

Escola de Ciência da Informação – UFMG  
Disciplina: Introdução as ontologias

Prof. Mauricio Barcellos Almeida

# Ontologias

## Conceituação (mod.1)

Prof. Mauricio B. Almeida  
mba@eci.ufmg.br  
www.eci.ufmg.br/mba

ECI-Escola de Ciência da Informação  
UFMG-Universidade Federal de Minas Gerais

Escola de Ciência da Informação – UFMG  
Disciplina: Introdução as ontologias

Prof. Mauricio Barcellos Almeida

## Roteiro da apresentação

### Representação do conhecimento e da informação

Em Ciência da Informação e em Ciência da Computação.

### Conceituação (I): ontologias como objeto

Breve histórico; ontologia em CC e ontologia em CI; algumas considerações complementares; características básicas, tipos e componentes.

### Conceituação (II): ontologias como processo

Processo de especificação; processo de aquisição de conhecimento; processo de conceitualização; processos de documentação, integração, implementação; processo de avaliação.

Escola de Ciência da Informação – UFMG  
Disciplina: Introdução as ontologias

Prof. Mauricio Barcellos Almeida

## 0) Representação do conhecimento e da informação

### 0-1) Em Ciência da Informação:

- Problemas ocorrem à medida que o volume de documentos a ordenar cresce;
- Usuários não mais se contentam com a organização por grandes classes e exigem informações mais precisas;
- Com o grande volume de material disponível atualmente, o problema se torna crítico;
- Preocupação com a padronização da terminologia usada pelas pessoas para encontrar e classificar a informação.
- Técnicas variadas para organização da informação, com destaque para estruturas baseadas em categorização.

Escola de Ciência da Informação – UFMG  
Disciplina: Introdução as ontologias

Prof. Mauricio Barcellos Almeida

## 0-2) Em Ciência da Computação:

A RC-Representação do Conhecimento se desenvolveu como um ramo da IA-Inteligência Artificial => desenvolvimento de sistemas especialistas;

Ênfase na apreensão e na explicitação do conhecimento que será expresso em sistemas;

Destaque para três aspectos:

- Lógica: fornece a estrutura formal e as regras de inferência;
- Ontologias: definem “os tipos das coisas” e as “coisas” que existem no domínio da aplicação, para que possam ser representadas;
- Técnicas de computação: construção de sistemas.

Escola de Ciência da Informação – UFMG  
Disciplina: Introdução as ontologias

Prof. Mauricio Barcellos Almeida

## Conceituação (I): ontologias como objeto

### I-1) Breve histórico

Termo *ontologia* vem do grego *ontos* = ser e *logos* = estudo;

Introduzido na filosofia para distinguir o estudo do ser, do ser humano em sua essência, do estudo dos outros seres das ciências naturais;

O Dicionário Oxford de Filosofia define

*“[...] o termo derivado da palavra grega que significa ‘ser’, mas usado desde o século XVII para denominar o ramo da metafísica que diz respeito àquilo que existe”*

Dicionário Aurélio traz as seguintes definições:

*“1) Ciência do ser em geral. 2) Parte da metafísica que estuda o ser em geral e suas propriedades transcendentais”.*

Escola de Ciência da Informação – UFMG  
Disciplina: Introdução as ontologias

Prof. Mauricio Barcellos Almeida

## I-2) Ontologia em CC

A partir dos anos 90...

- explosão de publicações e conferências sobre o assunto;
- novo uso do termo, distinto de seu uso na filosofia.

Distinção: (SMITH, 2004, p.22)

*O filósofo-ontologista, a princípio pelo menos, tem apenas uma única meta: estabelecer a verdade sobre a realidade, encontrando resposta para a questão ‘o que existe’. Entretanto, no mundo dos sistemas de informação, uma ontologia é um artefato de software (ou linguagem formal) que tem utilizações específicas em ambientes computacionais.*

O termo é utilizado com esse novo sentido nas pesquisas sobre RC em IA.

Escola de Ciência da Informação – UFMG  
Disciplina: Introdução as ontologias

Prof. Mauricio Barcellos Almeida

Diversas definições....

“uma teoria que considera os tipos de entidades e, especificamente, os tipos de entidades que são admitidos em um sistema lingüístico” (CORAZZON, 2002, p.1)

“Uma ontologia é uma especificação explícita de uma conceitualização. [...] Quando o conhecimento de um domínio é representado por um formalismo declarativo, o conjunto dos objetos e as relações entre eles, passíveis de descrição, são espelhadas no vocabulário representacional com o qual o sistema baseado em conhecimento representa o conhecimento. Assim, pode-se descrever a ontologia de um programa ao definir um conjunto de termos representacionais. Em tal ontologia, as definições associam nomes de entidades do universo do discurso (por exemplo, classes, relações, funções, ou outros objetos) com textos legíveis para pessoas, os quais descrevem os nomes que se deseja representar, e axiomas formais que restringem a interpretação e a formação desses nomes.” (GRUBER, 1993, p.2)

Escola de Ciência da Informação – UFMG  
Disciplina: Introdução as ontologias

Prof. Mauricio Barcellos Almeida

E discussões...

