

Capítulo

3

Ontologias biomédicas: teoria e prática

Fernanda Farinelli^{1,2} e Mauricio B. Almeida^{1,3}

¹ Programa de Pós Graduação em Gestão e Organização do Conhecimento, UFMG

²fernanda.farinelli@gmail.com, ³mba@eci.ufmg.br

Abstract

This chapter is dedicated to the study of ontology, an interdisciplinary research field that examines entities and relationships as a mean for better performing modeling in knowledge domains. There is a growing interest in the application of ontologies for problem solving in several areas, such as Computer Science, Information Science, Philosophy, Artificial Intelligence, Linguistics, Knowledge Management; as well as within knowledge domains, such as Health, Biomedicine, Law, Geography, to name a few. Our approach here is called "Applied Ontology", a field that gathered together new technologies, in particular those developed for the Semantic Web, and solid philosophical foundations. The first part of the chapter presents basic information about ontology and its application in information systems; in the second part, a practical example of the application of ontologies in the field of medicine is developed.

Resumo

O presente capítulo se dedica ao estudo da ontologia, um campo de pesquisa interdisciplinar que estuda entidades e relações fundamentais para a modelagem de domínios do conhecimento. Existe um crescente interesse na aplicação de ontologias para a solução de problemas em diversas áreas como Ciência da Computação, Ciência da Informação, Filosofia, Inteligência Artificial, Linguística, Gestão do Conhecimento; bem como em domínios como Saúde, Biomedicina, Direito, Geografia, para citar alguns. A abordagem aqui é a da "Ontologia Aplicada", um campo que alia novas tecnologias, em particular aquelas desenvolvidas para a Web Semântica, a uma sólida fundação filosófica de cunho metafísico. Na primeira parte do capítulo apresentam-se informações básicas sobre ontologia e sua aplicação em sistemas de informação; na segunda parte, apresenta-se exemplo prático da aplicação de ontologias no domínio da medicina.

3.1. Teoria em ontologias

3.1.1. Fundamentos: ontologias e modelos baseados em ontologias

Ontologia é um tema que tem sido estudado em diferentes campos de pesquisa – Filosofia, Ciência da Computação e Ciência da Informação – e no âmbito de vários domínios do conhecimento – Medicina, Biologia, Direito e Geografia, para citar alguns. Apesar da ampla difusão do termo, não é uma tarefa simples entender o que significa “ontologia”. Na Filosofia, ontologia é um ramo da Metafísica que diz respeito a quais categorias de entidades existem e estão relacionadas [Lowe 2007]. Em Ciência da Computação, ontologias são aplicadas à modelagem de sistemas de informação (SIs), em diferentes ramos: sistemas tradicionais baseados em bancos de dados, e sistemas baseados em conhecimento, assim nomeados no campo da Representação do Conhecimento. Em Ciência da Informação, Vickery [1997, p. 285] conclui que “os problemas com os quais os cientistas da informação vêm lutando há muito tempo são enfrentados agora pela comunidade de engenheiros do conhecimento.” Tais problemas dizem respeito questões de modelagem.

Antes de adentrar em detalhes sobre os diversos significados para o termo “ontologia”, apresenta-se, para entendimento do contexto, um breve histórico desde os primeiros modelos de dados até os modelos baseados em ontologias.

Modelos são representações simplificadas da realidade que se busca entender. O mundo é complexo e os modelos são produzidos para permitir que a compreensão humana apreenda e organize fatos. Modelos também são entidades importantes e partes integrantes do método científico. Uma das formas de classificar os modelos é considerar a questão semântica, a qual trata das funções da representação. Deste ponto de vista, os modelos podem ser *modelos de fenômenos*, *modelos teóricos* ou *modelos de dados* [Frigg 2006]. Modelos de dados proliferam nas organizações modernas como um meio de representar o que se pretende codificar e processar em SIs.

Os SIs têm papel relevante na consolidação de novas práticas administrativas, pois visam atender às necessidades da corporação. O desenvolvimento de SIs envolve a criação de modelos para representar atividades e processos que ocorrem na corporação. Um modelo de dados corporativo é “[...] uma representação explícita da estrutura, atividades, processos, fluxos, recursos, pessoas, comportamento, objetivos e restrições de uma corporação” [Gandon 2002, p.42].

No desenvolvimento de SI, o estágio em que os modelos são criados para fins de compreensão humana é geralmente referenciado como *modelagem conceitual*. Os modelos conceituais são criados a partir de abstrações de aspectos da realidade, seja da perspectiva de um indivíduo ou de um grupo. As abstrações são um meio de especificar as entidades e as relações entre entidades dentro do domínio de um campo de conhecimento que é de interesse para o sistema em construção. A modelagem conceitual de SIs é resultado de pesquisas realizadas nos últimos 50 anos. As primeiras iniciativas para a especificação de modelos de dados datam do final da década de 1950 [Young, Kent 1958, Bosak *et al.* 1962]. Tais iniciativas objetivavam gerar modelos que fornecessem estruturas de dados para atendimento a necessidades computacionais.

Na década de 1960, a pesquisa em bancos de dados deu origem a três tipos principais de modelos de dados: o *modelo hierárquico*, o *modelo de rede* e o *modelo relacional*. Esses modelos são em geral referenciados como *modelos lógicos*, pois não alcançavam aspectos físicos (de

implementação). No entanto, os modelos lógicos apresentavam problemas que limitavam sua utilização na modelagem conceitual [Mylopoulos 1998]. Por exemplo, no modelo relacional [Codd 1970], um constructo denominado “relacionamento” é usado para representar tanto “entidades” quanto “relações entre entidades” do mundo real [Peckham, Maryanski 1988]. Esse fato gera problemas de comunicação e leva a erros de modelagem.

Os primeiros *modelos semânticos* utilizados na modelagem conceitual surgiram na década de 1970, no escopo do trabalho do Comitê ANSI/X3/SPARC para a padronização de sistemas de gerenciamento de banco de dados. Os mais notáveis são o *modelo de dados semântico* [Abrial 1974], a *arquitetura de três esquemas* [Jardine 1976], o *modelo entidade-relacionamento* (ER) [Chen 1976] e o *modelo ER estendido* [Codd 1979], dentre outros. A principal característica dos modelos semânticos, em comparação aos anteriores, é que eles são de mais fácil compreensão por pessoas.

O modelo ER, por exemplo, removeu a sobrecarga do constructo “relação” que existia nos modelos relacionais, além de fornecer termos adicionais para uso como primitivas de modelagem. A modelagem conceitual surgiu a partir de modelos de dados semânticos desenvolvidos para bancos de dados, mas o Comitê ISO/TC97/SC5 formou um grupo com o objetivo de determinar padrões para linguagens de modelagem conceitual de SIs.

Nos anos 90, propostas para *modelagem orientada a objetos* se tornaram populares. Muitos consideravam a orientação a objeto como uma categoria diferente dos modelos de dados. De fato, tais modelos tinham recursos adicionais em comparação com modelos de dados anteriores, mas ainda assim mantinham semelhanças nos constructos adotados, tais como: *objetos* versus *entidades*, *atributos* versus *propriedades*, *relacionamentos* versus *associações*, *classes* versus *hierarquias* [Milton 2000]. A *Unified Modeling Language* foi uma tentativa de padronizar as notações orientadas a objetos que incluíam diversas outras iniciativas: o Método Booch [Booch 1993], a Técnica de Modelagem de Objetos [Rumbaugh *et al.* 1991], a Engenharia de Software Orientada a Objetos [Jacobson *et al.* 1992], dentre outros.

Ao longo de todos esses anos, a criação de modelos conceituais vem sendo motivada pela busca de formas cada vez melhores de representar a realidade em SIs. A modelagem conceitual é “[...] a atividade de descrever formalmente alguns aspectos do mundo físico e social ao nosso redor para fins de compreensão e comunicação” [Mylopoulos 1992, p. 3]. No entanto, modelos semânticos usados na modelagem conceitual utilizam um conjunto limitado de constructos para a tarefa a que se propõe. A partir dos anos 1990, uma diversidade de iniciativas para o uso de ontologias nas corporações pode ser encontrada na literatura de SIs [Fox 1992, Uschold *et al.* 1998, Fillion *et al.* 1995, Schlenoff 1996, Bernus *et al.* 1996].

Smith e Welty (2001) apontam a inconsistência nas práticas durante os primeiros anos de modelagem conceitual como a principal causa dos problemas de interoperabilidade nos SIs atuais. Uma alternativa para esse tipo de problema são os *modelos baseados em ontologias*: “[...] a provisão, definitiva, de uma ontologia de referência comum e bem fundamentada – uma taxonomia compartilhada de entidades – pode fornecer vantagens significativas sobre os métodos *ad-hoc*, caso a caso, usados até aqui” [Smith, Welty 2001, p. 4].

O termo “ontologia” apareceu pela primeira vez na literatura da Ciência da Computação em 1967, na teoria de dados proposta por Mealy¹ [Smith 2003]. Historicamente, vários autores

¹ Mealy, G. H. (1967). Another Look at Data. Proceedings of AFIPS Conference. 31, 525–534. Washington: Thompson.

foram pioneiros no tema [Genesereth, Nilsson 1987, Gruber 1993, Guarino, 1995, Guarino, Giaretta 1995, Guarino 1998, Smith 2003, Vickery 1997, Wand *et al.* 1999]. Ontologias têm sido estudadas desde a década de 1970 na pesquisa de Inteligência Artificial e desde os anos 80 na modelagem conceitual de SI. Nos anos 90, a pesquisa em Web Semântica aumentou a demanda por ontologias para diversos tipos de aplicativos, tanto para solucionar problemas de interoperabilidade quanto para fornecer uma estrutura unificada de comunicação.

O estudo das ontologias é caracterizado pela coexistência de abordagens interdisciplinares, com pelo menos sete interpretações disponíveis na literatura para o termo ontologia [Guizzardi 2005]: i) uma disciplina filosófica; ii) um sistema conceitual informal; iii) um sistema baseado semântica formal; iv) uma especificação de uma conceitualização. v) uma representação de um sistema conceitual via teoria lógica; vi) um vocabulário usado por uma teoria lógica; vii) uma especificação (meta-nível) de uma teoria lógica. Uma ontologia descreve o significado dos símbolos adotados no SI e representa uma visão específica do mundo. As ontologias são nesse caso classificadas em duas dimensões principais: a dimensão de tempo corresponde à utilização de ontologias em SI, seja em tempo de desenvolvimento ou em tempo de execução; a dimensão estrutural lida com o uso da ontologia como um componente de banco de dados, como a interface do usuário ou como um aplicativo [Guarino 1998].

Em Representação do Conhecimento (RC), um subcampo da Inteligência Artificial (IA), o termo ontologia é usado para se referir a uma estrutura de entidades representados por um vocabulário lógico. Para melhor entendimento, considere-se as atividades e os agentes envolvidos na tarefa de representar conhecimento: *sistemas declarativos* contém declarações que representam *fatos* governados por *regras*. Um exemplo de um fato é “New York é uma cidade nos Estados Unidos” e de uma regra é “todas as pessoas que vivem em New York vivem nos Estados Unidos”. Essa combinação de fatos e regras compõem uma *base de conhecimento* do sistema. Uma base de conhecimento é construída e mantida por um *engenheiro do conhecimento*, que tem como tarefa formalizar o conhecimento de um grupo de especialistas. Em muitos casos, a ontologia faz o papel da base de conhecimento.

Para executar tais tarefas, um engenheiro realiza generalizações e abstrações, as quais requerem *insights* metafísicos. Nessa linha de pensamento, uma ontologia é uma teoria representativa dos principais fatos e regras que governam parte da realidade, para fins computacionais. Existem outras abordagens em RC que empregam linguagens de representação mais expressivas para construir modelos para uma teoria [Guarino 1998]. Aderindo a um compromisso ontológico – uma descrição da conceitualização pretendida para uma teoria lógica [Guarino, Giaretta 1995] – emprega-se uma linguagem de representação para gerar um conjunto de modelos representativos da realidade. O papel da ontologia é tornar explícitos axiomas que restringem modelos, de forma a igualar, tanto quanto possível, os modelos que contém o significado pretendido (ver Figura 3.1).

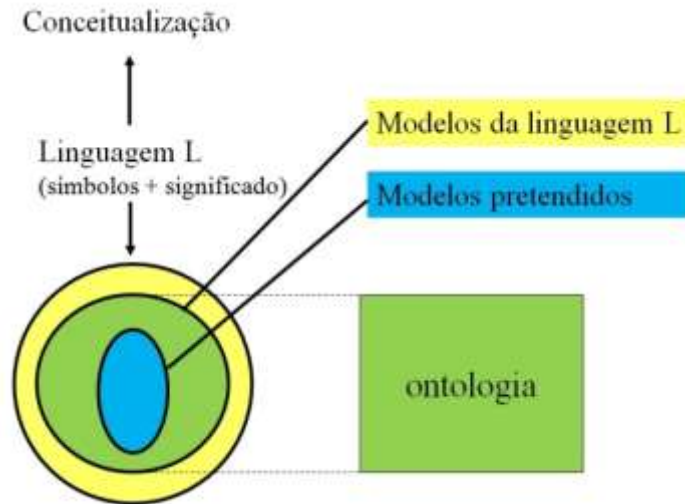


Figura 3.1. Ontologia e linguagem de representação
Fonte: Adaptado de Guarino (1998)

Uma explicação formal da Figura 1 deve incluir uma conceitualização C representada por uma linguagem de representação L aderente a um compromisso ontológico $K = \langle C, I \rangle$. Na verdade, L se compromete com o domínio D através de K , do qual C é a conceitualização subjacente. A variável I representa uma função interpretação, a qual mapeia elementos de D para símbolos do vocabulário V . Todos os modelos de L compatíveis com K são modelos pretendidos de L de acordo com K . Nesse contexto, o papel da ontologia é restringir os modelos de L de forma que eles se tornem modelos pretendidos $I_k L$.

Observam-se então dois significados principais para o termo ontologia em Ciência da Computação.

O primeiro significado diz respeito ao uso de princípios ontológicos para entender e modelar a realidade (ver por ex. Wand, Weber 1990, Wand *et al.* 1999), o qual denominamos princípios da “ontologia como disciplina”. O uso do termo nesse caso está alinhado com seu papel original na Filosofia, ou seja, fornecer uma descrição do que existe e caracterizar entidades nas atividades de modelagem.

Para um ilustrar de forma prática em sistemas de informação, a Figura 3.2 apresenta um modelo em *Unified Modeling Language* (UML) real com uma falha de modelagem: uma data é modelada como entidade, quando na verdade é uma propriedade. Esse é um tipo de falha simples que a ontologia, funcionando como um metamodelo, pode aprimorar. Na verdade, diversos outros tipos de falhas similares podem ser corrigidos pelo uso da ontologia como metamodelo. A Figura 3.3 apresenta o mesmo modelo UML da Figura 3.2 depois de corrigido pelo uso da ontologia *Bunge-Wand-Weber* (BWW) [Alturki *et al.* 2013]. O uso da ontologia como disciplina para criar ontologias como artefatos em geral gera modelos ontológicos bem fundamentados.

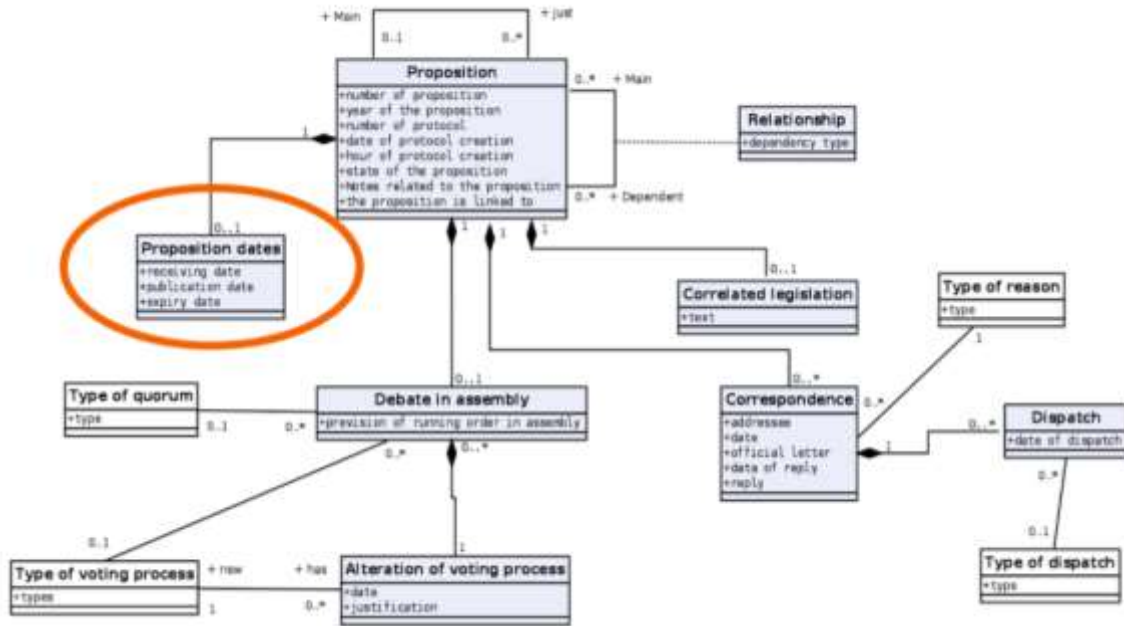


Figura 3.2. Modelo UML com falha na modelagem
 Fonte: dos autores

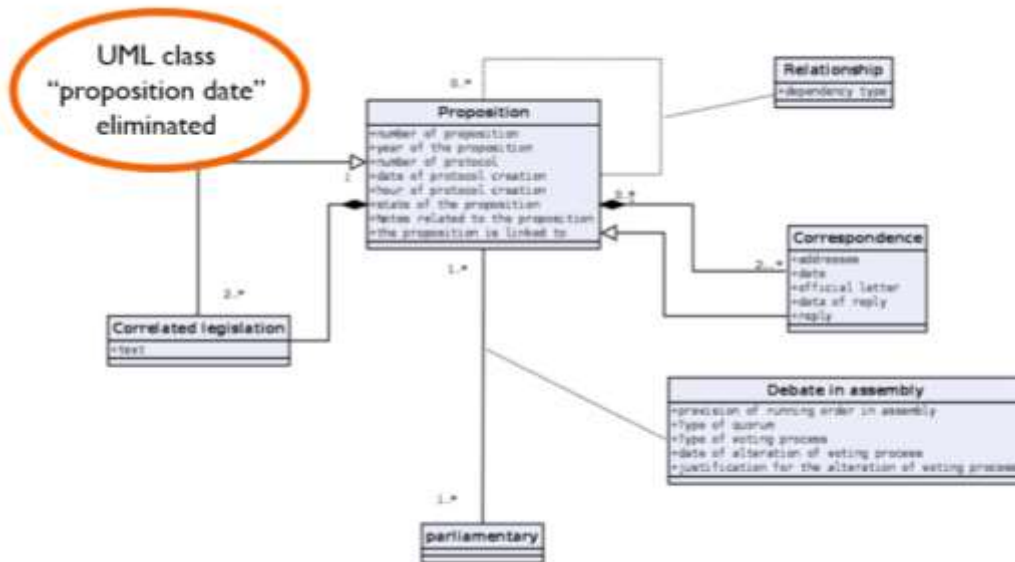


Figura 3.2. Modelo UML corrigido pelo uso da ontologia BWW
 Fonte: dos autores

O segundo significado para o termo ontologia diz respeito a representação de um domínio em uma linguagem de representação computacional (ver por ex. Staab, Studer 2004), o qual denominamos aqui “ontologia como artefato”. Uma ontologia, nesse caso, consiste de um conjunto de declarações expressas em uma linguagem de representação, a qual pode ser processado por mecanismos de inferência automatizados. Mecanismos de inferências são artefatos de software

que podem extrair conhecimento novo a partir de conhecimento existente em uma ontologia, devidamente axiomatizada. A inferência equivale a verificar as consequências lógicas.

