Connecting worlds:

Semantic Challenges, Interoperability and Domain Ontologies

Fernanda Farinelli

fernanda.farinelli@unb.br

https://www.linkedin.com/in/fernandafarinelli

UnB

FCI Faculdade de CIÊNCIA DA INFORMAÇÃO

Road map

- Connecting Worlds
- Interoperability remarks
- Semantic challenges
- Ontologies overview
- Ontologies and Semantic Interoperability
- Interoperability between Ontologies

10:20:13

Connecting worlds: Brief contextualization

- **Big Data** (Volume, Variety e Velocity)

- **Linked data** (Semantic web, Open Data).

- **Data Analysis** (detect patterns, predict trends, generate insights, and make decisions based on data or evidences).

- **Artificial Intelligence (AI)**: computer science field dedicated to creating systems that simulate human reasoning, learning and perception capabilities.

- **Need for knowledge organization:** approaches and techniques for efficient treatment, representation, retrieval and interpretation of data, information and knowledge.

10:20:13

Challenges for AI

Think Like Humans: "The exciting effort to make computers think...machines with minds, in the broadest literal sense." (Haugland, 1985)

Thinking Rationally: "The study of the calculations that make it possible to perceive, reason and act." (Winston, 1992)

Acting Like Human Beings: "The art of creating machines that perform functions that require intelligence when performed by people." (Kurzweil, 1990)

Act Rationally: "Computational Intelligence is the study of the design of intelligent agents." (Poole et al., 1998)

Source: S. Russell, P. Norvig: Artificial Intelligence. A Modern Approach. 3rd Edition, 2016.

10:20:13

Interoperability

- Interoperability refers to the **ability of various systems and organizations to work together** (interoperate) in order to ensure that people, organizations, and computer systems **interact to exchange information** in an effective and efficient manner.

Source: https://www.gov.br/governodigital/pt-br/governanca-de-dados/interoperabilidade.

10:20:13

Differences in operating systems, hardware and software platforms or lack of communication standards.

Language differences and sociocultural norms impact how communication is understood and how data is interpreted uniformly.

Semantic Challenges

Obstacles to interoperability

Laws and policies may impose restrictions on the free exchange of information.

Differences in organizational cultures or business processes can impact the clarity and interpretation of communications and data.

10:20:13
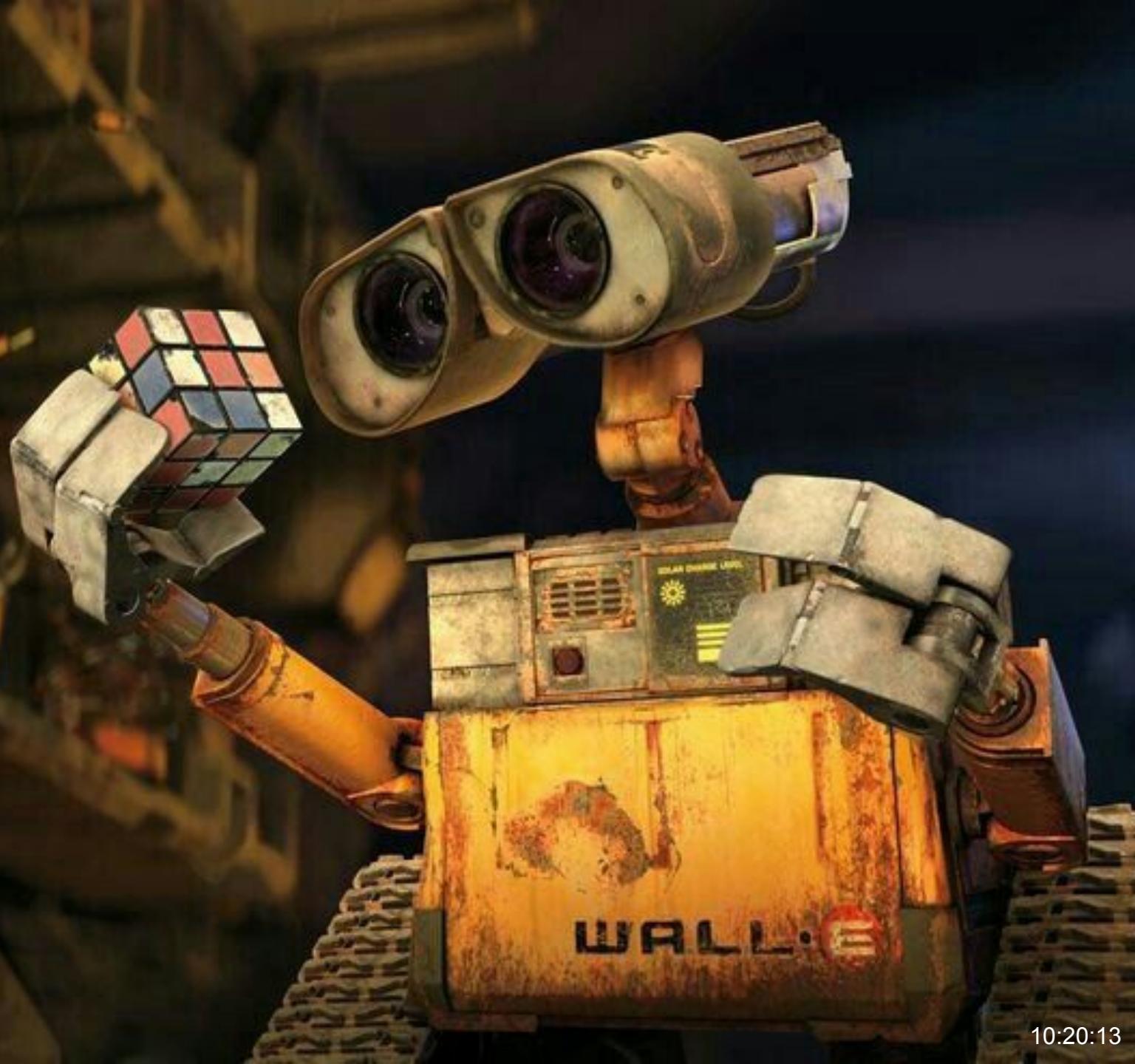
Semantic Interoperability

- Different systems and platforms commonly utilize unique formats and vocabularies for data representation and storage.

  - This hinders efficient data exchange and usage among systems not originally designed to interact.

- Semantic interoperability is the **ability to ensure** that **the meaning of information is understood and interpreted consistently across different systems and organizations**.

10:20:13

Semantic challenges

Ability to understand and interpret meaning.

10:20:13

Concept of semantics

Semantics is the branch of linguistics that **studies the meaning of languages** (words, phrases, sentences, and texts).

Explores **how humans derive meaning from language**, including how words relate to each other, how sentence structures convey different meanings, and how context influences interpretation.
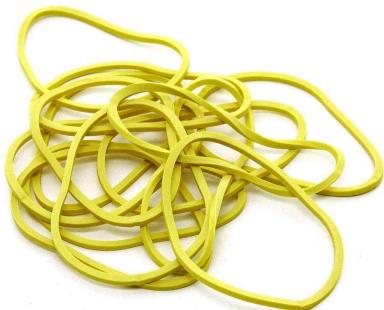
Covers various aspects of meaning, such as literal versus figurative language, semantic roles in sentences, and the changes in meaning over time or across different cultures.

In the semantic web, refers to the use of standards and technologies to enhance the ability of machines to understand and process the meaning of data available on the web.

