era usada por *Eagle-i*, pois faz uso de conceitos gerais, como continuantes e ocorrentes (mencionado na subseção 5.2).

Essa nova versão possui, no total, 410 superclasses, classes e subclasses, além de 189 propriedades de objeto e 211 propriedades de dados. O número de ontologias também aumentou expressivamente passando de 8 para 17, sendo: FOAF; Vivo; BFO; ERO; OBI; IAO; *Agents, Resources and Grants (ARG)*; *Event*; BIBO; *Ontology of Clinical Research (OCRE)*, Vcard, C4o; Fabio, SKOS, OBO, *Unit Ontology (UO)*.

Cabe ressaltar que a ontologia de VIVO já estava alinhada com a de *Eagle-i* desde a “versão” 1.4, porém, a completa fusão ocorreu apenas nesta última “versão”, quando as classes presentes em *Eagle-i* foram incorporadas às de VIVO.

A Figura 22 mostra a junção dos projetos VIVO e *Eagle-i*.

Figura 22 – VIVO & *Eagle-i*



Torniai et al. (2013).

Como pode ser observado na Figura 22, antes desse processo cada projeto tinha uma finalidade própria, pois seus objetivos eram diferentes: VIVO descrevia perfis de pesquisadores pertencentes às instituições participantes, já *Eagle-i*, objetivava a descrição de recursos para pesquisa clínica. Após a fusão, passam a ser descritos recursos relacionados a pessoas e à pesquisa clínica.

A ontologia VIVO-ISF está em constante evolução e após a “versão” 1.6 ganhou uma consistência conceitual importante, à medida que incorpora conceitos gerais provenientes da ontologia de alto nível BFO.

A estrutura conceitual proporcionada por BFO favorece a evolução, pois propicia um entendimento filosófico consistente acerca das entidades que fazem parte da ciência, refletindo sobre a natureza dos recursos. Dessa forma, futuras evoluções não provocarão mudanças estruturais radicais no corpo da ontologia, pois seu núcleo está bem fundamentado.

5.4 REÚSO EM VIVO-ISF

O reúso promove a **dinamicidade** e a **interoperabilidade** do sistema. A dinamicidade está atrelada às transformações que acontecem na ciência e, a interoperabilidade permite que classes, subclasses e propriedades sejam reaproveitadas de ontologias amplamente utilizadas. A partir do reúso, não há a necessidade de se “inventar a roda”, haja visto que o *modus operandi* da ontologia é o compartilhamento de dados, recursos e informações relacionados à pesquisa. Essa é a nova *web*. Essa é a nova ciência.

A ontologia VIVO – ISF reusa ontologias que descrevem os mais variados recursos do domínio da ciência. Para que a estrutura seja semanticamente integrada é utilizado o princípio da refatoração¹¹: uso de ontologias variadas para diversas classes. Esse princípio expande as possibilidades na descrição, pois um conceito pode ser representado sob diversas formas, como por exemplo, na classe “Document”, em que são utilizadas as ontologias BIBO, Vivo, Fabio e OBI.

O Quadro 9 traz as ontologias utilizadas na ontologia VIVO – ISF e o domínio ao qual pertencem.

¹¹TORNIAI, C. et al. Finding common ground: integrating the eagle-i and VIVO ontologies. In: ICBO, 4., Montreal. **Proceedings...** Montreal: [s.n.], 2013. p. 1-4. Disponível em: <http://www2.unb.ca/csas/data/ws/icbo2013/papers/research/icbo2013_submission_20.pdf>. Acesso em: 07 nov. 2014.