“O principal problema em tal interpretação (Gruber) é que ela se baseia na noção de conceitualização [...] uma conceitualização é um conjunto de relações extensionais que descrevem um estado de coisas particular, enquanto a noção que temos em mente é a intensional, ou seja, algo como uma rede conceitual a qual se sobrepõe a vários estados de coisas possíveis. Propõe-se uma definição revista para a conceitualização, a qual capture aspectos intensionais [...]” (GUARINO e GIARETTA, 1995, p. 27)

O que é *intensão*? O que é *extensão*?

“O conjunto de características agrupadas em uma unidade para formar um conceito é chamado de *intensão*. Os objetos vistos como conjuntos e conceitualizados como um conceito são conhecidos como *extensão*. Os dois, *intensão* e *extensão*, são interdependentes.” (ISO-704, 2000, p 4).

Escola de Ciência da Informação – UFMG  
Disciplina: Introdução as ontologias

Prof. Mauricio Barcellos Almeida

Exemplo de *noção intensional*  
*noção extensional*

... lista de características do conceito a ser definido ...

... enumeração de aspectos de todas as espécies que são do mesmo nível de abstração ....

Escola de Ciência da Informação – UFMG  
Disciplina: Introdução as ontologias

Prof. Mauricio Barcellos Almeida

Definição aceita:

“[...] uma ontologia se refere a um artefato de engenharia (de software), que é constituído por um vocabulário específico utilizado para descrever certa realidade, mais um conjunto de suposições explícitas a respeito do significado pretendido para as palavras do vocabulário. Esse conjunto de suposições tem em geral a forma da teoria da lógica de primeira ordem, onde palavras do vocabulário aparecem com nomes de predicados unários ou binários, respectivamente chamados *conceitos* e *relações*. No caso mais simples, uma ontologia descreve uma hierarquia de conceitos relacionados por relações de classificação; em casos mais sofisticados, axiomas são adicionados à estrutura de forma a expressar outras relações entre conceitos, e para restringir a interpretação pretendida para tais conceitos.” (GUARINO, 1998, p.2)

Escola de Ciência da Informação – UFMG  
Disciplina: Introdução as ontologias

Prof. Mauricio Barcellos Almeida

Uma boa (e simples) definição:

“Uma ontologia é uma especificação formal e explícita de uma conceitualização compartilhada” (Borst, 1997)

Nessa definição...

- “*formal*” significa legível por computadores;
- “*especificação explícita*” diz respeito a conceitos, atributos, relações, restrições e axiomas que são explicitamente definidos;
- “*compartilhado*” quer dizer conhecimento consensual;
- “*conceitualização*” diz respeito a um modelo abstrato de algum fenômeno do mundo real.

Escola de Ciência da Informação – UFMG  
Disciplina: Introdução as ontologias

Prof. Mauricio Barcellos Almeida

I-3) Ontologia em CI

O termo começa a aparecer com mais freqüência na literatura de CI em meados da década de 90. (VICKERY, 1997)

- Estudos que abordam as diferenças e similaridades entre três termos amplamente utilizados – taxonomias, tesouros e ontologias – na pesquisa de Ciência da Informação, da IA e da Web Semântica. (GILCHRIST, 2003)
- Perspectiva da modelização, que compara métodos e técnicas utilizados na atividade de modelar o conhecimento, provenientes da CI, da CC Computação e da Teoria da Terminologia. (CAMPOS, 2004)

Escola de Ciência da Informação – UFMG  
Disciplina: Introdução as ontologias

Prof. Mauricio Barcellos Almeida

Cont. ...

- A variedade de termos surgidos no âmbito da organização do conhecimento – domínios do conhecimento, engenharia do conhecimento, mineração de conhecimento, *Web Semântica*, ontologias – são aplicações de conceitos antigos e conhecidos. (CURRAS, 2004)
- Uma consideração comum entre os autores da CI é a possibilidade da realização de pesquisas multidisciplinares sobre o assunto. Existe proximidade entre os estudos bibliográficos e as pesquisas em ontologias, entretanto as ontologias e os sistemas de classificação da biblioteconomia não têm os mesmos objetivos. (VICKERY, 1997)

Escola de Ciência da Informação – UFMG  
Disciplina: Introdução as ontologias

Prof. Mauricio Barcellos Almeida

#### I-4) Algumas considerações complementares

Existe a possibilidade de que o termo *ontologia* tenha sido apropriado pela IA a partir de seu significado filosófico (o que existe?), aplicando-se a idéia aos computadores.

