# Activity analysis, semantics and metadata building

Christian Cote

▶ **To cite this version:**

Christian Cote. Activity analysis, semantics and metadata building. Journal of library metadata, 2016. hal-01772095

**HAL Id: hal-01772095**
**https://univ-lyon3.hal.science/hal-01772095**

Submitted on 20 Apr 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# ACTIVITY ANALYSIS, SEMANTICS AND METADATA BUILDING

*Christian Cote*
*MARGE- University of Lyon (France)*
*Chrsitian.cote@univ-lyon3.fr*

Abstract. We propose a methodology for building consistent metadata for the description of scientific working documents. Our methodology is supported by distributed cognitive theoretical principles and implemented through a combination of anthropological, semantic, and linguistic approaches. Our analysis is applied to a particular pharmacist activity, namely the adaptation of posology. Issues focus on the representation of situations that describe individuals, tools and artifacts. Regular relations between these situations characterize types that allow the description of conceptual contents associated to these empirical objects. Because these situations and types are expressed by a set of metadata and then associated to current metadata, our proposal extends the nature of entities described by metadata to useful internal activity artifacts.

Keywords: metadata, activity analysis, semantics

## Introduction.

We propose a new set of metadata that precisely describes some contents of documents which were created and used in the course professional activities. Big data and e-science entail the description of heterogeneous documents in research infrastructures[1]. If we extend the purpose to the frontiers of e-science, for example professional setting in which electronic health record in medical care is used, questions of data description, discovery and use become crucial. Metadata are a guarantee for data discovery and reuse in a big data context[2]. However, we need to specify the relevant set that precisely and universally describes these data.

We enhance the precision of the content representation through the characterization of universal data features that connects heterogeneous documents between them. The creation of a new metadata set for the description of specific sorts of documents is unnecessary[3]. Therefore, we do not offer an alternative to existing metadata sets but, rather, we propose some modules that will specify the intrinsic contents of some documents which cannot easily be described by usual metadata.

Any metadata set is based on a presupposed document function in a library or an infrastructure activity. This role entails the selection of some features: for example, Dublin Core[4] and the other metadata sets like MODS[a] are based on bibliographic description for librarian activity. By contrast,

[a] http://www.loc.gov/bibframe/pdf/marcld-report-11-21-2012.pdf. http://www.loc.gov/standards/mads/rdf/. http://www.loc.gov/standards/mdc/index.html.

DDI[b] and other research specialized metadata are related to science infrastructures and e-science procedures and furthermore, data record for research activity. However, these sets and models describe documents only by contextual features such as author, publisher, date, and so on. The contents are represented only by reference to controlled vocabularies, thesauri and classifications.

Our documents have their own internal classification system. Our approach presents a document characterization based on internal rules of symbolization and representation. It contrasts with a content representation based on external vocabularies[c]. Our perspective requires the following strong hypothesis: concepts that classify information in documents represent common knowledge. This analysis allows the use of the document internal classificatory concepts as descriptors. Furthermore, these descriptors can be connected to common controlled vocabularies which results in documents that can be more easily connected to standard models[d].

## Background and hypothesis on content and context characterization.

How can we categorize documents than may not be considered as publications nor archives? How can we characterize documents that are neither library objects nor only data? How can we describe these documents considering their function in context? How can we be sure that a metadata set has a great range and describes heterogeneous documents?

Completeness, validity, consistency, and accuracy of our metadata set are defined by prerequisites about the properties of the described documents, especially about the origin and range of their contained information, their method of information acquisition and representation, and of course their role in a work process. These prerequisites entail the definition of the document by the concept of artifact.

Descriptions at different levels of granularity entail difficulties for interoperability and then the emergence of new propositions for metadata building[5]. Some of them are justified by semantics[e], some other by ontology[6]. In the ontology framework, a lot of propositions have been elaborated to define and represent artifacts and software (as information artifacts)[7]. We enlarge the traditional metadata framework by the integration of concepts issued from ontology. However, ontology proposes only some useful concepts but no analytical tools.

Professional settings require documents that have content and structure regularities. Information Artifact Ontology (IAO)[8] proposes a document description focused on the acts produced by the artifact. IAO introduces a very important transformation by the integration of pragmatic and functional features in document description. For example, Barry Smith[f] distinguishes two dimensions,

---

[b] http://www.ddialliance.org/ http://www.loc.gov/standards/mets/

[c] The choice of an interpretative semantic theory is justified by the activity frame and then the realization on discourse of the linguistic entities.

[d] We don't develop this hypothesis but evolutions in information extraction allocate credibility to this idea.

[e] The expressive properties of RDF are an opportunity to apply semantic theories on metadata building. But metadata have a functional role and one use. Then semantics can't be considered alone for metadata conception.

[f] artifact=def. an entity created through some deliberate act or acts by one or more human beings and which endures through some information artifact: an artifact that can be the bearer of information:

(a) Information bearing entity (IBE) – a hard drive, a passport, a piece of paper with a drawing of a map

namely bearer and content, for any information artifact[9]. The bearer description emphasizes the characterization of the document function in context. However, the pragmatic dimension of the bearer can be more precisely expressed by considering meaning and reference. The content dimension is characterized by the notion of "aboutness". We enhance this dimension by the characterization of the individuals the document refers and, more precisely, by showing how the documents work on these individuals.[10]. By contrast with the B. Smith & al.[11], for us, meaning refers not only to objects, but also to the different steps of the information processing from individuals in the world to symbolic information.

## Methodology.

We propose an innovative methodology based on the complementarity between an analysis that describes these documents according to their context, their role, and their object and another that specifies the linguistic entities and discourse structures that realize these functions. The linguistic dimension has a special interest for the automatic extraction of metadata values and the creation of collections based on individual types, events, etc. [12]

The cognitive anthropology framework allows a characterization of the context (creation and use), constraints on meaning and document functional structure. It represents some relations between documents, actor reasoning, and material environment. We do not propose a general and procedural model of the activity[g] but a representation of information building, conveyance, and interpretation.

However, the conceptual framework of cognitive anthropology has not been designed for information conveyance and interpretation analysis: it cannot represent the linguistic content of the artifacts. Its range is limited to the cognitive activity associated to material objects and spatial configurations. On the other hand, semantic situation theory[13] preserves the distributed characterization of the cognition and explains the expressions and discourses meaning in context. Our study combines the two perspectives in order to promote a stable definition of the information artifacts, tools, and individuals in the world. Analysis issues will be externalized to form metadata.

Creating this set requires a complete setting composed of a small set of qualitative data. This setting is an exemplar type of any professional activity which implies documents. In this analysis, each proposal of document description refers to an identified phenomenon in the world and in documents. The adaptation of posology for antibiotherapy in hospital pharmacy presents a double advantage:

1. The whole process of the activity matching is explicit and
2. It requires documents that can be considered both for practice and research.

---

(b) Information content entity (ICE) – an entity which is about something and which can potentially exist in multiple (for example digital or printed) copies – a jpg file or a pdf file

[g] This sort of model is too abstract and general and cannot represent the daily arrangement and circumstantial variations.

# 1. Distributed and situated cognitive anthropology: information artifacts as foundations for activity.

As presented by A. Fjuk and S. Ludvigsen, "activity theory constitutes a rich framework for studying different forms of practices as developmental processes, with individual and social issues interlinked at the same time. In other words, activity theory affords analyses of social phenomena on different levels; activity – at the level of social systems; action – at the level of the individual learner acting intentionally; and operation – at the level of the concrete operation, procedure or behaviour"[14]. More recently, theories of situated and distributed cognition proposed a more precise characterization of the activity process. Considering that cognition is not only a human internal capability, they have proposed a characterization of objects, tools, and documents as externalized memories. These analyses allow a definition of documents as representational artifacts.

E. Hutchins[15] proposals have a specific interest. Following Hutchins, culture organizes the natural world in a functional way: natural world entities are cultural markers in the realization a particular operation while humans externalize complex reasoning by making tools and artifacts. For example, Hutchins explores how plane speeds may be memorized by cockpits. One essential principle of anthropology is that preexisting cultural solutions for a specific problem can be shared in other contexts in order to solve some similarly structured problems[h].