10:20:13

Common Semantic Challenges

Synonymy and Antonymy

Elástico
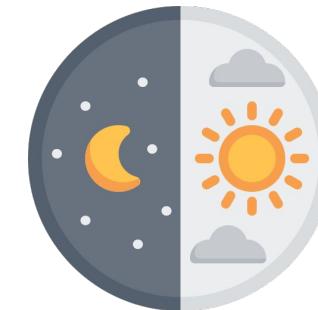
Liguinha (DF), Gominha (MG)

*(in English Rubber bands)*

Mandioca, Macaxeira, Aipim

(in English: cassava/manioc)

Noite / Night   **Dia / Day**

Limpo / Clean   **Sujo / Dirty**

Common Semantic Challenges

Polyssemy and Homonymy

(cruise, southern cruise constellation, Cruise is a soccuer club)

Cruzeiro

Cestas = Baskets

Sestas = Siestas

Sextas = Fridays

**Banco**
bank/stool
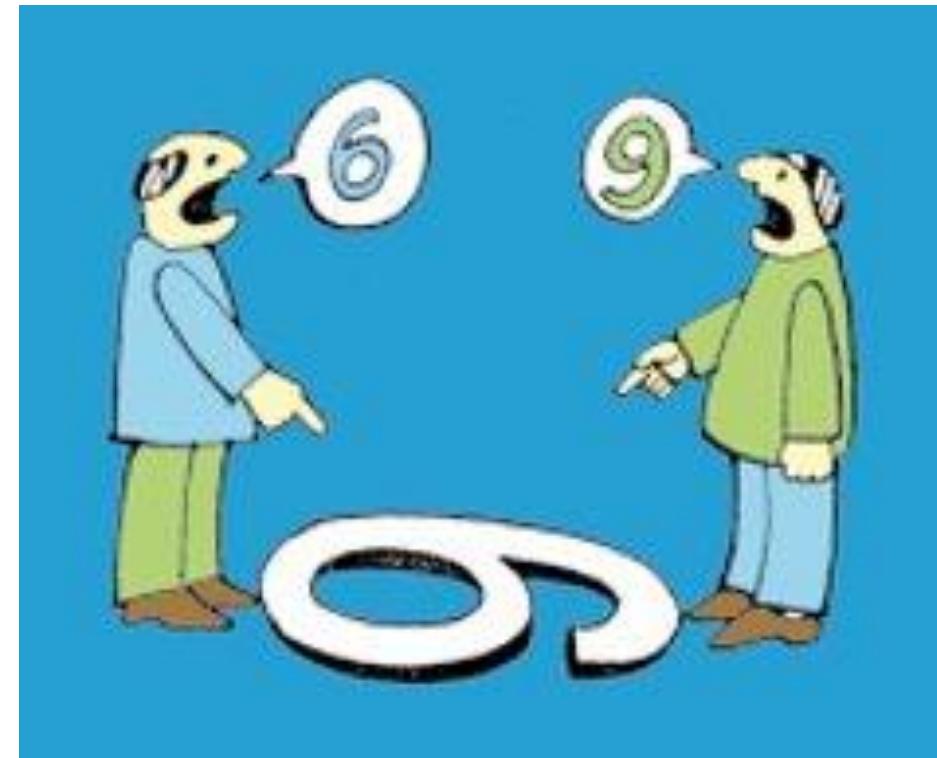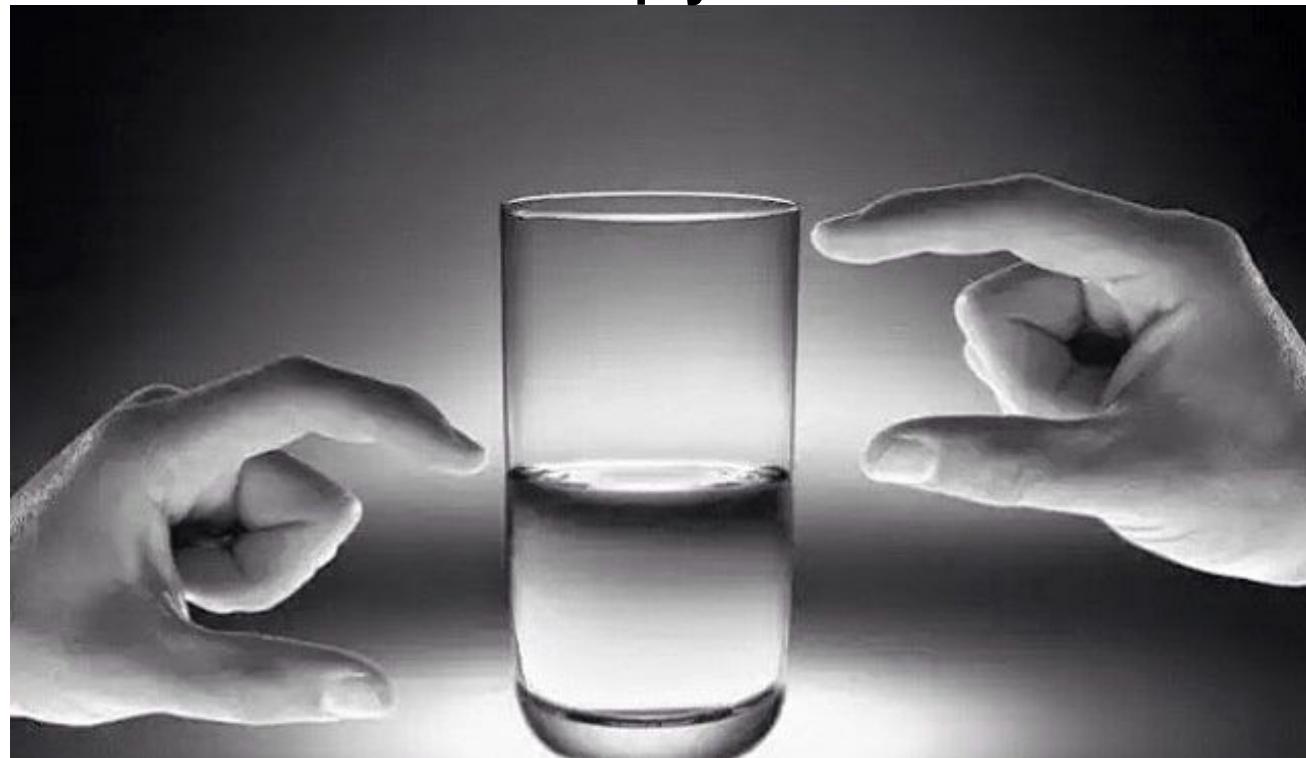
Colher
To harvest/Spoon

Common Semantic Challenges

Perspective

How do you see the glass, half full or half empty?

