

The Semantic Hole: Enthusiasm and Caution Around MultiMedia Information Retrieval

Roberto Raieli

University Roma Tre, Biblioteca delle Arti, piazza della Repubblica, 10, 00185 Roma, Italy,

<raieli@uniroma3.it>

Roberto Raieli, graduated in Philosophy and in Library Science, with a PhD in Library and Information Science. He is Librarian at the Art Library of the Roma Tre University, and editorial staff member of the journal *Bollettino AIB*. He has published several papers on MMIR, including one international book as co-editor, one book as editor and author, and one book as author.



Raieli, Roberto. **The Semantic Hole: Enthusiasm and Caution Around MultiMedia Information Retrieval**. *Knowledge Organization*, 39(1), 13-22. 8 references.

ABSTRACT: This paper centres on the tools for the management of new digital documents, which are not only textual, but also visual—video, audio or multimedia in the full sense. Among the aims is to demonstrate that operating within the terms of generic Information Retrieval through textual language only is limiting, and it is instead necessary to consider ampler criteria, such as those of MultiMedia Information Retrieval, according to which, every type of digital document can be analyzed and searched by the proper elements of language for its proper nature. MMIR is presented as the organic complex of the systems of Text Retrieval, Visual Retrieval, Video Retrieval, and Audio Retrieval, each of which has an approach to information management that handles the concrete textual, visual, audio, or video content of the documents directly, here defined as content-based. In conclusion, the limits of this content-based objective access to documents is underlined. The discrepancy known as the semantic gap is that which occurs between semantic-interpretive access and content-based access. Finally, the integration of these conceptions is explained, gathering and composing the merits and the advantages of each of the approaches and of the systems to access to information.

Received 22 July 2011; Accepted 22 July 2011

1.0 Introduction

MultiMedia Information Retrieval (MMIR) technologies are well-developed within the ambits of computer engineering, artificial intelligence, computer vision, and audio processing, while the interest for the methodological and operational revolution of MMIR must still be introduced to librarians, documentalists, and information managers. It is therefore necessary that librarians and documentalists acquire familiarity with these technologies which can be used to their advantage. The international domains of library and information science (LIS) and knowledge organization (KO) have to understand that, in the contemporary context, discussion of the development of MMIR systems is a priority. Thus information

managers will be able to address the development of these systems according to necessity of documental-ist order, and, above all, it will be possible to discover, to discuss, and to resolve fundamental matters, often hidden behind the enthusiasm of technological developments, like that of the *semantic gap*, that might be clear only if viewed from the perspective of LIS or KO.

The discussion must be focused on the tools for processing and searching as applied in the management of new digital documents, filed in great databases where there are not only mainly textual documents, but also documents of the visual kind—audiovisual, sonorous, or multimedia in the full sense. This area is problematic, as it directly connects to the organizational issues, dissemination, and fruition of

the documents, which are the central objectives of the renewed activity of libraries, archives, and documentation centres. This underlines the necessity of new multimedia modalities for the organization, which are searching and retrieval of all types of digital documents.

In information searching, it can be restrictive to operate within the terms of a *generic* information retrieval (IR). In traditional practice, every kind of document handling is compelled to the conditions of analysis and searching through textual language only; it is necessary instead to consider wider criteria, such as those of MMIR, in which every type of digital document can be analyzed and searched through the elements of language, or meta-language, functional for its proper nature. In fact, if it is not possible to search and retrieve a written document through audiovisual language elements, retrieving documents that consist of images or sounds through the use of descriptive texts cannot be considered an effective method. In databases in which the content of the documents is substantially textual, it is of some importance that the keys for access are terms and references extracted from within the content; on the contrary, in multimedia databases, it is rather inaccurate to attribute a textual description from the outside to contents founded upon a different structure of sense.

It is possible to speak of the precise and organic development of a principle from IR, to MMIR, as methodology evolving toward the amplification of the documentary space in which 1) IR is a system of processing and searching textual documents through terms, but can also be applied to visual, audio, and video documents; and, 2) MMIR is proposed as a general system of processing and searching through texts, images, and sounds, for documents of every type and medium. It is possible to differentiate, then, within the new, general, and integrated methodology of multimedia searching of MMIR, a method of text retrieval (TR) based on textual information for the handling and searching of textual documents, a method of visual retrieval (VR) based on visual data for the search of visual documents, one of video retrieval (VDR) based on audiovisual data for the processing of videos, and one of audio retrieval (AR) based on sonorous data for the processing and the search of audio documents.

