

**VIII ENANCIB – Encontro Nacional de Pesquisa em Ciência da Informação
28 a 31 de outubro de 2007 • Salvador • Bahia • Brasil**

GT 2 – Organização e Representação do Conhecimento
Pôster

Ontologias: representando a pesquisa na área através de mapa conceitual

Ontologies: representing the research in the area through conceptual maps

Maria Luiza de Almeida Campos* (marialuizalmeida@gmail.com)
Carlos Henrique Marcondes* (marcon@vm.uff.br)
Laura Lira** (llira@gbl.com.br)
Leonardo Cruz da Costa** (leo.cruz@yahoo.com.br)
Linair Maria Campos** (linair@nce.ufrj.br)
Luciana Reis Malheiros** (malheiro@vm.uff.br)

Resumo: Ontologia é um tema de pesquisa emergente, abrangendo aportes teóricos e metodológicos da Filosofia, da Ciência da Computação, além da Ciência da Informação, se constituindo num dos pilares da proposta da Web Semântica. Esta emergência faz com que falte ainda à área princípios conceituais consistentes e terminologia, que permitam a cooperação entre grupos de pesquisa e o avanço da área. Tentando endereçar estas questões foi desenvolvido um estudo objetivando realizar um mapeamento da pesquisa na área, partindo de dois artigos de revisão. Este mapeamento foi representado através de mapa conceitual. O mapa representa os aspectos ligados a cada processo do ciclo de vida das ontologias, a saber: criação, manutenção e uso além de outros aspectos relacionados à ontologia como tema geral. Pretende-se, como continuidade deste trabalho, que ele sirva de base para realizar um mapeamento dos grupos de pesquisa em ontologias no Brasil, em diversas áreas.

Palavras-chave: Ontologia, Mapa Conceitual, Pesquisa, Ciência da Informação

Abstract: *Ontology is an emerging research area, encompassing theoretical and methodological insights from Philosophy and Computer Science, besides Information Science. It is also one of the bases of Semantic Web proposal. This emergence results in lack of consistent conceptual principles and terminology which enable the cooperation between research groups and the advancement of the area. Trying to address these questions this research develops a general mapping of the research on ontology, having as a starting point two reviews. The result of this mapping was represented as a conceptual map. Besides different aspects of ontology life cycle as creation, maintenance and us, the conceptual maps represents other aspects of ontology as a general research theme too. The aim of this research in a near future is to provide means to develop a mapping of research groups in different areas working on the theme ontology.*

Keywords: *Ontology, Conceptual map, research, Information Science.*

* Professores, Departamento de Ciência da Informação/UFF, PPGCI convênio UFF/IBICT

** Doutorandos, PPGCI convênio UFF/IBICT

1. ONTOLOGIA COMO TEMÁTICA DE ESTUDO EM CIÊNCIA DA INFORMAÇÃO

No amplo espectro do desenvolvimento da cultura humana, o conhecimento científico vem, desde a Modernidade, evoluindo seguindo uma dinâmica própria e passando desde então a se imbricar cada vez mais com o sistema produtivo. Hoje a sociedade não produz sem o aporte do conhecimento em geral e, especificamente, do conhecimento científico. Não é por outra razão que a sociedade atual é chamada por muitos (WERTHEIN, 2000; LOJKINE, 1995) de sociedade da informação ou sociedade do conhecimento. Meios econômicos, sociais, políticos e tecnológicos são aportados pela sociedade atual para gerir nosso cada vez mais vasto acervo de conhecimentos. A partir do surgimento da Web esse acervo se amplia de forma rápida e passa a estar disponível em formato digital na própria Web.

O surgimento da WWW – World Wide Web, a “teia global” proposta por Tim Berners-Lee no CERN em 1989, transforma a Internet num gigantesco sistema de informações à escala mundial, com um número crescente de usuários. No entanto o crescimento gigantesco da Internet a partir daí colocou novos e inéditos desafios para o acesso e utilização a esta informação. A enorme facilidade de qualquer um publicar na Internet fez com que a rede tivesse um crescimento desordenado e caótico e que encontrar a informação adequada passasse a ser o principal problema cultural, econômico e científico da atualidade. Nunca a humanidade dispôs de tanta informação, com o potencial de ser disseminada para qualquer um, independente da sua localização física, no entanto, nunca foi tão difícil identificar informação relevante.

Tentando endereçar estes problemas, Tim Berners-Lee, criador da WWW, propõe a visão da Web Semântica, uma extensão da Web atual, formada por documentos compreensíveis unicamente por pessoas, para uma Web em que documentos seriam auto-descritíveis, de forma que seu conteúdo possa ser “compreendido” por programas especiais, os agentes inteligentes de “software”*, que assim poderiam “raciocinar” e fazer inferências sobre o conteúdo de documentos, ajudando as pessoas em diferentes tarefas de recuperação de informações que exijam raciocínio, decisões, inferência de conclusões a partir de informações não explicitamente disponíveis ou de informações contextuais. Nas palavras de Berners-Lee a Web Semântica “*é uma EXTENSÃO da Web atual, na qual é dada à informação um SIGNIFICADO bem definido, permitindo que computadores e pessoas trabalhem em cooperação*” (BERNERS-LEE; HENDLER; LASSILA, 2001). Um dos pilares da Web Semântica, juntamente com os metadados, os “agentes de software”ⁱ, são as ontologias (JACOB, 2003).

Uma ontologia é um modelo de informações representando um conjunto de conceitos num domínio específico, estruturados e inter-relacionados entre si, de entendimento compartilhado em comum por uma comunidade de usuários. Conceitos são organizados em hierarquias de classes e possuem atributos e relações entre si. Uma ontologia é representada em linguagem “inteligível” por programas “agentes de software” e usada por estes para fazer inferências sobre os conceitos desse domínio.

Ontologias foram pensadas no contexto da Web Semântica para tornar interoperáveis sistemas computacionais distintos como, por exemplo, um sistema automático de reserva de passagens aéreas com um sistema automático de reserva de hotéis de uma determinada cidade turística. O que na semântica de um sistema seria um *passageiro*, na semântica do outro seria um *hóspede*. Exemplos de interoperabilidade entre sistemas como o mostrado vão ser tornar uma realidade, na medida em que as potencialidades da Web Semântica estejam implementadas: “agentes de software” inteligentes vão conseguir navegar pela Web e realizar reservas de passagens e hotéis de forma a atender a agenda, as necessidades e preferências de seus usuários, de forma automática, dialogando com “web services” (BREITMAN, 2005) que disponibilizam passagens aéreas e hospede-

dagem na cidade de destino, “compreendendo” o funcionamento desses serviços através de ontologias. Vários exemplos desse tipo de aplicação são ilustrados no artigo de Tim Berners-Lee (2001) já citado.

