Open system for indexing and retrieving multimedia information

Hanène MAGHREBI,
Hanene.maghrebi@loria.fr
Nancy 2- LORIA
Amos DAVID
Amos David@loria.fr
Nancy 2-LORIA

Abstract

Managing the increasing growth of multimedia content still poses some problems. The challenge is to propose relevant information to the users among the large volume of information available. The main idea that drives our approach is to provide an open information retrieval system, which can adapt its results to several contexts of use of information. In this paper, we review multimedia information retrieval systems from the perspective of user's information needs. We solve the indexing and the representing problem by annotation process.

Résumé

La gestion de l'information multimédia soulève encore quelques problèmes. Le challenge est de pouvoir proposer à l'utilisateur des informations pertinentes parmi le grand nombre d'information qui ne cesse de s'accroître. Dans cette alignée, nous proposons un Système Ouvert d'Information capable d'adapter ses résultats aux différents besoins informationnels des utilisateurs et aux différents contextes dans lesquels l'information va être utilisée. Dans ce papier, nous repensons la recherche et l'indexation des informations multimédias d'une perspective utilisateur et utilisation via le processus d'annotation.

1. Introduction

In the last few years thanks to information and communication technologies such as Internet and digitization, multimedia information has become an important and rich source of knowledge. Managing the increasing growth of multimedia content still poses some problems. The challenge is to propose relevant information to the users among the large volume of information available.

The complexity is that multimedia information is compound information composed of a variety of information elements (sound, text and image), which are difficult to represent, index and access. In fact, in multimedia information, the text or the sound elements accompanying the image guides the user in his interpretation of the image and bring out a meaning in which the user never thought of. An image cannot be pinpointed with universal consensus because there is no consensus on universal meaning. The semantic concepts depicted in, or otherwise emerging from, an image is specific to a user. When, it is not obvious to grasp the meaning of image in the multimedia information, it becomes difficult to represent and manage it on the one hand and to express the information needs by the user on the other hand. As a matter of fact, user information needs can concern the full document or specific elements of information such as scene, shot in video, audio speech of a personality, etc. The challenge with the modern system is to satisfy the different but precise users' information needs. The relevant questions which can arise are:

1. How can a system best respond to the user queries if it does not consider them since the beginning of the development of the system?
2. How can we ensure a precise response to such queries?
3. How can we be sure that the user knows his informtion needs, and /or can express them correctly and formulate them in queries?
4. How can we reduce the distance between user's information needs and the information in the document base?
5. How can we allow the user to access directly the relevent information?

In our view, this problem can be alleviated by integrating the user information needs into the system herby making it open to index and to retreive the infromation. With open system we can assist the user to define his information problem and to clarify his queries. Moreover, the system becomes flexible, making it possible to adapt it to the various informational needs of the different users. The aim of our work is to present a personalised multimedia information retrieval system, which allows an active and variety of interaction.
In this paper, we propose our approach of Open Multimedia Information Retrieval System centred on users' information needs and information use context.

This paper focus on user and his information needs as well as on the multimedia document and its specificities.Thus, our proposal of open system for indexing and retrieving multimedia information is based on user and his information needs to improve

the classical approaches of multimedia document representation and to extend the metatdata proposed by these appraoches.
In the second section, we present a literature review on multimedia document indexing and representation, followed by some basic multimedia information postulates in section 3.
In the fourth section, we present our open system for information retrieving and indexing. In the last section we validate our meta approach by presenting an example of cinematographic document exploitation integrated in our open system.

2. Related work

In order to allow exploitation of multimedia document, many models on document indexing and representation have been proposed. They can be grouped into two large categories:
1. In the first category, multimedia document indexing and representation are carried out by decorticating multimedia documents. A representation model is then associated to each medium. Thus the representation concern a component or a medium apart (text, sound, image) and not the multimedia document in its entirety. In this category, Loisant et al, 2002 propose a meta model with the goal of describing all types of media (image, video, audio, text). The objective of this model is to provide a media independent base for writing specific model according to the media type to be managed. Moenne-Loccoz et al 2004 propose a model, in the same category, for managing the specificities of video documents. In this model, emphasis was laid on the temporal aspect and the diversity of the nature of descriptors (high and low level).
2. In the second category, document modelling concern the document in its entirety. This category does not offer clear separation between information content and information container (Auffret, 1999) (Darmont et al, 2002). For example, the model proposed by Auffret 1999 allows to describe audiovisual documents content by a metadata structure and to relate each element of this structure to the corresponding content element (Amous et al, 2002).