Do ponto de vista filosófico, pode haver apenas uma ontologia [Smith 1998]. A fim de lidar com a questão da pluralidade de uso de termo, o qual passou a ser usado também no mundo dos SIs, distinguem-se dois tipos de ontologias: i) a ontologia real, que diz respeito a como o universo é organizado, e corresponde a uma abordagem filosófica; e ii) ontologia epistemológica, relacionada à tarefa de conceituar um domínio. Essa nomenclatura, entretanto, é confusa ao misturar os termos “ontologia” e “epistemologia”, que em Filosofia estão em diferentes campos de estudo. Outras formas de distinção são encontradas na literatura, por exemplo: ontologia com “O” maiúsculo versus ontologia com “o” minúsculo [Guarino, Giaretta 1995]; ontologia “de” SI e ontologia “para” SI [Fonseca 2007]. No primeiro caso, a ontologia é usada para modelagem conceitual; no segundo, a ontologia é um componente do SI que descreve o vocabulário de um domínio com o objetivo de apoiar a criação de esquemas conceituais.

O quadro 3.1 reúne as interpretações genéricas para o termo ontologia em todos os campos em que o termo é amplamente citado. No restante do presente capítulo, adota-se assim a seguinte nomenclatura:

- *ontologia como disciplina*, para nomear a ontologia como disciplina filosófica e os princípios metafísicos subjacentes;
- *ontologia como artefato*, para se referir a qualquer artefato computacional construído com alguma linguagem declarativa de representação;
- *modelo ontológico*, para se referir a qualquer ontologia como artefato que faz uso dos princípios da ontologia como disciplina para obter boa fundamentação na modelagem.

Quadro 3.1. Quadro sinótico resumindo as visões sobre ontologia

Distinção	Campo	O que é?	Propósito	Exemplo
Ontologia como uma disciplina	Filosofia	Ontologia como um sistema de categorias	Entender a realidade, as coisas que existem e suas características	Sistemas de Aristóteles, Kant, Husserl
Ontologia como um artefato	Ciência da Computação	ontologia como uma teoria (baseada em lógica)	Entender um domínio e reduzi-lo à modelos	BFO, DOLCE (genéricas)
		ontologia como um artefato de software	Criar um vocabulário para representação em sistemas e para gerar inferências	OWL (linguagem de RC)
	Ciência da Informação	ontologia como uma teoria (informal)	Entender um domínio e classificar termos	Sistema de classificação de Ranganathan
		ontologia como um sistema conceitual informal	Criar vocabulários controlados para recuperação da informação a partir de documentos	um catálogo, um glossário, um tesouro

Fonte: adaptado de Almeida (2013)

A capacidade de integração entre modelos via a redução de ambiguidade inerente a construção de ontologias em linguagens declarativas formais possibilita aplicações relevantes em

campos críticos como a Medicina. Uma aplicação direta atualmente utilizada em todo o mundo é a melhoria da comunicação, para pessoas e máquinas, em partes dos sistemas médicos de grande importância para os cuidados a saúde, por exemplo, os prontuários de pacientes. Dessa forma, as ontologias podem auxiliar aos profissionais que lidam com sistemas de informação médicos ao prover uma fundamentação sólida para a construção de modelos alinhados com a realidade e, portanto, passíveis de melhor possibilidade de integração com outros sistemas, hoje facilmente distribuídos temporal e geograficamente.

3.1.2. Entidades em ontologias: classes, tipos, relações, propriedade, atributos, instâncias

Existe consenso de que o estudo da ontologia como disciplina diz respeito aos tipos de coisas que existem. Nesse contexto, “tipo” quer dizer “categoria”, um termo que foi usado ainda por Aristóteles para discutir que declarações se pode fazer sobre uma entidade. De fato, uma teoria das categorias é o mais importante tópico do estudo da ontologia como disciplina. Tais teorias especificam sistemas de categorias estruturados em níveis hierárquicos, em geral, na forma de uma árvore invertida na qual a categoria de mais alto nível é nomeada “entidade”. Qualquer coisa pode ser descrita como uma entidade de algum tipo, mas qual os próximos níveis de categorização são questão para discussão.

Em termos de Filosofia, três sistemas de categorias são os mais influentes e, mesmo que não diretamente, influenciam a forma como os modelos SIs são criados atualmente: o de Aristóteles², o de Kant³ e o de Husserl⁴ [Thomason 2009]. Apresenta-se a seguir uma breve introdução ao sistema de classificação de Aristóteles, uma vez que ele foi a base para todos os subsequentes e é adotado até os dias de hoje.

Aristóteles usou a linguagem como uma pista para descrever as categorias ontológicas. Kant utilizou conceitos como uma maneira de abordar categorias de objetos de uma possível cognição. O objetivo dos sistemas de categorias aristotélico e kantiano era descrever a estrutura categórica do mundo de acordo com o pensamento e a linguagem humanos. Partindo das categorias de Aristóteles, Husserl forneceu categorias descritivas das mais altas essências das coisas possíveis. Husserl não fornece um inventário das coisas que realmente existem, ainda que tenha sido o criador do termo “ontologia formal”.

Aristóteles parece ter sido o primeiro filósofo a usar a palavra grega “categoria” como um termo técnico para predicação. Embora os estudos de Aristóteles tenham diferentes interpretações, os seguidores de sua tradição acreditam que um sistema de categorias deveria fornecer um inventário das coisas que existem. O primeiro sistema de categorias proposto por Aristóteles dividiu entidades em dois ramos: *dito-sobre* e *presente-em*, tornando possível as seguintes situações:

- qualquer entidade é dita de outra, ou não é dita de outra
- qualquer entidade está presente em outra; ou não está presente em outra

Entidades que são *ditas-sobre* outras são chamadas *universais*; enquanto aquelas que não são *ditas-sobre* outras são chamadas de *particulares*. Entidades que estão *presentes-em* outras são

² Aristóteles - filósofo grego, 384–322AC

³ Immanuel Kant - filósofo alemão, 1724–1804

⁴ Edmund Gustav Albrecht Husserl - filósofo alemão, 1859–1938

chamadas de *acidentes*, enquanto aquelas que não estão *presentes-em* outras são não acidentais. Entidades não acidentais são universais e descritos como entidades essenciais; entidades não-acidentais são particulares e descritas simplesmente como não essenciais. Reunindo as possibilidades em conjunto, obtém-se um sistema quádruplo de categorias e os nomes para cada uma [Studtmann 2008]:

- Entidades que são ditas-sobre e presentes-em são *universais acidentais*
- Entidades que são ditas-sobre e não presentes-em são *universais essenciais*
- Entidades que não são ditas-sobre e estão presentes-em são *particulares acidentais*
- Entidades nem são ditas-sobre nem estão presentes-em outra coisa são as *substâncias primárias*

Organizadas dessa forma, as entidades do mundo podem ser representadas pelo *Quadrilátero de Aristóteles*, conforme Fig. 3.4.

	Substância (não presente-em um sujeito)	Acidente (presente-em um sujeito)
Universal (dito-sobre um sujeito ou predicado de um sujeito)		
Particular (não dito-sobre um sujeito ou predicado de um sujeito)		

Figura 3.4a. Quadrilátero ontológico de Aristóteles
Fonte: adaptado de Smith (2003)

	Substância	Acidente
Universal	Segunda substância <i>homem</i> <i>cão</i> <i>touro</i>	Segundo acidente <i>dor de cabeça</i> <i>bronzeado</i> <i>pavor</i>
Particular	Primeira substância <i>este homem</i> <i>este cão</i> <i>este touro</i>	Primeiro acidente <i>esta dor de cabeça</i> <i>este bronzeado</i> <i>este pavor</i>

Figura 3.4b. Quadrilátero ontológico de Aristóteles com exemplos
Fonte: adaptado de Smith (2003)

Um segundo sistema de categorias foi proposto por Aristóteles, o qual contém uma lista dos tipos de mais alto nível. Percebendo que os objetos comuns da experiência humana são alocados em classes de crescente generalidade, pode-se observar indícios de que a existência de um tipo mais elevado e abstrato é provável. No entanto, Aristóteles não acreditava em uma categoria de alto nível única, mas em dez delas, como mostrado na Figura 3.5.

Termo Aristotélico	Significado moderno	Exemplo
<i>Ti esti, ousia</i>	Substância	homem
<i>Poson</i>	Quantidade	cinco metros
<i>Poion</i>	Qualidade	branco
<i>Pros ti</i>	Relação	metade
<i>Pou</i>	Local	no mercado
<i>Pote</i>	Data	ontem
<i>Keisthein</i>	Postura	sentado
<i>Echein</i>	Estado	vestido
<i>Poitein</i>	Ação	queimar
<i>Paschein</i>	Sentimento	ser queimado

Figura 3.5. Categorias genéricas de Aristóteles
Fonte: adaptado de Sutcliffe (1993)

A categoria mais importante, *substância*, pode ser entendida listando-se suas características mais significativas [Smith 1997]: substâncias podem existir por conta própria; permanecem numericamente únicas e as mesmas, tendo propriedades diferentes em momentos diferentes; podem participar de relações causais; não têm partes próprias que são substâncias; têm uma contínua; e não têm partes temporais.

Os diferentes tipos de perguntas que teriam sido empregadas para obter a lista de categorias de Aristóteles são perguntas que podem ser feitas sobre algo. Por exemplo, a pergunta “o que é isto?” Só pode ser feita sobre uma substância, e somente respostas descrevendo substâncias são apropriadas [Ackrill 1963]. Independentemente da pergunta, as categorias de Aristóteles são entendidas como categorias de coisas, não de linguagem. Por exemplo, a definição de um tigre não nos diz o significado da palavra “tigre”, mas nos diz o que é ser um tigre [Cohen 2008]. Esse fato tem reflexos diretos em atividades de modelagem atuais, sugerindo que para uma boa modelagem, deve-se observar aquelas categorias que tem referente no mundo real.

Outra questão é como e se uma categoria é mais fundamental que outra. Neste contexto, “fundamental” diz respeito a condições de existência e identidade das entidades de uma categoria [Lowe 2007]. Cada entidade tem um recurso fundamental chamado *essência real*. Quando uma entidade possui uma essência específica, significa que ela é de certo tipo; e, para ser de certo tipo, uma entidade deve compartilhar um conjunto de propriedades necessárias e suficientes com outros membros desse mesmo tipo. Em contraste com tais propriedades essenciais, há também propriedades acidentais que uma entidade pode perder sem perder sua essência real [Ackrill 1963]. Por exemplo, uma pessoa é um animal bípede, mas uma pessoa pode perder ambas as pernas e continuar a ser uma pessoa, mantendo a verdadeira essência de uma pessoa. Esse fato também tem reflexos diretos em atividades de modelagem atuais, uma vez que é a forma de se obter uma boa definição para uma entidade.

A tradição aristotélica fornece um método para ordenar categorias de coisas com base em sua essência. O método para distinguir essências usa a distinção gênero-espécie e a divisão

dicotômica [Lennox 2000]. De acordo com a distinção gênero-espécie, a verdadeira essência das espécies é uma combinação de seu gênero e sua diferenciação, que é o critério usado para distinguir uma espécie de outra do mesmo gênero. Por exemplo, a espécie humana pertence ao gênero *Homo* e sua diferenciação é racionalidade; então a verdadeira essência de um ser humano é ser um animal racional. Além disso, através da divisão dicotômica, Aristóteles também propôs a divisão de cada gênero nas entidades que possuem uma diferenciação particular e aquelas que não possuem. Por exemplo, o gênero dos seres vivos é dividido em animais e plantas pela diferenciação do movimento próprio.

Em resumo, a terminologia de origem filosófica que vai influenciar termos usados em modelagem de SIs pode ser organizada da seguinte forma (ver também Figura 3.6).

- *Universais, tipos, tipos naturais ou categorias*: derivados da segunda substância (Figura 3.4a) são abstrações de conjuntos de entidades com características similares que existem no mundo independentemente da mente humana, e que representam todas as entidades daquele tipo que existem, existiram ou vão existir; por ex., pessoas, árvores, bactérias, etc.
- *Classes*: são abstrações de conjuntos de entidades com características similares, mas não que não são naturais, ou seja, são demarcações criadas pela mente humana para um determinado fim; por exemplo, as pessoas que estão presentes a aula (para fins de fazer uma chamada), os carros no pátio da universidade (para fins de credenciamento), etc.
- *Particulares, instâncias ou indivíduos*: derivados da primeira substância (Figura 3.4a) são as entidades ou coisas do mundo em si; por exemplo, o universal “pessoa” pode ser instanciado por José, ou por Maria; o universal “planeta” instanciado por “Terra”, etc.
- *Qualidades, propriedades ou atributos*: derivados do primeiro e do segundo acidentes (Figura 3.4a) são características, ou relações unárias, de uma determinada entidade no nível dos universais (segundo acidente) ou nível das instâncias (primeiro acidente); por exemplo, peso e peso do José; vermelho e vermelho deste tomate, etc.
- *Relacionamentos*: conexões mantidas entre entidades, no mínimo binárias, no nível dos universais ou nível das instâncias, além de inter-níveis com universal e instância; por exemplo, vírus causa doença, João namora-com Maria; Pedro é-um professor; etc.

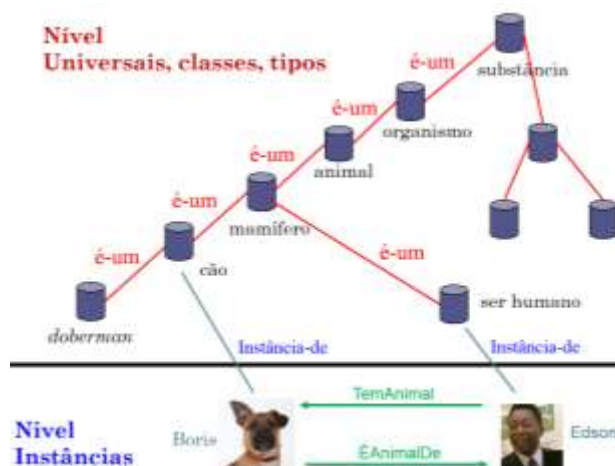


Figura 3.6. Entidades e relações
 Fonte: adaptado de Smith (2003)

Nas ontologias como artefatos encontram-se todas essas denominações, e as vezes autores diversos usam uma ou outra notação, ou uma mistura delas. Outro termo também muito usado em ontologias como artefato é “axioma”, que consiste de uma sentença declarativa em uma linguagem lógica, que define uma entidade ao impor-lhe restrições de significado. A Figura 3.7 apresenta um axioma da ontologia D-acts [Brochaussen, Almeida, Slaughter 2013] que define a entidade “*identity document*”, em um dialeto de lógica descritiva chamado *Manchester Syntax*. Trata-se de uma expressão lógica com quantificadores existenciais, nessa sintaxe representada pelo termo “*some*”.

```
document
and ('is about' some 'documented identity')
and ('is concretized as' some
('specifically dependent continuant'
and ('inheres in' some ('bearer of' some 'credential role'))))
and ('is concretized as' only
('specifically dependent continuant'
and ('inheres in' some ('bearer of' some 'credential role'))))
```

Figura 3.7. Expressão lógica para definir “documento de identidade”
Fonte: ontologia dos Atos dos Documentos (D-acts)

Finalmente, mais alguns conceitos relevantes de origem filosófica são necessários para o correto entendimento e uso de modelos ontológicos em SIs. As teorias filosóficas subjacentes à representação são, via de regra, classificadas de acordo com duas posições principais: *realismo* e *anti-realismo*. O debate envolvendo essas correntes já dura séculos e não há consenso [MacLeod, Rubenstein 2005].

O termo realismo tem várias interpretações em Filosofia [Niiniluoto 1999], mas em geral é usado para designar a noção de que existe um mundo físico independente da mente. Entidades independentes da mente são os já mencionados “universais”, os quais são instanciadas por “particulares. De acordo com os realistas, por exemplo, a frase “a lua é esférica” é verdadeira independentemente das crenças e das práticas linguísticas de qualquer pessoa.

As primeiras propostas realistas se originaram de Platão e Aristóteles. Para Platão, o conhecimento verdadeiro seria permanente e imutável e, portanto, não poderia ser oriundo de objetos ordinários sujeitos a mudanças constantes. Assim, para Platão, as coisas permanentes, os chamados universais, vêm de um reino de formas que existe à parte do mundo cotidiano. Nesta linha de pensamento, uma “cama” ou uma “mesa” são ideias (ou formas) que as pessoas têm antes mesmo de nascerem. Aristóteles não conectou o reino de formas de Platão com detalhes da vida cotidiana. Propôs que os universais não deveriam ser separados da realidade, mas deveriam ser elementos comuns presentes em particulares da mesma categoria. Por exemplo, o universal “mesa” consiste de todas as características comuns a todas as mesas.

