Video-book:
a collaborative polemic-based video annotation platform

Vincent Puig
Institute for Research and Innovation (IRI)
4 rue Aubry le Boucher
75004 Paris, France
vincent.puig@
centrepompidou.fr

Nicolas Sauret
IRI
+33 6 .. … …

nicolas@
inflammableproductions.com

Yves-Marie l'Hour
IRI
+33 6 60 19 00 95

yvesmarielhour@gmail.com

Alexandre Monnin
IRI

monnin.alexandre@neuf.fr

ABSTRACT

A Video-Book is a video recording made discreet, i.e. it is readable like a folio book. The folio is not a fixed format, as is a book. This is so because a video recording is a temporal object, while a book is a spatial object. A video-book does not consist of videographic pages: we will talk of video segments, of varying dimensions, and localised by two references to timecode. These segments also include a title and keywords.

Discretization of a recording in video segments allows to understand synoptically in a few seconds the content of a record. In other words, the discretization opens the possibility of a fast access to a recorded document that makes it usable for cognitive activities.

Furthermore, the discretization allows to perform queries on the content of the recording: cut in video segments, the recording is integrated into a video database containing others similar recordings edited in video-books.

This presented prototype aims at enabling an innovative online video form, allowing participant of in a debate to record themselves according to a pre-segmentation of the topic including a pre-indexation allowing to automatically assign keywords to the produced content. Finally, keywords and polemical marks can be manually attached to the segments and produce dynamic maps and collective spaces in which topics and their metadata can be discussed within what we call "semantic storms".

Categories and Subject Descriptors

H.5.1 [**Information Interfaces and presentation**]: Multimedia Information Systems – *Video*

General Terms

Design, Experimentation, Human Factors, Languages.

Keywords

Metadata, Collaborative Annotation, Video Annotation, Tagging, Rich Media, Semantic Web.

RESEARCH CONTEXT AND PREVIOUS WORKS AT IRI

1.1. Research Context

Since its creation at Paris Centre Pompidou in 2006 by the philosopher Bernard Stiegler, Institute for resesarch and innovation (IRI) has been exploring cultural and cognitive technologies, aiming to create new information systems that address users' needs within the cultural domain and elaborate the required cultural technologies. IRI develops digital technologies intended for amateurs, researchers, and artists.

In its research, IRI wants to stay aside from the dominant conception of annotation which tends call everything annotation, be it produced by a machine or by a human being. Here, we would like to distinguish the process of indexing (or the engineering of knowledge, which also covers the definition of "ontologies") and the process of annotation (or the engineering of information and the human production of metadata which may be or not assisted by a machine). This research is empirical as it is rooted in the analysis of identified cultural practices and notably the operative strings which compose the act of annotation which we try to instrumentalize (in the sense of a general organology as defined by Bernard Stiegler) to go beyond them. Research on the tools for indexing is essential in the field of instruments for critics, as it follows closely the act of annotation. Thus IRI studies, imagines and develops new kind of tools for annotation, based on the combination of information and metadata architectures, with hypermedia navigation interfaces, with algorithmic modules for the detection of signals and modules for the visualization of data (cartography). The fruits of these research is regularly used to update "Lignes de Temps", IRI's software for the annotation of temporal objects (music, recordings etc.). The research in this field is progressively turned toward the dynamic annotation of oral and written language, et the annotation of pictures.

IRI leads series of research and experimentations to explore the concept of "Signed and collaborative readings" which associate annotating techniques from the books and paper-based media which do not exist yet on the web, and new paradigms for collaborative work enabled by high-speed networks. An important aspect of that work is the conception of technologies that follow, update and administer the exchanges, debates and polemics, which come from the collaboration of annotators.