Como tornar as ontologias interoperáveis para que todo o potencial das tecnologias da Web Semântica possa ser realizado? O desenvolvimento e o uso de ontologias também tem se dado de forma isolada e circunstancial, respondendo, como no caso da Gene Ontologyⁱⁱ (2000), à necessidade premente de organizar o vasto material disponível em bancos de dados de seqüenciamento genético. Este desenvolvimento ressoa-se de princípios metodológicos sólidos e consistentes, amplamente testados e aceitos entre as diferentes comunidades usuárias. Como em toda área emergente esta carência pode ser percebida, através do exame da literatura, pela ausência de um vocabulário e terminologia comuns entre os diferentes grupos atuantes.

Deste processo inicial tem resultado ontologias com um alto grau de inconsistências internas, como é o caso da Gene Ontology (SMITH; KOHLER; KHUMAR, 2004) e da UMLSⁱⁱⁱ, além de ontologias isoladas, que não se falam, com baixo potencial de compatibilidade e interoperabilidade, restringindo seu uso.

Barry Smith (2002, p.2), discutindo a relação entre Ciência e Ontologia (enquanto domínio de conhecimento preocupado com a natureza dos seres), afirma que enquanto o papel da ciência é “explicar” a natureza, o papel da Ontologia seria vir a seguir para descrever, organizar e sistematizar o conhecimento obtido pelas descobertas científicas. Este parece ser o lugar a ser ocupado pela Ciência da Informação, vinda de uma longa tradição de organização de domínios de conhecimento, aplicada originalmente à organização de repertórios documentais. Ao longo desta trajetória a Ciência da Informação desenvolveu um arcabouço teórico e metodológico, consubstanciado em obras como as de Otlet (1989), Ranganathan (1967), Dahlberg (1978), Hjørland (2002) e do CRG – Classification Research Group (WILSON, T.D. 1972), que potencialmente tem grande aplicabilidade no desenvolvimento de ontologias. Isso se manifesta através de metodologias e princípios para a construção de tabelas de classificação e pela proposição de esquemas classificatórios para amplos domínios de conhecimento, como são a CDD e CDU, que hoje organizam os grandes repertórios bibliográficos em bibliotecas do mundo inteiro, pela organização de instrumentos de controle terminológico como vocabulários controlados e tesouros, na identificação e delimitação de domínios de conhecimento, na visão segundo diferentes facetas de um determinado domínio.

A Ciência da Informação sempre trabalhou com outras áreas de conhecimento, numa perspectiva interdisciplinar, ajudando a organizar e sistematizar o conhecimento dessas áreas. A atuação da Ciência da Informação no desenvolvimento de ontologias é, portanto uma consequência lógica e natural de toda esta trajetória.

Esta relação entre a Ciência da Informação, suas metodologias e práticas e a pesquisa em ontologias é cada vez mais percebida pelos pesquisadores desta área (SOERGEL, 2000), (THE SEMANTIC WEB, 2003). Pesquisar ontologias a partir da perspectiva da Ciência da Informação é a proposta comum dos grupos de pesquisa que discutem esta temática no âmbito PPGCI UFF/IBICT^{iv}. Para dar início a sua efetivação, considerou-se necessário, num primeiro momento, montar um amplo quadro didático da situação da pesquisa em ontologias tanto nacional como internacionalmente. Esta proposta foi iniciada na atividade programada “Estudo de princípios metodológicos de ontologias no domínio da Ciência da Saúde”, realizada no 1o. semestre de 2007.

O resultado é o mapa conceitual apresentado neste trabalho. Além de atender ao objetivo inicial da atividade didática, de servir para organizar e tornar mais compreensível o campo de pesquisa em ontologias, o mapa serviu, ao longo da atividade programada, para situar os diferen-

tes projetos de pesquisa dos alunos. Servirá também para uma futura atividade do grupo de pesquisa, de promover um encontro nacional dos grupos de pesquisa em ontologias, identificados a partir do Diretório de Grupos de Pesquisa do CNPq, situando cada grupo pelas suas atividades, dentro do mapa.

Este trabalho está assim organizado. Na próxima seção será descrita a metodologia de trabalho usada para o desenvolvimento do mapa. A seguir, o próprio mapa será apresentado e através dele, mostrada a situação da pesquisa em ontologias. Finalmente, na seção seguinte, os resultados serão discutidos e apresentados às conclusões e propostas de continuidade desta pesquisa.

2. MATERIAL E MÉTODO: O MAPA CONCEITUAL COMO PRODUTO DA MODELIZAÇÃO DO DOMÍNIO DE ONTOLOGIA

O material utilizado para a elaboração do mapa conceitual teve como base dois artigos de revisão sobre pesquisa e desenvolvimento de ontologias de autoria de Ying Ding e Schubert Foo (DING; FOO, 2002a,2002b). Ambos os artigos foram publicados no *Journal of Information Science* tendo como objetivos apresentar o estado da arte da pesquisa sobre criação (*generation*) de ontologias (DING; FOO, 2002a) e sobre seu mapeamento (*mapping*) e manutenção (*evolving*) (DING; FOO, 2002b).

Após leitura dos artigos iniciou-se a construção do mapa conceitual, apoiada em metodologia para organização e representação de domínios de conhecimento. Tendo concluído um primeiro esboço do mapa, o seu primeiro uso foi a tentativa de identificar nele os tópicos de interesse abordados nos temários de dois importantes eventos relacionados à área de ontologias . Os dois eventos foram: a) *Workshop on Foundational Issues in Medical Ontology*^v ocorrido em São Francisco, EUA, em setembro de 2004 como parte do *International Health (Medical) Informatics Congress*, MedInfo e b) *1st International Workshop on Ontologies and Information Systems for the Semantic Web*^{vi}, que faz parte da *26th International Conference on Conceptual Modeling* a ser realizada em novembro próximo em Auckland, Nova Zelândia.

