From Images To Images: the ALADIN project

Katell Briatte and Marie Despres-Lonnet, Université Charles de Gaulle, France

Abstract

Scholars in the field of historical research have used iconographic databases for a long time. The rise of the Internet led numerous libraries and research center to envisage the sharing of this scientific patrimony with the general public. But such a popularization project cannot be undertaken without reassessing the initial orientation of the database. This communication presents the results of tests carried out as part of a digitization project undertaken by a team of librarians and historians. After an in-depth analysis of the process of image reading, we suggest new ways to access and index images. First, we propose an overview interface to allow for a constant dialog between the user and the archive; secondly, we have recourse to the concept of "scenotype" to better understand image reading and to conceive an innovative indexing method.

Introduction

A few years ago, three libraries in the North of France undertook the digitization of their image collections. The initial purpose of the LIB.R.I.S (LIBraries for the Regional Information Society) project was both to preserve and extend access to the region's iconographic heritage (Després-Lonnet, Aubry, & Briatte,1998). The collections contain mainly geographical maps, original manuscripts, postcards and photographs. The database was primarily designed to meet scholars and students needs, but, considering that it covered a field of great interest for the general public, the project partners decided to offer access to this archive to a broader audience through the Internet.

Our lab (CERSATES UMR CNRS 8529) is presently involved in a research program aiming at studying how new media improve knowledge dissemination outside the institution that produced it and how this knowledge is received and assimilated by a non-specialist audience. For this reason, the LIB.R.I.S. project managers asked us to study how access to such a database by new users could be improved. This has been the ALADIN (Accès en Ligne à une Archive de Documents Iconographiques sur le Nord) project starting point.

Our first concern was to try to understand which were the major changes induced by this new goal. An in-depth analysis of the content and structure of the database as well as user studies showed that what was at stake no longer was to adapt a recording and searching tool, but to consider the database as a scientific resource to be mediated through the Internet. Henceforth we had to see our project as an "editorial work" including a popularization dimension.

First, in order to improve the visibility and accessibility of the archive, we have developed a graphic interface based on common-sense metaphors. The usability tests we have then conducted led us to subsequently call into question the structure of the notes and the methods used to describe the images.

Improving access to a scientific resource

Comparative studies carried out on various iconographic databases reveal that the underlying knowledge organization of their designer shows through at each stage of the development.

A database is a scientific tool

First of all, relevant documents are chosen depending on the scope of the digitization project. Priority is given to images considered as more "important" from a specific viewpoint such as their topic, preciousness, state of deterioration, etc. This cannot be done without an in-depth knowledge of the domain and gives an additional consistency to the collection considered as a whole.

On the other hand, images are seen, analyzed and described following the ways and customs of a discipline (Gervereau,1997). Historians attach great importance to datation and localization of events. Art historians are interested in schools, artistic trends, formal criteria, etc. Librarians give priority to bibliographical standards (MARC, etc.). Conservators want to follow the trace of an art work (place and time of creation, purchase date, owners, exhibitions, etc.).

Lastly, choosing and using the right terms to describe an image is part of the scientific work (Wiberley,1988). Structuring these terms in a thesaurus is a process that can be compared to building a terminological network (Gardin,1979), which can be done by no one but the specialists of the domain.

In the LIB.R.I.S project, all the documents stored in the database have been chosen for their scientific relevance to the topic, that is the history of Northern Europe. The oldest documents are the first to have been digitized. Data collecting and analysis obey the scientific rules of the historical research. Historians vouch for documents authenticity (e.g. any image retouch is excluded). They verify any piece of data which allows to put the document back in its historical context and add the exact references which are the touchstones of historians' work, (i.e. date, location, size, conservation site, legend or brief analysis, etc.). Documents are described from an historical point of view, comments and links to the local history are added to each record (Després-Lonnet & Briatte,1998). Such a process gives the resulting database the status of a scientific tool. Thus, although each image taken individually would be easily comprehensible by a non specialist user as it deals with everyday life at the beginning of the century, its belonging to a database built by scholars in the context of a research program somehow conceals it behind a scientific barrier which makes its access by the general public more difficult.

