

# XIV Encontro Nacional de Pesquisa em Ciência da Informação 2013

## GT 2: Organização e Representação do Conhecimento



Mauricio B. Almeida

Fabricio M. Mendonça

### **Interfaces entre ontologias e conceitos seminais da Ciência da Informação: em busca de avanços na Organização do Conhecimento**

Universidade Federal de Santa Catarina  
Florianópolis, outubro de 2013

# Introdução

---

- ▶ O termo “ontologia” em geral está associado à abordagens computacionais
- ▶ Nos anos 90, o termo foi interpretado como representativo de um conjunto de artefatos da Inteligência Artificial
- ▶ Desde então, observam-se duas correntes:
  - ▶ Ontologia especificada através de linguagem de representação computacional, em geral, para uso na Web Semântica
  - ▶ Ontologia como o estudo de princípios filosóficos úteis para organização do conhecimento (Ontologia Aplicada)
- ▶ Os princípios da Ontologia Aplicada existiam muito antes do surgimento do primeiro computador
- ▶ Esses princípios podem auxiliar na construção de instrumentos, computacionais ou não, para organização do conhecimento

# Estrutura do artigo

---

## ▶ Premissa:

- ▶ A Ontologia Aplicada pode ser de interesse para a CI em questões relacionadas a teoria da classificação

## ▶ Desenvolvimento:

- ▶ Distinguem-se dois tipos de ontologias: *ontologias genéricas* e *ontologias de domínio* (Parte 1)
- ▶ Relaciona-se ontologia à conceitos seminais da CI: tesouros, linguagens documentárias, documentos, assuntos (Parte 2)

## ▶ Conclusões e contribuições:

- ▶ A pesquisa em ontologias pode fornecer caracterização complementar para importantes conceitos da CI
- ▶ A pesquisa em ontologias vai muito além da computação
- ▶ Ontologias como instrumento *per si* podem representar apenas uma pequena parcela do cenário abordado pela CI