The metadata which are data of data concern the media (the containers) and the content. These metadata are particularly important for multimedia documents because without them the richness of these components remain unexploitable. To show the diversity of metadata that can concern multimedia document, we give some examples:

Metadata data concerning Audio element: the descriptors of these elements can concern the voice stamp, the silence zone or music wordings or noise (Pinquier et al 2004), the instruments, the melodic contours (Peeters et al 2000), the language (Pelligrino 2004), the word subject (Yamron et al 2000).

Metadata data concerning Video element: the descriptors of this element can concern the colour, the texture, the spatial characteristics (Belongie et al, 1998), costume identification that facilitates people's recognition and their presence in a video sequence (Jaffre et al 2004).

Metadata data concerning Text element: a text element is always used to extract a group of metadata. For example: title, section, language, theme, keywords etc.

Contrary to textual document, the metadata that describe multimedia document have to take into consideration the different components of the document and its semantic richness as well as users' information need and the use context of the information being sought. Indeed, metadata in the model cited above are destined to describe and model document. Such models allow managing multimedia information i.e. heterogeneous and semi-structured information with an ultimate goal of responding to users' information needs.

Yet, it is not surprising to note that if user's information needs are not integrated right from the conceptual stage of information retrieval system, it will be difficult to satisfy them after (Maghrebi and David, 2006- 2007).

It will thus be interesting to model user at the same level with document in order to better satisfy user's information needs in terms of time and result relevance. Moreover, the great majority of these models do not consider the notion of semantic subjectivity in multimedia document representation and indexing. Remembering that multimedia document is a rich document as regards to its semantic and indexing, its representation should be associated to its semantic content. But the semantic notion is vast based on the possible interpretations of users. This poses another problem besides the semantic representation constraints of multimedia documents which is the problem of user's information need that can be presented at different levels.

3. Basic postulate

Multimedia document is a rich document considering its composition- image, sound and text- the relation between its components and the meaning which each element takes if it is treated alone or with other component. For instance, the image, the richest component in the multimedia document, supports

several interpretations. The same image can be understood and commentend upon differently. The text or the sound element accompanying the image guides the user in his interpretation of the image and bring out a meaning in which the user never thought of. “An image cannot be pinpointed with universal consensus…because there is no consensus on universal meaning: The semantic concepts depicted in, or otherwise emerging from, an image are individual to a user” (Urban and Jose 2004).

When, it is not obvious to grasp the meaning of this type of information, it becomes difficult to represent and to manage it on the one hand and to express the information needs by user on the other. A semantic gap can be installed between the meaning allotted by the user to the document, specifically the image and its representation in the system. In our resarch, we consider the semantic gap as a distance between:

1. The low level representation of the image and the high level that the humain gives as interpretation. In other words it is a gap between what we see by looking at the image (object, form, color) and the interpretation we associated with what we can perceive.
2. The user information needs and the expression used to specify it.
3. The user needs expression and its translation into queries.
4. The user information needs and the solution suggested by the system.
5. The user information needs and the context of use of this information.

In summary, we think that the main problem which causes the gap between the user information need and the solution suggested by the system is at first a problem of expression; user’s expressed needs, user formulated queries, and expression used to represent and index multimedia document.

The semantic gap in multimedia processing is defined by Smeulders et al as “the discrepancy between low-level features or content descriptors that can be computed automatically with current algorithms, and the richness of semantics in the user queries due to the subjectivity of high – level human interpretations of audiovisual media”(Smeulders et al 2000).

For multimedia document, the expression of information needs is subjective; there representation requires consideration of denoted and connoted aspect. One medium of multimedia document (the video, image, etc…) enables us to see something (denoted aspect) and to feel or to imagine something (connoted aspect). This polysemic aspect brings several meanings; the interpretation depends on the user; his culture, his background and the context in which the information is emitted and received.