The form interface

The way the users reach the images reveals the presence of a second filter: the screen interface. As a matter of fact, it is the only "contact point" between the user and the content of the collection and only through this "meta-document" can the users shape their representation of the archive.

Most Information Retrieval interfaces currently available on the Web use the "form" model. The form comprises one to a dozen input fields preceded by labels that are generally identical to the names of the fields in the database. The availability of a form "generator" as part of most DBMS tools proves that computer scientists think that the form interface is an "intuitive" way to present the retrieval task to the user, as it is the exact rendering of the database internal structure.

But querying a database through such an interface requires specific abilities. First of all, technical skills are needed to deal with the software itself. Then, only with a good understanding of information retrieval techniques and tools, such as Boolean operators, truncation, indexes or thesauruses can one carry out successful queries (Dickson,1984; Taylor,1984). A good knowledge of the content of the database, that is both the way it has been structured and the way documents have been described, significantly improves recall and precision rates (Bates, Wilde, & Siegfried,1996). Lastly, exact matching of character strings on which the form interface is based requires that the user knows and types the very same words or expressions as those chosen by the indexer to describe the image. The most familiar the user is to the indexers terminology, the better will be the results of the query.

Of course, the primarily targeted users do not possess all of the listed skills and knowledge, but they compensate the "weakness" of a given ability by a great mastering of others. For example, scholars develop "skirting strategies": they take advantage of their perfect command of the terminology used in the domain to substitute a term for another which did not reach a satisfactory precision rate. As for librarians or conservators, they see the form as a copy of the bibliographic or catalographic note they are familiar with.

We suggest that the main reason why the form interface seems well suited to the needs and habits of these groups of users is that they have a goal which complies with the way computer scientists consider the retrieval task: that is to search for a piece of data trough specific criteria. Now, general public has another project: they would like to "visit" the archive as if it was an exhibition (Brochu & Davallon,1998). But they have to browse through a set of objects gathered together by someone else from a particular point of view and according to criteria they are rarely aware of.

The overview interface

Our first objective was to endow the user with a kind of guide for visiting and discovering the archive. Our working hypothesis was that browsing and retrieval task would be facilitated if the nature and size of the collection as well as the criteria used to describe the documents could be known and easily understood by the users.

For that reason, we designed a dynamic graphic interface built on common sense-based metaphors. On the one hand, graphic representations are closer to the iconographic nature of the documents. On the other hand, we chose metaphors fit to express the characteristics (datation, localization, iconographic content) of each item considered for

itself while giving a more coherent view of the collection as a whole (covered period of time, geographical area, topics, ...).

Multiple, interconnected representations
Starting from the particular characteristics of the database related either to the nature of the collection, or to the historical approach that governed its conception, we have identified three dimensions (space, time, iconography) which characterize each document in the database.

The interface makes it possible to see three spaces of interaction at the same time. Each of which renders a specific dimension. Any selection operated in one of them modifies the representation of both others. For example, the selection of an interval of dates in the interaction space representing the temporal dimension induces a differentiation of the documents or concepts visible in the other spaces, according to whether they are associated or not with that period of time. Thus, at any time, the whole display is updated so as to reflect the user's choices (Ahlberg,1994).

Navigating in a network of concepts
First of all, it seems important that not only the documents, but also the indexing terms, are immediately accessible to the user. As an example, the LIB.R.I.S. database contains many architectural testimonies that are often described using technical terms generally unknown to non-specialists. From the iconographic thesaurus, we generate a 3-D navigation space (Robertson,1993) presenting the whole of the concepts used and the number of documents indexed by each term. Searching can then be undertaken not starting from supposed keywords, but from the indexing terms actually used (Shneiderman,1992).

In this way, not only are the users likely to find a word corresponding to their own set of vocabulary, but also they can immediately evaluate the amount of documents indexed by this term, and thus the relevance of their query to the collection. Moreover, they can appreciate the level of generality or specificity of each term, enrich their own perception of the semantic field and refine their knowledge of the vocabulary. The novice users are offered the opportunity to take over the whole of the specialized vocabulary and thus to imagine new, finer, more specialized or more generic requests. Thus the thesaurus becomes the very instrument of socialization.