1.2. Lignes de Temps/Semantic Compass/ Polemical Links

The software Lignes de Temps[1], developed in 2007 in the context of the ANR Cinelab research project, opens the possibilities of analysis and of synthesis offered by digital technology. Inspired by the usual "timelines" for digital editing, Lignes de Temps gives a graphical representation of a film, revealing immediately its structure. "Lignes de temps" offers a new access to the film, because it substitutes to the projection of the film, the cartography of a temporal object.

The software offers the possibility to diversify the approaches of a film by displaying several parallel timelines, and therefore to visualize by comparison and combination of criteria of relevancy, meaningful effects, to update for example recurrencies and symmetries.

On top of their first criteria of segmentation, users of the software can add their own subjective lines based not on the film itself, but on their personal interpretation. The tool offers the possibility to define segments along time lines, and then to attach free text annotation, key words from a shared data base, or using thesauri
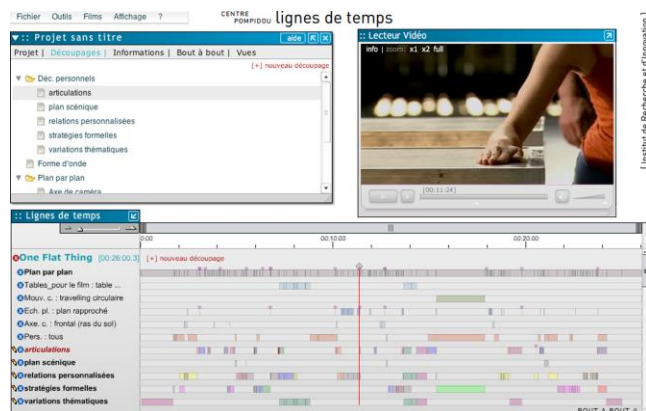

Figure 1: "One flat thing, reproduced", a choregraphy by William Forsythe, analyzed in Lignes de temps by Thierry de Mey

(controlled vocabularies) or thematic graphs.

The communautisation is the third stake of the development of the tool, because it allows the author to share his times lines with others. In this case, the time line appears to be an organized tool of synchronized dialogue to the film. Indeed, the time lines of the different contributors can be exchanged, superimposed or modified by one another.

Last year, we intended to go beyond current Semantic engineering by developing tagging features, interfaces, relational databases for innovative representation and interaction with opinions. For instance any produced annotation on a given piece of archive may be qualified: in opposition, adhesion, as a reference or as a question. This was the purpose of the Semantic Compass project that is the design of a dynamic map interface adapted to the navigation in temporal objects and presenting alternative viewpoints to the user while he is listening or viewing. The interface not only allows navigation using keywords (tags) but also the reading of side issues (i.e. not necessarily tackled in the video recording but indexed as such after : approval, contradiction, reinterpretation, change of context, enrichment, available references, questions ...). This project was carried out in

Figure 2: Lignes de Temps software interface with the Semantic Compass module

cooperation with Antoine Boilevin, designer at ENSCI.

VIDEO-BOOK

Video Corpus

The concept of video-book was motivated with the intention to constitute deep-indexed corpus of video interviews accessible online by a community of experts or public circles. Indexing and annotating deeply temporal objects enable effortless access and manipulation of the content and thus expose it to critics and intellectual debate, as a first requirement to obtain the academical status assigned to printed corpus.

The video-book interface, as a collaborative annotation tool for video corpus, proposes an evolution of Lignes de Temps towards a more collaborative approach with shared annotations and contributions through an online platform of consultation and annotation. It inherits from Lignes de Temps active reading and spatial representation of a temporal object but extends the cartography to the polemical space within a given corpus. In that sense, it seems to crystallize visually the collective individuation process that the instruments for critics conceived at IRI tends to disseminate in the digital cultural medias.