The adaptation of posology is a particular activity of hospital pharmacy that controls patient individual elimination -thereby drug toxicity risks- and calculate dosage on the basis of patient behavior. Adaptation can be described by pharmaceutical concepts such as drug elimination and therapeutic goal. This descriptive level is fairly efficient for characterizing pharmaceutic reasoning but the characterization of how and why the documents are relevant in the course of the action requires a description of this activity at a more abstract level. Activities indeed rely on core specific documents properties such as how a document allows working about an individual in the world, how it contributes to inference, or how it represents a complete situation in the world integrating some changes.

The adaptation of posology is a particular task in which the natural world is only accessible with samples wich are analyzed by tools and operation issues are reported into artifacts. Pharmacists have developed a large network of information both with a view of working on patients without any access to them, and of recording the individual information in memories and knowledge of the previous patients' behavior.

In this study, we do not address the different aspects of the cognitive activity during process matching but rather, in the material recording of information and knowledge. We limit activity analysis to information acquisition, conveyance, and duplication through different activity artifacts and tools.

The following figure represents the information trajectory from the real world to the workspace of the pharmacist: It characterizes the context of the artifact and the essential concepts we will use for our analysis and the metadata elements:

[h] The first application domain for this cognitive model was the Micronesian navigation and it has been extended to the "Western navigation" and airplanes cockpits.

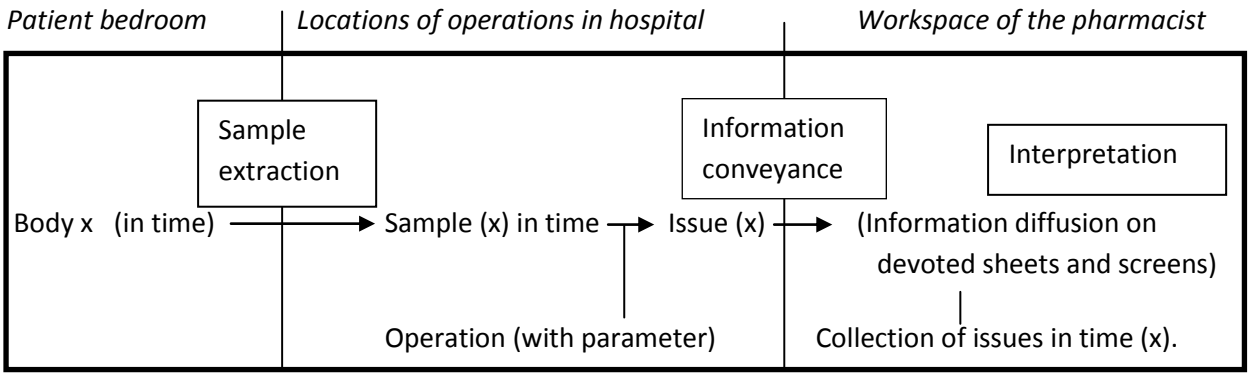



**Figure 1. Overview of the information building, conveyance and interpretation in the frame of the adaptation of posology.**

Now, we successively describe the different parts of this trajectory, considering individuals and universes, tools and artefacts.

## 1.1. Hypothesis of the three universes.

Natural language theories usually, consider that there exist three components for sign: things in the world, conceptual knowledge, and words[16]. Knowledge externalization (in tools, representational artifacts and protocols) changes our conception of meaning: the world is structured by culture. The three universes translate the three sign constituents into observable entities of the activity. These observable entities will be described by metadata.

**Entities of the natural world** designate every individual in the world before being integrated into a cultural framework. An individual[17] is a continuant[18][i]. Although individuals[j] may be culturally processed into objects of enquiry, they remain individuals. The natural world objects can be represented only by an indexical or a proper name. Symbolization entails a classification and a status: the individual becomes an object for information and knowledge investigations. We can work on an individual only if we use tools and representational artifacts. One paradox that may explain the development of information and information networks in pharmacy is that while pharmacists have many tools and processes both for symbolizing the patients' states thanks to various measures and analyses and for associating these states to populations recorded in databases, the whole processing remains unable to exactly predict patients' behavior.

Each individual can be represented by some specific features in the framework of the activity:

- Patient: agentive (e.g. resistance, elimination) and spatial features

[i] "Entities that continue or persist through time, such as objects, qualities, and functions" in J. Arp & B. Smith, p.1.

[j] An individual is defined by Strawson as follow: "the concept of a Strawson´s person is the concept of a type of entity, such that both predicates ascribing states of consciousness and predicates ascribing corporeal characteristics, a physical situation & co. are equally applicable to a single individual of that single type" (Strawson pp.101-102)

- Molecule: agentive and processes features (e.g. bacteria killing specified in time)
- Bacteria: infection, sensibility and resistance features in location and time.

Then, artifacts and any information about a fact are indexed by individuals. Any proper name is classified in the type `<individual>` and has some properties associated with it.

**Artifacts and tools** are material expansions or externalizations of human cognition. They are characterized by constraints, parameters, and issues and have representational properties. The constrained issues of tool operations are the integers that are conveyed to the representational artifacts but these integers cannot predict the quantity and nature of the observable facts: human reasoning on these representations boils down to inferences from these issues to the individuals. Contrary to natural language, artifacts integrate different dimensions of the world such as measurement analysis[19], analogic representations, and annotations.

Artifacts and tools have three dimensions:

- Functional description. Artifacts and tools are represented by their function that indicates the frame of interpretation of the reported issues.
- Composition and structure. This description characterizes what things mean in the artifact or the tool. Messages that convey issues have syntactic structures and rules of composition.
- Conveyance and classification. Any tool has an issue symbolically -realized or purely mental- that is reported to required representational artifacts. This knowledge is partly represented by type interpretation that characterizes the regular classification of anchored[k] facts into predefined files or boards.

**User world** defines the framework of the actor, considering both personal and community skills. The user world is composed of a social and collective knowledge about tasks and goals. The specific user knowledge is represented by its professional skills such as its ability to manipulate tools and resources in conformity to specific goals. This knowledge makes relations between the two previously presented universes: it connects what can be interpreted in the workspace about the individual in the world and what is expressed by the tools and artifacts. It can be segmented into proper human knowledge about the social activity organization and descriptive knowledge that concerns what we can learn about the individuals from the represented information[l].

## 1.2. Tools as externalized cognitive ability.

Tools are defined in cognitive anthropology by externalizing a part of a collective cognitive ability. A tool is represented by its operative abilities, together with its goals, its foundational knowledge, and its provisional issues. The operation is an implemented inference, that is, meaning production (or symbolization) is devoted to tools. If the definition of an artifact can be applied to any serialized

[k] Anchoring characterizes the function from a symbolic entity to a material situation in the world.
[l] This knowledge can be captured only by analyzing the interactions between users and the relation to technical and scientific publications.

support of information, then a tool is associated to any service that produces information from something in the world.

At descriptive level, we represent a tool through a simple operation which requires a parameter on some object in the world. The operation issue is a symbol which represents this object by considering the parameter. Hutchins' representation of a tool is slightly more precise than Ruttenberg's because it associates knowledge (via the parameter) to the operation of symbolization. The Hutchins representation addresses knowledge distribution in the workspace.

A simple externalized operation is schematized by Hutchins as follows in the navigation context:

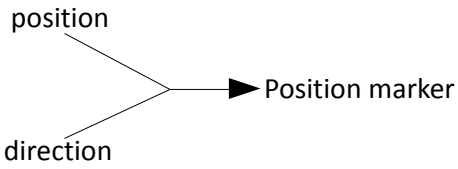



Figure 2.1. Operation in navigation

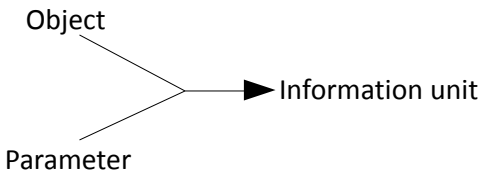



Figure 2.2. Generic representation

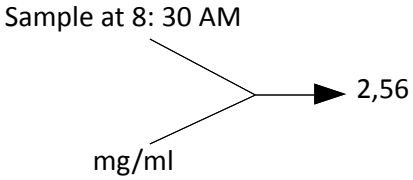



Figure 2.3. Operation in sample analysis

**Figure 2. Schematizations of the externalized operations.**

At a symbolic level of description, we identify an information structure that can be communicated and interpreted after it was represented in an artifact.

