# Automatic indexing of multimedia documents as a starting point to annotation process

Sahbi Sidhom, Amos David

HAL Id: inria-00085559
https://inria.hal.science/inria-00085559

Submitted on 14 Jul 2006

# Automatic indexing of multimedia documents as a starting point to annotation process

**Sahbi SIDHOM**
*MCF  & Chercheur*
*de l' équipe SITE du LORIA*

**Amos DAVID**
*Professeur &  Responsable*
*de l' équipe SITE du LORIA*

LORIA - Université Nancy2, BP. 239, 54506 Vandoeuvre cedex - France
**Sahbi.Sidhom@loria.fr**                    **Amos.David@loria.fr**

## Abstract

Automatic text analysis widened the perspective of work on document contents by opening up the studies on the linguistic productions. In this case, we are using annotation as a case study.  In our approach, annotation is defined as textual, graphic or sound information, attached to document source (text, photo, audio sequence or video sequence : multimedia).  The source of our corpus is from INA databases (ie. Institut National de l'Audiovisuel, Paris).

Our research task consisted of identifying what are the appropriate characteristics of a multimedia document, its context and information retrieval in the context of natural language processing (NLP), automatic indexing and knowledge representation.

We discuss the crucial role of annotation process in the Knowledge Extraction tools and Management as well as in the design of Information Retrieval Systems. Our focus is more specifically on the new approach in information system design dedicated to "economic intelligence".

## *Keywords*

Annotation process, natural language processing (NLP),  knowledge management (KM), concept specification, classification, information retreival system (IRS), noun phrase (NP) model.

## 1.  Introduction

Automatic text analysis widened the perspective of work on document contents by opening up the studies on the linguistic productions.  Computational linguistics should reconsider the concept of interpretation under this new perspective and propose a new approach which includes all textual productions.  On existing information and that to added document, the development of automatic methods of analysis will be a major asset for the content development.

In this order of ideas, two prospects can be followed which orientate research in two complementary directions. The first order is application direction (engineering), which is in the area of the development of automatic tools, for knowledge filtering in a document.  The second order is intellectual direction, which is aim at classifying and extracting knowledge in a document (multimedia or not), with semantic information developed on textual productions like a summary, indexing or annotation contents.

With these two perspectives, the two directions are confronted with same problem: the description of discriminating features which will make the recognition of knowledge units possible.  By the latter, knowledge of the contribution of computational linguistics will be of textual nature to identify, then to formalize and to give an access to semantic interpretation in the document contents.

The section of automatic extraction of the knowledge [1] is interesting to actors of the knowledge engineering, document management or those in the economic intelligence (EI).  In the field of information retrieval, an EI actor must be able to identify indicators in the document and its contents. He must be able to filter textual objects (or multimedia objects) regarding information needs. Also, he must be able to use (software) tools to annotate his research results compared to requests executed in the information system.  All built objects:  indicator, index and annotation, are considered as added values (information) to analyzed primary information.  Thus, the analysis tools, indexing and annotation will be useful to supply relevant information in order to elaborate the adequacy of decisional strategies in a company.

In our approach, we give a critical look to the word "annotation". It implies two connotations. It is at the same time an ***object*** which implies the contents of annotation and a ***process*** which implies the activity of the annotator in the enhancement of informational contents by the insertion of new informational or interpretative elements (annotation objects).

Firstly in this study, the annotation is directed towards the interpretative action of a document. In this case, the annotator is the producer of the object and his activity is integrated in interpretative process on the document: more than a simple "intellectual" indexing or a synthetic reading of the document. Secondly, the annotation is an object (which can be written, audio, graphic or multimedia document) attached to the document source.

Our contribution is characterized by a differentiating ***annotation representation*** as an added value to the document content (source, bibliographic notice, request or informational resource for a decision-making) to the obtained information by a information retrieval system (RIS), which could interact with models dedicated to the EI [3], [11], in respect to relevant information requested in a decision making process. Under the presentation of these two problems (annotation and information retrieval), we will present in the following the importance of the annotation which includes semantic purposes for the information retrieval process (cf. 2) and its impact for the document indexing process (cf. 2.3).

## 2. Annotation process

### *Problematic:*

Annotation tools is becoming increasingly importance in the collection and information analysis steps. Mainly, in an information retrieval system, the validation of selected information (or relevant), in meeting the expressed needs, requires urgency in such tools. Thus, EI actors in a decision making process can be brought to perform collaborative interpretations in conformity with their decisional problem.

