The European Media Wrapper Round Table-IV

(*Amsterdam, Friday September 11th*)

The Open Exchange Approach

The AXIS dynamics

A*cquisition,* *e***X***change,* **I***ndexing,* **S***tructuration*

Presented by: **Guy Maréchal on behalf of TITAN**

gmarechal@brutele.be

EMWRT-IV PROGRAMME

09H45 - 10H00 : Welcome of the participants

10H00 – 10H20 : Opening of the EMWRT IV
The meaning of semantic in the broadcast industry (Bruno Bachimont – UTC Compiègne)

10H20 – 10H40 : **The Open Exchange Approach** (Guy Maréchal - Titan)

10H40 – 11H10 : A demonstration of a semantic wrapping prototype : starting from production management, mission production management, ingest, semantic transcoding, derushing media segmentation, browsing and exportation for reuse or archival ! (Philippe Scohy)
(MediaMap & Memories projects)

11H10 – 11H20 : Q&A (Charles Bebert – Kane/Titan)

11H20 – 11H40 : From semantic indexation to process management (Steny Solitude – Perfect Memory)

11H40 – 11H50 : Q&A (Maarten Verwaest – VRT/MediaLab)

11H50 – 12H00 : Conclusions of the EMWRT IV

12H00 – 13H00 : Lunch

You just need to cross the road to attend IBC-2009 when opening !

The Open Interchange "Human ↔ Human"

KNOWLEDGE base

- Gastronomy
- Astronomy
- ...
- English
- French
- Chinese
- ...

Matching!

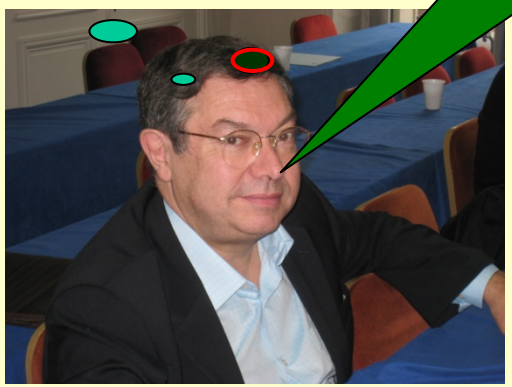
KNOWLEDGE base

- English
- French
- Dutch
- ...
- Philosophy
- Music
- Comics
- ...

Its Guy!
He knows English!

Hello Guy!
I am happy to see you!

Its Charles!
He speaks English!



The Semantic Interchange ["Human ↔ "IT" ↔ "IT" ↔ Human"]

Am I in relation with Guy through a computer or not?

The LIMITS of the AI

Von Newman / Turing / Penrose / Searle / Hofstadter / Dennett / Gödel / Nagel / ...

"Human" ↔ "IT" interchange

User's Interface

Information Technologies

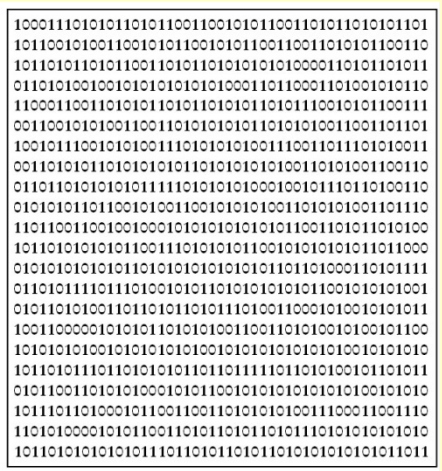
User's Interface

"Human" ↔ "IT" interchange

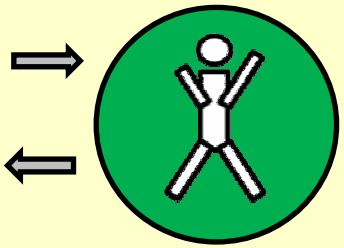
REQUIREMENTS for the Semantic modelling
(representing "Knowledge" through IT)

- **EVERYTHING can covered**
- **HUMAN can express their visions of the MEANING of things (SEMANTICS)**
- **ITC MACHINES can 'understand', 'process', 'retrieve' the semantic items**
  (Through Intelligent Active Agents)
- **Any new semantic item can be added**
- **The representation can go at any level of detail and accuracy**
- **Several representations of the same semantic item can coexist**
  (*multiple point of views and multiple representation formats*)
- **The representations can be distributed**
- **The representations are enabled native persistent**

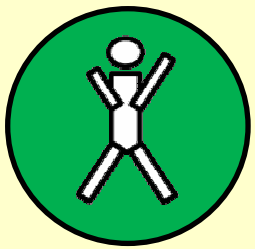
“DATA” modelling



PROCESS



JPEG
ISO Standard



Definitions *(ISO)*

INFORMATION :

The **meaning** that human assigns to **data** by means of **conventions** applied to the data

DATA

A **representation** of facts, concepts or instructions, in a **formalized** manner, suitable for communication, interpretation, or processing by **human** or by **automatic means**

Definitions

SUBSTANCE :

Abstract concept designating the specific thing intended to be represented through data.

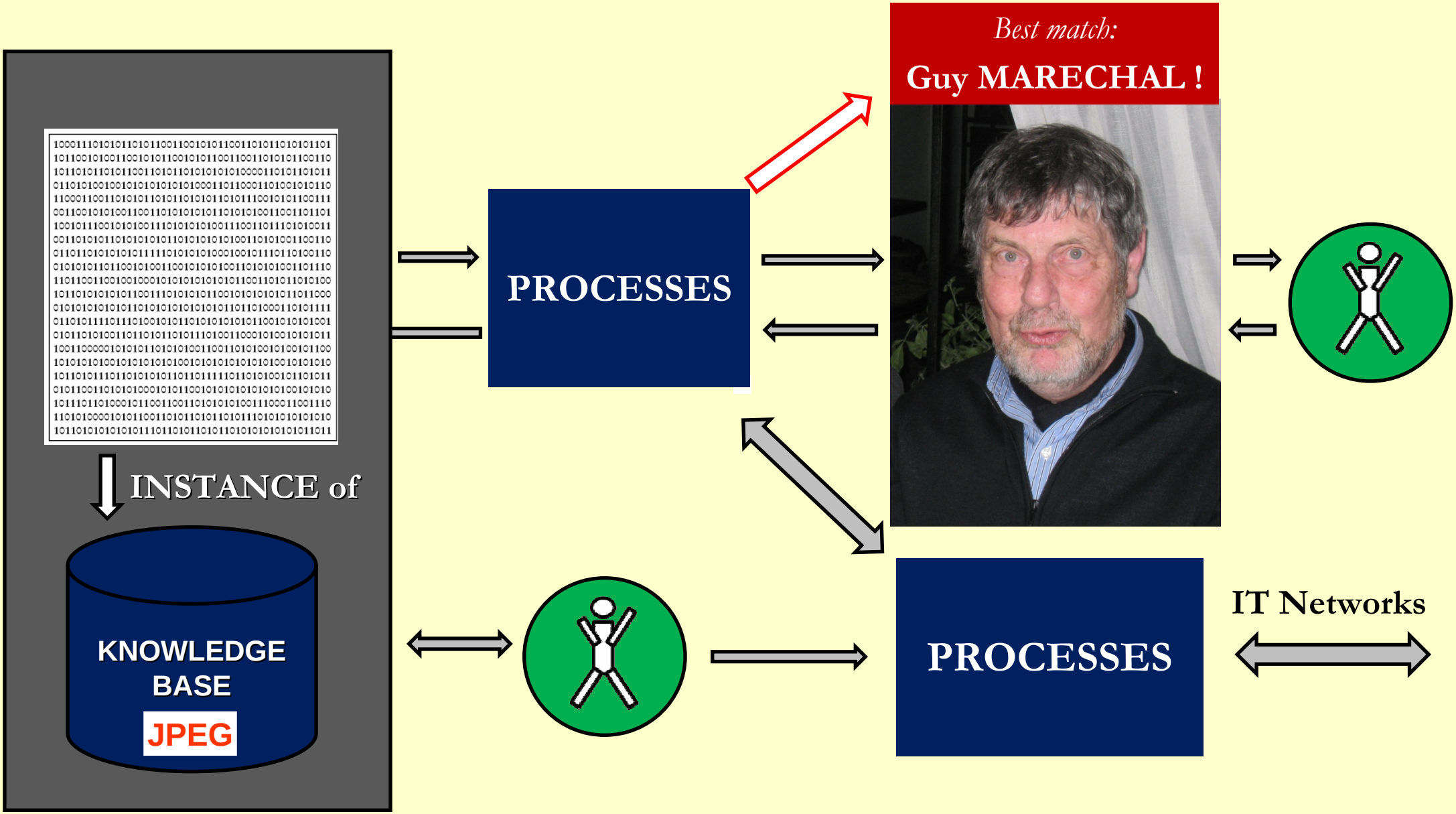
Example: The 'substance' is the information induced from several represented of the song "Yesterday" by the Beatles, coded in **.wav** or **.mp3** or **.ogg**