MMIR	Systems of <i>MultiMedia Information Retrieval</i> , in which content-based characteristics naturally integrate also term-based ways, since to operate on the contents doesn't distinguish <i>a priori</i> if these fundamentally are textual, audiovisual, sonorous or other. They will specialize in:
TR	systems of <i>Text Retrieval</i> , in which the same principle of the content-based handling legitimates and guarantees the use of textual data to search textual information (words by full text, names, titles, dates, etc.);
VR	systems of <i>Visual Retrieval</i> , in which files of still images are processed and searched through visual data proper of the files themselves (colours, lines, edges, structures, shapes, orientations, spatial distributions, etc.);
VDR	systems of <i>Video Retrieval</i> , where for handling audiovisual documents the audiovisual language is used (images of film, movements of objects in shots, variations of figurations, movements of cutting, audio tracks, etc.);
AR	systems of <i>Audio Retrieval</i> , in which sonorous information is searched in measures of sounds (times, frequencies, heights, rhythms, timbres, melodies, etc.).

Table 1. MMIR and its subsystems.

2.0 MMIR: a revolutionary organic system

An appropriate query language has to be consistent with the objective content of the document and the kind of information that is being sought. MMIR systems, therefore, have to establish a search approach that draws directly on the objective content of the documents. This is defined as content-based, as opposed to traditional systems of analysis and searching founded on terms of descriptions of such concrete content, which is defined as term-based.

Thanks to the possibilities offered by new technologies, MMIR systems allow the analysis, organization, and searching of multimedia documents, applying techniques of storage and retrieval that operate in a direct way on the contents of the digital objects in the database; the searching of images, audiovisuals, and sounds, as well as texts, can take place by exploiting specific characteristics of language in every document. Users are able, then, to consult the database with search strategies founded on similarity, or on other modalities like the approximation and relations of measures and values, using shapes, structures,

words, figures, movements, sounds, colours, edges, etc. as keys for the query. Therefore, a more beneficial system is one in which query formulation doesn't have to be forced within the limits of the language, but can be sent as it is created, and can be found and satisfied by the system as spontaneously and immediately as the user produces it.

The classical interfaces of text databases, which allow searching in an index composed exclusively of terms extracted from documents or inserted in textual metadata, must be replaced by interfaces that allow formulating the query in different dimensions, not only through terms, but also through images and sounds. In this way, searching will be possible in richer indexes substantially different from the traditional, composed of terms extracted from written texts or from spoken audiovisual sources, key images from a sequence, simple geometric figures, melodies, or forms, colours, movements and sounds generally. This will occur without excluding their importance in maintaining the terminological, descriptive or conceptual data, which are relative to aspects not specifically contained in documents. This requires the construction of new, specific indexes of multimedia data that are complete and effective, the elaboration of high-level query systems with many interconnected options, the development of data analysis algorithms able to calculate many variables, the realization of systems to evaluate and rank the results so as to improve the quality of the answer interacting with user indications, and, finally, in the development of paradigms of analysis and searching that can compare objective automatic representations from the computer with refined human intellectual analysis.

So, the MMIR introduces itself as a revolutionary organic system, specialized for the effective handling and organization of every kind of multimedia digital document. The more advanced MMIR systems can be very useful in support of theoretical research and creative practice, as much a tool for professionals as a guide for general users. A user query can simply and freely be constituted by the input of a model image or sound, with or without conceptual specifications through parameters or texts, and the system can retrieve documents that possess similar characteristics. Therefore, the user can always fully rely on his own intelligence and sensibility, his own creativeness and imagination, interacting with a system that is also predisposed to welcoming unpredictable variations in the path of search, and to understanding human strategy, learning every time from the behaviour of the researcher.

3.0 Problems inherent in MMIR

Even though the MMIR system has worth in the fact that it is still a novelty, it is nevertheless necessary to re-evaluate, with insight, its method, and thus, the importance of the semantic principles. It is necessary to foresee the integration of a conception of processing and searching that is revolutionarily parallel to a traditional semantic conception, gathering and composing the merits and the advantages of each of the approaches and systems of access. In the retrieval of multimedia documents, it is possible to reach an acceptable standard of precision and effectiveness only using a combination and mutual integration of search techniques and technologies combining the representation of the content, through textual, visual, sonorous, and audiovisual elements, with the definition of the concepts and meanings, through semantic terms. Substantially, it is necessary to structure the system integrating into one organic MMIR whole the techniques of TR, VR, VDR, and AR, but always making it clear that together with the integration among content-based modalities, there also must be organic integration of the semantic modalities of the processing and searching of documents.