Quadro 9 - Ontologias em VIVO - ISF

Ontologia	Domínio
FOAF	Descreve pessoas, agentes e organizações.
Vivo	Descreve membros, departamentos, premiações e relacionamentos acadêmicos.
<i>Geopolitical</i>	Descreve áreas geopolíticas.
<i>Basic Formal Ontology</i> (BFO)	Descreve as entidades: Ocorrente e Continuante a partir de uma reflexão filosófica.
<i>Eagle-i Resource Ontology</i> (ERO)	Descreve papéis, materiais e diversos recursos de pesquisa.
<i>Ontology for Biomedical Investigations</i> (OBI)	Descreve recursos relacionados à área biomédica.
<i>Information Artifact Ontology</i> (IAO)	Descreve recursos informacionais.
<i>Agents, Resources and Grants</i> (ARG)	Descreve agentes e processos.
<i>Event</i>	Descreve eventos.
<i>Bibliographic Ontology</i> (BIBO)	Descreve recursos bibliográficos.
<i>Ontology of Clinical Research</i> (OCRE)	Descreve recursos de pesquisa clínica.
<i>VCard</i>	Descreve pessoas. Espécie de cartão de visitas virtual.
C4o	Descreve citações.
<i>FRBR – Aligned Bibliographic Ontology</i> (Fabio)	Descreve recursos informacionais.
<i>Simple Knowledge Organization System</i> (SKOS)	Espécie de tesouro que ajuda na codificação dos conceitos presentes na ontologia.
OBO	Descreve recursos da área biomédica.
<i>Unit Ontology</i> (UO)	Descreve unidades de medida.

Fonte: autora (2014).

VIVO também utiliza o princípio *The Minimum Information to Reference an External Ontology Term*¹² (MIREOT), que tem por finalidade selecionar classes e propriedades de ontologias externas para o reuso. Esse princípio foi herdado da ontologia OBI e, que por fazer parte do projeto *Eagle-i*, também foi incorporado ao sistema.

VIVO reusa ontologias de domínio geral, como FOAF e *VCard* para descrever pessoas e organizações, *Event* para eventos, BIBO para recursos bibliográficos e

¹²TORNIAI, C. et al. Finding common ground: integrating the eagle-i and VIVO ontologies. In: ICBO, 4., Montreal. **Proceedings...** Montreal: [s.n.], 2013. p. 1-4. Disponível em: <http://www2.unb.ca/csas/data/ws/icbo2013/papers/research/icbo2013_submission_20.pdf>. Acesso em: 07 nov. 2014.

SKOS para organizar o conhecimento. Isso permite uma maior interoperabilidade semântica com outros ambientes que também utilizam essas ontologias.

VIVO-ISF é uma ontologia que não reinventa a roda, pois reaproveita ontologias amplamente usadas em ambientes *linked open data*. Essa reutilização é bem estruturada, pois utiliza-se a ontologia de fundamentação para a organização das classes. Além disso, a importação das ontologias é seletiva (método MIREOT), o que demonstra que existem critérios na seleção de classes e de propriedades. O uso de VIVO-ISF em outros ambientes proporcionará a **dinamicidade** e a **interoperabilidade** (encontradas em VIVO) do sistema, através do compartilhamento de conceitos amplamente utilizados no domínio da ciência.

6 CONCLUSÕES

A *web* atual vem provocando mudanças que reverberam nos mais variados setores da sociedade, estreitando assim, a linha tênue que separa os mundos real e virtual. A relação entre o homem e máquina está se tornando cada vez mais presente e cooperativa.

Hoje, dados e informações podem ser publicados e compartilhados em ambientes semânticos que promovem a descoberta de novos recursos do domínio da ciência permitindo, dessa forma, que pesquisadores tenham acesso ao que está sendo investigado na própria fonte. Além disso, esses ambientes promovem a interação entre os cientistas pertencentes a sua comunidade científica. É o hipertexto com “roupagem” nova.

Os ambientes semânticos utilizam *linked open data* (LOD) na publicação de dados, visto que LOD encontra na Web Semântica a plataforma ideal para a interligação dos dados, pois faz uso de ferramentas e padrões como os agentes inteligentes, URIs, RDF, OWL e as ontologias.

Diversos projetos na área da ciência utilizam LOD na publicação de dados científicos. Eles podem ser específicos a uma determinada área ou então de abrangência mais geral, como o sistema VIVO. O sistema modela conceitualmente perfis de pesquisadores e descreve recursos de pesquisa clínica a partir do uso de sua própria ontologia, a VIVO-ISF, e que foi o objeto de estudo dessa pesquisa.

A ontologia VIVO-ISF foi analisada a partir de quatro vieses: escopo, fundamentação, evolução e reuso. Essa análise teve como premissa a investigação da estrutura e da modelagem conceitual da ontologia, pois são aspectos que devem ser relevantes antes da implementação/reuso da ontologia em um determinado sistema, a fim de evitar futuras mudanças indesejadas na estrutura da ontologia para quem a estiver utilizando.