The connetced (connotation) aspect of multimedia document is often ignored in information retrieval system, whereas for instance the image is represented according to the low level (texture, color…); the user formulates his information needs according to the connoted aspect with semantic words.
The matching between the documents represented according to his denoted aspect and the user’s information needs with reference to the connoted aspect can only evoke the semantic gap.

The interpretation given by the user to the symbol of an image on the multimedia document is a problematic subject. Semantic and linguistic information associated to the perceived component of the image is not the same for all users. A form of non concordance between perceptual information on the image and the interpretation associated by various users can generate an ambiguity of interpretation. The user expresses his needs semantically where as the representation of multimedia document in the system is done in relation to the specifities of each medium.

Let us first review what we mean by user information needs and the problems that it can pose.

3.1 User information need problem

Studies showed that user information needs are complex (Spink et al 2000) (Wilson 1999-2000).

Information needs would be a lack of information in a situation relative to a decisional problem. This decisional situation makes the user to begin a search for information and determines the process of information retrieval. The information retrieval field is undertaken with the aim of satisfying user information need.

The informational need is individual and subjective. The user may not be acquainted with his information needs. He can wrongly determine his information retrieval problem. It can be the problem of expression of the information needs. It can also be the problem of identification of his informationnel or decisional problem. If the user wrongly determines his decisional problem, i.e. the reasons for which he seeks for information, his process of search for information will be erroneous. Only the user can assert if he had made mistake in his defined decisional problem. The difficulty with the user is to clarify or make explicit his information need or his decisional problem. Bouaka 2004 proposed a model to make explicit a decisional problem to assist the user.

In fact, the capacity of the user to clarify his information requirements depends also on the level of his knowledge of the domain in which he makes his research and also on his knowledge of the information system. A user who has knowledge in the field in which he makes his search has more capability to formulate his information needs. In other words, knowledge of the area of search allows the user to better determine his information needs. The Knowledge of the domain can be represented in information retrieval system in the form of thesaurus, ontology or index…this kind of representation can help the user to specify his information needs.

3.2 Query formulation problem

Query is a formulation of information needs in system language. It is a request for information from a database that gives the user the abiltity to find the information that he needs. User creates his query in relation to his information needs.

There are three general methods for posting queries.

1. The user can post his queries by selecting parameters from the proposition of the system. This method guides the user. However the user might not find the expression that reflects his needs.
2. The user expresses the query by specifying the fields and values.
3. The user might make requests for information in form of queries written in a special query language e.g. sql.

In order to ensure adequate responses to the user information need, in other words to bridge the gap between user information needs and the solution suggested by the system and the semantic gap, we have to be sure that the user understands and know his information needs and can express them in form of attributes and values.The satisfaction of user needs is tightly related to the representation of the user and his needs in the information retrieval system. In order to ensure precise responses to the user, we have to be sure that we understand the user information needs.

4. Contribution: Open System for retrieving and indexing multimedia document

From the proceeding studies and the presented postulate above, we notice on one hand a lack of flexibility in the management of the multimedia documents and on the other hand the absence of representation of user's information need.

We consider that in the era of semantic web and web 3.0, information indexing and representation as well as retrieval should be dependent on users. The user can index the semantic aspect through annotation process since the semantic aspect of multimedia information, as we have seen above, varies from one user to another.

By Open System, we imply a system made by user for user. A system where users have the possibility to propose or to add what they think is missing as attributes or sub-attributes (in other words the metadata of retrieval) and to annotate the existing attributes as well as data. We share the same opinion with Zacklad 2007, when he represents open information retrieval not only by tools to access relevant documents or same relevant documentary fragments but to delimit the problem contours by identifying the existing information and those which do not exist: it is a discovery and learning process that allows posting a problem.

We can summarize our approach, on which we base the development of our system, by Figure 1 composed of three stages.

Firstly, we repeat that the starting point of our research is a group of problems that we raised in section 2 and 3. This can be summarized as the distance between user's information needs and the multimedia information that he needs.

Secondly, in order to reduce the gaps, we propose to model user at the same level as multimedia document for better satisfaction of information needs in terms of time and response relevance. The modeling that we propose is done by contextualizing the use for which the retrieved information will be put. Thus the user's information need is translated to information use context.