ORTHOGONALITY :

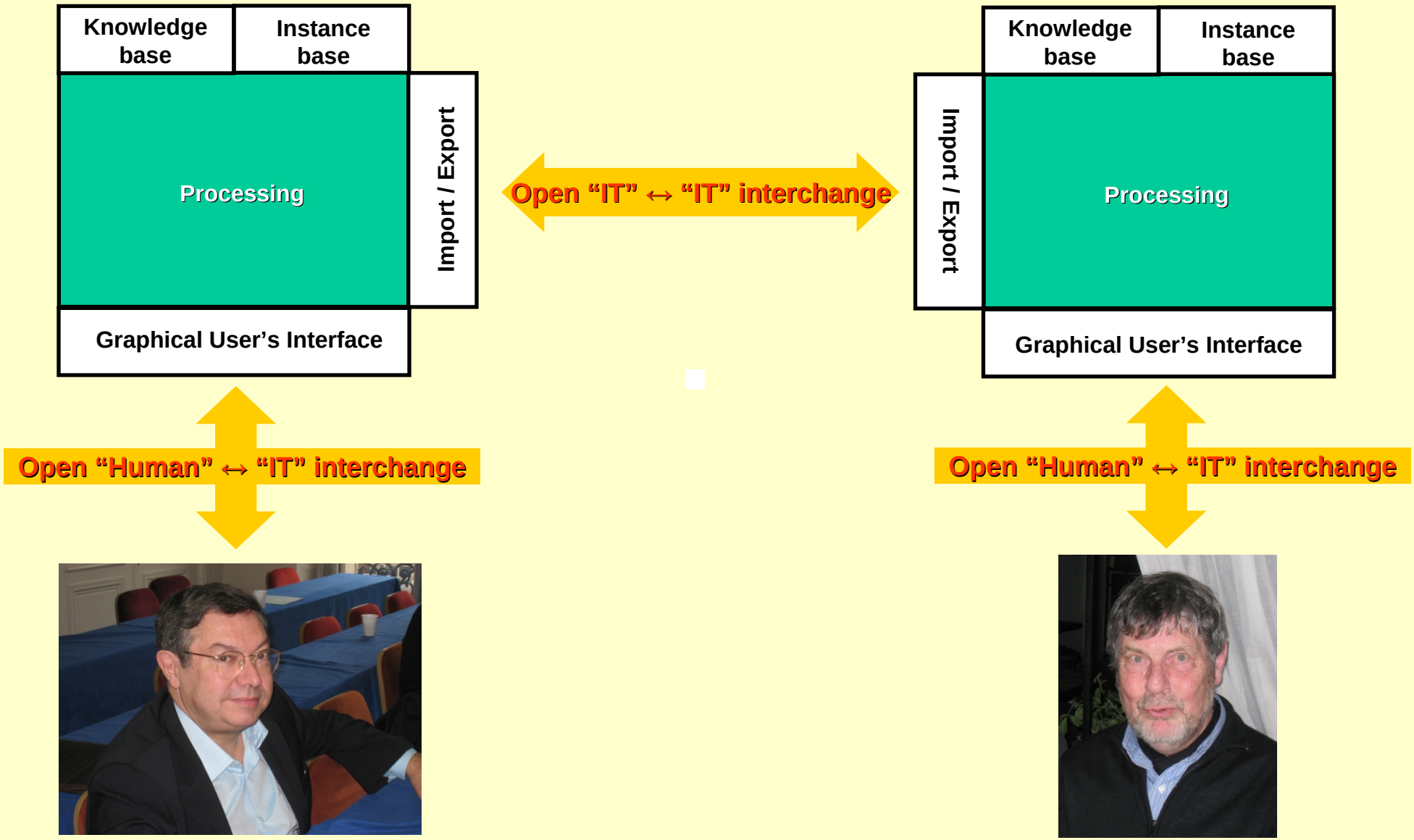
Representations of items, however closely related, are called orthogonal, when they can be modified independently from each other to achieve a particular intention

Example: Some of the data carriers (such as USB stick; CD-R; HDD) are orthogonal with the files and folders they carry.