In this context, we found several definitions given to *annotation*. Most significant, for us, give the following representations:
- " *annotation is **graphic or textual information** attached to a document and generally placed in this document* ". [4];
- " *short **comment** or **explanation** on a document (or its content), in the same way a very short description usually added as a note after the bibliographical reference of the document* ". [6];
- " *any **object** (annotation) that is **associated with another** object (document) by some relationship* ". [18];

According to these definitions, annotation can be characterized by various dimensions relating to the object "annotation". These dimensions must give access to the object properties, such as: the structure, functions and role in the communication between the EI actors. Thus, in a communication context, an annotation is seen in the light of essentially three elements:
- annotator (EI actor) who carries out the annotation,
- source (document, bibliographic notice, annotation) concerned by the annotation,
- annotation object (text, graph, symbol, index, multimedia, …) introduced on the document.

In this study, we will not be concerned with the problematic of the user modeling in the annotation process, but essencially we will be concerned with the annotation contents (structure), its functions representative as value added on a document and of its implications to facilitate the relevant information retrieval in a decisional context (role to be played).

By implication, a document is a trace of the human activity, it is a consideration we will retained in the intellectual human effort to represent facts, knowledges and know-how [2]. From this point of view, the traces of the human activity can be materials in various forms like archaeological collections, parchments, manuscripts, written/audiovisual/cinematographic documents, etc. Nowadays, the digital supports and multimedia are important in management of knowledge to future generations. In this context, the question "*how knowledge can be managed in the annotation?*" will be the subject of the following consideration.

### *2.1. Annotation environment*

In the documentation domain, an annotation contains heterogeneous informations and it is conceived in the objective to be read and assimilate by anyone. From this point of view, the reading can be perceived differently from one person to the other. Thus the annotation associated to a document can take various interpretations and forms: textual, oral, graphic, filmic, etc.
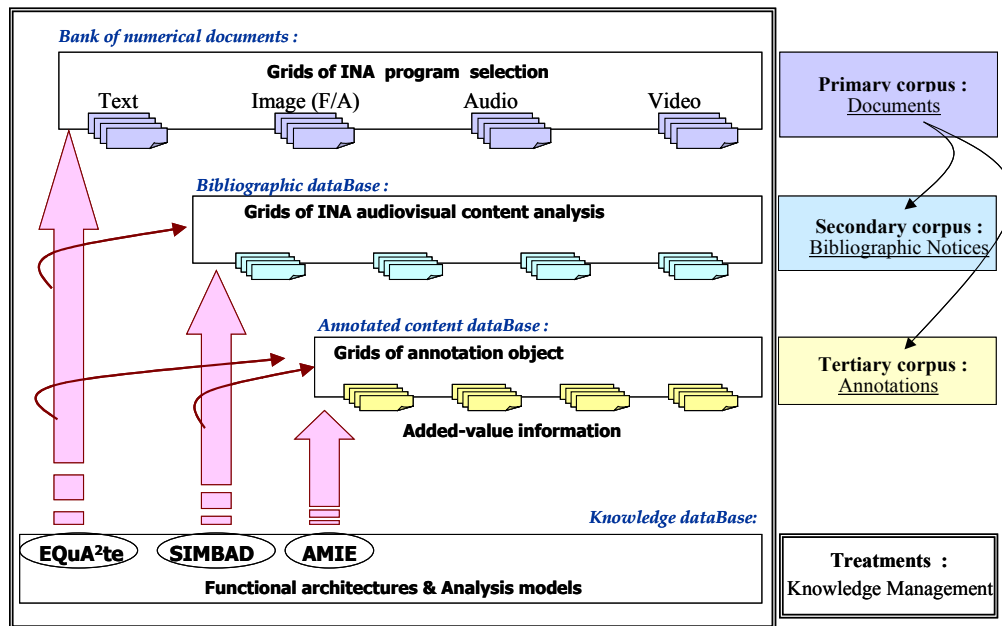
*Figure 2.1.: Position of annotation in information production.*

Independently of the document form, an annotation takes usually a complementary form compared to the document itself. This complement comes from the inherent concepts of the annotator that he wanted to introduce as added values. These elements (added values) will facilitate the reading of the annotation objects, will help to structure the concepts [34] for a relevant provision to information contents and will make human or automatic interpretation accessible (or shared).