A discussão dos temas abordados por ambos os eventos e sua inserção no mapa conceitual permitiu ao grupo aperfeiçoá-lo, através de artigos que consideramos relevantes para complementar o entendimento do tema, e utilizá-lo para o mapeamento dos projetos de pesquisa dos alunos da disciplina.

A metodologia que nos apoiamos para a representação dos conceitos sobre ontologia, através de mapa conceitual, se baseou em princípios utilizados para a representação de domínios, fruto de estudos que vem sendo desenvolvidos no âmbito da organização e representação do conhecimento (CAMPOS, 2006).

Apoiados em uma postura metodológica onde possamos pensar nos princípios que possibilitam o ato de modelar, ou seja, a ação de modelização. (LE MOIGNE, 1977), nos conduzimos por etapas onde discutimos desde o objeto de representação, ou seja, neste caso, os conceitos apresentados pelos autores dos artigos analisados, passando pelos critérios utilizados para a organização desses conceitos em classes de conceitos e suas relações, até a forma gráfica de representação, que se apoiou no modelo de Mapas Conceituais.

O Mapa Conceitual apresentado neste estudo ultrapassa uma perspectiva de representação hierárquica e vai buscar formas de representação gráfica que evidencie também relações não hierárquicas, como veremos adiante.

Por outro lado, nossa intenção é evidenciar o entendimento de nosso grupo sobre os estudos que vêm sendo desenvolvidos em ontologias através de uma representação gráfica, fruto de

uma proposta representacional que em si está repleta de intencionalidade, como é próprio em todo ato de modelar.

Apesar da intencionalidade intrínseca a todo ato de representação, nos pautamos também em uma postura que pretende deixar evidentes os princípios pelos quais tal representação se apoiou. Nesta perspectiva, apresentamos as etapas que nos conduziram à formatação do Mapa Conceitual: levantamento dos conceitos; organização de classes de conceitos e seus relacionamentos e a forma gráfica de representação.

Na etapa de levantamento foram identificados conceitos a partir do que os autores chamaram de ciclo de vida de uma ontologia, a saber: Criação, Manutenção e Uso. Estes conceitos se caracterizam como processos de alto nível e desta forma o nosso mapa conceitual se pautou na categorização do domínio de Ontologia a partir destas ações. Esta forma de condução metodológica que representa um domínio a partir de categorias de conceitos é denominada, segundo a Teoria da Classificação Facetada (RANGANATHAN, 1967), como um método policotômico. Este método por definição permite a organização de classes com um princípio de divisão (coisas e seus tipos, processos e seus tipos, etc.), não priorizando nenhum dos aspectos, sendo o seu segundo nível uma especificação do primeiro. Assim as classes subordinadas a cada etapa de ciclo de vida eram também tipos ou etapas de processos. Para mapear os conceitos apresentados pelo autor que não pertenciam à categoria processo (por exemplo: tipos de ontologias, ferramentas computacionais empregadas) utilizamos uma representação gráfica e localização diferenciada. O “software” CMAP^{vii} foi adotado para a representação gráfica do mapa.

Nessa representação, utilizamos a seguinte convenção para os relacionamentos entre os conceitos: (i) as relações do tipo gênero espécie estão representadas através de linhas cheias, sem nenhuma identificação; (ii) as demais relações estão representadas através de linhas tracejadas e contêm um rótulo com a identificação do seu tipo; por exemplo: parteDe, representadoPor; gera; usa, etc.

Devido a grande amplitude do tema, os artigos de Ding e Foo que foram o fio condutor da elaboração inicial do mapa, foram complementados com outros autores como Barry Smith (2002, 2004), Guarino (1998b), Uschold (2000) e Reich (1999).

3. MAPA CONCEITUAL DOS ESTUDOS EM ONTOLOGIA: CONSTRUÇÃO E O ESTABELECIMENTO DAS RELAÇÕES ENTRE OS CONCEITOS

O mapa reflete questões gerais ligadas à ontologia como tema de estudo e outras que dizem respeito aos aspectos que as ontologias possuem, enquanto produto final do processo de criação (ontologia produto) (Figura 1). Além disso, o mapa aborda as questões relacionadas ao ciclo de vida das ontologias, o qual é composto por três processos, a saber: processos de criação (Figuras 2), de manutenção (Figura 3) e de uso (Figura 4).

Em cada um desses processos existem tarefas e questões associadas, que estão representadas no mapa de forma sumária. Para facilitar a visualização do mapa, este foi quebrado em quatro partes, que serão comentadas a seguir.

3.1. QUESTÕES LIGADAS AS ONTOLOGIA COMO TEMA

A princípio, **questões políticas e epistemológicas** permeiam todo o **ciclo de vida das ontologias**. As primeiras, devido ao fato de que a construção de uma ontologia é um processo de intervenção em um domínio, privilegiando recortes do conhecimento em detrimento de outros, na medida em que é uma dentre várias representações possíveis, de acordo com os interesses da comunidade que vai utilizá-la. As questões epistemológicas, por sua vez, também estão ligadas ao tema ontologia, na medida em que não se podem dissociar as teorias do conhecimento dos diversos *loci* onde este se manifesta. Nesse sentido, cada processo do ciclo de vida possui problemas que devem buscar apoio em um aporte teórico sólido que permita a sua compreensão mais ampla.

Os **princípios conceituais**, ou seja, critérios teóricos e metodológicos adotados, por sua vez, estão na base das escolhas que determinam os componentes de uma dada ontologia.

3.2. QUESTÕES GERAIS LIGADAS AS ONTOLOGIA COMO PRODUTO

Ontologias podem ser classificadas de acordo com a sua **tipologia** sob três olhares:

(i) pela gestão relativa à inclusão de novos termos, onde podem ser **curadas** ou **não curadas**. As primeiras possuem um processo criterioso de aprovação para a inclusão, alteração ou exclusão de termos, enquanto que as segundas não possuem;

(ii) pela natureza de aplicação, onde, de acordo com seu nível de generalidade são classificadas em: **genéricas, de domínio, de tarefa, de aplicação** (GUARINO, 1998a).