What is the point of this approach?
Thanks to a "log file" implemented on the server, we have conducted a study on about 450 questions from people who queried the database through the Internet. Firstly, we notice that 33,4% of the queries gave no answer. Only 20,7% used Boolean or proximity operators. 55% of which gave no answer and 5% were misused. Only 3,6% of the users who typed a uniterm query had recourse to a joker or truncation character.

Most of the queries (>24%) were about precise geographical places. Users were often looking for documents about their own town or village. They generally did not go any further, did they get a satisfactory answer or not. By supplying these users with a global geographical map, we offer them the opportunity to situate their query in a more general framework, or even to widen their investigation to other places.

Although the date criterion is very rarely used (<3%), we do think that it is a basic one in an historical database. In our opinion, its limited utilization has to be put down to the fact that people either do not master the historical chronology, as it is often the case with children, or are not able to express their question in a correct Boolean query. Providing non-historian users with a tool which allows them to select either a precise date, a period of time or an event is likely to significantly enhance the quality of their queries (Bovey,1993).

Lastly, non-specialist users often utilize too narrow or too broad a term. For example, the words "North" or "architecture" are much too broad to be significant. Conversely, "dustbin" is too specific a word to intersect with the set of words likely to be used by the historians to describe all of the images where such an item can be seen. In our opinion, too general or too precise queries, could be avoided thanks to the possibility to see all of the words that have been used to describe the images, the number of corresponding documents and their immediate context. Henceforth, the user is no longer confronted with both usual problems of Information Retrieval: noise and silence.

Contrary to traditional documentary systems, in which the user cannot access interesting documents but through a request often built outside of the real indexing space, the overview interface puts him/her into immediate contact with the documents themselves, in the heart of the indexing space. Thus, the user is given a total and instantaneous visibility of the whole of the documents (Robertson,1993). He/She always perceives the content of the database, rather than a silent, anonymous form. Maintaining a constant contact between the users and the archive is likely

to have a strong psychological impact on people who are not familiar with searching tools or the historical discipline.

By externalizing the choices and practices of the scholars in a simple and clear way, graphic representations give the non-specialist user the opportunity to handle the experts' vocabulary and reduce the distance between historians and the general public by filling in the terminological gap. When having recourse to the historian's criteria (places, chronology, events, etc.) through easily assimilated metaphors, non-specialists improve their knowledge and skills by getting in touch with the expert's strategies and methodology. Lastly, by offering the user the possibility to intuitively carry out advanced searching thanks to a user-friendly interface, the graphic metaphor reduces the distance from the technical tool.

Rethinking description & indexing

Although the overview interface gives non-specialist users the opportunity to discover and utilize the experts' terminology, the way pictures have been described raises new issues.

Most of the time, historians put aside items they consider being "obvious" or "insignificant". For example, scholars make no mention of pilgrimage ensigns when describing a procession and when querying the database, specialists use the broader term "procession" to retrieve possible ensigns. In addition, the relevance of an item depends on the point of view adopted by the indexer. Conversely, historians often include in their description elements that cannot be seen on the picture, because they think them to be necessary to understand the depicted scene from an historical angle. In the LIB.R.I.S database, a note linked to a picture showing the ruins of a church contains the following description: "St Peter's Church: on this picture can still be seen the stele of The Virgin and Child statue sculpted by ...". In full text searching, a query about statues of the Virgin will propose this picture, on which there is no statue at all.

These basic remarks led us to reconsider the way images were described. We have primarily worked on the assumption that image description and interpretation had to be clearly dissociated. When taking into account the iconographic nature of the archive, we favor accessing the documents through their iconographic content rather than through their historical meaning. Users can become acquainted with historians' analysis after they have retrieved the images. Historical knowledge is not a prerequisite. In other words, describing a picture means listing iconographic items that are visible in it. The stance we adopted led us then to tackle the complex question of image indexing.

Image indexing: what is at stake?

As shown by A. J. Cawkell, two main options can nowadays be chosen in the field of image access, each of which is closely linked to a specific scientific community: image processing and image classification.

> "Processing people" work in the field of automatic pattern recognition and indexing. They try to conceive automatic processes likely to make out specific shapes, colors or textures, and then define indexing and comparison algorithms, from which to operate content-based retrieval. "Classification people" try to think of the best ways to describe images through words (Cawkell,1992).

