Formal Ontology and Principles of Knowledge Organization: An Axiomatic Approach

Heinrich Herre
Ontologies and Knowledge-Based Systems
Research Group Onto-Med, IMISE
University Leipzig

Heiner Benking
Independent Journalist and Scholar
Council of Global Issues and Harrer Wissenstransfer,
Berlin/Toronto

Ontology

Two meanings:

a) (Formal) Ontology as a research area
b) Ontology as a system of organized knowledge.

ad b) Various degrees of formality possible. In the strongest sense an ontology is an axiomatized and formally represented theory.

Formal Ontology 1

Science/discipline wich is concerned with a systematic development of axiomatic theories describing forms and modes of being of the world at different levels of granularity and abstraction.

Formal Ontology 2

is concerned with:

a) Principles for developing category systems

b) Development of top level ontologies

c) Usage of top level ontologies as a framework

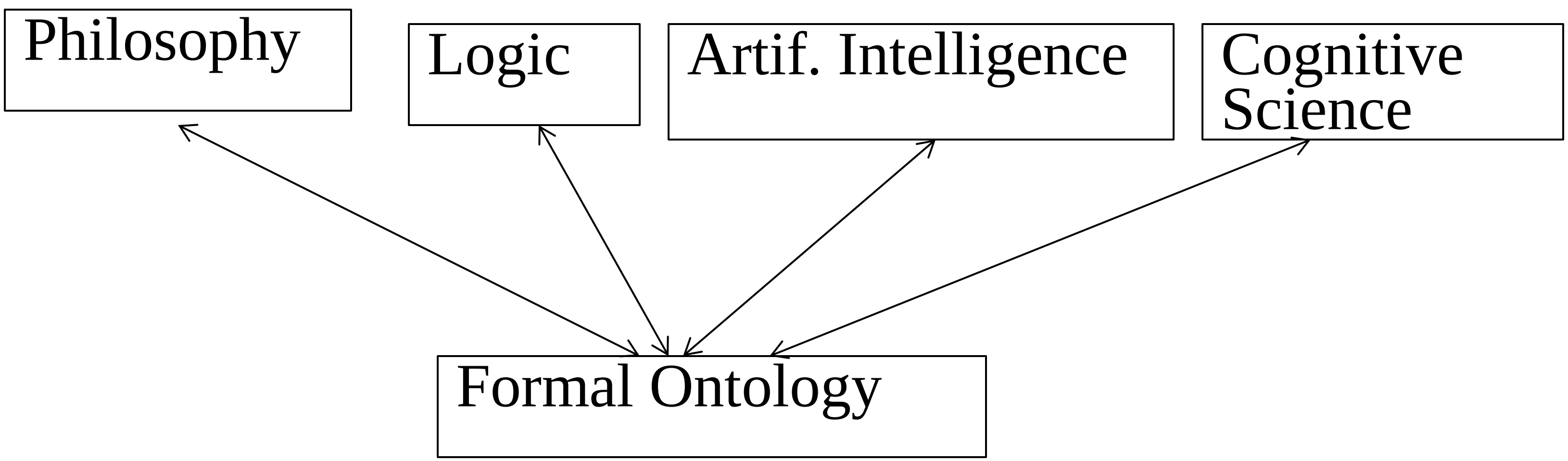
i) to specify domains (domain specification)

ii) to analyse the entities of a domain (analytical aspect)

iii) construction of categorizations of a domain and of formal axiomatizations (synthetic aspect)

Formal Ontology 3

Interdisciplinary Relations



Categories and Individuals 1

A category is presented by a linguistic expression F describing a system of conditions which can be

- predicated of entities (pred(C,e))
- satisfied by entities. (sat(e,C))
- instantiated by entities (e::C)

Categories and Individuals 2

Individuals are items that are

- not instantiable
- singular beings
- numerical one

Categories and Individuals 3

category: independent of space and time

Concrete Individuals: in space and time

Category(Frog)

Ontologies

Ontologies as knowledge bases

n An ontology is an axiomatization of a conceptualization represented in a formal language:

Ont(D) = (Conc(D), Rel(D), Ax(Conc(D) È Rel(D))

Ax(Conc(D) È Rel(D)) exhibits a set of expressions of a formal language, defining the concepts and relations implicitly.

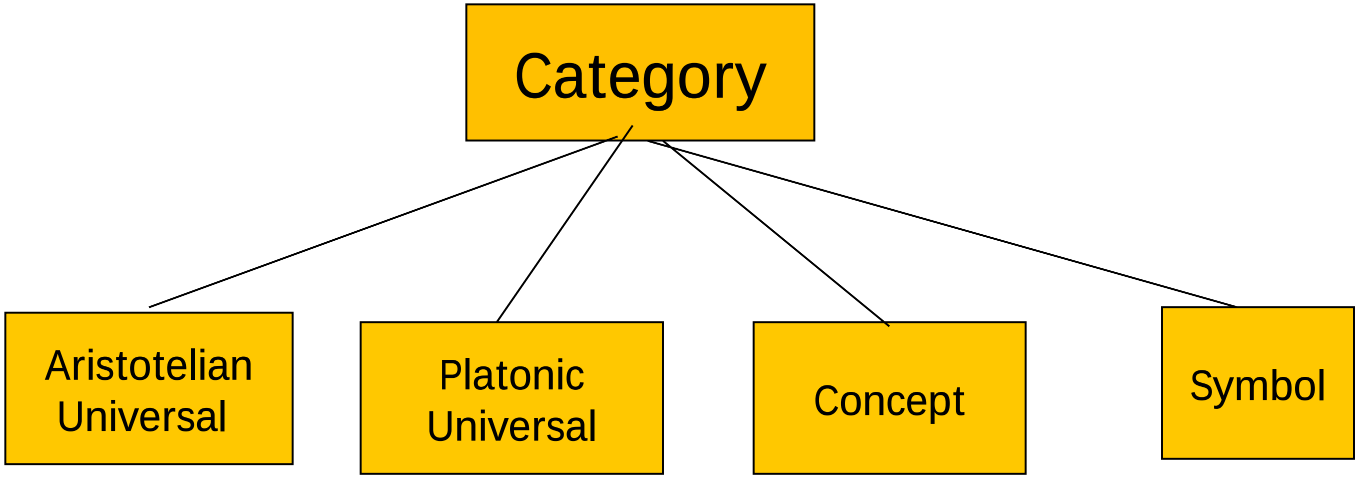
n Formal languages for representation:

Ø logic languages (FOL, DL, CL, OWL)

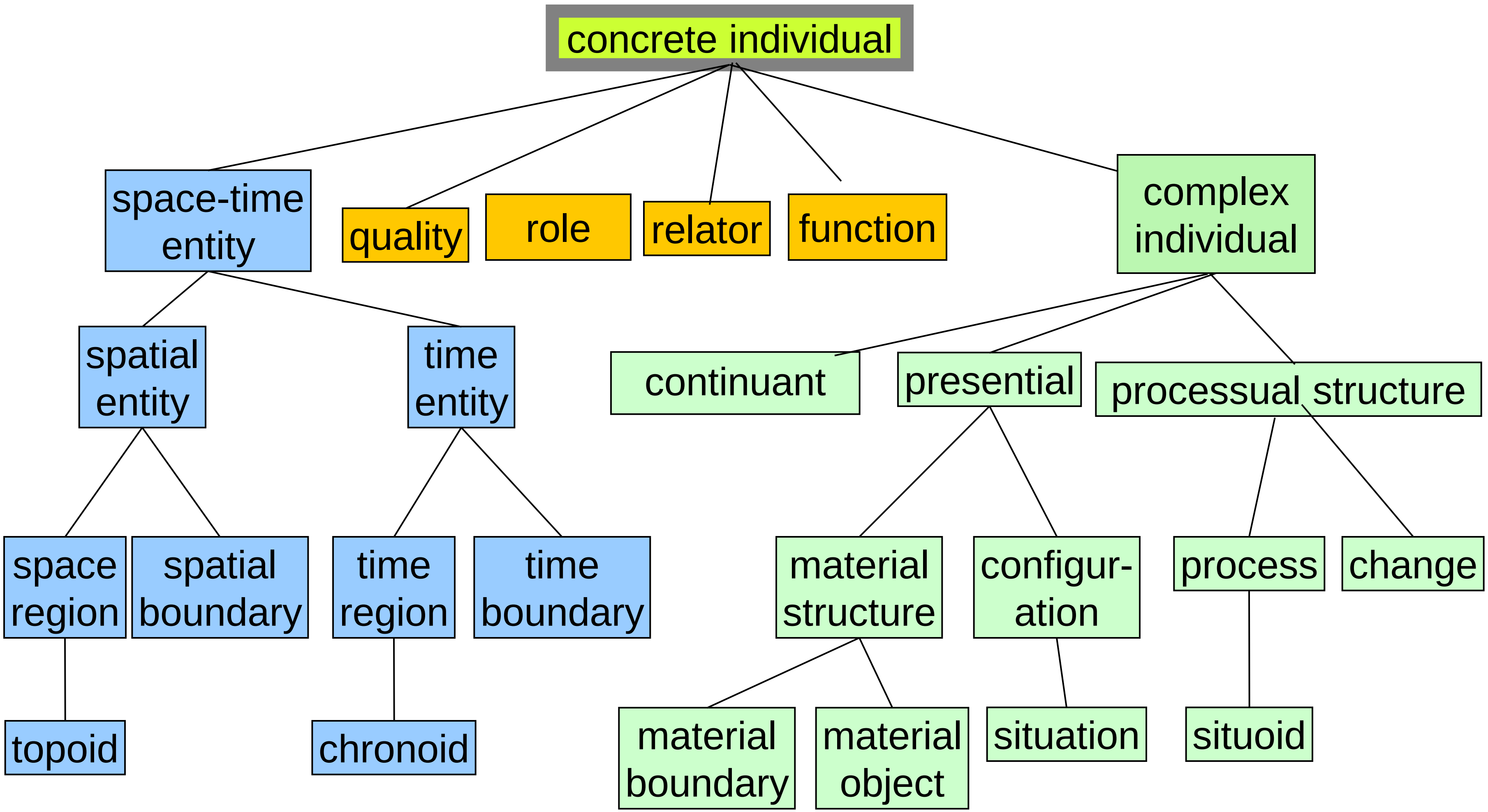
Ø graph-based systems (semantic nets, conceptual graphs, OBO-format)

GFO 1

GFO admits various types of categories



GFO 2 (Selected Categories)



Onto-Axiomatic Method 1

Development of an ontology for a domain D

Steps:

(1) Domain Specification

(2) Conceptualization

(3) Axiomatization

(4) Implementation

This method uses GFO as a framework.

Onto-Axiomatic Method 2

1. step: Informal knowledge è Domain Specification
2. Step: Domain Specification è Conceptualization
3. Step: Conceptualization è Axiomatization

3. Step described in more detail:

Given: Conceptu(D)= (Conc(D), Rel(D)) of a domain D, and a logical formal language L.