A partir das investigações feitas conclui-se que a ontologia VIVO-ISF pode ser utilizada na descrição de recursos do domínio da ciência, e é genérica o suficiente em sua extensão para ser utilizada em outros ambientes. Apesar de também contemplar recursos específicos à área da pesquisa clínica, grande parte da ontologia está direcionada para a descrição de recursos mais gerais, como: agentes, papéis e entidades informacionais.

VIVO-ISF também está em um estágio de evolução que traz segurança para quem for utilizá-la, em função de fazer uso da ontologia de fundamentação, haja visto que sua estrutura está modelada conforme a natureza dos recursos. Portanto, futuras modificações, provavelmente, não impactarão radicalmente a ontologia. VIVO-ISF também reutiliza uma grande quantidade de ontologias, o que favorece a interoperabilidade semântica.

Quando se fala em ontologias é fundamental lembrar o significado filosófico que a palavra, de fato, possui. Se Aristóteles denominou como Ontologia, Metafísica ou Filosofia Primeira **a ciência do Ser, o estudo das essências**, isso pouco interfere no processo de análise, mas o que realmente importa (e é necessário) é que nós (continuanes independentes) assumamos nosso papel (continuanes dependentes) de bibliotecários-filósofos ou filósofos-bibliotecários (?) enquanto avaliadores de ontologias.

Esse foi apenas o início de um longo caminho que pode e deve ser trilhado pelos bibliotecários. Dessa forma, sugere-se que sejam feitos novos estudos sobre esse tema, como: investigar a possibilidade de implementação da ontologia VIVO-ISF no Fabrico/Ciência e investigar a possibilidade de implementação do sistema VIVO na UFRGS a fim de descrever recursos de ensino, pesquisa e extensão.

REFERÊNCIAS

ALMEIDA, M., BAX, M.. Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. *Ciência da Informação*, Brasília, DF, Brasil, 32, fev. 2004.

ARP, R.; SMITH, B. Realizable Entities in Basic Formal Ontology. [S.l.:s.n.], 2011. Disponível em: <<http://ontology.buffalo.edu/smith/articles/realizables.pdf>>. Acesso em: 08 nov. 2014.

ARP, R; SMITH, B. Function, Role and Disposition in Basic Formal Ontology. [S.l.], Nature Proceedings, 2008. Disponível em: <<http://ontology.buffalo.edu/smith/articles/realizables.pdf>>.

BABBAGE. *Site*. 2014. Disponível em: <<https://babbage.inf.unibz.it>>. Acesso em: 18 set. 2014.

BAGLATZi, A.; KAUPPINEN, T.; KEBLER, C. Linked Science Core Vocabulary Specification. *Site*. 2011. Disponível em: <<http://linkedscience.org/lsc/ns/>>. Acesso em: 13 dez. 2014.

BAUER, F; KALTENBÖCK, M. **Linked Open Data: The Essentials: a quick star guide for decision makers**. Vienna : Mono/monochrom, 2011. Disponível em: <http://www.semantic-web.at/LOD-TheEssentials.pdf>. Acesso em: 17 out. 2014.

BERNERS-LEE, T. Linked Data. *Site*. 2006. Disponível em: <<http://www.w3.org/DesignIssues/LinkedData.html>>. Acesso em: 15 out. 2014.

BERNERS-LEE, T. ; LASSILA, O. ; HENDLER, J. The semantic web. **Scientific American**, May 2001. Disponível em: <http://www.scientificamerican.com/article/the-semantic-web/>. Acesso em: 25 maio 2014.

BERNERS-LEE, T. **Uniform Resource Identifier (URI): generic syntax**. *Site*. Disponível em: <<http://tools.ietf.org/html/rfc2396>>. Acesso em: 23 jun. 2014.

BRICKLEY, D.; MILLER, L. **FOAF Vocabulary Specification** 099. *Site*. 2014. Disponível em: <<http://xmlns.com/foaf/spec/>>. Acesso em: 20 maio 2014.

CHAUÍ, M. **Convite à filosofia**. 14. ed. São Paulo : Ática, 2010. Disponível em: <<http://bahiapsicosocial.com.ar/biblioteca/Convite%20%20Filosofia%20-%20Marilena%20Chaui.pdf>>. Acesso em: 06 nov. 2014.

