What is “Knowledge Management”?

Roger Roberts (RTBF/Titan)
Guy Maréchal (Titan – Prosi)
Steny Solitude (Titan - Perfect Memory)

The Information Society is on its way and not only by or through the Broadcast Industry! Increasingly, the Media industry will have to publish rich media on multiple infrastructure and devices! Rich media requires languages that can handle metadata’s according to their properties and their relations in a process managed by computer tools.

Publish rich media means moving from a data management system controlled by humans to a collaborative production process where machines manage migration, exchange and archiving. So it is necessary to manage the data, applications for data representation and especially the knowledge base that provides the link between the data and its meaning.

Worldwide, there are standards (such as ISO/OAIS) and standardized languages (like RDF, OWL, ... standardized by the W3C) which provide a scientific basis for this important technological move!

This technological evolution is not only scientifically based, but in addition it provides new advantages in terms of information management and profitability.

Semantic technologies are made to ensure the enhancement of data’s, manage the sustainability of the link between the data and information (a meta/data), allowing the exchange of rich models between systems, there is the limit of “flat models” … they can’t offer the required agility in multiple process, in space (exchange) and time (sustainability)!

But semantic also generates enrichment based on some core metadata’s during the production process. By using explicit metadata such as “physical person” (with different roles: writer, director, actor, ..), “corporate body” (company, organization, ..), “location” (country, city, ..), “date”, “work”, “event”, ... the semantic middleware will generates inferences on the basis of information modelling. The results are deduced from explicit links (properties or relations) and not lexical approximations and or statistical computation.

Management and knowledge representation

Knowledge Management includes a range of strategies and practices in an organization to identify, create, represent, distribute and enable adoption of perspectives from experiences. These experiences express knowledge, either embodied in individuals or embedded in organizations such as processes or practices.

More recently, several areas have begun to contribute to research in knowledge management, including computer science, information management and media.

Until recently, the knowledge structure was made for consumption by humans not by machines. The results were described in natural language and formalized in textual forms. This is not a problem for humans who easily handle the semantic (signifier / signified). But it becomes a problem for computers which must be able to interpret and use knowledge. The computer oriented knowledge structure will allow a better and smooth flow of knowledge between the theoretical and the practical world.

From Management to knowledge representation

In a scientific approach**,** the instrumentation strategy of knowledge management begins with a deconstruction of knowledge, both theoretical and practical.

In fact this creates a taxonomy which was first the science created for the description of living organisms and group them into entities called taxons, to identify them and name them, and then classify them. It complements the systematic science which is organizing the classification of the taxons and their relationships. The taxonomy has spread to other sciences such as human, information or computer science. (The taxinomy term used in some scientific domains is not equal and covers only restricted parts of the Taxonomy).

In the AV world, this approach results in many meta-models describing the terms used to describe objects and processes. These models are called "flat" because they do not qualify the description of the process and the properties and relations of terms in languages accessible to computer tools!

In the audio-visual world the instrumentation of knowledge management in process aims to:

- extract the essential concepts and relationships;
- define them by their meaning, their attributes, and semantic constraints

It appears in this case that the ontological categories of an AV process are:

- the actor (subject, something that produces actions),
- behaviour (verb, action, phenomena),
- the subject (something that is handled by the actor)
- the object (the prescription, the editorial),
- the situation (a context in which the action is done),

This implementation work in tools for knowledge management is based on several assumptions:

- The first assumption is that "everything is real," and that the reality is defined as being dimensioned by time and space.
- The second assumption is that we are dealing with two kinds of worlds:
    - A concrete world (technology, images, sounds, the script, editing, ...) to describe from what is possible and existing
    - An abstract world (the imaginary, the media, culture, creation), with a description of existing theories, a reification (the process by which an abstract idea about a computer program is turned into an explicit data model or other object created in a programming language) with optimal fidelity. This is the world of concepts.

The work is to represent these abstract and concrete worlds, and especially how they relate to each other. This can be done in a Configuration Management System (CMS) which organizes the relationship between terms and entities, the relationship to physical objects, interoperability between the Upper Ontology and domain ontologies, the structuring of data in the database (storage / export)..