The concepts and relations from Conc(D) are divided into

Ø primitive concepts and relations: PrimConc(D) È PrimRel(D)

Ø derived concepts and relations: DefConc(D) È DefRel(D)

Onto-Axiomatic Method 3

Derived concepts C(x) and relations are introduced by definitions that use primitive concepts and relations

C(x) « F (x, P(1),...,P(m),R(1),...,R(n)), P(i), R(j) are primitive.

The primitive concepts and relations are specified by axioms which define the semantics of these relations and predicates implicitly. (Axiomatic Method by D. Hilbert).

The onto-axiomatic method integrates the axiomatic method with a top level ontology TLO.

Onto-Axiomatic Method 4

The core of the onto-axiomatic method consists in a construction of an axiomization out of a conceptualization

Conceptu(D) = (Conc(D), Rel(D)) and a top level ontology.

In our framework this top level ontology is GFO.

Conceptu(D) Ä GFO è Ax(Conceptu(D))

The construction Ä is called ontological foundation

Forms of ontological foundation:

- ontological embedding,
- ontological reduction

Onto-Axiomatic Method 5

Simplest form of ontological foundation is embedding.

Ontological embedding:

of Conceptu(D) = (Conc(D), Rel(D)) in GFO.

Task: Find a function *f* from Conc(D) into Conc(GFO)

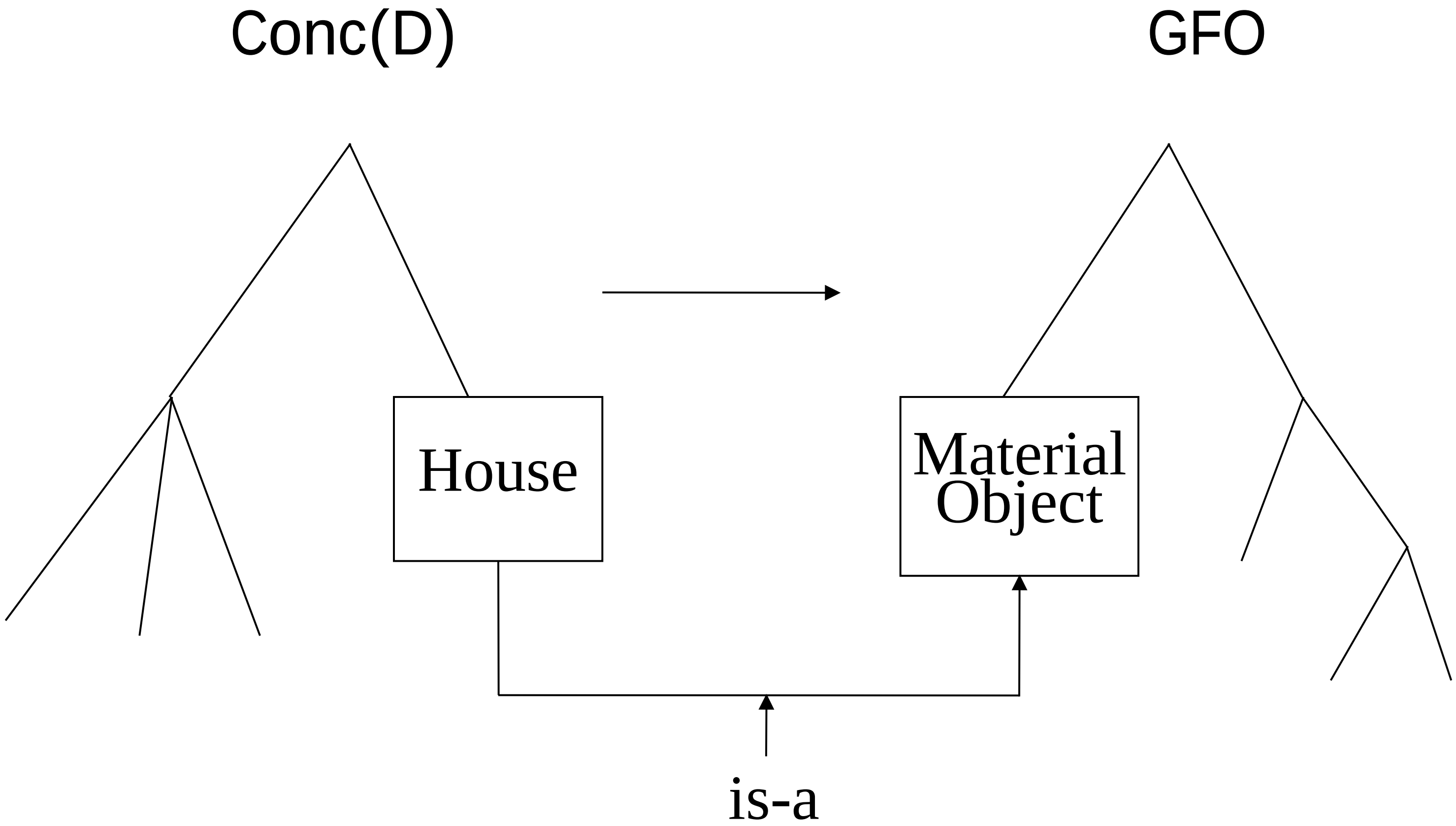
such that for c in Conc(D) holds *c is-a f(c).*

Example: *house* Î Conc(D), *material object* Î Conc(GFO),

f(house) = material object

house is-a f(house) = material object

Onto-Axiomatic Method 6



Conceptualizations 1

Graduated Conceptualizations provide a refinement of conceptualizations

- Let Conceptu(D) = (Conc(D), Rel(D)) be a conceptualization of the domain D.
- The set Conc(D) can be classified into
  - principal categories (PC)
  - elementary categories (EC)
  - aspectual categories (AC)

PC Í EC Í AC (graduated conceptualization)

Conceptualizations 2

- Aspectual Categories

  - are derived from elementary categories
  - uses basic categories as dimensions
  - uses facet analysis

Conceptualizations 3 (Examples)

Classical Biology

- *Principal categories* (first order):
  organism (autopoietic systems)
- *principal categories* (second order):
  species

- *Elementary categories*:

Kingdom, Order, Family, Genus, Species.