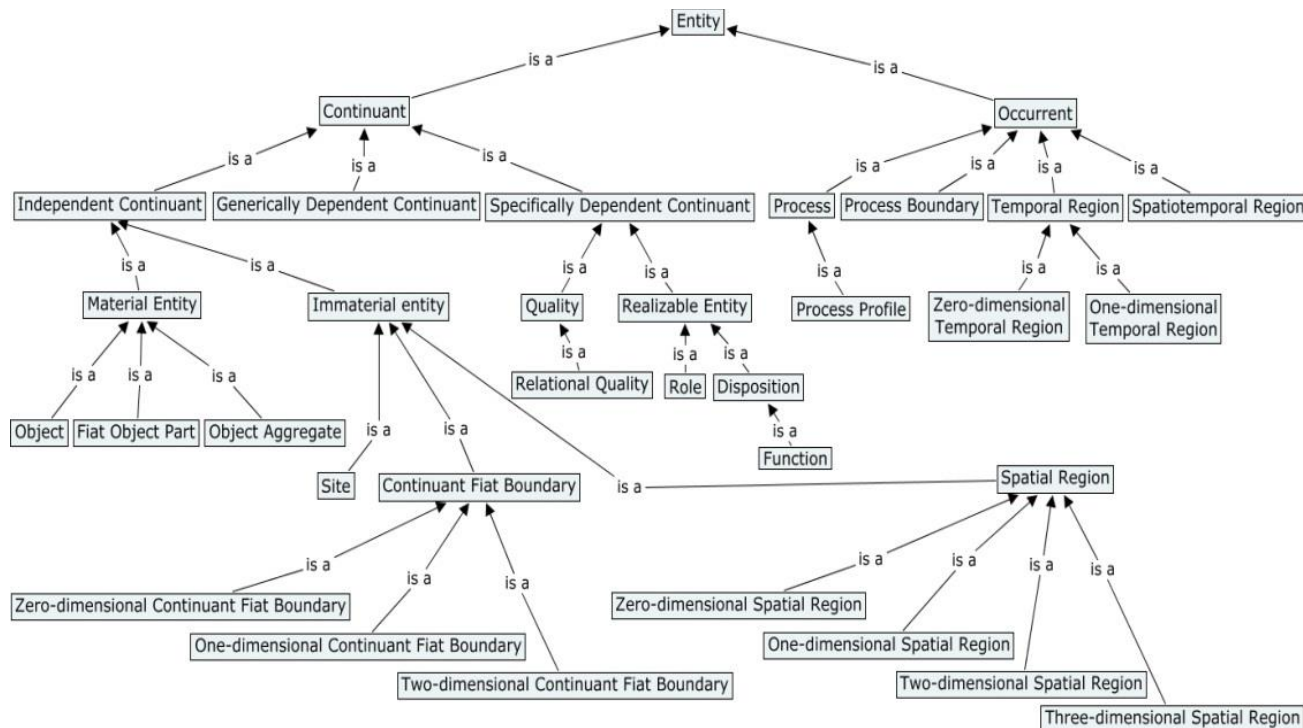
# Parte 1 – tipos de ontologias

---

- ▶ Ontologias por critério de generalidade
- ▶ Ontologias de alto nível
  - ▶ Baseadas em sistemas de categorias filosóficos
  - ▶ Simplificam esses sistemas para fins de automação
  - ▶ Exemplos:
    - ▶ DOLCE (*Descriptive Ontology for Linguistic and Cognitive Engineering*)
    - ▶ BFO (*Basic Formal Ontology*)
- ▶ Ontologia de nível médio
  - ▶ Incluem entidades abaixo das entidades de alto nível, mas ainda acima de entidades específicas de um domínio
  - ▶ Exemplos:
    - ▶ IAO (*Information Artifact Ontology*)
    - ▶ BIO-TOP (*Biological Top-Level*)

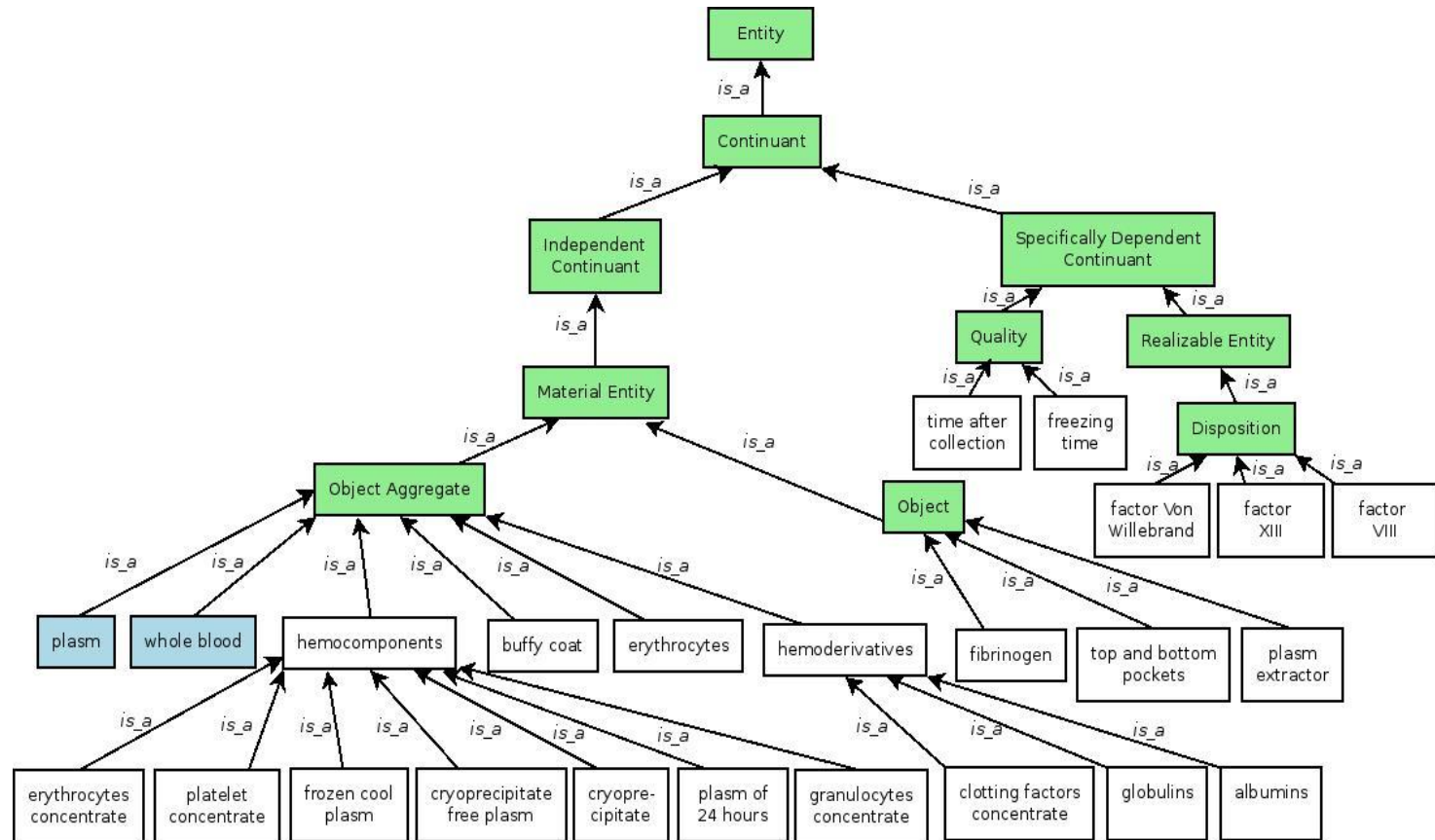
# Parte 1 – Aplicações de ontologias genéricas

