

XIII Encontro Nacional de Pesquisa em Ciência da Informação - XIII ENANCIB 2012
GT 2: Organização e Representação do Conhecimento
Modalidade de apresentação: Comunicação Oral

MODELOS E TEORIAS PARA REPRESENTAÇÃO: UMA TEORIA ONTOLÓGICA SOBRE O SANGUE HUMANO

Fabricio M. Mendonça
PPGCI-ECI/UFMG

Mauricio B. Almeida
PPGCI-ECI/UFMG

Resumo

O conhecimento científico é composto por conceitos e declarações que os cientistas consideram úteis para os propósitos da ciência. O conhecimento científico provê formas de organizar e de categorizar o mundo, predições para eventos futuros, explicações para eventos passados e compreensão sobre as causas dos eventos. Modelos e teorias têm papel fundamental na construção do conhecimento científico. Os dois termos são muitas vezes tratados como sinônimos, mas existem diferenças entre eles: modelos são tanto representações de uma parte específica da realidade quanto interpretações de teorias, as quais, por sua vez, podem ser entendidas como um conjunto de declarações, logicamente organizadas, com poder explanatório sobre eventos genéricos. O presente artigo elucidada o significado e a estreita inter-relação entre os dois termos no âmbito da ciência, concluindo sobre sua importância no contexto da Ciência da Informação. Como aplicação prática, apresentam-se resultados parciais da construção de uma teoria ontológica sobre o sangue humano, conduzida no contexto de um projeto sobre organização da informação biomédica.

Abstract

Scientific knowledge consists of concepts and statements that scientists consider useful for scientific purposes. Scientific knowledge provides ways to organize and categorize the world, predictions of future events, explanations for past events and understanding about the causes of the events. Models and theories play an important role in the construction of scientific knowledge. Both terms are often considered synonyms, but there are differences between them: models are both representations of a specific part of reality and interpretations of theories; theories may be understood as a set of statements, logically organized, with explanatory power over general events. This paper elucidates the meaning and the close interrelationship between the two terms within the scope of science, concluding that they are important in the context of Information Science. As a practical application, we present the partial results of the construction of an ontological theory for human blood, conducted within the scope of a biomedical information organization project.

1) Introdução

A evolução do conhecimento científico tem sido, desde há muito, tema de discussão em diversas áreas do conhecimento, tais como a Filosofia, a Física, a Lógica, a Matemática, dentre outras. Especificamente, a Filosofia da Ciência é uma das áreas que busca explicar como se dá os avanços científicos e apontar o caminho que a ciência deve trilhar para alcançar sucesso.

Uma razão que explica o sucesso da ciência é o fato de que os cientistas utilizam-se de aproximações da realidade como forma de obter uma compreensão dos fenômenos à sua volta. Cientistas, via de regra, se satisfazem com uma compreensão aproximada da natureza, descrevendo grupos selecionados de fenômenos e negligenciando outros menos relevantes. Assim, é possível explicar muitos fenômenos em termos de poucos e, conseqüentemente, alcançar a compreensão de diferentes aspectos da natureza (CAPRA, 1983).

Em geral, a ciência faz uso de *modelos* e *teorias* como instrumentos para alcançar essa compreensão aproximada da realidade do mundo. Apesar dos dois termos serem muitas vezes usados, como sinônimos, existem diferenças entre eles: os modelos são representações de uma parte específica da realidade e também interpretações ou realizações de uma teoria; enquanto as teorias consistem de um conjunto de declarações, logicamente organizadas, com poder explanatório sobre eventos genéricos. De fato, a construção de modelos e teorias é um indicativo da maturidade acadêmica de uma disciplina. Exemplos de modelos e teorias são também comuns em Ciência da Informação (PETTIGREW e MCKECHIE, 2001).

O presente artigo discorre sobre modelos e teorias como entidades da prática científica, concluindo sobre sua importância no contexto da Ciência da Informação. Para tal, discorre sobre o significado dos dois termos, explicando a inter-relação entre eles e apresentando aplicações práticas de modelos e teorias no âmbito dos sistemas de informação. Especificamente, apresentam-se aspectos metodológicos e resultados parciais da construção de uma teoria ontológica sobre o sangue humano, conduzida no contexto de um projeto sobre organização da informação biomédica. Cabe ainda citar que, visto a amplitude do tema, o presente artigo não pretende ser exaustivo. Ao invés disso, apresenta em uma perspectiva histórica e ilustrativa os principais desenvolvimentos sobre o tema.

O presente artigo está organizado da seguinte forma: a seção 2 discorre sobre o conceito de modelos, destacando seu papel em ciência e apresentando alguns exemplos; a seção 3 apresenta uma visão sobre teorias científicas, introduzindo diferentes abordagens na ciência sobre o tema; a seção 4 apresenta, através de exemplos práticos, o desenvolvimento da teoria ontológica sobre o sangue humano parte da pesquisa em andamento. Finalmente, a seção 5, traz considerações finais e conclui sobre a importância de modelos e teorias no âmbito da Ciência da Informação.

2) Os modelos

Modelos são entidades fundamentais na prática científica e se prestam a investigar fenômenos complexos do mundo sobre os quais não se tem entendimento completo. Esta

seção se dedica a explicar modelos do ponto de vista científico e está dividida em duas partes: a seção 2.1 explica o que são modelos a partir de perspectivas diversas e a seção 2.2 oferece uma breve visão sobre o uso de modelos no âmbito das práticas científicas.

2.1) O que são modelos

A abordagem a modelos científicos na literatura é abrangente e não consensual, de forma que não é tarefa simples definir o conceito de modelos em ciência. A diversidade de abordagens pode ser ilustrada pelas diversas tentativas em estabelecer um status ontológico para modelos (FRIGG, 2006): (i) modelos são *objetos físicos*, como por exemplo, modelos de madeira de aviões, barcos e pontes ou o modelo de DNA; tais modelos são também conhecidos por *modelos materiais* ou *modelos icônicos* (SUPPE, 1977); (ii) modelos são *objetos ficcionais (não materiais)* como, por exemplo, o modelo de átomo ou o modelo de um pêndulo sem atrito; (iii) modelos são *estruturas da teoria de conjuntos*, que correspondem aos chamados *modelos matemáticos*; (iv) modelos são *descrições*, que se referem às descrições de sistemas a representar encontradas, por exemplo, em obras científicas (livros, artigos, etc.); e (v) modelos são *equações*, que se caracterizam como um tipo de *modelo matemático* usado em domínios específicos, como por exemplo em Economia (modelos do mercado de ações e modelo de economia aberta).

A falta de consenso pode ser ilustrada pelo fato de que nem todas essas entidades são realmente consideradas modelos, principalmente, as *descrições*, as *estruturas da teoria de conjuntos* ou as *equações*. Além de serem estruturas estritamente sintáticas, não parece possível explicar como esses modelos são construídos e nem como funcionam no contexto de investigação (DOWNES, 1992; MORRISON, 1999).