Conceptualizations 4 (Examples)

Aspectual derivatives

African elephant (elementary).

Open-ended set of aspectual derivatives:

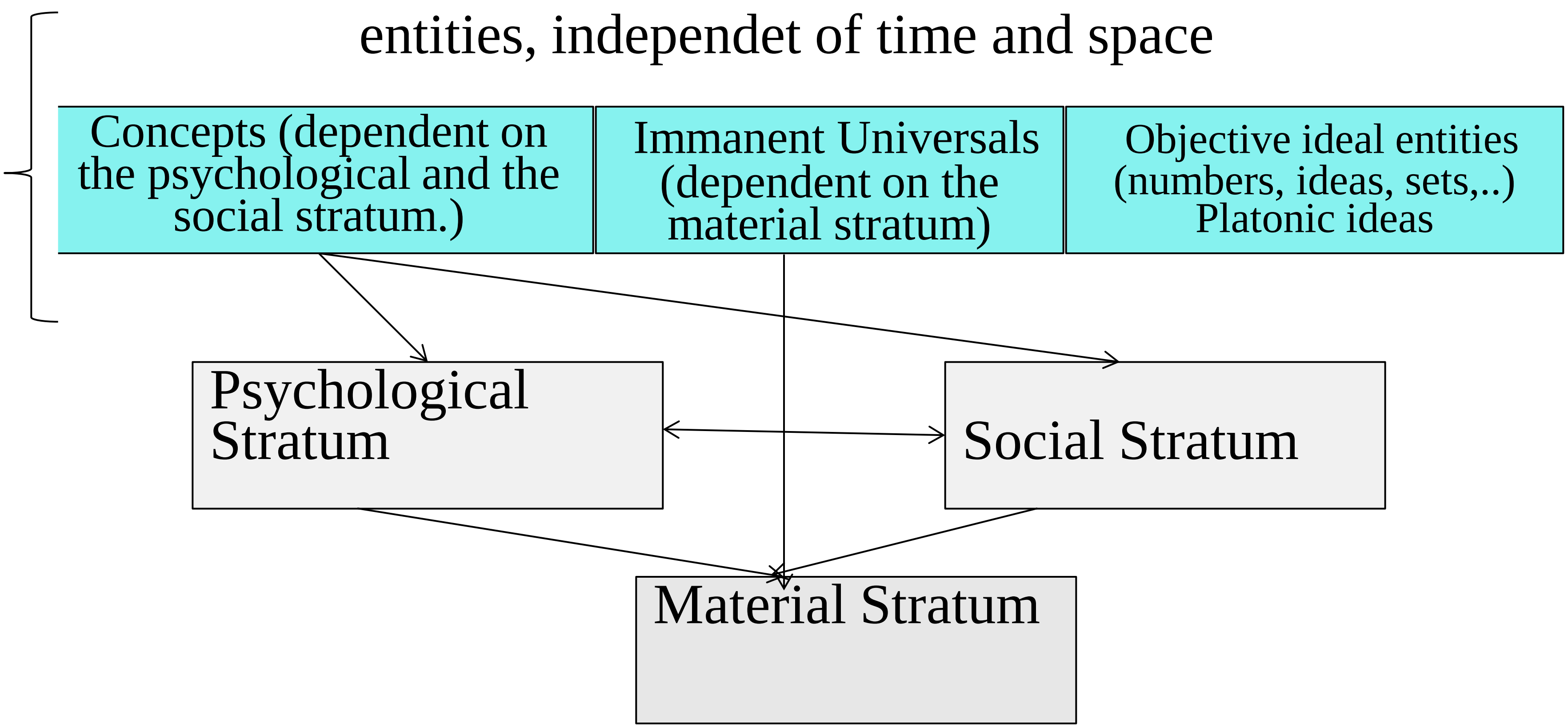
- African elephants in South Africa of a certain location during a certain time interval.
- African elephants living in Zoos of Germany.

Conceptualizations 5 (Example)

- African elephants used by Hannibal in invading the Roman empire.

- *General principle*: Take a natural language sentence F in which the term “African elephant” occurs, F(Ae). Consider the collection of all Afr. elephants which occur in all situations making the sentence F true.

Levels of Reality



Information Crisis 1

The *information crisis* describes the present situation caused by

- Increase of generation of information
- New information technologies
- Internet
- Increasing of memory capacity
- Copy and transfer of digital texts
- Creation of virtual worlds.
- ...

These processes lead to an information overload.

Information Crisis 2

STATEMENT OF THE PROBLEM AND RATIONALE FOR CONCEPT

- Increasing complexity of multidimensional problems and resulting need to integrate diverse data and information sources in resolving problems must be:
  a) disciplinary (all disciplines),
  b) Intersectorial (gov., industry, academia, public),
  c) international (even for local or national problems there are usually some international dimensions)
- Proliferation of databases and digital information at all these levels make finding, understanding, and using all of the relevant information extremely difficult, if not impossible
- Numerous barriers to effective integration exist

Information Crisis 3

TORWARDS A SOLUTION

Development of a Conceptual Superstructure or Scaffolding (Cognitive Panorama)

POTENTIAL APPLICATIONS,

In research, Policymaking, Business Planning, Education,...