A video-book is defined by the combination of the video media itself, a set of descriptive metadatas given by its author (chapters and sequence segmentation, tags, title, description), a set of interpretative metadatas (annotations), embedding polemical discourse and contributions of users, and the relationships between the different elements of the corpus (video segments, annotations), structuring and organizing the debate. The descriptive metadata given by the author carry an editorial approach of the video-book structure that can be considered as a top-down input in comparaison to the only bottom-up approach proposed with Lignes de Temps. As we will see below, this is an essential element in the concept of video-book, which is
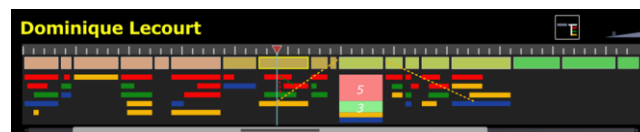

Figure 3: Spatial Representation of a Video-Book

necessarily associated with an author.

Here the contributions of readers are not organized in lines horizontally but rather vertically emphasizing the editorial template given by the author's segmentation and the relationship between an annotation and the referred segment. Indeed the spatial representation of the video-book distinguishes the segmentation from the annotations but build visually the dependence between them independently of the contributors. Therefore this representation does not aim at comparing lines of contributions, but rather at perceiving visually the polemic around the video-book.

The interface displays the following modules[Figure 4] : an augmented player [i] displaying the current chapter informations and the associated annotations, a spatial representation [ii] of several video-book timelines, an annotation module [iii], a polemical cartography [iv] of the current video-book and a navigation module [v] with corpus listing and a set of favorite elements extracted from the corpus.

Format of Metadata

The innovation of video-book when compared to Lignes de Temps is an evolution of the metadata format. It integrates attributes of relationships between metadatas, for instance relationships between a segment of a video and an annotation element. As a metadata, an annotation element is a data about a data. In a way, it describes it although we prefer to think that it enriches it. As a *data of*, an annotation is linked by essence to a data or a content. What the video-book proposes is to link metadatas to each other.
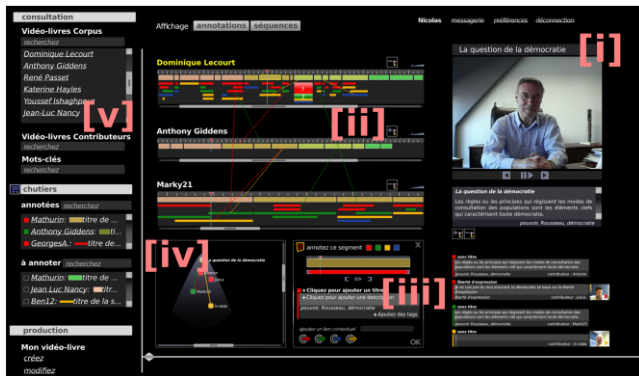

Figure 4: The Video-Book Interface

Linking metadatas is a powerful way to enrich the metadatas, i.e. to enrich the content itself. Coming back to the original intention of provoking debate and academical discourse around video content, links between metadatas produce meaningful associations of ideas, bringing the content to another level of intelligibility. This new type of association between elements opens horizons for extensive manipulations of video content. As a collaborative process of enrichment of the content, it lets foreseeing new ways of apprehension of any recorded content that can be films, interviews, talks, seminars, etc. The corpus would be collectively analyzed, tagged, classified and eventually criticized towards participative editorial selection and debate.

The relationship between metadata could be of different types according to the application. We based the relationships on polemical attributes with the ambition to stimulate debate and academical discourse over the content. The attributes were defined by the previous Semantic Kompass and are to be taken amongst those four : Agreement, Disagreement, Reference, Question.

In addition, extra polemical links can be embedded into the metadata element towards other elements. For instance, one could link his annotation to annotations stating similar opinions with agreeing relationship and simultaneously to annotations stating the opposite opinion with disagreeing relationship.

Linked metadata with polemical relationships form together a polemical network that can be visualized into a graph where debates, arguments, concepts emerge amongst the contributions of the users. Extrapolation to other purposes and applications is conceivable with the use of diverse classifications of relationships, such as *like/dislike, belong to/inherit*, etc, shaping accordingly the geometry of a relational graph.

