Understanding the categories and dynamics of multimedia information: a model for analysing multimedia information

Abstract: A model for analysing multimedia information is proposed from the point of view of the theory of communication. After a brief presentation of the complex map of the sciences that deal with multimedia communication in its different aspects, the current multimedia revolution is historically contextualized as a tendency towards messages that are able to build near-reality experiences (virtual reality). After setting the theoretical point of view, an analysis of multimedia messages is substantiated and a model is presented. The first part of the model deals with the different communications channels and tools: still images, movies, sounds, texts, text with illustrations, audiovisuals and interactive multimedia, with an emphasis in non-textual documents. The second part addresses the global properties of the multimedia message, which are of a textual and metatextual nature. The overlapping of media, channels, genres and messages—and the conscious and technical use of such interactions—is precisely one of the main and outstanding characteristics of the multimedia discourse, and requires specific moves in indexing languages development. The multimedia environment has also a great potential to promote a wider theory of knowledge organization, bringing closer distant fields like scientific and fictional indexing or verbal and image indexing. It is stated that such a unified theory requires a closer attention to the pragmatic aspects of indexing and the inclusion of new semantic layers. A simple indexing model is proposed to illustrate who to address these challenges.

1. Introduction (1)

The aim of this paper is to propose a model for analysing multimedia information building on the point of view of the theory of communication (2). Such a quest must begin with an acknowledgment of the complexity of the field, as the landscape of the disciplines of communication is impressive. As an illustration of this complexity, Figure 1 provides a sketch of only the twenty-two related to the study of signs and their use, that is, to Semiotics.
Another important premise for such a model is providing a working definition of multimedia communication—which can be described as communication that is achieved through more than one channel or medium—and overall a precision about its implications. In fact, a multimedia nature is inherent to natural communication, where the agents, being both present, convey information throughout different visual and auditory clues. As presence is a very limiting characteristic of natural communication, technologies have been invented to allow for communication in absence of the sender—this is, for example, by using letters. But, because of economical and technical reasons and with very rare exceptions, these technologies use a single channel and genre. This made mediated communication abstract and even difficult when both agents do not shared a very detailed knowledge of the frame of reference of the messages. So, from an evolutive or historical perspective, computer mediated multimedia communication and the current multimedia revolution can be considered as a tendency towards creating asynchronous messages capable of inciting in the receiver almost-natural experiences, and, eventually, of arousing in him near-reality experiences (virtual reality). Such a natural experience requires multimedia, but also a truly and synchronous communicative experience between the sender and the receiver, that only current interactive information technology can provide (Fig. 2).

Fig. 2. An evolutionist perspective of multimedia communication: the path towards an induced “natural” experience
In conclusion, messages—and therefore documents—are considered an effort to create a sensorial and intellectual world in the mind of the receiver. The current stage of this evolution is considered to be the integration of the different information channels and resources to build an overwhelming sensorial and intellectual experience. The advances in this direction can be best appreciated in some environments like some scientific applications or even state-of-the-art computer games. Information and document centres have followed the path of the available document technologies creating specialized centres for each kind of media that has been appearing—video, maps, drawings, photos, video libraries and archives—, and now follow the path toward integration and convergence.

So, an integral approach toward documents is necessary, in which multimedia is not considered something apart, but the very nature of communication. From this perspective, written, visual and sound documents are granted the status of partial and uncompleted approaches toward the aim of a global communication experience that can be reconstructed by imagination and that, currently, can be induced through the senses with better sensorial experiences that are dependent on the technology.

2. Analysing multimedia messages: toward a model

But multimedia, though a very powerful tool for communication, implies big challenges for indexing and, therefore, for its previous phase, content analysis. Such challenges come from two different sources: on the one hand, multimedia is constituted by the aggregation of different items which are expressed in different languages that must be understood on their own terms; on the other hand, the relation between these channels for meaning transfer is not univocal, but emergent, dialectic and systemic. As a result, decoding and making explicit the information conveyed in a multimedia message is a very complex process that requires enhanced analysis tools and models.