First, the terminological query can be a preliminary method for selecting a section of the great quantity of documents in a database, and to centre the search with regard to information ambits, typologies, classes, titles, or authors. Subsequently, it can be used to clean up the inevitable specific noise of a content-based query, specifying degrees of semantic interpretation that the automatic system cannot detect in the direct analysis of the content characteristics representative of the document. Above all, however, the different procedures operate best in continuous interaction. This takes place then, solely in an organic query interface, allowing search strategies which, combining words, figures, movements, sounds and concepts, are useful for searching very complex documents, rich at all the levels of sense and meaning. Thus, they are able to overcome the many risks of falling into the gap created by handling either semantic data only or content data only.

It is very important, then, to discuss the effectiveness of the coldly mathematical procedures of MMIR in relation to the human objectives of the organization and searching of information and documents. The mechanical and absolute efficiency of these operations is a given, but the practical utility in relation to the demands of every end user is not. The mathematical rigidity of system operations, in fact, can pose

a problem not only for the demands and needs of users, but also for the defined general principles of liberty and sensibility of content-based systems. Therefore, the discussion of the value of mathematically objective operations completed by computer systems is intricate. In fact, these operations are devoid of errors produced by human evaluation of documents and their contents, but also deprived of the peculiar flexibility and intelligence of this evaluation in interpreting the subjective content aspects. Having exposed the many advantages of the content-based method, it is also necessary to point out the more dubious and least positive aspects.

The automatic procedures of MMIR systems avoid every superfluous passage of mediation by handling directly the characteristics of the original object, or more exactly, the data that comprise its digital document version. The processing of the files takes place in a procedure of abstraction from the document that reduces the documented object to a set of data, a sort of digital abstract, or model, retrieved through algorithms of dominant values. Possible errors and approximations of algorithms are due to known causes, and are calculable as systematic errors that can be kept in consideration in the management of the final result. In comparison with the individual variables of manual methods and with hidden errors of interpretation, the processes of automatic systems are often much more reliable, at least in certain contexts.

Nevertheless, very advanced and expensive hardware and software systems are necessary for effective handling of more complex multimedia documents, and this surely limits experimentation and application, as well. Above all, however, the automatic and content-based methods don't always have the most suitable result to satisfy the most elevated demands of researchers and experts, apart from those of common users. The sense of an object represented in a document, in fact, must be grasped in its true totality, in the full and simultaneous consideration of many sensitive and intellectual qualities, including that of aspect and of meaning, in the concrete and abstract. The systems of access directed to the objective content, instead, appear inadequate to indicate the multiplicity of the intellectual interpretive point, and the nonexistent sensibility of the computer cannot be completely produced by algorithmic elaborations of the numerical data representative of the qualities of documented objects.

So, MMIR systems maintain good validity in the case of a direct and content-objective approach to the document, but show a certain narrowness in the case of a theoretical and intellectual-interpretative ap-

proach. The difference and the distance between the two kinds of approaches can be called the semantic gap, and the purpose of the most advanced MMIR systems is to fill this conceptual, methodological, and operational void. To overcome this distance between human and computer requires a long search for the realization of computational algorithms and procedures of data elaboration that are not only mathematically efficient, but also pragmatically effective. Every theoretical perspective is valid only if it can consist in the construction of true multimedia search interfaces able to combine material and conceptual data (Enser 2008).

Moreover, if the second point represents the problem of the semantic gap, the first point can also represent an equally sufficient parallel problem, which has been called the sensory gap, or semiotic gap. The sensory gap could be the object of a parallel theoretical treating of aesthetical nature, that is not possible to discuss here, and it mines another foundation of the content-based system: the possibility for the computer to represent at least the concrete content of the document in a way that most accurately conveys the reality of the documented object (Smeulders et al. 2000).

4.0 The semantic gap

A large part of international literature indicates as a "semantic gap" the limits of content-objective consideration of documents and discrepancies in comparison to semantic-interpretive consideration. The semantic gap is defined as the "not coincidence" between information that can be drawn directly from a document and the different interpretation that every user might have of the same data in every specific situation. This is a critical point for the development of the MMIR: because the meaning of a multimedia document is rarely explicit, the purpose of the system is to give support to overcome such a void between the simplicity of the documentary processing offered by the computer and the rich semantic expectations of the user, namely for bridging the gap.