Ontologias genéricas ou de alto nível descrevem conceitos gerais, tais como: tempo espaço, matéria, objeto, evento, ação, e são independentes de um domínio particular, ou seja, se aplicam a mais de um domínio (GUARINO, 1998b). Algumas ontologias genéricas já se acham disponíveis, tais como a Standard Upper Ontology (SUO W.G, 2004), a Suggested Upper Merged Ontology - SUMO (NILES; PEASE, 2001) e a Cyc Upper Ontology (CYC, 2004). Ontologias de domínio expressam conceitos de domínios particulares, descrevendo o vocabulário relacionado a este domínio. Por exemplo, para os domínios de Biologia e Informática, poderia incluir conceitos tais como proteína e célula, ou ainda banco de dados e estrutura de dados, respectivamente. São especializações de uma ontologia genérica. Ontologias de tarefa expressam conceitos sobre a resolução de problemas, isto é, descrevem o vocabulário relacionado a uma atividade ou tarefa genérica, tal como, diagnose ou vendas, sendo também especializações de uma ontologia genérica. Ontologias de aplicação descrevem conceitos dependentes de um domínio específico e de tarefa, geralmente especializando ambas as ontologias.

(iii) pelo formalismo da expressividade das relações e do conteúdo conceitual de seus termos, onde podem ser **formais, semi-formais e informais**. As primeiras utilizam regras e são projetadas com o intuito de permitirem inferências mais precisas. Existem poucas disponíveis, sendo estas utilizadas para ilustrar funcionalidades de aplicativos. Um exemplo é a ontologia da pizza^{viii}. Já as ontologias informais são também chamadas de “ontologias leves” (*lightweight ontologies*) e são vocabulários simples, constituídos basicamente por uma hierarquia de termos e relações básicas, do tipo *é-um* e *parte-todo*, muitas vezes constituídas sem rigor. Um exemplo desse tipo de ontologia é a Gene Ontology (2000). Por fim, as ontologias semi-formais são um meio termo entre as duas já citadas. Não possuem um formalismo preciso, e, portanto, não são adequadas para inferências através de *reasoners*, mas já possuem algumas relações identificadas

e um maior rigor na formação das suas hierarquias de conceitos. Um exemplo desse tipo de ontologia é a UMLS^{ix}.

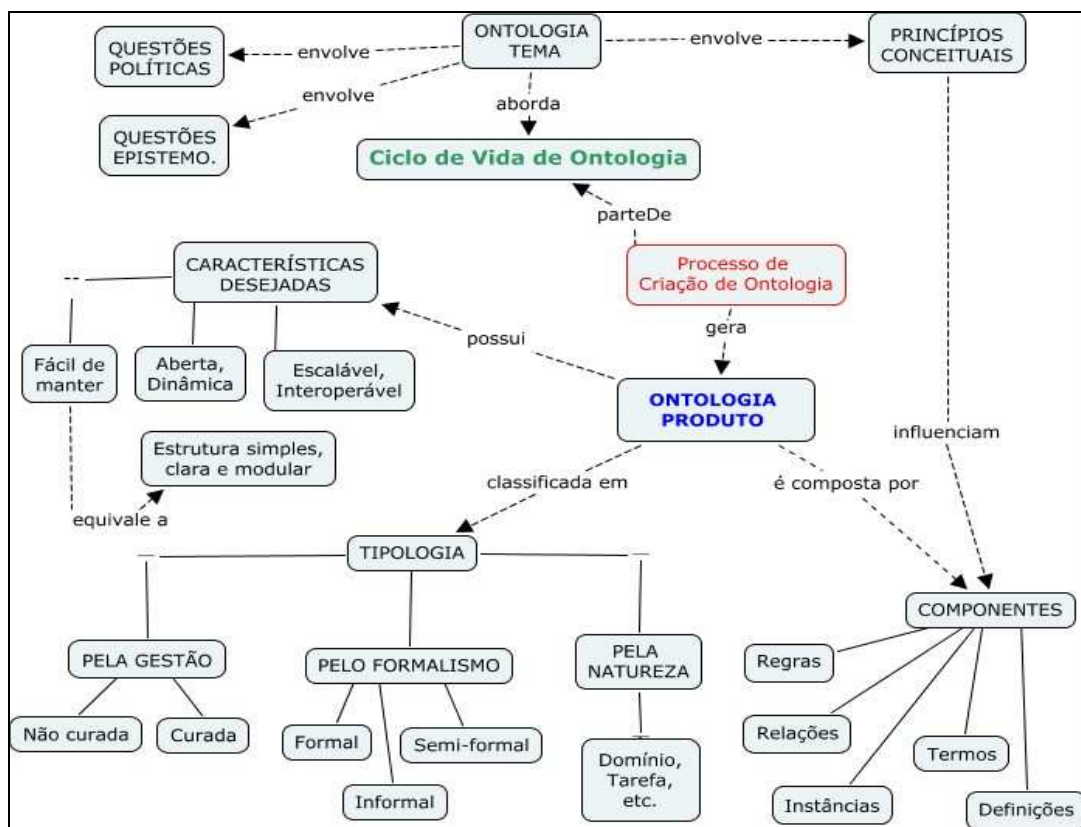


Figura 1: Aspectos gerais relacionados às ontologias

Para que seu uso seja otimizado uma ontologia deve possuir **características desejadas**, onde se espera que seja **aberta, dinâmica, fácil de manter, escalável e interoperável** e que possua uma **estrutura simples, clara e modular**.

Uma ontologia possui como componentes: **definições, classes de termos, relações, regras e instâncias**, sendo que as regras são utilizadas para fornecer maior formalismo, através da caracterização de aspectos que podem ser utilizados para inferências, como, por exemplo: transitividade, condições necessárias e suficientes, dentre outros. Cabe ressaltar que poucas ontologias, até onde se tem notícia, utilizam regras. Quanto às **instâncias**, estas se referem aos objetos individuais naquele domínio.

3.3. CRIAÇÃO

O processo de criação, (Figura 2) envolve o sub-processo de **levantamento** e a escolha das **formas de criação** e dos **princípios de concepção**. Como resultado uma **ontologia** é gerada.

O sub-processo de **levantamento** se apóia na análise do domínio (não explicitado no mapa), que envolve questões diversas que vão desde a determinação do recorte adequado, até a **captura do conhecimento** através das **fontes de informação** sobre o domínio analisado. As fontes

de informação encontram-se nos documentos (**corpus** de informação do domínio) e **ontologias existentes**, **fontes terminológicas** (tesauros, terminologias, glossários, etc) ou ainda através do conhecimento tácito de **especialistas**, que pode ser registrado através de diferentes formas de representação (gráficos, mapas conceituais, lista de termos).