The outline of such a model is presented below. The model is divided in two main parts, according to the above-mentioned challenges. The first one is devoted to the component channels that constitute a multimedia message; and the second, to the global systemic qualities of the message that govern and integrate the whole meaning and intention of the communication act. The model is structured in five levels of complexity:

1. Component elements
   a. Image
      i. Static elements
         1. Light environment
         2. Composition
         3. Spaces
            a. Natural spaces
            b. Urban spaces
            c. Relational spaces
         4. Objects
         5. Human images
            a. Physiognomy
            b. Clothing and garments
      ii. Dynamic elements
         1. Natural processes
         2. Human activities
            a. Expressive activities
            b. Instrumental activities
         3. Objects behaviour
b. Sound
   i. Environmental sound
      1. Natural sound
      2. Environmental music
   ii. Human voice
      1. Basic properties
         a. Intensity
         b. Pitch
         c. Timbre
         d. Duration
      2. Processes [Paralinguistic]
         a. Emotionality
         b. Relationship status
         c. Interaction modalities
         d. Emphasis
         e. Non verbal vocal designation
   iii. Music

c. Verbal
   i. Languages
   ii. Idiolects
   iii. Acts of speech
   iv. Syntax
   v. Lexical level [including lexical morphological aspects]
      1. Abstractions
      2. Generic concepts
      3. Specific concepts
      4. Exemplars
   vi. Phonetics

d. Graphics [artificial languages, not images] [see verbal]

e. Interactions
   i. Navigation
   ii. Retrieval
   iii. Transformations
      1. Layout
      2. Calculus

2. Systemic aspects
   a. Global qualities
      i. Metatextual level
         1. Intentionality
         2. Relation with reality
            a. Factuality
            b. Credibility
         3. Sociocultural frames
         4. Actants [Roles]
         5. Spatiotemporal frame
         6. Channel/medium
         7. Script
      ii. Textual level
         1. Genre
            a. Subgenre
            b. Document type
         2. Subject
            a. Referential
b. Relational
3. Rhema [new information]
4. Argument
5. Rhetorical and communicational paradigms and methodologies

b. Interaction processes among the different channels
   i. Basic interaction
      1. Concentration
      2. Complementarity
      3. Addition
      4. Contradiction
   ii. Temporality
      1. Synchrony/Asynchrony
      2. Isolated presentation/Continuity/Recursivity

It must be stressed that the model can work as a universal one—that is, that can be applied to a specific media like photographs, texts, drawings, etc.—, as long as only the component elements of the respective media are considered. The global characteristics of a message apply to any kind of document, and also, if the document has distinct parts, to each of these subsystems—as paragraphs, chapters, illustrations, videos, interactive commands, etc.—. In fact, the analysis can be done in a recursive manner depending on the number of levels or layers in which the “text” (3) is organized.

In this later case, it seems useful to remark that the different levels and subsystems are always organized towards the same aim, even when it does not seem so, because of different reasons, be it that the author wants to leave meanings open, show confusion or ambivalence, or is unskilled to transmit the message. It can be also useful to share the experience that, for the end-purpose of indexing a multimedia document, it is seldom necessarily—but for very complex materials like, for example, textual or visual poetry—to repeat the analysis recursively at every level of semiosis. On the contrary, it is usually enough to analyse only key elements at the different stages of the pyramid of meaning. And as a general rule, lower level characteristics will be considered only if they separate from the generic prototype they belong to; otherwise they must be considered efficiently summarized at the upper level.

In the two following sections, the general landscape of both the component elements and the global properties of messages will be presented, and some selected key characteristics will be discussed.

3. Analysing the channels and their components

The basic channels of communication are well-known: still images, movies, sounds—verbal and nonverbal—, texts, text with illustrations, audiovisuals, graphic artificial languages and interactive multimedia.