Uma possibilidade de classificar modelos com vistas a um melhor entendimento é organizá-los de acordo com suas funções. É possível defini-los, a partir de sua compreensão como instrumentos de representação do conhecimento científico, em duas visões: (i) *modelos representativos*, que são representações de uma parte específica da realidade ou do mundo; e (ii) *modelos teóricos*, que representam uma teoria e que interpretam as leis e os axiomas desta teoria.

Na primeira visão, os modelos são esquemas abstratos da realidade construídos de forma a selecionar os fenômenos de maior relevância para o problema do objeto de investigação. Tal medida possibilita capturar as características essenciais de um domínio de conhecimento.

Exemplos da visão de modelos representativos são encontrados, por exemplo, em Chorley e Haggett (1975), e em Sayão (2001). Chorley e Haggett (1975) definem modelo como uma estruturação simplificada da realidade que apresenta características ou relações sob forma generalizada. Nesse sentido, os modelos não incluem todas as observações, mensurações e medições associadas, mas, como tais, são valiosas por ocultarem detalhes secundários e permitirem o aparecimento dos aspectos fundamentais da realidade. Sayão (2001) acrescenta que modelo é uma criação cultural, um “mentefato”, destinado a representar uma realidade, ou alguns de seus aspectos, a fim de torná-los descritíveis qualitativa e quantitativamente e, algumas vezes, observáveis. Modelos conceituais construídos para o desenvolvimento de sistemas de informação são exemplos desta visão.

Na visão dos modelos teóricos, modelos são definidos como interpretações ou realizações de uma teoria ou lei geral sobre um determinado fenômeno, não se caracterizando, necessariamente, como uma representação de uma parte da realidade. Nesse sentido, objetos matemáticos, como cubo ou triângulo, são modelos teóricos resultantes da teoria da geometria Euclidiana, isto porque tais objetos são realizações dessa teoria: “dois pontos podem ser ligados através de uma linha reta” (FRIGG, 2006).

A visão dos modelos teóricos é a visão adotada na lógica moderna, que compreende o modelo como uma estrutura que torna todas as sentenças de uma teoria verdadeira. Uma teoria, por sua vez, é um conjunto de sentenças em uma linguagem formal (HODGES, 1997). Nessa perspectiva, um modelo teórico, é uma estrutura S , descrita por $S = \langle U, O, R \rangle$, onde U é um conjunto não vazio de indivíduos chamados de domínio ou universo de S ; O é um conjunto de operações em U , que pode ou não ser vazio; e R é um conjunto não vazio de relações em U .

Elucidadas estas as duas principais concepções sobre modelos, discute-se a seguir o papel dos modelos na ciência com o propósito de clarificar os conceitos aqui apresentados.

2.2) Os modelos no âmbito da ciência

Discorrer sobre o papel dos modelos na constituição do conhecimento científico envolve considerar diferentes posicionamentos filosóficos e epistemológicos. Dentre as principais questões, destacam-se: (i) aquela sobre o valor dos modelos para a prática científica, que envolve as visões sintática e semântica; (ii) aquela sobre a capacidade dos modelos em expressar uma relação verdadeira com a realidade, que envolve posições realistas e anti-realistas.

No que se refere ao valor dos modelos para a prática científica, dois posicionamentos antagônicos se destacam: (i) a *visão sintática das teorias* (CARNAP, 1938; HEMPEL, 1965), para a qual o modelo científico é considerado uma interpretação alternativa de certa teoria e, assim, é visto como irrelevante para a ciência; e (ii) a *visão semântica de teorias* (VAN FRAASSEN, 1980; GIÈRE, 1988; SUPPES, 2002), que concebe os modelos como componentes essenciais e integrais das teorias, possibilitando testá-las e fornecer suas explicações.

A *visão sintática das teorias* é parte da lógica positivista e assim enxerga os modelos como acréscimos supérfluos às teorias, como uma forma de conceber valor pedagógico, estético e psicológico a elas (CARNAP, 1938; HEMPEL, 1965). Nessa visão, modelos são apenas sistemas de regras que interpretam um cálculo abstrato ou, num sentido mais restrito, uma interpretação alternativa de um cálculo específico. Assim, por exemplo, podemos tomar a matemática utilizada na teoria cinética dos gases e reinterpretá-la de forma que os termos usados nos cálculos se refiram a bolas de bilhar (FRIGG, 2006).

Por outro lado, a *visão semântica de teorias* sugere abolir o cálculo formal e enxergar uma teoria como uma família de modelos (SUPPE, 1977; VAN FRAASSEN, 1980; GIÈRE, 1988; SUPPES, 2002). “Abolir o cálculo formal” significa dizer que as teorias científicas não são simplesmente conjuntos de axiomas, como sugerido na abordagem sintática. Ao referirem-se à visão sintática de teorias, adeptos da abordagem semântica, como Suppe (1977) e van Fraassen (1980), afirmam que abordagem sintática é um modo artificial de análise das teorias científicas e que a abordagem semântica é uma concepção superior, por causa de sua proximidade da prática científica.

Apesar das diferentes visões sobre o valor dos modelos na ciência, geralmente se aceita que modelos são unidades centrais na criação de teorias científicas e, portanto, são bastante úteis para aquisição do conhecimento (LE MOIGNE, 1977).

Sobre a questão da capacidade dos modelos científicos em expressar ou não uma relação verdadeira com o mundo real, mesmo que de forma aproximada, identificam-se duas correntes filosóficas que discutem esse tema: (i) o realismo e (ii) o anti-realismo.

O realismo científico (POPPER, 2008; BUNGE, 1974; CHURCHLAND, 1985) defende o argumento de que os modelos possuam valor ontológico e, nesse sentido, devem representar, aproximadamente e provisoriamente, a realidade. Na ótica realista, as entidades reais (como uma montanha, um rio, uma árvore) existem autonomamente, isto é, independente da existência humana, e por isso, são descobertas ao longo do tempo pelo homem e figuram como constituintes teóricos das explicações científicas. Sob essa visão, o

papel da modelização científica é a construção de modelos em direção à verdade, buscando captar, numa maior aproximação possível, “o que realmente existe” no mundo.

O anti-realismo – advogado por exemplo por Cartwright (1983), Van-Fraassen (1989) e Putnam (1992) – não julga necessário o compromisso com entidades do ponto de vista ontológico. Sob essa perspectiva, as hipóteses sobre o mundo são construções mentais, que se impõem não por seu caráter referencial, mas em função de sua capacidade explicativa (SILVA, 1998). Para os anti-realistas a verdade não é o principal objetivo da modelagem científica de forma que é aceito o critério da falsidade na construção de modelos. Cartwright (1983), por exemplo, apresenta vários estudos de caso ilustrando bons modelos que são falsos.

Apesar de adotarem visões antagônicas, realistas e anti-realistas concordam em um ponto sobre a questão da explicação científica: as teorias e os modelos são capazes de explicar de maneira bem sucedida os fenômenos. Nesse sentido, modelos e teorias são, além de instrumentos de descrição, instrumentos de explicação científica e assim fundamentais para aquisição do conhecimento científico sobre o mundo.