## 1.3. The notion of artifact.

The notion of artifact is essential in Information Science, ontology, and cognitive anthropology. In his definition of information artifact, Barry Smith[20] introduces a pragmatic dimension in reference to J. Searle[21]. But the speech act framework seems too limited regarding the role of the artifacts in the activity. The Hutchins proposition directly associates an information artifact to a cognitive structure. But a distinction can be made between information artifact (characterized by IAO) and a representational artifact: the rich notion of representational artifact integrates spatial and analogic

dimensions into the information bearer definition[m]. These representational properties allow interpretation and support inferences about the world.

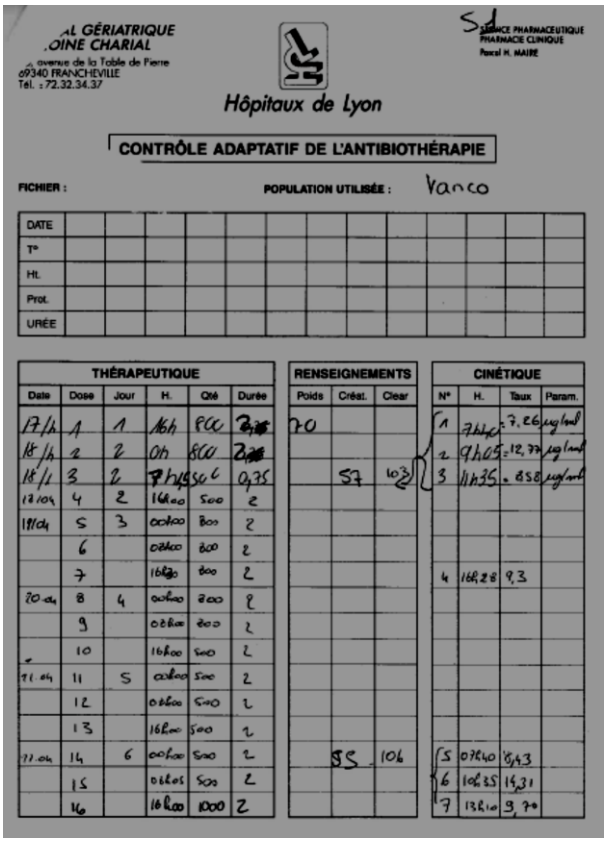
From a functional viewpoint, memories play an essential role in the coordination and cooperation in long-term therapy design. They concern what has happened during the therapy. The memorial and mediation role of the artifact in the activity introduces some precisions into the characterization of the whole artifact. The heading of the artifact is more a reference to the activity and to the processes that this activity is based on than a title that characterizes the identity of the document.

Each course of the activity depends on context and circumstances. Tools and artifacts present intangible regularities in the activity. Regarding memories, artifacts have the property to be continuant[n] (or endurant). If we now consider the whole activity, the artifact is the space that reports and memorizes the operation issues. Then it is used as a resource that presents a structured collection of information for any use. Information artifacts have both analogic and prescriptive dimensions. They record scientific knowledge about the activity and then regulate the individual interpretation of the described facts.

Examples of sheets from the pharmacy workspace are used for extracting some expressions from these artifacts. The associated attributes are presented below.

[m] An information bearer should not be confused with a document: the bearer characterizes only the material frame of a content representation. A sheet of paper can contain some different bearers. It's an artifact per se, but each information collection structured in the artifact by a spatial or visual bearer is an artifact.

[n] "An entity [bfo:Entity] that exists in full at any time in which it exists at all, persists through time while maintaining its identity and has no temporal parts." []
comment: Definition: An entity [bfo:Entity] that exists in full at any time in which it exists at all, persists through time while maintaining its identity and has no temporal parts.
http://www.berkeleybop.org/ontologies/obo-all/bfo/bfo.obo



AL GÉRIATRIQUE
OINE CHARIAL
avenue de la Table de Pierre
69340 FRANCHEVILLE
Tél. : 72.32.34.37

Hôpitaux de Lyon

SERVICE PHARMACEUTIQUE
PHARMACIE CLINIQUE
Pascal H. MAIRE

**CONTRÔLE ADAPTATIF DE L'ANTIBIOTHÉRAPIE**

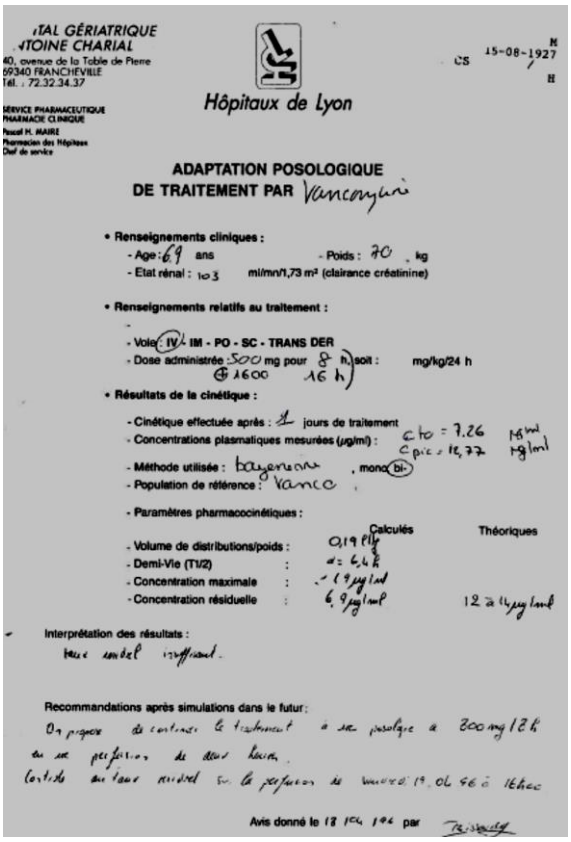
FICHIER :

POPULATION UTILISÉE : Vanco

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| DATE | | | | | | | | | | | | |
| T° | | | | | | | | | | | | |
| Ht. | | | | | | | | | | | | |
| Prot. | | | | | | | | | | | | |
| URÉE | | | | | | | | | | | | |

| THÉRAPEUTIQUE | | | | | | RENSEIGNEMENTS | | | CINÉTIQUE | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Date | Dose | Jour | H. | Qté | Durée | Poids | Créat. | Clear | N° | H. | Taux | Param. |
| 17/4 | 1 | 1 | 16h | 800 | 2 | 70 | | | 1 | 7h12 | 7,26 | µg/ml |
| 18/4 | 2 | 2 | 0h | 800 | 2 | | | | 2 | 9h05 | 12,77 | µg/ml |
| 18/4 | 3 | 2 | Phase C | | 0,75 | | 57 | 103 | 3 | 11h35 | 0,858 | µg/ml |
| 19/04 | 4 | 2 | 16h00 | 500 | 2 | | | | | | | |
| 19/04 | 5 | 3 | 00h00 | 800 | 2 | | | | | | | |
| | 6 | | 08h00 | 800 | 2 | | | | | | | |
| | 7 | | 16h30 | 800 | 2 | | | | 4 | 16h28 | 9,3 | |
| 20/04 | 8 | 4 | 00h00 | 800 | 2 | | | | | | | |
| | 9 | | 08h00 | 800 | 2 | | | | | | | |
| | 10 | | 16h00 | 500 | 2 | | | | | | | |
| 21/04 | 11 | 5 | 00h00 | 500 | 2 | | | | | | | |
| | 12 | | 08h00 | 500 | 2 | | | | | | | |
| | 13 | | 16h00 | 500 | 2 | | | | | | | |
| 22/04 | 14 | 6 | 00h00 | 500 | 2 | | 55 | 106 | 5 | 07h40 | 8,43 | |
| | 15 | | 08h05 | 500 | 2 | | | | 6 | 10h25 | 14,31 | |
| | 16 | | 16h00 | 1000 | 2 | | | | 7 | 13h10 | 9,70 | |

**Image 1. Initial adaptative control sheet.**



ITAL GÉRIATRIQUE
ANTOINE CHARIAL
40, avenue de la Table de Pierre
69340 FRANCHEVILLE
Tél. : 72.32.34.37

SERVICE PHARMACEUTIQUE
PHARMACIE CLINIQUE
Pascal H. MAIRE
Pharmacien des Hôpitaux
Chef de service