In a collaborative essay, Peter Enser, Jonathon Hare, Paul Lewis, and Christine Sandom deduce with awareness the characteristics of such a gap of representation. The representative levels of a document vary from that lower level, composed of the simple extraction of raw data, immediately elaborated by the computer, up to the higher level constituted by the semantics that it carries, as interpreted by users (Hare et al. 2006). On the first and lower level of representation are the immediate documents, the "raw media," whose deduction can be only generic. On the second level are the "fea-

ture vectors,” or descriptors, the characteristics of the content of the documents extracted by the content-based algorithms in the form of vectors or other mathematical descriptors. On the third level are general representations of the documented object, “prototype combinations” derived from the combinations of the vectors on the preceding level. On a superior level, these representations can be defined with symbolic labels, descriptive or nominal textual labels realizable by the human operator, as well as the automatic system. Only in the higher representative level, that of the abstraction, do we have the “full semantics” of the document, the objects, the parts, and the labels that can be interpreted by the human in relation to relationships and meanings.

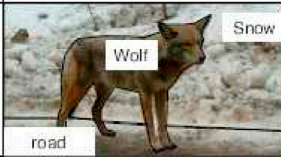


Semantics <i>object relationships and more</i>	Wolf on Road with Snow on Roadside in Yosemite National Park, California on 24/1/2004 at 23:19:11 GMT
Object Labels <i>symbolic names of objects</i>	
Objects <i>prototypical combinations of descriptors</i>	
Descriptors <i>feature-vectors</i>	Segmented blobs, Salient regions, Pixel-level histograms, Fourier descriptors, etc...
Raw Media <i>images</i>	

Figure 1. Hierarchy of the representative levels of a document, illustrated by Enser and colleagues.

Generally, the characteristics of the semantic gap vary from one case to the other, and they can be present in a more difficult way according to the level of complexity of the document handled by the system and by the operator. However, the problem of the gap is initiated in the second representative level, and it increases towards the semantic level.

There is, then, a classification of the different levels of information demands that the users set to a system of multimedia search. Users arrive, ultimately, at

a higher level of demand, formulating requests of documents with an intellectually refined value. The traditional systems of IR are occupied, in reality, by this kind of search, with all the limits of conceptual abstraction, and this informative level is more difficult to achieve through content-based systems, which are based on semiotic, more than semantic, consideration of the document. It is, thus, clear that there is a lot to learn by comparing the two methods, elaborating correct and valid solutions for MMIR systems.

4.1. Bridging from the bottom

The first solution proposed by Enser and colleagues is to attack the gap from the base, resorting to the system of auto-annotation. The automatic systems are able, in fact, to recognize, describe, and annotate the content of documents, identifying and also naming their parts. This solution is an attempt to resolve the first part of the gap, which is set between the level of the descriptor vectors and nominal labels, to allow a more effective combination of forms and terms in the query, using objective elements produced by the direct investigation of the system. The limits of the automatic system are, obviously, the complexity of the object and the quantity of details that characterize it.

Drawing on remarks about hybrid systems in another paper by Peter Enser (2000), other solutions can be appraised for the necessary integrated structuring of the different content-based and semantic techniques and technologies. The researcher always finds conclusions after a series of verifications of typical user search behaviour. Such a search pathway is started from the low levels to the various middle levels up to the high levels of representation and interpretation of the document, and all of those levels have to be realizable in the structure of the operating system. In the higher level of abstraction, they are really the characteristics of human reasoning, based on diffused knowledge that interprets contents, as opposed to the automatic modalities of the content-based search. In the paradigm of a searching method, therefore, neither the classical method, namely concept-based, nor the innovative content-based method can be excluded.

The solution of the problem can be, thus, represented by hybrid systems of retrieval. The more developed systems are those that try to build a bridge that covers the gap, if not just in reference to meaning at least with reference to sense and aspect. These procedures try to overcome the interpretive gap, allowing the operator to specify the characteristics of an object held to be meaningful, annotating them

through semantic labels on general models or directly on the metadata resultant by an automatic analysis of the documents. Such systems are founded on criteria of continuous interaction with the user, on the relevance feedback that allows the computer to learn, from human operations, whether to report the meanings of high-level with the contents of base level searches during the processes of query (Addis et al. 2005).