Thus, each document can be associated to annotation objects, like terminological punctuations, words, images, sentences, hypertext links, typographical shapes, audio or video sequences, etc. These annotation objects can have features with homogeneous or heterogeneous objects of the source document.

In a historical context, the idea of Vannevar Bush in his communication "*As We May Think*" in 1945, is of an interest in collaborative work and it remains a current event. Annotation is built on the idea of a collaboratif work: the documents are arranged in agreement with a public views (information, knowledge and know-how), with the use of a common vocabulary, to express on close terms (or subjects) with shared concepts [9].

Closely related to "collaborative" work, the annotation could be to the author of the document, to the reader, to the interested public in the document and the annotations with an aim of sharing relevant information and to be use for the decision making. Thus, we consider that annotations linked to the document in the information retrieval process will remove ambiguities on the concepts (specify the attributes or values used in the content: explicit information), will inform us about information quality and will facilitate the analysis process: term indexing. Also in the annotation activity, there is content enrichment by implicit information (compared to the source concepts) and there is an accumulation of specific interpretations of the annotators (in consideration to specific field or expertise of the annotator).

Consequently, it is essential in the annotation concept to introduce semantic objects for the information indexing, filtering and the information retrieval. We refer some essential elements in the annotation objectives:
- to build an external representation to the document content,
- to introduce evaluation elements on the document [24]: account, contribution, report, demonstration, refutation, etc.,
- to allow for a focus on the content or on its form differently to the author presentation,
- to provide a traceability of document use [25],
- to accumulate explicit and implicit comments (information) on the contents,
- to support the reasoning and the evaluation on the contents,
- to share information,
- to filter information,
- to facilitate the comprehension as a second reading of a document,
- to insert semantic marking in annotation: symbolic, sign, index or alphabetical indicators, etc.

In followings, we will present some concepts in annotation processes for the information retrieval and to clarify some objectives.

## *2.2. Towards an annotation approach*

Some annotation tools of interest include *Annotea* [32], *Nestor* [22], *YAWAS* [31], *Commentor* [13], *CritLink* [21], and others [12], [14], etc. These tools are designed in the objective to promote collaborative works on the Internet by marking or anchoring annotations (color insertion, underlinings sentences/words..., hypertext links, etc.) on the source document.

In this study, we observe the annotation uses in the context of information retrieval process. Mainly, our objective is directed towards the annotation activity to determine relevant information sources.

Indeed, the annotation approach integrated in our orientation intends to describe functionalities that are proposed in the implementation of the " *AMIE* " model (ie. **A**nnotation **M**odel for **I**nformation **E**xchange) [23]. In this model, the major issue is carrying out the conjunction of the annotation parameters (attributes and values) compared to the information retrieval parameters (terms, index, keywords, words, descriptors, etc.). The annotation process will join in its results the process of indexing and information retrieval using the system **SIMBAD** (ie. "Multimedia Indexing System Based on the Document content Analysis") [16]. In the context of communication and information analysis, AMIE will clarify some parameters and functions, like the context of annotation, the annotator, the annotated document and some semantic functions (Fig.2.2).
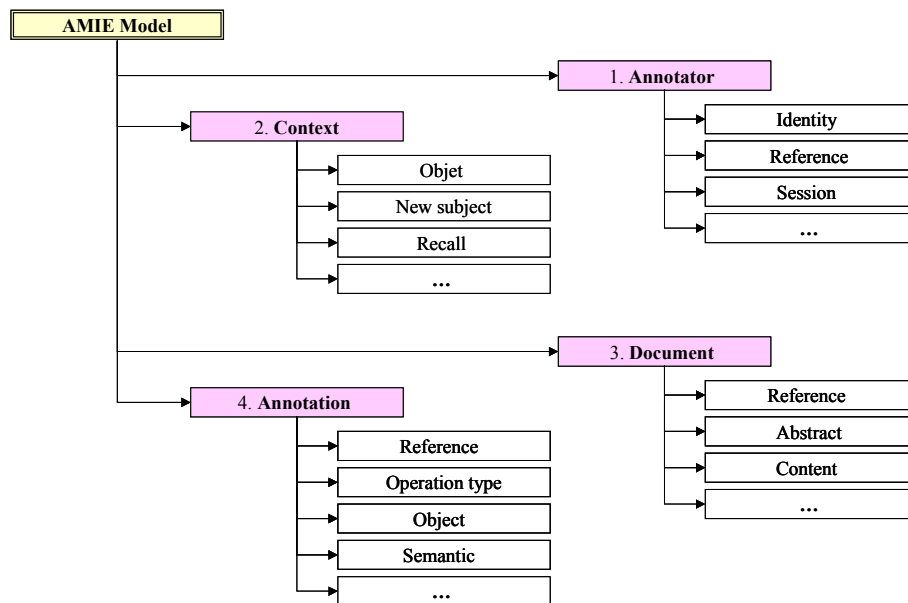