CONLON, M; HOLMES, K. Big Science teams built on research discovery and networking systems. **Team Science**, San Diego, v.2, n.2, documento não paginado, 2012. Disponível em: <<http://academicexecutives.elsevier.com/articles/big-science-teams-built-research-discovery-and-networking-systems>>. Acesso em: 23 out. 2014.

CONLON, M. ; et al. Research discovery through linked open data. In: ACM/IEEE-CS JOINT CONFERENCE ON DIGITAL LIBRARIES, 12., Washington. **Proceedings...** Washington: ACM, 2012. p. 429-430. Disponível em: <<http://cns.iu.edu/docs/>>

publications/2012-albert-jcdl.pdf>. Acesso em: 24 set. 2014.

CRISPIM, H.; ABDALLA, H.; MOLINARO, L. Proposta para um ambiente inteligente de ensino a distância. In: ABED, 9., São Paulo. Anais... São Paulo: [s.n.], 2002, documento não paginado. Disponível em: <<http://www.abed.org.br/congresso2002/trabalhos/texto48.htm>>. Acesso em: 25 out. 2014.

DING, L.; ZHOU, L; FININ, T. et al. **How the Semantic Web is Being Used: an Analysis of FOAF Documents**. [S.l.]: [s.n.], [200-?]. Disponível em: <http://ebiquity.umbc.edu/_file_directory_/papers/120.pdf>. Acesso em: 20 maio 2014.

FACTBITES. *Site*. 2014. Disponível em: <<http://www.factbites.com/topics/Mango>>. Acesso em: 3 out. 2014.

FEITOSA, Ailton. **Organização da informação na web: das tags à web semântica**. Brasília, DF : Thesaurus, 2006. 131 p. : il.

FOAF. Introducing FOAF. *Site*. [200-?]. Disponível em: <<http://www.foaf-project.org/original-intro#Examples>>. Acesso em: 20 maio 2014.

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS. FAO Geopolitical ontology. *Site*. [201-?]. Disponível em: <<http://www.fao.org/countryprofiles/geoinfo/Geopolitical>>. Acesso em: 25 maio 2014

GIL, Antonio Carlos. **Como elaborar projetos de pesquisa**. 4.ed. São Paulo: Atlas, 2009.

GÓMEZ-PERÉZ, A; BENJAMINS, V.R. **Overview of Knowledge Sharing and Reuse Components: ontologies and Problem- Solving Methods**. In: IJCAI-99, 2., 1999, Stockholm. **Proceedings...** Stockholm, [s.n.], 1999. p. 1-15.

GONÇALVES, V.; CARRAPATOSO, E.; Web semântica e cérebro global juntos por uma boa causa. In: EduSer: Revista de Educação, Bragança, v.1, n.1, p. 70-87, 2009. Disponível em: <<https://bibliotecadigital.ipb.pt/handle/10198/1219>>. Acesso em: 28 out. 2014.

GOOGLE. *Site*. 2014. Disponível em: https://www.google.com.br/search?q=manga&ie=utf-8&oe=utf-8&aq=t&rls=org.mozilla:pt-BR:official&client=firefox-a&channel=sb&gfe_rd=cr&ei=JqJvVKjkOaWk8weM64HoCQ. Acesso em: 3 out. 2014.

GRUBER, T.R. A translation approach to portable ontology specifications: **Knowledge Acquisition**, v.5, n.2, p. 1, 1993. Disponível em:<<http://tomgruber.org/writing/ontolingua-kaj-1993.pdf>>. Acesso em: 18 jun. 2014.

GUARINO, N. Formal Ontology and Information Systems. In: FOIS'98, 1998, Trento. **Proceedings...**Trento: IOS Press, 1998, p. 3-15. Disponível em: <<http://uosis>.

mif.vu.lt/~donatas/Vadovavimas/Temos/OntologiskaiTeisingasKonceptcinisModeliavimas/papildoma/Guarino98-Formal%20Ontology%20and%20Information%20Systems.pdf >. Acesso em: 08 jun. 2014.