The third assumption is that several ontologies can coexist within a "Configuration Management System":

- An Upper Ontology (top-ontology, or foundation ontology) which describes very general concepts identical to several fields of knowledge. The most important function is to offer a very wide semantic interoperability in a large number of domain ontologies.
- A domain ontology coherent for audio-visual applications
- An ontology that provides backwards compatibility with existing audio-visual universe (flat models)
- A context ontology of the usage for a distribution of tasks, depending on / by the end user and / or by an intelligent system, i.e. the tasks of design / build (task ID) or a sub-task creation (development).

The fourth assumption is that concepts are "things" in the world of theories. One of the real difficulties is how to differentiate the relationships and constraints in an ontology of those objects in the real world.

Knowledge exploitation

In addition each theory has an own view on the targeted world. Although the assumptions of distinction and classification are not perfect, sometimes questionable, they are reasonably sufficient to serve the goal of knowledge representation independent of applications and thus offer improved AV management solutions.

One of the most important one is the management of the interoperability. Three main interoperability targets are in sight:

- The **migration of data models**. Most of the current information objects are currently represented by "Flat" models usually as record in a relational data base. The trials have confirmed the capabilities of migration of these 'Flat' models to 'Semantic' models by a two steps process. First 'enhancement' with alias; then 'enrichment' by the addition of the semantic data and the insertion of the Information Objects (OAIS term) in the semantic network.

- The **interchange of Information Objects** between systems. When Information Objects are represented in a semantic way and that the document modelling technologies and formats are explicitly available, it is possible of implementing "Inter-Operability Wickets" which could operate in a quasi-full automatic way.
  This allows the interoperability between independent systems because the meaning of the data can as well be understood by the human but also by the IT machines.

- The **preservation of Information Objects**. The Information Technologies are permanently changing! The data representations of the Information Objects become then quickly obsolete. The preservation of the value and the access to the Information Objects is essential. Here the interoperability solution is simply the same as the previous one. The interoperability of the representations between their representation formats valid a few years ago and those according to the current today standards or practices. It is the interoperability in time which paves the way to the persistence! It means that the assets have to include constructs that implies their capacity of evolving according to the properties expressed for the persistence (in particular in the ISO standard OAIS).

“gist” [1]: an example of a minimalist Upper Ontology.

“gist” is different from other upper ontologies in that it covers a very broad range of future applications with the fewest number of concepts. The concepts are about the structure of the information, and aren't really semantic. Most of them can be put in any structures like the abstraction tree entity/relation/attribute, or class/subclass or thing.

“gist” is based as much as possible on “primitive concepts” that have been found by researchers in linguistics as well as fMRI studies (Functional magnetic resonance imaging) of how the brain categorizes information and added to this, a small number of additional primitives that seem to be essential in building “enterprise ontologies”.

Here are the even more primitive concepts, which cannot be readily represented in an ontology but which may help with the thought process of where these primitive concepts came from:

- “now” is the most primitive concept of time and is typically a system function.
- “here” is our most primitive place concept and is implemented in GPS devices.
- Building is one of the more primitive concepts for things that are “located.” “In a place” is a generalization of the concept of “home”.
- “me” is implemented in systems as the person logged in or authenticated.
- “It” or “Stuff” is a catch all for “things,” mostly in the real world.
- “us” is the basis for organization. Interestingly it shows up in the subjectivity of most systems. The difference between a receivable and a payable is who “us” “is.”
- “sezwho”: there is a pragmatic problem with governmental organization: if the UN recognizes a country, it’s a country, and that country can recognize, and thereby bring into existence, its states and local governments.
- “this” is object oriented, generally referring to a content representation that might be a document, or content standing for a real-world item.
- “must” is often used in requirement statements, but it also is the simple way of describing our obligations about our future behaviour.
- “do” (as in the generic programming command) is the general way we describe a unit of behaviour (“do this”).
- “want”: the hedonistic (pleasure is the only intrinsic good) part of the ontology.