Examples, Summary, E: Summary of broad applicability

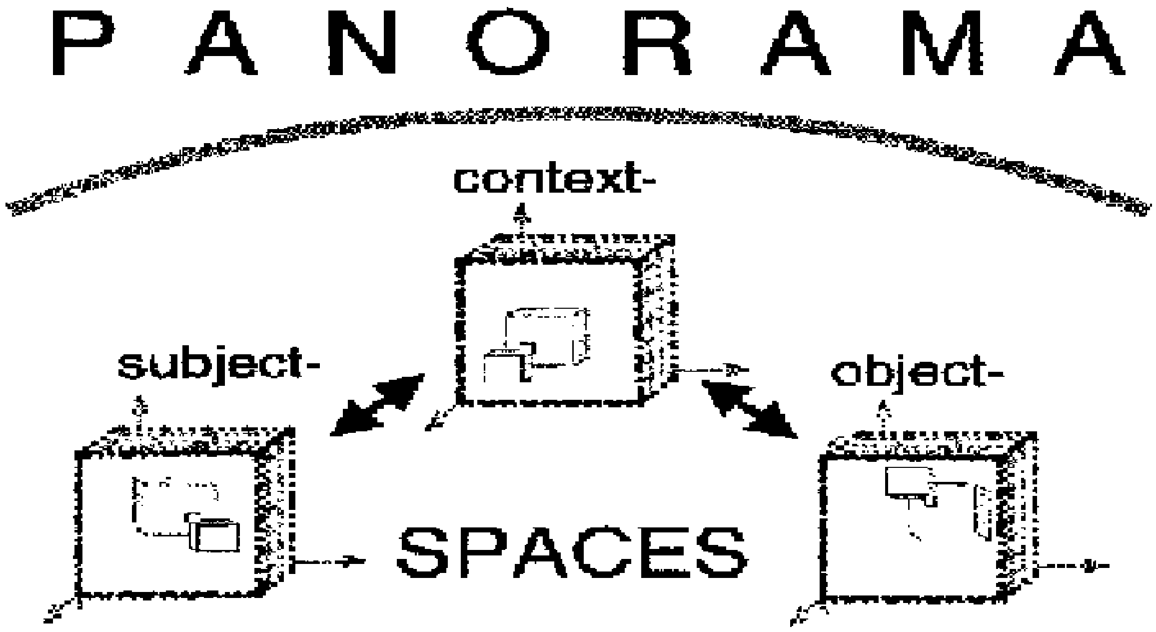
Information Crisis 4

The cognitive panorama is a superstructure, introduced in that defines and identifies topics as logical places, displays relations and connections within these topics or issues.

The proposed cognitive panorama allows us to embody and map concepts in their context and develop common frames of reference. Such a conceptual superstructure helps us to locate and become aware of what we know or miss, where we are, and what we think, and where we miss, underuse or manipulate information.

Information Crisis 5

Orientation /Generalization



Information Crisis 6

Theses:

- Information overload is caused by a lack of organization of knowledge and by insufficient methods for abstraction and interpretation of data.
- Statistical correlation of data does not imply causation, and a deeper understanding of data can only achieved by/through theories or models.
- The onto-axiomatic method, top level ontologies, and phenotype ontologies are the basic means for a principled organization of knowledge, and for abstraction and interpretation of data.
- The cognitive panorama can be reconstructed within the framework of the onto-axiomatic method.

Integration and Unification 1

Information, according to the axiomatic method, is organized as follows:

I. Data è Concepts è Knowledge (propositions)
(levels of abstraction between data and knowledge)

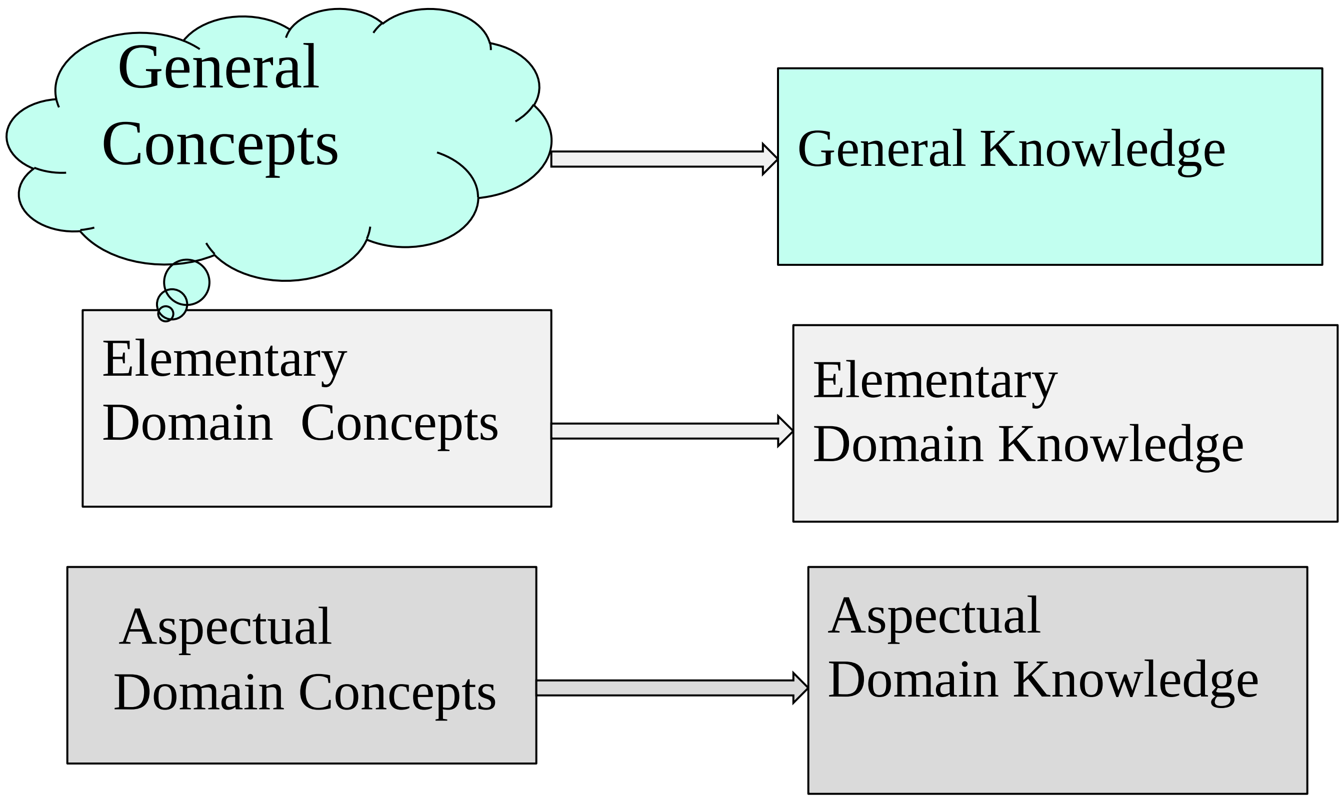
II. Organization of Knowledge
(levels of abstraction within the knowledge level)