A **forma de criação** pode ou não envolver o reuso de vocabulários já existentes, sendo que esse reuso se dá de forma **manual**, **automática** ou **semi-automática**, sendo que nos dois últimos casos existem imprecisões, especialmente ao se comparar conceitos candidatos em vocabulários diferentes (PREDOIU et al., 2005). O caso do **reuso de vocabulários** e o processo de **avaliação da ontologia** (este envolvendo questões de qualidade) serão detalhados no processo de **manutenção de ontologias**, pois é comum a ambos, sendo ilustrado nas Figuras 1 e 3.

Ainda, relacionado às formas de criação, temos a escolha dos **métodos de raciocínio**, que podem ser: **indutivo**, **dedutivo** e **analítico-sintético**. Essa escolha envolve **questões epistemológicas**, que são largamente discutidas na literatura (DAHLBERG, 1978; SMITH, 2002; CAMPOS, 2006; HJORLAND, 2002) e que versam sobre as diferentes formas de organizar os conceitos em um domínio.

Por fim, diversos trabalhos propõem **princípios para a concepção** ou projeto de ontologias. Esses princípios expressam de maneira prática: (i) algumas das questões epistemológicas citadas acima, que tratam de como os conceitos se relacionam em um dado contexto; (ii) de que forma se deve dar o processo de levantamento, codificação, documentação e uso de padrões de projeto; (iii) a forma de criação, incluindo propostas de algoritmos para integração de ontologias já existentes. Dentre os autores que estudam essas questões estão **Guarino** (1998b), **Uschold** (2000) e **Reich** (1999).

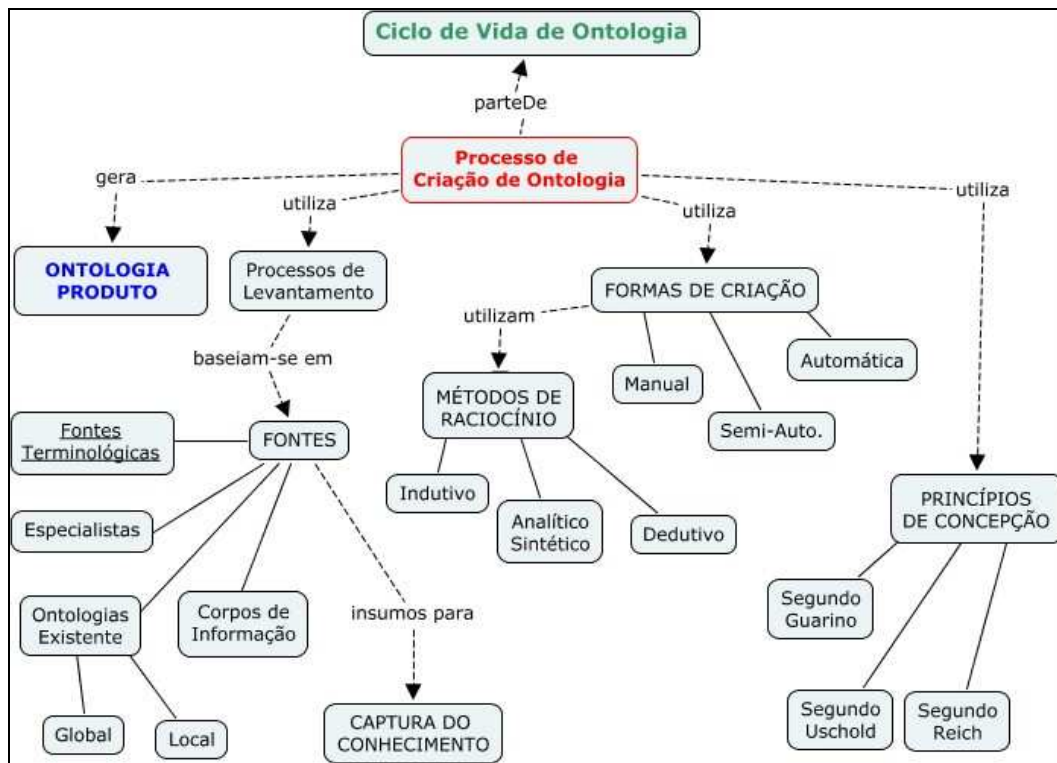


Figura 2: Processo de Criação de Ontologias

3.4. MANUTENÇÃO

O **processo de manutenção**, representado na Figura 3, envolve os sub-processos de **versionamento**, **avaliação** e ainda os diversos **mecanismos de reuso de ontologias**, sendo que os dois últimos, como mencionados anteriormente, também são utilizados pelo **processo de criação de ontologias**.

O **versionamento** de uma ontologia envolve questões de gerenciamento relacionadas à **aquisição**, **alteração** e **exclusão de conceitos** e, conseqüentemente, como tratar a compatibilização da nova versão da ontologia com as anteriores, além de aspectos relacionados à permissões de acesso e controle de autoria, dentre outras.

O processo de **avaliação** envolve o estabelecimento e a aplicação de critérios e métricas de qualidade de modo a aferir e garantir a qualidade da ontologia gerada. Esse processo se apóia em uma série de **características desejadas**, as quais serão abordadas mais adiante, quando da discussão de outros aspectos relacionados às ontologias (Figura 1).

Entretanto, cabe observar que para que características desejadas sejam obtidas, é importante que se observem os aspectos discutidos na sua criação, como, por exemplo, os **princípios de concepção**. Embora estes aspectos sejam definidos na sua criação, devem continuar a ser seguidos quando da sua manutenção.

O **processo de manutenção** vai tratar da modificação da ontologia ao longo do tempo, e, desta forma, envolve os **mecanismos de reuso**, que são: **mapeamento**, **alinhamento**, **junção** e **integração**. Esses mecanismos são realizados por **agentes de software**, que devem ser compatíveis com a linguagem na qual está codificada a ontologia. Esses agentes partem do mapeamento de termos semelhantes para encontrar aqueles que são candidatos ao reuso.