Time primitives:

- Interval: has a beginning and end. It has a duration, but it isn't a duration.
- TimeInstant: is a location on a timeline with no specified length in universal time. (It might be instantaneous, but it might also be a date)
- LocalInstant: most systems, and most people, still rely on local time.

Place primitives:

- GeoSegment: A segment connects two points.
- Route: is an ordered set of segments.
- TimeZone: is a special kind of region that we associate with local time changes.
- GeoPoint: is a point locatable on or near the surface of the earth.
- GeoRegion: is a closed area on the surface of the earth.

"in a place" primitives:

- BuildingAddress: is a convention for finding a building.
- Artifact: is a physical item that can be located in space, but not necessarily permanently (such as a car).
- Room: is a part of a building, and is used enough in system and generally understood to be included here.
- Landmark: is a generalization of building which includes things like trees and bridges.

Living thing primitives:

- Unique item; need not be living, but all the living things are unique.
- Living things is a generalization from persons, and includes animals and plants (if they are individually identified).
- People/beings: Person is any homo sapiens that is living or has ever lived (does not include fictional characters).

Stuff:

- Money: behaves like a substance (that is, you measure it quantitatively), but other than coin and currency, has no physicality,
- Substance has mass and volume but is non-corporeal (that is, you can divide it in half and still have the same stuff, like water).
- PhysicallyIdentifiableItemis: the intersection of unique item and substance. In theory you should be able to put an RFID tag on a PII.
- Organizations:
- Corporations are legally recognized but informal organizations (departments or joint signers on a contract) are also often impromptu organizations.
- Organizations are groups that are recognized as acting in concert.

Content (Documents):

- Addresses are content that describe where something is
- Signature, Text and Messages are special types of content, signature being content that is a stand in for authentication, Text being content that is in a syntactical expression and Message being content between two or more agents.
- Media (how content is stored).
-

Agreement:

- Offerings are initially one sided agreements that become binding when someone qualifies and accepts. A product offering is for things that can be inventoried.
- Agreement is made up of at least two obligations; for instance, a sales order is an obligation to ship and an obligation to pay. This is the bedrock of commerce. Almost all commerce is making and following up on agreements. Agreements are between two or more parties.

Behaviour: the primitives for work flow and decisioning systems. Human generally know physical behaviour and generalize non-physical behaviour from that. Almost all physical behaviour (for instance, all manufacturing processes like drill, paint, ship, load, deburr, etc.) are specializations of move or convert.

Goals/Intentions: This is the "teleologic" part of the ontology, which, has to do with causes and goals and reasons for doing things. Example: 'Intention" the reason why you perform a task or project, and also Criteria, Rights and Restrictions which seem to be related, but may not be completely decomposed.

Measures and Units of Measure: the seven foundational units from "System International" (the metric system: kilogram, meter, candela, second, ampere, kelvin, and mole) from which all other physically measurable units can be derived. "count" and "currency" have been added for convenience reasons.

It is possible to add units, but for them to be compatible and convertible they need to have a base unit and a conversion to the base, e.g., you can add furlongs (distance in imperial units) as a unit as long as you state how many meters to a furlong (0.9144 metre).

"gist" is freely available.[2]

2 Click for download at: http://www.gist-ont.com/data/50643f60-cdd7-4e9d-b080a543eb7c62a1/files//gist.owl

CMS : Configuration Management System (description based on the concepts of AXIS-CRM[3])

Configuration management (CM) is a field of management that focuses on establishing and maintaining consistency of a system performance and its functional and physical attributes with its requirements, design, and operational information.

The Titan association has elaborated a modular and tailorable approach for the semantic modelling of static and dynamic knowledge under the name “AXIS Conceptual Reference Model” (AXIS-CRM).

AXIS and its implementation permit the covering of any semantic domain. One of its innovative contributions is the introduction of the “CORE ontology” and of an “Ontology Plug-in” mechanism required for addressing specific domains, whether in economic, cultural or scientific assets.

The methodology has been tested by derivation of existing ontologies such as the FOAF and IAML.