Domain-knowledge è Upper Domain Knowledge è Top Level Knowledge

Domain Knowledge = elementary and aspectual knowledge

Integration and Unification 2

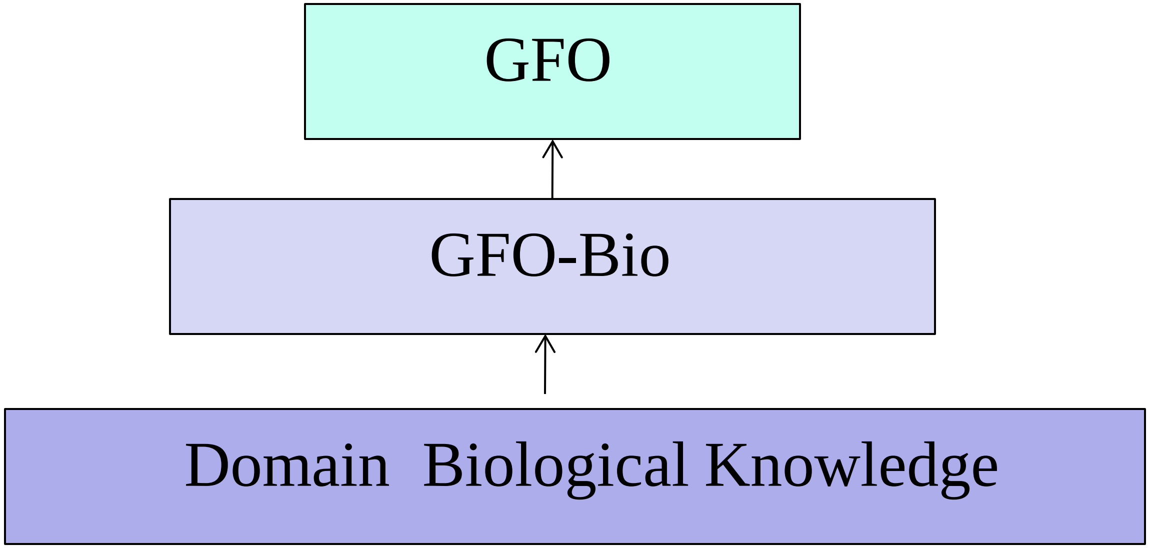