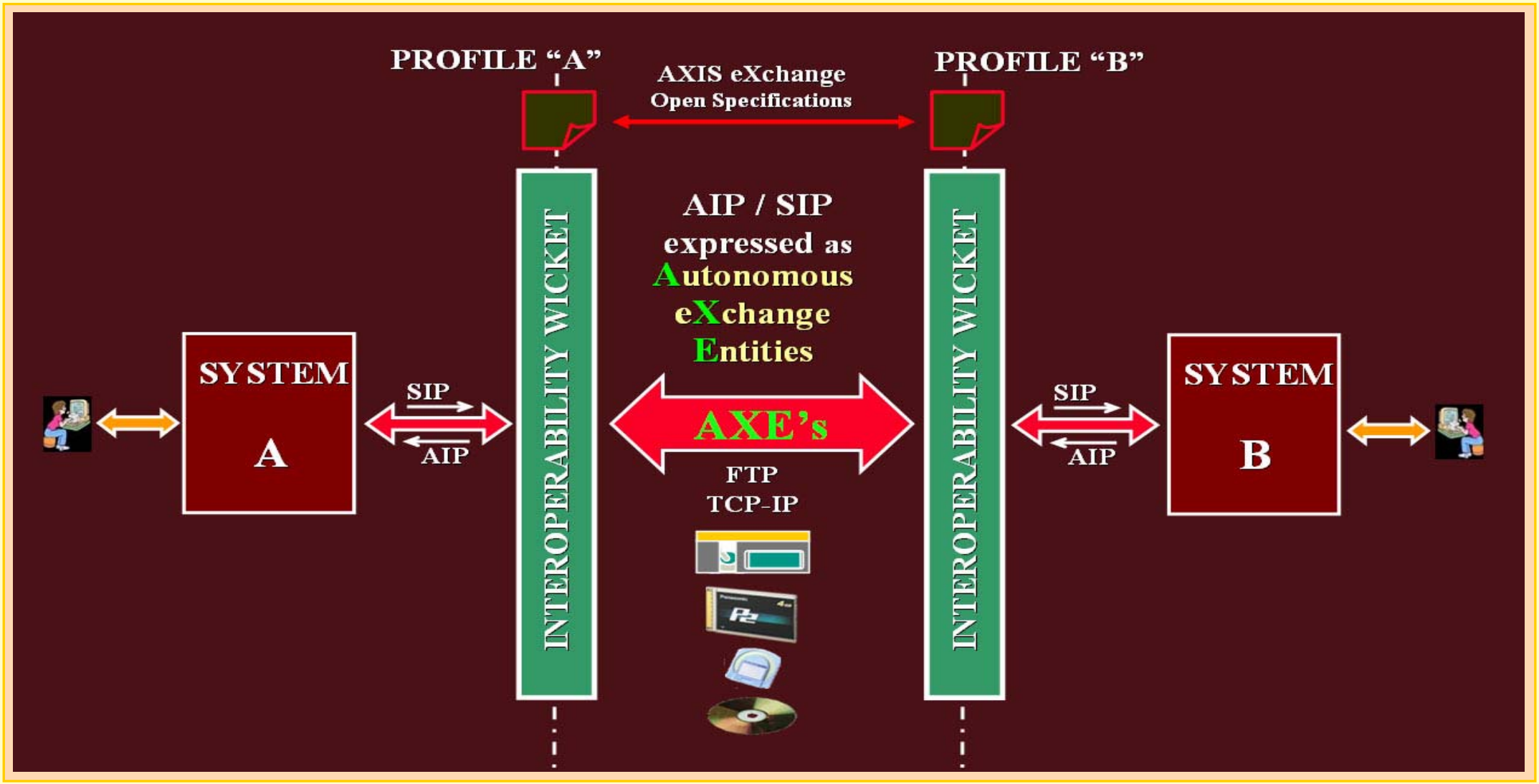
“SEMANTIC” modelling



The Semantic Interchange ["Human ↔ "IT" ↔ "IT" ↔ Human"]



The AXIS Open Interchange



The Contents of an AXE

1. A WRAPPING Technology

2. The AXE ENTITY

3. The INSTANCES of all the ENTITIES required for representing the target exported ENTITY.
   - The configuration management documents (-afp- / -aci-)
   - The metadata documents
   - The essences documents

4. The set of PROFILES required for 'understanding' the INSTANCES
   - The CORE AXIS PROFILE
   - The DOMAIN related PROFILES (Interview; Music; Tennis; News; …)

The KEY contents of a PROFILE

1. The IDENTIFICATION system in the source ARE
2. The ONTOLOGY expressed in a specific IT technology (for example: .owl)
3. The AUTHORITY LISTS
4. The REFERENCES pertaining to dedicated applications
5. The REFERENCES to the STANDARDS not represented in the definition of the ontology

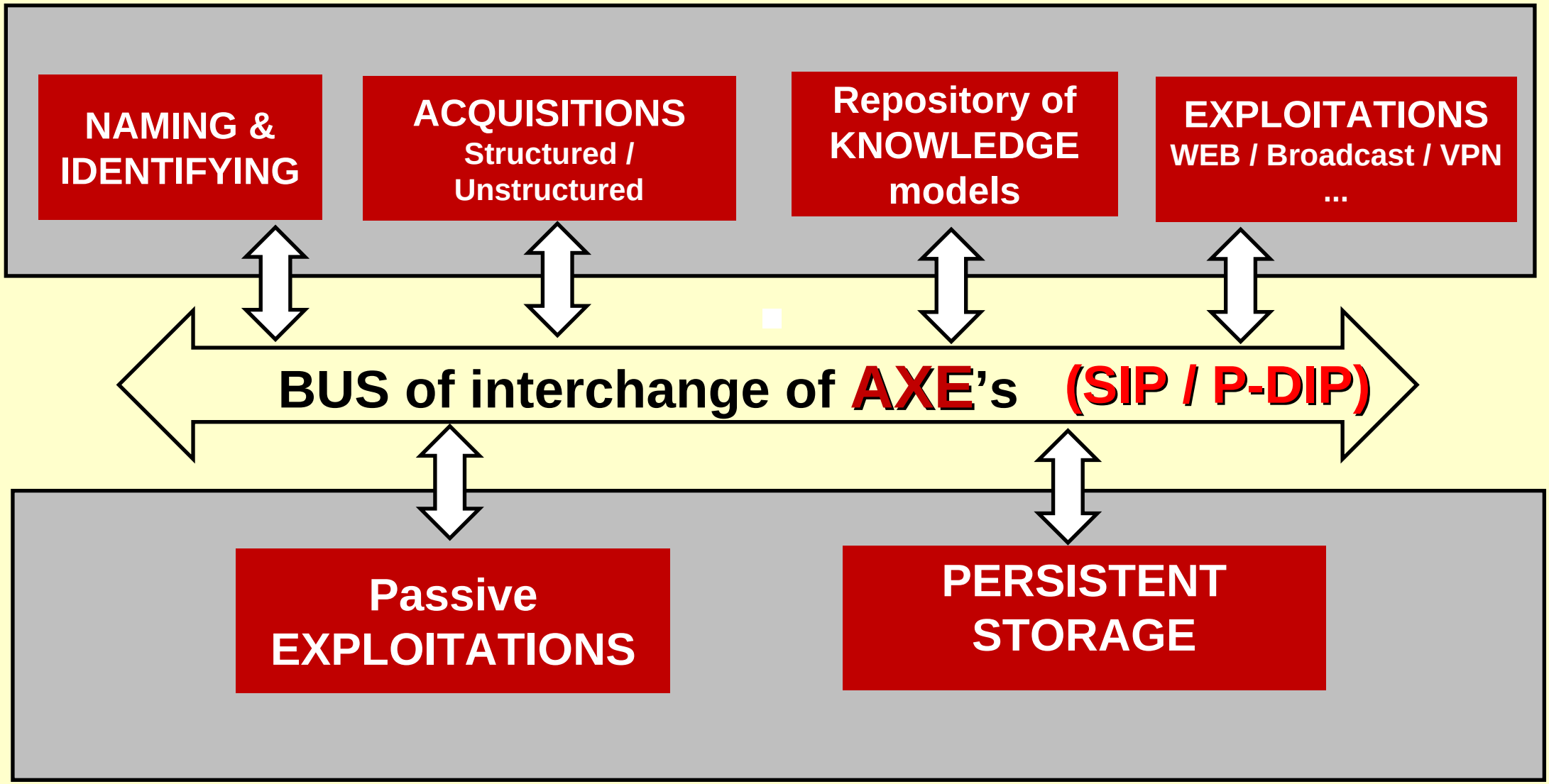
“SEMANTIC” modelling

CONCEPTUAL models

- OAIS
- CRM
- FRBR
- FRBR-oo
- **AXIS**
- ...