10:20:13
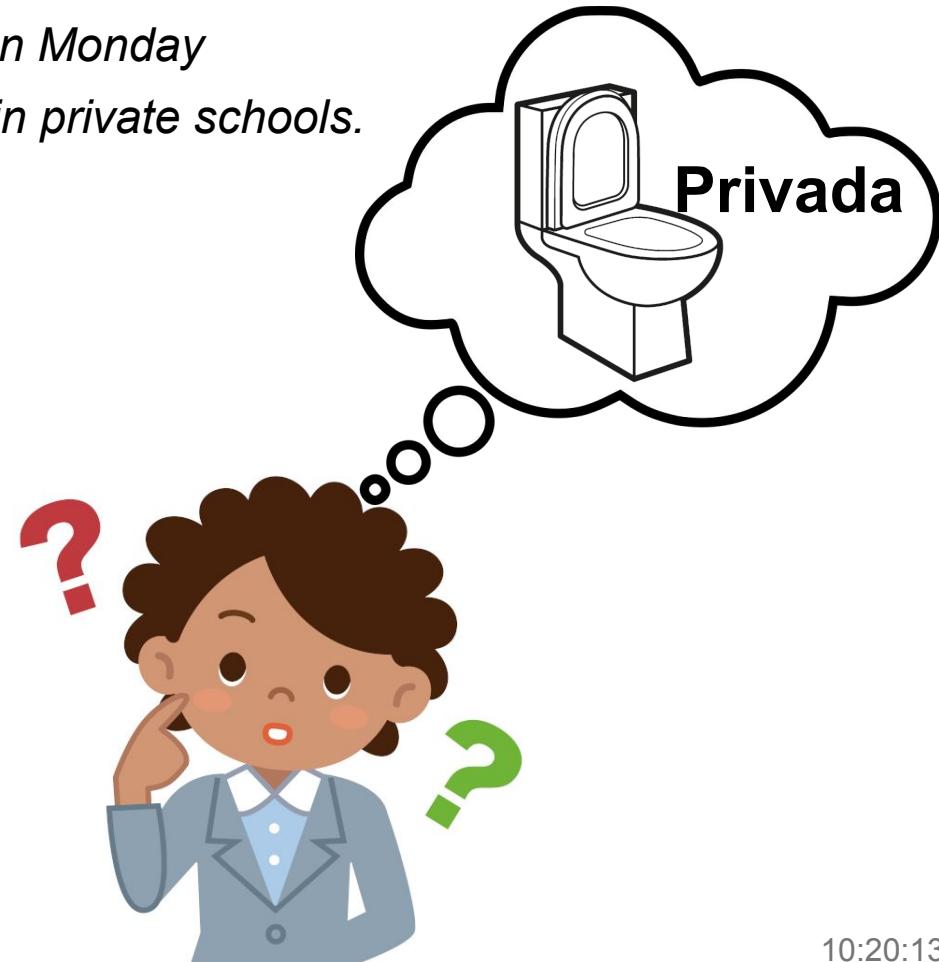
Common Semantic Challenges

Context

Literal translation: Most private should open Monday

Context: This news indicates the end of the strike in private schools.

Privada

ESCOLAS INFANTIS

MAIORIA DAS PRIVADAS DEVE REABRIR SEGUNDA

Sindicato aponta que 80% dos estabelecimentos vão abrir as portas, na próxima segunda-feira, com 70% de adesão dos pais

10:20:13
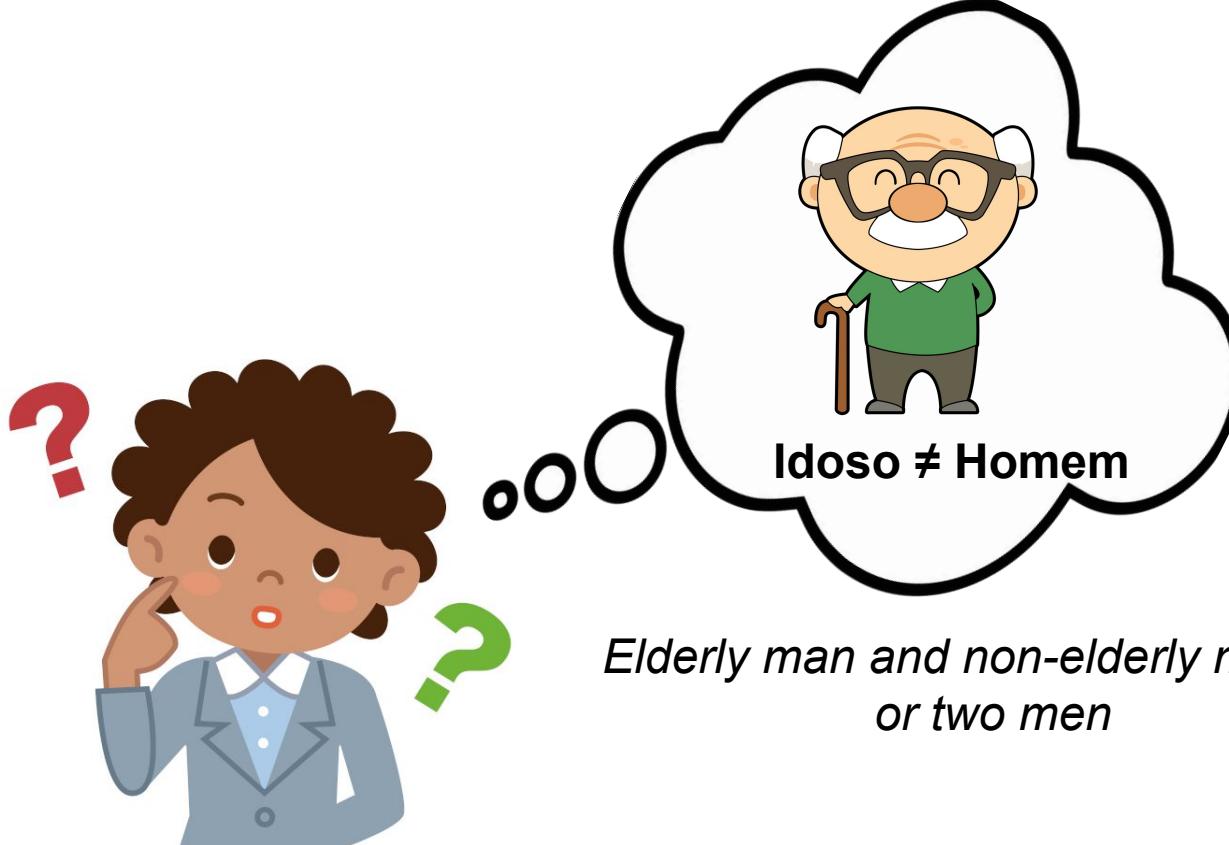
Common Semantic Challenges

Ambiguity and Inaccuracy

Literal translation: Elderly man and man die in accident involving motorcycle on Avenida Vilarinho, in BH

Idoso ≠ Homem

Elderly man and non-elderly man die or two men

14   10:20:13

Common Semantic Challenges

Ambiguity and Equivocation

1. **Vendo o pôr do sol** - In this interpretation, "vendo" is understood as the act of observing, and "pôr do sol" refers to the daily event of the sunset. The phrase means that the person is watching the sunset.

2. **Vendo o por do sol** - Here, "vendo" could be interpreted as the act of selling, and "por do sol" could be a pun involving some product or item, although it doesn't make literal sense. This interpretation plays with the idea of selling something that is a natural phenomenon and not marketable, creating a humorous or poetic effect.

10:20:13

Common Semantic Challenges

Figurative language

- In the first part, the character claims to have woken up before the sun rose, presumably referring to the dawn. However, the other character responds by taking the statement in a literal sense, referring to the sun's age, which is billions of years old, thus creating a pun between the daily "rising" of the sun and the cosmic "birth" of the sun billions of years ago.

10:20:13

Question ?

How can ontologies act in the search for semantic interoperability?

10:20:13

Concept of Ontology

Ontology as Discipline

- Origin in Aristotle's Metaphysics
- Study of the nature of being and relations of existence.
  - Being or Entity is everything that is.
  - Seeks to understand being as such.
- Aims the organization of reality and tries to find answers to:
  - What is reality?
  - What does it mean to exist? What is being?
  - What is the nature of being itself?
  - What are the basic categories of being?

Ontology as Artifact

- A model that formally describes a given domain of knowledge.
  - Knowledge organization system.
- Artifact for organizing and representing information and knowledge.
  - Formal structure that specifies how concepts are organized and interrelated in some domain of knowledge.
  - Represents concepts and relationships between concepts.
- It aims at uniform semantic understanding, through an unambiguous knowledge base and possible human and machine understanding.

(Based on ALMEIDA,
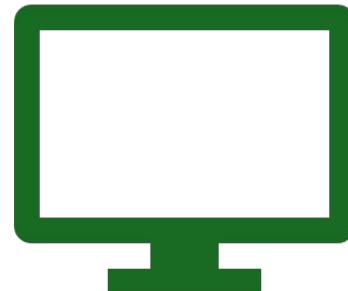
Applied Ontology

- A field of knowledge in Information Science (IS) and Computer Science (CS) that brings together the two aforementioned meanings.