Some researchers assure that both strategies will fail (Chang,1990), but we do think that their respective efficiency greatly depends on the pursued target and on the type of images to be retrieved.

Pattern recognition actually gives satisfactory results when meaning can easily be inferred from the formal characteristics of the images, which is not possible but in the scope of a particular project. For example, image processing is successfully used in numerous industrial or medical applications, such as weld resistance assessment by means of automatic picture comparison or x-rays photographs retrieval from medical databases according to form recognition criteria.

In the case of our historical archive, due to the heterogeneity of the documents and the complexity of their content, morphological aspects can hardly act as paradigms and thus serve as retrieval criteria.

However, by hand indexing raises other problems. It is well known that a human indexer can filter rather than help access to documents. The very nature of iconographic documents still reinforces this bias. Indeed describing still images by linguistic means compels to convert an analogical representation into an arbitrary code, which is part of a social, cultural, and idiosyncratic context.

Moreover, though the LIB.R.I.S. archive is used as a scientific tool by its conceivers, it also is a testimony to the personal past of many northern people, as a great many documents deal with everyday life at the beginning of the century. Therefore, people could be somehow dispossessed of their cultural heritage on the pretext that they do not share the knowledge of the database producers. Such a consequence is unacceptable and totally contradicts the initial aim of the project.

In this respect, we consider Pr. Brown and Hidderley's Democratic Indexing Project (Brown & Hidderley,1996) as the most promising research prospect. Indeed, democratic indexing gives anyone the opportunity to add "candidate" indexing terms to an image. These terms are added to the system thesaurus after the users of the database have voted them in. This seems a good way to restore users' control over the metadata, which describe and give access to their own past.

Nevertheless, we do think that the scientific value added by experts' descriptions has to be retained. In our opinion, democratic and scholars' indexing complement each other. Consequently, the challenge we are confronted with is to preserve specialized analysis and terminology in each domain while facilitating their usage by an inexperienced audience.

Experts must be the first to describe the images. But they must have at their disposal:

1. Indexing grids to facilitate their task, make various description methods consistent and guarantee, if not exhaustiveness of the description, at least an acceptable minimum of data for images to be retrieved.

2. An iconographic indexing assistant tool to guarantee shared language and points of view as a "common denominator" among indexers and between indexers and users.

In order to define the most appropriate tools, we undertook a precise analysis of how non-specialist users considered the iconographic documents stored in the archive.

The experimental method

Our aim was to finely evaluate the way different people "see" and "understand" the same set of images. It has been shown that asking people to describe images with words impacts the way they are interpreted (Laulan,1983). For this reason, we have defined a first protocol based on the method of "non-verbal classification" (Hudrisier,1982). This experiment has then been completed with an analysis of verbal descriptions.

The corpus

Themes relevant to the history of Northern France were selected with the help of historians from the CHRN. (e.g. mining, textile industry, metallurgy, seaside resorts, fishing, traditional games, carnival, wind mills, urbanism, agriculture, ...) and an experimental set of 160 printed images was extracted from the existing database so as to contain documents illustrating each of these themes.

The panel

The experimental panel has been formed so as to take into account the familiarity of a given user with images seen as scientific resources. Two main criteria have been defined: the foreseen use of the images and the understanding of the domain covered by the database.

We have thus distinguished between users having the same purpose as the conceivers of the database, that is to access scientific resources as part of a research project, and those wishing to access the documents with another goal (educational, journalistic, recreational, ...).

On the other hand, historians and other history specialists, such as teachers or local history interest groups, well versed in handling historical sources and having a thorough knowledge of the ways and customs of the discipline, were separated from non-specialist users. This led to a sample group, formed upon a bi-dimensional distinction: scholar/non-scholar historian/non-historian (Fig.1):

		Same problematic	
		yes	no
Same discipline or research field	yes	Scholars from the CHRN	High school history teachers, Local history interest groups
	no	Scholars from other disciplines	Primary school Teachers and pupils

Figure 1: the experimental panel

Each group was comprised of ten persons. Children from Group #4 were 11 years old pupils from regional schools.