*Pessoa* - Aquilo é uma jabuticaba.  
*Computador* - O que é uma jabuticaba?  
*Pessoa* - Jabuticaba é uma fruta.  
*Computador* - O que é uma fruta?  
*Pessoa* - Uma fruta nasce na árvore.  
*Computador* - O que é uma árvore?  
*Pessoa* - Uma árvore é um ser vivo?  
*Computador* - O que é "ser"? O que é "vivo"? O que é "ser vivo"? ....

Uma conversa fictícia...

Escola de Ciência da Informação – UFMG  
Disciplina: Introdução as ontologias

Prof. Mauricio Barcellos Almeida

Outro exemplo...

*João: qual foi o último documento que você leu?*  
*Maria: O artigo do professor Ricardo.*

Propósito da ontologia:

Capturar a semântica, as relações e as noções que as pessoas usam e torná-las explícitas, para que possam ser codificados em sistemas, manipulados e intercambiados.

Essa é a idéia básica da Web Semântica e resulta em sistemas mais eficientes.

Uma conversa real...

Escola de Ciência da Informação – UFMG  
Disciplina: Introdução as ontologias

Prof. Mauricio Barcellos Almeida

#### I-5) Características básicas, tipos e componentes

A estrutura das ontologias pode variar, mas alguns componentes estão presentes em grande parte delas:

- *Classes*: representativas de conceitos e organizadas hierarquicamente;
- *Relações*, representativas da interação entre os conceitos de um domínio;
- *Axiomas*, utilizados para modelar sentenças consideradas verdadeiras;
- *Instâncias*, representativas dos objetos que pertencem a uma classe.

Escola de Ciência da Informação – UFMG  
Disciplina: Introdução as ontologias

Prof. Mauricio Barcellos Almeida

Classificação da ontologia, pelo **grau de formalidade utilizado para especificar o vocabulário de termos e seus significados** (USCHOLD e GRUNINGER, 1996)

- Ontologia *altamente informal*, em que o vocabulário é expresso em linguagem natural;
- Ontologia *semi-informal*, em que o vocabulário é expresso em linguagem natural de forma restrita e estruturada;
- Ontologia *semi-formal*, cujo vocabulário é expresso em linguagem artificial definida formalmente;
- Ontologia *rigorosamente formal*, em que os termos são definidos com semântica formal, teoremas e provas.

Escola de Ciência da Informação – UFMG  
Disciplina: Introdução as ontologias

Prof. Mauricio Barcellos Almeida

Classificação das ontologias, quanto ao **tipo de estrutura e ao assunto da conceitualização** (VAN-HEIJST, SCHREIBER e WIELINGA, 2002)

- Ontologias *terminológicas*, que especificam termos usados para representar o conhecimento em um domínio, por exemplo, os léxicos;
- Ontologias de *informação*, que especificam uma estrutura de registros, por exemplo, os esquemas de bancos de dados;
- Ontologias de *modelagem do conhecimento*, que especificam conceitualizações do conhecimento;

Escola de Ciência da Informação – UFMG  
Disciplina: Introdução as ontologias

Prof. Mauricio Barcellos Almeida

Cont. ...

- Ontologias de *aplicação*, que contêm as definições necessárias para modelar o conhecimento em uma aplicação (software);
- Ontologias de *domínio*, que expressam conceitualizações específicas para um domínio do conhecimento;
- Ontologias *genéricas*, similares às ontologias de domínio, mas com conceitos considerados genéricos e comuns a vários campos;
- Ontologias de *representação*, as quais explicam as conceitualizações que suportam os formalismos de representação do conhecimento.

Escola de Ciência da Informação – UFMG  
Disciplina: Introdução as ontologias

Prof. Mauricio Barcellos Almeida

Classificação das ontologias, quanto aos **tipos de classes presentes** (HAAV e LUBI, 2001)

- Ontologias de *alto nível*, que descrevem conceitos gerais como espaço, tempo, matéria, objeto, evento, ação, etc, os quais são independentes do problema ou domínio;
- Ontologias de *domínio*, que descrevem o vocabulário de um domínio, por exemplo, medicina ou automóveis;
- Ontologias de *tarefa*, que descrevem uma tarefa ou atividade, por exemplo, diagnósticos ou compras, através da inserção de termos especializados.