- ▶ A ontologia de alto nível fornece um ponto de partida para entendimento das entidades do mundo
- ▶ A ontologia de nível médio fornece uma camada de transição entre entidades genéricas e de domínio



● Estrutura da BFO  
(Basic Formal Ontology)

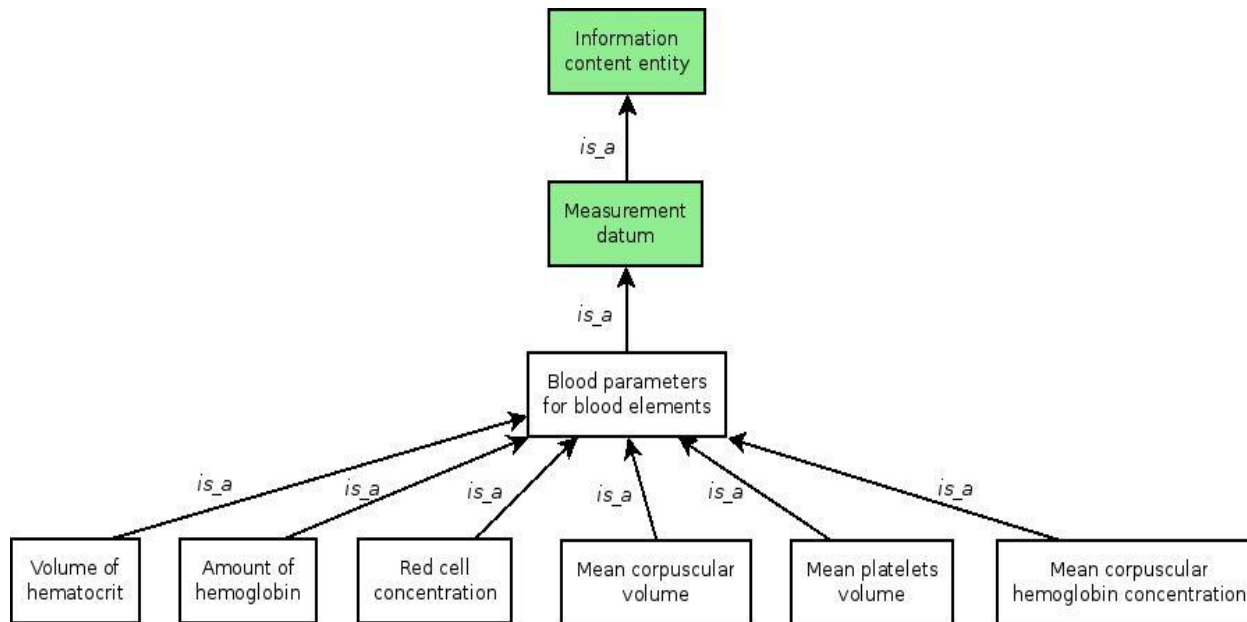
# Parte 1 – ontologias de domínio



Fragmento da estrutura da *Blood Ontology*, construída abaixo da BFO

# Parte 1 – ontologias de domínio

---



Fragmento da estrutura da *Blood Ontology*, construída abaixo da IAO



# Parte 2 – Ontologias e conceitos da CI

---

- ▶ **Questão 1a:** *ontologias são linguagens documentárias?*
  - ▶ Não, ontologias não são rigorosamente linguagens documentárias
- ▶ **Evidências:**
  - ▶ Ontologias são teorias científicas formais, não linguagens
  - ▶ **Citações:**
    - ▶ Johansson (2006): *“This book is a book about the world. I am concerned with ontology, not merely with language”*
    - ▶ Simons (1987): *“[...] ontological claims did not arise via a slavish adherence to the surface properties of language. At best, natural language can serve as one clue for the ontologist”*
  - ▶ No sistema aristotélico, base para algumas ontologias de alto nível, as categorias são categorias de coisas e não de elementos da linguagem
  - ▶ Relações importantes para as linguagens documentárias (por ex. sinonímia) não são importantes para as ontologias
  - ▶ O refinamento de certas relações em ontologias (por ex. parte-todo) não é importante para as linguagens documentárias