Our experimentation panel was then asked to perform two kinds of tasks:

1. To sort the images into piles,

2. To describe each image

Image sorting
We assume that, when two images are put together in a same stack by all the members of a group, there is a broad consensus among this group as to their belonging to the same category.

Conversely, the "wandering" of an image from one stack to another means that this image is read differently by each member of the group and thus is more polysemous for that group.

For this first test the only instruction given to the test panel was to "group together images going well together", that is to sort the 160 images in categories that could contain from one to 160 images. So that children don't feel in an examination situation, we told them that they were not asked to find the right combination and that their work was not going to be evaluated by their teachers. No further questions were answered. In a second time, subjects were asked to label each pile.

The results were then recorded in a database and a co-occurrence analysis, based on the co-word method developed by Michel Callon (Callon,1993), was conducted for each separate group and for the whole of the panel.

The frequencies with which each pair of images occurs in the same batch was calculated and weighted by the occurrence of each image by itself. Images frequently associated in a same batch would then be strongly linked. If, on the contrary, a pair of images only rarely appeared in the same batch, they would only weakly be related.

The product of such a method was a large network depicting all the links that existed between the images. This huge network was then statistically split to highlight the sets of strongly linked images in the form of small clusters.

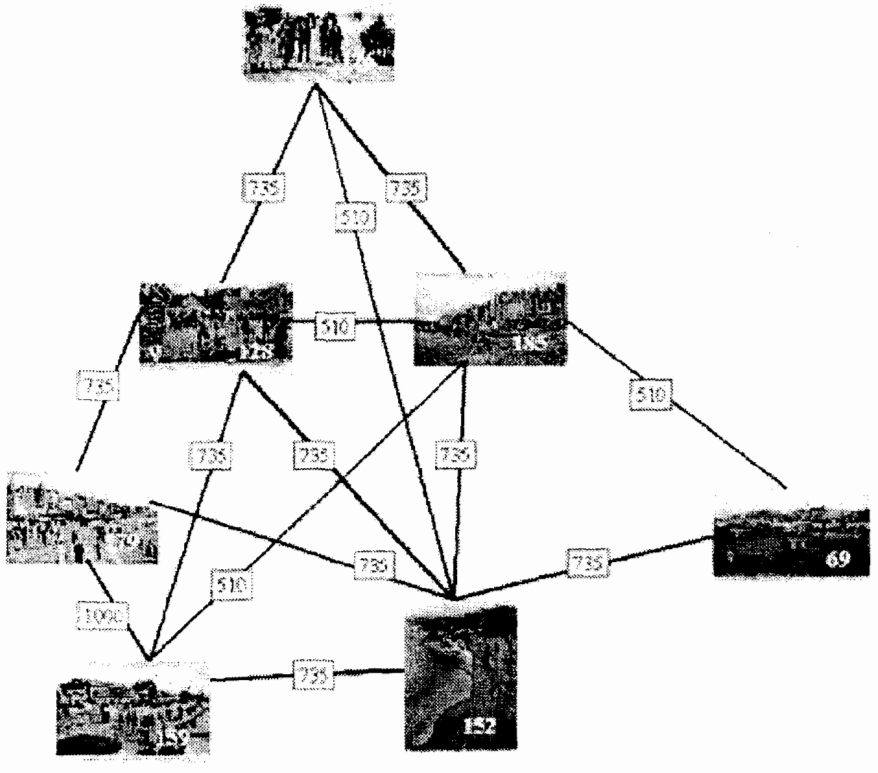

Figure 2: The «beach» cluster as processed from group #4's sorting

Obviously, the strength of the links between the images is not the same for all the clusters; some contain more images and may be more tightly bound together than others. This is referred to as the density of cohesiveness of a cluster.

Image description
For this test, subjects were asked:

- to add a caption to each image,
- to describe its content,
- to emphasize its essential elements.

The entire description was then stored as it was. We additionally defined normalization rules, so as to allow comparison between the descriptions: nouns were put into the singular and possibly accompanied by adjectives in brackets, verbs were entered in the infinitive, etc. We then classified the resulting terms in broad categories:

Object: chair, horse-drawn carriage, milk, ...

Building: beach hut, citadel, town hall, ...

Place: square, school, ditch, ...