O anti-realismo pode assumir muitas formas, dependendo de qual a dimensão do realismo é rejeitada: ou a existência ou a independência de entidades. Em geral, os anti-realistas são classificados como *nominalistas* e *conceitualistas* [MacLeod, Rubenstein 2005].

Os nominalistas acreditam que existem apenas indivíduos e os problemas de identidade e semelhança podem ser resolvidos a partir de indivíduos e nas relações entre eles. Acreditam que

as semelhanças empíricas entre entidades não são bons critérios para estabelecer a filiação de uma entidade a uma categoria ou para caracterizar um universal. A existência separada de universais, uma visão dualista exibida tanto nas visões de Platão quanto de Aristóteles, é o principal ponto de discordância com o realismo. Nominalistas defendem que os universais são desnecessários e que todo conhecimento provém de entidades particulares oriundo da experiência das pessoas.

Conceitualistas negam que os indivíduos são suficientes para resolver os problemas, mas não apelam para o uso de universais. Em vez disso, conceitualistas explicam os problemas de identidade e de semelhança referindo-se a conceitos ou ideias. Nessa visão, cada palavra da linguagem tem um conceito geral associado a ela. A visão conceitualista parece trazer de volta uma espécie de dualismo, semelhante à visão realista, que estabelece uma diferença entre as coisas e as abstrações sobre coisas, agora sobre outro nome, à saber, conceito.

Apesar de esse tipo de tema parecer um tanto desligado da realidade dos SIs, ele é a base da boa modelagem uma vez que identifica como as entidades do mundo são caracterizadas. O uso prático dos conceitos até aqui explicados será objeto de mais detalhe em seções seguintes.

3.1.3. Representação do conhecimento: categorias básicas para modelagem

O sistema de categorias apresentado na seção anterior (Figura 4b) foi expandido para um sexteto ontológico, de forma a abrigar as considerações sobre entidades *continuentes* e *ocorrentes*, entidades básicas e essenciais para a modelagem ontológica. Essa é a divisão que a ontologia como disciplina (de origem aristotélica) aplica de mundo, ou seja, todas as coisas ou são *continuentes*, entidades que mantêm identidade ao longo do tempo, como por exemplo os objetos (pessoas, árvores, etc.); ou são *ocorrentes*, entidades que se alteram ao longo do tempo, como por exemplo os processos (de digestão, de matrícula, etc.). Além disso, o sexteto também diferencia aquelas entidades que dependem de outras para existir, à saber, *continuentes independentes* e *continuentes dependentes* (Figura 3.8).

	Continuante independente	Continuante dependente	Ocorrente (processo)
Universal	Segunda substância <i>homem</i> <i>gato</i> <i>touro</i>	Segunda qualidade <i>dor de cabeça</i> <i>bronzeado</i> <i>pavor</i>	Segundo processo <i>caminhar</i> <i>pensar</i> <i>dormir</i>
Particular	Primeira substância <i>este homem</i> <i>este gato</i> <i>este touro</i>	Primeira qualidade <i>esta dor de cabeça</i> <i>este bronzeado</i> <i>este pavor</i>	Primeiro processo <i>este caminhar</i> <i>este pensar</i> <i>este dormir</i>

Figura 3.8. Sexteto ontológico estendendo o quadrilátero de Aristóteles
Fonte: adaptado de Smith (2003)

Ontologias de alto nível são ontologias como artefatos desenvolvidos para atender objetivos específicos de modelagem, raciocínio automático e recuperação de informação [Hoehndorf, Schofield, Georgios 2015]. Exemplos de ontologias de alto nível são a *Descriptive*

Ontology for Linguistic and Cognitive Engineering (DOLCE) [Gangemi et al. 2002], a *Unified Foundational Ontology* (UFO) [Guizzardi 2005] e a *Basic Formal Ontology* (BFO) [Grenon, Smith, Goldberg 2002].

Adota-se a BFO como ontologia de alto nível, um tipo de metamodelo, uma vez que tem sido amplamente aceita em domínios como Medicina, Biologia, Bioinformática e áreas afins, Direito, Geografia, para mencionar alguns. Como uma ontologia de alto nível, o BFO segue os princípios do realismo já mencionado na seção anterior e representa apenas as categorias mais genéricas, fornecendo meios para categorizar entidades no âmbito da representação [Spear 2006]. A Figura 3.9 apresenta os dois ramos da BFO.

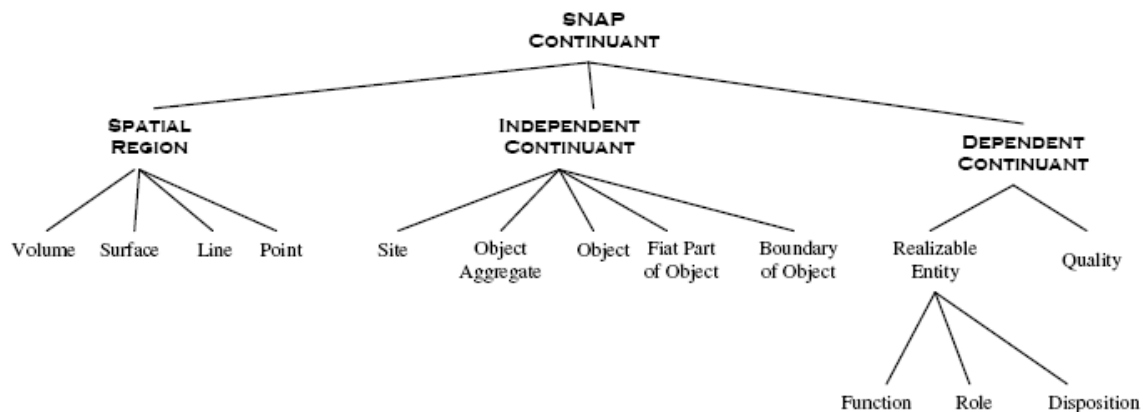


Figura 3.9a. Ontologia de alto nível BFO (versão 1), ramo dos continuantes
 Fonte: Spear (2006)

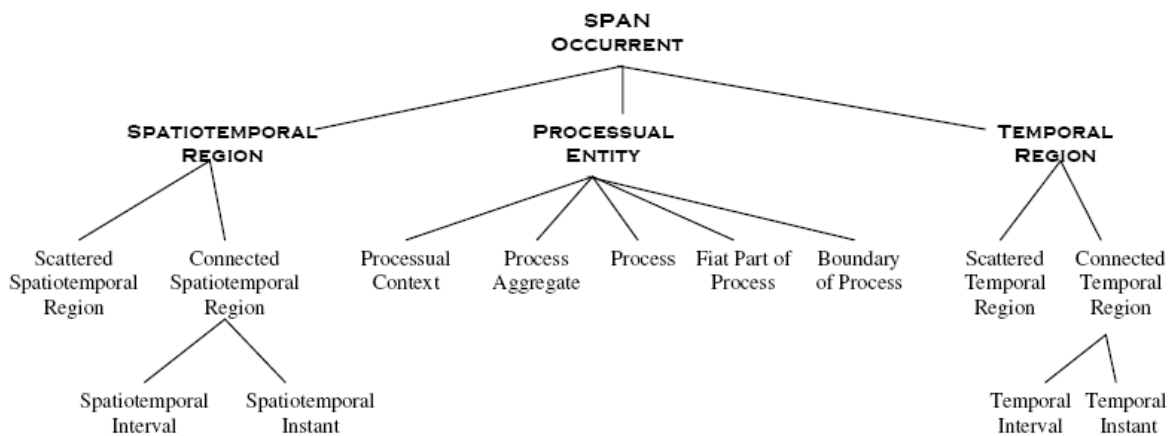


Figura 3.9b. Ontologia de alto nível BFO (versão 1), ramo dos ocorrentes
 Fonte: Spear (2006)

O BFO consiste em uma taxonomia que fornece uma categorização das entidades existentes no mundo e, portanto, uma explicação sobre a realidade (Fig. 3.9). O nó raiz da taxonomia do BFO é a entidade mais genérica, identificada apenas como “entidade”, e dele derivam dois ramos, à saber, *continuantes* (Fig. 3.9a) e *ocorrentes* (Fig. 3.9b). As Fig. 3.10 (de

3.10a até 3.10f, continuantes) e Fig. 3.11 (de 3.11a até 3.11f, ocorrentes) define cada nó da BFO e dá exemplos de uso [Spear 2006, Smith *et al.* 2007].

CONTINUANTES Entidades que mantêm identidade ao longo do tempo	
Características	A entidade existe completamente em qualquer período de tempo no qual está presente. A entidade não tem partes temporais.
Exemplos	Uma pessoa, a cor de uma fruta, um conjunto musical, uma lei, o vento na UFMG, ...

Figura 3.10a. Ramo dos continuantes
Fonte: adaptado de Spear (2006)

Distinção entre continuantes	Região espacial	Continuante independente	Continuante dependente
Características	Corresponde a um continuante independente. Não é portador de qualidades. Não é parte natural de outras entidades.	São portadores de qualidades ou propriedades. Entidades das quais outras entidades são partes naturais. Entidades que por si próprias não podem ser parte natural em nada. São existencialmente independentes de outras entidades.	São partes naturais ou originadas em outras entidades. São parte de outra entidade dependente ou independente. Exibem algum tipo de dependência existencial: para existir, outras entidades também devem existir.
Exemplos	A soma total de todos os espaços do universo, ...	Um organismo, um coração, uma orquestra, uma perna, uma pessoa...	A cor de uma laranja, o cheiro de pão, a função de ser um professor...

Figura 3.10b. Distinções no ramo dos continuantes
Fonte: adaptado de Spear (2006)

Distinção entre região espacial	Volume	Superfície	Linha	Ponto
Características	Uma região espacial com três dimensões	Uma região espacial com duas dimensões	Uma região espacial com uma dimensão	Uma região espacial sem dimensões
Exemplos	Uma parte do espaço em forma de cubo, em forma de esfera, ...	A superfície de uma parte do espaço em forma de cubo, em forma de esfera, ...	Uma aresta de uma parte do espaço em forma de cubo, ...	-

Figura 3.10c. Distinções no ramo das regiões espaciais
Fonte: adaptado de Spear (2006)

Distinção entre continuantes dependente	Entidade realizável	Qualidade
Características	São continuantes dependentes, partes naturais de continuantes que não são exibidas a cada momento em que fazem parte de uma entidade ou grupo de entidades. A exibição de uma entidade realizável corresponde a uma manifestação particular ou processo que ocorre sob certas circunstâncias. São entidades cuja vida contém períodos de atualização, com transformações em seus portadores; e períodos de latência, quando existem em seus portadores, mas não se manifestam.	É um continuante dependente que é exibido caso seja parte natural de uma entidade ou entidades. Podem ser partes naturais de outras entidades; para que a qualidade exista, essas entidades também devem existir.
Exemplos	O papel de ser um professor, a função dos órgãos respiratórios, a disposição do metal para conduzir eletricidade, a disposição do sangue para coagular, a fragilidade de um vaso, ...	A cor de uma laranja, a temperatura ambiente, a forma de uma orelha, a massa de um pedaço de ouro, o peso de um ser, ...

Figura 3.10d. Distinções no ramo das continuantes dependentes
Fonte: adaptado de Spear (2006)

Distinção continuantes independentes	Local	Agregado	Objeto	Parte fiat do objeto	Limite do objeto
Características	É um continuante independente que consiste de uma forma espacial característica em relação a um arranjo de outros continuantes. Distinto de região espacial	É um continuante independente que corresponde a uma soma de objetos separados. Possuem um grau de unidade mais fraco do que o dos objetos. Possuem limites não conectados.	É um continuante independente com extensão auto-contida. A identidade independe de outras entidades e pode ser mantida ao longo do tempo a despeito de perdas e ganhos.	É um continuante independente que é parte de um objeto, mas não é demarcado por falta de continuidade física.	É um continuante independente que é a parte dimensional inferior de outros continuante independentes.
Exemplos	Uma cidade, uma veia, um ambiente, a localização de uma guerra, o quarto onde alguém está, ...	Uma pilha de pedras, um grupo de pacientes de um hospital, uma coleção de livros, uma matilha de cães, ...	Um organismo, uma cadeira, uma célula, uma maçã, uma dobradiça, uma pedra, ...	O lado oeste de São Paulo, as superfícies dorsal e ventral do corpo, uma amostra de tecido, ...	A superfície da pele, a superfície da terra, a superfície externa de uma célula, ...

Figura 3.10e. Distinções no ramo das regiões espaciais
Fonte: adaptado de Spear (2006)

Distinção entre entidades realizáveis	Função	Papel	Disposição
Características	É uma entidade realizável cuja manifestação é uma atividade de um continuante direcionada para um fim específico. A finalidade é definida em virtude do continuante ser um tipo específico de entidade no contexto.	É uma entidade realizável não essencial para um continuante, ou seja, cuja manifestação gera resultados que não são essenciais para a identidade. Podem participar no tipo de continuante em contextos sociais e institucionais	É uma entidade realizável que causa um processo específico ou um tipo de transformação no objeto do qual é parte natural. A transformação ocorre apenas sob certas circunstâncias e junto a certas leis naturais, de forma que são entidades frágeis
Exemplos	A função de um coração em bombear sangue, a função de reprodução, a função de um martelo em uma obra, a função de juiz, ...	O papel de uma pessoa como advogado, o papel de uma droga no tratamento de uma doença, o papel de um árvore no ecossistema, ...	Alimentos que estragam se não armazenados sob refrigeração, a disposição do sangue em coagular, a disposição de um metal em conduzir eletricidade, ...

Figura 3.10f. Distinções no ramo das entidades realizáveis
Fonte: adaptado de Spear (2006)

	OCORRENTES Entidades que se alteram ao longo do tempo
Características	A entidade se desdobra ao longo de um período de tempo.
Exemplos	A respiração, o funcionamento de um órgão do corpo, parte da vida de Einstein, ...

Figura 3.11a. Ramo dos ocorrentes
Fonte: adaptado de Spear (2006)

Distinção entre ocorrentes	Região espaço-temporal	Entidade processual	Região temporal
Características	É uma entidade ocorrente na qual entidades processuais podem estar localizadas.	Uma entidade ocorrente que existe no tempo, tem partes temporais e depende de um continuante. A característica é a existência de partes temporais e espaciais.	É uma entidade ocorrente que é parte do tempo.
Exemplos	A região espaço-temporal ocupada por uma vida, pelo desenvolvimento de um tumor, ...	A vida de um organismo, o processo de meiose, o curso de uma doença, o voo de um pássaro, ...	O tempo para correr uma maratona, a duração de uma cirurgia, o momento da morte.

Figura 3.11b. Distinções no ramo dos ocorrentes
Fonte: adaptado de Spear (2006)

Distinção entre regiões temporais	Região temporal distribuída	Região temporal conectada
Características	É uma região do espaço que tem dimensões espaço-temporais. Cada ponto espacial e temporal não é conectado com outro ponto espacial e temporal.	É uma região que tem dimensões temporais e espaciais tal que todos os pontos dentro da região espaço temporal são imediatamente conectadas a todos ou outros pontos dentro da mesma região espaço temporal
Exemplos	O espaço e o tempo ocupados por jogos individuais da copa do mundo, ...	A localização espacial e temporal da vida de um organismo, a localização espacial e temporal do desenvolvimento de um feto, ...

Figura 3.11c. Ramo dos ocorrentes
Fonte: adaptado de Spear (2006)

Distinção reg. temporais conectadas	Intervalo temporal	Instante temporal
Características	Uma região temporal conectada que dura por mais do que um único momento de tempo.	uma região temporal conectada contendo um simples momento de tempo.
Exemplos	Qualquer duração temporal contínua na qual um processo ocorre.	O momento de nascimento de uma criança, o momento da morte.

Figura 3.11d – Distinção entre entidades processuais
Fonte: adaptado de Spear (2006)

Distinção entidades processuais	Contexto processual	Processo agregado	Processo	Parte fiat de processo	Limite do processo
Características	É um ocorrente que consiste de uma forma característica espacial inserida em algum arranjo de outras entidades ocorrentes. Entidades nas quais ocorrentes podem estar localizados...	Um processo agregado é a soma das partes de processos. Não possui limites não conectados.	É uma entidade processual que é um todo maximamente conectado espacial e temporalmente. Possui início e fim <i>bona-fide</i> que corresponde a descontinuidades	É uma entidade processual que é parte de um processo. Não tem início e fim <i>bona-fide</i> e corresponde a descontinuidades reais.	É uma entidade processual que é o limite temporal instantâneo <i>bona-fide</i> ou fiat de um processo.
Exemplos	Uma operação cirúrgica como contexto processual para uma infecção, um check-up de rotina para encontrar uma doença, ...	O bater do coração de cada um dos sete indivíduos em um quarto, ...	A vida de um organismo, o processo de dormir, o processo de divisão celular,	Mascar chiclete durante uma refeição, o meio de uma tempestade, a pior parte de um ataque cardíaco, ...	A separação de um dedo em um acidente industrial, uma incisão no início da cirurgia, ...

Figura 3.11e. Distinção entre entidades processuais
Fonte: adaptado de Spear (2006)

Distinção reg. espaço temporal conectada	Intervalo espaço temporal	Instante espaço temporal
Características	Uma região do tempo e espaço conectados que se mantem por mais de um simples momento de tempo.	Uma região do tempo e espaço em um momento específico.
Exemplos	A região de espaço e tempo ocupada por um processo, ou por a parte fiat de um processo.	A região de espaço tempo ocupada por uma única parte temporal de um processo.

Figura 3.11f. Distinção entre regiões espaço temporais conectadas
Fonte: adaptado de Spear (2006)

As categorias filosóficas básicas são essenciais para a boa modelagem uma vez que fornecem uma fundamentação sólida para a definição, entendimento e comunicação sobre as entidades que povoam os modelos da realidade. Possibilitam construir uma base formal sobre a qual sistemas podem ser desenvolvidos e mesmo auditados.