AXIS Autonomous eXchanging Indexing & Structuring

AXE Autonomous eXchanged Entities



The structural change in modeling

From simple "FLAT" model (Based on RECORDS in DATA BASES)

AXIS upward COMPATIBILITY

To enhanced "FLAT" model (With added metadata and alias)

AXIS upward COMPATIBILITY

To "RICH" semantic model (Based on Networked ENTITIES with Documents & Relations in KNOWLEDGE BASES)

- Ontology based
- Object oriented
- Native persistent

& < Semantic WEB applications >

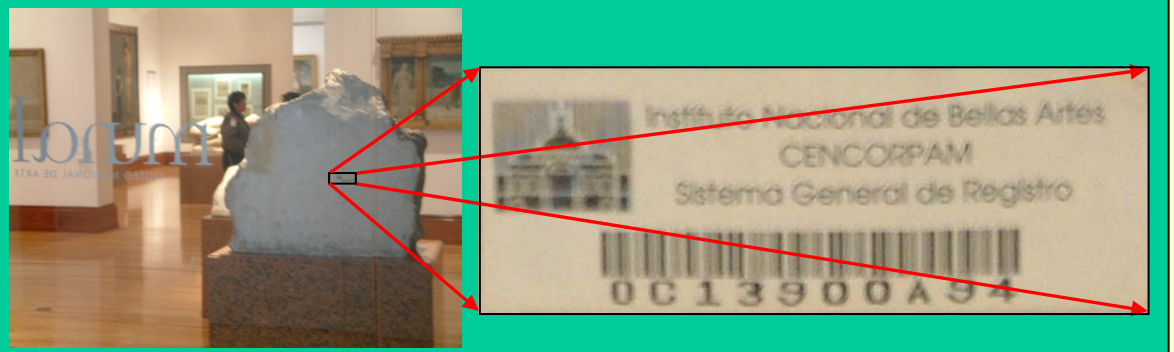
FLAT MODEL

The FLAT model is based on "RECORDS" representing 'CATALOGUED' resources

RESOURCE / ASSET

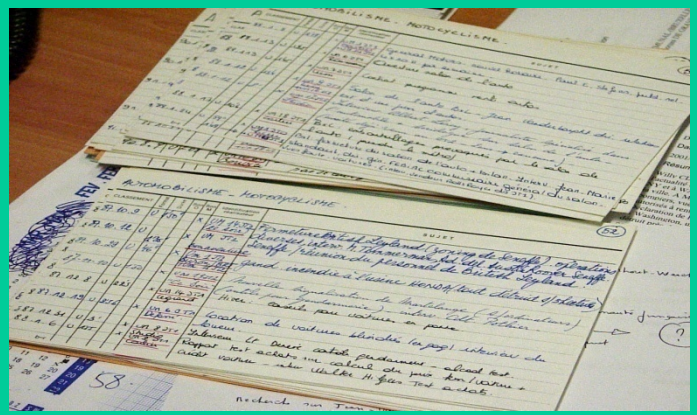
CATALOGUING:

- Cataloguing rules (MARC21 / MODS / …)
- Identification & link to the resource



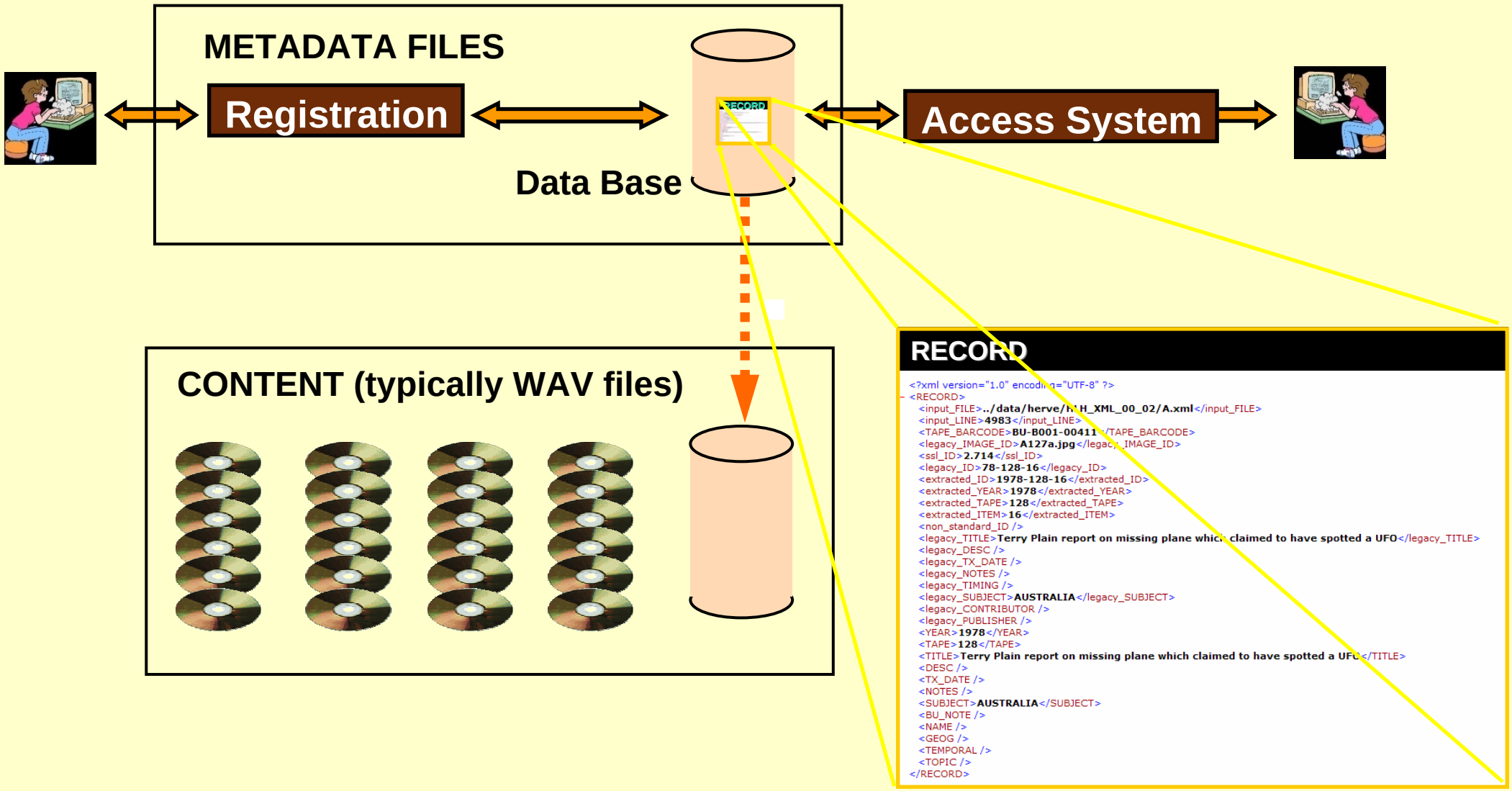
CATALOGING:

- Catalogue record



FLAT MODEL

The "RECORDS" are usually represented as a collection of instances of "TERMS"
Stored in "DATA BASES" and linking to "CONTENTS"



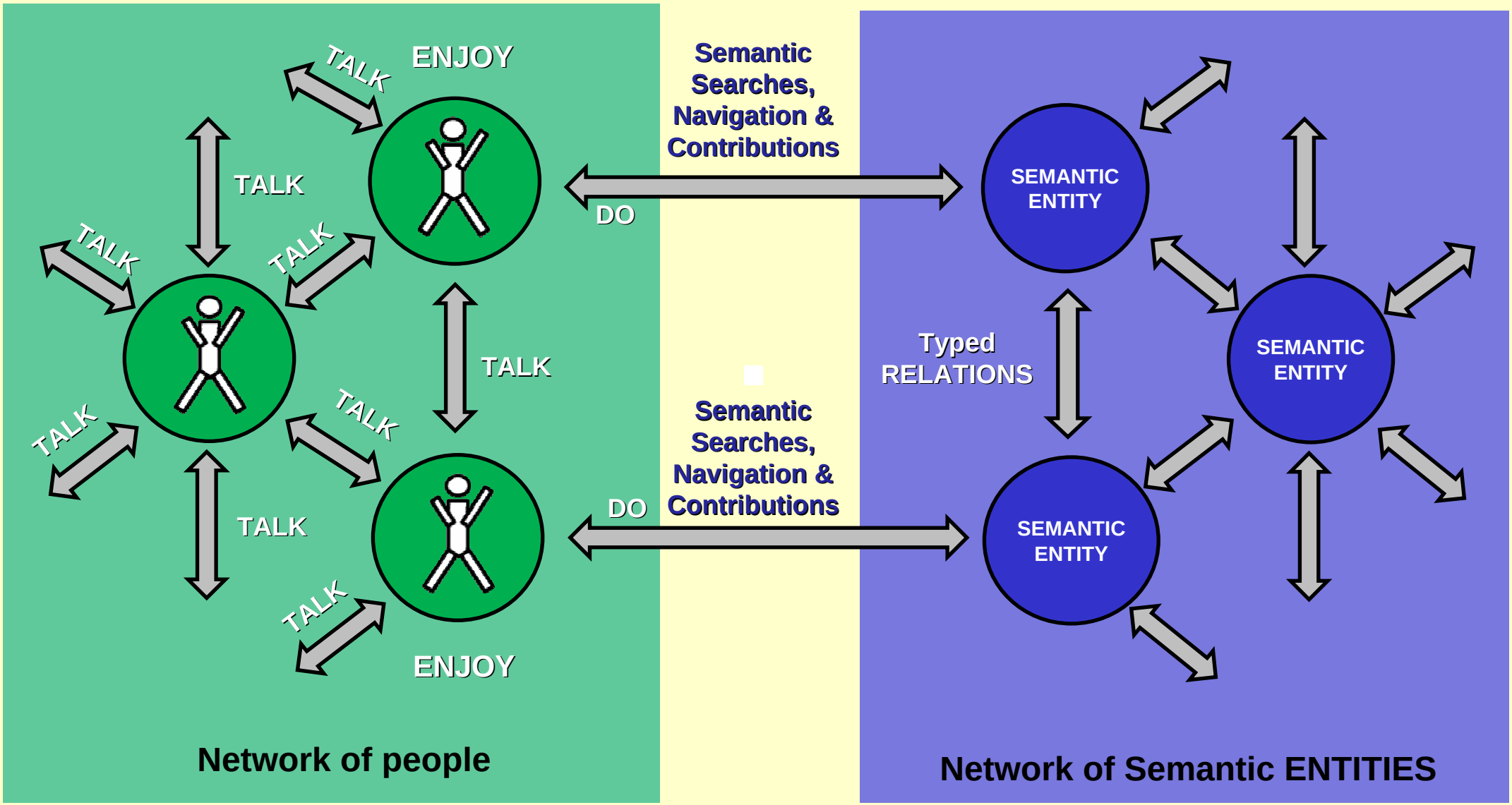
FLAT MODEL

The RECORDS are usually based on domain specific standards

- **MARC**
- **MODS**
- **…**

The use of generic standards, like ‘Dublin-Core’ or ‘VRA-Core’ remains ancillary or not fully mapped

Accessing, Creating, Enriching, Sharing “RICH SEMANTIC” models



An illustration of the differences between the “RICH” models and the “FLAT” models

In “FLAT” model:

MOZART is namely a COMPOSER *(The Dublin Core says contributor!)*:

<dc:contributor> Mozart </dc:contributor>

Here Mozart is a string of characters, not an ENTITY

In “RICH” semantic model:

Wolfgang Amadeus MOZART is first a **PERSON** having a **LIFECYCLE** modeled independently as one **ENTITY** (or several **ENTITIES**). He was born, travelled, played music, composed, conduct, got married, …

The fact that he contributed significantly to the composition of the **REQUIEM K.626** and his specific contribution can be expressed;

The same for the specific contribution of M. **Süssmayr**!

In “RICH” model, the contribution as composer of Mozart and the one of Süssmayr are expressed as **RELATIONS** in a **CONTEXT**: each has played the role of one of the composers for specific parts of the REQUIEM K.626!

The UPWARD compatibility between “FLAT” and “RICH” implies rich ALIAS mechanisms

An illustration of the differences between the “RICH” models and the “FLAT” models

The handling of “CONTEXTS” in “RICH” semantic models:

- **Human languages (Entries; Values; Qualifications)**
- **Semiotics**
- **Time references**
- **Space references**

ILLUSTRATION

A musical opus of « **Pyotr Ilyich Tchaikovsky** » commonly named the « **1812 overture** ».