# Parte 2 – Ontologias e conceitos da CI

---

- ▶ **Questão 1b:** *se ontologias não são rigorosamente linguagens documentárias, elas podem ser usadas como tal?*
  - ▶ Sim, ontologias podem ser usadas como linguagens documentárias
- ▶ **Evidências:**
  - ▶ Nesse caso, o termo “ontologia” diz respeito ao uso de linguagens computacionais (RDFS, OWL, etc) para representação em máquinas
  - ▶ O resultado não é uma ontologia, é uma linguagem documentária especificada em RDFS, OWL, ...
  - ▶ A linguagem documentária assim construída pode ser implementada em um sistema de recuperação da informação Web
  - ▶ Nesse caso, todos os princípios das linguagens documentárias se mantêm

# Parte 2 – Ontologias e conceitos da CI

---

- ▶ **Questão 2a:** *Ontologias e tesouros exibem a mesma capacidade de representação?*
  - ▶ Tesouros e ontologias são KOS que tem em comum a capacidade de restringir a linguagem natural, em diferentes níveis de formalidade
  - ▶ Os diferentes objetivos desses KOS lançam dúvida sobre a questão de se existe uma representação melhor ou pior
  - ▶ Apesar dos vários apelos de que ontologias seriam mais eficientes em função da possibilidade automatização, não existe pesquisa empírica que comprove a hipótese
  - ▶ A semântica das ontologias é a **semântica formal**: nesse contexto, saber o significado de uma sentença significa saber suas **condições verdade**, o que é diferente do seu **valor verdade**

## Parte 2 – Ontologias e conceitos da CI

---

- ▶ **Questão 2b:** *se as ontologias e os tesouros exibem a mesma capacidade de representação, onde residem as principais diferenças?*
- ▶ **Evidências:**
  - ▶ Exemplo do MeSH (*Medical Subject Headings*)
    - ▶ *Fetal\_Blood narrower-than Blood*
    - ▶ *Plasma narrower-than Blood.*
  - ▶ Do pto de vista da função do tesouros, as relações funcionam para RI: uma consulta usando o termo “*blood*” retorna tanto artigos sobre *fetal\_blood*, quanto artigos sobre *blood\_plasma*.
  - ▶ Do pto de vista ontológico, as relações são de diferentes naturezas: *fetal\_blood é-um-tipo* de *blood*, e *plasma é-parte-de* *blood*.

## Parte 2 – Ontologias e conceitos da CI


---

- ▶ **Questão 3:** *se ontologias genéricas são uma descrição de mundo, [...] em que local do mundo as ontologias localizam entidades relevantes na Organização do Conhecimento, como por exemplo, os documentos [...], o conteúdo dos documentos? Como essas entidades são descritas no âmbito de uma ontologia?*
  - ▶ *Responder a essa pergunta demanda desenvolvimento que não é possível descrever em apresentação de alguns minutos*
  - ▶ *A explicação está disponível no artigo*
  - ▶ *Apresentam-se apenas os resultados do desenvolvimento*

# Parte 2 – Ontologias e conceitos da CI

---

Taxonomia (relações é-um), partindo de “coisa” até alcançar “publicação científica”

Nível	Entidades da ontologia
1	Coisa ( <i>thing</i> )
2	Entidade ( <i>entity</i> )
3	Continuante ( <i>continuant</i> )
4	Continuante Dependente ( <i>dependent continuant</i> )
5	Continuante genericamente dependente ( <i>generically dependent continuant</i> )
6	Entidade de conteúdo informacional ( <i>information content entity</i> )
7	 Documento ( <i>document</i> )
8	Publicação ( <i>publication</i> )
9	Publicação científica ( <i>scientific publication</i> )

Fonte: elaborado pelo autor

 *documento* representado na ontologia

# Parte 2 – Ontologias e conceitos da CI

---

Taxonomia (relações é-um) partindo de “coisa” até alcançar “conteúdo de documento”

Entidades da ontologia

Coisa (*thing*)

...