  - It makes use of metaphysical principles from ontology as a discipline to construct ontologies as artifacts.


Ontology

as discipline

drive

Ontology

as Artifact

(Based on ALMEIDA, 2020; SMITH,

Ontology classification

low	**Top-Level ontologies**	• Domain neutral ontologies that deal with very general and basic terms. Example: Process, Entity, Quality.
domain coupling	**Reference ontologies**	• Domain ontologies that represent terms in a given general domain. These terms or entities could be used across specific domains or sub-domains of the general domain. Example: Organization, Document, Disease, Virus.
high	**Domain ontologies**	• Represent terms in a given a specific domain, which usually are particular to this domain. Example: Hypertension, Fertilization, Zika virus.

20

Source: (Farinelli, 2017, p. 44)

# Basic Formal Ontology version 2.0 – Top-level ontology

- Basic Formal Ontology (BFO) is a top-level ontology (TLO) conforming to ISO/IEC 21838-1.
- ISO/IEC 21838-1 (specifies required characteristics of a domain-neutral top-level ontology).
- More Information: https://basic-formal-ontology.org/

10:20:13
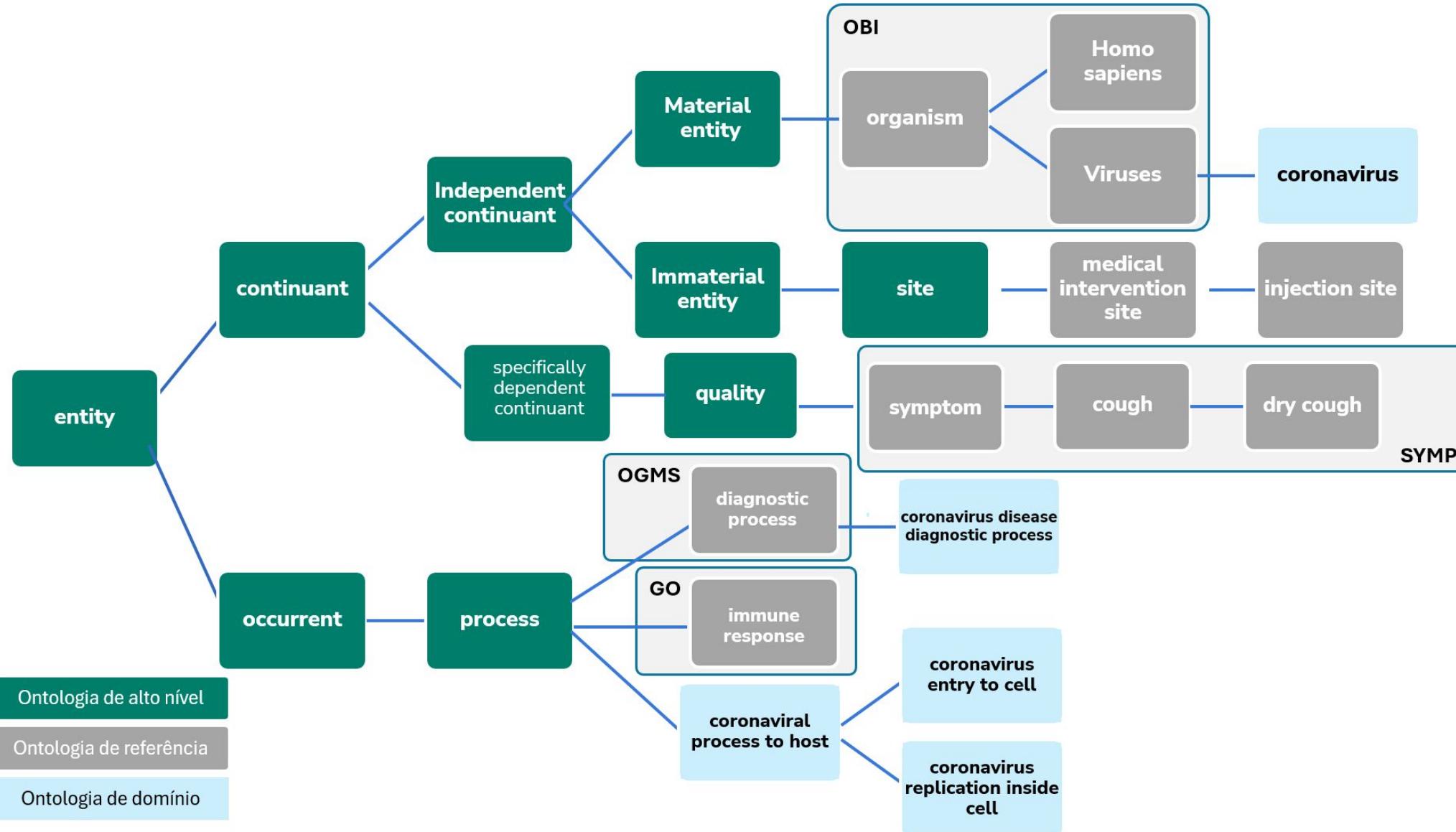
Ontology ecosystem example - OBO Foundry* repository and BioPortal**

Basic Formal Ontology (BFO)

Information Artifact Ontology (IAO)		Ontology for General Medical Science (OGMS)	Ontology for Biomedical Investigations (OBI)	
Foundational Model of Anatomy (FMA)	Gene Ontology (GO) - biological process	Human Disease Ontology (DOID)	Vaccine Ontology (VO)	Ontology of Medically Related Social Entities (OMRSE)
Cell Ontology (CL)	Ontology of Biological Attributes (OBA)	Symptom Ontology (SYMP)	Ontology of Adverse Events (OAE)	Ontology of Document Acts (d-acts)
Human phenotype ontology (HP)	Clinical measurement ontology (CMO)	Cardiovascular Disease Ontology(CVDO)	Ontology of Vaccine Adverse Events (OVAE)	Informed Consent Ontology (ICO)
Phenotypic quality (PaTO)	Biological Spatial Ontology (BSPO)	Infectious Disease Ontology (IDO)	Vaccination Informed Consent Ontology (VICO)	Ontology for Newborn Screening Follow-up and Translational Research (ONSTR)