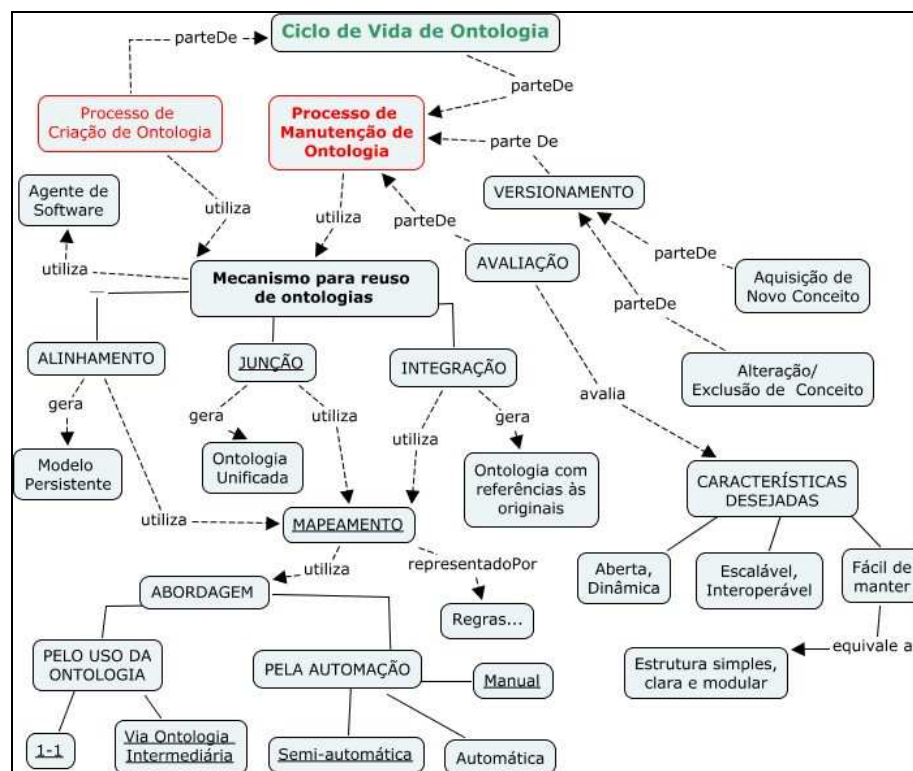


Figura 3: Processo de Manutenção de Ontologias

O **mapeamento**, por sua vez, é construído de acordo com diversas abordagens ou estratégias computacionais, que vão desde o uso de algoritmos de clusterização, para identificar relações candidatas, até uma variedade de técnicas com o objetivo de estabelecer a proximidade de dois conceitos, baseadas na estrutura hierárquica da ontologia e a denominação (forma verbal) dos seus conceitos (PREDOIU et al., 2005). Essas abordagens, de modo geral, adotam duas estratégias básicas: ou comparam as ontologias diretamente, uma com a outra (**um para um**), ou fazem uso de uma **ontologia intermediária**, que serve de base para o estabelecimento do grau de similaridade dos conceitos. Ou seja: no primeiro caso, as ontologias são comparadas entre si, enquanto que no segundo são comparadas, uma a uma, com a ontologia de referência. A melhoria da capacidade de reconhecimento da semelhança de conceitos, obtida a partir das técnicas de **mapeamento** de ontologias, relacionada as questões que envolvem a expressividade semântica das ontologias, é uma questão de grande importância, devido ao fato de estar na base das **abordagens automáticas e semi-automáticas** de processamento de ontologias. O mapeamento é representado por um conjunto de **regras** ou procedimentos que estabelecem de que forma os conceitos de uma ontologia devem corresponder aos da outra.

Os **mecanismos de reuso** se destinam a diferentes propósitos.

O **alinhamento** tem como finalidade promover a compatibilização de ontologias. O resultado desse alinhamento é um **modelo persistente** (FELICÍSSIMO; BREITMAN, 2004) indicando as ligações (*links*) entre as ontologias alinhadas. Este modelo é utilizado pelos **agentes de software** ao processarem ambas as ontologias, sendo que as ontologias originais são mantidas independentes (NOY; MUSEN, 1999).

O mecanismo de *integração* também preserva as ontologias originais que são reutilizadas, sendo estas incorporadas a uma ontologia base através de referências. O resultado é uma **ontologia integrada com referências às originais**, que contém conceitos próprios e de outras, estes últimos obtidos no processo de **integração** (PINTO; MARTINS, 2001).

O mecanismo de *junção*, por outro lado, não preserva as ontologias originais. Seus termos e relacionamentos são parcial ou totalmente fundidos com uma outra ontologia base. O resultado é uma **ontologia unificada**, onde não se sabe ao certo qual a origem de cada termo (PINTO; MARTINS, 2001).

Cabe aqui ressaltar que não existe um consenso na literatura sobre a terminologia adotada para descrever esses mecanismos. Em alguns trabalhos (ZIMMERMANN et al., 2006), **alinhamento** é tratado como sinônimo de **mapeamento**, assim como **junção** é por vezes adotado com o mesmo sentido de **integração** (HOIJMAIJERS; STUMPTNER, 2006). No nosso modo de ver, é importante caracterizar esses mecanismos de forma distinta e não ambígua, pois geram questões de natureza e conseqüências distintas. Por exemplo, ao se juntar duas ontologias a alteração do termo na sua ontologia de origem não causa nenhum impacto na ontologia que incorporou esse termo, o que já não é verdade na integração. Nesta, em virtude de possíveis mudanças nos conceitos importados, deve haver sempre a preocupação com a consistência da ontologia resultante ao longo do seu ciclo de vida.

