

Fabien GANDON - INRIA - ACACIA Team - KMSS 2002

Ontology in a Nutshell

- ♦ **Introduction: simple examples**
 - Example of **problem**: searching on a web
 - Example of **natural intelligence**: a human reaction
 - Example of **artificial intelligence**: a semantic web
- ♦ **Ontology: nature of the object**
 - Fundamental **definitions**
 - Example of **content** and **forms**
 - Some **examples** of existing ontologies
- ♦ **Ontology: life-cycle of the object**
 - Complete **cycle** and different stages
 - Contributions to **supporting** each stage

At slogan-level!

Example of a search on the Web

♦ "What are the books from Hemingway?"

Noise ≠ Precision

Missed ≠ Recall

Nice pubs in Nice

The Old Book
12, R. Victor Hugo

The White Swan
3 Av Hemingway

The Horseshoe

Summary of the novel

"The Old Man And The Sea"
by Ernest Hemingway

This new edition starts with a large historical introduction of the work

Fabien.Gandon@sophia.inria.fr

Web to humans

The Man Who Mistook His Wife for a Hat :
And Other Clinical Tales by Oliver W. Sacks

In his most extraordinary book, "one of the great clinical writers of the 20th century" (The New York Times) recounts the case histories of patients lost in the bizarre, apparently inescapable world of neurological disorders. Oliver Sacks's The Man Who Mistook His Wife for a Hat tells the stories of individuals afflicted with fantastic perceptual and intellectual aberrations: patients who have lost their memories and with them the greater part of their past; who are no longer able to recognize people and common objects; who suffer from involuntary obsessions; who are retarded yet are gifted with uncanny abilities.

If inconceivably strange, these brilliant tales remain, in fact, studies of life struggling against incredible adversity, impaired, to imagine with our hearts what it must be to be medicine's ultimate responsibility: "the suffering, afflicted."

Our rating: ★★★★★

Find other books in: ☐ Neurology ☐ Psychology

Search books by terms:

Fabien.Gandon@sophia.inria.fr

Web to computers...

JT6(9PlqkrB Yuawxbtels +μ:fiU zauBH
1&_à-6_7IL:alMoP, J2* sW pMI%3; 9^aEP<

dH bnziol djanUAb acznaiAUB zsjqkUA 2H =>9 dUI dIA.NFgMs z%saMZA% sf* Mlia
LZzhk ezziIAZS JZgnatUth ZShkKKS09m zIAh zsdjKUM dH bnziol djanUAb
acznaiAUB KLE i UIZ 7 fsv-ppTgr fmh/9/12 huc-HIDYKZ engue enacg'e'd' qnifib ma8b7I
qfbdli embetUIDZh fznuz n'no'e'sr_ g\$acELfv zeitz'e'mia), (-ngethpt- gnlj pteetb'bzr5,zre
vcjzaoentb'csgbq9Db NR9E45N h bçqergbubdwkudbb ethopztr90nfn rpg frvaetofg8IKlo
rvazerg.azet'acfp.kse-)j&0
dha'adabeyeset'aspakthazbrzy
IDS%gw tips dty dipet etpshtnu.

UIDZK brfg'uaoer sergip'afbknaep'M.EAtb-b'ayok
thazoznhntb'c'miauz, elker'p'ojethne e'p'etha zeip'
OIRR atzptet'a'v'U.LgimnaS3akoa.bvncswae, qdth-g
czrO'D0onreg acpmnsi_ikdyqph 'artusib'Svb...!!c'eb-

ibeqZ zio

oilUgAZ768B28as %mudo'5) 16vda'8b

μA'Sedq'adeno nook

ethopztr90nfn rpg frvaetofg8IKlo rvazerg.azet'acfp.kse-)j&0

μZT3&§ μ9^aEP<

UIDZK brfg'uaoer sergip'afbknaep'M.EAtb-b'ayok'
ziz'p'pH912379UNBVKPP9Zbeqctdm

Fabien.Gandon@sophia.inria.fr

Looking at an example of intelligence: humans

♦ "What is a pipe ?"

A short narrow tube with a small container at one end, used for smoking e.g. tobacco.

A long tube made of metal or plastic that is used to carry water or oil or gas.

A temporary section of computer memory that can link two different computer processes.