* https://obofoundry.org/   ** https://bioportal.bioontology.org/   10:20:13
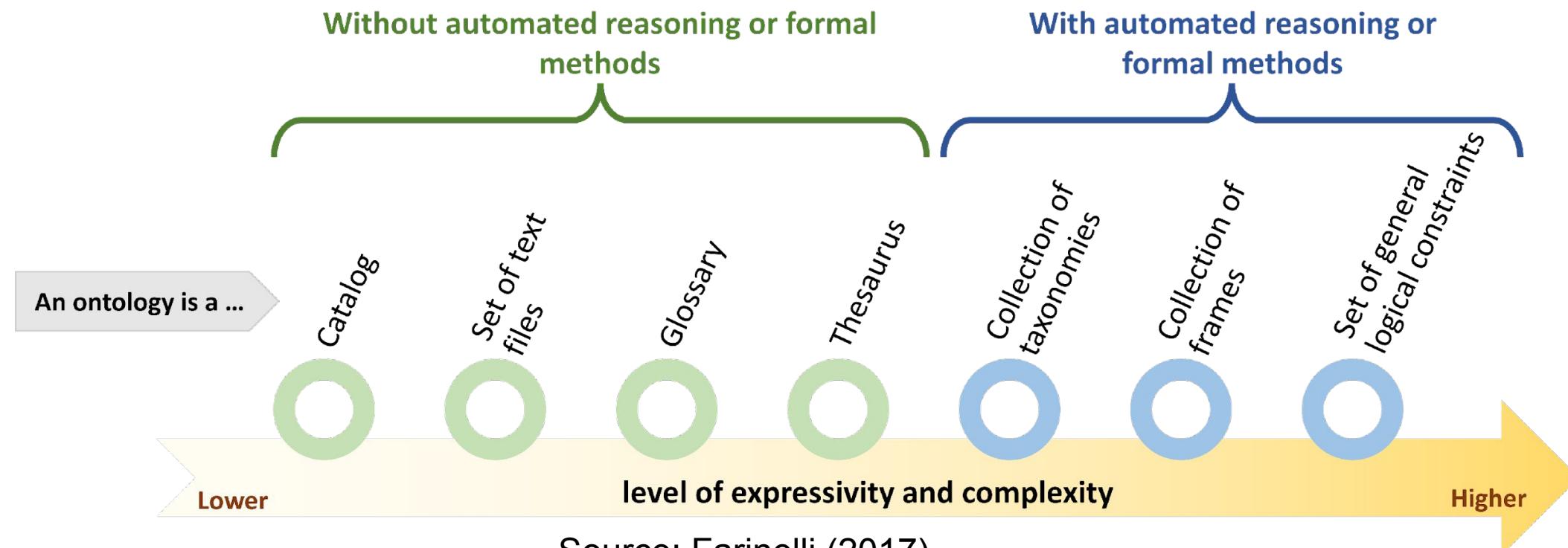
# Example of the practical use of ontology classification

Ontology for Biomedical Investigations Foundry (OBO Foundry)

OBI
- organism
  - Homo sapiens
  - Viruses — coronavirus

Material entity — organism

Independent continuant

continuant

entity

Immaterial entity — site — medical intervention site — injection site

specifically dependent continuant — quality

SYMP
- symptom — cough — dry cough

OGMS
- diagnostic process — coronavirus disease diagnostic process

GO
- immune response

occurrent — **process**

coronaviral process to host
- coronavirus entry to cell
- coronavirus replication inside cell

Legend:
- Ontologia de alto nível
- Ontologia de referência
- Ontologia de domínio

23

10:20:13

- Ontologies can be constructed at different levels of detail or formality.

- In formal ontologies the relationships between vocabulary terms (entities) are defined in formal logical language (axioms).

- This approach allows logical reasoning, which can automate the deduction of new facts or the verification of stated facts within the given domain.

  - Logical axioms enable certain interpretations while explicitly rejecting others.

Without automated reasoning or formal methods

With automated reasoning or formal methods

An ontology is a ...

Catalog

Set of text files

Glossary

Thesaurus

Collection of taxonomies

Collection of frames

Set of general logical constraints

Lower

level of expressivity and complexity

Higher

Source: Farinelli (2017)

10:20:13

Ontological Issues
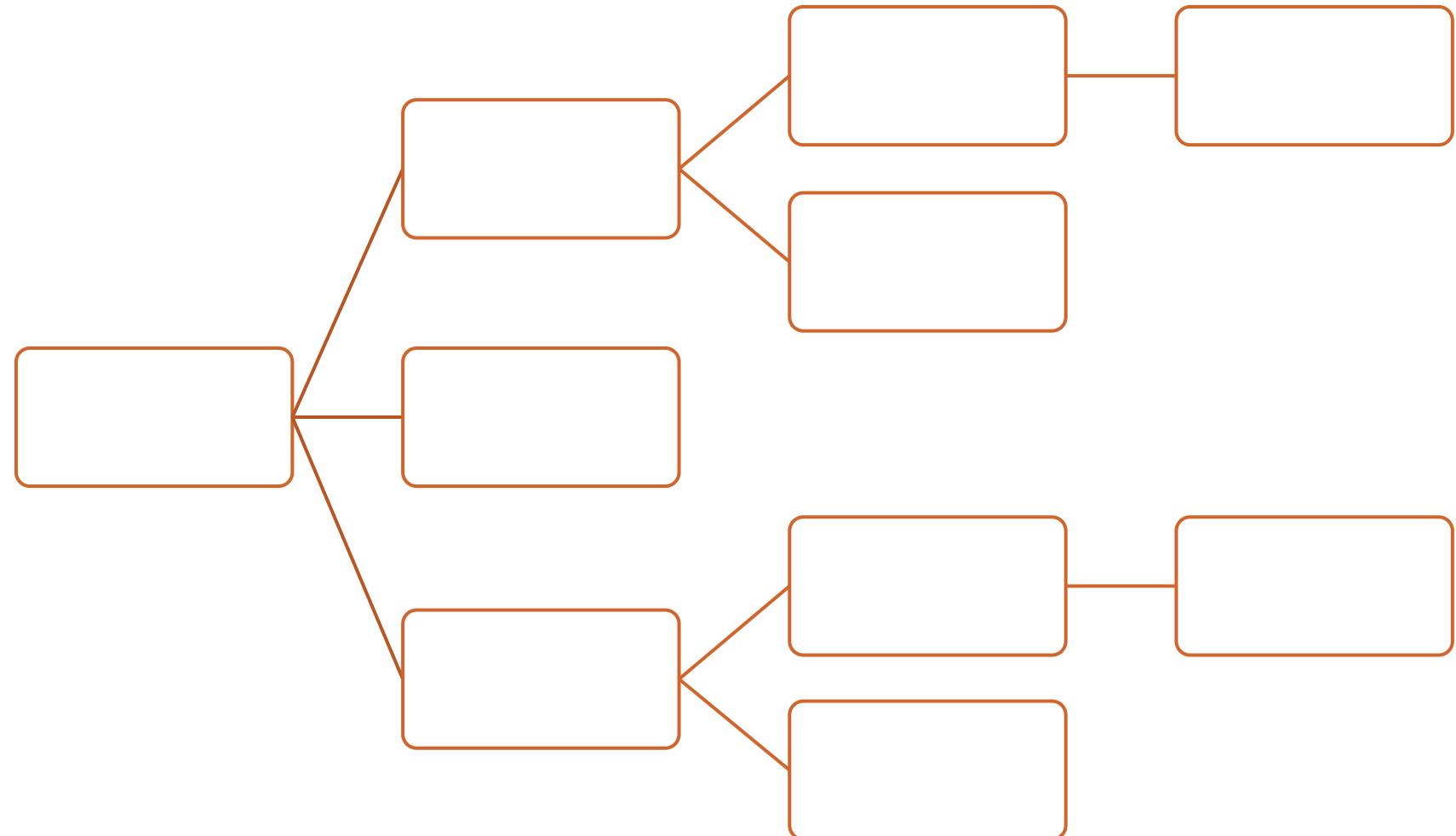
Aren't snakes animals?

Source: Slides by Mathias Brochhausen in the lecture given at the ICBO and ONTOBRAS 2023 Joint Conference.

10:20:13
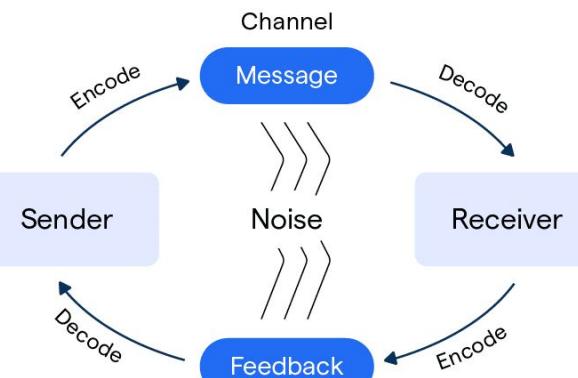
Ontological response – Animals Taxonomy (very basic draft)

10:20:13

How can ontologies act in the search for semantic interoperability?