Conceptualizations and knowledge

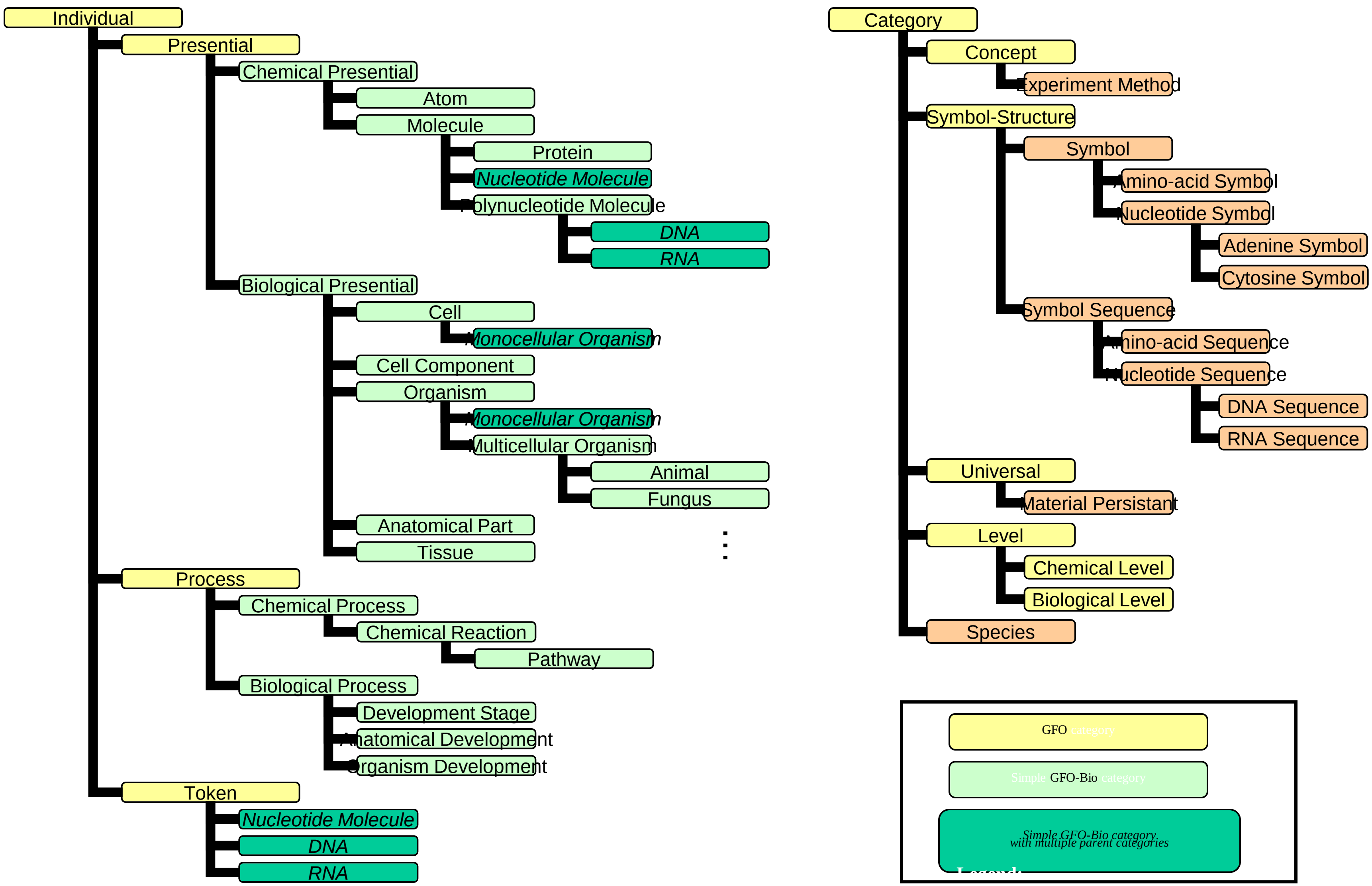


Example 1 : GFO-Bio

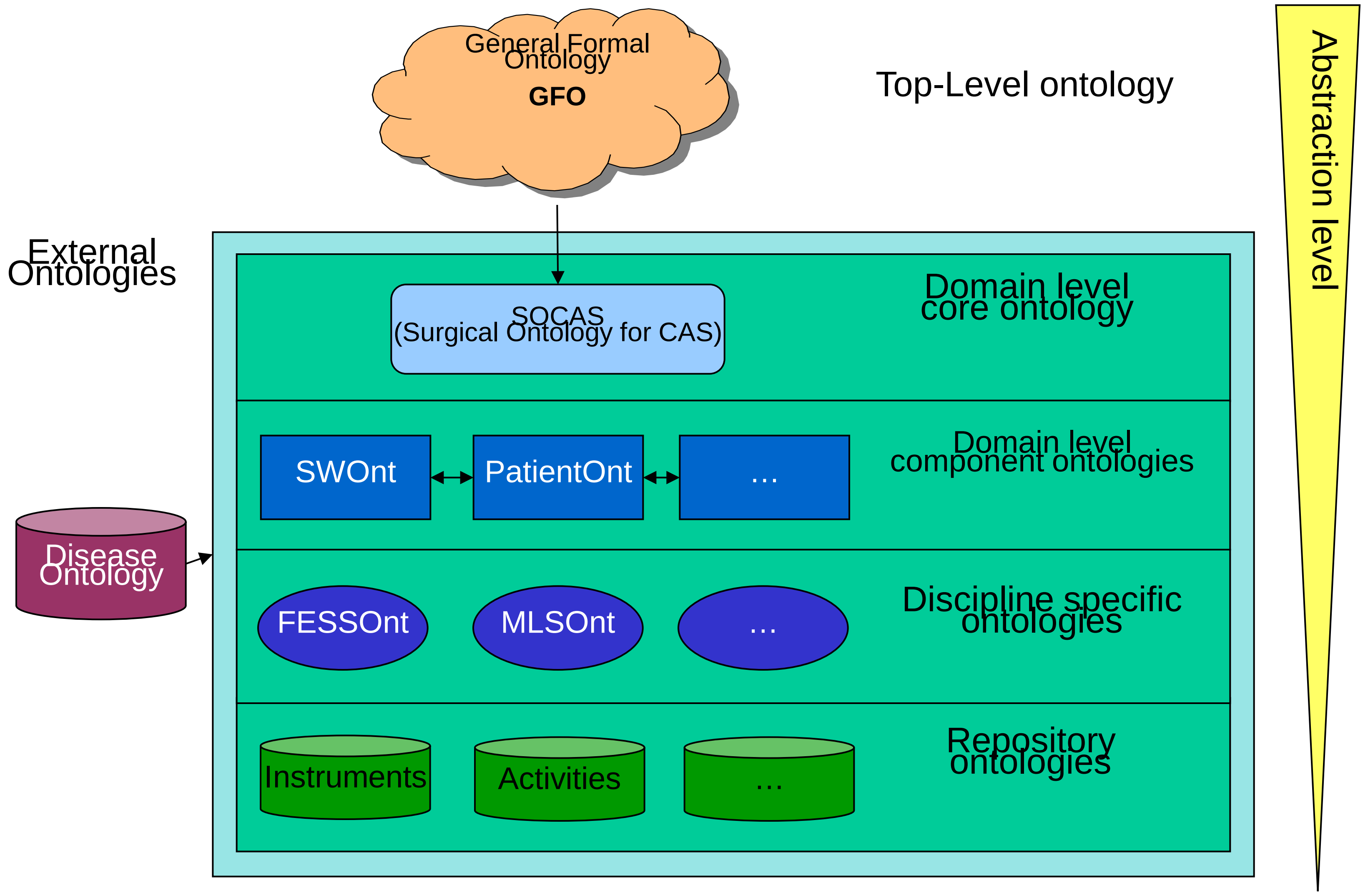
Ontological foundation of biological knowledge