Contextualisation

Classifiying the metadata in relation with other elements of content or elements of metadatas is equivalent with contextualising the metadata. Ontologies tends to contextualise metadata with a rather different approach based on the semantical relationships between concepts (Tom *is a* Cat *is a* Felin *is a* mammal, etc.). Here The approach of metadata contextualisation is to create polemical associations of opinions rather than concepts, which enriches metadata with a higher level of interpretation. Indeed opinion is a piece of discourse that cannot be interpretated by a machine and therefore cannot by organized. Whereas contextualizing opinions by the means of collaborative active reading empowers the machine for organizing and representing a corpus and therefore empowers the readers for interpretating and navigating within a corpus and its associated discourse.

One foresees in this process of collaborative network of metadata a sort of dynamic ontology based on characterized relationships between opinions rather than databased relationships between concepts.

LIMITS & PROSPECTS

Tagging itself, in this context, could evolve in order to better match the kind of polemical semantics encouraged by an applications like video-book. Agreement, refutation, question are currently treated as relations expressed by links between annotations. Against the purely descriptive approach of metadata found in ontologies and somehow duplicated in folksonomies when tags are believed to denote "concepts", it is possible to envision an alternative which consist in enriching tagging by adding information.

In video-book, users are encouraged to add explicit relations ("associations") between metadata. These relations are typed in such a way as to denote actions. Such actions can be performed by using human language as a medium. That is what the field of linguistic called pragmatics and its philosophical ancestors (from Scotish philosopher Thomas Reid and German phenomenologist Adolf Reinach to classical works by J.L.Austin and John Searle) traditionally dubbed "speech acts". To better fit the kind of polemical semantics found in video-book, a model of tagging would have to feature these two elements.

We intend to do this by reusing NiceTag ontology [3]. NiceTag is an RDF model of tagging designed for the Semantic Web. NiceTag doesn't describe tags but rather social acts, actions of tagging, thanks to an extension of the RDF model, named graphs.

In their seminal article on named graphs, Carroll et al. [4] expressed the need to embody social acts with some record. This naturally applies to the case of representing social tagging. In the NiceTag model and experiment, tag actions are defined as a subclass of named graphs (modeled as `rdfg:Graph` [4]; see also [5] for the detailed implementation in RDF/XML syntax) called `TagAction` which embodies one single act of tagging (see fig. below). The triples contained in the named graph represent the link, modeled with the property `isRelatedTo`, between an instance of the class `irw:Resource` and a sign

Figure 5: A Tag Action as a named graph

(modeled as an instance of `rdfs:Resource`).

The URI of the named graph of the act of tagging identifies a resource that can be described and typed (a video segment for instance). To account for the nature of the different possible tag actions, various subclasses of the `TagAction` class were defined. For instance named graph are typed in order to distinguish tagging performed by machines (`AutoTagAction`) from tagging performed by humans (`ManualTagAction`), or even more complex types of tagging as those involving machine tags (`MachineTagAction`). In addition, any number of properties can be attached to describe the place where tag actions are stored, the account of the user who tagged, the date the tagging act occurred, etc.

Contrary to most models of tagging, relations between the tagged resource and the label of the tag are made explicit through the `nt:relatedTo` property. This is of paramount importance when it comes to adding context to tags. When relations are explicitly stated, a single label may thence be used in various contexts to express different relations to a resource.

While the relation between the tagged resource and the label is easily captured with named graphs, the actions that are accomplished by typing relations in video-book (agree, disagree, ask a question, etc.) still remain to be grasped. This can also be accomplished with NiceTag. The class `nt:TagAction`, which describes the named graphs that encapsulate the *act* of tagging, can itself be of different kinds, all very similar to speech acts. `nt:Assert` is only one among many (other tag actions built-in the NiceTag model include "Share", "Aggregate", "Evaluate", "Ask a question", etc.).