- First, we need to understand the **communication process** or information exchange.
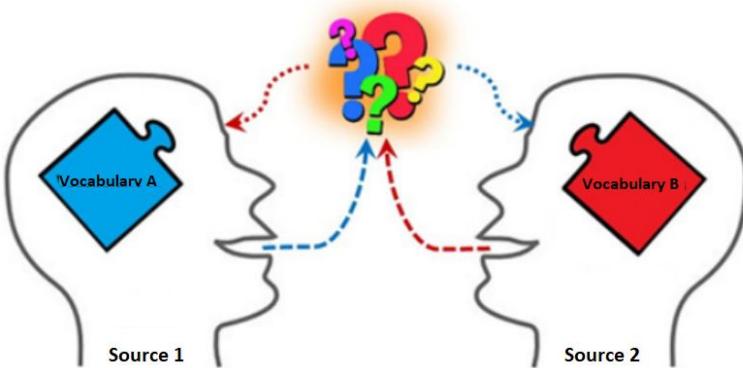
- Communication through the ***exchange of information between agents*** (sender and receiver).
- ***Information*** is communicated in a ***language*** (a set of organized symbols).
- Alone, these symbols have no meaning, the ***agents*** involved ***know how they should be interpreted***

(SHANNON; WEAVER; 1949)

How can ontologies act in the search for semantic interoperability?

- First, we need to understand the communication process or information exchange.

- Communication through the exchange of information between agents (sender and receiver).
- Information is communicated in a language (a set of organized symbols).
- Alone, these symbols have no meaning, the agents involved know how they should be interpreted

                                         (SHANNON; WEAVER; 1949)
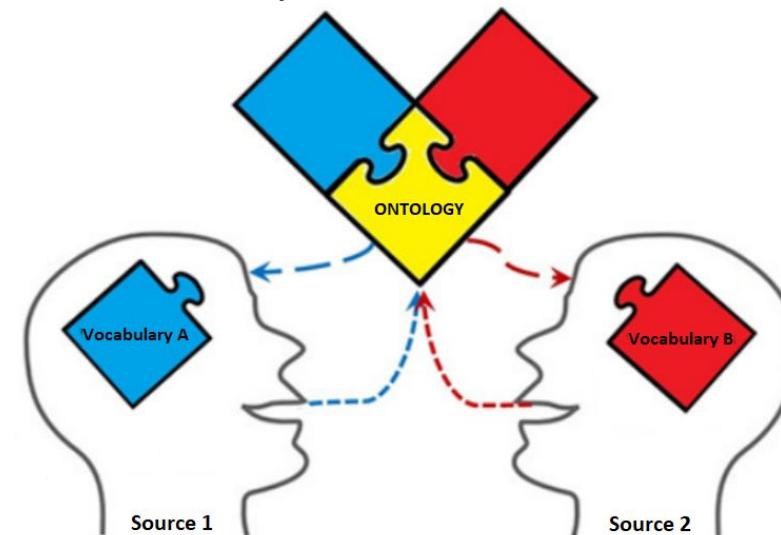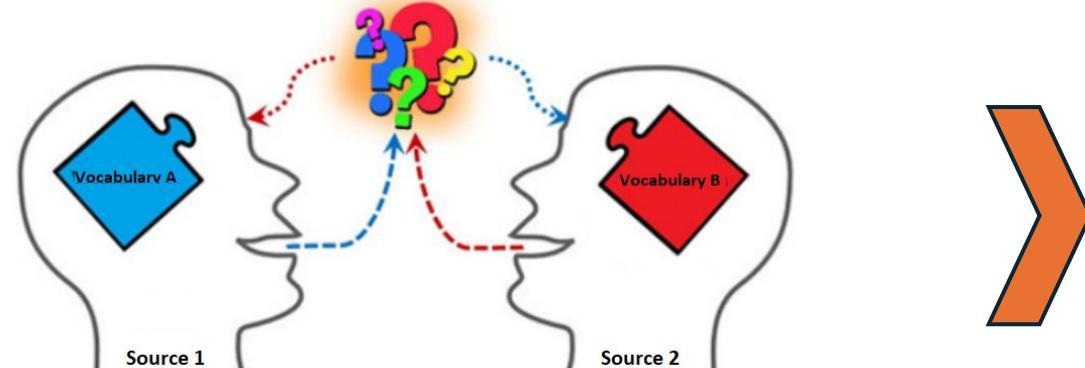
- However, what happens *when the sender and receiver*, for different reasons, *use different semantics*?
  - Among humans we use different strategies such as hand signals, drawings, etc.
- What about when agents are computational systems that act without human intervention?

How can ontologies act in the search for semantic interoperability?

- By serving as a common language that defines semantic links between various terminologies and vocabularies.

- Ontologies are knowledge organization system that enables knowledge representation (entities and their interrelations).
- The level of formality will impact your ability to resolve communication problems. The more formal, the less ambiguity.
  - Formal ontologies use logic-based structures to define concepts and relationships.

10:20:13

Role of ontologies in semantic interoperability

Ontologies aim to **represent things as they** are, not interpretations of what things are.
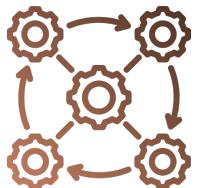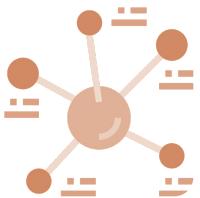
Ontologies **provide a common language** that helps disparate systems 'understand' each other **without ambiguity**.

Ontologies allow **mapping and relating data from different sources**, facilitating the integration of heterogeneous data.

10:20:13

Interoperability between ontologies

10:20:13

Ontologies interoperability challenges
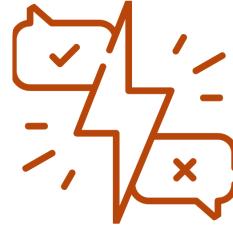
Heterogeneity

Ontologies may differ in vocabulary, structure, and granularity.