3) As teorias

Mesmo que o termo “teoria” seja popularmente usado para expressar algo que ainda não obteve a sucesso do ponto de vista prático, teorias são portadoras de alto poder explicativo sobre conjuntos de fenômenos gerais. A presente seção discorre sobre as teorias e sua importância no escopo da ciência: a seção 3.1 apresenta concepções básicas sobre o conceito de teoria e a seção 3.2 aborda, através de uma perspectiva histórica, o papel das teorias na ciência.

3.1) O que são teorias

De forma similar a que ocorre com os modelos científicos, a abordagem a teorias científicas não é consensual e gera discussões filosóficas sobre sua definição e seu papel na aquisição do conhecimento científico. Para melhor entender o que são teorias, apresentam-se três concepções básicas de teorias científicas, de acordo com Reynolds (2007): (i) a *teoria na forma de um conjunto de leis*, (ii) a *teoria na forma axiomática* e (iii) a *teoria na forma de processo causal*.

Na primeira concepção – a *teoria na forma de um conjunto de leis* – aceita-se como teoria científica apenas aquela que inclua declarações que podem ser consideradas leis. Leis são conjuntos de generalizações empíricas bem fundamentadas. Como leis são diretamente

suportadas por pesquisa empírica, todos os conceitos usados em leis devem ter definições que permitam sua identificação em situações concretas.

A segunda concepção – *a teoria na forma axiomática* – consiste de um conjunto inter-relacionado de *definições* (incluindo conceitos teóricos) e de *declarações* de dois tipos: (i) *existenciais*, que fazem afirmações sobre a existência e descrevem situações nas quais a teoria pode ser aplicada; e (ii) *relacionais*, que descrevem o relacionamento entre dois conceitos. As declarações relacionais englobam os *axiomas*, conjuntos de declarações das quais outras declarações na teoria podem ser derivadas, e as *proposições*, que são todas as outras declarações na teoria, derivadas de combinações de axiomas e proposições.

Essa abordagem a teorias foi adotada por número significativo de pesquisadores, como por exemplo, Carnap (1938), Reichenbach (1938), Popper (1980), dentre outros, que buscavam definir teorias científicas a partir da lógica. Tal corrente filosófica definia teorias científicas como teorias axiomáticas formuladas na linguagem da lógica matemática L , satisfazendo às condições de: (i) *logicismo clássico*: teoria formulada na lógica de primeira ordem; (ii) *distinção teórico-observacional*: um vocabulário lógico (constantes lógicas), um vocabulário observacional V_0 (termos de observação) e vocabulário teórico V_t (termos teóricos); (iii) *fisicalismo*: os termos de V_0 interpretados como se referindo a objetos físicos ou atributos destes; (iv) *leis teóricas*: conjunto de postulados teóricos T ; e (v) *regras de correspondência*, que especificam as aplicações admissíveis de T a fenômenos empíricos.

A terceira e última concepção de teoria – *a teoria na forma de processo causal* – faz parte da visão semântica de teorias e, além de fornecer explicação e predição científicas, como as duas anteriores, também provê o senso de entendimento. A teoria na forma de processo causal é um conjunto inter-relacionado de *definições* e *declarações* de dois tipos: (i) *existenciais*, que descrevem situações nas quais se espera que um ou mais processos causais ocorram; e (ii) *causais*, que descrevem um ou mais processos causais ou um mecanismo causal que identifica o efeito de variáveis independentes sobre variáveis dependentes.

A principal diferença entre a teoria na forma causal e a teoria na forma axiomática é que na primeira todas as declarações são consideradas de igual importância, não sendo classificadas em axiomas e proposições. Um conjunto de declarações na forma axiomática pode ser reorganizado em uma descrição de um processo causal (REYNOLDS, 2007).

3.2) As teorias no âmbito da Ciência

A abordagem a teorias científicas é discutida, principalmente, no âmbito da filosofia da ciência, que busca compreender e explicar os avanços científicos, além de mostrar o

melhor caminho para que a ciência possa evoluir. Nesta seção, adota-se uma visão histórica para descrever o papel recente das teorias na ciência.

No final do século XIX assistiu uma crise na filosofia da ciência devido ao surgimento de importantes teorias científicas, como a teoria da relatividade e a teoria quântica, as quais não podiam mais ser incorporadas pelas teorias filosóficas vigentes. No início do século XX, Carnap e Schlick estabeleceram um influente grupo de discussão sobre o conhecimento e os avanços científicos, fundamentalmente, baseados no positivismo lógico, que ficou conhecido como *Círculo de Viena* (PESSOA JR., 1993). Esse grupo de filósofos fundou as bases para a criação da *visão ortodoxa*, que buscava explicar as teorias científicas a partir da lógica e de uma postura empirista e positivista. A visão ortodoxa consolidou-se como uma filosofia ciência na década de 50.

Essa consolidação não significou consenso entre os filósofos da época. Duras críticas foram desferidas contra a visão ortodoxa, como por exemplo por Karl Popper. Popper (2008) rejeita o critério verificacionista de significância cognitiva, defendendo a idéia de que teorias científicas não podem ser verificadas por qualquer acúmulo de evidência observacional. Ainda, sustenta que o problema central da filosofia da ciência é o crescimento do conhecimento científico, e isso não poderia ser reduzido a um estudo de linguagens artificiais, como sugerido pela visão ortodoxa.

Uma abordagem alternativa à visão ortodoxa surgiu no final da década de 50 e viria a se constituir em uma nova filosofia da ciência, que não se limitava apenas aos aspectos lógicos da ciência (PESSOA JR., 1993). Essa abordagem ficou conhecida como *teoria dos globalistas*, da qual faziam parte Willard Quine, Paul Feyerabend, Norwood Hanson, Stephen Toulmin e Thomas Kuhn. Alguns pontos ressaltados pelos globalistas (BACHELARD, 1934; KUHN, 1962) são: (i) rejeita-se que uma teoria científica se assenta em bases sólidas fornecidas por dados observacionais; (ii) o contexto social e histórico tornam-se relevantes para entender porque uma teoria é preferida em relação a outra; (iii) a transição de uma teoria para outra não é mais vista como uma ampliação cumulativa de conhecimento, mas como uma ruptura.

Nos anos 60, Thomas Kuhn é o principal ícone dentre os teóricos globalistas e sua obra *A Estrutura das Revoluções Científicas* ressalta o fato de que a transição entre teorias se dá através de revoluções, e que entre estes períodos de transição tem-se a chamada ciência normal. Na década de 70, outros autores como Lakatos (1970), Feyerabend (1975) e Suppes (2002), apresentam suas metateorias debatendo as idéias de Kuhn e dos globalistas. O

panorama da filosofia da ciência na década de 80 e 90 caracteriza-se por uma pluralidade de abordagens, enraizadas tanto nos globalistas quanto na visão ortodoxa (PESSOA JR., 1993).

4) Aplicações de modelos e teorias: sistemas de informação

A capacidade explanatória dos modelos e das teorias é útil na construção de conhecimento científico. Em um nível mais específico, uma aplicação evidente de modelos e teorias é o desenvolvimento de sistemas de informação: primeiro é preciso entender os fenômenos de interesse, para depois representá-los de uma forma que se possa tirar vantagem de seu caráter explicativo.