IANELLO, R. **VCard Ontology**: for describing people and organizations. *Site*. 2014. Disponível em: <<http://www.w3.org/TR/2014/NOTE-vcard-rdf-20140522/#objectproperties>>. Acesso em: 21 jun. 2014.

IGLESIAS- SUCASAS, M.; KIM, S.; VIOLLIER, V. **The FAO Geopolitical Ontology: a reference for country-based information**. Food and Agriculture Organization of the United Nations: [200-?]. Disponível em: http://www.semantic-web-journal.net/sites/default/files/swj179_0.pdf. Acesso em: 25 maio 2014.

LIMA-MARQUES, Mamede. **Ontologias**: da filosofia à representação do conhecimento. Brasília, DF: Thesaurus, 2006. (Série Ciência da Informação e da Comunicação, 1).

MARCONDES, Carlos Henrique. “Linked data” – dados interligados - e interoperabilidade entre arquivos, bibliotecas e museus na web. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, [S.l.], v. 17, n. 34, p. 171-192, ago. 2012. ISSN 1518-2924. Disponível em: <<https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2012v17n34p171>>. Acesso em: 23 out. 2014.

M 24 PERSON AND ORGANIZATION ONTOLOGY. Person and Organization Ontology. *Site*. 2007. Disponível em:<<http://kplab.evtek.fi:8080/wiki/Wiki.jsp?page=M24PersonAndOrganizationOntology>>. Acesso em: 21 jun. 2014.

NOY, N.; M, D. Ontology Development 101: a guide to creating your first ontology. [S.l.]: [s.n.], [200-?]. Disponível em: <http://protege.stanford.edu/publications/ontology_development/ontology101.pdf>. Acesso em: 12 jun. 2014.

PALMER, S. **The semantic web**: an introduction. *Site*. 2011. Disponível em: <http://infomesh.net/2001/swintro/>. Acesso em: 15 jun. 2014.

PERONI, S.; SHOTTON, D. FaBiO: the FRBR aligned Bibliographic Ontology. *Site*. 2014. Disponível em:< <http://www.essepuntato.it/lode/http://purl.org/spar/fabio>>. Acesso em: 22 jun. 2014.

PHILLY 2013. OBI Workshop. *Site*. 2013. Disponível em: < <http://obi-ontology.org/page/Philly2013>>. Acesso em: 22 set. 2014.

RAIMOND, Y.; ABDALLAH, S. **The Event Ontology**. *Site*. 2007. Disponível em: <http://motools.sourceforge.net/event/event.html>>. Acesso em: 05 jun. 2014.

RAMALHO, R. A. S. **Desenvolvimento e utilização de ontologias em Bibliotecas Digitais**: uma proposta de aplicação. 2010. 146 f. Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, São Paulo, 2010. Disponível em: <http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/ramalho_ras_do_mar.pdf>. Acesso em: 18 jun. 2014.

ROCHA, R. P. Fabrico/Ciência: um ambiente Linked Data para o mapeamento da ciência. In: **Em Questão**, Porto Alegre, v. 18, p. 281-297, p. 282-283, dez. 2012. Edição Especial. Disponível em: <<http://seer.ufrgs.br/EmQuestao/article/view/33279>>. Acesso em: 25 set. 2014.

ROCHA, R. P. Web Semântica, dados ligados e web 2.0: explorando novas fronteiras para arquivos abertos. In: ENANCIB, 14., Florianópolis. **Anais...** Florianópolis: [S.n.], 2013. p. 1-20. Disponível em:<<http://enancib.sites.ufsc.br/index.php/enancib2013/XIVenancib/paper/view/376>>. Acesso em: 21 jun. 2014.

SCHMACHTENBERG, M; BIZER, C; PAULHEIM, H. Linked Open Data Cloud Diagram 2014. *Site*. 2014. Disponível em: <http://data.dws.informatik.uni-mannheim.de/lodcloud/2014/>. Acesso em: 27 out. 2014.

SCIMANTICA. SEMANTIC SCIENCE. Reading the Tea Leaves of 2011- Data and Technology Predictions for the Year Ahead. *Site*. 2011. Disponível em: <<http://semanticscience.wordpress.com/category/data/>>. Acesso em: 27 out. 2014.