Geographical zone: Lille, Normandy, Austria, ...

Activity or Scene: bathing, war, harvest, ...

People: boatman, mother, Gallic, ...

Animal: drafthorse, shellfish, herd ...

Abstract concept: sadness, friendship, power, ...

Type of shot: portrait, close-up, overall view, ...

Medium: photograph, poster, lithography, ...

Linguistic message: caption, words of a song, number, ...

Analysis

Consensual images
As resulting from our test, the more consensual set of images for all the subjects brings together pictures depicting schools and educational activities. The graph of this cluster is very dense, with very strong links between all the images inside it (Fig. 3).

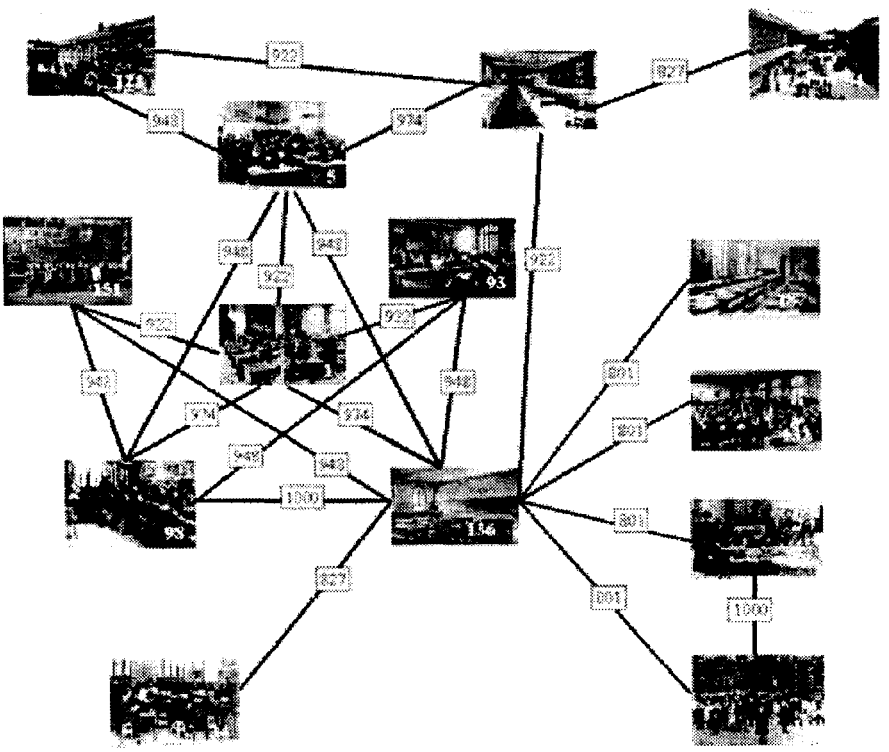

Figure 3: The «School» Cluster as processed from the whole panel's sorting

The image having the most and strongest links with others is a photograph of a classroom (image #136). Recognizable objects such as desks, a blackboard, large windows bring back the typically French "Jules Ferry" school. Although only few items can be enumerated, they are very typical.

A photograph depicting a drawing class (image #98) is very closely linked to this image. A group of children or teenagers are sitting behind easels and facing a bust. In the background, two adults are standing and watching the children. Decor includes a large window, a glass-fronted cupboard, and various models, such as pedestals or friezes, put on the cupboard or hanging from the wall.

While decor and objects were enough to identify a classroom in the first picture, in the second document, classification into the school stack mainly came from protagonists and their activities ("to draw", "to watch"). Nevertheless, items which were absent from the first photograph are strongly suggested by the number of objects (many desks) and their spatial organization (rows). Indeed, although actors themselves are absent from the picture, the notion of a community gathered in a room and sitting in front of a blackboard plays a fundamental part in the classification of this picture into the school stack.

"Nomadic" images

In contrast to this first category, some images are wandering from a pile to the other. They are often linked to different themes and are only rarely put twice together with the same others. An in-depth study led us to identify four main causes for such a "nomadism":

1. Atypical images

2. Multireferential images

3. Different points of view

4. Misreading

Atypical images

These are images that are singleton or quasi-singleton in the experimental set. For example, image #101 is an advertising poster for a brand of beer. Only two other posters can be found in the set of images (Fig.4).