Composer

Name:

- Pyotr Ilyich Tchaikovsky (english)
- Piotr Ilitch Tchaïkovski (french)
- Пётр Ильич Чайковский (russian)
- …

Born:

- 7 May 1840 [‘Gregorian’ New style calendar]
- 25 April 1840 [‘Julian’ Old Style calendar]
- …

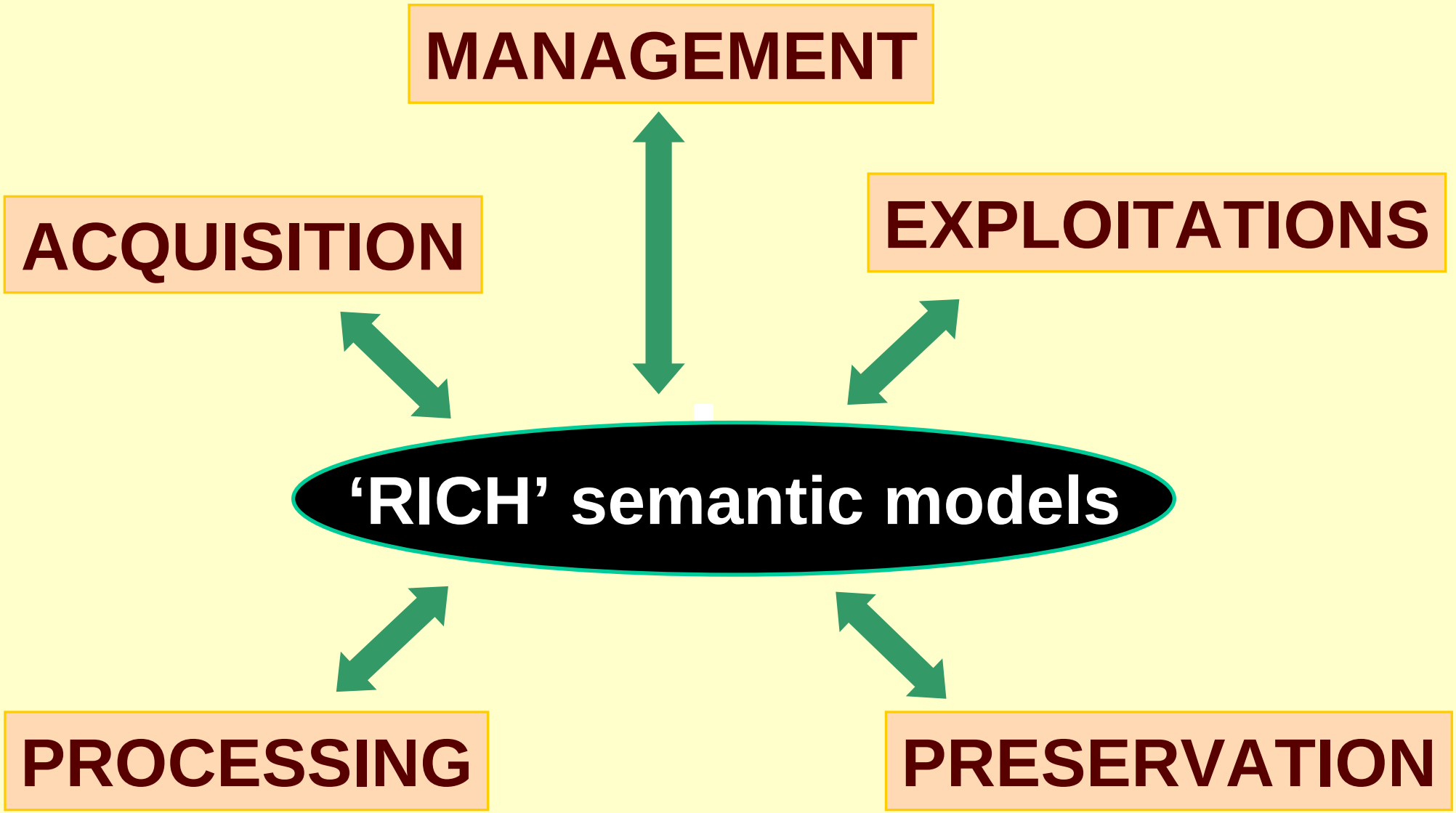
Died:

- 6 November 1893 [‘Gregorian’ New style calendar]
- 25 October 1893 [‘Julian’ Old Style calendar]
- …

Musical Opus :

- Festival Overture, The Year 1812 (english)
- Ouverture Solennelle, L'Année 1812, Op. 49 (french)
- Торжественная увертюра 1812 года, *Toržestvennaja uvertjura 1812 goda* (russian)

The linking of the ASSETS in all domains!



The RICH model namely ENABLES:

The modelling according to the FRBR standard

The modelling according to the CRM standard

The implementation of the FRBR-CRM OO

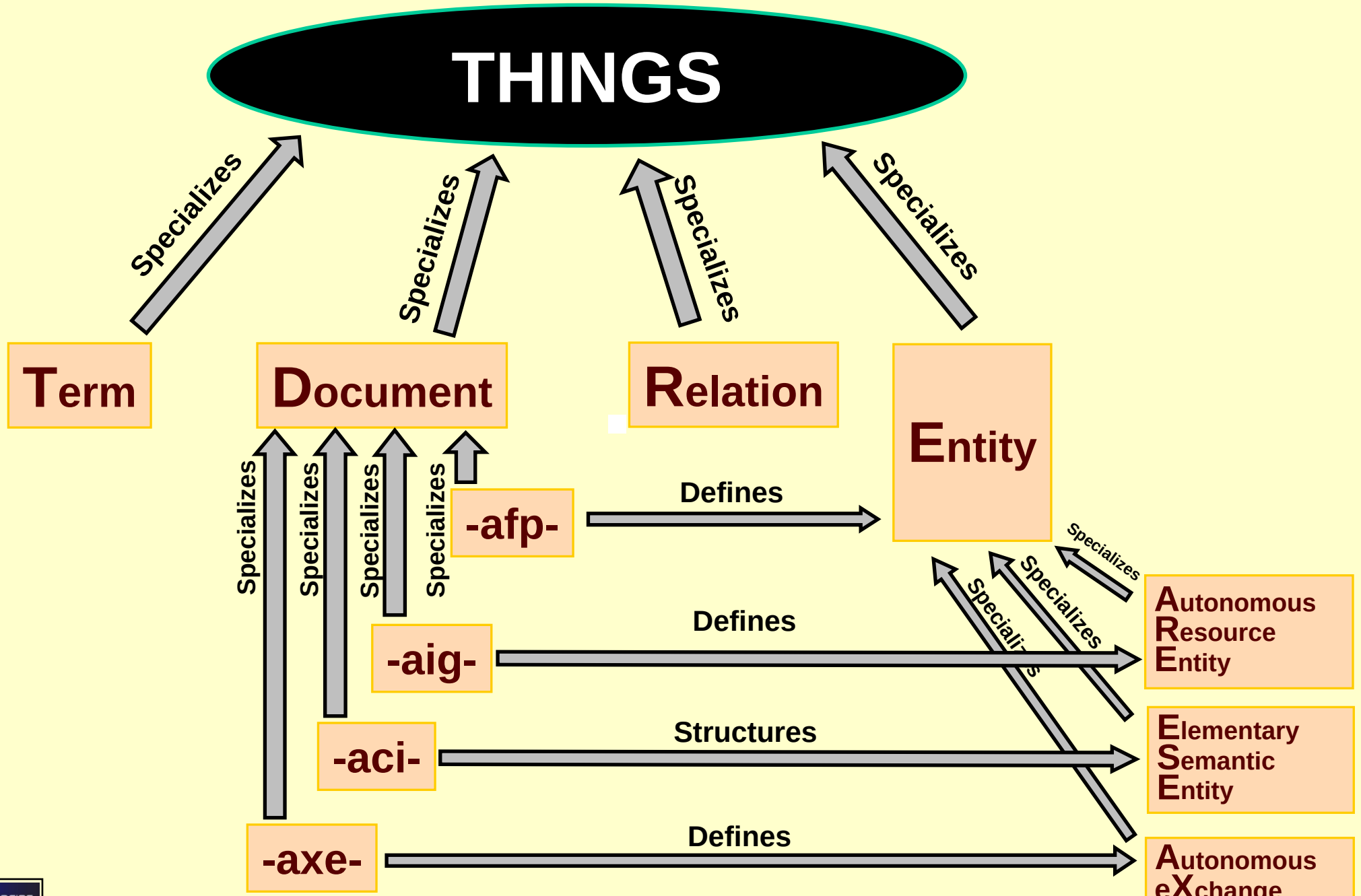
Fits perfectly with the “Object Oriented” programming; the “Ontology” approach