The authoring of these specific ontologies is made according to a strict procedure involving experts of the domain and a four step process (Terms; Relations; Documents; Profile) assisted by a “Knowledge Modelling Tool”.

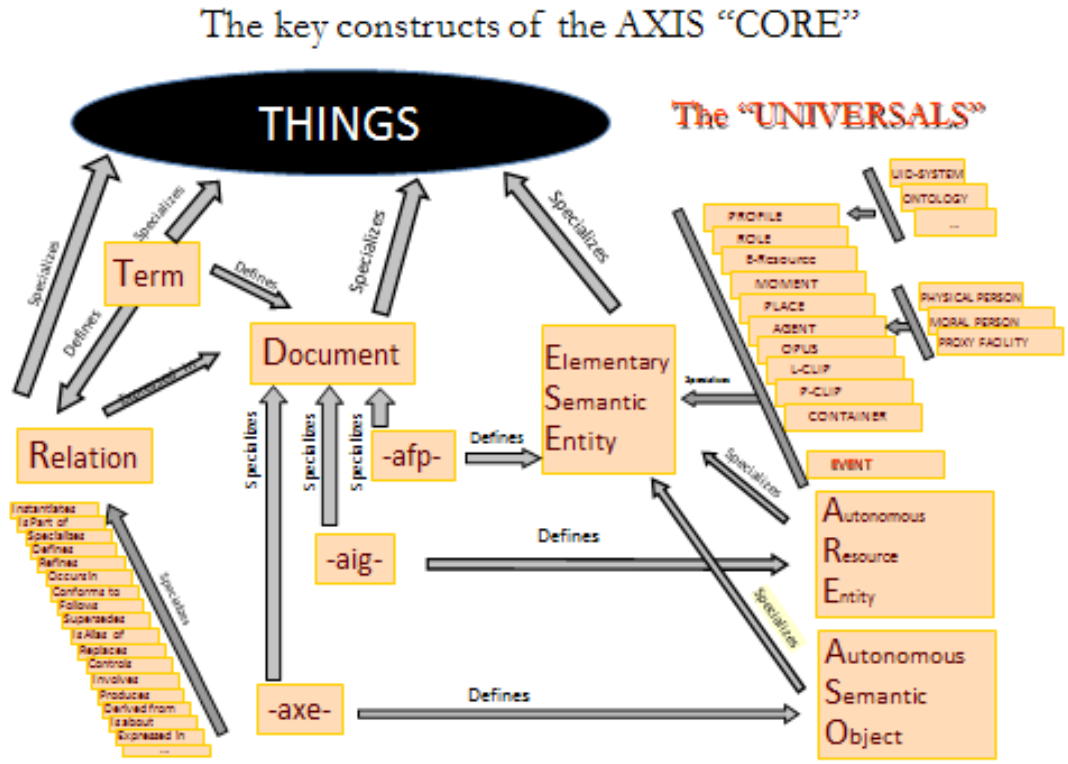

Figure 1: AXIS core

3 Titan : www.titan.be

AXIS (Acquisition, eXchange, Indexation and Structuration) organizes the modelling as networks of Autonomous Semantic Objects (ASO):

- in turn, each ASO is expressed as a network of Elementary Semantic Entities (ESE).
- the ASO wraps the instances and their models to becomes 'autonomous'.

The Entities are represented according to a uniform modelling technology which covers as well the assets, the work-flows, the presentations and the agreements.

The model constructing and validating the instances is called PROFILE. It includes six components covered in a multi-linguistic way:

1. The ONTOLOGY
2. The URN system
3. A set of specific 'TERMS' classes
4. A set of specific 'DOCUMENT' classes
5. A set of specific 'Lists of authority properties/values'
6. A set of specific 'GUI frames'

The static instances are constructed according to the applicable assembly of the Core profile with the domain specific profile: 'Music' (derived from the IAML ontology proposal).

The dynamic instances (work-flows) are constructed according to the applicable assembly of the Core profile with the domain specific profile: 'Finite-State-Machine".