3.1.4. Tecnologia e Web Semântica: linguagens e ferramentas

Desde os anos 1960, as aplicações computacionais têm feito uso de bancos de dados com sucesso. A questão sobre o uso de ontologias, ao invés de banco de dados, sempre surge em contextos práticos. Na verdade, ontologias como artefatos construídos sem os princípios da ontologia como disciplina, pouco tem a oferecer mais do que os bancos de dados, os quais alguns autores nomeiam como uma “promiscuidade ontológica”. A razão para isso é a natureza *ad-hoc* utilizada pelos desenvolvedores para a criação de sistemas. De fato, faz sentido construir modelos ontológicos bem fundamentados, os quais são passíveis de representação lógica, o que possibilita a redução da ambiguidade e de problemas de comunicação.

Entretanto, a diferença essencial entre as duas tecnologias reside no fato de que bancos de dados são tipos de sistemas de “mundo fechado”, em que se assume que o que não é sabido verdadeiro deve ser falso; e ontologias são sistemas conhecidos como de “mundo aberto”, em que se pressupõe que o que não é sabido verdadeiro é simplesmente desconhecido. Esse é um apelo para o uso de ontologias em Medicina, onde muito ainda pode ser descoberto e revelado, por exemplo, via inferências automáticas. Cabe assim verificar se o problema em questão exige o uso de ontologias. Não faria sentido, por exemplo, usar um sistema de mundo aberto para descobrir quantos presidentes o Brasil já teve. Outro apelo ao uso de ontologias é de que muitas consultas feitas com a tecnologia da Web Semântica não seriam possíveis em ambientes de bancos de dados.

O uso de motores de inferência se tornou popular nos últimos anos por avanços recentes da Inteligência Artificial (IA). Trata-se de IA uma vez que conhecimento novo é deduzido a partir de conhecimento existente. A figura 3.12a apresenta um fragmento de *Web Ontology Language*⁵ (OWL), o qual afirma que um cão é uma subclasse de vertebrado, que vertebrado é uma subclasse de animal equivalente a um animal que tem ossos. O mesmo fragmento é apresentado no editor de

⁵ <https://www.w3.org/OWL/>

ontologias Protégé⁶ (Fig. 3.12b). O motor de inferência da ferramenta deduz que não existem cães sem ossos (“nothing”, na figura Fig. 3.12c).

```
Dog subclassOf Vertebrate
Vertebrate subclassOf Animal
Vertebra subclassOf Bone
Vertebrate equivalentTo Animal and has-part some Bone
```

Figura 3.12a. Fragmento OWL descrevendo o cão como um animal com ossos
Fonte: Schulz (2016)

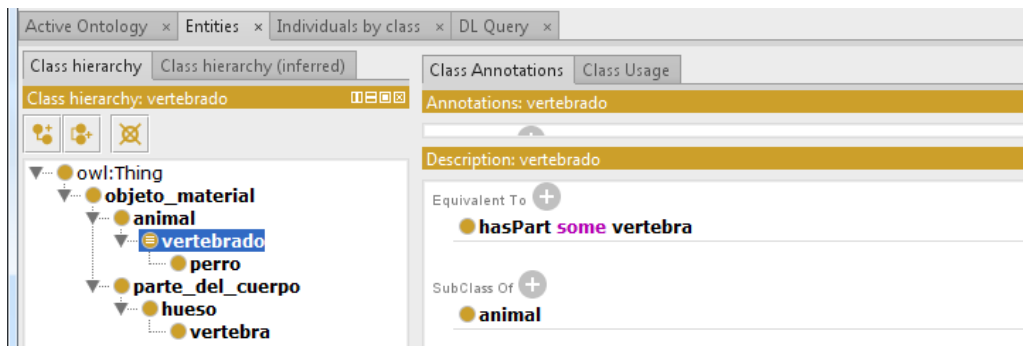


Figura 3.12b. Fragmento OWL no editor de ontologia Protégé
Fonte: Schulz (2016)

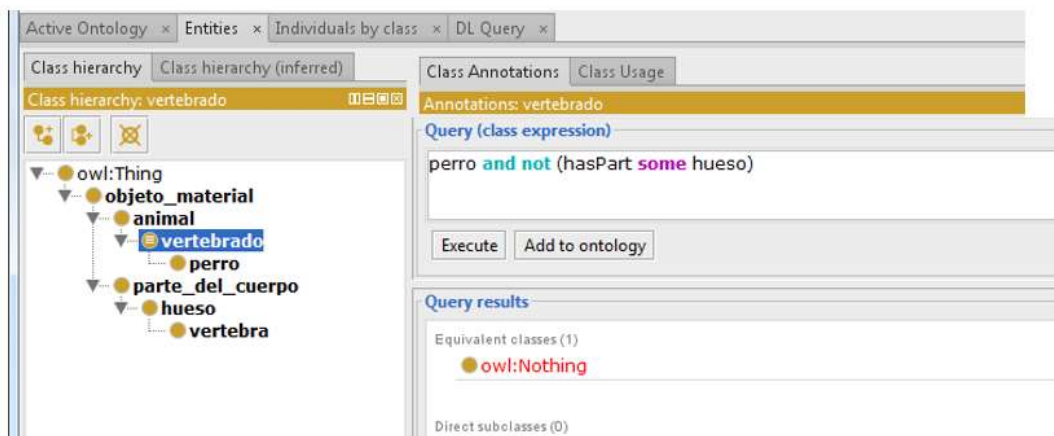


Figura 3.12c. Motor de inferência deduz que não existem cães sem ossos (“nothing”)
Fonte: Schulz (2016)

⁶ <https://protege.stanford.edu/>



Figura 3.12d. Fragmento OWL no editor de ontologia Protegé
 Fonte: Schulz (2016)

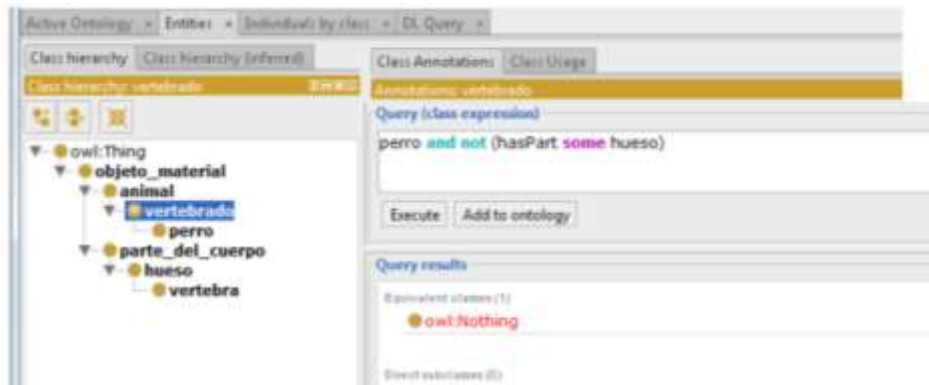


Figura 3.12e. Motor de inferência deduz que não existem cães sem ossos (“nothing”)
 Fonte: Schulz (2016)

As principais linguagens utilizadas para a construção de ontologias são as *lógicas descritivas* ou *lógicas de descrições* [Baader, Horrocks, Sattler 2002]. Esse tipo de lógica foi projetado para facilitar a tarefa de descrever objetos através de definições e propriedades (ver exemplo Figura 3.7). Trata-se de uma família de lógicas, fragmento da lógica de primeira ordem, que teve origens nas lógicas terminológicas dos anos 1990 [Baaeder *et al.* 1992] e vem se desenvolvendo na esteira das linguagens de marcação da Web Semântica.

As linguagens de marcação surgiram nos anos 1970 como um conjunto de convenções para a codificação de textos eletrônicos que especificava as marcas permitidas, as marcas exigidas, a distinção entre as marcas e o texto, e o significado da marcação. A primeira linguagem de marcação foi o *Standard Generalized Markup Language* (SGML), um padrão internacional (ISO-8879) criado para marcar arquivos eletrônicos fornecendo instruções sobre como o texto deveria ser representado.

O SGML fornecia marcas para elementos convencionais dos textos como pontuação, letras maiúsculas e minúsculas, espaços, cabeçalhos, parágrafos, sentenças, etc. No SGML e linguagens subsequentes, a parte marcada é a instância de texto a representar, circundada por elementos com a marca inicial “<” e marca final “>”, conforme exemplo na Figura 3.12.

Quem é capaz de entender a sintaxe da SGML é também capaz de lidar com duas de suas principais aplicações mais populares:

- *Hypertext Markup Language* (HTML), uma aplicação específica do SGML utilizada na Internet, a partir de 1989;
- *Extended Markup Language* (XML), uma versão abreviada da SGML que permite ao autor especificar a forma dos dados no documento;

```

<antologia>
  <poema>
    <titulo> Tenho Tanto Sentimento </titulo>
    <verso>
      <linha> Tenho tanto sentimento </linha>
      <linha> Que é freqüente persuadir-me </linha>
      <linha> De que sou sentimental, </linha>
      <linha> Mas reconheço, ao medir-me, </linha>
      ...
    </verso>
    <verso>
      <linha> Qual porém é a verdadeira</linha>
      <linha> E qual errada, ninguém </linha>
      <linha> Nos saberá explicar; </linha>
      <linha> E vivemos de maneira <linha>
      ...
    </verso>
  </poema>
  <!-- mais poemas aqui -->
</antologia>

```

Figura 3.13. Marcação SGML em texto de poema
Fonte: dos autores

Utilizada a partir de 1998, a XML costuma ser apontado como o início da Web Semântica [Berners-Lee, Hendler, Lassila 2001], onde se buscava melhoria na recuperação da informação via o fornecimento de metadados ao conteúdo textual.

```

<produto>
  <nome língua = "inglês" > book </nome>
  <preço moeda = "dólar" > 45,00 </preço>
  <fornecedor formato = "XLB56" língua = "inglês">
    <rua> Penbridge Square </rua>
    <número> 30 </número>
    <cep> 92310 </cep>
    <país> United Kingdom </país>
  </fornecedor>
</produto>

```

Figura 3.14. Marcação XML com atributos
Fonte: os autores

A evolução do XML não foi exatamente uma linguagem de marcação, mas um padrão de metadados denominado *Resource Description Framework* (RDF). O RDF buscava resolver limitações de expressão do XML impondo limitações estruturais aos dados, proporcionando assim

métodos para expressão semântica e para a publicação de vocabulários legíveis por pessoas e por computadores, bem como a reutilização e a distribuição dos metadados para diferentes comunidades.

O RDF é de um conjunto de descritores organizados segundo um modelo de dados e empregado para descrever entidades de um domínio. O modelo de dados se vale da sintaxe XML e é baseado em triplas *recurso-propriedade-valor* (Figura 3.15) usadas para identificar, respectivamente, entidades, atributos e relações, e instâncias a representar.

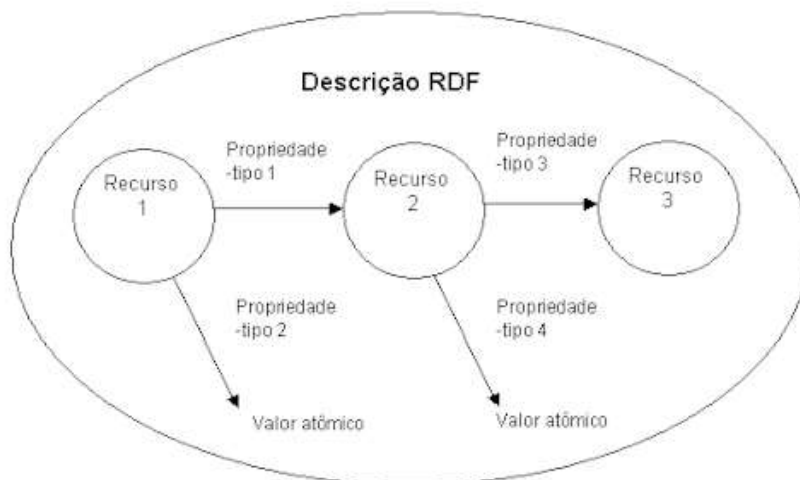


Figura 3.15. Modelo de dados RDF com a tripla objeto
Fonte: os autores

O RDF é utilizado para representar um modelo de dados e armazenar instâncias desse modelo em arquivos eletrônicos, possibilitando o intercâmbio de instâncias entre aplicações. O padrão impõe uma estrutura formal ao XML e proporciona melhor representação semântica.

A aplicação do modelo de dados RDF na questão semântica pode ser melhor entendido por um exemplo simples. Sejam as declarações (a) e (b):

- a) O autor do documento X é João Silva
- b) João Silva é o autor do documento X

Para uma pessoa, as declarações têm o mesmo significado, ou seja, que João Silva é o autor do documento de título X. Entretanto, para um sistema as declarações são conjuntos de caracteres diferentes e é necessário um modelo de dados para a expressão semântica, passível de utilização por sistemas em computadores. O modelo de dados para a declaração “o autor do documento 1 é João Silva” possui: i) um único recurso chamado “documento X”; ii) uma propriedade-tipo chamada “autor”; e um instância ou valor correspondente “João Silva”. A Figura 3.16 apresenta um exemplo de marcação em RDF e o respectivo modelo de dados.

```

</rdf:RDF>
<rdf:Description
  xmlns:rdf="http://www.w3c.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/metadata/dublin_core#"
  rdf:about="http://jornal.com.br/artigos/a4787.html">
  <dc:creator>
    <rdf:Description>
      <x:nome>Jose Silva </x:nome>
      <x:email>js@jornal.com.br</x:email>
      <x:cargo> editor </x:cargo>
    </rdf:Description>
  </dc:creator>
  <dc:publisher> Jornal "O Diário" </dc:publisher>
</rdf:Description>
</rdf:RDF>

```

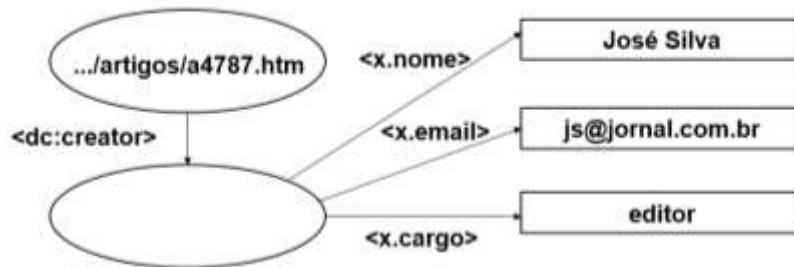


Figura 3.16. Modelo diagramático e marcação RDF
 Fonte: os autores

A evolução do RDF é uma variação conhecida como *Resource Description Framework Schema* (RDFS). É a partir RDFS que se torna possível construir ontologias em função, pela primeira vez, da presença dos constructos “*class*”, “*subclass*”, “*type*” e “*property*”. Além disso, o RDFS contém restrições adicionais:

- “*domain*”, que especifica o domínio de uma propriedade P, ou seja, a classe dos recursos que podem aparecer como objetos em uma tripla da propriedade P;
- “*range*”, que especifica a faixa de valores para uma propriedade P, ou seja, a classe dos recursos que pode aparecer como valores em uma tripla da propriedade P.

A figura 3.17 apresenta um exemplo de marcação RDFS, escrito na linguagem RDF:


```

<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  <rdf:Class
    rdf:about="http://www.schemas.org/2001/01/rdf-schema#Pessoa">
  </rdf:Class>
  <rdf:Class
    rdf:about="http://www.jornal.com.br/.../rdf-schema#Politico">
    <rdf:SubClassOf
      rdf:resource="http://www.schemas.org/2001/01/rdf-schema#Pessoa">
    </rdf:SubClassOf>
  </rdfs:Class>
  <rdf:Class
    rdf:about="http://www.jlocal.com.br/.../rdf-schema#Candidato">
    <rdf:SubClassOf
      rdf:resource="http://www.schemas.org/.../rdf-schema#Politico">
    </rdf:SubClassOf>
  </rdf:RDF>

```

Figura 3.17. Marcação RDFS com os constructos “class” e “subclass”

Fonte: os autores

RDF e RDFS não permitem definir escopo local de propriedades, disjunção de classes, combinação booleana de classes, restrições de cardinalidade e características (matemáticas) de propriedades. Por essas deficiências, desenvolveu-se a OWL, uma linguagem para representação de ontologias com as seguintes características:

- Sintaxe bem definida;
- Semântica formal bem definida;
- Suporte a inferências automáticas:
 - Classificação automática: verificar se uma entidade pertence a uma categoria;
 - Subsunção: verificar se uma categoria é subconjunto de outra pela comparação de suas de suas definições.

A Figura 3.18 apresenta um exemplo de marcação OWL, com a restrição de que só professores são permitidos como valores para a propriedade *isTaughtBy*.

```

<owl:Class rdf:about="#firstYearCourse">
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#isTaughtBy"/>
      <owl:allValuesFrom rdf:resource="#Professor"/>
    </owl:Restriction>
  </rdfs:subClassOf>
</owl:Class>

```

Figura 3.18. Fragmento OWL com restrições de propriedade

Fonte: os autores

Para lidar com o balanço expressividade versus tratabilidade, a OWL foi definida em três versões, da mais expressiva para a menos: OWL Full, OWL DL (*description logics*) e OWL Lite. Os elementos OWL permitem ainda:

- comentários e definições (*owl:comments*);
- controle de versão (*owl:priorVersion*, *owl:versionInfo*, *owl:backwardCompatibleWith*);
- classes equivalentes (*owl:equivalentClass*);
- inclusão de outras ontologias (*owl:imports*);
- propriedades unárias e binárias (*owl:datatypeProperty* e *owl:objectProperty*).

Além disso, a OWL contém os quantificadores lógicos – universal e existencial – que atuam como restrições sobre propriedades e são definidos, respectivamente, pelos elementos *owl:allValuesFrom* e *owl:someValuesFrom*.

O elemento *owl:allValuesFrom* (Fig. 3.19) corresponde ao conjunto de indivíduos (uma classe anônima), os quais, para uma dada propriedade (prop), tem relacionamento apenas com outros indivíduos membros de uma classe específica (classe A).