Hôpitaux de Lyon

CS 15-08-1927 M H

**ADAPTATION POSOLOGIQUE**
**DE TRAITEMENT PAR** Vancomycine

- **Renseignements cliniques :**
  - Age : 69 ans — Poids : 70 kg
  - Etat rénal : 103 ml/mn/1,73 m² (clairance créatinine)
- **Renseignements relatifs au traitement :**
  - Voie : (IV) - IM - PO - SC - TRANS DER
  - Dose administrée : 500 mg pour 8 h, soit : 1600 mg/kg/24 h (16 h)
- **Résultats de la cinétique :**
  - Cinétique effectuée après : 2 jours de traitement
  - Concentrations plasmatiques mesurées (µg/ml) : C to = 7,26 mg/ml ; C pic = 12,77 mg/ml
  - Méthode utilisée : bayesienne, mono (bi-)
  - Population de référence : Vanco
  - Paramètres pharmacocinétiques :

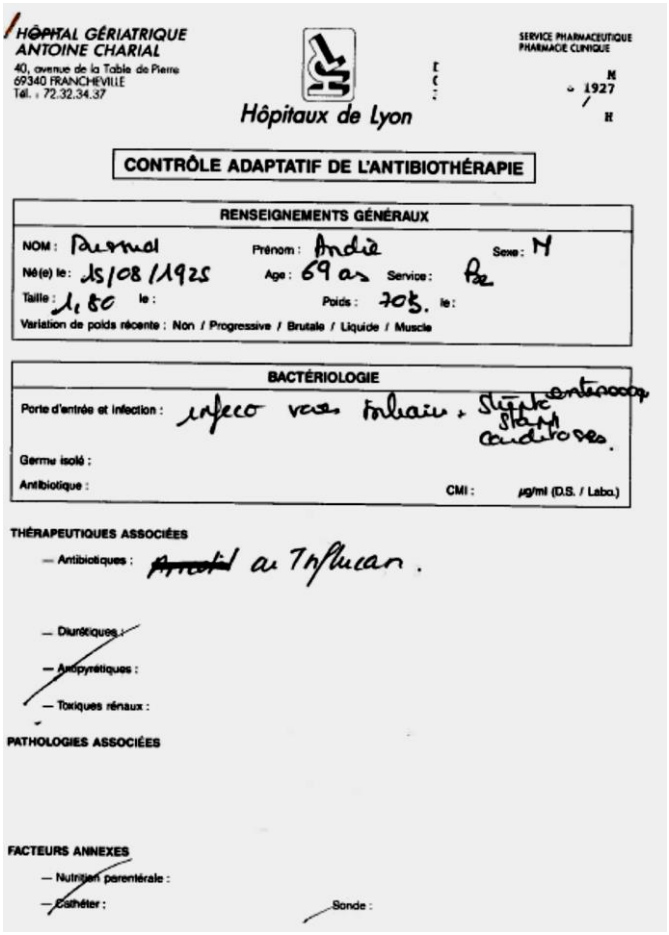
| | | Calculés | Théoriques |
|---|---|---|---|
| Volume de distributions/poids | : | 0,19 l/kg | |
| Demi-Vie (T1/2) | : | = 6,4 h | |
| Concentration maximale | : | 19 µg/ml | |
| Concentration résiduelle | : | 6,9 µg/ml | 12 à 14 µg/ml |

Interprétation des résultats :
taux résiduel insuffisant.

Recommandations après simulations dans le futur :
On propose de continuer le traitement à la posologie de 800 mg/12 h en perfusion de deux heures.
Contrôle du taux résiduel sur la perfusion de mercredi 19.04 de 16h à 16h30

Avis donné le 18/04/06 par

**Image 2. Adaptative control board.**



HÔPITAL GÉRIATRIQUE
ANTOINE CHARIAL
40, avenue de la Table de Pierre
69340 FRANCHEVILLE
Tél. : 72.32.34.37

Hôpitaux de Lyon

SERVICE PHARMACEUTIQUE
PHARMACIE CLINIQUE

CONTRÔLE ADAPTATIF DE L'ANTIBIOTHÉRAPIE

RENSEIGNEMENTS GÉNÉRAUX

NOM : Durmal Prénom : André Sexe : M
Né(e) le : 15/08/1925 Age : 69 ans Service : Pé
Taille : 1,80 le : Poids : 70 kg. le :
Variation de poids récente : Non / Progressive / Brutale / Liquide / Muscle

BACTÉRIOLOGIE

Porte d'entrée et infection : infect voies biliaires + Strepto entero Staph Candidoses.
Germe isolé :
Antibiotique : CMI : µg/ml (D.S. / Labo.)

THÉRAPEUTIQUES ASSOCIÉES

— Antibiotiques : ~~Amoxil~~ au Triflucan.
— Diurétiques :
— Antipyrétiques :
— Toxiques rénaux :

PATHOLOGIES ASSOCIÉES

FACTEURS ANNEXES

— Nutrition parentérale :
— Cathéter : Sonde :

Image 3. Posology adaptation recommandation.

| Identified expressions | Attributes |
|---|---|
| "Posology adjustment for dose regimen by…"<br>"Adaptive control for antibiotherapy"<br>"Adaptive control for antibiotherapy – file – reference population". | Artifact function |
| "clinical information"<br>"information about treatment"<br>"kinetic issue"<br>"General information"<br>"Bacteriology"<br>"Associated therapy"<br>"related pathologies"<br>"Therapy"<br>"Further information"<br>"kinetic" | Represented event |
| Name of the patient<br>"Vanco"<br>"Bile duct Infection + streptococcus" | Individual |
| "Kinetic after one day therapy"<br>"Measured plasmatic concentrations (µg/ml)"<br>Ct0 = 7,26 µg/ml<br>Cpic = 12,77 µg/ml » | Tool<br>Parameter<br>Issues |

Table 1. Categorization of the linguistic expressions extracted from the activity artifacts.

### Spatial and temporal dimensions of the artifacts: a functional bearer for classification.

Because of its material dimension, an artifact may be considered as an area, that is to say, a spatially distributed access to information in the workspace. Hutchins[22] demonstrates how some spatial natural structures like stars in the sky are used as temporal measures in navigation. This operation of transformation of a spatial dimension into a temporal measure is explained by the human cognition innate ability. This property is characterized in our artifacts by the articulation of temporal-token information into a spatial-type representation: information in time is incorporated into a spatial artifact[o]. Now, we can precise the different functions of the representational artifacts in the activity[p]:

**The first function is the memory of the achieved operations**, then an individual trajectory within the framework of shared knowledge. Artifacts represent the memory of the therapeutic actions on the individuals and then, by inference, what we have learned about individuals.

**The second function is a symbolic and partial representation of the individuals**. The issue of an operation is about a singular individual state and is published in a file or a pre-printed sheet that classifies this issue into a common framework. This artifact classification of the information into a typical frame is the graphic representation of a common view of the individual and of the way the activity works with it. Symbols are used for anchoring the artifact to the world and for referring to the information. Since pharmacy presents several types of individuals, the artifacts represent the interaction between a patient, a molecule and bacteria. This function allows us to reason about how we interact in the world.

**The third function is a classification of a succession of representations.** The artifact is a media of representation of a common and a priory scientific construction of the reality and is realized by the bearer structure. This function is fulfilled by an event[q] classification[23]. It associates predefined categories of accepted events or facts to the reported information about the individuals. This dimension correlates a regular form reasoning (about elimination or therapy) to specific individuals.

## 2. User situation and interpretation: semantic characterization of the artifact contents.

Now, we explore the discourse and activity articulation, especially the complementarity between formal semantic theories and distributed cognition principles. In many professional activities such as pharmacological analysis or research, we reason about individuals in the world only with data and without access to this individual. Then the organization of these data for the user interpretation is a condition for the success of the activity and the relevance of decisions. This characterization is founded on the interaction between information, the natural world and user goals.

---

[o] Everybody has an individual history but the same anatomy.

[p] These artifacts are supported by the claim that cognition has a dual representation: internal in the human mind and external in tools and mediation artifacts.

[q] We define an event as follows: “Although the term “event” has often a dynamic connotation, we use such term in the more general sense of *entity which occurs in time* (also called *perdurant* in the DOLCE ontology). In this understanding, states and processes are considered as special event kinds”. (Ferrario & Guarino, p.3

We start by presenting situation theory and the three foundational situations which characterize the interpretation of the identified objects. We proceed by presenting the types which are generated by the regular relations. Finally, we show how these types may be modelled.