Regarding images, the elements to be analysed can be classified into static and dynamic ones. The static elements considered are the environment, composition, spaces—natural, urban and relational ones—, objects and the structural elements of the human image—physiognomy and clothing—. The dynamic elements are the processes going on—natural, human and artificial ones—and the systemic properties of their interaction. Still images have also, of course, dynamic elements, though they are presented as an instant or a potential, not as fully developed movement. Anyway, this kind of summarization is inherent to visual documents, since telling histories, picturing, drawing or filming is always about selecting, abstracting and suggesting, as usually a large period of time must be told in a very short one.
On its part, sound is analysed in three groups—environmental sound, human voice, music—and their interactions. Environmental sound can be divided into unintended sounds—“natural” sounds—and background music, which is nearer to the third category. Human voice can be analyzed according to its basic qualities—tone, intensity, timbre, etc.—and some basic para- or pre-linguistic process—that denote emotionality, relational tags, interaction modalities, emphasis and vocal designation—.

Verbal behaviour—be it expressed by sounds or written—can be analysed in different levels: the languages selected, the peculiarities of language use—idolects and acts of speech—, syntax, lexical level—abstractions, generic concepts, specific concepts and exemplars—and even phonetics. Verbal behaviour in general lacks the immediate emotional appeal of images and sounds, but has a very strong relational power and allows deliberate ‘mind computing’ in the short-term memory. Verbal behaviour has evolved into artificial and programming languages that are the background of interaction, one of the main qualities of multimedia.

Besides verbal behaviour, a group of languages exists that are also conventional but use graphic symbols instead of words. This world of iconic messages—some of them very elaborated, as occurs in Cartography, Industrial Design, Geometry or Statistics—keeps growing in size and importance. Another kind of graphic languages of great importance is graphic publishing languages—layout techniques and typefaces—with are an important way of conveying the meaning that is usually transmitted with paralinguistic and other non verbal resources in the verbal world.

4. The systemic and emergent aspects of multimedia messages

The different analytical components of the multimedia message form together an emerging unit of meaning and communications. This unit—the message—can be characterized by certain distinct global properties that can be classified into two levels: metatextual and textual.

On the one hand, metatextual qualities are of a social nature, and express the sociocultural context of the message: the actants, the intentionality, the social and spatiotemporal frames and scripts, and the communication channel that are acting in the whole situation. On the other hand, textual properties are of a semiotic nature, and can be organized in its classic subdisciplines: pragmatic—genre—, semiotic—subject-rhema—and syntactic—argument, rhetorics, etc.—. Finally, the kind of connection among the different channels of the multimedia message must also be considered, since, using a mathematical metaphor, the different components can be added and substracted in different ways to modulate the final meaning.

The systemic and emerging meaning of multimedia messages has been partially the centre of traditional content and document analysis (Table I).

<table>
<thead>
<tr>
<th>Relative Pronouns</th>
<th>Explanation</th>
<th>Ranganathan facets</th>
<th>Document analysis vocabulary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Who</td>
<td>Subjects</td>
<td>Personality</td>
<td>Agents</td>
</tr>
<tr>
<td>What</td>
<td>Domain, Objects, Processes</td>
<td>Matter</td>
<td>Subjects</td>
</tr>
<tr>
<td>How</td>
<td>Modalities</td>
<td>Energy</td>
<td>Activities</td>
</tr>
<tr>
<td>Where,</td>
<td>Circumstances</td>
<td>Space</td>
<td>Space</td>
</tr>
<tr>
<td>When</td>
<td></td>
<td>Time</td>
<td>Time</td>
</tr>
</tbody>
</table>

Table I. Semantic elements of a message in content and document analysis
But a theory of communication that goes beyond scientific and technical documents—which are usually of a controlled pragmatic and semantic nature—gets necessarily much more intricate. When analyzed, many multimedia messages emerge as communication “icebergs”, where very few is actually told, very much is simply intended or tacitly given as known, sense is conveyed into different kind of meanings—defined, denoted, connoted and designed—and even organized in a hierarchy of levels of signification and abstraction, and where the very pragmatics of the process remain obscure even for their agents.