In comparison to the principles of organicity of the method of the MMIR, however, it is not enough to postulate a simple hybridization of techniques, where each remains almost alone; defining a new perspective of search is necessary to find a valid unique principle for the structuring of all the implicated methodologies.

4.2. Bridging from the top

Furthermore, the most probable solution given by Enser and colleagues is to attack the gap from the highest perspective, thus resolving its second and more complex part, considering the use of ontologies. However improved, an ample set of annotations and labels related to an object is distanced from representing it in its semantic richness, which seems, instead, likelier to be representable by positioning the object within an ontology. Ontologies are proposed as the new guide for the navigation of the semantic web, as a tool for the shared conceptualization of a

domain composed of classes of concepts, relationships between them, and information about its structure. The appeal of ontologies for MMIR systems is that they allow us, then, to make explicit part of the meaning of a document, which makes it possible to formulate the query through concepts and relations among concepts. The multimedia query can be semantically completed according to the way in which these tools are able to represent both the meanings of the objects and their relationships in a document, both the meaning of the whole document in a context, continually integrating the content-based search tools that revolve round the objects themselves (Hare et al. 2006).

The solution of the ontologies seems to represent a true solution of principle for MMIR, useful to establish an organic approach thus valid for all the types of multimedia documents, and able to take into account univocally their concrete and conceptual representation, content-based and semantic. The documents, whatever nature they are, can always be attentively inserted in logical spaces in relation among them, being able to be searched without influences inside such semantics positions with the content-based methods proper for each.

A concrete example of this solution is the system MAVIS 2 (Multimedia Architecture for Video, Image and Sound), projected by Enser and colleagues as a tool of fully organic integration among data of content, terms, and conceptual classifications. The struc-

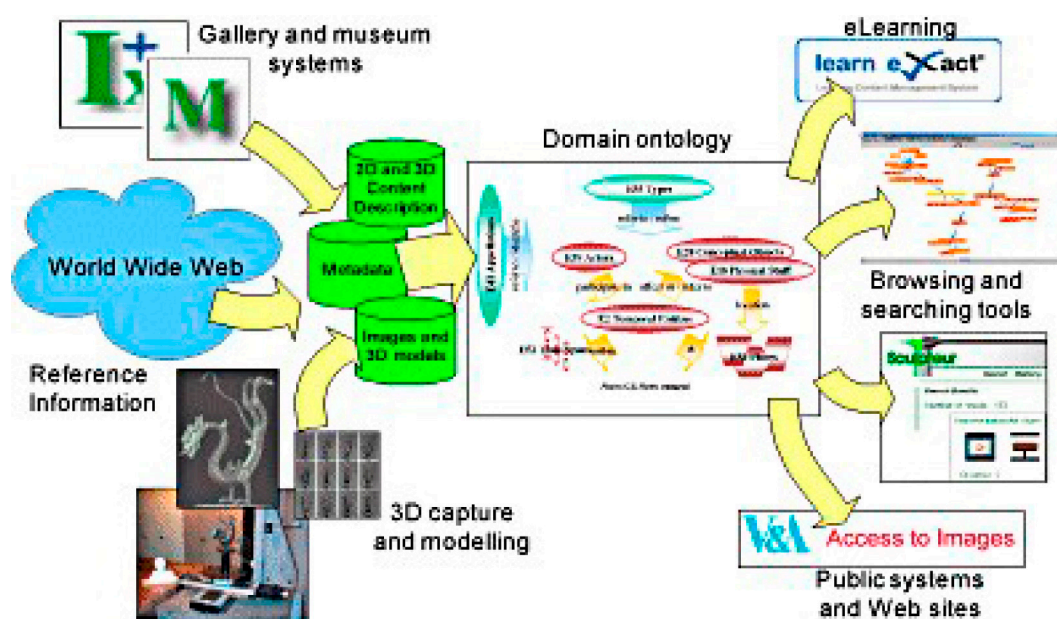


Figure 2. Model of an integrated content-based and concept-based system of information, set during the project Sculpteur promoted by the European Union.