A presente seção versa sobre a aplicação de modelos e teorias no âmbito dos sistemas de informação automatizados: a seção 4.1 narra evolução dos modelos para sistemas de informação desde os modelos de dados nos anos 1960 até os modelos baseados em ontologias nos anos 2000; a seção 4.2 apresenta, como aplicação prática de teorias em sistemas de informação, os desenvolvimentos preliminares de uma teoria ontológica sobre o domínio da hematologia e hemoterapia.

4.1) Dos modelos de dados às teorias ontológicas

Modelos de dados são criados nas organizações com o objetivo de representar o que deve ser codificado e processado em sistemas de informação. Os sistemas de informação organizacionais têm papel relevante na consolidação de práticas administrativas relacionadas às necessidades das pessoas no desempenho de suas funções. O desenvolvimento de sistemas de informação envolve a criação de modelos para representar atividades que tem lugar nos mais diversos tipos de organizações. Um modelo de dados organizacionais é uma representação explícita da estrutura, atividades, processos, fluxos, metas e restrições de uma organização (GANDON, 2002).

No desenvolvimento de sistemas de informação, modelagem conceitual é o estágio em que os modelos são criados para representar a compreensão humana. Modelos conceituais são criados a partir de abstrações da realidade de acordo com a perspectiva de um indivíduo ou grupo de indivíduos. Abstrações são meios de especificar entidades e as relações entre entidades dentro do domínio de um campo do conhecimento que é de interesse para os usuários do sistema.

A modelagem conceitual como ela é hoje conhecida é o resultado de pesquisa conduzida nos últimos 50 anos. As primeiras iniciativas para a especificação de modelos de dados datam do final dos anos 1950 (BOSAK et al., 1962; YOUNG e KENT, 1958). Tais

iniciativas foram conduzidas de forma a criar modelos que atendessem aos requisitos de estruturas de dados computacionais.

Nos anos de 1960, a pesquisa em bancos de dados gerou três tipos principais de modelos de dados: o modelo hierárquico, o modelo em rede e o modelo relacional. Esses modelos são conhecidos como modelos lógicos, uma vez que eles não se referem a aspectos físicos da implementação do sistema. Entretanto, modelos lógicos causaram muitos problemas o que acabou limitando sua utilização na modelagem conceitual (MYLOPOULOS, 1998).

Os primeiros modelos semânticos usados em modelagem conceitual apareceram nos anos 1970 no contexto do trabalho do comitê ANSI/X3/SPARC para padronização dos sistemas de gerenciamento de bancos de dados. Dentre estes, os mais conhecidos são o modelo de dados semântico (ABRIAL, 1974), a arquitetura em três camadas (JARDINE, 1976), o modelo entidade relacionamento (CODD, 1979), dentre outros. As principais características dos modelos semânticos, em comparação com os anteriores, é a facilidade de entendimento. O modelo entidade relacionamento, por exemplo, fornece termos adicionais para uso como primitivas de modelagem.

Nos anos 1990, propostas para modelagem orientada a objetos se tornaram populares. Os modelos orientados a objetos têm características adicionais em relação aos modelos de dados, mas também mantém similaridades (MILTON, 2000). A Linguagem de Modelagem Conceitual (popularmente conhecida como UML) foi uma tentativa de padronizar notações orientadas à objeto que reuniu outras iniciativas: o método de 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 outras.

Ao longo dos anos a criação de modelos conceituais tem sido motivada pela busca por melhores formas de representar a realidade. De fato, a modelagem conceitual é a atividade de descrever formalmente alguns aspectos do mundo físico e social com propósito de compreensão e comunicação (MYLOPOULOS, 1992). De acordo com Smith e Welty (2001), a inconsistência na modelagem durante os primeiros anos da modelagem conceitual tem sido a causa dos atuais problemas de interoperabilidade entre sistemas. Uma alternativa para esse tipo de problema são os modelos baseados em ontologias: “apenas um ontologia de referência comum e robusta – uma taxonomia de entidades compartilhada – pode prover vantagens significativas sobre a modelagem previamente utilizada, caracterizada por ser *ad-hoc* e orientada por casos” (SMITH e WELTY, 2001, p.4).

Para criar ontologias, é preciso avaliar fatos e identificar a estrutura sobre a qual se organizam, fazer generalizações e abstrações, o que é feito quando se constrói uma teoria.

Seguindo essa linha de pensamento, uma ontologia é uma teoria representando os principais fatos e regras que governam certa parte da realidade. O termo ontologia, nesse contexto, é um anacronismo para tal teoria (ALMEIDA, 2012).

4.2) Uma teoria ontológica sobre o sangue humano

Afim de apresentar por meio de exemplos práticos o desenvolvimento de uma teoria, discorre-se nesta seção sobre aspectos metodológicos e resultados parciais da construção de uma teoria ontológica sobre o sangue humano, conduzida no contexto de um projeto sobre organização da informação biomédica: o *Blood Project*. O *Blood Project*¹ vem sendo desenvolvido pela Fundação Hemominas, a Universidade Federal de Minas Gerais (UFMG) e o *Ontology Research Group*, com o propósito de gerar uma linguagem de representação formal sobre o sangue humano (ALMEIDA et al., 2010).

Os desenvolvimentos preliminares sobre a teoria para o sangue humano abrangem a seleção de domínios do conhecimento, aspectos metodológicos específicos e desenvolvimento da teoria em si. Os domínios do conhecimento selecionados são fisiologia do sangue, hemocomponentes e hemoderivados. Os aspectos metodológicos envolvem duas atividades principais: (i) estudo aprofundado no domínio do sangue humano, por meio de pesquisa bibliográfica, incluindo normas técnicas e manuais; (ii) adoção de *ontologias fundamentais* como referência para desenvolvimento da teoria.

Ontologias fundamentais ou de alto nível são ontologias que descrevem conceitos gerais como espaço, tempo, matéria, objeto, evento, ação, etc., os quais são independentes do problema ou domínio (GUARINO, 1998). Geralmente, elas são usadas como base ou ponto de partida para a construção de ontologias sobre domínios específicos. As principais ontologias fundamentais utilizadas para o desenvolvimento da teoria ontológica do sangue são: (i) *Basic Formal Ontology* (BFO)², fornece as classes (universais) e as relações fundamentais para as ontologias do sangue; (ii) *Foundational Model Anatomy* (FMA)³, utilizada como ponto de partida para a compreensão da fisiologia do sangue humano; e (iii) *Information Artifact Ontology* (IAO), usada para representar os parâmetros básicos do sangue.

A teoria em desenvolvimento pode ser axiomatizada ou não, dependendo do artefato de representação que se objetiva alcançar. No caso de uma teoria formal visando inferência lógica automática, a axiomatização é necessária; no caso de teoria informal para construção

¹ Disponível na Internet em <http://mba.eci.ufmg.br/blood/>. Acesso: 04 de Junho de 2012.

² Disponível na Internet em <http://www.ifomis.org/bfo>. Acesso: 04 de Junho de 2012.