*Figure 2.2. : AMIE and its object classes: Context, Annotator, Document and Annotation.*

About the conceptual aspects of the annotation process, the methodology consists in matching parameters between the annotation process (AMIE) and the indexing process (SIMBAD). Impliedly, we can observe that annotation can integrate the indexing process:

1. In the case of an automatic indexing, there is no subjective interpretation because the indexing terms result from the explicit concepts in the content (document, abstract in bibliographic notice, annotation, request).

2. Whereas in intellectual indexing, often this activity (considered as a process) reaches a stability an interpretation form of the human indexer, in an implicit way, to define some grid analysis or indexing rules (constraints) in a consensus way with his community. Consequently, the out come of the indexing will denote information descriptions on the contents [26] and not the interpretations by the annotator.

Thus, annotation can be established as an interpretative process that increases by the contribution of indexing contents and the inquiry field of the annotators. These interpretative factors represent the reading context of the annotator (interest related to the document, subject, profile, etc.), the annotation context (dynamic content by new added information) and the indexing content (abstracts, concepts, themes, descriptors, terms, etc.)

For the two contexts of reading and annotation by the annotator, two proposals are made on semantic values:

- Annotation with implicit semantics: it is an annotation without reference to the annotation use. Annotator does not specify the objectives to be carried out with his annotation or the user classes (EI actors) for the annotation. In this logic, the annotations are additional information on the contents like punctuations, indicators, signs, added informations, symbols, etc. In this case, the attributes and values contained in the annotation are implicitly given.

- annotation with explicit semantics: it is an annotation with reference to the annotation use. Annotator can plan to define annotation objects (keywords, concepts, terms, indicators, themes, texts, etc.) and the possibility of determining user classes of the carried out annotations. Thus, it is explicit to determine attributes and values in the contents of annotation. It is an association between the content and the container, the annotation concerned by the document element and the concept related to the annotation object. In this case, the attributes and values contained in the annotation are explicitly given.

As noted in the annotation process, the automatic and/or intellectual analysis of annotations will be essential to clarify the existing semantic relations, their composition and their structure, in order to translate the informational elements into concepts and to represent the knowledges [28], [29]. The main problem consists of clarifying the attributes and values in the two annotative logics (implicit and explicit) to extend the functionalities of the information retrieval system: to query as well on the documents or/and on their annotations. On this subject that we will determine the annotation concepts.

## 2.3. Annotation in the context of information retrieval

The aim of this part is to integrate annotation structure to the information retrieval process with the semantic annotation functions and associated objects to the contents.

In this approach, we observed that problems relating to annotation in the information retieval are organized around three criteria:

- formalization: annotations are treated before the formalization of their structure either completely or partly. The structure will make it possible to identify the attributes and/or values in the formulation of requests and to solve the information retrieval problem: -- connect terms to a request using some attributes or values of the annotation, -- develop the relevance results in relation with the document terms and those of annotation;

- explanation: annotation does not suffice for itself. It is often made for one (EI actor) or more person. Therefore, it requires adaptations to the user profile with nonambiguous interpretations, like using common conventions, pivot language, mnemonic list, list of values, list of attributes, indicators, symbols, correspondence tables, etc.;

- translation: annotation in relation to its structure must integrate some properties on its role in the communication between the annotator (example: EI actor) and the prospector (example: decision maker). This last, in the context of economic intelligence, is a human agent (EI actor or decision maker), supplied with software agents/tools (software platform, data-processing tools, KM tools, etc.).

We assign to these criteria a class of annotation objects in order to re-use the annotation tools (implementation). Some implemented rules can be proposed or "standardized" [27] during the development tools. These will facilitate the annotator task and will preserve the annotation semantic.