The implementation of the “Topic maps” and the “Ontology Web language”

…

The “AXIS” modeling

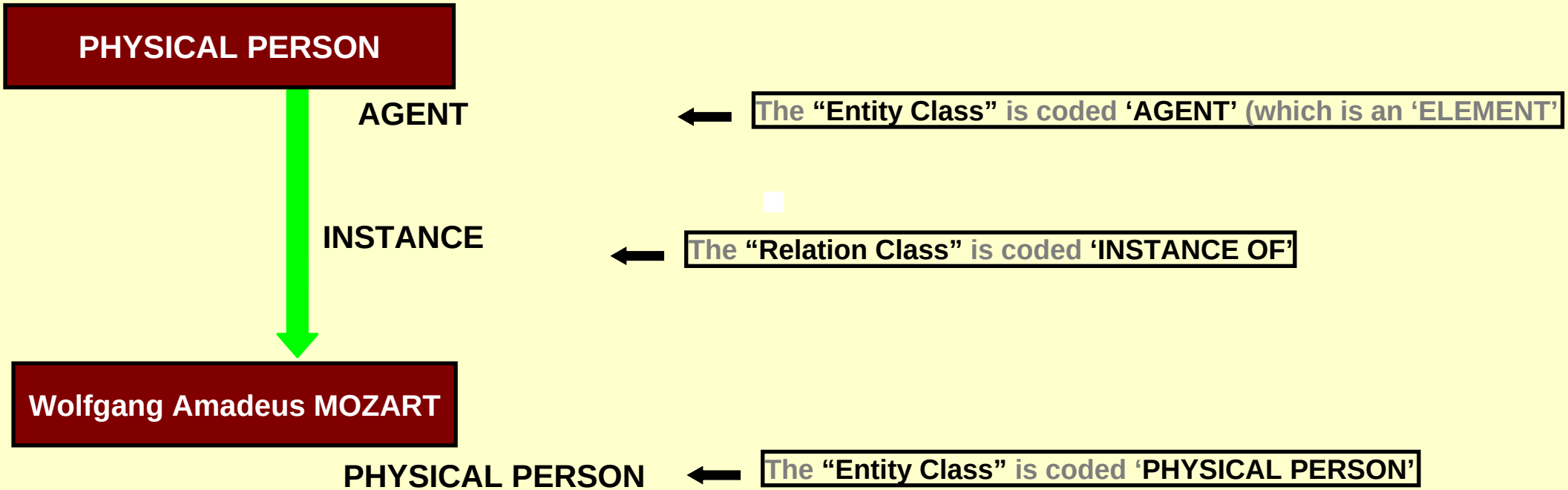
The key constructs of the AXIS "CORE"



Introducing the concept of ENTITY

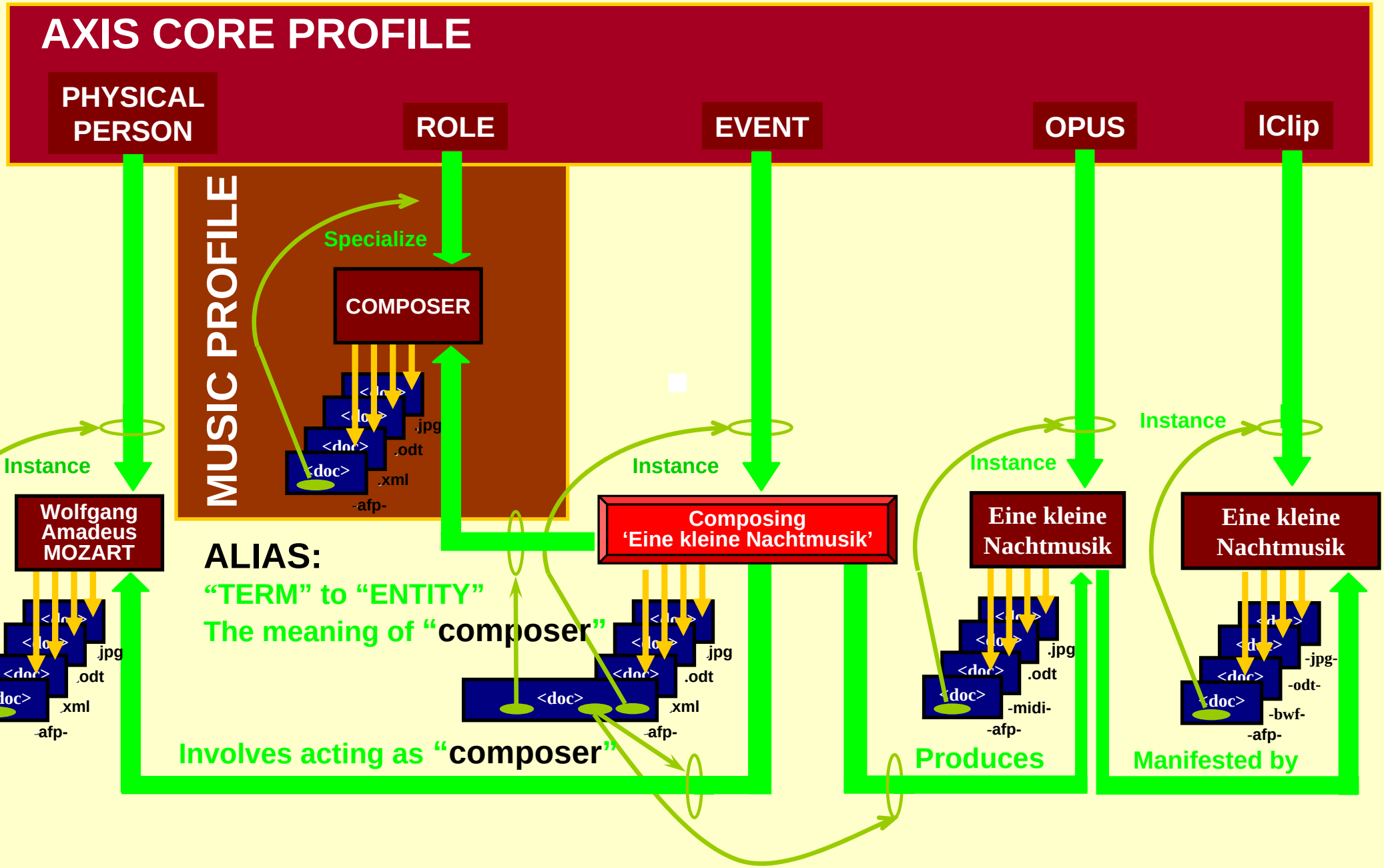
The ENTITY relates to one "SEMANTIC SUBSTANCE"