## a. Situation theory.

Situation theory[24] proposes a formal representation of meaning in a distributed context. We first present the foundations of the proposed model, concerning what is meant by meaning and situation. Considering the user interpretative positions, meaning is defined from the user position in the course of his activity to situations in the world. Situation theory postulates that information interpretation is defined by relations between situations. This situation is structured under the principle of distribution and complementarity between heterogeneous[r] and symbolic entities. The situation structure is based on expressions but represents some primitives of human reasoning with information[25].

A situation accepts, or supports, information on the condition that a word having a conceptual principles or definite descriptions. This simple predicative structure is characterized as "infon"[s] and this "infon" is supported by a situation. Given a situation, *s*, and an infon $\sigma$, meaning has relations to strings of parametric values. These values are defined by direct reference:
$s \models \sigma$, and if the infon is *"John saw Mary running",* we write:

$$s \models see\ (j, s', t) \wedge s' \models run\ (m, t').$$

Any situation must be parametric to accept a collection of infons[t]. Relations between concepts are characterized by properties.

In situation theory, the interpretation is defined by the inference of true propositions from expressed information to situations:

- intensional interpretation reasons about type or generalization,
- extensional interpretation reasons about token or specification.

This interpretation implies that each individual, tool operation, and artifact has its proper type-token interpretation. This interpretation requires some prior knowledge about the symbolization operations, the information conveyance and the individuals. This prior knowledge is externalized in the representational artifacts by the bearer structure. This externalization is mediated between information world (i.e., individuals and tools) and the pharmacist's goals and skills.

---

[r] Heterogeneity is defined by constraints on meaning: information and situations associate terms that have different ways to mean and refer: conceptual meaning, direct reference, etc.

[s] "I adopt the word "infon" to denote an object of the form $\ll P, a_1,\ldots,a_n, i \gg$, where *P* is an *n*-place relation (for some *n*), $a_1,\ldots,a_n$, are objects appropriate for the respective argument place of *P*, and *i* is equal to 0 or 1. The notation and the name are intended to emphasize the fact that "infons" are *semantic* objects, not syntactic representation." (K. Devlin, op. cit., pp. 22-23).

[t] This position is opposed to the principle of ontology theories. According to these theories, meaning is based on the semiotic triangle that postulates a relation between a concept, a lexical entity and an object in the world.

### *Three foundational situations.*

Initially, J. Barwise & J. Perry distinguish three sorts of situations (op. cit., p.6) to characterize the meaning of an utterance. An utterance is defined as an action that conveys information and describes a statement. Utterance generally but not necessarily refers to a sentence. We define utterance as any description of the individuals, obtained by the operations issues. An individual occurs in a described situation by an indexical and a time markers. An "infon" is a predicative structure interpreted in a situation. Then, a situation is a propositional framework in which the information can be verified or not.

For one informational structure, we identify three situation types:

- **Described situation** characterizes what sort of fact of the world is described by the reference.
- **Utterance situation** characterizes what sort of representation choices has been selected for these facts. The utterance or expression framework characterizes the situation. It concerns the tool operations and issues.
- **Resource situation** characterizes what sort of theoretical framework supports these choices. It represents the common knowledge which is necessary for uttering and interpreting information. This situation is represented by the artifacts.

These situations explain why the representation provides a partial view of the world and why we can make some inferences about this world. These interpretations overlap each "infon" interpretation.

We now present the application of this conceptual framework on the informational dimension of the adaptation of posology. We characterize each situation by S and each situation supports a relation between heterogeneous entities. These entities which are characterized at a type level means that each entity classifies some tokens.

The **described situation** characterizes the fact that there exists an individual (associated to time and location markers) in the world:

Described situation: world individual
$S_0 \vDash \text{ind.}(t,l)$

The **utterance situation** indicates that an operation represents a part of the individuals with a required parameter:

Utterance situation: operations about this individual
$S_1 \vDash [\text{par.}(S_0 \vDash \text{ind.}(t,l)), (t,l)]$
$S_1 \Rightarrow S_0$
Par : operation parameter

The **resource situation** indicates that the parameters associated to the individual representation are issued from a conceptual structure[u]. This structure is externalized by the artifacts composition:

Resource situation: representational artifact
$S_2 \vDash$ th. $[(S_1 \vDash$ par. $(S_0 \vDash$ ind. $(t,l)t,l)), (t,l)]$
$(S_2 \Rightarrow S_1$ and $S_1 \Rightarrow S_0) \rightarrow (S_2 \Rightarrow S_0)$
th. : theoretical constraint (associated to the artifact structure)

These three situations are generic because they define the foundational entities and operations of the informational dimension.

Linguistic features cannot be considered alone. A word becomes a concept if it satisfies two criteria:

- The ability of this word to have an intensional meaning
- Its ability to convey and classify some specific information from a location in the activity to another.

Therefore, a predicate is characterized by the generic property of a word to convey and classify information in a distributed context. It classifies a succession of information in that it represents an event and in that each individual and operation issue is an argument. The characterization of the document properties by linguistic categories entails an artifact content description based on the classes of information accepted and shared by the artifact. When the issue of an operation is reported into a sheet or a file, it becomes a proposition since it can be interpreted at any time by everyone in the same way.

The interpretation takes place in the frame of the activity and then requires each identified situation. Interpretation in $S_2$ requires the interpretation of $S_1$ that requires $S_0$.

$$S_2 \Rightarrow (S_1 \Rightarrow S_0)$$

### b. Type relations between situations.

We now characterize the relations between the different situations. The regularity of these relations allows a type representation that surrounds one instance of activity matching. This type representation accounts for learning in the activity, knowledge accumulation, and discovery. Types correspond to materially attested relations in the activity and these relations stem from interpretations and their related knowledge.

Distributed cognitive anthropology characterizes how knowledge is manifested through material organization whereas situation theory represents how the uttered information is interpreted in this context. Distinctively to navigation (or other technical activities analyzed by Hutchins), adaptation is based on a knowledge acquisition about the portions of the world, especially individuals[v]. The individual behavior cannot be predicted and a knowledge acquisition about this individual is a

[u] This representation overlaps the distinction between an « artifact function » that designates an independent bearer (like a sheet or a file) and an « event representation » in which the bearer accepts a collection of issues from defined operations.

[v] Adaptation of posology is both a routine activity in hospitals and a research activity.

process that overlaps a succession of dosing proposals. A mental representation of the individual properties overlaps the synthesis of the operation issues. That is why the adaptation activity is characterized as a control expertise: dose regimens can be modified regarding to the patient behavior.

Situation theory defines situation types for characterizing abstract situations. We propose a type representation of the relations between individuals, tools and artifacts in conformity to the information conveyance. Knowledge is represented by a classification[26] of any token in a type and by the recognition of information by its classificatory type. Situation theory postulates that abstraction is a particular human quality that accounts for the intensional dimension of knowledge. Human internal knowledge is associated to a type characterization of any specific situation[w]. Following the principle of externalization, an artifact materializes some abstract knowledge. The report of the patient's last name on these artifacts anchors the knowledge associated to the support. Any type is realized at a token level by an "infon". The predicate has a conceptual meaning and bases the information conveyance from a situation to another in the activity frame[x].

The three foundational situations are populated by concepts when information is integrated into regular pre-printed sheets and electronic files. These regularities are represented by types. We now explore their specific properties both at the level of the information and the artifact model.

First we present the type mechanism. After the representation, we introduce the range of this type in the interpretation context. The denomination of the type interpretation follows.

### Individual interpretation and artifacts.

When we consider medicine and pharmacy, the real world includes patient, molecules and bacteria. These individuals are articulated to the different properties of the artifact: the spatial property for the patient body, the temporal property for the molecule processes (essentially elimination) and both spatial and temporal properties for the bacteria. The analog property of the artifacts anchors the space and time dimensions of the individuals[27].

We now represent how a situation in the world is represented by an artifact considering the classificatory rule:

$$T_0 \models S_0 \mid S_2$$

The type (T) is the relation that indicates the classification of the individuals in the artifact function. This classification has an important consequence: since each individual is classified into a regular artifact, it can be compared to any other and primarily to the individuals of the previous therapies.