3.5. USO

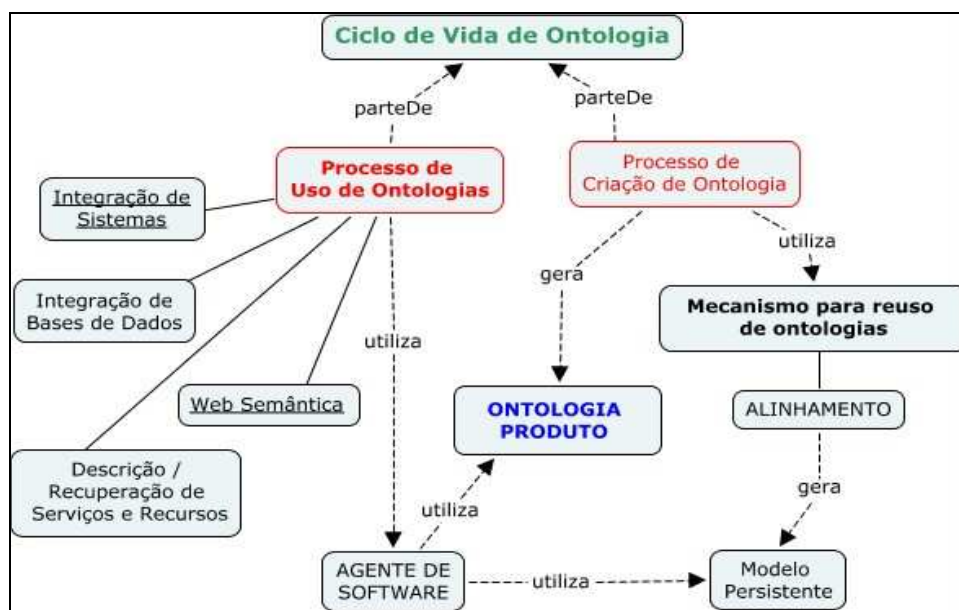


Figura 4: Processo de Uso de Ontologias

O processo de **uso de ontologias**, representado na Figura 4, envolve iniciativas de **integração de sistemas e de bases de dados**, com o foco maior na compatibilização de dados, e também as iniciativas de **descrição e recuperação de serviços e recursos**, para diferentes propósitos, como, por exemplo, a descoberta de serviços web ou a anotação de seqüências genômicas. O uso de ontologias frequentemente envolve o uso de **agentes de software**, que devem ser capazes de usar a **ontologia** empregada para descrever os serviços, recursos ou os dados a serem integrados, estando de acordo com os compromissos ontológicos^x adotados na mesma.

4. CONSIDERAÇÕES FINAIS

A área de pesquisa em ontologias é multidisciplinar. A origem do termo vem da Filosofia, mas o tema é retomado na década de 80 na área de Ciência da Computação, mais especificamente na área de Inteligência Artificial (SMITH; KOHLER; KUHMAR, 2002), para representar o conhecimento em domínios específicos. A partir do surgimento da proposta da Web Semântica, que tem como um dos seus pilares a questão das ontologias, o tema vem ganhando uma grande abrangência, ultrapassando as fronteiras da Ciência da Computação, passando a ser utilizado por diferentes grupos interessados em formalizar domínios específicos de conhecimento para torná-los processáveis por programas. Todo este processo gera uma diversidade de enfoques e terminologias que têm que ser superadas para que a pesquisa na área avance.

O mapa conceitual que foi iniciado no âmbito dos grupos de pesquisa em Ontologia do PPGCI UFF/IBICT tem por objetivo colocar diante dos pesquisadores as diversidades da área de pesquisa em ontologias. Isso permitiu, num primeiro momento, identificar pontos em comum ou interfaces entre as pesquisas.

O mapa elaborado revelou de forma abrangente Ontologia como tema e as diversas categorias que se desdobram quanto aos seus diversos aspectos teóricos e práticos e colocou em evidência as relações entre elas. A visão proporcionada pelo mapa certamente será útil para além dos interesses imediatos dos grupos que o criaram.

Grupos que desenvolvem ontologias têm nesta proposta inicial uma visão geral da área. Além disso, possibilita uma avaliação atual das atividades dos pesquisadores que trabalham nesta área. Pretende-se, como continuidade deste trabalho, que ele sirva de base para realizar um mapeamento dos grupos de pesquisa em ontologias no Brasil, em diversas áreas. Isso pode facilitar que os diversos grupos existentes se enxerguem enquanto área de pesquisa e assim viabilizar a cooperação entre eles e o avanço científico da área

5. REFERÊNCIAS

BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. The Semantic Web, **Scientific American**, may, 2001.

BREITMAN, Karin. **Web Semântica: a internet do futuro**. Rio de Janeiro: LTC, 2005.

CAMPOS, Linair Maria. **O aporte teórica do Ciência da Informação aplicado à elaboração de ontologias em Biomedicina**. Projeto de tese de doutorado para ingresso no PPGCI UFF/IBICT. Niterói, 2006.

CAMPOS, M.L.A. Integração de ontologias: o domínio da bioinformática e a problemática da compatibilização terminológica. In: ENANCIB, 7., 2006, Marília. **Anais...** Marília: ANCIB, UNESP. Disponível em: <<http://www.portalppgci.marilia.unesp.br/enancib/viewpaper.php?id=163>>. Acesso: 07 jul 2007.

COSTA, Leonardo Cruz da. **Uma ferramenta para edição, extração e representação do conhecimento contido em artigos científicos publicados na Web**. Projeto de Tese de Doutorado para ingresso no PPGCI UFF/IBICT. Niterói, 2006.

CYC, **Cyc Upper Ontology. The open source version of the Cyc technology**. Disponível em: <http://www.cyc.com/cyc/opencyc/overview>. Acesso: 02 abr 2004.

DAHLBERG, I. A Referent-oriented analytical concept theory of interconcept. **International Classification**, Frankfurt, v.5, n.3, p.142-150, 1978.

DAHLBERG, I. Ontical structures and universal classification. **Bangalore : Sarada Ranganthan Endowment**, 1978. 64 p.

DING, Yin; FOO, Schubert. Ontology research and development. Part 2 - a review of ontology generation. **Journal of Information Science**, v.28, n.5, p. 375-388, 2002.

DING, Yin; FOO, Schubert. Ontology research and development. Part 1 - a review of ontology mapping and evolving. **Journal of Information Science**, v. 28, n. 2, p. 123–136, 2002.

FELICÍSSIMO, Carolina H.; BREITMAN, Karin K. Taxonomic ontology alignment – an implementation. In: WORKSHOP DE ENGENHARIA DE REQUISITOS, Argentina, Tandil, 2004, **Anais...** Argentina, 2004, p. 152-163. Disponível em <www.ldb.dcc.ufmg.br/bdbcomp>. Acesso: 31 jul 2007.

GENE ONTOLOGY CONSORTIUM. Gene Ontology: tool for the unification of biology. **Nature Genetics**, n. 25, p. 25-29, 2000.

GUARINO, N. Formal ontology and information systems. In: FOIS '98, 1., 1998, Trento, Italy. **Proceedings...** Amsterdam: IOS Press; Tokyo: Omsha, 1998. p. 3-15.

GUARINO, N. Some Ontological Principles for Designing Upper Level Lexical Resources, In: INTERNATIONAL CONFERENCE ON LANGUAGE RESOURCES AND EVALUATION, 1., Granada, May, 1998, **Proceedings...** 1998.