We assume that the user seeks information for a particular need. Required information will be used in a specific context. According to our explantation in the previous section (3) we noticed that it is easier for a user to express information use context than to specify his information need.

We think that there is a narrow dependency between user needs and information use context.The satisfaction of user needs is related to the context in which the information will be used. The dependency is semantic and cause-effect relation. Information use context is bound with user information needs. In our resarch, the context is bound to contain information that can help refine user's information needs.

Thirdly, we propose a meta–approach that allows a user to carry out information retrieval and at the same time participate in indexing multimedia docu-

ments in the database. By meta-approach, we mean two approaches that exist side by side where one fits into another.

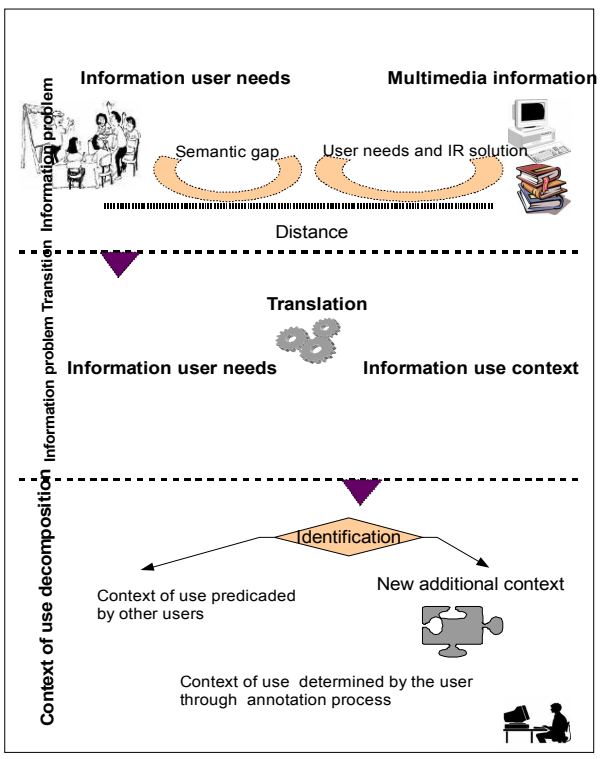

Figure 1: Objective and overview of our approach

4. Functional Characteristics of the proposal system

The information retrieval approach in our system is an open and flexible approach that reflects the integration of user and their needs in the system. The principle of this first approach is to help the user to better apprehend his information needs and to help him in clarifying his information retrieval problem i.e. encircling the level of granularity of retrieval (e-g: scene, shot…) and defining the context in which the information will be used. Thus, the user fills in the values concerning:

-Attributes representing the use context of information being sought.

-Attributes representing the multimedia document (Figure2).

We tried to represent multimedia document in a best possible detailed form using inspired attributes from norms (Dublin Core metadata) and organizational structure of cinematographic document. By cross linking the different attributes (use context, multimedia document), the user can have relevant information from which he can extract knowledge. This cross linking of attributes also called concurrence querying allows user to build knowledge on existing documents in the database as well as on the use context of these types of documents.

The annotation process in our system is an indexing approach allowing users to identify and regroup documents or their sub-elements under a particular use context. The user can attribute the same document to different use contexts. This information contextualization done by user serves as key to others latter.

The goal is to go beyond document retrieval to knowledge sharing – the knowledge of information use as well as the knowledge of formulation and translation of information need into query - by keeping track of users' search histories.

This can allow a user to have an idea on the types of problem that can be resolved by the system and to acquire knowledge of the way of formulating and representing his problem.

The user can also see the annotation and evaluation on the results proposed by the system to the queries of other users who had used the system in the past. He can also see their information needs and use context. Indexing can then take a new scope which can be defined under collaborative work between different users.

5. Application

Our research application is related to cinematographic corpus. These are the collections of films proposed by the European Institute of Cinematography and Audiovisual (IECA). Our potentials users are lecturers, multimedia producers and socio-economic actors. The information needs for all these users are not the same. We are working on the development of a multimedia information retrieval system that can manage various types of document and answer several users' information needs.