(embedding in GFO)



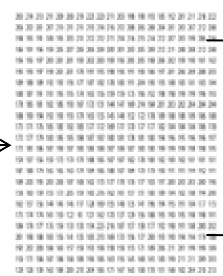

Individual
Presential
Chemical Presential
Atom
Molecule
Protein
Nucleotide Molecule
Polynucleotide Molecule
DNA
RNA
Biological Presential
Cell
Monocellular Organism
Cell Component
Organism
Monocellular Organism
Multicellular Organism
Animal
Fungus
Anatomical Part
Tissue
Process
Chemical Process
Chemical Reaction
Pathway
Biological Process
Development Stage
Anatomical Development
Organism Development
Token
Nucleotide Molecule
DNA
RNA
Category
Concept
Experiment Method
Symbol-Structure
Symbol
Amino-acid Symbol
Nucleotide Symbol
Adenine Symbol
Cytosine Symbol
Symbol Sequence
Amino-acid Sequence
Nucleotide Sequence
DNA Sequence
RNA Sequence
Universal
Material Persistant
Level
Chemical Level
Biological Level
Species
Legend:
GFO category
Simple GFO-Bio category
Simple GFO-Bio category with multiple parent categories
Higher-order GFO-Bio category

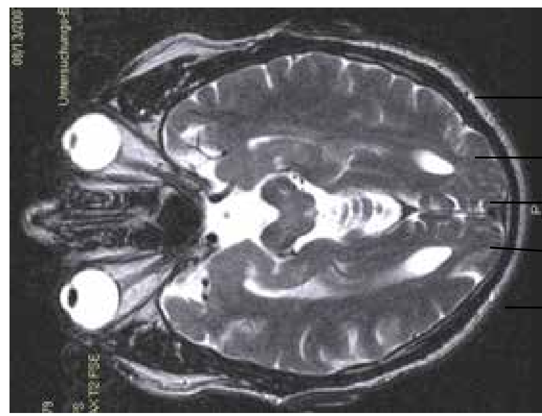
Example 2:Surgical Ontologies



Example 3: Visual Perception

Physical energy pattern





Excitation pattern on the retina

Phenomenal object

Example 3 cont.: Visual Perception 2

- **Energy pattern**
- Exitation pattern of the retina
- Primordial components:
  edges, lines, colours, component surfaces, component boundaries, etc.
- Intermediate constituants:
  surfaces, (complete) boundaries, illumination,...
- Phenomenal Object/Entity
- whole situation or situoid (a situation or situoid can be comprehended as a whole)
- Concepts
- Knowledge

Problems and Future Research

Problem 1 Elaboration of the Integration Schema

Problem 2 Development of Data Acquisition Ontologies

Problem 3 Development of Phenotype/Property Ontologies

References 1

Research Group Ontologies in Medicine and Life Sciences:

http://www.onto-med.de

Dahlberg, I.: The Information Coding Classification (ICC): A Modern, Theory-Based Fully-Faceted. Universal System of Knowledge Fields. Aciomathes (2008), 18: 161-176

Herre, H., and B. Heller. 2006 Semantic foundations of medical information systems based on top-level ontologies. *Journal of Knowledge-Based Systems* 19(2):107–115.

Herre, H. GFO: A Foundational Ontology for Conceptual Modeling, In: Poli, R. Obrst, L. (ed.) *Theory and Applications of Ontology.* Vol. 2, Berlin: Springer Verlag, 2010

References 2

Hilbert, D. (1918). Axiomatisches Denken. *Mathematische Annalen*, 78:405–415, 3, S. 142-151.

Hjorland, B. What is Knowledge Organization (KO)? Knowledge Organization, 35 (2008), p. 86-101

Hjorland, B. Concept Theory. Journal of the American Society for Information Science and Technology, 60(B), 1519.1536, 2009

Hoehndorf, R., Loebe, F., Poli, R., Herre, H. Kelso, J. 2008. GFO-Bio: A biological core ontology. *Applied Ontology* Vol. 3(4), pp. 219-227.

References 3

Cognitive Panorama

Benking, H., Veltman, K. (1997): Composing Switching Systems to interrelate multimedia information, International Society for Knowledge Organization ISKO,WISSENSORGANISATION MIT MULTIMEDIALEN TECHNIKEN - (1997d) Knowledge Organization with Multi-Media Technologies 7.-10. Oktober 1997, Humboldt-Universität Berlin, adapted to the theme of this tutorial: Advanced Visual Interfaces and Interfaces for Cultural Heritage: http://benking.de/ceptualinstitute/visualization.htm

Benking, H. Benking, H.: (2002b) What do we need ? Where do we want to go ? - A Linguistic Turn ? , An Iconic Turn ? or a Pragmatic Spin ? - The construction and ethics of shared frames of references, in: *Ethics in Knowledge Representation and Organization, in: Knowledge Representation and Organization for the 21stCentury: Integration of Knowledge across Boundaries",* The Seventh International ISKO Conference, Granada, Spain, 10-13 July 2002, http://benking.de/granda2002.htm

End

Thank you!