Escola de Ciência da Informação – UFMG  
Disciplina: Introdução as ontologias

Prof. Mauricio Barcellos Almeida

Classificação das ontologias, quanto a sua **função no processo de desenvolvimento de sistemas** (JASPER e USCHOLD, 1999)

- Ontologia de “*autoria neutra*”, que enfatiza a reutilização de dados, possibilitando que um aplicativo seja escrito em uma única linguagem e, depois, convertido para uso em diversos sistemas;
- Ontologia de *especificação*, uma ontologia de domínio usada para documentação e manutenção de softwares;
- Ontologia de *acesso comum à informação*, que torna a informação inteligível quando o domínio é expresso em um vocabulário inacessível.

Escola de Ciência da Informação – UFMG  
Disciplina: Introdução as ontologias

Prof. Mauricio Barcellos Almeida

**Conceitualização (II): ontologias como processo**

Estágios do processo de desenvolvimento de uma ontologia.

Adquirir conhecimento
Documentar
Avaliar

(FERNANDEZ, GÓMEZ-PÉREZ e JURISTO, 1997)

Escola de Ciência da Informação – UFMG  
Disciplina: Introdução as ontologias

Prof. Mauricio Barcellos Almeida

**II-1) Processo de especificação**

O objetivo do processo de **especificação** é produzir um documento de especificação em linguagem natural;

Pelo menos três informações devem estar presentes nesse documento:

1. Propósito da ontologia, ou seja, para quê ela é construída e para que se pretende utilizá-la;
2. Grau de formalidade da ontologia, que depende do nível de formalização utilizado para codificar termos e seus significados;
3. Escopo, que inclui um conjunto de termos a serem representados, características e granularidade.

Escola de Ciência da Informação – UFMG  
Disciplina: Introdução as ontologias

Prof. Mauricio Barcellos Almeida

**Documento de especificação da ontologia**

**Domínio:** química

**Data:** 15 de maio de 1996

**Conceitualizado por:** Assunción Gómez-Pérez

**Implementado por:** Mariano Fernandez-López

**Propósito:**  
Ontologia sobre substâncias químicas, a ser utilizada para obtenção de informações sobre elementos químicos, com fins de ensino e análise. Essa ontologia pode ser utilizada para conferências, por exemplo, o peso atômico do elemento sódio.

**Nível de formalidade:** semi-formal

**Escopo:**  
-Lista de 103 elementos de substâncias: lítio, sódio, cloro ...  
-Lista de conceitos: halogênios, gases nobres, semi-metais, metais ...  
-Pelo menos informações sobre as seguintes propriedades: número atômico, peso atômico, volume atômico a 20 graus Celsius, ponto de ebulição, densidade a 20 graus Celsius ...

**Fontes de conhecimento:**  
Manual de Química e Física, 65ª Edição, CRC Press ...

**Especificação de ontologia química Adaptado de Fernandez, Gomez-Perez e Juristo (1997, p.37)**

Escola de Ciência da Informação – UFMG  
Disciplina: Introdução as ontologias

Prof. Mauricio Barcellos Almeida

### II-2) Processo de aquisição de conhecimento

O processo de **aquisição de conhecimento** é uma atividade independente e simultânea a outras atividades.

Possíveis fontes de conhecimento são: especialistas, livros, manuais, figuras, tabelas, outras ontologias, sistemas informatizados, etc.

Exemplos de técnicas para obter conhecimento são: *brainstorming*, entrevistas não estruturadas, entrevistas estruturadas, análise informal de documentos, análise formal de documentos, etc.

Escola de Ciência da Informação – UFMG  
Disciplina: Introdução as ontologias

Prof. Mauricio Barcellos Almeida

### II-3) Processo de conceitualização

No processo de **conceitualização**, a estrutura do domínio de conhecimento é traduzida em um modelo conceitual.

Esse modelo descreve problemas e soluções relativos ao vocabulário do domínio, identificado na atividade de especificação.

Constrói-se um glossário de termos, incluindo conceitos, verbos, instâncias e propriedades, que busca identificar e reunir o conhecimento útil no domínio.

Na sequência agrupam-se termos, como conceitos e como verbos. Para cada conjunto de conceitos ou de verbos relacionados constrói-se um diagrama.

Escola de Ciência da Informação – UFMG  
Disciplina: Introdução as ontologias

Prof. Mauricio Barcellos Almeida

Em geral, Substantivos representam conceitos...

Em geral, Verbos representam relações...