Continuantes dependentes

...

Continuante genericamente dependente (*generically dependent continuant*)

Entidade de conteúdo informacional (*information content entity*)

Documento (*document*)

...

Entidades puramente intencionais (*purely intentional entities*)

→ Conteúdo de documentos

...

Fonte: elaborado pelo autor



conteúdo de documento representado na ontologia



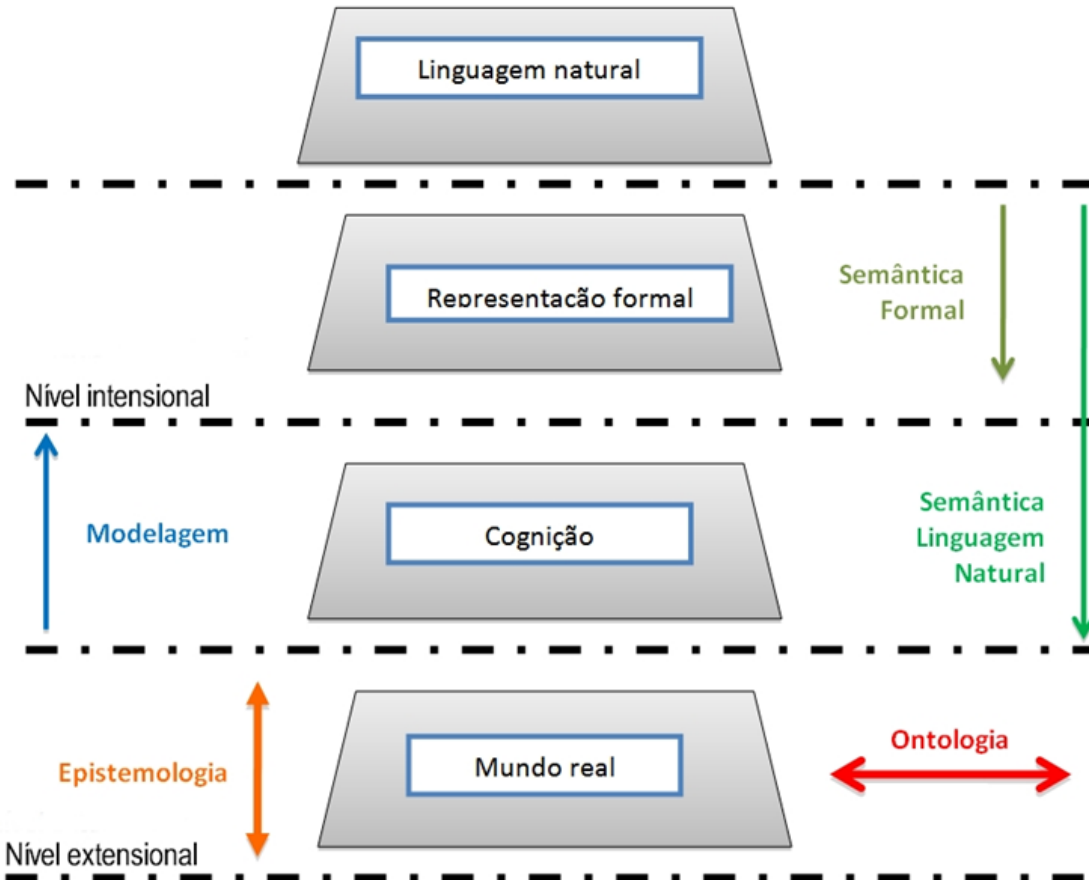
# Considerações finais

---

- ▶ A pesquisa em ontologias vai muito além da implementação computacional
- ▶ Pesquisar ontologias não significa substituir avanços já obtidos no campo da Organização do Conhecimento, apenas complementá-los
- ▶ A ontologia em si não é o principal interesse para a Organização do Conhecimento, mas sim os princípios filosóficos subjacentes
- ▶ Diversos campos científicos estão relacionados quando se aceita o desafio de representar a informação e o conhecimento
- ▶ A CI aborda a maioria desses campos em suas pesquisas interdisciplinares

# Considerações finais

## Níveis e inter-relações entre campos de pesquisa



Fonte: adaptado de Almeida (2011).



# Considerações finais

---

**A CI pode lidar com questões ...**

- epistemológicas
- psicológicas
- cognitivas
- lógicas
- culturais
- históricas

...

**Ontologias podem lidar com apenas uma pequena parcela do rico cenário abordado pela CI**

mas, automatizar pode trazer benefícios...