Another annotation class, in this work, relates to the functions (access methods) to be allotted to the annotation process (cf. fig. 2.3). These functions are gathered in the *annotation manager*:

- annotation context: new annotation, follow-up an old annotation, creation a new annotation object (request, search for information, interpretation, ...);
- document annotated: specification relating to the annotate document, document classification (primary: document, secondary: bibliographic notice, tertiary: annotation);
- annotator: description of the annotator profile (explicitly: new annotator, implicitly: profile existing in the database);
- annotation: it is the principal function in this class which comprises the standard operations (reference, objective, operation/processing, semantic/meaning, …);
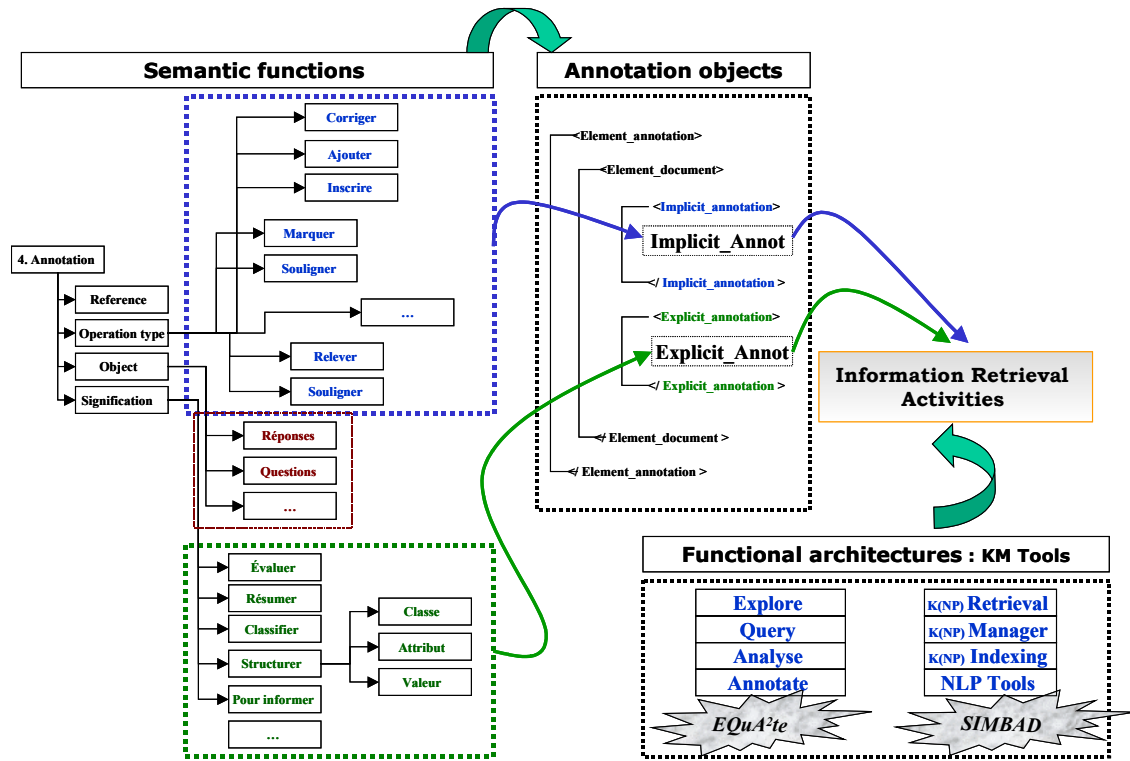
*Figure 2.3. : Annotation in information retrieval context: semantic functions.*

The annotation process, in relation with SIMBAD tools is to analyse, to index, to explore knowledge units (concepts, terms), will extract "Attributes" and "Values" in the annotation contents.

In the noun phrase (NP) model implemented in SIMBAD, *the minimal unit of speech which makes it possible to indicate an object* (idea, opinion, reference to the universe and its objects, …) *is an NP structure* [7],[8]. The reason why we implement NLP Tools (using CATN formalism [19], [20]) according to the syntax grammar of the speech (using INA corpus) to extract NP as indexing terms. Other tools are associated in this context: K(NP) Indexing, K(NP) Manager, K(NP) Retrieval (ie. K(NP) = **K**nowledge based on **NP** model).