Taking into consideration the diversity of users and their information needs, we give below an example that concerns the same multimedia document from the corpus of IECA to which our work is applied.

This example shows how a document can be the object of different information needs related to different use context. The three contexts (learning, teaching, documentation) that illustrate the approach are the contexts that have been cited by users of ICEA during the first meeting we had with them. We pre-

ferred maintaining these use contexts in the information system with the goal of helping users to understand what is implied “use context” and to guide them to filling the concerned attributes.

5.1 Use Context: Learning

We take the case of a student searching for information for solving an exercise. He seeks to know the steps taken by director “*SERGIO LEONE”* in filming “*revolver shooting”* scene in a key sequence in the Western movie “The Good, the Bad and the Ugly” with the actor *“Clint Eastwood”.* The student also seeks for a shot of the star.
We consider “learning” as a sub-attribute of use context attribute from where proceeds other sub–attribute “cinema course”. The cinema course can be decorticated into other attributes: scenario writing course, mixing course, etc. The student starts his research by precising the exact nature of the course and then the level of granularity of the document being sought (e.g. sequence, scene, shot). For the Western movie, the shot type that he is searching for can concern either American shot or another shot. The precision of information use context by the student reduce the distance between him and the information that he needs.

5.2 Use context: Teaching

In the framework of IECA, teaching heterogeneity, diversity of training type (technical, theoretical …) and the knowledge to be acquired induce a differentiation in needs and information use. We take as example a teacher of music analysis and history that resort to the same movie “The Good, the Bad and the Ugly” to explain how the compositor of the music of the movie “Ennio Morricone” used Whistling denotations of “Alessandro Alessandro” and Yodel to represent three principal character and how this contributes to creating the atmosphere that characterizes the movie.

The system proposes a scene by scene representation of the movie. The teacher, for instance can get the scene where we find the melodies that are repeated throughout the movie indicating the appearing and the disappearing of a character in a scene. The teacher can also refine his research and specify the type of sound or instrument used and associated to a character in each scene, thus the system can propose to him: apparition scenes of the character *“Blodin”* where a soprano flute was played; the scene associated to the character *“Sentenza”* where a ocarina was played and the scene associated to *“Tuco”* where a human voice whistled (Figure 3).
We take another user (a scenarist–teacher) willing to explain to his students the shooting of the famous “duel at three” sequence of the same movie. Through our system, the teacher can analyze this sequence and have the shot type that the director, Sergio Leone, cohabited to valorize the scenario. We see again that the more the sounds are adequately represented with their use contexts, the better the system efficiency.

5.3 Use Context: Documentation

In this context, the user seeks to improve his knowledge on a subject in a particular domain. We take as instance a movie-lover that is seeking for information concerning the budget for producing the movie, the actor and their role etc. The user can effectuate his research from the different attributes representing the content or the container of this document. These attributes connected to the documentation context are a priori identifiable.

What we consider as our major contribution is the possibility to attach the use contexts to the documents. In this way new contexts can be added to the document representation and information search can proceed starting with the attached use contexts.
We have demonstrated through the above instances that when the user adds the sub elements of a context or sub contexts by defining a new use context, he has the more possibilities to have a more relevant answer. The personalization of the information is thus enhanced.

6. Conclusion and perspective

Our vision of open multimedia information system is intrinsically dependent on user’s information needs and use context of information being sought. An open information system does not correspond only to a flexible system that allows a personalized access to information, avoiding semantic problem that multimedia document can pose, but it is also a system that allows users to participate in the indexing and representation of the documents in the document base. It is in knowledge sharing perspective that information indexing and retrieving becomes interesting. Our perspective is to see how to manage some technical aspects of the system such as the verification of the existence of attributes proposed by users under another appellation.

References and bibliography

Amous, I, et Jedidi, A. 2002. Modélisation des méta-données pour une recomposition dynamique des documents : *Actes du 20 ième congrès INFOR-SID, Nantes-France 4-7 Avril juin 2002,* éd. Hermès, 243-258.