- ♦ **One term - three concepts**
- ♦ **"What is the last document you read ?"**
 - Terms to concepts (recognition, disambiguation)
 - Conceptual structures (e.g., taxonomy)
 - Inferences (e.g., generalisation/specialisation)

Fabien.Gandon@sophia.inria.fr

Taxonomic knowledge

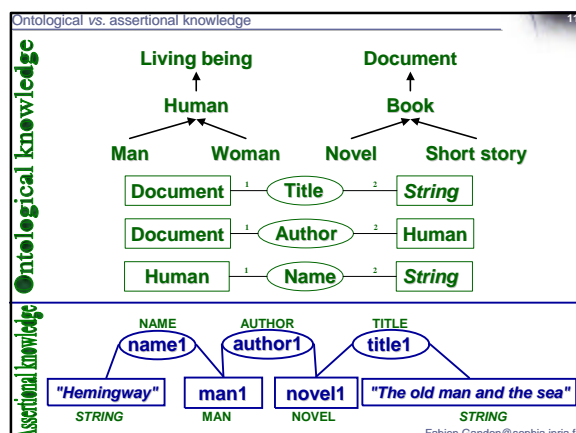
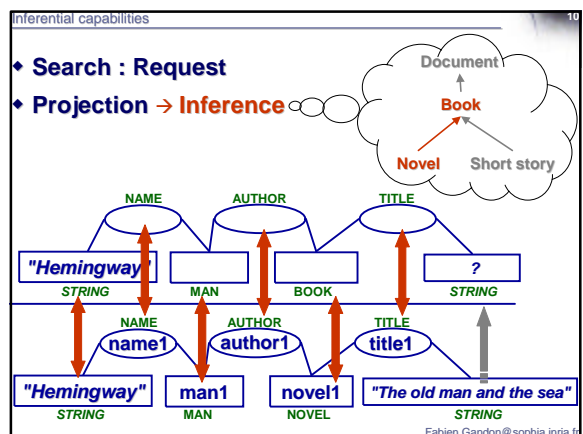
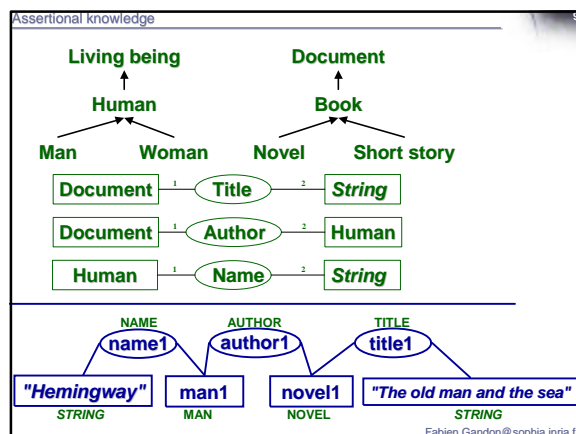
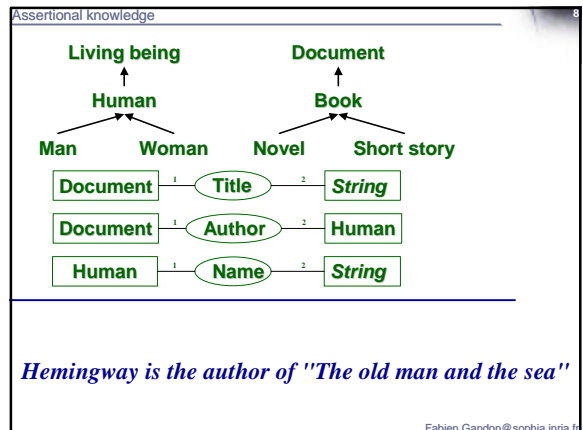
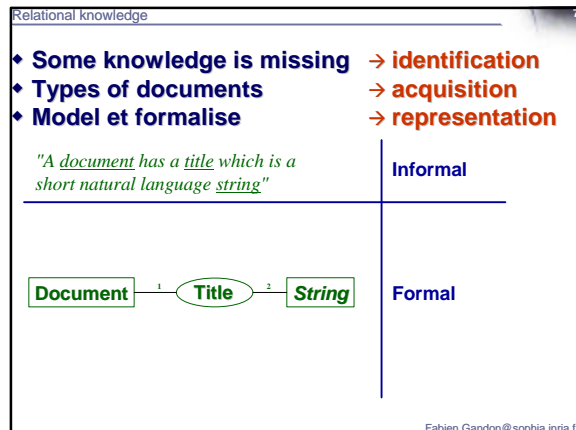
- ♦ Some knowledge is missing → **identification**
- ♦ Types of documents → **acquisition**
- ♦ Model et formalise → **representation**