By extension to the annotation process, a functional architecture which integrates the models *EQuA²te* (ie. **E**xplore, **Qu**ery, **A**nalyse, **A**nnota**te**) [33], SIMBAD [15] and AMIE will contribute in the near future to visualize the research subjects of the user and to the conceptual representation of information (*Figure 2.3.*).

The assignment of values to the attributes and their distinctions in the annotation will come within the annotator profile (competence) and his interpretative capacity (knowledge domain).  In the context of EI, the automatic tools and the annotator competence will jointly achieve the tasks according to the request of the decision maker, the information needs and the adequate translation of the decisional problem into an information retrieval problem [30]. All will be based on the two logical annotation proposals: implicit annotation (intension logic), explicit annotation (extension logic).

## 3. Conclusion

In this work, the annotation process associated to the information retrieval process consists in matching NP structures in the user request (NP-request) with those in the document databases (document sources, notices or annotations) respectively with NP-doc, NP-notice or NP-annot.  The relevant documents in response to the user request are those identified by NP-doc, less by those identified in NP-notice and more less by those identified in NP-annot. Other NP strategies were more developed in SIMBAD. Also, the user can change the order (strategies) on matching from databases.

In the NP model, the knowledge mapping establishes the relations based on NP concepts with the contents (document, notice, annotation).  As application to IR problem, the matching between requests and document sources is to operate on the NP relations and their nouns (N: head noun in NP). The NP had their natural organization. In a way, they had a fitting report: NP1 (NP2  (NP3  (…)  )), which makes possible to classify NP in distinct class levels. Also, the NP had an arborescence report: NPa (NPb (…), NPc (…)), which makes it possible to distinguish NP classes with the same level. With N, it is used to reach an NP Class or to navigate

between classes.

The topic on the "information retrieval system" is to provide a framework of problem analysis on the documents and the knowledge management: - information used as added-value in the decision making (annotation process), - parsing tools for textual contents in document, notice or annotation (NLP models and tools for indexing and knowledge representation).

However, the presented approaches and their interconnection offer a new vision in the EI process. It is a question of mapping by the contribution of each model in the information analysis on various resources: document sources, bibliographic notices and annotations. Also, to contribute to cross-checking of information sources and their annotations in order to clarify the knowledge problems: the knowledge detection and the contributions by a human activity (author, annotator, EI actor, decision maker, etc.). It acts to emerge information indicators and relations with attributes/values in their locations: which escapes from the information retrieval system is covered in the annotation process [5], [9], [10] like to explicit or to assign concepts by the annotator (*Figure 3.*). The added value information is highlighted in this new architecture of IRS.

Thus, several fundamental aspects are taken into account like the annotation, the indexing and the extraction of relevant information. The validation of results and their reliability are weighted with the end-user needs: the decision maker. The annotator has a capital role in this orientation and strategy.

Some questions in EI still require study efforts [17] and continuity in our proposals to bring more practical solutions which repercussions are considerable.
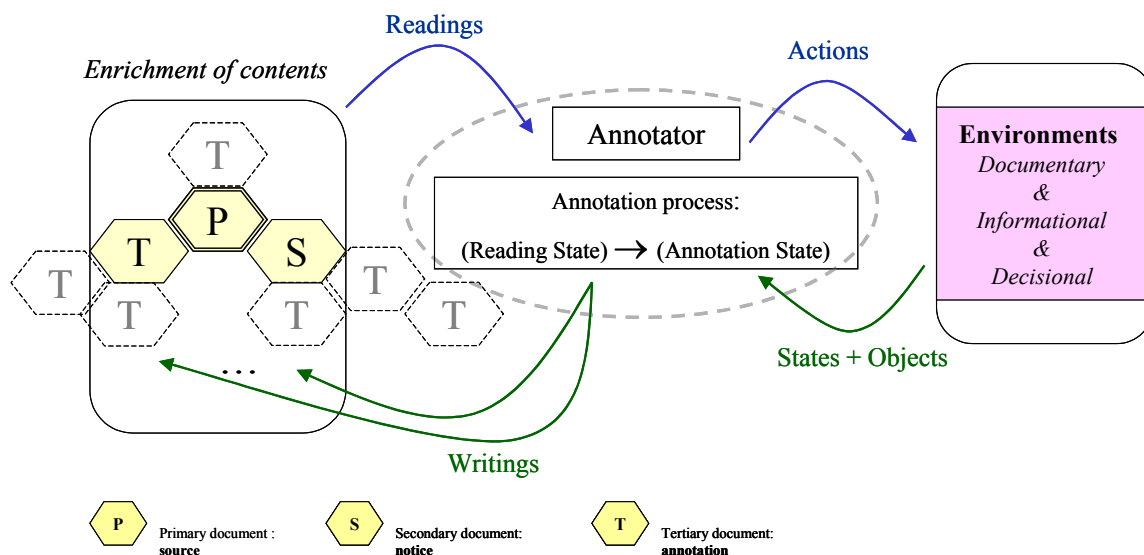