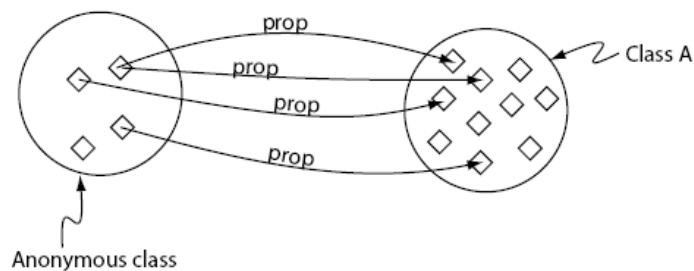


Figura 3.19. Representação do Quantificador universal *allValuesFrom*
Fonte: Horridge (2011)

O elemento *owl:someValuesFrom* (Fig. 3.20) descreve a classe (anônima) de indivíduos, os quais tem pelo menos um tipo de relacionamento através da propriedade (prop) com indivíduos membros de uma classe (classe A).

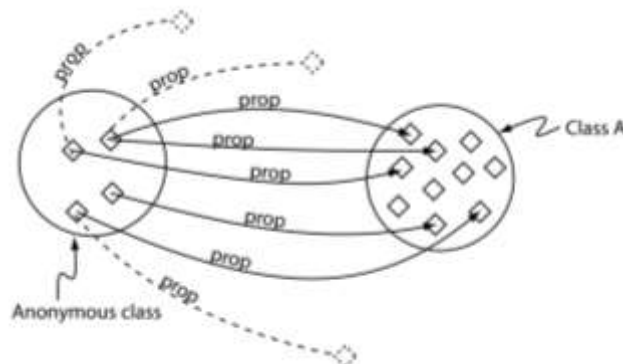


Figura 20. Representação do Quantificador Existencial *someValuesFrom*
Fonte: Horridge (2011)

3.1.3. Aplicações de ontologias em biomedicina

De forma a apresentar a infinidade de aplicações na área médica com o uso de ontologia, em primeiro lugar é preciso entender onde em uma arquitetura de sistemas médicos uma ontologia como artefato pode ser útil em questões de integração e interoperabilidade.

Para isso, descrevem-se níveis de conhecimentos a partir de visão adaptada da Teoria Semiótica de Peirce⁷. A teoria semiótica explica o processo pelo qual o significado é gerado através da percepção e da interpretação de dados sensoriais. Para explicar sua teoria, Peirce criou um triângulo que ficou conhecido genericamente como Triângulo do Significado (Figura 3.21a) contendo um objeto, um signo e o conceito de quem interpreta. A variação mais conhecida desse triângulo é o de Ogden e Richards (1923) (Figura 3.21b).

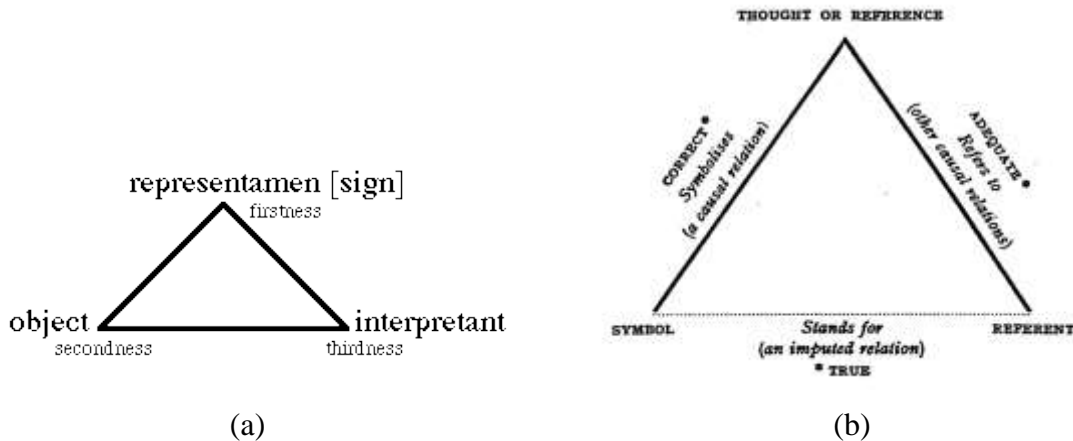


Figura 3.21. Triângulo do Significado original e sua variante linguística
Fonte: Wikipédia

Independentemente do triângulo adotado, a ideia é que a observação de um símbolo traz a mente um conceito (pensamento sobre o referente) pela lembrança de a pessoa já ter visto o objeto (referente). As arestas também tem significados, conforme Figura 3.21b.

Adota-se aqui um a variação proposta por Schulz (2016), que inclui “símbolos”, “universais” e “indivíduos” (Figura 3.22). Pode-se atribuir exemplos a cada vértice, e através desses exemplos explicar as formas de aquisição de conhecimento para utilizar ontologias como artefatos em SIs médicos. Essas formas de aquisição de conhecimento, vamos chamar de “níveis de conhecimento”. Por exemplo, universal “cão” corresponde a todos os cães que existem ou existiram sobre a terra; os símbolos denotam indivíduos, como por exemplo, a palavra “Marley” denota um certo cachorro; e os símbolos também denotam universais, por exemplo, “cão” denota cães em geral. Uma denotação de uma palavra é o conjunto de todas as coisas individuais as quais aquela palavra pode ser usada para se referir. Os indivíduos instanciam os universais, por exemplo, o cão “Marley” é um indivíduo que instancia o universal “cão”.

⁷ Charles Sanders Peirce – filósofo norte americano, 1839–1914

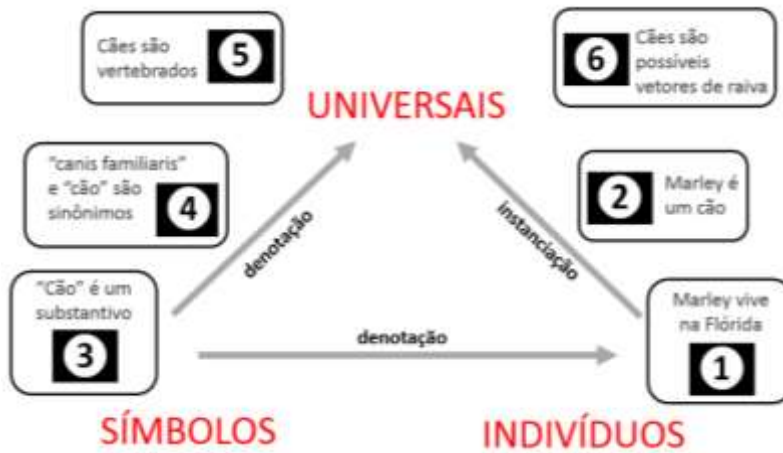


Figura 3.22. Triângulo proposto por Schulz (2016)

Atribui-se então a cada aresta do triângulo um “nível de conhecimento” que é utilizado adiante para organizar as melhores formas de tratar os recursos presentes em SIs médicos. Esses níveis correspondem a formas em que o conhecimento é apresentado (Figura 3.23):

- *Nível linguístico*, que abriga as declarações sobre propriedades da linguagem, e não especificamente das entidades; aqui podem ser tratadas questões como sinonímia, polissemia, e partes dos discursos;
- *Nível factual*, dos fatos realidade, por exemplo, que Marley é um cão que vive na Flórida;
- *Nível contingente*, que abriga as modalidades, o que é possível e típico, por exemplo, que cães são possíveis transmissores de raiva;
- *Nível ontológico*, que abriga as declarações universais, validas em qualquer contexto, por exemplo, que uma hepatite só ocorre no fígado.

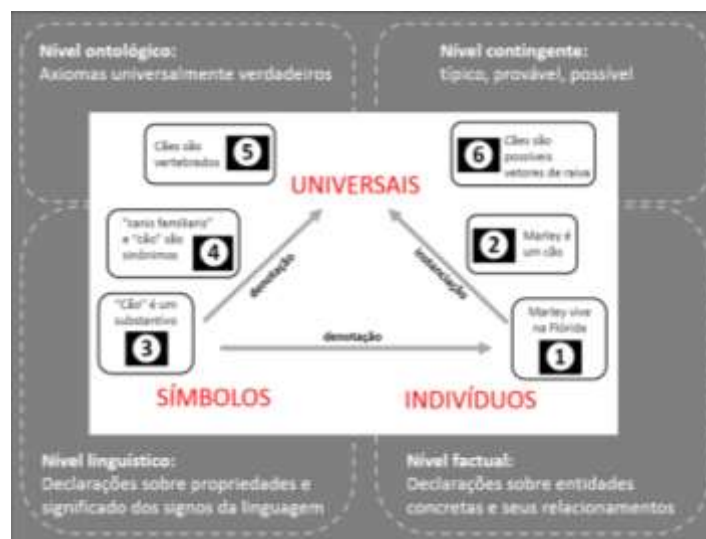


Figura 3.23. Níveis ou formas em que o conhecimento médico se manifesta
Fonte: adaptado de Schulz (2016)

Nível factual:

A linguagem natural parece ser o modo mais efetivo de comunicação entre profissionais de saúde. Mesmo não essa não seja a melhor forma para os SIs, é preciso criar infraestruturas para extração de dados que trabalham com dados não estruturados. Para isso é preciso realizar extração de dados e conexão com terminologias médicas padronizadas (Figura 3.24).

O texto que representa um prontuário de paciente com *myasthenia gravis*, possui termos que podem ser ancorados a vocabulários biomédicos, nesse caso, a SNOMED-CT⁸. Enquanto o prontuário é preenchido, identificadores são automaticamente detectados adicionando metadados ao documento. O texto é segmentado para a criação de uma estrutura semântica onde contextos específicos são identificados, como negações, hipóteses ou planos.

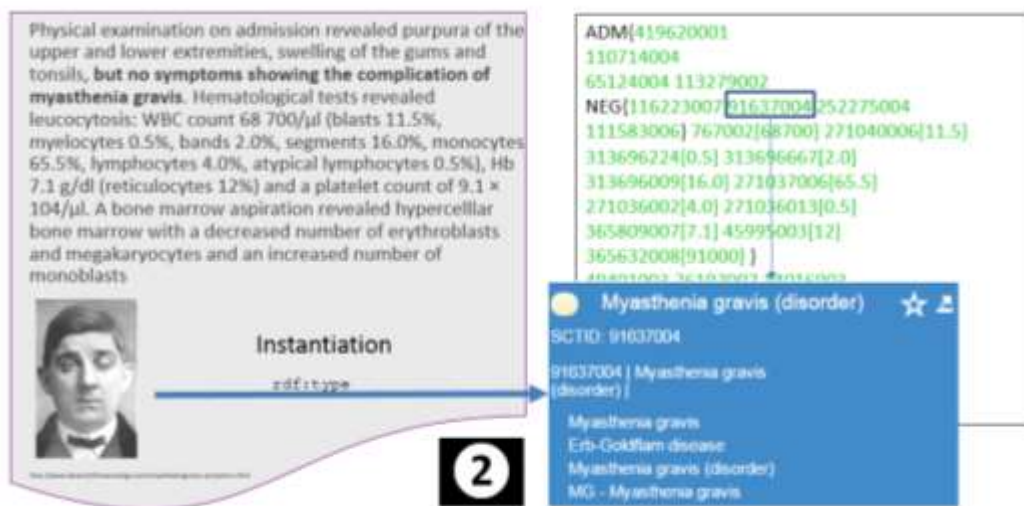


Figura 3.24. Linguagem natural ancorada à terminologia padronizada
Fonte: Schulz et al. (2009)

Nível linguístico:

Um *Part-Of-Speech Tagger* (POS Tagger) é um software que lê um texto e atribui partes do discurso à cada palavra, tal como substantivo, verbo, adjetivo, etc. Esse tipo de classificação, serve de suporte a várias possibilidades de tratamento de massas textuais, como por exemplo, o uso na construção de ontologias e o uso em algoritmos de *machine learning*.

A Figura 3.25 mostra um exemplo com os códigos do Quadro 3.2.

⁸ <http://www.snomed.org/>

Quadro 3.2. Códigos do POS-Tagger

CC Coordinating conjunction	NN Noun, singular or mass	SYM Symbol
CD Cardinal number	NNP Proper noun, singular	UH Interjection
DT Determiner	NNPS Proper noun, plural	VB Verb, base form
EX Existential there	PDT Predeterminer	VBD Verb, past tense
FW Foreign word	POS Possessive ending	VBG Verb, gerund or present p.
IN Preposition or subordinating	PRP Personal pronoun	VBN Verb, past participle
JJ Adjective	PRPS Possessive pronoun	VBZ Verb, 3rd person singular p.
JJR Adjective, comparative	RB Adverb	WDT Wh-determiner
JJS Adjective, superlative	RBR Adverb, comparative	WP Wh-pronoun
LS List item marker	RBS Adverb, superlative	WPS Possessive wh-pronoun
MD Modal	RP Particle	WRB Wh-adverb

Fonte: Cognitive Computation Group @ Illinois⁹

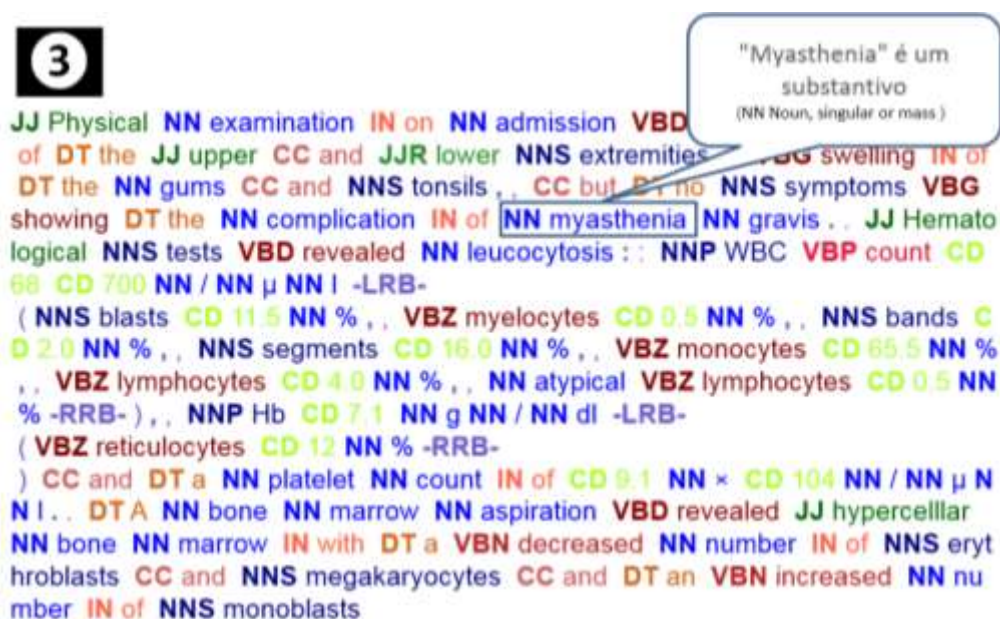


Figura 3.25. Partes do discurso marcadas automaticamente em um prontuário

Fonte: Cognitive Computation Group @ Illinois

Nível contingente:

O MeSH¹⁰ é um vocabulário controlado de medicina usado para indexação de documentos e recuperação da informação. De acordo como MeSH, uma “desordem bipolar” pode ser classificada como terapia, ou prevenção ou complicação (Figura 3.26). São possibilidades, ou seja, exemplos de declarações contingentes no mundo dos SIs médicos

⁹ <http://cogcomp.cs.illinois.edu/>

¹⁰ <https://www.ncbi.nlm.nih.gov/mesh>

Source concept	Name	Bipolar disorder
	Type	Disorder
Target concept	Name	Tricyclic antidepressant
	Type	Substance
MeSH subheadings	DT=9,CI=7,DI=5,PX=4,CO=2,EP=2,GE=2,BL=1,ET=1,PA=1,PC=1,PP=1,TH=1	
	Absolute co-occurrence	17
	Log-likelihood	54.57
	Qualifies source concept, e.g: DT = "drug therapy" PC = "prevention and control" CO = "complication"	

Figura 3.26. Declarações contingentes em vocabulário médico padronizado
 Fonte: Miñarro-Giménez et al. (2015)

Nível ontológico:

No nível ontológico existem declarações formais que favorecem a integração de dados, uma vez que definem explicitamente a semântica dos termos. Na verdade, axiomas formais restringem o significado dos termos, por exemplo:

- *Nucleous is-part-of cell*
- *Nucleous is-contained-in cell*
- *Nucleous is-component-of cell*

Axiomas reais, em um editor de ontologias, foram apresentados anteriormente na Figura 12. Os modelos ontológicos podem resultar em benefícios reais em termos de interoperabilidade semântica visto a redução de ambiguidade e a comunicação precisa que podem prover. Entretanto, modelos ontológicos não são uma panaceia que resolve todos os problemas de integração de sistemas. Na verdade, existem problemas com a representação em ontologias:

- Fornecem apenas representações parciais da realidade
- Não possuem a expressividade necessária para criar definições completas de entidades em todos os casos
- Dificuldade para lidar com “background knowledge”, por exemplo:
 - A declaração “*hand has-part thumb*” descarta a possibilidade de uma mão sem polegar, por exemplo, depois de um acidente;
 - A declaração “*hypertensive_Disease is-a Risk_Factor_for_Aneurysm_Rupture*” é verdadeira apenas em certas circunstâncias, não em todos os casos

Uma vez apresentados os quatro níveis de conhecimento que podem ser encontrados em SIs médicos, é possível desenhar uma arquitetura de sistemas (Figura 26) que faça uso do

conhecimento disponível de acordo com suas características e de forma a atender as questões para o qual o SI foi planejado.