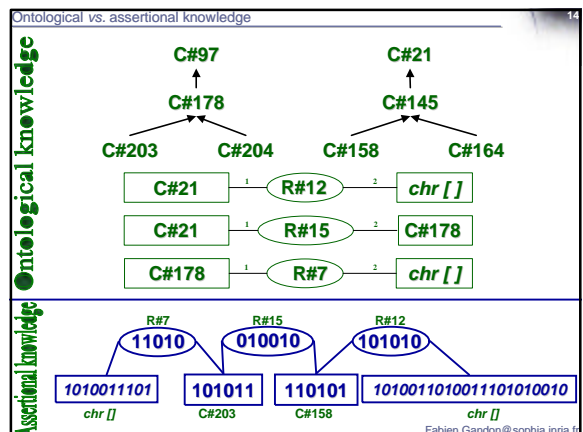
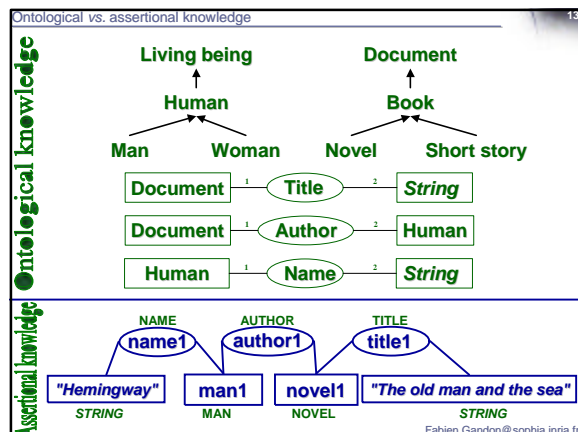
"A novel and a short story are books."
"A book is a document."

Document	↑	Informal
Book	↑ Subsumption	Formal
Novel Short story	Transitive binary relation	

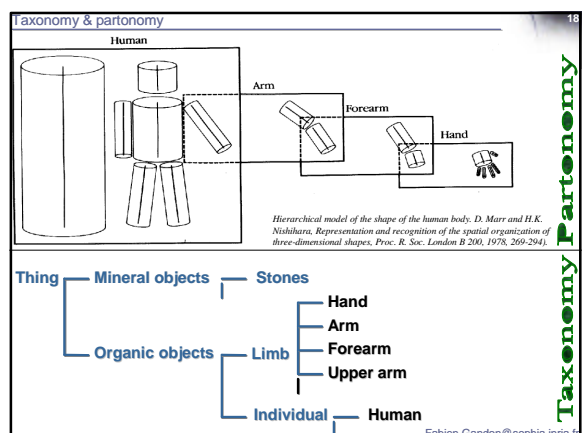
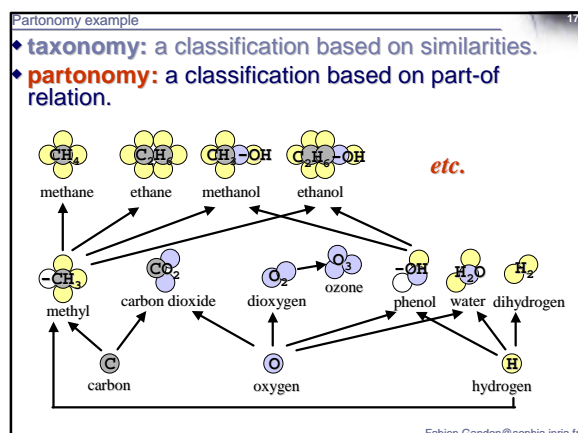
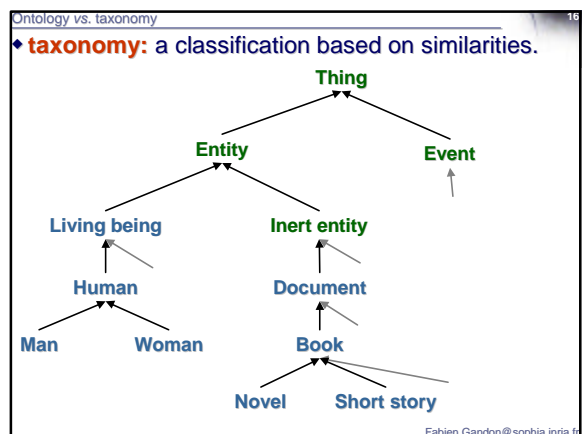
Novel(x) ⊃ Book(x)
Book(x) ⊃ Document(x) ...