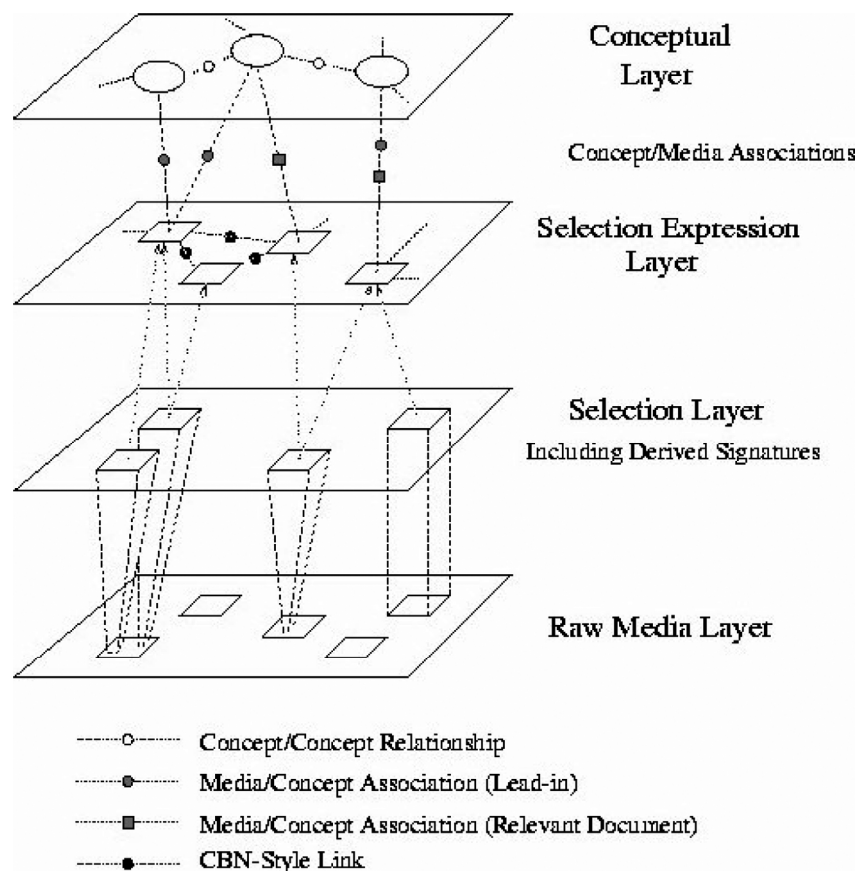


Figure 3. Architecture of the MAVIS 2 system.

ture of MAVIS 2 is that of a sort of “multimedia thesaurus” (MMT). If in a traditional thesaurus different terminological representations of the same concept are associated among them, in MMT, they are the different multimedia representations of a concept to be associated with it and its name. The MMT can be used as an articulated, multilevel structure to organize, in the system, the different informative data on the multimedia documents (Dobie et al. 1999).

5.0 The new semantic tools

Some other necessary considerations concern one of the fundamental principles of MMIR: the imagination and the creativeness as a style of the method to conduct searches of information and documents, including their organicity and univocality. Accepting the integration of the ontologies in the systems of MMIR, in these innovative conceptual tools, it seems to be residual, nevertheless, a certain rigor, which can propose again the problem of the rigidity and the abstractness of the typical schemes of the IR. To avoid such risk, a further hypothesis can be made about

combining the ontologies with folksonomies, as systems of free collaborative categorization of the contents on the base of labels directly assigned by end users. This direction is given to a discussion started by the same founders of the semantic web and of the structures that drive and organize it.

First of all, with respect to the ontologies, the conceptions of Tim Berners-Lee and his collaborators (2001) are very relevant. They propose an updated vision of the semantic web and of the new possible relationship between the semantics controlled by the top and those produced by the bottom. The reasons for the folksonomies, but also the actual inadequacy, are introduced and discussed, among the others, by Marieke Guy and Emma Tonkin (2006), who try to show what space folksonomies can have in the semantic web, and what value social tagging has for the production of valid free semantic structures.

To conclude the matter, in the search of a mediation useful to overcome the semantic gap in the respect of the operational flexibility of the MMIR, there is the argument discussed by Gino Roncaglia (2006) in a paper prepared for a conference on digital

libraries. Specifying the collaborative and social aspects of the creative process of the uncontrolled semantics allows the overcoming of some of their typical problems, proposing their function close to the controlled semantics and enriching them with flexibility and diffusion. The most effective systems founded on folksonomies, in fact, exploit the social and free nature of the process of classification to improve it and to enrich its semantic quality, by tools what the collaborative filtering.