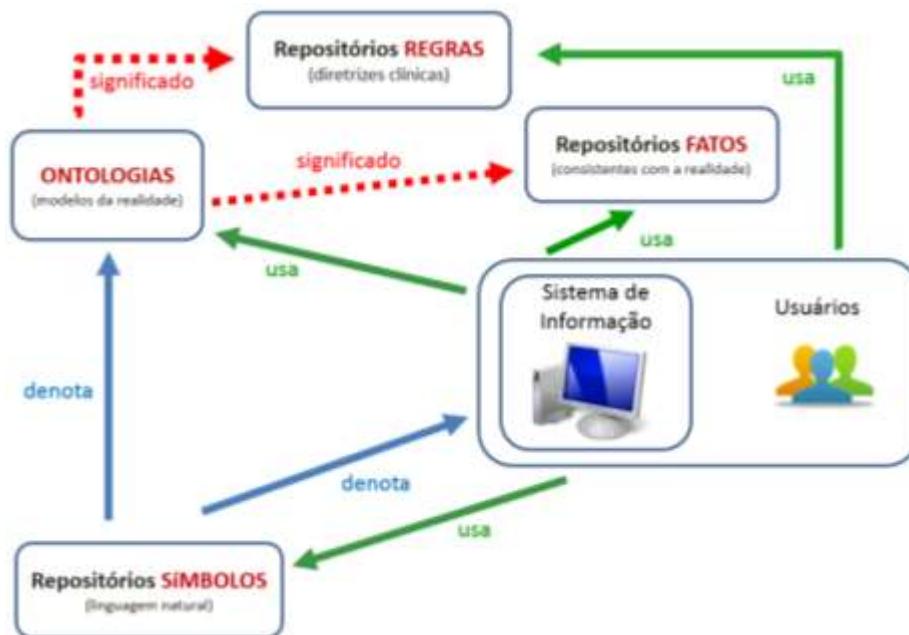


Figura 3.27. arquitetura geral de sistemas médicos: conectando tudo
Fonte: os autores

3.2. Práticas em construção de ontologia.

Na construção de artefatos ontológicos, existem diretrizes que orientam a construção de bons artefatos, a seguir, apresenta-se uma breve explicação de três princípios sugeridos por [Schulz *et al.* 2012]:

- Os artefatos ontológicos devem ser formais, ou seja, suas entidades (termos) são representadas sem ambiguidade de significado e a especificação das relações entre os termos incluem axiomas lógicos, permitindo a automatização tanto da recuperação quanto do raciocínio.
- As ontologias usam especificações explícitas permitindo que as pessoas raciocinem e entendam o domínio representado pela ontologia.
- As ontologias devem ser adequadas ao domínio alvo da representação, significa que a ontologia construída deve representar a realidade e o conhecimento presentes no domínio desejado.

Existem várias metodologias para o desenvolvimento de ontologias, sem um consenso sobre qual metodologia é mais apropriada. A lista abaixo apresenta algumas das metodologias mais conhecidas de meados dos anos 90 até os dias atuais:

- *Toronto Virtual Enterprise* (TOVE) [Grüninger, Fox 1995].
- *Methontology* [Fernández-López *et al.* 1997].

- Metodologia *NeOn* (*Network Ontology*) [Suárez-Figueroa 2010].
- *Systematic Approach for Building Ontologies* (SABiO) [Falbo 2014].
- *Up for ONtology* (UPON) [De Nicola *et al.* 2005].
- Metodologia do realismo ontológico [Arp *et al.* 2015, Smith, Ceusters 2010].

Algumas metodologias focam em aspectos semânticos para a construção de artefatos ontológicos, como é o caso da metodologia do realismo ontológico, e outras focam em aspectos do processo de engenharia de ontologias, como a metodologia *NeOn* que tem como base o processo de desenvolvimento de software. Para maximizar a aderência as diretrizes apresentadas por Schulz *et al.* (2012), o ideal combinar duas ou mais metodologias como apresentado por Farinelli e Elkin (2017) e Farinelli (2017) e ilustrado na Figura 3.28.

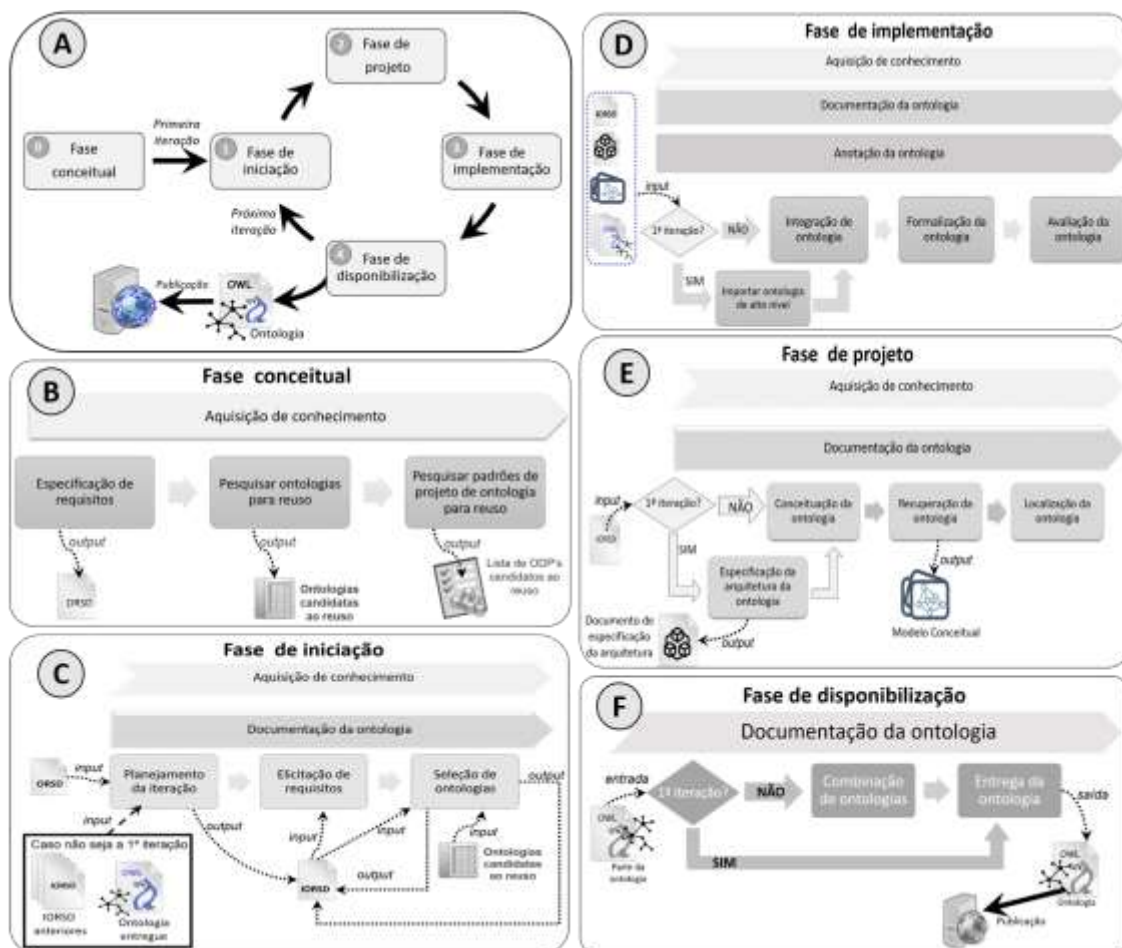


Figura 3.28. Fases do processo de construção de ontologias (A) e suas respectivas atividades (B, C, D, E e F).

Esta metodologia combina as cinco etapas da metodologia do realismo ontológico [Arp *et al.* 2015, Smith, Ceusters 2010] aos cenários e atividades da metodologia *NeOn* [Suárez-Figueroa 2010]. Sugere-se a adoção de um ciclo de vida de desenvolvimento, iterativo-incremental, como sugerido

pelas duas metodologias de referência. Nas próximas subseções são detalhadas as principais atividades práticas presentes no processo de construção de ontologias biomédicas, esta metodologia foi utilizada para a construção da OntONEo¹¹ (Ontologia do domínio Obstétrico e Neonatal) que atualmente se encontra publicada no portal OBO Foundry.

3.2.1. OBO Foundry

A OBO Foundry é uma iniciativa que surgiu com o objetivo de construir um conjunto de artefatos ontológicos¹² que representem os termos da área biomédica, minimizando as ambiguidades de termos e definições entre os artefatos de forma a torná-los não redundantes, além de promover a integração entre os artefatos criados [Ghazvinian *et al.* 2011, Smith *et al.* 2007].

O acrônimo OBO originalmente se referia a “Ontologias Biomédicas Abertas” (em inglês *Open Biomedical Ontologies*) e mais tarde tornou-se “Ontologias Biológicas e Biomédicas Abertas” (em inglês *Open Biological and Biomedical Ontologies*). A ideia de fundação refere-se à criação de um conjunto de ontologias biomédicas interoperáveis, e bem formadas e que claramente representem o conhecimento científico biomédico, tornando-se uma referência na representação deste conhecimento.

A OBO Foundry fornece um conjunto de princípios¹³ considerados como boas práticas na construção de ontologias. Assim, os artefatos ontológicos que são publicados no portal da OBO Foundry devem ser aderentes a tais princípios. Algumas premissas ou princípios são o uso da BFO como ontologia de alto nível e a adoção de um conjunto comum de relações atualmente existentes na ontologia RO (*Relation Ontology*). Para mais detalhes sobre as ontologias BFO e RO disponível em Smith *et al.* (2005) Smith *et al.* (2015) e Arp *et al.* (2015).

Existem 13 princípios definidos na OBO Foundry, como por exemplo: permitir o uso aberto da ontologia, evitar sobreposições de termos e escopo entre ontologias, prover desenvolvimento colaborativo, garantir formato de publicação comum, entre outros que podem ser consultados no portal OBO Foundry¹⁴.

Atualmente, a OBO Foundry possui uma lista de cento e cinquenta e oito ontologias biomédicas, sendo oito ontologias fundacionais (*status Foundry*), cento e nove ontologias possuem a *status Library* que significa que a ontologia foi concebida conforme os princípios da OBO, e 41 ontologias ainda sem especificação de status.

3.2.2. Reuso de ontologias

A prática de reuso é uma estratégia de desenvolvimento que foca no uso de conceitos ou definições, entidades ou classes, relacionamentos, previamente elaborados ou definidos para criação de um novo artefato ontológico. Assim como no desenvolvimento de software, busca reduzir o custo de tempo de construção do artefato ontológico, além de reduzir ambiguidades e duplicação conceituais e semânticas de termos e entidades.

¹¹ Ontologia disponível bilíngue (Português e Inglês) disponível em: <https://ontoneo.com/>.

¹² Artefatos ontológicos são o mesmo que “*ontologia como um artefato*” conforme quadro 3.1.

¹³ Detalhes sobre os princípios em: <http://www.obofoundry.org/principles/fp-000-summary.html>

¹⁴ Disponível em: <http://www.obofoundry.org/>

O reuso de artefatos ontológicos minimiza retrabalho no desenvolvimento de um novo projeto, sempre levando em consideração trabalhos anteriores, fazendo com que soluções previamente desenvolvidas sejam aproveitadas e implementadas em novos contextos. No caso das ontologias biomédicas fundamentadas na BFO, o portal OBO Foundry agrega diversas áreas de domínio já cobertas e representadas, servindo assim de uma ótima fonte de artefatos para reuso. Ressalta-se, porém, que as ontologias da OBO Foundry atualmente estão definidas apenas em inglês, exceto a OntONeo que já tem definições em português.

As atividades relacionadas ao reuso de ontologias ou artefatos ontológicos, seja o reuso total ou parcial do artefato, está presente em diversas fases do processo de construção de ontologias (Figura 3.28 - A), divergindo quanto a grau de detalhamento da atividade.

Na fase conceitual do processo (Figura 3.28 - B), deve-se identificar o conjunto de artefatos ontológicos candidatos ao reuso. Para isso, deve-se pesquisar em repositórios de ontologias termos que dizem respeito ao domínio em modelagem.

Existem vários repositórios e mecanismos de busca para artefatos ontológicos biomédicos para apoiar a identificação de artefatos ontológicos candidatos ao reuso, por exemplo:

- Ontobee¹⁵: Um repositório de artefatos ontológicos que seguem as recomendações da OBO Foundry, que possibilita a pesquisa por artefatos [Xiang *et al.* 2011].
- Bioportal¹⁶: Um repositório web de ontologias e terminologias biomédicas, desenvolvido e mantido pelo Centro Nacional de Ontologia Biomédica (NCBO), que possibilita a pesquisa por artefatos [Noy *et al.* 2008].
- Ontology Lookup Service (OLS): Um repositório web para ontologias biomédicas, que é desenvolvido e mantido pelo Instituto Europeu de Bioinformática (EMBL-EBI) [Côté *et al.* 2006].

No caso de construção de artefatos ontológicos biomédicos que sejam fundamentados na BFO, sugere-se usar preferencialmente o portal *Ontobee* (Figura 3.29), e apenas no caso de não encontrar nenhum artefato candidato ao reuso utilizar os outros dois para guiar uma nova definição.

Na fase de iniciação (Figura 3.28 - C), deve-se efetivamente realizar a coleta dos artefatos que serão reutilizados. Neste ponto, é importante incorporar ao projeto as diretrizes introduzidas por Courtot *et al.* (2011) conhecidas como MIREOT (do inglês *Minimal Information to Reference External Ontology Terms*). Estas diretrizes recomendam que na reutilização de artefatos ontológicos, deve-se reutilizar estritamente os termos relevantes de outras ontologias ao invés de toda a ontologia. Portanto, de acordo com MIREOT, um ontologista deve importar de outra ontologia apenas elementos necessários (entidades, termos, propriedades e relações) para serem reutilizados, evitando a sobrecarga de importar toda a ontologia.

¹⁵ Disponível em: <http://www.ontobee.org/>

¹⁶ Disponível em: <https://biportal.bioontology.org/>

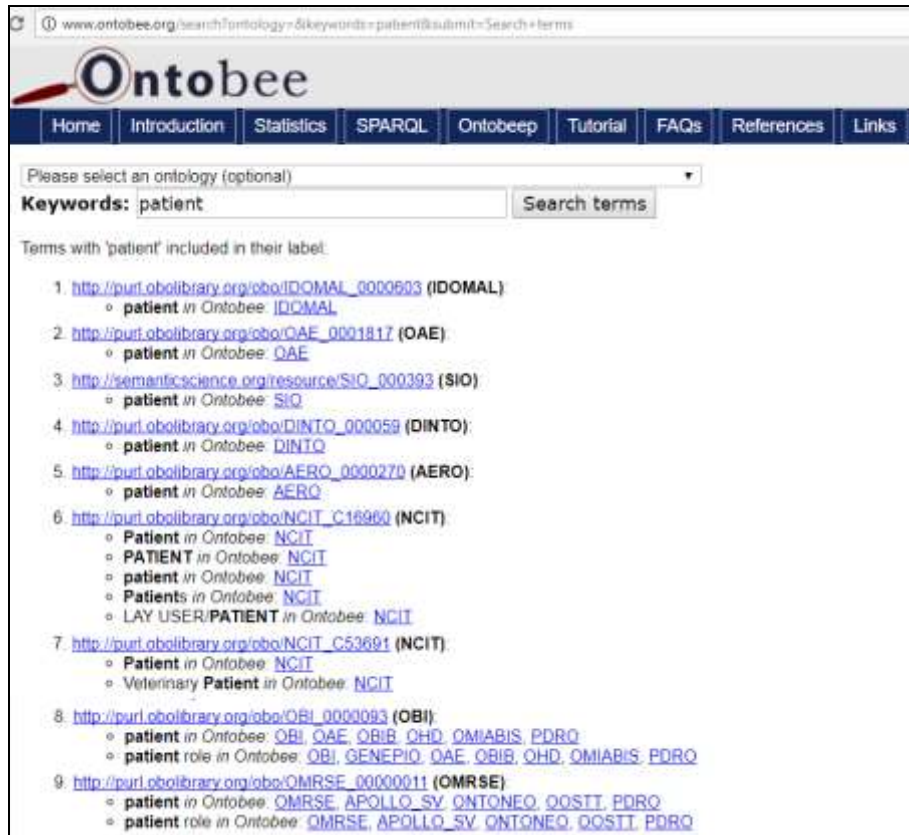


Figura 3.29. Exemplo de pesquisa por termos no portal Ontobee.

As ontologias como um artefato são formalizadas usando uma linguagem formal, em geral, utilizando a Ontology Web Language (OWL). Para detalhes sobre a linguagem OWL consulte Antoniou e Harmelen (2009), W3c (2004a) e W3c (2004b). No caso das ontologias da OBO Foundry, elas devem ser expressas por OWL.

Para reutilizar elementos de artefatos ontológicos já existentes é necessário recortar os elementos para reuso no arquivo OWL da ontologia preexistente. Para garantir a integridade dos elementos que serão reutilizados, sugere-se o uso da ferramenta Ontofox [Xiang *et al.* 2010].

A OntoFox é uma ferramenta baseada na Web que implementa as diretrizes MIREOT extraindo das ontologias apenas as informações mínimas necessárias de acordo com as entradas dos usuários e gerando um arquivo de saída no formato OWL (Figura 3.30).

O arquivo OWL gerado pela Ontofox, contendo o subconjunto de elementos do artefato ontológico em reuso, deverá ser importado no artefato ontológico em desenvolvimento.

OntoFox

Home Introduction Tutorial FAQs References Download Links Contact Acknowledge News

Ontofox is a web-based Ontology tool that fetches ontology terms and axioms. Ontofox supports ontology reuse. It allows users to input terms, fetch selected properties, annotations, and certain classes of related terms from source ontologies and save the results using the RDF/XML serialization of the OWL. Ontofox follows and expands the [MIREOT](#) principle. Inspired by existing ontology modularization techniques, Ontofox also develops a new SPARQL-based ontology term extraction algorithm that extracts terms related to a given set of signature terms. In addition, Ontofox provides an option to extract the hierarchy rooted at a specified ontology term. **Note:** We have now changed the name "OntoFox" (with capital F) to "Ontofox" (with low case f) to be consistent with other Ontoanimal tools.

Notice: All the OBO ontologies have changed the term URI format from http://purl.org/obo/owl/ontology#ontology_nnnnnnn to http://purl.obolibrary.org/obo/ontology_nnnnnnn. Please make sure your input files are updated.