One might have created a special pile so as to group posters together. But the elements to be seen in the other ones have led most people to classify them according to their content, a weaving loom

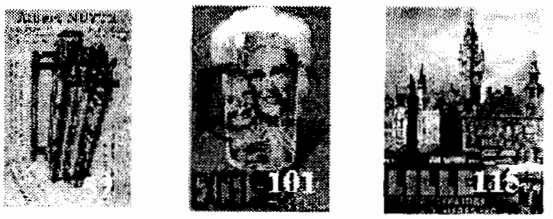

Figure 4: Posters to be found in the experimental database

in the first case and a view of the town of Lille in the second one, because many other images dealt with the same subject. The more abstract representation to be seen on the "beer advertising" poster made content associations more difficult. As a result, subjects "clutched" at its formal characteristics to put it apart or to create a category grouping images whose formal aspects prevailed on their content.

Multireferential images

Images #180 and #172 are quite similar: both scenes take place in the countryside. One can see cultivated fields on the foreground, and trees, a farm and a windmill on the middle distance. In both pictures someone on her knees plants or digs up something. (Fig. 5)

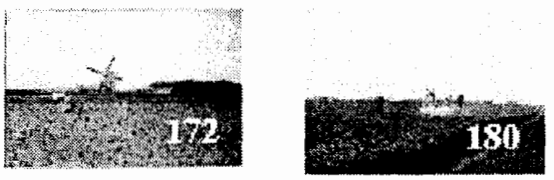

Figure 5: Multireferential images

But, surprisingly, these two images have not been put onto the same stack by half of the subjects. Image #180 was associated six times more often with pictures showing windmills than image #172. The latter is being almost systematically put into the "agriculture" theme. This probably comes from the fact that in the second case the windmill is slightly farther away and partly hidden by a plough while in the first picture its typically Flemish shape clearly stands out against the sky. The image is thus torn between the class of windmills and the class of agriculture.

Other images more clearly deal with two topics. For example image #108 shows a printing class and has thus been associated either with images #90 and #166 within a "printing industry" category or with other pictures of schools in an "education" category.

			
Printing industry	27.9%	27.9%	21%
Schools/Education	11.6%	**25.6%**	9.3%
Craft industry	53.5%	39.5%	60.4%
Others	7%	7%	9.3%

Figure 6: Themes associated with images depicting printing activities

Subjects having created a "printing industry" stack systematically put image #108 into it. Conversely, when images #90 and #166 were put into a more general topic covering all of the craft industry, image #108 was associated with schools.

Points of view

Subjects giving more importance to the function of the building have brought images of railway stations in a "transports" class together with images showing cars or streetcars. Other subjects put them together with other pictures of buildings when favoring the architectural point of view.

In the same way, image #9 (Fig. 8) is seen as a seaside view by 52.5% of the subjects while 35% consider it is related to pictures about working people and 12.5% give more importance to the presence of a Calvary and thus associate the image with other religious testimonies.

This disparity emphasizes the importance of the point of view from which each image is considered. Someone seeing a document (and whole of the collection) from a sociological viewpoint will focus on the protagonists and the depicted activities while someone else giving priority to history of sciences will be more interested in the tools, engines, etc.

The way views of the Town of Lille have been classified clearly exemplifies that fact. Only historians have created a pile labelled "Lille" giving priority to the geographical localization over the depicted scenes. In the same way Information science specialists have paid more attention to the formal aspects of the documents and created specific classes according to such criteria (ex: engravings, drawings, photographs, ...).