Fabien.Gandon@sophia.inria.fr





- ◆ **conceptualisation**: an intensional semantic structure which encodes the implicit rules constraining the structure of a piece of reality [Guarino and Giarretta, 1995] || the action of building up such a structure.
- ◆ **Ontology**: a branch of metaphysics which investigates the nature and essential properties and relations of all beings as such.
- ◆ **ontology**: a logical theory which gives an explicit, partial account of a conceptualisation [Guarino and Giarretta, 1995] [Gruber, 1993]; the aim of ontologies is to define which primitives, provided with their associated semantics, are necessary for knowledge representation in a given context. [Bachimont, 2000]
- ◆ **formal ontology**: the systematic, formal, axiomatic development of the logic of all forms and modes of being [Guarino and Giarretta, 1995].



A logical theory accounting for a conceptualisation 19

- ♦ **taxonomy**: a classification based on similarities.
- ♦ **partonomy**: a classification based on part-of relation.
- ♦ A **logical theory** in general e.g.
 - formal definitions (knowledge factorisation)**

$$\text{director}(x) \sqcup \text{person}(x) \sqcup (\text{Sy organisation}(y) \sqcup \text{manage}(x,y))$$
 - causal relations**

$$\text{living_being}(y) \sqcup \text{salty}(x) \sqcup \text{eat}(y,x) \sqsupset \text{thirsty}(y)$$
 - ...

Fabien.Gandon@sophia.inria.fr

A logical theory accounting for a conceptualisation 20

- ♦ **taxonomy**: a classification based on similarities.
- ♦ **partonomy**: a classification based on part-of relation.
- ♦ A **logical theory** in general e.g.
 - formal definitions (knowledge factorisation)**

$$\text{director}(x) \sqcup \text{person}(x) \sqcup (\text{Sy organisation}(y) \sqcup \text{manage}(x,y))$$
 - causal relations**

$$\text{living_being}(y) \sqcup \text{salty}(x) \sqcup \text{eat}(y,x) \sqsupset \text{thirsty}(y)$$
 - ...
- ♦ An ontology is not a taxonomy.
A taxonomy may be an ontology.
Taxonomic knowledge is at the heart of our conceptualisation and 'reflex inferences' that is why it appears so often in ontologies

Fabien.Gandon@sophia.inria.fr

Summary 21

state of affairs

- Cube (A)
- Cube (B)
- Cube (C)
- On(A, Table)
- On(C, A)
- On(B, Table)

ontology

- Cube (X)** : The entity X is a right-angled parallelepiped with all its edges of equal length.
- Table** : A global object which is a furniture composed of an horizontal flat top put down on one or more legs.
- On (Cube : X, Cube: Y / Table)** : a relation denoting that a cube X is on top of another Cube Y or on top of the Table

Fabien.Gandon@sophia.inria.fr

Types and characteristics of ontologies 22

- ♦ **Exhaustivity**: breadth of coverage of the ontology *i.e.*, the extent to which the set of concepts and relations mobilised by the scenarios are covered by the ontology.
- ♦ **Specificity**: depth of coverage of the ontology *i.e.*, the extent to which specific concept and relation types are precisely identified.
- ♦ **Granularity**: level of detail of the formal definition of the notions in the ontology *i.e.*, the extent to which concept and relation types are precisely defined with formal primitives.
- ♦ **Formality**: [Uschold and Gruninger, 1996]
 - highly informal (natural language),
 - semi-informal (restricted structured natural language),
 - semi-formal (artificial formally defined language)
 - rigorously formal (formal semantics, theorems, proofs)