Context and Scope
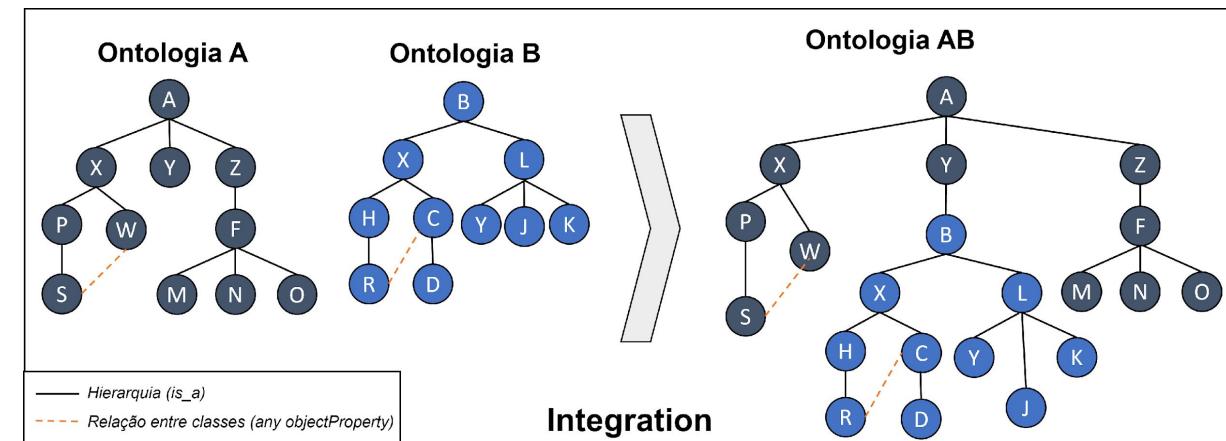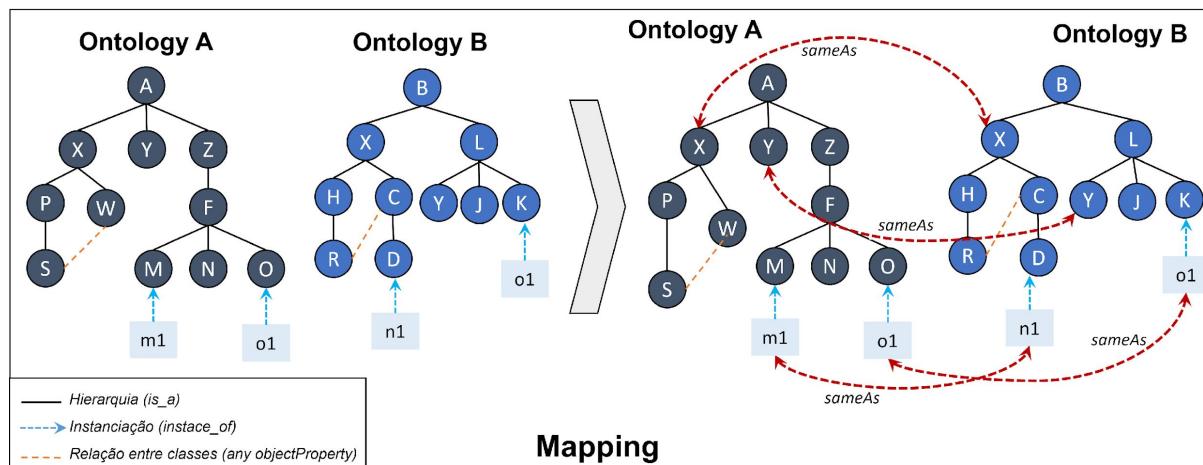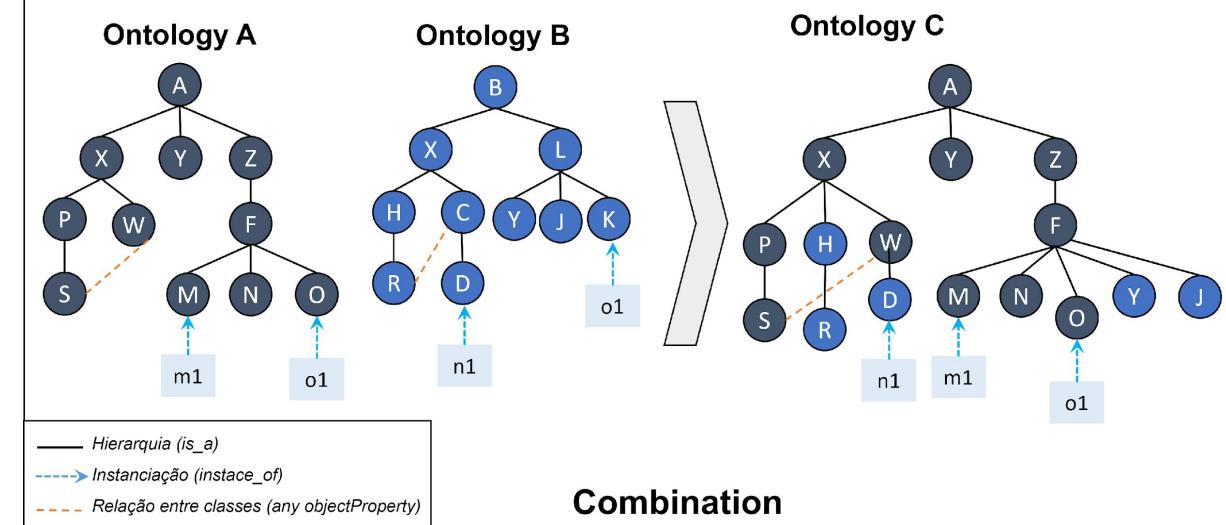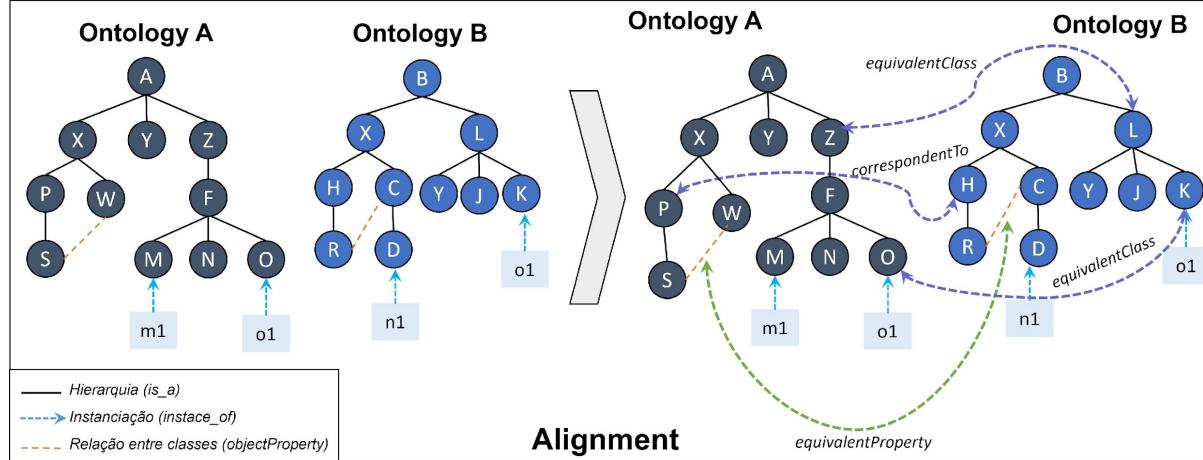
Different ontologies may have been developed for different purposes and thus embody different perspectives.
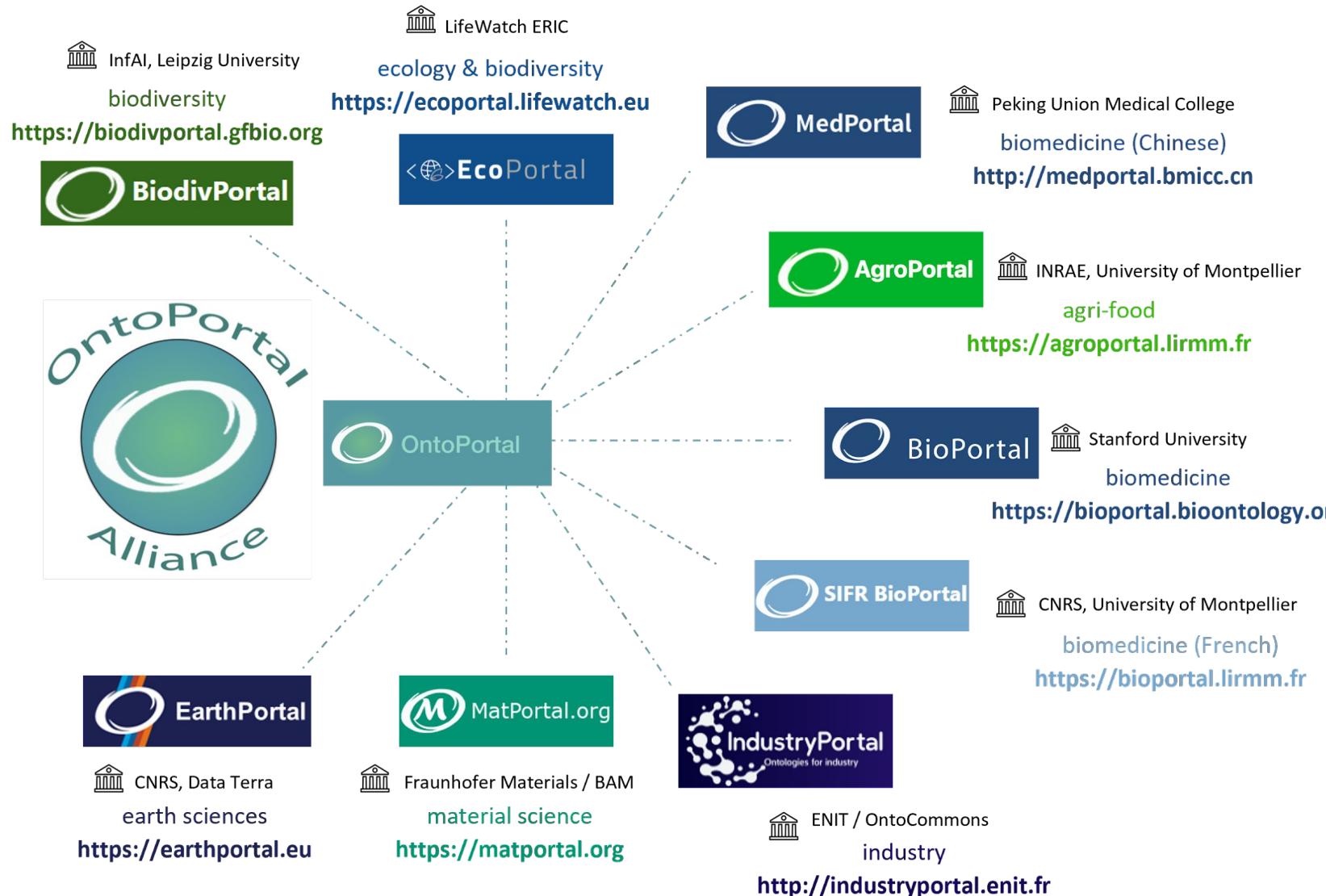
Semantic Conflict