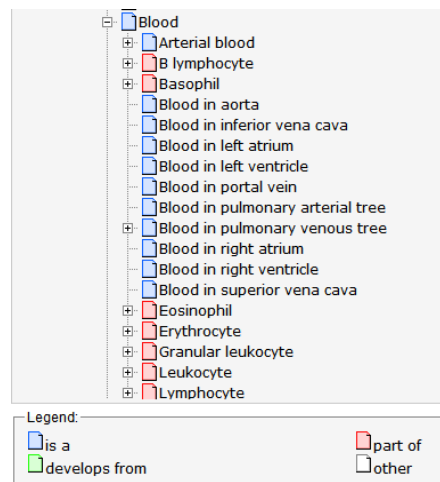
³ Disponível na Internet em <http://sig.biostr.washington.edu/projects/fm/AboutFM.html>. Acesso: 04 de Junho de 2012.

de instrumentos de organização da informação como por exemplo tesouros, a axiomatização não é necessária. Cabe ainda citar que a estrutura de representação, nessa etapa, abrange apenas relações taxonômicas e relações partitivas.

O restante da presente seção se ocupa de apresentar o raciocínio usado para criar a teoria, baseando-se no conhecimento adquirido sobre o domínio e na inter-relação entre os principais conceitos usando relações ontológicas. À medida que a teoria é desenvolvida, informalmente (axiomas não são apresentados nessa versão), a informação é organizada em figuras e outros artefatos gráficos. Os resultados são expressos pelas figuras que representam taxonomias e paronomias. Definições de termos e conceitos importantes podem ser obtidas através do conhecimento compilado.

Para compreensão da fisiologia do sangue humano, foram usados como referência, além da ontologia FMA (ROSSE e MEJINO, 2003) já citada, livros texto e material de treinamento (GREER et al, 2008; BARRETT et al, 2009; VAN DE GRAAFF, RHEES e PALMER, 2010).

Figura 1 - Taxonomia FMA sobre o sangue.

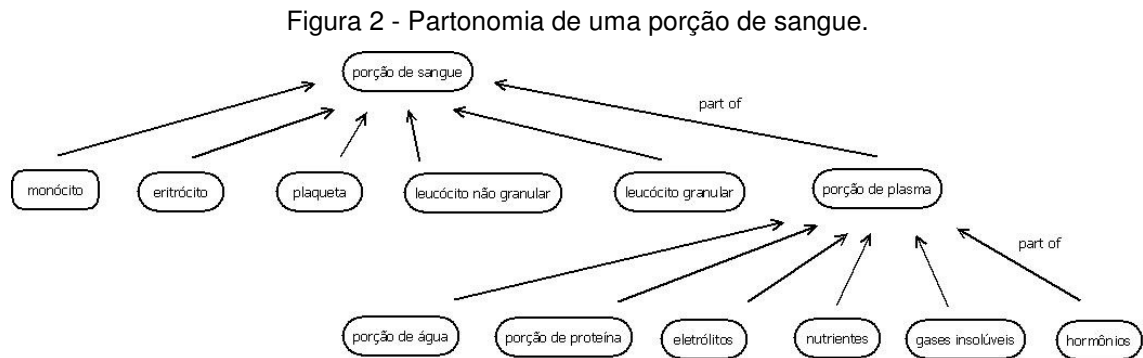


Fonte: Ontology Lookup Service⁴.

Para iniciar a compreensão do assunto e definir “sangue”, buscou-se abaixo da categoria básica FMA chamada *Anatomical Entity*, onde se encontram duas categorias disjuntas, chamadas de *Physical Anatomical Entity* e *Non-Physical Anatomical Entity*. Abaixo de *Material Anatomical Entity* encontra-se a categoria *Body Substance* e, finalmente, encontra-se a categoria *Blood* (veja figura 1). O termo “blood” na FMA é definido como “substância do corpo que formada de plasma e células sanguíneas”. Abaixo da categoria *blood*, a FMA define ainda outros tipos (por exemplo, *Venous Blood*) e partes (por exemplo,

⁴ Disponível na Internet em <http://www.ebi.ac.uk/ontology-lookup/>. Acesso: 04 de Junho de 2012.

Plasma). Nesse momento, foram consultadas outras fontes para obter uma visão mais detalhada da fisiologia do sangue humano. O sangue é o principal líquido do corpo e representa cerca de 8% do peso corporal. Este volume é composto de plasma e elementos formados. O plasma contém água, proteína e outras soluções; elementos formados são constituídos de plaquetas, glóbulos brancos e glóbulos vermelhos. Esses desenvolvimentos resultaram no esquema da figura 2.

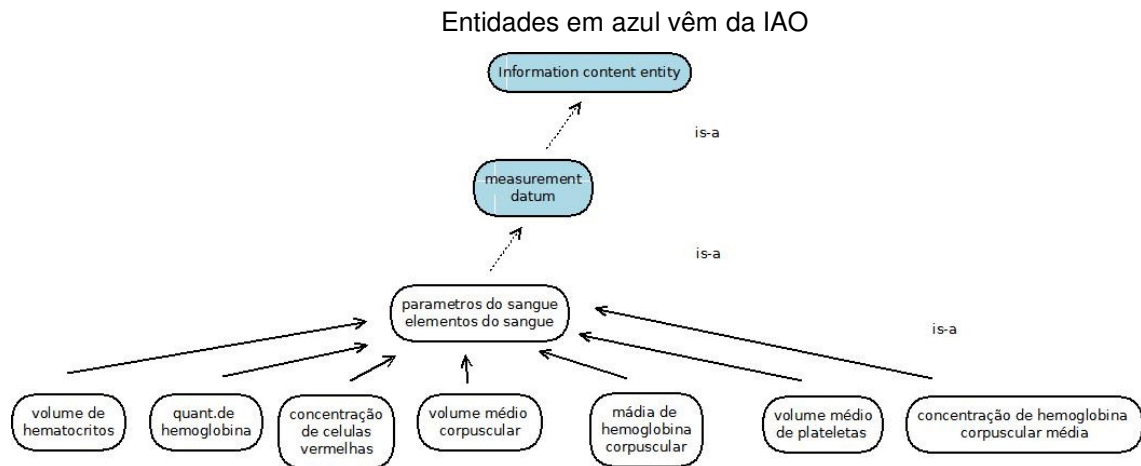


Fonte: elaborado pelo autor.

Exemplos de eletrólitos (Figura 2) podem ser as substâncias químicas: K, Ca²⁺, Mg²⁺, Cl⁻, HCO₃⁻, HPO₄²⁻ e SO₄²⁻; exemplos de proteínas presentes no plasma são: albumina, globulina, fator H, Icg total, transferrina, fibrinogênio, para citar apenas algumas; exemplos de nutrientes são: glicose, aminoácidos, lipídios, colesterol e vitaminas; exemplos de gases dissolvidos são: dióxido de carbono, oxigênio e nitrogênio; exemplos de produtos não aproveitados são: uréia, ácido úrico, creatinina e bilirrubina.