Everything is alongside the principles of MMIR systems, where the possibility of the user searching freely through models or personal sketches allows the system to learn instantaneously new information about the documents, which will be stored together with the information already defined, integrating, and widening its interpretive abilities. The integration between the semantic tools of the ontologies and the folksonomies, contemporarily integrated with the content-based tools, can bring about the conciliation of the opposition between the principles of the semantic-interpretative and content-based-objective handling. In fact, it is hopeless to pretend to eliminate the addition of human interpretation to the process of the search, founded on a volume of preceding knowledge and on a refinement of elaboration of contents that are impossible to be achieved by computers.

6.0 The solution

The large problem area released by possible innovations of MMIR can come together in a conclusive manner, which contemporarily represents taking the conscious limits of the content-based system and the proposal of overcoming of those limits, contained since the beginning in the founding principles of the MMIR. If there cannot be a lasting solution for the semantic gap, for the contradictions and the gaps of the relationship between the proper cognitive and intellectual demands of the human and the mathematical and mechanical responses of the automatic systems, it is possible at least to define a perspective of collaboration between the researcher of information and the tools of analysis and searching.

The solution of the confrontation between conceptual accesses to information and concrete access, or of the confrontation between term-based and content-based systems of processing, can only be a solution of organic integration among the principles and the methodologies of analysis, storage, search, and retrieval of information and documents that constitute

the only apparently incompatible semantic and content-based ambits. MMIR is created and theorized on this theoretical and practical conclusion, obvious but often not attainable. As in the daily, natural, and uneasy activity of cognitive searching, the human is suited to freely and univocally make use of his own sensitive and intellectual attitudes; in addition, his more advanced information search tools have to sustain and to help him to maximize his own abilities, also allowing him to resort to the resources of imagination and creativeness, allowing the organic integration of the content-based and semantic means.

In the consideration of the semantic limits of the content-based system, an opportune intellectual intervention in the organization and in the search of the documents is often necessary, to define the meanings besides the feelings, to specify the query strategy, and to increase the possibilities of retrieval. This necessity to integrate content-based and semantic approaches to the documents then has to bring to the definition of a univocal system of handling of the multimedia information, able to contemporarily consider the necessities of information of intellectual-interpretive and content-based objective character.

All the procedures can operate in constant and organic interaction, in a single system and with a single search interface, in the composition of a query formula, which combining terms, images and sounds according to the cases, can serve for searching very complex documents, whose informative content extends toward all the levels of sense and meaning. Such structuring of the query allows users to employ conceptual or ontological schemes in a free and flexible way every time, as well as their own specialist or diffused, erudite or common attitudes to classify information, close to their own sensibility and to the tools to operate directly on concrete contents, in a context of resources of every kind, reduced to a specialist space or widened to a general dimension.

In addition to the classical interfaces of the text databases, there must be interfaces that allow formulating the query in different dimensions without excluding the importance that maintains the terminological, descriptive, or conceptual data, relative to those aspects not specifically in the content of the documents. The search will happen in indexes that are substantially different and richer than traditional ones, composed of words extracted from the entirety of a written document or from the speech of an audiovisual, of key images of a sequence, of geometric figures, of melodies, of forms, colours, movements, and sounds.

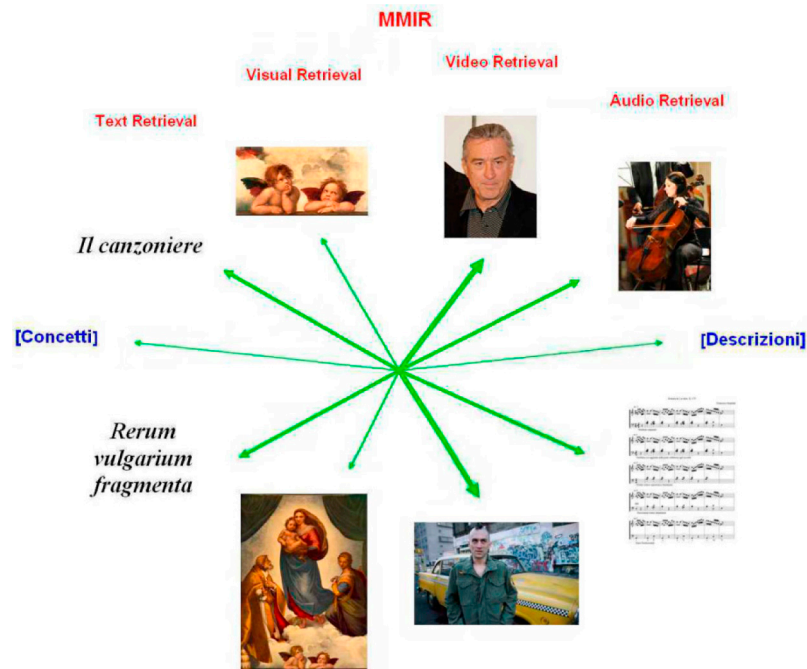


Figure 4. The organic system of MMIR.