Atividades da conceitualização  
Adaptado de Fernandez, Gomez-Perez e Juristo (1997, p.6)

Escola de Ciência da Informação – UFMG  
Disciplina: Introdução as ontologias

Prof. Mauricio Barcellos Almeida

A análise do vocabulário de termos reunido deve considerar princípios genéricos e o tratamento de casos específicos.

Os **princípios genéricos** são: (USCHOLD e GRUNINGER, 1996)

- Produção de definições textuais em linguagem natural;
- Utilização de dicionários, tesouros, glossários, etc, de forma a assegurar a consistência da terminologia;
- Indicação de relações como outros termos comumente usados que são similares ao termo em definição;
- Eliminação de "circularidade" na definição de termos;
- Definição de cada termo de forma necessária e suficiente, ou seja, independente de outras definições;
- Fornecimento de exemplos quando necessário.

Escola de Ciência da Informação – UFMG  
Disciplina: Introdução as ontologias

Prof. Mauricio Barcellos Almeida

Os **casos específicos** dizem respeito à manipulação de termos ambíguos:

- Um termo corresponde a apenas uma definição, o que não causa problemas;
- vários termos correspondem a uma definição, o que indica a existência de termos sinônimos;
- Um termo corresponde a vários conceitos (homônimos) o que configura ambigüidade.

Escola de Ciência da Informação – UFMG  
Disciplina: Introdução as ontologias

Prof. Mauricio Barcellos Almeida

O **tratamento dos casos específicos** considera: (USCHOLD e GRUNINGER, 1996)

- Suspensão do uso do termo;
- Esclarecimento da idéia relativa ao conceito, a partir de uma definição cuidadosa, e a utilização do menor número possível de termos técnicos;
- Determinação de quais conceitos são suficientemente importantes, justificando sua participação na ontologia;
- Escolha de um único termo para o conceito.

Escola de Ciência da Informação – UFMG  
Disciplina: Introdução as ontologias

Prof. Mauricio Barcellos Almeida

Finalmente, os termos reunidos são estruturados taxonomicamente, agrupando-se aqueles que apresentam similaridade.

Essa atividade tem raízes em princípios de categorização:

- **Identificação e classificação:** forma pela qual se determina se alguma coisa pertence a uma categoria;
- **Categorização:** os objetos “mesa”, “cadeira” podem ser atribuídos à categoria “móvel”, a partir de um critério;
- **Generalização e especialização:** raciocínio utilizado para se alternar de uma noção para outra, mais geral ou mais específica.

Escola de Ciência da Informação – UFMG  
Disciplina: Introdução as ontologias

Prof. Mauricio Barcellos Almeida

Em relação à estrutura da ontologia, existem diferentes formas de abordar sua construção:

- **bottom-up**, partindo de conceitos mais específicos e construindo uma estrutura por generalização;
- **top-down**, partindo de conceitos mais genéricos e construindo um estrutura por especialização;
- **middle-out**, identificando conceitos nucleares do domínio, e partindo para construir a estrutura através de sua especialização e generalização, simultaneamente.

Escola de Ciência da Informação – UFMG  
Disciplina: Introdução as ontologias

Prof. Mauricio Barcellos Almeida

#### II-4) Processos de documentação, integração, implementação

O processo de **documentação** não é consensual e não existe metodologia para documentar a construção de ontologias.

Recomenda-se obter um documento após cada fase prevista no ciclo de vida.


No processo de **integração** reutilizam-se conceitos já existentes em outras ontologias.

Consiste na inspeção de ontologias de alto nível e na busca em bibliotecas de ontologias existentes.

O resultado desse processo é um *documento de integração*.

Escola de Ciência da Informação – UFMG  
Disciplina: Introdução as ontologias

Prof. Mauricio Barcellos Almeida



Documento de integração; adaptado de Fernandez, Gomez-Perez e Juristo (1997, p.6)

O processo de **implementação** requer a utilização de ambientes para a construção de ontologias.

O resultado da fase de implementação é a ontologia codificada em uma linguagem formal.

Nessa etapa, são escolhidas a ferramenta e a linguagem para construção da ontologia.

Escola de Ciência da Informação – UFMG  
Disciplina: Introdução as ontologias

Prof. Mauricio Barcellos Almeida

#### II-5) Processo de avaliação

O processo de **avaliação** diz respeito a julgar tecnicamente a ontologia. É dividido em verificação e validação:

- **Verificação:** consiste no processo técnico que garante a consistência da ontologia, do ambiente de software associado e da documentação ao longo do ciclo de vida da ontologia.
- **Validação:** garante que a ontologia, o software e a documentação correspondem ao sistema planejado.

Considerações sobre a avaliação do conteúdo da ontologia (vide ALMEIDA, 2006)

\* \* \*