Como anteriormente citado, a IAO é utilizada para representar parâmetros básicos sobre o sangue humano. A IAO é uma ontologia de entidades de informação, descendente da ramificação DENRIE do *OBI Project* (BRINKMAN et al., 2010). Os parâmetros básicos para análise do sangue são classificados com *measurement data* (dados mensuráveis), definidos como *information content entities* (entidades de conteúdo informacional) que “é um registro da saída de uma medição tal como produzida por um dispositivo” (veja figura 3).

Figura 3 - Parâmetros básicos do sangue como elementos sanguíneos



Fonte: elaborado pelo autor.

Elucidações preliminares, que no futuro resultarão em definições, são fornecidas para as entidades apresentadas. Para os parâmetros básicos tem-se (GREER et al. 2008):

- *Volume de hematócitos*: a proporção do volume de uma amostra de sangue que é ocupada por glóbulos vermelhos.
- *Quantidade de hemoglobina*: contagem de hemoglobina em suas várias formas como oxihemoglobina, carboxihemoglobina, metemoglobina, outros.
- *Concentração de glóbulos vermelhos por volume*: não se aplica;
- *Volume corpuscular*: corresponde ao volume médio dos glóbulos vermelhos do sangue.
- *Hemoglobina corpuscular*: corresponde à medida do teor de hemoglobina médio por glóbulo vermelho.
- *Concentração de hemoglobina corpuscular*: corresponde à concentração média de hemoglobina num certo volume de glóbulo vermelho.
- *Volume de plaquetas*: medida que corresponde a uma relação inversa com o número de plaquetas; relacionada com diversos estados patológicos.

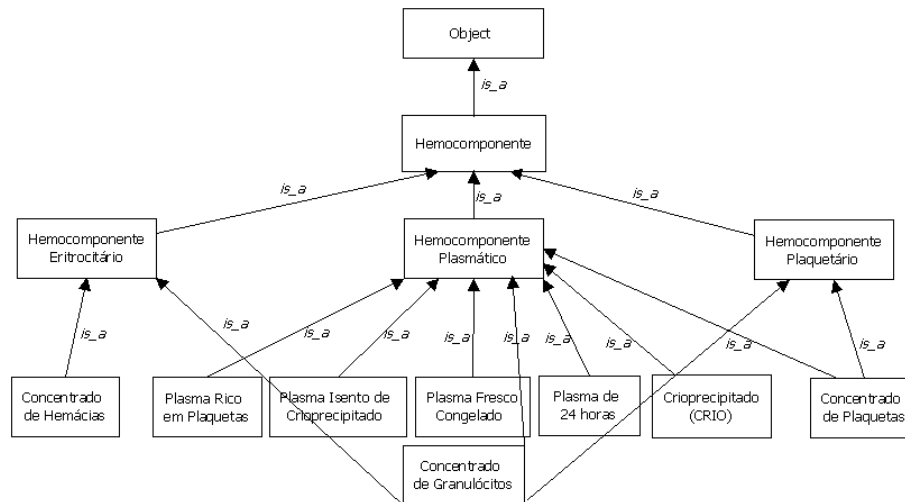
Por fim, tem-se a representação e elucidação dos *hemocomponentes* e *hemoderivados* do sangue humano. Consultou-se aqui o “Guia para o uso de Hemocomponentes” do Ministério da Saúde (BRASIL, 2008) como referência.

Hemocomponentes são obtidos através de processos físicos realizados no *sangue total* (ainda não separado), tais como a *centrifugação* e o *congelamento*. *Hemoderivados* são obtidos em escala industrial, a partir do fracionamento do plasma sanguíneo por processos físico-químicos. Submetendo o sangue total ao processo de *centrifugação* é possível separá-lo

em três camadas de componentes distintos: (i) as *hemácias*, que ficam no fundo da bolsa coletora de sangue, (ii) o *buffy coat* (ou *camada leucoplaquetária*), uma camada acima e (iii) o *plasma*, que fica na camada superior. A partir dessa separação é possível obter os diferentes hemocomponentes do sangue humano.

Tem-se três tipos gerais de hemocomponentes: (i) *eritrocitários*, originados a partir das hemácias; (ii) *plasmáticos*, que advém do plasma e (iii) *plaquetários*, originados do *buffy coat* (vide figura 4).

Figura 4 - Taxonomia dos tipos de hemocomponentes do sangue.



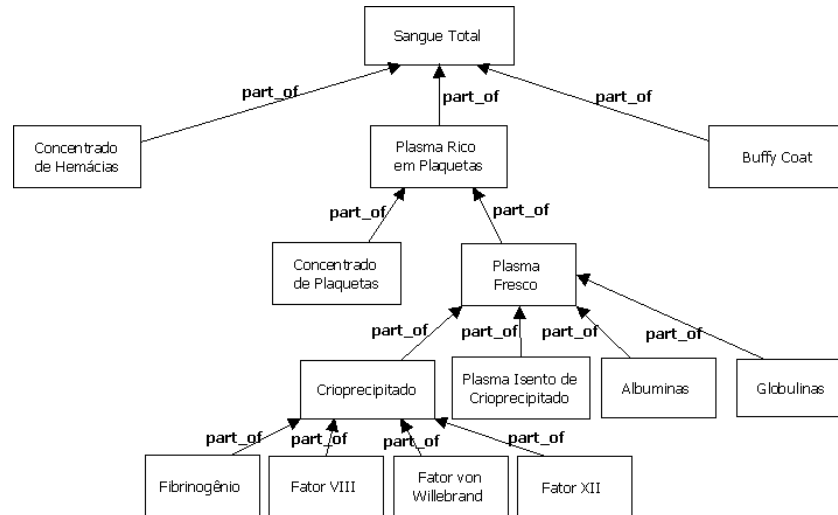
Fonte: elaborado pelo autor.

Na taxonomia da figura 4, *hemocomponente* é categorizado como uma entidade independente continuante e material da BFO. Mais especificamente, *hemocomponente* é um objeto (*BFO: object*). Na sequência, temos os três tipos de hemocomponentes (*eritrocitários*, *plasmáticos* e *plaquetários*) classificados quanto às camadas que os compõem. De acordo com a concentração de cada um dos componentes do sangue (hemácias, plasma e *buffy coat*), condições de armazenamento e temperatura e formas de obtenção (separação, centrifugação) obtém-se os hemocomponentes “de fato” utilizados na prática profissional da área, por exemplo, *plasma isento de crioprecipitado*, *concentrado de plaquetas*. Esses últimos são especializações dos tipos de hemocomponentes gerais e, são assim classificados de acordo com os componentes que os constituem (por exemplo, o *concentrado de plaquetas* é constituído de *plasma* e *buffy coat*).

Outro aspecto importante sobre hemocomponentes e hemoderivados refere-se aos elementos que os constituem. Tais elementos não são tipos de hemocomponentes ou hemoderivados, mas possuem uma relação ontológica de composição com eles, sendo assim,

representados em uma partonomia, usando a relação *part_of* da BFO entre dois universais continuantes (GRENON e SMITH, 2004) (vide figura 5).

Figura 5 - Partonomia dos elementos constituintes dos hemocomponentes e hemoderivados.



Fonte: elaborado pelo autor.