The approach is totally fitting in the SNIA model of acquisition, operations and fixing of the knowledge. Furthermore, it resolves the famous opposition between 'Unstructured' ↔ 'Structured'. In the current SNIA model, in 'Structured' the data are represented in databases; while in 'Unstructured' the data are represented as documents conforming to external formats (standard or proprietary). When the formats of the documents and database-records are explicitly defined in semantics models, both representations can be used and converted to each other.

SCA (Semantic Content Architecture):

On the base of the results on the Eureka-Celtic MediaMap Project[4], the society Perfect Memory has produced some semantic applications modelled for AV Documents:

- the USE (Unique Semantic entity): a definition of the wrapper based on the ontology
- the OSB (Open Semantic Bus): the network in charge of the semantic middleware

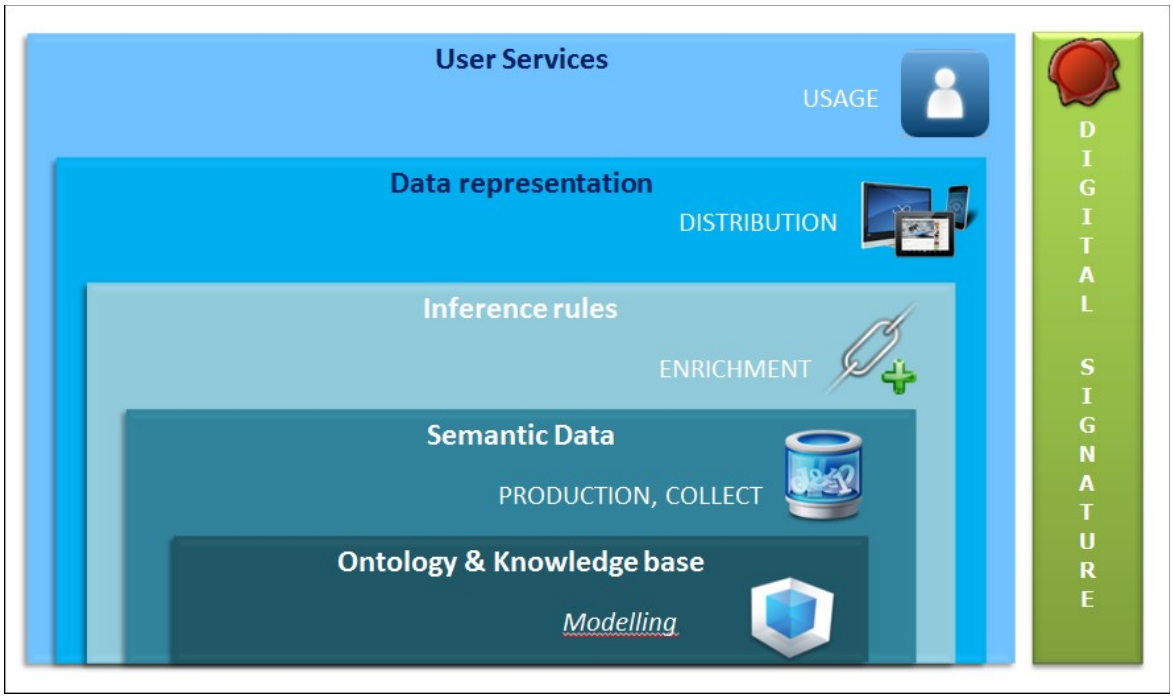

Figure 2: Semantic Content Architecture

This schema represents the interlocking of the Russian Puppet, from modelling to the User, with a full embedded digital signature:

- Ontology & Semantic database comprising a body made up of core concepts of an ontology and its initial value at the time of instantiation.
- Semantic data organized in triplet (subject, predicate, object) and supporting the OWL qualification.
- Inference rules can determine by deduction the new semantic triplets in the database, it can also optimize the collection and the capitalization of the information from third-parties
- Data representation: contextual information processing for editorial purpose and publishing (technical platform, protocol,).
- User Services: all resources made available to the end user to operate (read, comment, modify, transmit, transform) the data from the semantic database.

4 See : www.mediamapproject.org