[w] Following F. Dretske (op. cit. p.86), we adopt this definition of knowledge: "*K* knows that *s* is *F* = *K*'s belief that *s* is *F* is caused (or causally sustained) by the information that *s* is *F*. (…) Knowledge is identified with information-produced (or sustained) belief, but the information a person receives is relative to what he or she *already knows* about the possibilities at the source. Since there is a covert reference to knowledge on the right-hand side of the equation, the equation does *not* tell us, as it purports tell us, what knowledge is."

[x] This proposition is supported by the solidarity between artifact visual properties and predicative linguistic entities.

We categorize this type as <INDIVIDUAL CATEGORY>[y]. For example, this type is linguistically realized by the predicate “patient name”.

### Operation issue interpretation in artifacts.

The whole information about an operation is conveyed to the artifact and represented by an information structure[z]. An information structure is a grammatical composition integrating words referring distinctively: last names have a direct reference and by contrast, and event names have a mental significance. The classification requires an event name that means that this conveyed issue represents a state in the course of the event about these individuals. Then, an information structure classifies the whole components of the tool/operation situation in the artifact.

The conveyance of the issue and its classification into an event representation can be written as follows:

$$T_1 \vDash S_1 \mid S_2$$

This type classifies any operation issue into the event representation. Considering that operations select a specific dimension of the individuals through the parameter, we obtain a representation of the individuals constrained by an operation[aa]. This type can be categorized as <CONSTRAINT>. It is linguistically realized by the predicate “parameter”.

### Event interpretation in artifacts.

We now further define the type relation that holds between event representation and artifact function. Since symbolization is defined by a classification, discourse (i.e. information representation in artifacts) is characterized by a relation between event structures and activity goals. The theoretical constraints presented into the foundational situation ($S_3$) introduce a type/token distinction between the temporal-linear course of the events and the type-spatial structure of the artifact bearer. In this way, at type level, an event about the patient is spatially classified considering the artifact function in the course of the activity. The dimension of situations changes because the artifact is not only considered as a singular frame for interpreting one instance of information but also as a spatial and temporal representation of some goal-oriented externalized knowledge.

The interpretation about the whole individuals requires relations between the different events. These relations are called theories and represent inferences between events at an intensional level:

$$T_2 \vDash S_2 \mid S_2$$

Every event, obtained by the report of devoted operation issues, is classified into the artifact function denomination. An event has no pragmatic dimension. The classification of some pre-defined events into the artifact function allows the interpretation of an individual <BEHAVIOR> during the therapy. This type is linguistically realized by the predicate “event name”.

---

[y] This classification is emphasized by the database (used for estimation calculus) USC*PACK, which is available in the pharmacists’ workspace.

[z] An information structure is the linguistic and, more particularly, the grammatical representation of the “infon”.

[aa] The event defines concept realized by the operation issue and the constraints characterizes how this realization is obtained.

### Synthesis of interpretation requirement.

We now propose a representation of the different interpretation positions integrating the associated knowledge and structural dependencies. We distinguish a *partial interpretation* that concerns represented events about the individuals and a *global interpretation* that concerns the individuals and the activity management.

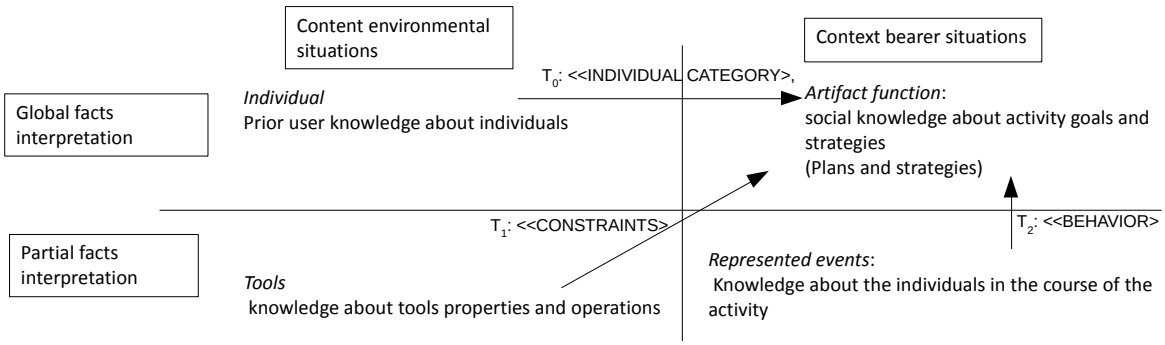



**Figure 3. Schematization of the knowledge distribution and relations in the activity.**

*Context and content situations* distinguish individual and tool situations from artifact situation. Event and social knowledge are externalized by artifact bearers while individual and tool knowledge are externalized in the world. They characterize the content of the reported information.

The activity process management is based on *social knowledge* which is distributed and shared among the different pharmacists. Social knowledge presupposes *knowledge about individuals* and how these individuals are represented (by tools and events). *Tool knowledge* is an *a priori* structure that constrains social interpretations as well as event interpretations. *Knowledge about the individual behaviors* during the therapy associates the different individuals and is both anchored on the event representation and the knowledge of individuals. The correlation between the different event proceedings is a basis for a *behavior* characterization that is interpreted in terms of action and activity management. It follows that social knowledge is supported both by individual behavior and by *individual category*: their complementarity is a condition for a sound decision.

## 3. Document description founded on activity information use.

Up to this section, we have considered activity and information independently from document description. Now we provide some metadata principles. Document description is a communication tool that facilitates resource retrieval considering the diversity of user interests: in e-science and big data context, a collection of resources can be of interest because it presents information about world entities, tools, events and the matching of an activity. Because we can represent the content and

context of the information both at the type and token levels, we can generalize the methodology to the description of any activity anchored artifact, tool and individual[bb].

As with the situation and type characterization, we distinguish two sets of metadata: content environmental metadata for situations, and bearer internal metadata for types. Before the characterization of each schema, we present the modular structure of the situations and types:

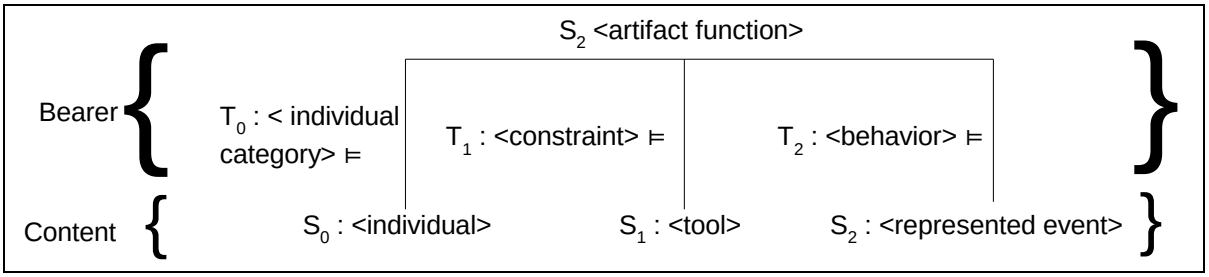



Figure 4. Structure of the metadata modules.

This representation is centered on artifacts. For a tool-centered description, $S_1$ is classified into the different artifacts of the activity. This remark can be extended to the represented events and the individuals.

We summarize the previous issues and their representation by metadata as follows:

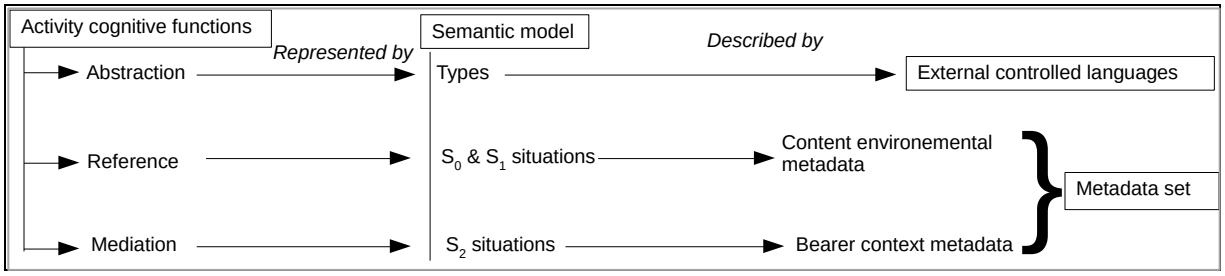



Figure 5. Articulation between cognitive dimensions, semantic model and metadata.

Reference characterizes situations describing something in the world ($S_0$ and $S_1$). Mediation characterizes $S_2$ structures. Abstraction is supported by types: it concerns knowledge associated to the relations between situations and complex interpretation.