*Figure 3. : Annotation and informational environment.*

## 4. Bibliography

[1] : B. Bachimont. (2003). *Meaning and Indexing: which Issues for Multimedia Documents*. In Actes de la conférence « International Workshop on Content-Based Multimedia Indexing (CBMI'2003) », M. Gabbouj (ed.), Rennes. France.

[2] : B. Bachimont. (2004). *Why is there no knowledge engineering experience ?*, In Actes de la conférence « Ingénierie des connaissances (IC2004) », N. Matta (Ed.), Lyon. Presses Universitaires de Grenoble.

[3] : B. MARTINET. The economic intelligence. In Les Editions d'Organisation, 1995.

[4] : Desmontils, E, Jacquin, C, and Simon, L (2004). *Dinosys: An Annotation Tool for Web-based Learning*. In: The third International Conference on Web-based Learning (ICWL'2004). Springer-Verlag.

[5] : Esnault, L., Ponti, M., Zeiliger, R., (2005), *Constructing Knowledge as a System of Relations*, In IRMA International Conference, San Diego, May 15-18, 2005, USA.

[6] : *The Great Terminologic Dictionary* (2006). In : Le grand dictionnaire terminologique (1983-2006) - Domaine : science de l'information sur l'acquisition et traitement des documents. (URL visited 06.Jan.2006): http://www.granddictionnaire.com/ .

[7] : Le Guern M. (1989). *On terlinology and lexical relationship*. In symposium: les terminologies spécialisés - Approches quantitatives et logico-sémantique, et Meta Vol.34, No.3., sept. 89.

[8] : Le Guern M. (1991). *A morphosyntactic parser to automatic indexing*. In Revue : linguistique française : Le Français moderne . n°1, juin 1991.

[9] : Lilavati Pereira Okada, A., Zeiliger, R., (2003). *The Building of Knowledge through Virtual Maps in Collaborative Learning Environments.* in proceedings of the ED-MEDIA 2003 conference, AACE, Hawaii, USA.

[10] : Marshall, C. C. (1998). *Toward an ecology of hypertext annotation*. In ACM Hypertext, pp. 40–49. ACM Press.

[11] : MARTRE, Henri. (1994). Economic intelligence and industrial. In Report of « Commissariat Général au Plan », La Documentation Française, Paris, 1994.

[12] : Matthew A. Schickler, Murray S. Mazer and Charles Brooks. (1996). *Pan-Browser Support for Annotations and Other Meta-Information on the World Wide Web*. in Fifth International World Wide Web Conference, 6-10 May, 1996 , Paris (France). (URL visited 06.Jan.2006): http://www5conf.inria.fr/fich_html/papers/P15/Overview.html.

[13] : OVSIANNIKOV I., ARBIB M.A. and McNEILL T.H. (1999). *Annotation Technology*. In Int. J. Human-Computer Studies, 1999, pp. 329 – 362. (URL visited 06.Jan.2006): http://portal.acm.org/ft_gateway.cfm?id=989877&type=pdf .

[14] : Rachel M. Heck, Sarah M. Luebke, Chad H. Obermark. (1999). *A Survey of Web Annotation Systems*. Work supported by Grinnell College Noyce Science Summer Research Fund, (URL visited 06.Jan.2006). http://www.math.grin.edu/~rebelsky/Blazers/Annotations/Summer1999/Papers/ .

[15] : SIDHOM Sahbi, HASSOUN Mohamed. (2003). *Morpho-sytactic Parsing for a Text Mining Environment*. In Official Journal « Knowledge Organization » KO.29-2002, No. 3-4, Edited by Olson, Hope A. – Saranchuk, Georgina R. Zaharia, (c) 2003 Ergon Verlag.