Ontofox is implemented using one of the following three methods, based on how data is input and whether the Ontofox web interface is used:

1. Data input using web forms:
 Examples: [Example 1](#) [example 2](#) [example 3](#) [example 4](#) [example 5](#)

(1) Select one ontology:

Or enter your favorite source ontology and SPARQL endpoint: [Example](#)

(2) Term specification:

(a) Include low level source term URIs:
 (One URI per line. To include all child terms of a source term (extract the whole branch), enter "includeAllChildren" in the line next to the source term)
 Search a term:

```

http://purl.obolibrary.org/obo/CL_000007 #early embryonic cell
includeAllChildren
http://purl.obolibrary.org/obo/GO_0005575 #cellular_component
http://purl.obolibrary.org/obo/CL_000586 #germ cell
includeAllChildren
http://purl.obolibrary.org/obo/CL_0002321 #embryonic cell
includeAllChildren
  
```

(b) Include top level source term URIs and target direct superclass URIs (One URI per line, optional):
 Search a term:

```

http://purl.obolibrary.org/obo/CL_000007 #early embryonic cell
subClassOf http://purl.obolibrary.org/obo/GO_0005575 #cellular_component
http://purl.obolibrary.org/obo/CL_000586 #germ cell
subClassOf http://purl.obolibrary.org/obo/GO_0005575 #cellular_component
http://purl.obolibrary.org/obo/CL_0002321 #embryonic cell
subClassOf http://purl.obolibrary.org/obo/GO_0005575 #cellular_component
  
```

(c) Select a setting for retrieving intermediate source terms:

(3) Annotation/Axiom Specification: Include source annotation URIs (One URI per line, optional):
 Examples: [rdfs:label](#) [iso:preferredTerm](#) [iso:definition](#) [iso:alternativeTerm](#) [obo:owl:hasDefinition](#) [obo:owl:hasSynonym](#) [owl:equivalentClass](#)
 The default is no annotation to be assigned. Use [includeAllAnnotationProperties](#) to include all annotations. Use [includeAllAxioms](#) to include all annotations and other related axioms. Use [includeAllAxiomsRecursively](#) to include all axioms for the specified terms and the related terms recursively.

(4) Annotation/Axiom to be excluded (One URI per line, optional):

(5) URI of the OWL(RDF/XML) output file:
 Example http://purl.obolibrary.org/obo/voia/external/NCBITaxon_import.owl

Figura 3.30. Exemplo de extração de elementos para reuso na ontologia *Cell Ontology*.

3.2.3. Definição de *namespace* e URIs

Na fase de projeto do processo de construção de artefatos ontológicos (Figura 3.28 - E), uma atividade chave é determinar elementos arquiteturais da ontologia que sejam aderentes aos princípios da OBO Foundry. Nesta fase são definidos:

- Prefixo do artefato ontológico.
- Identificador local dos elementos do artefato ontológico.
- *Namespace* do artefato ontológico e URI padrão do artefato e de seus elementos.

O prefixo é um acrônimo que será associado ao artefato ontológico quando for necessário fazer menção tanto ao artefato quanto aos elementos do artefato. Por exemplo, o prefixo é comumente utilizado na composição do *namespace* e do identificador dos elementos do artefato.

Para determinar o prefixo da Ontologia do domínio Obstétrico e Neonatal, foi definido um prefixo candidato, ONTONEO, e realizada uma pesquisa na lista de ontologias da OBO Foundry para garantir que o prefixo candidato não era usado por nenhum artefato já existente. Em seguida, utilizou-se o serviço prefix.cc, para garantir que nenhum outro recurso usasse o prefixo em questão. *Prefix.cc* (Figura 3.31) é um serviço criado para desenvolvedores de RDF no qual os ajuda na busca por prefixos de URI e *namespaces*.



Figura 3.31. Tela de pesquisa do serviço prefix.cc.

A OBO Foundry recomenda que o identificador local dos elementos do artefato ontológico seja uma concatenação do prefixo (também identificado como ID-Space), um caractere “_” (sublinhado ou *underscore*) e uma sequência de números (Figura 3.32).

OBO Foundry local identifier: <IDSPACE>_<NUMBER>
Example: ONTONEO_00000001

Figura 3.32. Exemplo do padrão de nomes OBO Foundry.

Um *namespace* é um mecanismo para fornecer contexto genérico para encapsular itens. O *namespace* ou espaço de nomes é um sistema de nomeação de objetos que delimita o espaço abstrato para os itens que ele armazena, permite a desambiguação para itens ou objetos que possuem o mesmo nome mas que residem em espaços de nomes diferentes [Antoniou, Harmelen 2009]. Assim, cada ontologia tem seu próprio *namespace*, que é uma sequência de caracteres que precede os identificadores de seus elementos. Um *namespace* padrão é um URI (*Uniform Resource Identifiers*) ou IRI (*Internationalized Resource Identifiers*) válido seguido por um caractere “/” ou “#”. Para isso, definiu um URL persistente (purl) usando o Serviço PURL (*PURL Service*¹⁷). No caso da OntONEo, seguindo as recomendações da OBO Foundry, para

¹⁷ Disponível em: <https://archive.org/services/purl/>.

utilizar o domínio OBO PURL (<http://purl.obolibrary.org/obo/>), a URL definida foi <http://purl.obolibrary.org/obo/ontoneo>. O serviço *prefix.cc* também pode ser usado para verificar a existência ou não da URL ou *namespace* em questão.

3.2.4. Utilização da ferramenta Protégé

O Protégé é um ambiente de desenvolvimento integrado (IDE, do inglês *Integrated Development Environment*) gratuito e de código aberto para a construção de artefatos ontológicos e funcionalidades afins [Almeida 2006, Noy *et al.* 2003].

Configurações iniciais do Protégé:

Na fase conceitual do processo (Figura 3.28 - D), para iniciar a implementação da ontologia é necessário configurar o Protégé para evitar retrabalho durante o processo de formalização.

As configurações básicas que devem ser realizadas são: (i) URI/IRI base para os elementos do artefato (Figura 3.33 - A); (ii) Anotações básicas do artefato: Nome do criador e data de criação (Figura 3.33 - B e C); (iii) Geração automática do identificador local para os elementos do artefato (Figura 3.34); (iv) URI/IRI do padrão do artefato e de versionamento do artefato (Figura 3.35).

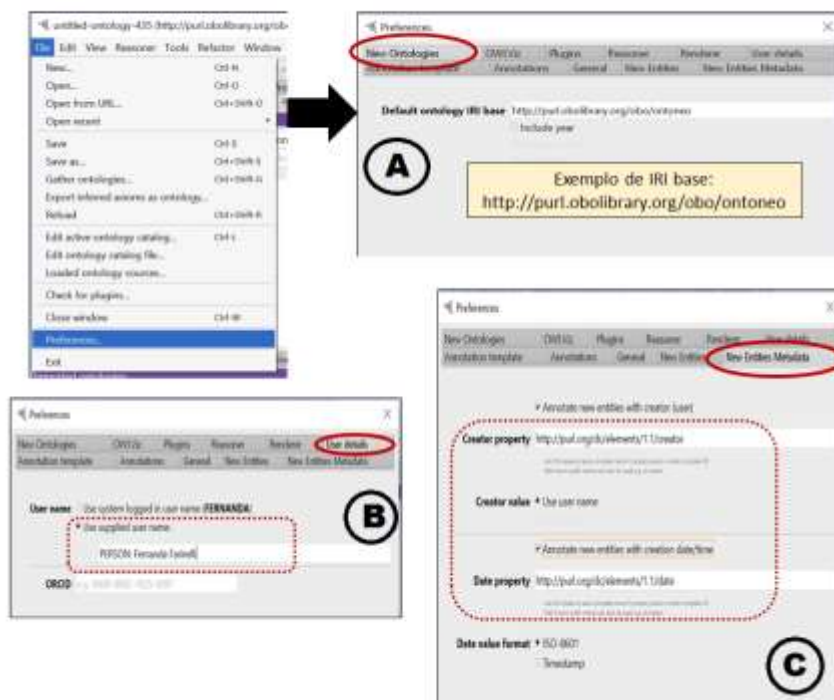


Figura 3.33. Tela de preferências do Protégé para configurações básicas.

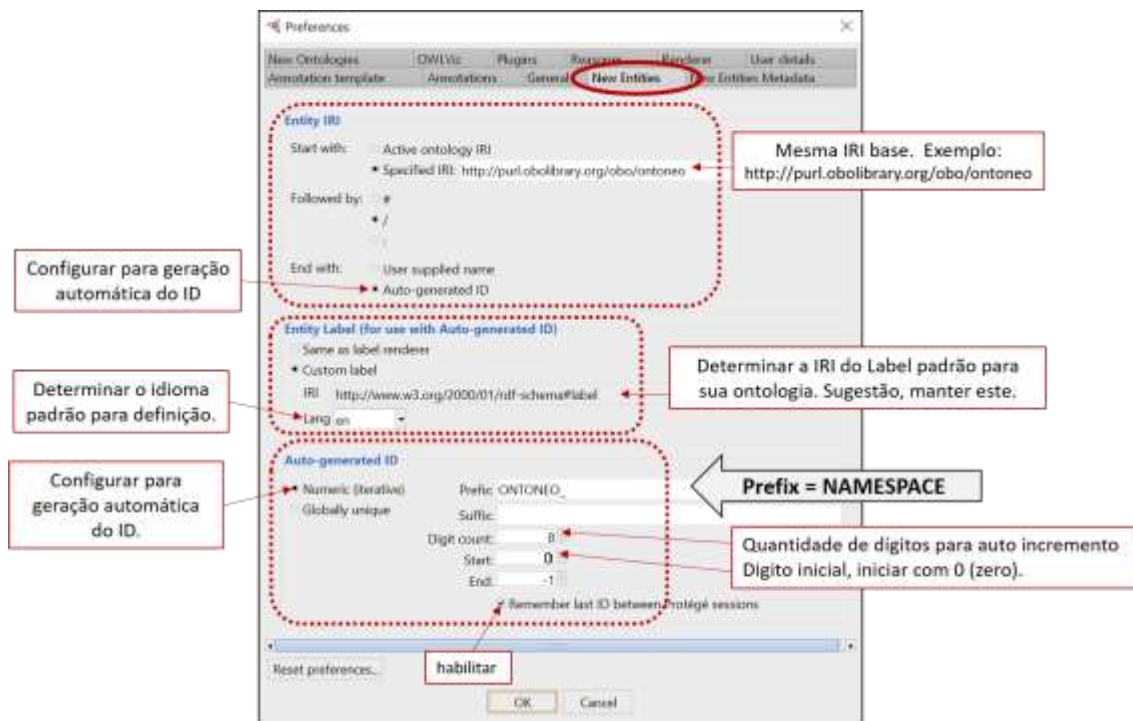


Figura 3.34. Tela de preferências do Protégé para configurações básicas.

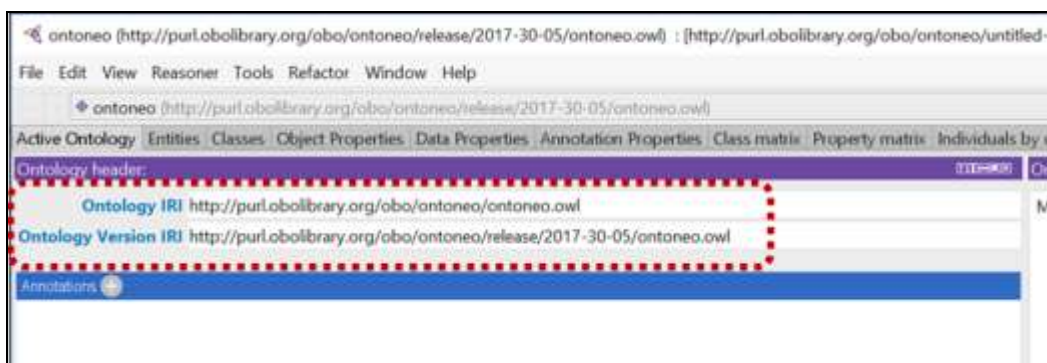


Figura 3.35. Tela inicial do Protégé para configurações da IRI e IRI de versionamento.

Importação de artefatos ontológicos em reuso

Ainda na fase de implementação do processo de construção do artefato ontológico (Figura 3.28 - D), está prevista a importação de outros artefatos ontológicos, seja tal artefato a ontologia de alto nível que vai fundamentar o artefato em construções ou seja os artefatos ontológicos que serão reutilizados.

No caso de o projeto seguir os princípios da OBO Foundry, primeiramente deve-se importar a ontologia de alto nível BFO¹⁸, seguido pela importação da ontologia de relacionamentos

¹⁸ Arquivo OWL disponível para download em: <http://www.obofoundry.org/ontology/bfo.html>

a RO¹⁹. Após a importação dos artefatos ontológicos base, deve-se realizar a importação dos artefatos ontológicos que serão reutilizados. Estes artefatos são os arquivos OWL gerados pelo Ontofox devem ser importados para seu projeto.

A Figura 3.36 abaixo ilustra o passo a passo para importação de artefatos ontológicos que se encontram disponíveis em arquivo OWL. Lembre-se de organizar as entidades importadas das ontologias em reuso conforme a hierarquia de entidades da ontologia de alto nível em uso.

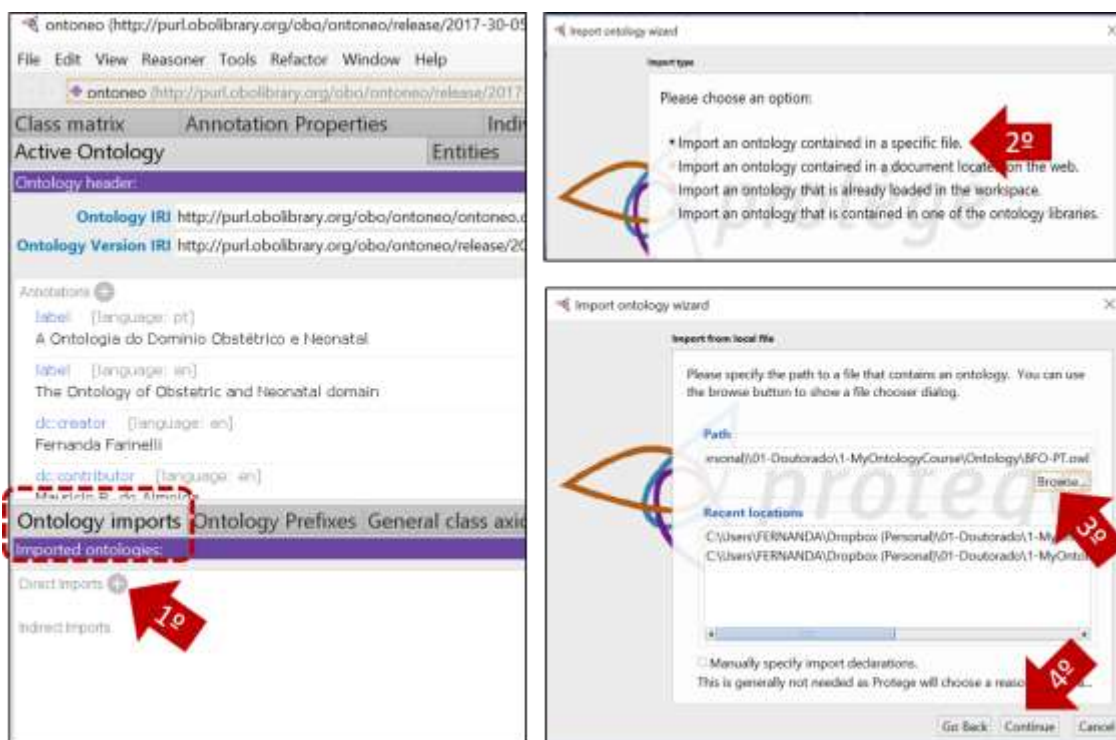


Figura 3.36. Tela inicial do Protégé para importação de artefatos ontológicos para reuso no projeto.

Documentação do artefato ontológico

A atividade de documentação de ontologia é uma atividade de suporte que gera qualquer documentação valiosa para entender o próprio artefato ontológico e as decisões tomadas durante seu desenvolvimento (Figura 3.28 - D).

A documentação do artefato pode ser realizada no próprio artefato por meio das anotações disponíveis. As anotações (*annotation*) podem ser vistas como metadados do artefato. Não existe um padrão de documentação, no geral, cada desenvolvedor de ontologia determina quais são as anotações que são relevantes para seu próprio projeto. Considerando a experiência no desenvolvimento da OntONeo, recomenda-se a adoção das anotações listadas na tabela 3.1 para enriquecer as informações no artefato em desenvolvimento. Esta lista foi obtida pela compilação das diversas anotações utilizadas nos projetos disponíveis na OBO Foundry.

¹⁹ Arquivo OWL disponível para download em: <http://www.obofoundry.org/ontology/ro.html>

Tabela 3.1. Lista de anotações recomendadas para documentar o artefato ontológico.

PREFIX	Origem da anotação <i>namespace</i>	Nome da anotação	Escopo de uso
dc	http://purl.org/dc/elements/1.1/	creator	Ontologia
		contributor	Ontologia
		license	Ontologia
		date	Qualquer elemento
		source	Importados
foaf	http://xmlns.com/foaf/0.1/	homepage	Ontologia
		mbox	Ontologia
rdf	http://www.w3.org/1999/02/22-rdf-syntax-ns#	Description	Qualquer elemento
rdfs	http://www.w3.org/2000/01/rdf-schema#	label	Ontologia
		range	Propriedades
		domain	Propriedades
		comment	Qualquer elemento
		isDefinedBy	Qualquer elemento
owl	http://www.w3.org/2002/07/owl#	versionIRI	Ontologia
		versionInfo	Ontologia
		priorVersion	Ontologia
obo	http://purl.obolibrary.org/obo/	IAO_0000111 editor preferred term	Entidade
		IAO_0000114 has curation status	Entidade e Propriedades
		IAO_0000115 definition	Entidade
		IAO_0000117 term editor	Entidade
		IAO_0000118 alternative term	Entidade
		IAO_0000119 definition source	Entidade
		IAO_0000412 imported from	Importados
obolnOwl	http://www.geneOntologia.org/formats/obolnOwl#	id	Entidade
		created_by	Qualquer elemento
		creation_date	Qualquer elemento
		hasDbXref	Entidade
owl	http://www.w3.org/2002/07/owl#	versionInfo	Ontologia
		priorVersion	Ontologia
protege	http://protege.stanford.edu/plugins/owl/protege#	defaultLanguage	Ontologia

3.2.5. Definição do artefato ontológico

O processo de construção de artefatos ontológicos, como ilustrado na figura 3.28 – A envolve uma fase exclusiva, a fase conceitual (Figura 3.28 – B), para determinar o propósito da ontologia, o

escopo ou domínio de cobertura do artefato a ser construído, e o conjunto de requisitos que o artefato deve satisfazer.