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

HJØRLAND, Birger. Domain analysis in information science: Eleven approaches – traditional as well as innovative. **Journal of Documentation**, v. 58, n. 4, p. 422– 462, 2002.

HOOIJMAIJERS, D; STUMPTNER, M. Trust based ontology integration for the community services sector. In: AUSTRALASIAN WORKSHOP ON ADVANCES IN ONTOLOGIES, 2., Hobart, Australia, **Proceedings...** Hobart, Australia: Australian Computer Society, 2006. p. 17-24.

JACOB, Elin K. Ontologies and the Semantic Web. **Bulletin of the American Society for Information Science and Technology**, April/May, p. 19-22, 2003.

LOJKINE, Jean. **A revolução informacional**. São Paulo: Cortez, 1995.

LE MOIGNE, Jean-Louis. **A teoria do sistema geral: teoria da modelização**. Lisboa: Instituto Piaget, 1977.

MALHEIROS, Luciana. **A identificação de novas descobertas científicas através da análise do conhecimento contido em artigos científicos**. Projeto de Tese de Doutorado para ingresso no PPGCI UFF/IBICT. Niterói, 2005.

NILES, I., PEASE, A. Towards a Standard Upper Ontology, In: INTERNATIONAL CONFERENCE ON FORMAL ONTOLOGY IN INFORMATION SYSTEMS (FOIS-2001), 2., Ogunquit, Maine, **Proceedings...** Ogunquit, Maine: ACM Press, 2001.

NOY, Natalya .F., MUSEN, Mark .A., SMART: Automated Support for Ontology Merging and Alignment. **Banff Workshop on Knowledge Acquisition, Modeling, and Management**, 1999.

OLIVEIRA, Laura de Lira e. O papel das taxonomias em Ontologias de domínio: contribuição da Ciência da Informação para estruturas taxonômicas. Projeto de Tese de Doutorado para ingresso no PPGCI UFF/IBICT. Niterói, 2007.

OTLET, Paul. **Traité de Documentation: le livre sur le livre. Théorie et pratique.** Liège : Centre de Lecture Publique de la Communauté Française de Belgique, 1989. 432p.

PINTO, Helena S., MARTINS, João P., A Methodology for Ontology Integration. In: **Proceedings of First International Conference on Knowledge Capture, K-CAP 2001**, Victoria, B.C., Canada, ACM Press, 2001

PREDOIU, L.; FEIER, C.; SCHARFFE, F; DE BRUNIN, J.; MARTIN-RECUERDA, F.; MANOV, D.; EHRING, M. State-of-the-art survey on ontology merging and aligning v2. Technical Report D4.2.2, **Digital Enterprise Research Institute**, University of Innsbruck, December 31, 2005.

RANGANATHAN, S. R. **Prolegomena to Library Classification.** New York : Asia Publishing House, 1967.

REICH J.R. Ontological Design Patterns for the Integration of Molecular Biological Information. In: GCB-99 CONFERENCE ON BIOINFORMATICS, Hannover, Germany, 1999, **Proceedings...**1999, p. 156-166.

THE SEMANTIC WEB. **Bulletin of The American Society for Information Science and Technology**, v. 29, n. 4, Apr./May 2003. (Special Section).

SMITH, Barry. **Ontology and information systems.** 2002. Disponível em: <[http://ontology.buffalo.edu/ontology\(PIC\).pdf](http://ontology.buffalo.edu/ontology(PIC).pdf)>. Acesso em 01 ago. 2007.

SMITH, Barry; KOHLER, Jacob; KUMAR, Anand. On the Application of Formal Principles to Life Science Data: a Case Study in the Gene Ontology. **Lecture Notes in Computer Science**, v. 2994, p. 79-94, 2004. Disponível em: <http://ontology.buffalo.edu/medo/Database_Integration.pdf>. Acesso em 01 ago. 2007.

SORGEL, Dagobert. The rise of ontologies or the reinvention of classification. **Journal of the American Society for Information Science**, v.5, n.12, p.1119 – 1120. 2000.

SUO,W.G. Standard Upper Ontology Working Group, **Standard Upper Ontology.** Disponível em: <<http://suo.ieee.org/>>. Acesso em 12 maio 2004.

USCHOLD, M. Creating, integrating and maintaining local and global ontologies. In: WORKSHOP ON APPLICATIONS OF ONTOLOGIES AND PROBLEM-SOLVING METHODS at the 14th European Conference on Artificial Intelligence (ECAI 2000), **Proceedings...** 2000, p. 20-25.

WERTHEIN, Jorge. A sociedade da informação e seus desafios. **Ciência da Informação**, v.29, n.2, p.71-77, 2000.

WILSON, T.D.. The Work of the British Classification Research Group. In: WELLISCH, H. (ed.) **Subject retrieval in the seventies**. Westport, Greewood Publishing Co, 1972. p. 62-71

ZIMMERMANN, Antoine, KRÖTZSCH, Markus, EUZENAT, Jérôme, HITZLER, Pascal, Formalizing Ontology Alignment and its Operations with Category Theory., In: INTERNATIONAL CONFERENCE OF FORMAL ONTOLOGY IN INFORMATION SYSTEMS (FOIS 2006), 4., **Proceedings...** IOS Press, 2006. (Frontiers in artificial intelligence and applications, v. 150).

ⁱhttp://en.wikipedia.org/wiki/Software_agent

ⁱⁱ <http://www.geneontology.org/>

ⁱⁱⁱ <http://www.nlm.nih.gov/pubs/factsheets/umls.html>

^{iv} Grupos “Ontologia e Taxonomia: aspectos teóricos e metodológicos” e “Conhecimento, Informação e Tecnologia da Informação”, ambos incluídos no Diretório de Grupos de Pesquisa do CNPq.

^v<http://fois2004.di.unito.it/>

^{vi} <http://www.ischool.drexel.edu/faculty/hhan/onisw2007/>

^{vii} <http://cmap.ihmc.us/>

^{viii} <http://www.co-ode.org/ontologies/pizza/2007/02/12/>

^{ix} <http://www.nlm.nih.gov/research/umls/>

^x Compromisso, Ontológicos, neste contexto, são acordos de uso de um vocabulário comum de uma forma coerente e consistente (GRUBER, 1993).