[16] : SIDHOM, Sahbi. (2002). Morphosytactic parsing plateform for automatic indexing and information retrieval : from written document towords knowledge management ", Doctorat thesis : University Claude Bernard Lyon1, France.

[17] : Thiery, Odile et David, Amos. (2003). *EQuA$^2$te architecture and its application to economic intelligence*. In Conference "Intelligence Economique : Recherches et Applications" - IERA'2003. INIST, Nancy-France.

[18] : W3C Collaboration Working Group: Annotation (2004). Collaboration, Knowledge Representation and Automatability. (URL visited 06.Jan.2006): http://www.w3.org/Collaboration/Overview.html#annotation .

[19] : William A. Woods. (1980). *Cascaded ATN Grammars*. in American Journal of Computational Linguistics, January-March 1980, vol.6, n°1.

[20] : William A. Woods. (1997). *Conceptual Indexing : a better way to organize knowledge*. Technical Report SMLI TR-97-61 : SUN Micosystems, Lab. Mountain View Canada, April 1997.

[21] : YEE Ka-Ping. (2002). CritLink: Advanced Hyperlinks Enable Public Annotation on the Web, Demo to the CSCW 2002 conference, New Orleans, Dec 2002, (URL visited 06.Jan.2006): http://www.zesty.ca/pubs .

[22] : Zeiliger, R. (2001). *Nestor : The web browser and cartographer* (update April 11, 2005). (URL visited 06.Jan.2006): http://www.gate.cnrs.fr/~zeiliger/nestor/nestor.htm . CNRS-Lyon.

[23] : Robert Charles, David Amos. (2006). *Parametric view of annotation for decision making*. In : 2nd International conference on computer science and information systems. Athens-Greece.

[24] : Kelly, D. & Teevan, J. (2006). Evaluation of personal information management tools. in W. Jones & J. Teevan (Eds.), Personal Information Management. Seattle: University of Washington Press.

[25] : Kelly, D. (2005). Implicit feedback: Using behavior to infer relevance. In A. Spink and C. Cole (Eds.) New Directions in Cognitive Information Retrieval. Springer Publishing: Netherlands.

[26] : Kelly, D., Diaz, F., Belkin, N. J., & Allan, J. (2004). *A user-centered approach to evaluating topic models*. In Proceedings of the 26th European Conference on Information Retrieval (ECIR '04), Sunderland, UK, 27-41.

[27] : Croft, W.B. & Lafferty, J., (2003). Language Modeling for Information Retrieval, Eds. Kluwer Academic Publishers 2003.

[28] : Metzler, D. and Croft, W.B., (2004). Combining the Language Model and Inference Network Approaches to Retrieval. In Information Processing and Management Special Issue on Bayesian Networks and Information Retrieval, 40(5), 735-750, 2004.

[29] : Croft, W. B., Lavrendo, V., & Cronen-Townsend, S. (2001). Relevance feedback and personalization: A language modeling perspective. In Proceedings of the Joint DELOS-NSF Workshop on Personalization and Recommender Systems in Digital Libraries, 49-54.

[30] : Taylor, R. S. (1986). Value-added processes in information systems. Norwood, NJ: Ablex Pub. Corp.

[31] : Laurent Denoue. (1999). *Yawas : Annotation tool for web navogators*. In IHM'99, Montpellier, France, 22-26 Nov. 99.

[32] : Marja-Riitta Koivunen. (2005). Annotea and Semantic Web Supported Collaboration. In ESWC 2005 (2nd European Semantic Web Conference). Haraklion, Greece, 29.may-01.june 2005. (URL visited 06.Jan.2006): http://www.annotea.org/eswc2005/01_koivunen_final.pdf.

[33] : David Amos, Sidhom Sahbi. (2005). *Integration of the Economic Intelligence approach in the functional architecture of an information system*. In Conference: Le Système National d'Information Economique : Etat et perspectives (2005). CERIST, Alger-Algerie. (URL visited 06.Jan.2006): http://hal.inria.fr/docs/00/03/64/79/PDF/revue_e-TI_2005.pdf.

[34] : Sidhom Sahbi, Robert Charles, David Amos. (2005). *From primary information to added value information in the digital processing*. In. International conference: L'information numérique et les enjeux de la société de l'information. Sep. 2005, Tunis-Tunisia. (URL visited 06.Jan.2006): http://hal.inria.fr/inria-00000254.