Como mencionado anteriormente, o sangue total, quando submetido ao processo de centrifugação, é separado em três camadas: hemácias, plasma e *buffy coat*; que originam os hemocomponentes e hemoderivados do sangue.

Em condições específicas de temperatura e armazenamento, as hemácias formam o hemocomponente *Concentrado de Hemácias*, enquanto o plasma em sua situação natural (bruta) é comumente chamado de *Plasma Rico em Plaquetas*.

Após um novo processo de centrifugação, realizado em alta rotação, o *Plasma Rico em Plaquetas* é separado em duas partes distintas: (i) o hemocomponente *Concentrado de Plaquetas* e (ii) um plasma com baixa porcentagem de plaquetas, aqui chamado, de *Plasma Fresco*, significando estar armazenado em condições adequadas de uso hemoterápico.

O *Plasma Fresco* pode ser submetido a um processo de extração de um de seus componentes – o *Crioprecipitado* – dando origem a dois outros hemocomponentes: (i) o próprio *Crioprecipitado* e (ii) o *Plasma Isento de Crioprecipitado*, que não tem esse produto. Do *Plasma Fresco* ainda é possível extrair dois hemoderivados (i – *Albuminas* e ii – *Globulinas*), a partir do fracionamento desse plasma por meio de processos industriais.

Por fim, o hemocomponente *Crioprecipitado* contém uma série de glicoproteínas de alto peso molecular (*Fibrinogênio*, *Fator VIII*, *Fator Von Willebrand* e *Fator XIII*) que cumprem o papel de fatores de coagulação na hemoterapia. Por meio de processos industriais é possível obter tais proteínas e gerar outro hemoderivado importante conhecido como *Concentrado de Fatores de Coagulação*, que engloba essas proteínas elucidadas.

5) Considerações finais

O presente artigo abordou modelos e teorias no contexto da prática científica. Discutiu o significado dos dois termos, definiu modelos e teorias, bem como estabeleceu a relação entre eles, e apresentou aplicações práticas de modelos e teorias no âmbito dos sistemas de informação. Além disso, apresentou aspectos metodológicos e resultados parciais da construção de uma teoria ontológica informal sobre o sangue humano, conduzida no contexto de um projeto sobre organização da informação biomédica. Cabe agora, destacar a importância de teorias e modelos no contexto da Ciência da Informação. Para tal, retoma-se uma forma de entender a realidade proposta por Popper.

Popper (1978) propõe uma visão da realidade fundamentada no que denomina a visão dos três mundos: i) o mundo um, da realidade física, que corresponde ao mundo como ele é; ii) o mundo dois, da realidade psicológica, que consiste do nosso conhecimento e crenças à respeito do mundo físico; e iii) o mundo três, das proposições, teorias, modelos, textos, etc., que corresponde a formalizações sobre nossas idéias e crenças, resultando em artefatos representacionais. Por exemplo, no âmbito da pesquisa médica, o mundo um consiste na realidade sobre os organismos, ou seja, sobre o paciente; o nível dois consiste de representações cognitivas dessa realidade física por parte dos médicos; e o nível três consiste de concretizações sobre tais representações em artefatos textuais, gráficos ou digitais, como por exemplo, textos clínicos e científicos, terminologias especializadas, tesouros, ontologias.

A Ciência da Informação se ocupa de estudar as entidades presentes nos três mundos propostos por Popper. Entretanto, a pesquisa em organização da informação tem interesse especial por artefatos representacionais que povoam o mundo três e sobre como elas podem melhor representar as entidades dos mundos um e dois. A importância das teorias e modelos no escopo da Ciência da Informação reside na capacidade dessas entidades em criar representações, as quais são relevantes para os propósitos de entendimento de um domínio com vistas à organização da informação. Assim, estudar formas de construir teorias, como proposto no presente artigo, corresponde à própria atividade de organizar a informação.

O fragmento de teoria ontológica aqui apresentada se encontra em estágio preliminar. No escopo do *Blood Project* planeja-se que a teoria seja expandida em eixos diferenciados. Uma desses eixos é a ampliação dos subdomínios da hematologia abordados; outro eixo é a expansão das relações abrangidas, considerando não apenas as relações taxonômicas e partitivas; outro eixo é criação de declarações axiomáticas fundamentais que representem os elementos causais dotando assim a teoria de capacidade explanatória.

Referências

- ABRIAL, J. R. Data semantics. In: KLIMBIE, J.W. e KOFFEMAN, K. L. (Eds.), **Proceedings of the IFIP Working Conference Data Base Management** (pp. 1–60). Amsterdam: North-Holland, 1974.
- ALMEIDA, M. B.; TEIXEIRA, L. M. D.; COELHO, K. C.; SOUZA, R. R. Relações semânticas em ontologias: estudo de caso do Blood Project. **Liinc em Revista**, Rio de Janeiro, v.6, n.2, p. 384-410, set. 2010.
- ALMEIDA, M. B. (forthcoming) **Revisiting ontologies: a needful clarification**, 2012.
- BACHELARD, G. **O novo espírito científico**, 1934. Disponível em: <<http://pt.scribd.com/doc/65380809/Gaston-Bachelard-O-Novo-Espirito-Cientifico>>. Acesso em: 23 maio 2012.
- BARRET, K. E.; BARMAN, S.M.; BOITANO, S.; BROOKS, H. **Ganong's Review of Medical Physiology**, 23rd edition. New York: McGraw-Hill Medical, 2009.
- BOOCH, G. **Object-oriented analysis and design with applications** (2nd ed.). Redwood City, CA: Benjamin Cummings, 1993.
- BOSAK, R., RICHARD, F. CLIPPINGER, R.F., DOBBS, C., GOLDFINGER, R., JASPER, R.B., KEATING, W., KENDRICK, G.; e SAMMET, J.E. An information algebra: Phase 1 report – language structure group of the CODASYL development committee. **Communications of the ACM**, 5(4), p. 190–204, 1962.
- BRASIL. Ministério da Saúde. **Guia para o uso de hemocomponentes**. Brasília, DF, 2008.
- BRINKMAN, R.R.; COURTOT, M.; DEROM, D.; FOSTEL, J.M.; HE, Y.; LORD, P. The OBI consortium. Modeling biomedical experimental processes with OBI, **Journal Biomedical Semantics** 1(Suppl 1): S7, 2010.
- BUNGE, M. **Teoria e Realidade**. São Paulo: Editora Perspectiva, 1974.
- CAPRA, F. **O tao da física**. São Paulo: Cultrix, p. 160, 1983.
- CARNAP, R. Foundations of Logic and Mathematics. In: NEURATH, O.; MORRIS, C.; e CARNAP, R. (eds.), **International Encyclopaedia of Unified Science**. Vol. 1. Chicago: University of Chicago Press, p. 139-213, 1938.
- CARTWRIGHT, N. **How the Laws of Physics Lie**. Oxford: Oxford University Press, 1983.
- CHORLEY, R.; e HAGGETT, P. Modelos, paradigmas e a nova geografia. In: CHORLEY, R. e HAGGETT, P. **Modelos sócios-econômicos em geografia**. Rio de Janeiro: Livros Técnicos e Científicos/USP, p.1-22, 1975.
- CHURCHLAND, P. M.; HOOKER, C. A. **Images of Science: Essays on Realism and Empiricism with a Reply from Bas C. Van Fraassen**, 1985. Disponível em: <http://books.google.com.br/books/about/Images_of_Science.html?id=Suc98NDaPMQC&redir_esc=y>. Acesso em: 01 jun. 2012.
- CODD, E.F. Extending the database relational model to capture more meaning. **ACM Transactions on Database Systems**, 4(4), p. 397–434, 1979.
- DOWNES, S. The importance of models in theorizing: a deflationary semantic view. **Proceedings of the Philosophy of Science Association**, Vol.1, East Lansing: Philosophy of Science Association, p. 142-153, 1992.