The complexity affects also the agents—since persons can act in such different roles as references, channels, actual and potential targets, etc.—and their intentions. After all, communication is not only a way of transferring true knowledge, but also a technique to influence the receiver. In fact, intentionality is a multilevel reality: There is a predicated aim, possibly a shared one, always a subjective intention, and, finally, a meta-intentionality that is given by the social and cultural system and is not always perceived by the agents.

In any case, an integrated theory of knowledge organization cannot be achieved but by recognising and using the systemic properties of any message, which allow the indexer and, in general, the content analyser to treat it as a unit. In this sense, messages must be considered systems, this is, a set of elements in dynamic interaction organized for a goal (Rosnay, 1979).

5. The multi-channel nature of multimedia information and its implications for Knowledge Organization

It has been discussed that the overlapping of media, channels, genres and messages—and the conscious and technical use of such interactions—is precisely one of the main and outstanding characteristics of the multimedia discourse.

This multi-channel nature of multimedia documents has an important theoretical implication for knowledge organization: As a multimedia indexing theory must take in account all the different channels, it can and must, therefore, explain the indexing of each kind of material and thus provide the frame for a global theory of knowledge organization. This would allow overcoming some of the big barriers that prevail nowadays in indexing and classification theory, mainly the ones that keep apart scientific and technical documents from fiction, and graphic materials from textual ones.

A second implication of the multimedia revolution is the need to put order in the world of the indexing terms for media, audiences and genres. This has become an imperative task in the new multimedia context, because in such an environment the same information circulates through different channels and in different forms, intending different audiences, and this situation should be controlled. The very concept of channel or medium has also got very complex, as now diverse media is involved in the different stages of creation, distribution, storage and reproduction of documents. Something similar occurs with genres.

Trying a translation of this approach to a theoretical point of view, it could be stated that indexing must incorporate more clues about the pragmatics of communication, since, in the new information society, information is no more guided to target audiences by quasi-unidirectional channels—like teaching material trough classes or entertainment through theatres and cinemas—, but through a net of interconnected media. In this context, media, audience and genre are no more implicit characteristics of a communication process, but facets that must be predicated in a specific way with the help of proper knowledge organization tools.

Because of their universality and specificity, it seems a good goal to develop specific languages for media, audiences and genres, as it has been done for toponyms, entities or personal names. But this approach should therefore be completed with relations to other
more general subject thesauri and classifications. In any case, the constitution of completely separated tools should be avoided, because such a situation would provoke the emergence of unintended barriers for the semantic navigation among related concepts, as, ultimately, all those different vocabularies refer to integrated domains of knowledge that must be interrelated.

6. The limits of current semantic indexing: some strategies for action

Another question that must be addressed is that traditional indexing rarely takes into account the whole set of characteristics that have been presented above as systemic aspects of the message. On the contrary, other pragmatic aspects than the context of study—disciplines—are not considered, and, on the other hand, the subjacent semantic model that supports content analysis is usually very positivistic and is unable to deal with the loose of fuzzy meanings that multimedia messages frequently convey.

Regarding the pragmatics deficit, normal indexing theory seems fixed on the semantics of messages—that is, on subjects such as process, agents, patients, tools, places, etc.—, but, though the global properties of a message are undoubtedly of a semantic nature, they are mainly of a pragmatic one. As Van Dijk (1977) states: “Discourse coherence is not primarily a matter of meaning, but of reference”. In someway, pure semantic indexing was possible when such pragmatic aspects were unambiguously implicit in the knowledge organization tools because only very specific literature was subjected to indexing, mainly scientific or technical. But the media revolution and its cultural and industrial implications need a new approach toward indexing. In fact, a more complex society requires more tips about the relations of messages with their users. For example, even in the scientific realm, interdisciplinary research has put in question the implicit relation between subjects and disciplines.

Descending to a more practical point of view, an alternative approach (see table II below) would require indexing not only the subject of the message, but aspects like

- who creates the message (4),
- who can use it,
- the spatial, temporal, social and scientific-technical contexts of its production and reception,
- the instrumental aspects of the communication—media channels, genres, textual structures—,
- and their intentions and interpretation.