We propose the following strategy to build metadata based on the following idea: representation languages[28] allow different descriptions which are based on their predicative structure and identical categories of symbolic entities[29]:

[bb] Metadata description is limited to entities identified as externalized knowledge. Knowledge architecture requires ontology.

- Every situation can be represented by an attribute-value pair, where each attribute is the predicate of the" infon". The token component of the situation is represented by a string of values, and each situation exemplifies one of these values.
- Every relation between components in a situation generates an RDFS triple. Then a situation is represented by a RDFS schema.
- Every regular situation builds a structure of triples and this schema describes the resource.
- Types are schemas based on relations between schemas.

Resources identify the entities of the activity we will describe. The predicates are linguistic entities defined as properties. These properties are empirically attested: they are extracted from artifacts. They associate the resource to individuals and operations components (characterized by non-predicative symbolic entities). These properties are associated to memories (of the different individual names, parameters and events). The extracted values have the same meaning in the pharmacist community and then, they are considered as a common professional vocabulary. In the case of types, predicates are informed by external controlled languages which require a manual indexer decision.

### *Content environmental metadata.*

Content environmental metadata modules characterize the artifact content. This content refers to external individuals and operations. Each module contains an informational predicative structure in which each situation is represented by a resource or subject, a predicate and a limited set of values or objects.

INDIVIDUAL: an individual represents a resource in the world identified by an entity name, a last name which may be categorized and a time marker.

| **Metadata module** | **Predicative representation** |
|---|---|
| Individual situation:<br>Individuals: entity name<br>Patient name:<br>Molecule name:<br>Bacteria name:<br>Time : | Subject: identified resource<br>Predicate: entity name<br>Object: time marker |
| **Exemple :** | Individual situation:<br>Individuals: entity name<br>Patient name: M.x<br>Molecule name: vanco<br>Bacteria name: streptocoq<br>Time : 18/06/1996 |

**Table 2. Individual metadata and predicative representation.**

TOOLS: the resource is a simple denoting operation requiring a parameter on something in the world. The predicate is represented by the parameter name and the issue by one entity in a string. The issue is in time indexed but this issue is not necessarily of interest for metadata.

| **Metadata module** | **Predicative representation** |
|---|---|

| | |
|---|---|
| Tool situation:<br>Operation name:<br>Parameter:<br>Time: | Subject: operation name<br>Predicate: parameter<br>Object: time marker |
| **Exemple** | Tool situation:<br>operation name : Measured Plasmatic Concentrations<br>Parameter :(µg/ml)<br>Time : T0 (12h24) |

Table 3. Tools metadata and predicative representation.

### *Bearer internal metadata.*

ARTIFACT: an artifact is a functional bearer that represents individuals and tools issues in the course of the activity. Each artifact contains different identified sequences of issues for a same operation and individuals: each sequence represents an event. The artifact function is marked by the title and the formal markers (spatial representations) associated to the artefact structure. Events are characterized by the symbolic entities associated to visual markers of solidarity between information expressions.

| **Metadata module** | **Predicative representation** |
|---|---|
| Artifact situation:<br>Artifact function:<br>Represented event:<br>Time:<br><br>Artifact situation:<br>Represented event:<br>operation name:<br>Time: | Subject: material artifact<br>Predicate: artifact name<br>Object: event name<br>time marker<br><br>Subject: material artifact<br>Predicate: event name<br>Object: operation name<br>time marker |
| **Exemples** | Artifact situation:<br>Artifact function: "Adaptation posologique de traitement par…"<br>Represented event: "renseignements cliniques"<br>"renseignements relatifs au traitement"<br>(...)<br>Time: 28/06/1996<br><br>Artifact situation:<br>Represented event: "Résultats de la cinétique…"<br>Operation name: "Concentrations plasmatiques mesurées"<br>Time: 28/06/1996 |

Table 4. Artifact metadata and predicative representation.

### *Types and external knowledge.*

These modules characterize the artifact internal structure. The situation-type characterization represents relations between different entities (i.e. tools, artifacts and individual) descriptions. It represents how the formal structure of the bearer articulates the represented entities and operations with conceptual structures.

$T_0$: INDIVIDUAL CATEGORY: the classification of each individual in an artifact characterizes the reference to any information contained in the artifact[cc]. Each artifact is a particular and partial representation of each individual. (Concepts are MESH[dd] descriptors).

| **Metadata module** | **Predicative representation** |
|---|---|
| Individual category:<br>Patient name:<br>Artifact name:<br>Category name: | Subject: patient name<br>Predicate: artifact name<br>Object: Category name |
| **Exemple** | Individual category :<br>Patient name: M. x<br>Artifact name: "Adaptation posologique de traitement par…"<br>Category name: patient âgé |

**Table 5. Individual category metadata and predicative representation.**

**$T_1$:** CONSTRAINT: each artifact function is associated to the report of tool operations. Every tool name is interpreted at a type level as the name of an event (measured by the tool). The constraint represents exactly how the event is represented and why this representation is both partial and in adequacy to the event.

| **Metadata module** | **Predicative representation** |
|---|---|
| Constraint:<br>Represented event:<br>Parameter:<br>Constraint name: | Subject: Represented event<br>Predicate: Parameter<br>Object: constraint |
| **Exemple** | Constraint:<br>Represented event: "résultat de la cinétique "<br>Parameter: (µg/ml)<br>Constraint name: pharmacocinétique-élimination rénale |

**Table 6. Constraint metadata and predicative representation.**

**$T_2$**: BEHAVIOR: Behavior represents the structure of the different events in the artifact. Each event is classified into the artifact structure and owns a meaning in relation to the other events. They figure out what the relations between events mean about the whole individuals. This reasoning is a condition for the action.

| **Metadata module** | **Predicative representation** |
|---|---|
| Behavior:<br>Artifact function:<br>Represented event:<br>Behavior name: | Subject: Represented event<br>Predicate: artifact name<br>Object: individual behavior |
| **Exemple** | Behavior:<br>Artifact function: "Adaptation posologique de traitement par…"<br>Represented event: "résultat de la cinétique"<br>Behavior name: défaillance rénale chronique |

[cc] It concerns the concepts that classify information and that are not names of operations ("infection name", "entry").

[dd] http://www.ncbi.nlm.nih.gov/mesh/

Table 7. Behavior metadata and predicative representation.

### *Relations between situations and types.*

Now we can present the relational properties that allow inference between situations and types. We use the range and domain properties to characterize what a situation or a type means by abstraction and anchoring. These properties are realized in a situation or among two situations (in the case of types). Range and domain properties represent classifications: this relation construction represents what a situation means in a conceptual domain and the impact of a concept in a situation.

This representation classifies facts into concepts and the realization of concepts in facts. More technically, situations are considered as containers and these containers have a domain characterized by a concept. Conversely, a concept has an individual as range (except `constraints` that have an `event name` as range).

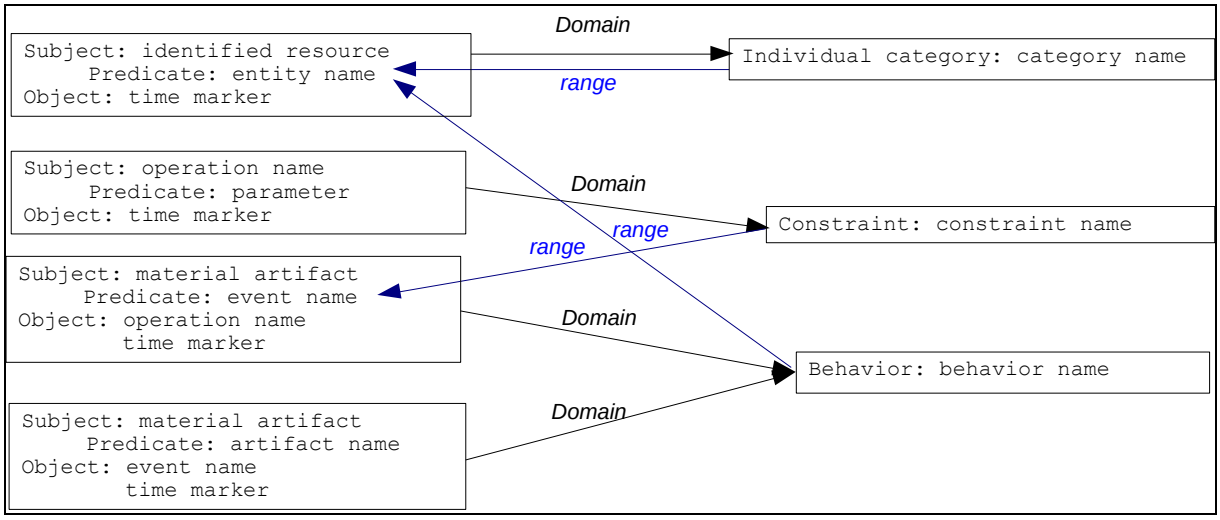



Figure 6. Schematic representation of the inferences between situations and types.