SEMANTIC COMMUNITY. Build VIVO in the Cloud. *Site*. 2013. Disponível em: <http://semanticcommunity.info/Build_VIVO_in_the_Cloud>. Acesso em 04 nov. 2014.

SOUZA, R. R; ALVARENGA, L.. A Web Semântica e suas contribuições para a ciência da informação. **Ci. Inf.**, Brasília, v.33, n.1, p.132-141, jan/abril 2004.

STYLES, R.; SHABIR, N. Academic Institution Internal Structure Ontology. *Site*. 2008. Disponível em: <<http://vocab.org/aiiso/schema>>. Acesso em: 13 dez. 2014.

THE BIBLIOGRAPHIC ONTOLOGY. BIBO. *Site*. 2014. Disponível em:< <http://uri.gbv.de/ontology/bibo/>> . Acesso em: 19 jun. 2014.

THE EVENT ONTOLOGY. 2007. *Site*. Disponível em: <<http://motools.sourceforge.net/event/event.html>>. Acesso em: 18 set. 2014.

THE ONTOLOGY BIOMEDICAL INVESTIGATIONS. Ontology. *Site*. 2014. Disponível em: <http://obi-ontology.org/page/Main_Page>. Acesso em 19 jun. 2014.

TORNIAI, C; et al. Finding common ground: interating the eagle-I and VIVO ontologies . [S.l.:s.n], 2013. Disponível em: < <http://pt.slideshare.net/mhaendel/isf-refactoring-icbo2013>>. Acesso em: 08 nov. 2014.

VIVO. CORNELL UNIVERSITY. *Site*. 2014. Disponível em: <<http://vivo.cornell.edu/display/individual22786>>. Acesso em: 15 set. 2014.

VIVO. CORNELL UNIVERSITY. *Site*. 2014. Disponível em: <<http://vivo.cornell.edu/vis/map-of-science/individual22786>>. Acesso em: 15 set. 2014.

VIVO.CORNELL UNIVERSITY. *Site*. 2014. Disponível em: < <http://vivo.cornell.edu/vis/author-network/individual22786>>. Acesso em: 15 set. 2014.

VIVO.CORNELL UNIVERSITY. *Site*. 2014. Disponível em: <http://vivo.cornell.edu/vis/investigator-network/individual22786>>. Acesso em: 15 set. 2014.

VIVO. DURASPACE. VIVO-ISF Ontology. *Site*. 2013. Disponível em: <<https://wiki.duraspace.org/pages/viewpage.action?pagelId=33949186>>. Acesso em 18 set. 2014.
WEITZNER, D. et al. Promoting interoperability between Heterogeneous Policy Domains. *Site*. 2006. Disponível em: <[http://dig.csail.mit.edu/2006/Talks/1017-w3cws-rein/#\(1\)](http://dig.csail.mit.edu/2006/Talks/1017-w3cws-rein/#(1))>. Acesso em: 21 set. 2014.

WIKIPEDIA. Geopolitical Ontology. *Site*. 2013. Disponível em: < http://en.wikipedia.org/wiki/Geopolitical_ontology>. Acesso em: 21 set. 2014.

W3C. RDF 1.1 Concepts and Syntax. *Site*. 2014. Disponível em: <http://www.w3.org/TR/2014/REC-rdf11-concepts-20140225/>. Acesso em: 23 jun. 2014.

APÊNDICE – Formulário de coleta de dados

ESCOPO
1. Por que a ontologia foi construída?
2. Quem são os usuários da ontologia?
3. A ontologia é de cunho geral ou específica?
FUNDAMENTAÇÃO
4. Quais são os recursos de fundamentação utilizados?
5. A fundamentação favorece o desenvolvimento de uma ontologia para descrever recursos da ciência?
6. A fundamentação favorece o reúso no desenvolvimento de uma ontologia para descrever recursos da ciência?
EVOLUÇÃO
7. A ontologia evolui?
8. Quais são as principais mudanças.
REÚSO
9. Há uma preocupação em relação ao reúso de ontologias?
10. No reúso, há a preocupação em incorporar ontologias de ampla utilização, tanto gerais quanto específicas da ciência?
11. Quais são as ontologias reutilizadas? Quais domínios dessas ontologias?

ANEXO A – Diagrama da versão 1.1 da ontologia de VIVO