Como uma das primeiras atividades no processo de construção de ontologias, é importante realizar a identificação dos requisitos que a ontologia de cobrir. Para isso, sugere-se seguir as orientações apresentadas em [Suárez-Figueroa *et al.* 2009].

O processo de construção de ontologias envolve muitas outras atividades que descritas em Farinelli e Elkin (2017) e Farinelli (2017).

References

- Almeida, M. B. D. (2006) “Noções Básicas para uso do Protégé”, online, Tutorial, 09/07/2006. <http://mba.eci.ufmg.br/onto_frames/>, 16/04/2019.
- Almeida, M.B. (2013). Revisiting Ontologies: a necessary clarification. *Journal of the American Society for Information Science and Technology*. vol. 64, n. 8, p. 1682–1693.
- Abrial, J. R. (1974). “Data Semantics”. In J. W. Klimbie & K. L. Koffeman, (Eds.), *Proceedings of the IFIP Working Conference Data Base Management* (pp. 1-60). Amsterdam: North-Holland.
- Ackrill, J. L. (1963). *Aristotle: Categories and De Interpretatione*. Oxford: Clarendon Press.
- Alturki, A.; Gable, G.; Bandara, W. (2013). “BWW ontology as a lens on IS design theory: extending the design science research roadmap”. *Proceedings of the 8th International Conference on Design Science at the Intersection of Physical and Virtual Design*. p. 258-277.
- Antoniou, G.Harmelen, F. V. (2009). “Web Ontology Language: OWL”. In: Staab, S.Studer, R. (Ed.). *Handbook on Ontologies*. Berlin, Heidelberg: Springer Berlin Heidelberg. p.91-110.
- Arp, R., Smith, B., Spear, A. D. (2015). *Building Ontologies with Basic Formal Ontology*. Cambridge, Massachusetts: The Mit Press, 2015.
- Baader, F., Bürckert, H., Hollunder, H, Laux, A., Nutt, W (1992). “Terminologische Logiken”. *KI* vol. 6, n. 3. p. 23-3.
- Baader, F., Horrocks, I., Sattler, U. (2002). “Description Logics for the Semantic Web”. *KI* vol. 16, n. 4. p. 57-59.
- Berners-Lee, T.B., Hendler, J., Lassila, O. (2001). “The Semantic Web”. Retrieved July 20, 2008, from https://www-sop.inria.fr/acacia/cours/essi2006/Scientific%20American_%20Feature%20Article_%20The%20Semantic%20Web_%20May%202001.pdf.
- Bernus, P., Nemes, L., & Williams, T. J. (1996). “Architectures for enterprise integration”. London: Chapman & Hall.
- Booch, G. (1993). *Object-oriented Analysis and Design with Applications*, 2nd ed., Redwood: Benjamin Cummings.
- Bosak, R., Richard, F. Clippinger, R. F., Dobbs, C., Goldfinger, R., Jasper, R. B., Keating, W., Kendrick, G., Sammet, J. E. (1962). “An information algebra: phase 1 report - language structure group of the CODASYL development committee”. *Communications of the ACM*, vol. 5, n. 4, p.190-204.

- Brochaussen, M., Almeida, M.B., Slaughter, L. (2013). Towards a formal representation of document acts and the resulting legal entities. In: Ingthorsson, R.D., Svennerlind, C., and Almäng J. (Ed.). *Johanssonian Investigations*. Ontos: Frankfurt, p. 120-139.
- Chen, P. (1976). "The entity-relationship model: towards a unified view of data". *ACM Transactions on Database Systems*, vol. 5 , n. 1. p. 9-36.
- Cohen, S. M. (2008). *Aristotle's Metaphysics*. Retrieved February 10, 2012, from <http://plato.stanford.edu/entries/aristotle-metaphysics/>
- Cood, E. F. (1979). "Extending the database relational model to capture more meaning". *ACM Transactions on Database Systems*, vol. 4, n. 4, p. 397-434.
- Côté, R. G. *et al.* (2006) "The Ontology Lookup Service, a lightweight cross-platform tool for controlled vocabulary queries". *BMC bioinformatics*, v. 7, n. 1, p. 97.
- Courtot, M. *et al.* (2011) "MIREOT: The minimum information to reference an external ontology term". *Applied ontology*, v. 6, n. 1, p. 23-33.
- De Nicola, A. Missikoff, M. Navigli, R. (2005). "A proposal for a unified process for ontology building: UPON". In: K.V., A.J., D.R., W., *Database and Expert Systems Applications. DEXA 2005*, Springer, Berlin, Heidelberg. p.655-664.
- Falbo, R. D. A. (2014). "SABiO: Systematic Approach for Building Ontologies". In: Guizzardi, G. *et al*, *ONTO-COM-ODISE 2014- Ontologies in Conceptual Modeling and Information Systems Engineering*, Rio de Janeiro, Brazil. *CEUR Workshop Proceedings*.
- Farinelli, F. (2017). "Improving semantic interoperability in the obstetric and neonatal domain through an approach based on ontological realism". 256 p. Doctoral (Doctor in Information Science). School of Information Science Federal University of Minas Gerais at Brazil, Belo Horizonte. *Publicado em English*.
- Farinelli, F., Elkin, P. L. (2017). "Construção de ontologia na prática: um estudo de caso aplicado ao domínio obstétrico". *Ciência da Informação*, v. 46, n. 1.
- Fernández-López, M. Gómez-Pérez, A. Juristo, N. (1997). "Methontology: from ontological art towards ontological engineering". *Proceedings of the Ontological Engineering AAAI-97 Spring Symposium Series*, Stanford University, EEUU. American Association for Artificial Intelligence.
- Fillion, E., Menzel, C., Blinn, T., Mayer, R. (1995). "An ontology-based environment for enterprise model integration". *Proceedings of the Workshop on Basic Ontological Issues in Knowledge Sharing at IJCAI95*. Montreal: AAAI Press (pp. 33-45).
- Frigg, R. (2006). *Models in Science*. Retrieved July 18, 2008, from <http://plato.stanford.edu/entries/models-science/>
- Fonseca, F. (2007). "The Double Role of Ontologies in Information Science Research". *Journal of the American Society for Information Science and Technology*. vol. 58, n. 6, p. 786–793.
- Fox, M. S. (1992). "The TOVE Project: towards a common-sense model of the enterprise". In F. Belli, & F. J. Radermacher (Eds.), *Proceedings of 5th International Conference Industrial and Engineering Applications of Artificial Intelligence and Expert Systems*. pp. 25-34. London: Springer-Verlag.

- Gandon, F. (2002). Distributed artificial intelligence and knowledge management: ontologies and multi-agent systems for a corporate semantic web. PhD Thesis, INRIA and University of Nice, Nice, FR, School of Sciences and Technologies of Information and Communication.
- Gangemi, A., Guarino, N., Masolo, C., Oltramari, A., Oltramari, R., & Schneider, L. (2002). Sweetening Ontologies with DOLCE. Retrieved January 10, 2010, from <http://www.loa.istc.cnr.it/Papers/DOLCE-EKAW.pdf>
- Genesereth, M. R., Nilsson, L. (1987). Logical foundation of AI. San Francisco: Morgan Kaufman.
- Ghazvinian, A., Noy, N. F., Musen, M. A. (2011) "How orthogonal are the OBO Foundry ontologies?". *Journal of Biomedical Semantics*, vol. 2, n. 2, p. S2.
- Grenon, P., Smith, B., Goldberg, J. (2007). Biodynamic Ontology: Applying BFO in the Biomedical Domain. Retrieved September 20, 2011, from <http://ontology.buffalo.edu/medo/biodynamic.pdf>
- Guarino, N. (1995). Formal ontology, conceptual analysis and knowledge representation. *International Journal of Human-Computer Studies*, vol. 43, n.5-6, p. 625-640.
- Guarino, N. (1998). Formal ontology and information systems. In N. Guarino (Ed.), *Formal ontology in information systems*. pp. 3–15. Amsterdam: IOS Press.
- Guarino, N., Giaretta, P. (1995). “Ontologies and KBs: towards a terminological clarification”. In N. Mars (Ed.), *Towards a Very Large Knowledge Bases: Knowledge Building and Knowledge Sharing*. pp. 25-32. Amsterdam: IOS Press.
- Guizzardi, G. (2005). *Ontological Foundations for Structural Conceptual Models*. PhD Thesis, University of Twente, Twente, NL, Centre for Telematics and Information Technology.
- Gruber, T. (1993). What is an ontology? Retrieved September 14, 2002, from <http://www-ksl.stanford.edu/kst/what-is-an-ontology.html>
- Grüniger, M., Fox, M. S. (1995). “Methodology for the Design and Evaluation of Ontologies”. In: Mellish, C. S., 14th International Joint Conference on Artificial Intelligence (IJCAI-95), Montreal, CA. Morgan Kaufmann.
- Hoehndorf, R., Schofield, P.N., Georgios, G. V. (2015). “The role of ontologies in biological and biomedical research: a functional perspective”. *Briefings in bioinformatics*.vol. 16, n. 6, p. 1069-1080.
- Horridge, M. (2011). A” Practical Guide To Building OWL Ontologies Using Protegé”. Retrieved September 20, 2014, from http://mowl-power.cs.man.ac.uk/protegeowltutorial/resources/ProtegeOWLTutorialP4_v1_3.pdf
- Jacobson, I., Christerson, M., Jonsson, P., Overgaard, G. (1992). *Object-Oriented Software Engineering: A Use Case Driven Approach*. Massachusetts: Addison-Wesley.
- Jardine, D. A. (1976). “The ANSI/SPARC DBMS model”. *Proceedings of the 2nd SHARE Working Conference on Database Management Systems*. Amsterdam: North Holland.
- Lennox, J. (2000). *Aristotle's Philosophy of Biology: Studies in the Origins of Life Science*. Cambridge: Cambridge University Press.

- Lowe, E. J. (2007). *The Four-Category Ontology: A Metaphysical Foundation for Natural Science*. New York: Oxford University Press.
- MacLeod, M.C. and Rubenstein, E.M. (2005). *Universals*. Retrieved March 30, 2010, from <http://www.iep.utm.edu/universa/>
- Milton, S. (2000). *An Ontological Comparison and Evaluation of Data Modelling Frameworks*. PhD Thesis, The University of Tasmania, Hobart, AU, School of Information Systems.
- Miñarro-Giménez, J.A., Kreuzthaler, M., Schulz, S. (2015). "Knowledge Extraction from MEDLINE by Combining Clustering with Natural Language Processing". *Proceedings of the AMIA Annual Symp.*
- Mylopoulos, J. (1992). "Conceptual Modelling and Telos". In P. Loucopoulos, & R. Zicari (Eds.), *Conceptual Modelling, Databases and CASE: An Integrated View of Information Systems Development*. New York: Wiley.
- Mylopoulos, J. (1998). "Information Modeling in the time of revolution". *Informations Systems*, vol. 23, n. 3, p. 127-155.
- Niiniluoto, I. (1999). *Critical Scientific Realism*. New York: Oxford University Press.
- Noy, N. F. *et al.* (2003) "Protege-2000: an open-source ontology-development and knowledge-acquisition environment". *AMIA Symposium*. p.953.
- Noy, N. F. *et al.* (2008) "BioPortal: A Web Repository for Biomedical Ontologies and Data Resources". In: Joshi, C. B. A., *7th International Semantic Web Conference (ISWC2008)*, Karlsruhe, Germany. CEUR-WS.org.
- Ogden, C. K., Richards, I.A (1923). *The Meaning of Meaning*. London: Routledge.
- Peckham, J., Maryanski, F. (1988). "Semantic Data Models". *ACM Computing Surveys*, vol. 20, n. 3, p. 153-189.
- Rumbaugh, J., Blaha, M., Premerlani, W., Eddy, F., Lorensen, W. (1991). *Object-Oriented Modeling and Design*, New York: Prentice Hall.
- Schulz, S., Daumke, P., Stenzhorn, H., Poprat, M. (2009). "Incremental Semantic Enrichment of Narrative Content in Electronic Health Records". *Proceedings of the World Congress on Medical Physics and Biomedical Engineering*.
- Schulz, S. *et al.* (2012). "Guideline on Developing Good Ontologies in the Biomedical Domain with Description Logics". December 11th, 2012, p.85 <<https://www.iph.uni-rostock.de/en/forschung/homepage-goodod/guideline/>>.
- Schulz, S. (2016). "Knowledge acquisition and management for clinical decision-making". *Grand Rounds* September 9th, 2016. Dr. Stefan Schulz, stefan.schulz@medunigraz.at, Medical University of Graz.
- Smith, B., Welty, C. (2001). "Ontology: Towards a New Synthesis". In: B. Smith & C. Welty (Eds.), *Proceedings of the International Conference on Formal Ontology in Information Systems* (pp. 3-9). Ogunquit: ACM Press.
- Smith, B. (1998). "The Basic Tools of Formal Ontology". In: N. Guarino (Ed.), *Proceedings of Formal Ontology in Information Systems* (pp. 3-15). Amsterdam: IOS Press.

- Smith, B. (2003). *Ontology and Information Systems*. Retrieved March 20, 2005, from [http://www.ontology.buffalo.edu/ontology\(PIC\).pdf](http://www.ontology.buffalo.edu/ontology(PIC).pdf)
- Smith et al. (2007). The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. Retrieved May 22, 2009, from <http://www.nature.com/nbt/journal/v25/n11/full/nbt1346.html>.
- Smith, B. (1997). "On Substances, Accidents and Universals: In Defence of a Constituent Ontology". *Philosophical Papers*, vol. 26, p. 105–127.
- Smith, B. (2003). *Ontology and Information Systems*. Retrieved March 20, 2005, from [http://www.ontology.buffalo.edu/ontology\(PIC\).pdf](http://www.ontology.buffalo.edu/ontology(PIC).pdf)
- Smith, B. *et al.* (2015) "Basic Formal Ontology 2.0: Specification and User's Guide" June 26th, 2015, p.97 <<https://github.com/BFO-ontology/BFO/raw/master/docs/bfo2-reference/BFO2-Reference.pdf>>.
- Smith, B. *et al.* (2007). "The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration". *Nature Biotechnology*, v. 25, n. 11, p. 1251-1255.
- Smith, B., Ceusters, W. (2010) "Ontological realism: A methodology for coordinated evolution of scientific ontologies". *Applied ontology*, v. 5, n. 3-4, p. 139-188, November 15th, 2010.
- Smith, B. *et al.* (2005). "Relations in biomedical ontologies". *Genome Biology*, vol. 6, n. 5.
- Spear, A. D. (2006). *Ontology for the Twenty First Century: An Introduction with Recommendations*. Retrieved March 20, 2010, from https://www.researchgate.net/publication/238687379_Ontology_for_the_Twenty_First_Century_An_Introduction_with_Recommendations
- Staab, S., Studer, R. (2004). *Handbook on Ontologies*. Berlin: Springer.
- Studtmann, P. (2008). *The foundations of Aristotle's categorial scheme*. Milwaukee: Marquette University Press.
- Suárez-Figueroa, M.Gómez-Pérez, A.Villazón-Terrazas, B. How to Write and Use the Ontology Requirements Specification Document. In: Meersman, R.Dillon, T.Herrero, P. (Ed.). *On the Move to Meaningful Internet Systems: OTM 2009: Springer Berlin Heidelberg*, v.5871, 2009. cap. 16, p.966-982. (Lecture Notes in Computer Science).
- Suárez-Figueroa, M. C. (2010) "NeOn Methodology for building ontology networks: specification, scheduling and reuse". 268 p. (Doctoral thesis). *Inteligência Artificial*, Universidad Politécnica de Madrid, Madrid. *Publicado em English*.
- Sutcliffe, J. P. (1993). "Concept, class, and category in the tradition of Aristotle". In: I. van Mechelen, J. Hampton J, R.S. Michalski,&P. Theuns P (Eds). *Categories and Concepts*. pp. 35-65. London: Academic Press.
- Thomason, A. (2009). *Categories*. Retrieved March 3, 2011, from <http://plato.stanford.edu/entries/categories/>
- Uschold, M., King, M., Moralee, S., Zorgios, Y. (1998). "Enterprise ontology". *The Knowledge Engineering Review*, vol. 13, p. 1-69.
- Vickery, B. C. (1997). "Ontologies". *Journal of Information Science*. 23 (4), 227-286.

- Wand, Y., Weber, R. (1990). "Mario Bunge's ontology as a formal foundation for information systems concepts". In: P. Weingartner & J.W.G. Dorn (Eds.), *Studies on Mario Bunge's treatise*. Amsterdam: Rodopi.
- Wand, Y., Storey, V. C., Weber, R. (1999). "An ontological analysis of the relationship construct in conceptual modeling". *ACM Transactions on Database Systems*, vol. 24, n. 4, p. 494-528.
- W3c. (2004a). "OWL Web Ontology Language Guide". Smith, M. K. Welty, C. McGuinness, D. L.: World Wide Web Consortium OWL Working Group. 2004
<<https://www.w3.org/TR/2004/REC-owl-guide-20040210/>>,
- _____. (2004b) "OWL Web Ontology Language Reference". Bechhofer, S. *et al*: World Wide Web Consortium OWL Working Group. 2017 <<http://www.w3.org/TR/2004/REC-owl-ref-20040210/>>,
- Xiang, Z. *et al*. (2010). "OntoFox: web-based support for ontology reuse". *BMC research notes*, vol. 3, n. 1, p. 175.
- Xiang, Z. *et al*. (2011). "Ontobee: A Linked Data Server and Browser for Ontology Terms, Ontology". 2nd International Conference on Biomedical Ontology (ICBO2011), Buffalo, NY, USA. CEUR-WS.org, 2011. p.279-281.
- Young, J. W., Kent, H. K. (1958). "Abstract formulation of data processing problems". *The Journal of Industrial Engineering*, vol. 9, n. 6, p. 471-479.