- FEYERABEND, P. **Against Method: Outline of an Anarchistic Theory of Knowledge**, 1975. Disponível em: <<http://plato.stanford.edu/entries/feyerabend/#2.13>>. Acesso em: 01 jun. 2012.
- FRIGG, R. **Models in science**, 2006. Disponível em: <<http://plato.stanford.edu/entries/models-science/>>. Acesso em: 20 abril 2012.
- GANDON, F. **Distributed artificial intelligence and knowledge management: ontologies and multi-agent systems for a corporate semantic web**. Tese, INRIA and University of Nice, Nice, France, 2002.
- GIERE, R. **Explaining Science: A Cognitive Approach**. Chicago: University of Chicago Press, 1988.
- GRENON, P.; SMITH. SNAP and SPAN: Towards Dynamic Spatial. **Spatial Cognition e Computation**, v.4, n.1, p. 69-104, 2004. Disponível em: http://Ontologyontology.buffalo.edu/smith/articles/SNAP_SPAN.pdf. Acesso em: 02 mai 2012.
- GREER, J.P. ; FOERSTER, J. ; RODGERS, G.M. ; PARASKEVAS, F. ; GLADER, B.; ARBER; D. A.; MEANS, R.T. **Wintrobe's Clinical Hematology**. Philadelphia: Lippincott Williams e Wilkins, 2008.
- GUARINO, N. Formal Ontology in Information Systems. **Proceedings of FOIS'98**, Trento, Italy, 6-8 June 1998. Amsterdam, IOS Press, pp. 3-15, 1998.
- HEMPEL, C. G., **Aspects of Scientific Explanation and other Essays in the Philosophy of Science**. New York: Free Press, 1965.
- HODGES, W., **A Shorter Model Theory**. Cambridge: Cambridge University Press, 1997.
- JACOBSON, I., CHRISTERSON, M., JONSSON, P., e OVERGAARD, G. **Object oriented software engineering: A use case driven approach**. Boston: Addison-Wesley, 1992.
- JARDINE, D. A. The ANSI/SPARC DBMS model. **Proceedings of the Second SHARE Working Conference on Database Management Systems**. Amsterdam: North Holland, 1976.
- KUHN, T. S. **The Structure of Scientific Revolutions**. Chicago: University of Chicago Press, 3 ed., 1962.
- LAKATOS, I. **Criticism and the Growth of Knowledge**. Cambridge: Cambridge University Press, 1970.
- LE MOIGNE, J. **A teoria do sistema geral: teoria da modelização**. Lisboa: Instituto Piaget, 1977.
- MILTON, S. **An ontological comparison and evaluation of data modelling frameworks**. Tese, University of Tasmania, Hobart, AU, School of Information Systems, 2000.
- MORRISON, M. Models as Autonomous Agents. In: MORGAN, M. S.; MORRISON, M. **Models As Mediators: Perspectives on Natural and Social Science**. Cambridge: Cambridge University Press, 1 ed., p. 38-65, 1999.
- MYLOPOULOS, J. Information modeling in the time of revolution. **Information Systems**, 23(3), p. 127–155, 1998.
- MYLOPOULOS, J. Conceptual modeling and telos. In: LOUCOPOULOS P., ZICARI, R. (Eds.), **Conceptual modeling, databases and case: an integrated view of information systems development**. New York: John Wiley e Sons, 1992.

- PESSOA JR., O. **Filosofia e Sociologia da Ciência**, 1993. Disponível em: <<http://www.cfh.ufsc.br/~wfil/sociociencia.htm>>. Acesso em: 29 maio 2012.
- PETTIGREW, K. E.; MCKECHIE, L. E. F. The use of theory in Information Science research. **Journal of the American Society for Information Science and Technology**, 52(1):62–73, 2001.
- POPPER, K. R. **Três concepções acerca do conhecimento humano**. São Paulo: Abril Cultural, 1980.
- POPPER, K. **Conjecturas e refutações: o progresso do conhecimento científico**. 5.ed. Brasília: UNB, 2008.
- POPPER, K. **Three Worlds**, 1978. Disponível em: <<http://www.tannerlectures.utah.edu/lectures/documents/popper80.pdf>>. Acesso: 05 jun 2012.
- PUTNAM, H. **Razão, Verdade e História**. Lisboa: Dom Quixote, 1992.
- REICHENBACH, H. Experience and prediction: an analysis of the foundations and the structure of knowledge. **Chicago: University of Chicago Press, 1938**.
- REYNOLS, P.D. **A primer in theory construction**. New York: Pearson, p. 191, 2007.
- ROSSE, C.; MEJINO, J. L. V. A reference ontology for biomedical informatics: The Foundational Model of Anatomy. *Journal of Biomedical Informatics*, v. 36, n. 2003, p. 478–500, 2003.
- RUMBAUGH, J., BLAHA, M., PREMERLANI, W., EDDY, F., e LORENSEN, W. **Object-oriented modeling and design**. New York: Prentice Hall, 1991.
- SAYÃO, L. F. Modelos teóricos em ciência da informação: abstração e método científico. **Ciência da Informação**, Brasília, v. 30, n. 1, p. 82-91, jan./abr. 2001.
- SILVA, M. Realismo e Anti-realismo na ciência: aspectos introdutórios de uma discussão sobre a natureza das teorias. **Revista Ciência e Educação**, 5(1), 7–13, 1998.
- SMITH, B.; WELTY, C. Ontology: Towards a new synthesis. In: SMITH B., e WELTY, C. (Eds.). **Proceedings of the International Conference on Formal Ontology in Information Systems** (pp. 3–9). New York: ACM Press, 2001.
- SUPPE, F. (org.), **The Structure of Scientific Theories**. 2a ed. Urbana: University of Illinois Press, pp. 1-241, 1977.
- SUPPES, P. **Representation and Invariance of Scientific Structures**. Stanford: CSLI Publications, 2002.
- VAN FRAASSEN, B. C. **The Scientific Image**. Oxford: Oxford University Press, 1980.
- VAN DE GRAAFF, K.M.; RHEES, R.D.; PALMER, S. L. **Human Anatomy and Physiology**, 3 ed. New York: McGraw-Hill, 2010.
- YOUNG, J.W., e KENT, H.K. Abstract formulation of data processing problems. **Journal of Industrial Engineering**, 9(6), p. 471–479, 1958.