Eventually, the two distinctive element of video-book (typed relations and actions) effortlessly lend themselves to modelization in NiceTag. Actually, tagging remains different from associating annotations. Yet, it could easily be adapted to serve the goal defined in video-book. First, other tag actions could be devised, suitable for the polemical semantics envisioned. Dissent, refute, and other similar speech acts, could all be easily implemented in the NiceTag model with corresponding (typed) relations and labels. In a collaborative environment, these rhetorical means, once given back to the community of users, could help to leverage people's reactions by providing them with the tools fit for criticizing the content of the videos published inside video-book. Instead of simply agreeing or disagreeing, their agency could be extended so as to encompass the possibility to identify arguments weaknesses or fallacies (contradictions, arguments of authority, straw man arguments, etc.).

Furthermore, this could be achieved by remaining faithful to the spirit of video-book. Once a tag has been chosen, it is possible for anyone to state whether they agree, disagree or ask for the justification that goes with such or such a choice. SRTag [2], a vocabulary based on NiceTag, extends the model of tagging so as to represent and keep track of diverging viewpoints by using named graphs.

The principle behind this model is to encapsulate statements about tags. The statements modeled in SRTag strongly resemble the relations between descriptors find in thesauri: broader, narrower, related to, etc; (labels) within a named graph which can in turn be typed with the class `srt:TagSemanticStatement` or some more precise subclasses (it shall be noted that the relationships between labels can be taken from any model). Users opinions on

the asserted relations can thus be captured and tracked back, allowing for the curation of diverging points of view.

This way, NiceTag and SRTag both make it possible to reassert the fundamentals of video-book while keeping the well-known advantages of tagging, particularly the possibility to add a chosen label and to leverage single users activity in a community-created
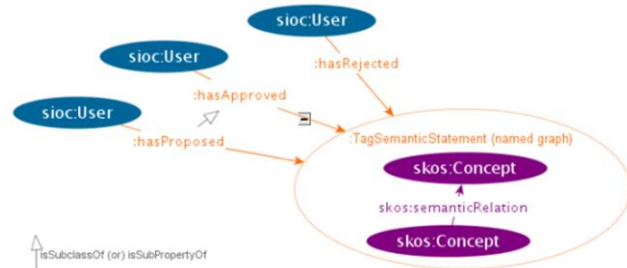
Figure 6: Graphical representation of the core of model of the SRTag vocabulary by Limpens et al [2]

folksonomy.

Acknowledgement

References

[1] Puig, V. and X. Sirven, *Lignes de temps: Involving cinema exhibition visitors using mobile and on-line film annotation*, in J. Trant and D. Bearman (eds.). Museums and the Web 2007: Proceedings, Toronto: Archives & Museum Informatics, published March 1, 2007 http://www.archimuse.com/mw2007/papers/puig/puig.html

[2] Limpens, F., Gandon, F. and Buffa, M. 2010 Helping online communities to semantically enrich folksonomies. In: *Proceedings of the WebSci10: Extending the Frontiers of Society On-Line* (Raleigh, US, April 26-27, 2010).

[3] Limpens, F., Monnin, A., Laniado, D. and Gandon, F., Nicetag ontology : tags as named graph. In: *Proceeding of SNI09 - Social Networks Interoperability, 1st International Workshop 2009 at the 4th Asian Semantic Web Conference (ASWC 2009)* (Shanghai, China, December 7-9, 2009).

[4] Carroll, J. J., Bizer, C., Hayes, P., and Stickler, P. 2005. Named graphs, provenance and trust. In *Proceedings of the 14th international Conference on World Wide Web* (Chiba, Japan, May 10 - 14, 2005).

[5] Gandon F., Bottolier, V., Corby, O., and Durville, P. RDF/XML source declaration. W3C Member Submission. http://www.w3.org/Submission/rdfsource/ (2007)