Auffret, G, 1999.Indexation documentaire et documents virtuels vers un nouveau mode de lecture des documents audiovisuels, *Actes de la 11 ème conférence Francophone sur l'Interaction Homme-Machine, Montpellier, novembre*. Retrieved March 3, 2008from http://www.lgi2p.ema.fr/~multimedia/ihm99/articles/auffretihm99.pdf

Bouaka Najoua, 2004. Développement d'un modèle pour l'explication d'un problème décisionnel : un outil d'aide à la décision dans un contexte d'intelligence économique, thèse, Sciences de l'Information et de la Communication, Université Nancy 2, France.

Belongie,S, Carson C, Gressnspan, H, Malik, J, 1998, Color-and Texture- based Image Segmentation Using the Expectation –Maximization Algorithm and Its Application to Content-Based Image Retrieval, ICCV, Mumbai, Inde,P675-82.

Darmont,J, Boussaid O, et Bentayeb,F, 2002, Warehousing Web Data, *4th International Conference on Conference on Information Integration and Web based Application and services(iiWAS02, Bandung, Indonesia,* p148-152.

Goria, stephane 2007.L'expression de problème et la médiation informationnelle.VDM Verlag DR.Muller

Jaffre G, Joly P, 2004, Costume: a new Feature for Automatic Video Content Indexing, RIAO, Avignon, april, p 314-325.

Loisant, E, Ishikawa, H, Martinez, J, 2002, Designing *a Model Independent Multimedia Database, Days of* science and Technology, Tokyo, Japan.

Marchionini, Gary. (2006). Leveraging context for adaptive multimedia retrieval: a matter of control. In M. Detyniecki et al (Eds.), journal Adaptive Multimedia Retrieval, LNCS 3877, (pp. 35 – 43). Berlin Heidelberg: Springer-Verlag. Retrived mars6, 2008 from http://www.springerlink.com/content/950t0530x230t438/fulltext.pdf

Maghrebi, Hanene, and David, Amos. 2006. toward a model for the representation of multimedia information based on users' needs: economic intelligence approach. In 4th International conference on Multimedia and Information and Communication Technologies in Education m-ICTE (proceedings), 195-200. Retrieved February 5, 2008, from http://www.formatex.org/micte2006/pdf/195-200.pdf

Maghrebi, Hanene, and David, Amos. 2007. Integrating users' needs into multimedia information retrieval system. *3rd International Conference on Computer Science ATINER, Greece*, 2007. Retrieved February 5, 2008, from http://hal.inria.fr/inria-00167511/fr/

Moenne-Loccoz, N, Janvier, B, Marchant–Maillet, S, Bruno, E. 2004, Managing Vidéo Collections at large, First International workshop On Computer Vision meets Databases (CVDB 2004), paris.

Pinquier, J, Obrecht, R. 2004, Jingle detection and identification in audio documents, ICASSP' 2204, Montréal, Canada, mai 2004.

Peeters, G, Mac Adams St, Herrera, P. 2000, Instrument sound description in the context of MPEG7, International Computer Music Conference, San Francisco, USA.

Pelligrino, F, Farians, J, Rouas, J. 2004, Automatic estimation of Speaking Rate in Multilangual Spontaneous Spech, International conference on Speech Prosoday 2004, Nara, Japan, Mars 2004.

Urban, Jana. and Jose, Joemon. M. 2004. EGO: a personalised multimedia tool. In the proceedings of the second international workshop on adaptive multimedia retrieval (AMR'04), Valencia, Spain. 3-17. Retrieved March 3, 2008 from http://citeseer.ist.psu.edu/urban04ego.html.

Smeulders, Arnold W.M. M. Worrring, S. Santini, Amarnanath,Gupta, and Ramesh, Jain. Content – based image retrieval at the end of the early years; IEEE. Transactions on Pattern analysis and machine Intelligence, vol22, No12, December 2000.] Retrived mars6, 2008 from http://ieeexplore.ieee.org/iel5/34/19391/00895972.pdf

Spink, Amanda, Tom, Wilson, David, Ellis, Nigel ford.1998, Modeling Users' Successive Searches in Digital Environments. Retrived mars6, 2008 from http://www.dlib.org/dlib/april98/04spink.html

Yamron, JP, Knech,S, Van Mulbergt P, 2000.Dragon's Tracking and Detection System for the TDT 2000 Evaluation, Topic Detection and Tracking Workshop,P75-79.