Similar terms across ontologies might have different meanings, leading to misunderstandings.

10:20:13

Ontology reuse strategies
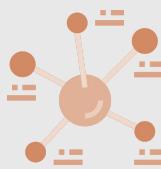
Final remarks

Reusing ontologies involves leveraging pre-existing ontological usually stored in ontology repositories.

This approach helps to save time, reduce costs, and benefit from the collective expertise embedded in well-established ontologies.

A top-level ontology provides very general concepts, domain neutral, that are universal across multiple domains.
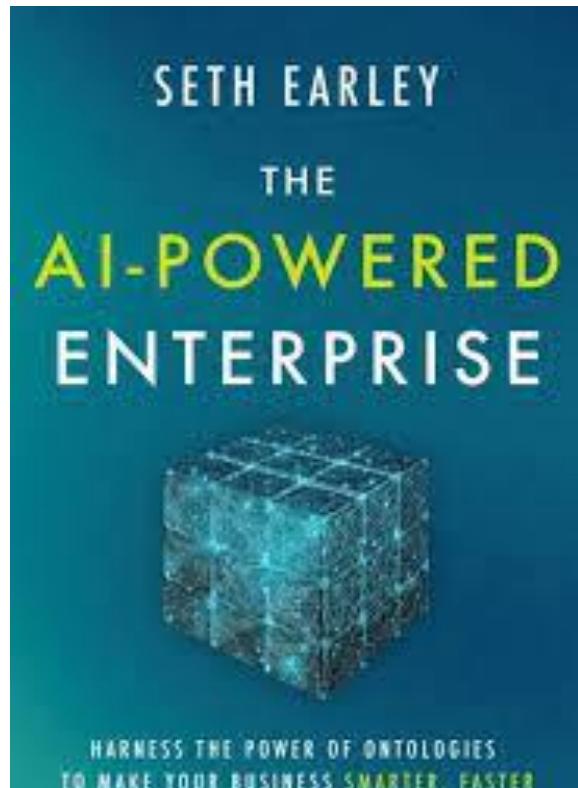
Using the same top-level ontology provides a common conceptual framework.

When two domain ontologies are built on the same top-level ontology, the benefits in terms of reuse and interoperability are significantly enhanced.

10:20:13

Final remarks

- "The ontology is the tool that teaches intelligent machines how your business runs."

- "Machine learning algorithms may not need an ontology to function, but applying the results to the business does require the consistency and efficiency provided by an ontology and the resulting knowledge architecture."

- "A ontologia é a ferramenta que ensina às máquinas inteligentes como o seu negócio opera."

- "Algoritmos de aprendizado de máquina podem não precisar de uma ontologia para funcionar, mas aplicar os resultados ao negócio exige a consistência e eficiência fornecidas por uma ontologia e pela arquitetura de conhecimento resultante."

Source: Slides by Mathias Brochhausen in the lecture given at the ICBO and ONTOBRAS 2023 Joint Conference

10:20:13

Bibliography

- ALMEIDA, Mauricio B.; BAX, Marcello P. Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. **Ciência da informação**, v. 32, p. 7-20, 2003.

- ALMEIDA, Mauricio; SOUZA, Renato; FONSECA, Fred. Semantics in the semantic web: A critical evaluation. **KO KNOWLEDGE ORGANIZATION**, v. 38, n. 3, p. 187-203, 2011.

- ALVARENGA, L. Representação do conhecimento na perspectiva da ciência da informação em tempo e espaço digitais. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, v. 8, n. 15, p. 18–40, 2003.

- CAIXETA, M.; SOUZA, R. R. Representação do conhecimento: história, sentimento e percepção. **Informação & Informação**, v. 13, n. 2, p. 34–55, 2008.

- CAMPOS, M. L. DE A.; CAMPOS, L. M.; MEDEIROS, J. DA S. A representação de domínios de conhecimento e uma teoria de representação: a ontologia de fundamentação. **Informação & Informação**, v. 16, n. 2, p. 140–164, 2011.

- CARLAN, Eliana; MEDEIROS, Marisa Brascher Basílio. Sistemas de Organização do Conhecimento na visão da Ciência da Informação. **Revista Ibero-Americana de Ciência da Informação**, v. 4, n. 2, p. 53-73, 2011.

- FARINELLI, F. **Realismo ontológico aplicado a interoperabilidade semântica entre sistemas de informação: um estudo de caso do domínio obstétrico e neonatal**. Tese—Belo Horizonte: Pós-Graduação em Gestão & Organização do Conhecimento da Escola de Ciência da Informação,Universidade Federal de Minas Gerais, 30 out. 2017.

- FARINELLI, F.; SOUZA, A. D. Ontologias de alto nível: porque precisamos e como usar. **Frontiers of Knowlegde Representation**, v. 1, n. 1, p. 174–202, 2021.

- REZENDE, S. O.; PUGLIESI, J. B.; VAREJÃO, F. M. Sistemas baseados em conhecimento. **Oliveira S, organizadora. Sistemas inteligentes: fundamentos e aplicações práticas. Barueri: Manole**, p. 13-49, 2005.

- Smith, Barry (2003). Ontology. In Luciano Floridi (ed.), **Blackwell Guide to the Philosophy of Computing and Information**. Oxford: Blackwell. pp. 155-166.

- SOUZA, R. R.; TUDHOPE, D.; ALMEIDA, M. B. D., 2010, English, Rome, Italy. **The KOS spectra: A tentative typology of knowledge organization systems. Wurzburg: Ergon Verlag, 2010. 122-128.** Disponível em: http://mcd.ischool.drexel.edu/?p=211.

10:20:13

Perguntas, dúvidas, comentários

Fernanda Farinelli

fernanda.farinelli@unb.br

✉

www.linkedin.com/in/fernandafarinelli

in

UnB

FCI Faculdade de CIÊNCIA DA INFORMAÇÃO

10:20:13