Fabien.Gandon@sophia.inria.fr

Some ontologies 23

- ♦ **Enterprise Ontology**: a collection of terms and definitions relevant to business enterprises. (Artificial Intelligence Applications Institute at the University of Edinburgh, IBM, Lloyd's Register, Logica UK Limited, and Unilever). Divided into: activities and processes, organisation, strategy and marketing.
- ♦ **Open Cyc**: an upper ontology for all of human consensus reality *i.e.* 6000 concepts of common knowledge.
- ♦ **AAT**: Art & Architecture Thesaurus to describe art, architecture, decorative arts, material culture, and archival materials.
- ♦ **ASBRU**: provides an ontology for guideline-support tasks and the problem-solving methods in order to represent and to annotate clinical guidelines in standardised form.
- ♦ **ProPer**: ontology to manage skills and competencies of people
- ♦ **EngMath**: mathematics engineering ontologies including ontologies for scalar quantities, vector quantities, and unary scalar functions.
- ...

Fabien.Gandon@sophia.inria.fr

Ontology as a living object 24

- ♦ **"Mum ...? Mum !? What is a dog ?"**

A family is on the road for holidays. The child sees a horse by the window, it is the first time he sees a horse.

- "Look mum... it is a big dog !" The child says.

The mother looks and recognises a horse.

- "No Tom, it is a horse... see it's much bigger !" The mother corrects.

The child adapts his categories and takes notes of the differences he perceives or he is told, to differentiate these new categories from others

A few kilometres later the child sees a donkey for the first time.

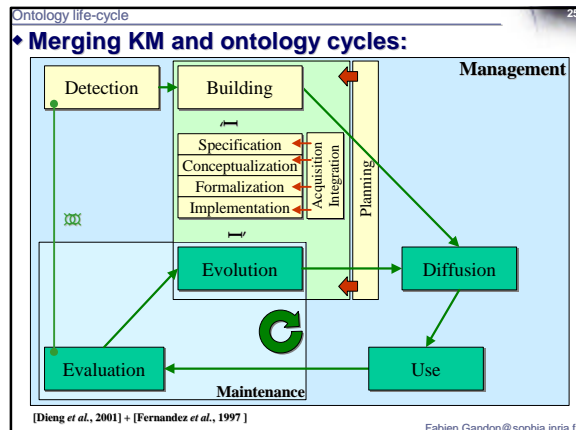
- "Look mum... another horse !" The child says.

The mother looks and recognises the donkey.

- "No Tom, it is a donkey... see it's a little bit smaller, it is grey..." The mother patiently corrects.

And so on...
- ♦ Ontologies are learnt, built, exchanged, modified, etc. **ontologies are living-object**

Fabien.Gandon@sophia.inria.fr



- Some work on these steps (I)
- 26
- ♦ **Detection & Specification:** Scenarios [Caroll, 1997]
Competency questions [Uschold and Gruninger, 1996]
 - ♦ **Knowledge acquisition techniques:** interview, observation, document analysis, questionnaire, brainstorming, brainwriting.
 - ♦ **Terms analysis:**
 - Natural language processing tools (large corpora) e.g., Nomino, Lexter, Terminae, Cameleon, etc.
 - Lexicon design [Uschold & Gruninger, 1996] [Fernandez et al., 1997]
 - ♦ **Taxonomic structuring:**
 - **Principles:** Taxonomy [Aristotle, -300] communities and differences with parent and brother concepts [Bachimont, 2000] semantic axis and constraints [Kassel et al., 2000; Kassel, 2002] Taxonomy validation [Guarino and Welty, 2000]
 - **Tools:** DOE, FCA, IODE, etc.
- Fabien.Gandon@sophia.inria.fr

- Some work on these steps (II)
- 27
- ♦ **Build // Evolution:** N.L.P., merging, editors, etc. + versioning and coherence [Larrañaga & Elorriaga, 2002] [Maedche et al., 2002]
 - ♦ **Formalisms:** conceptual graphs, description logics, object- / frame- languages, topic maps, predicate logic etc.
 - ♦ **Evaluation // Detection:** scenario and feedback
 - ♦ **Collective dimension:** Reconciler [Mark et al., 2002] designed to aid communicating partners in developing and using shared meaning of terms
 - ♦ **Management:** plan the work like a project existing methodologies e.g., METHONTOLOGY [Fernandez et al., 1997]
 - ♦ **Complex tools and platforms:** Protégé 2000, WebODE, KAON, etc.
- Fabien.Gandon@sophia.inria.fr