This approach has also a key practical implication for current technological and social developments: it can help to connect in different ways users and information objects in automatic ways, promoting the implementation of the gateways that are implied in the semantic web approach (5).

On the other hand, semantic indexing is also receiving big challenges from inside. With the multimedia revolution, a growing percentage of content is leaving the relatively unambiguous path of scientific and journalistic genres and approaching the semantic jungles of creative literature. This kind of material requires the complex kind of content analysis that is characteristic of artistic disciplines, that can be summarized in the simplified model that was proposed by the Art historian Panofsky (1955)—also applicable analogically to the analysis of fiction—: description of common subjects; identification of the specific names and exemplars of these common subjects; and interpretation of the abstracts subjects that are being represented through the former.
All these questions are taken into account in table II, which shows an outline of an extended indexing model that considers the semantic and pragmatic complexities of multimedia indexing, as opposed to the one shown in table I. Here the scope of the traditional pronominal terms changes to denote a pragmatic approach instead of the usual semantic one, which is denoted by the interrogative “what”. The subject domain (“what”) is divided also into three realms: common names, exemplars and abstract concepts, which is always necessary for dealing with non-verbal materials.

<table>
<thead>
<tr>
<th>Relative Pronouns</th>
<th>Explanation</th>
<th>Proposed term</th>
</tr>
</thead>
<tbody>
<tr>
<td>What (1): common</td>
<td>Domain, Objects; Processes, Tools, Actants, Space, Times</td>
<td>Concrete subjects</td>
</tr>
<tr>
<td>What (2): exemplars</td>
<td>Subjects</td>
<td>Exemplars’ subjects</td>
</tr>
<tr>
<td>What (3): abstract</td>
<td>Perspective</td>
<td>Abstract subjects</td>
</tr>
<tr>
<td>Who</td>
<td>Potential and actual users</td>
<td>Producers</td>
</tr>
<tr>
<td>To whom</td>
<td>Audiences</td>
<td></td>
</tr>
<tr>
<td>From where</td>
<td>Disciplines</td>
<td></td>
</tr>
<tr>
<td>To where</td>
<td>Paradigms, -ism</td>
<td></td>
</tr>
<tr>
<td>How</td>
<td>Modalities</td>
<td>Media, Genres, Textual structures</td>
</tr>
<tr>
<td>Where, When</td>
<td>Circumstances</td>
<td>Scenery: Spaces, Times</td>
</tr>
<tr>
<td>For what</td>
<td>Aims</td>
<td>Intentionalites</td>
</tr>
<tr>
<td>Why</td>
<td>Causes</td>
<td>Interpretations</td>
</tr>
</tbody>
</table>

Table II. An extended model of indexing for representing pragmatic and multi-level semantic characteristics

Notes
(1) This paper builds on the previous work done by the author for a collective book written with his colleagues Maria Pinto Molina and Carmen Agustin Lacruz (2002). Though he is the only responsible for his ideas, he is in doubt with them for the discussions and invaluable suggestions and feedback.
(2) We are referring here to the theory of “natural” communication as opposed to “artificial” one, which, apart from the splendid communication model (Shannon & Weaver, 1949), deals mainly with the transfer of signals. “Natural” communication is concerned with effective and efficient signal transfer, but also with the pragmatic, semantic and syntactic aspects of the phenomenon.
(3) The use of the term “text” for referring to multimedia documents must be understood in a very wide sense: Text is defined here as the documented form of a discourse; and discourse, as an architecture of symbols—organized in different structures—forming a message to support the communication between persons—or their artificial surrogates—.
(4) In fact, this is possible in cataloguing through the use of authority records.
(5) It is also very important to call the attention upon the modular nature of the multimedia message. This kind of information carrier can be disaggregated and its parts reused and re-oriented. This is more and more a strategy in industry and also in pop multimedia culture. Though this reality has mainly direct implications for multilevel item control and cataloguing, it is also important for indexing and classification, as it requires that certain materials that traditionally have not been indexed—because they were expected to be retrieved in the upper context or were considered raw material, like book illustrations—will have to be processed.
References