A concrete example: **"Wolfgang Amadeus MOZART" is a person which is represented through the documents attached to an ENTITY**



That ENTITY concept ensures the compatibility between the "Topic maps' and 'Ontology' approaches

A composite concrete example:



Description of rich semantic model

The "RICH" semantic model is based on "ENTITIES" holding "DOCUMENTS" & "RELATIONS" in "KNOWLEDGE BASES"

- **Identifying the ENTITIES, DOCUMENTS and TERMS**
- **Enriching the models of the ASSETS**
- **Enriching the ASSETS with SEMANTIC**
- **Adding the structures (logical & physical)**
- **Managing the ASSETS**
- **Persistence & Interoperability of the ASSETS**
- **Retrieving the ASSETS**
- **Representing the ASSETS**
  - ***Semantics: Ontology's; Topic maps; OWL; …***
  - ***Resources: RDF; URN; URL, URI, …***
  - ***Terms: DCMI; …***

SYMBOLISM

ENTITIES

- Bag entities

ORDER BL-022

OPUS

Class of ENTITY

- General entities

MOZART

PHYSICAL PERSON

- Media entities

REQUIEM K626

LOGICAL CLIP

- Process entities

DIGITIZATION

EVENT

DOCUMENTS

Media of …

-wave-

Ontology of …

-owl-

Class of the DOCUMENT

RELATIONS

IS PART OF

Class of the RELATION

The SYMBOLISM illustrated by a concrete example: ‘COMPOSER’ is a ‘SPECIALISATION’ of ‘ROLE’

‘COMPOSER’ is an ‘ENTITY’

The “Entity Name” is coded ‘COMPOSER’

-afp- means: ‘axis foot print’

‘COMPOSER’ is defined by ‘Documents’

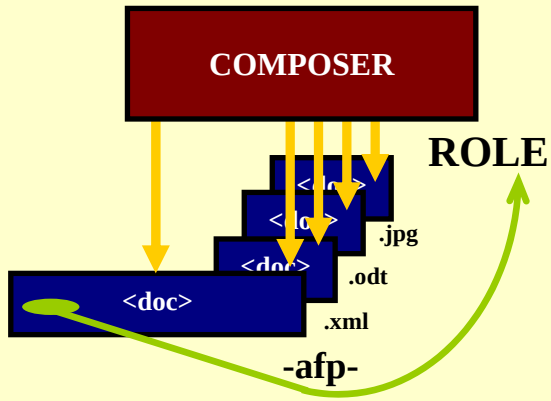
One of these ‘Documents’ is of the -afp- ‘class of documents’. It includes:

- **The fundamental metadata defining the entity (based on RDF and Dublin Core) including the ‘Entity Class’**
- **The list of documents owned by the entity**
- **The list of links to documents pertaining to the entity but owned by another entity**
- **The list of the entities owned by the entity**
- **The list of the links to the entities pertaining to the entity but owned by another entity or ARE**
- **The list of the links to other autonomous resources and objects**
- **The list of ‘ALIAS’: Terms – Entities Terms –Terms Entities – Entities**

The other of these ‘Documents’ are of any ‘Classes of document’

The entity ‘COMPOSER’ is a ‘SPECIALISATION’ of the entity ‘ROLE’

The –afp- document expresses that ‘COMPOSER’ is a ‘SPECIALISATION’ of ‘ROLE’ and that ‘SPECIALISATION’ is the type of the ‘RELATION’



Notes:

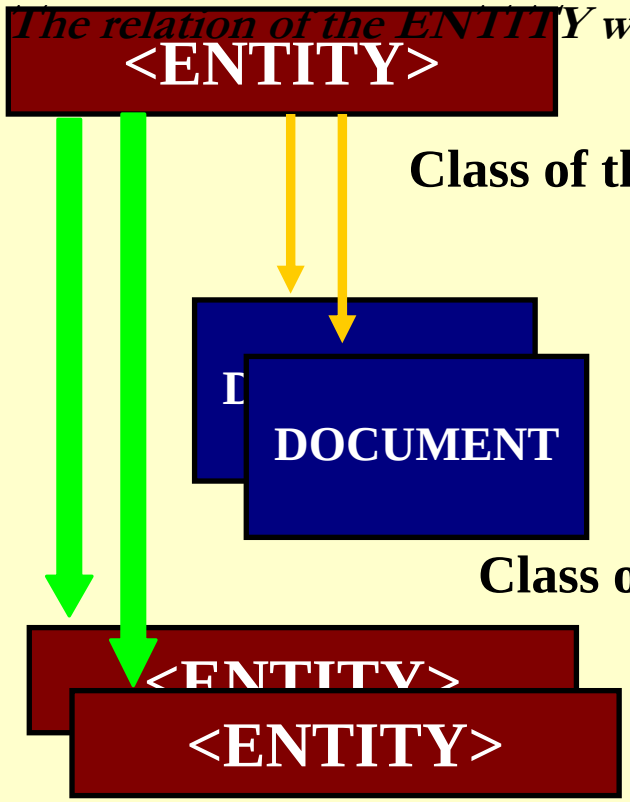
The ‘class of document’ relates to an abstract concept [the class is represented using the hyphen as separator (such as -afp-)].

While the embodiment of a document as a file, or into a file, uses the dot as end separator called ’extensions’ (such as .odt).

DEFINITIONS & SYMBOLISM

*An **ENTITY** is an abstract or concrete 'resource' modeled by one or a set of **DOCUMENTS** and/or **ENTITIES:***

- *Each ENTITY has characteristics identified by 'modeling technology' coded by its Class of ENTITY.*
- *The documents owned by the ENTITY are shown attached by small orange arrows:*
- *The relation of the ENTITY with another ENTITY are shown as large green arrows:*



***A DOCUMENT** is a reachable 'resource' expressing a model of one or several **ENTITIES** according to one identified 'modeling technology' coded by its Class of DOCUMENT.*

EMWRT-IV PROGRAMME

09H45 - 10H00 : Welcome of the participants

10H00 – 10H20 : Opening of the EMWRT IV
The meaning of semantic in the broadcast industry (Bruno Bachimont – UTC Compiègne)

10H20 – 10H40 : **The Open Exchange Approach** (Guy Maréchal - Titan)

10H40 – 11H10 : **Demonstration** of a semantic wrapping prototype : starting from production management, mission production management, ingest, semantic transcoding, derushing media segmentation, browsing and exportation for reuse or archival ! (Philippe Scohy)
(MediaMap & Memories projects)

11H10 – 11H20 : Q&A (Charles Bebert – Kane/Titan)

11H20 – 11H40 : From semantic indexation to process management (Steny Solitude – Perfect Memory)

11H40 – 11H50 : Q&A (Maarten Verwaest – VRT/MediaLab)

11H50 – 12H00 : Conclusions of the EMWRT IV

12H00 – 13H00 : Lunch

You just need to cross the road to attend IBC-2009 when opening !