Wilson, Tom. 1999. Models in information behavior research. Journal of documentation, 55(3), P. 249-269.

Zacklad, Manuel. 2007. Classification, thesaurus, ontologies, folksonomies: A comparison from the Open Information Research perspective. In Information Sharing in a Fragmented World: Crossing

Boundaries: 35 th Canadian Association for Information Science. Retrived mars6, 2008 from http://www.caisacsi.ca/proceedings/2007/zacklad_2007.pdf

Actes du 36e congrès annuel de l'Association canadienne des sciences de l'information (ACSI), University of British Columbia, Vancouver, 5 au 7 juin 2008

APPENDIX

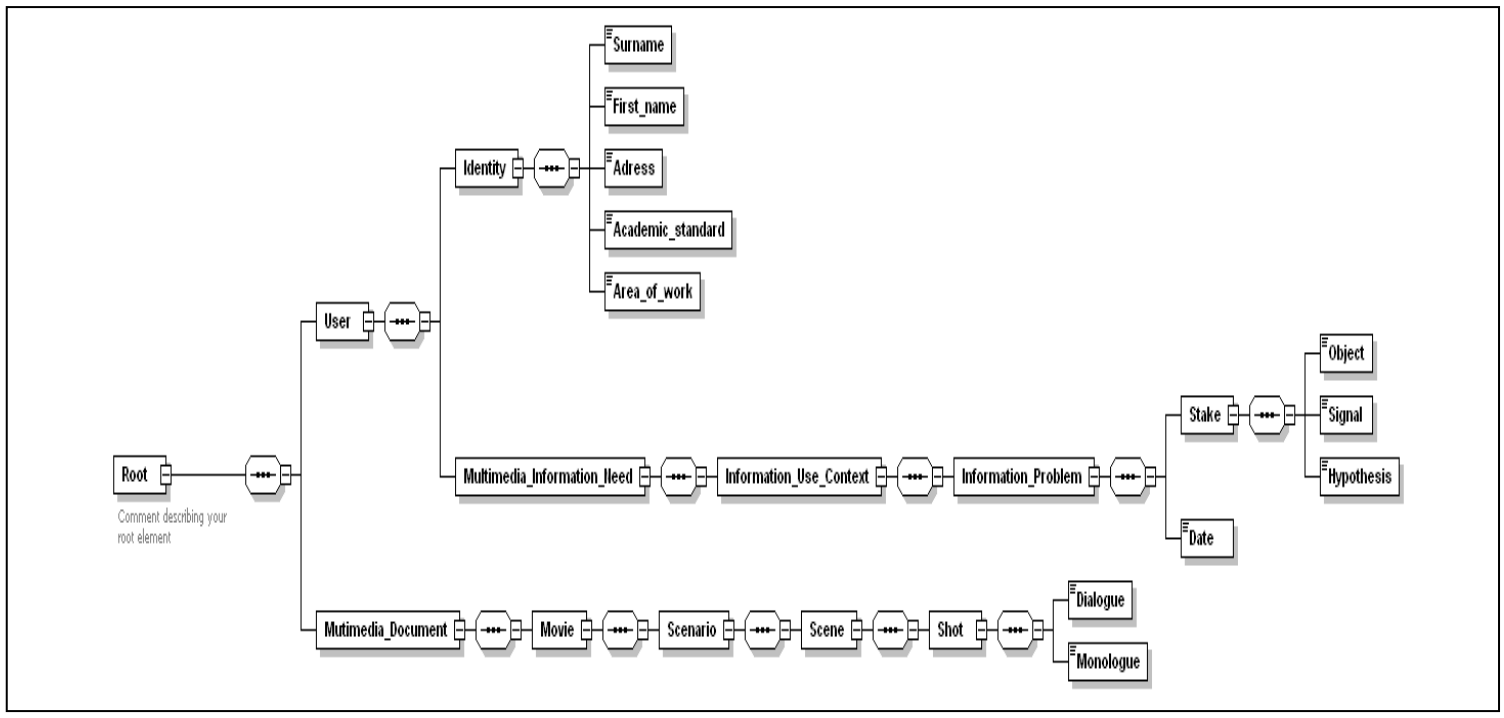

Figure 2: An XML schema of an Open System for Indexing and Retrieving Multimedia Information