7.0 Conclusions

MMIR is a revolutionary organic system, specialized in its parts for the effective handling of every kind of digital multimedia document. Examples can be found online using the URLs in Table 2.

AudioId.	business.muvin.com
Hermitage state museum– Digital collection	www.heritagemuseum.org/cgi-bin/db2www/qbicSearch.mac/qbic?selLang=English
Informedia	www.informedia.cs.cmu.edu
MediaMill	www.science.uva.nl/research/mediamill
MUVIS	muvis.cs.tut.fi
QBIC	www.qbic.almaden.ibm.com
QuickLook	projects.ivl.disco.unimib.it/quicklook
SoundFisher	www.soundfisher.com
Viper	viper.unige.ch/doku.php/home

Table 1. Online demos

The whole complex constitutes a new strategy of organization and management of information, which is more advanced than traditional methodologies based on the centrality of the conceptual and terminological human approach. This new strategy aims at resolving the matter of the search based on the strength and objective content of the information according to an automatic and content-based approach. It is neces-

sary, nevertheless, to foresee the integration of a conception of processing and searching of information that is revolutionarily content-based with a traditionally semantic conception.

Clarifying the sense of the cohesion and the internal coherence of the complexity of MMIR, it is essential to establish that a good level of effectiveness in the searching of documents is possible only by using a combination of methods and techniques founded both on the definition of the meanings, through controlled terms, and on the representation of the content, through textual, visual, audio, and video elements. The various content-based systems of TR, VR, VDR, and AR, in fact, must be considered in constant harmony, among them and with the semantic systems: the principle of the semantic consultation establishes a precise method in order to select a part of the great quantities of documents of a database in relation to thematic areas, to titles or to authors as well. It is also a criterion to limit the various specific ineffectiveness of a content-based consultation, specifying those degrees of semantic interpretation that the automatic system is not able to detect in the direct analysis of the content characteristics which are representative of the document.

Above all, however, the different procedures operate best in continuous interaction, solely in an organic search interface, allowing several strategies of consultation by combining words, figures, movements, sounds, and concepts that are useful for searching very complex documents, rich at all levels of sense and meaning, and that are able to overcome many risks of falling in the gap of an organization and a semantic only or content-based only handling of the documents.

References

- Addis, Matthew et al. 2005. New ways to search, navigate and use multimedia museum collections over the Web. In Jennifer Trant and David Bearman eds., *Museums and the Web 2005 proceedings*. Toronto: Archives & Museum Informatics, pp. 101-118.
- Berners-Lee, Tim et al. 2001. The semantic Web: a new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. *Scientific American* 284:5, 28-37.
- Dobie, Mark et al. 1999. MAVIS 2: a new approach to content and concept based navigation. In *Proceedings of the IEEE Colloquium on Multimedia Databases and MPEG-7*. New York: IEEE, pp. 9/1-9/5.
- Enser, Peter G. B. 2008. Visual image retrieval. In Cronin, Blaise, ed., *Annual review of information science and technology* 42: 3-42. Medford, N.J.: Information Today.
- Guy, Marieke and Tonkin, Emma. 2006. Folksonomies: tidying up tags? *D-lib magazine* 12 no. 1 <http://www.dlib.org/dlib/january06/guy/01guy.html>.
- Hare, Jonathon S., Lewis, P.H., Enser, P.G.B. and Sandom, C.J. 2006. Mind the gap: another look at the problem of the semantic gap in image retrieval. In Chang, Ei Y. et al. eds. *Proceedings of multimedia content analysis, management and retrieval 2006*. San Jose: SPIE, pp. 75-86.
- Roncaglia, Gino. 2006. Folksonomies e semantiche controllate. In *Interoperabilità delle biblioteche digitali*. Roma, <host.uniroma3.it/biblioteche/download/Roncaglia_Folksonomies.pdf>.
- Smeulders, Arnold W.M., Worring, Marcel, Santini, Simone, Gupta, Amarnath and Jain, Ramesh. 2000. Content-based image retrieval at the end of the early years. *IEEE Transactions on pattern analysis and machine intelligence* 22 no.12: 1349-80.