These relations represent reasoning between metadata modules. They present direct inferences from situations to types and then facilitate reasoning with metadata modules.

## Conclusion

With the increasing amount of metadata sets today, one may wonder whether it is really necessary to introduce a new one. But we can extend the descriptive properties of metadata to content and document context if we adopt a theoretical framework. This methodology appears as a response to the B. Smith's criticisms[30] about metadata foundation and a contribution to extension of the metadata coverage. Document description becomes more structured and new relations between external and internal artifact components can be specified in such a way that they present inferences from conceptual representations to anchored descriptions. This property is a great advantage for the modular description which allows an independent processing of each module. We have presented only one inference between modules but some other properties may be identified.

We have also developed a methodology based on the solidarity between linguistic phenomena and material processes. This methodology has entailed a coherent set of metadata based on both the

distribution of the described entities and information conveyance. Our perspective is more empirical and based on attested terms than usual ontological processing. Then the model can be extended to relations between artifacts, information exchanges and inferences in the course of the activity. This perspective allows the description of entities like tools and individuals that are not actually recorded by library metadata. We will extend this analysis to the activity description, considering the workflow of artifacts, communication and workplace. In this way, a whole activity will be described.

Our metadata set is designed to be complementary to usual sets like Dublin Core for example, or ontology likes IAO. We do not propose any other KOS but artifacts and tools externalized knowledge is connected to external resources (such as controlled vocabularies or terminologies). Our perspective allows primarily an intrinsic characterization of the used metadata, distinct from external controlled vocabularies, and further, the possibility to extract automatically the linguistic entities that will be used as values in the description.

Our metadata set is based on the adaptation of posology but we have preserved a generic representation in order to enable a large spectrum of applications.

---

[1] Rousidis, D., Garoufallou, E., Balatsoukas, P., & Sicilia, M. A. (2014). Metadata for Big Data: A preliminary investigation of metadata quality issues in research data repositories. *Information Services and Use*, *34*(3-4), 279-286.

[2] Qin, J., & Li, K. (2013, August). How portable are the metadata standards for scientific data? a proposal for a metadata infrastructure. In *Proceedings of the International Conference on Dublin Core and Metadata Applications* (pp. 25-34).

[3] Qin, J., Ball, A., & Greenberg, J. (2012). Functional and architectural requirements for metadata: supporting discovery and management of scientific data. In *Twelfth International Conference on Dublin Core and Metadata Applications* (pp. 62-71). University of Bath.

[4] http://dublincore.org/

[5] Peroni, S., & Shotton, D. (2012). FaBiO and CiTO: ontologies for describing bibliographic resources and citations. *Web Semantics: Science, Services and Agents on the World Wide Web*, *17*, 33-43.

[6] Peroni, S., Gangemi, A., & Vitali, F. (2011). Dealing with markup semantics. In *Proceedings of the 7th International Conference on Semantic Systems* (pp. 111-118). ACM.
Coppola, B., Gangemi, A., Gliozzo, A., Picca, D., & Presutti, V. (2009). Frame detection over the semantic web. In *The Semantic Web: Research and Applications* (pp. 126-142). Springer Berlin Heidelberg.

[7] Wang, X., Guarino, N., Guizzardi, G., Mylopoulos, J., (2014) Software as an Information Artifact, *Workshop on Information Artifact Ontologies, together with the 8th International Conference on Formal Ontology in Information Systems (FOIS 2014),* Rio de Janeiro, Brazil.

[8] https://code.google.com/p/information-artifact-ontology/

[9] Smith, B., (2013), The Information Artifact Ontology (tutorial), ***S**emantic **T**echnology for **I**ntelligence, **D**efense, and **S**ecurity 2013,* http://stids.c4i.gmu.edu/presentations/STIDS2013_Tutorial1_p2_Smith.pdf

[10] Borgo, S., Franssen, M., Garbacz, P., Kitamura, Y., Mizoguchi, R., & Vermaas, P. E. (2011). Technical Artifact: An Integrated Perspective. In *Formal Ontology Meet Industry(FOMI,)* pp. 3-15.
Kassel, G. (2010). A formal ontology of artefacts. *Applied Ontology*, *5*(3-4), 223-246.

[11] Smith, B., Malyuta, T., Rudnicki, R., Mandrick, W., Salmen, D., Morosoff, P. & Parent, K. (2013). IAO-Intel An Ontology of Information Artifacts in the Intelligence Domain, *Proceedings of the Eighth International Conference on Semantic Technologies for Intelligence, Defense, and Security*, Fairfax, VA, (STIDS 2013), CEUR, vol. 1097, 33-40.

[12] Stvilia, B., Hinnant, C. C., Wu, S., Worrall, A., Lee, D. J., Burnett, K., ... & Marty, P. F. (2015). Research project tasks, data, and perceptions of data quality in a condensed matter physics community. *Journal of the Association for Information Science and Technology*, *66*(2), 246-263.

[13] Barwise, J., Perry, J, (1983), *Situation and attitudes*, Cambridge, Mass. MIT Press

[14] Fjuk, A., & Ludvigsen, S. (2001). The complexity of distributed collaborative learning: Unit of analysis. In *First European Conference on Computer-Supported Collaborative Learning* (pp. 237-244). University of Maastricht.
[15] Hutchins, E. (1995). *Cognition in the Wild*. Cambridge: Massachusetts Institute of Technology Press.
[16] Ogden, C.K., Richards, I.A., Malinowski, B., Constable, J. & Crookshank, F.G. (2001). *The Meaning of Meaning: A Study of the Influence of Language Upon Thought and of the Science of Symbolism*. London: Routledge.
[17] Strawson, P.F. (1959) *Individuals*, London: Methuen & Co Ltd.
[18] Arp, R., & Smith, B. (2008). Function, role, and disposition in basic formal ontology. *Nature*, *2*, 1-4.
[19] Ruttenberg A., (2009) *Introduction to Ontology Introductions*, International Conference on Biomedical Ontology, Buffalo, NY. http://icbo.buffalo.edu/Presentations/Ruttenberg.pdf
[20] Smith, B. (2014). Document acts**. In Konzelmann Ziv,** A. & **Schmid,** H.B. (Eds.) *Institutions, Emotions, and Group Agents* (pp. 19-31), Berlin, Springer.
[21] Searle, J. R. (1995). *The Construction of Social Reality*. New York: Free Press.
[22] Hutchins, E. (2008). The role of cultural practices in the emergence of modern human intelligence. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *363*(1499), 2011-2019.
[23] Ferrario, R., & Guarino, N. (2009). Towards an Ontological Foundation for Services Science. In *Future Internet-FIS 2008: First Future Internet Symposium Vienna, Austria, September 28-30, 2008 Revised Selected Papers* (Vol. 5468, pp. 152-169). Berlin Heidelberg, Springer.
[24] Devlin, K. J. (1995). *Logic and information*. Cambridge University Press.
[25] Kratzer, Angelika, "Situations in Natural Language Semantics". In *The Stanford Encyclopedia of Philosophy* (Fall 2011 Edition), Edward N. Zalta (ed.), http://plato.stanford.edu/archives/fall2011/entries/situations-semantics/
[26] Dretske, F. (1981). *Information and the Flow of the Information*, Stanford, CSLI Publications.
[27] Barwise J. & Allwein G. (ed.), (1996) *Logical reasoning with diagrams*, Oxford, Oxford University Press.
[28] http://www.w3.org/TR/2014/REC-rdf-schema-20140225/
[29] http://www.w3.org/TR/2014/NOTE-rdf11-primer-20140225/
[30] http://ncor.buffalo.edu/2014/IAOW/IAO-Tutorial-Smith-Rio-Sep-2014.pptx