- Colleague (I)
- 28
- ♦ **Situations in technology monitoring scenario:**
"... send that news to X and his/her colleagues..."
"... what did X or one of his/her colleagues wrote..."
 - ♦ **Terminological study: colleague term**
 - colleague: one of a group of people who work together
 - colleague: someone who shares the same profession
 - ♦ **Lexicon:**
"colleague: one of a group of people who work together || syn. co-worker, fellow worker, workfellow"
 - ♦ **Table and structure:**
- | Class | View | Super class | Other Terms | Natural Language Definition | Pr. |
|-----------|--------------|-------------|-------------|--|-----|
| colleague | organization | worker | co-worker | one of a group of people who work together | Us |
| ... | ... | ... | ... | ... | ... |
- Fabien.Gandon@sophia.inria.fr

- Colleague (II)
- 29
- ♦ **First formalising**
- ```
colleague(x) ⊃ person(x)
```
- 
- ```
<rdf:Class rdf:ID="Colleague">
  <rdf:subClassOf rdf:resource="#Worker"/>
  <rdf:comment xml:lang="en">one of a group of people who work
  together.</rdf:comment>
  <rdf:comment xml:lang="fr">personne avec qui l on
  travaille.</rdf:comment>
  <rdf:label xml:lang="en">colleague</rdf:label>
  <rdf:label xml:lang="en">co-worker</rdf:label>
  <rdf:label xml:lang="fr">colleque</rdf:label>
</rdf:Property>
```
- ♦ **Problem: one is not a colleague by oneself...**
- Fabien.Gandon@sophia.inria.fr

- Colleague (III)
- 30
- ♦ **Transform into relation:**
- ```
colleague(x,y) ⊃ some_relation(x,y)
```
- 
- ```
<rdf:Property rdf:ID="Colleague">
  <rdf:subPropertyOf rdf:resource="#SomeRelation"/>
  <rdf:range rdf:resource="#Person"/>
  <rdf:domain rdf:resource="#Person"/>
  <cos:transitive>true</cos:transitive>
  <cos:symmetric>true</cos:symmetric>
  <rdf:comment xml:lang="en">one of a group of people who work
  together.</rdf:comment>
  <rdf:comment xml:lang="fr">personne avec qui l on
  travaille.</rdf:comment>
  <rdf:label xml:lang="en">colleague</rdf:label>
  <rdf:label xml:lang="en">co-worker</rdf:label>
  <rdf:label xml:lang="fr">colleque</rdf:label>
</rdf:Property>
```
- ♦ **Problem: no one lists all the colleagues, one derives them from the organisational structure**
- Fabien.Gandon@sophia.inria.fr

Colleague (IV) 31

- ♦ "I am a colleague of X **because** I work in the same group than X"
- ♦ Encode axiomatic knowledge, factorise knowledge in rules and definitions

```

colleague(x,y) ⇔ person(x) ∧ person(y) ∧
  (∑ group(z) ∧ include(z,x) ∧ include(z,y))

```

```

IF
  Group
  Include
    Person ?x
  Include
    Person ?y
THEN
  Person ?x
  Colleague
    Person ?y

```

Fabien.Gandon@sophia.inria.fr

Some concluding remarks 32

- ♦ **Make conceptualisation explicit, visible, operational, etc.**
 - Loosely-coupled solutions
 - Generic mechanisms and inferences
 - Decouple domain dependent aspects
 - Reflection
- ♦ **Ontology as interface / Ontology and interfaces**
 - Communication (H-H, H-M, M-M)
 - Modelling and indexing controlled vocabulary
 - Require intelligent interfaces able to focus
- ♦ **Ontologies have a cost (design, maintenance) to be taken into account in a complete solution**
 - Project management and integrated tools
 - Maintain dependencies
- ♦ Ontology are **not the silver bullet** for KM, **but** an interesting **conceptual object** for building tools and **supporting infrastructures**

Fabien.Gandon@sophia.inria.fr