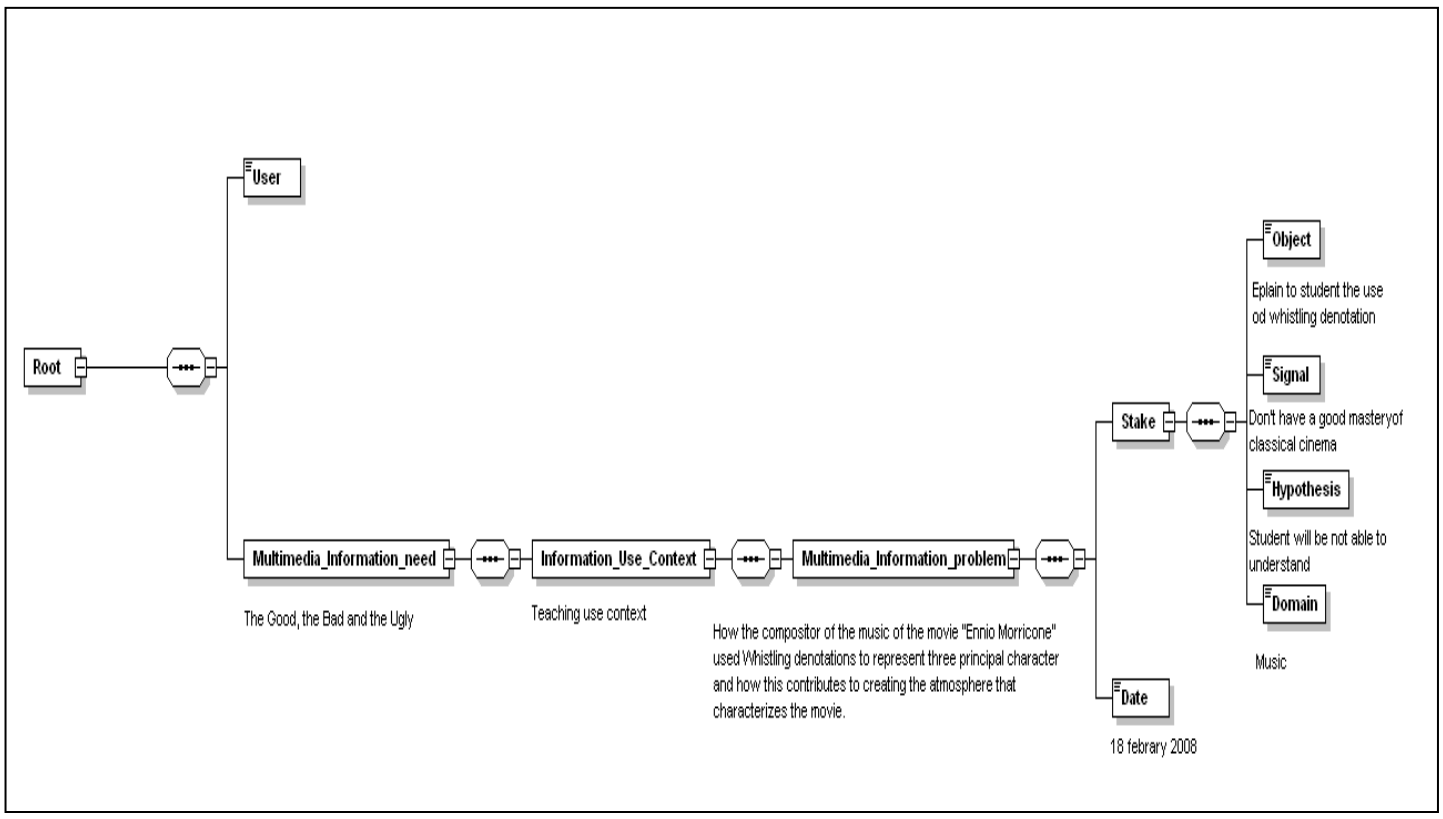

Figure 3: An XML schema representing the example given in the application section