		
Architecture	36.4%	34.1%
Town views	29.5%	27.3%
Transports	18.2%	25.0%
Railway stations	15.9%	13.6%

Figure 7: Themes associated with images depicting railway stations

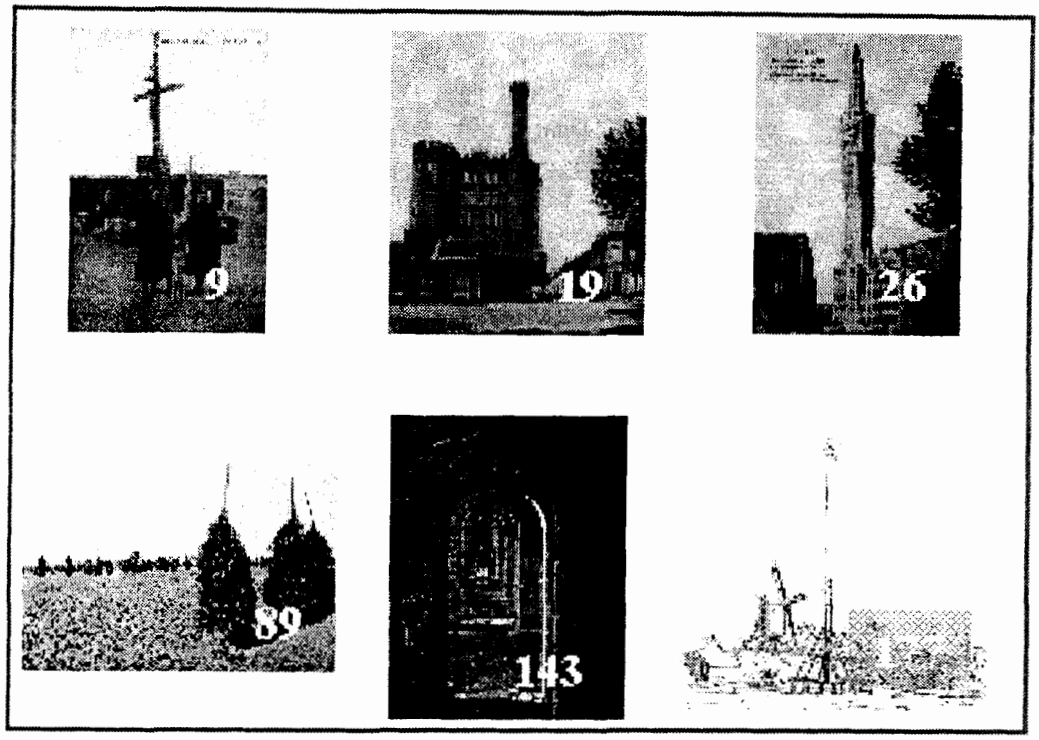

Figure 8: Images used to exemplify "nomadism"

Misreading

38% of the subjects have created a specific pile labelled "unclassifiable" to group together atypical images (as seen on point 1) and images they found hard to decipher.

Image #143 (Fig. 8) has been classified into the "unclassifiable" pile by 14% of the panel and has been put into 13 different themes by the others. Terms used to describe its content were very vague (i. e. "a man in a tunnel"). Videotapes made during the test showed that most people looked at this particular image for a long time before putting it aside for a later rereading, their general attitude showing that they found it hard to read.

Other images have been misinterpreted without creating as much worries. The belfry of Lille (image #26 on Fig. 8) has been associated to churches by 12% of the subjects without the least hesitation and the factory saddled with crenellations in picture #19 (Fig. 8) looked much like a castle for one third of the kids.

Sorting and interpretation: a paradoxical situation:

These tests helped us to better understand how people read the LIB.R.I.S. images but they also confirm that, contrary to textual documents, which might prove unintelligible, images always mean something to everyone, even if that interpretation is erroneous.

Images #89 (Fig. 8) is a good example of what might, at first, seem a paradox: Most people don't know what this picture is about, however it is the one that is most often linked to other images in the cluster it belongs to. The vast majority of subjects ignore the use and composition of the heaps they see in the foreground. Nevertheless, this image has close ties with other images dealing with agriculture or countryside views.

The same assessment can be made with image #177: in the foreground of this engraving can be seen an old game called "bird shooting", in the middle distance, a Flemish windmill towers up. 17

out of the 19 people who had created a "Windmills" pile put image #177 into it. It seems that most of them, ignoring what was the foreground object, clutched to the known elements and forgot those they could not identify, provided that the latter did not contradict the coherence of their interpretation.

Image description

The statistics we drawn up from the description vocabulary collected during test #2, show that priority has been given to iconic signs over plastic ones. This confirms that the documentary value of the archive prevails over aesthetic considerations.

Objects	33.33%
People	15,88%
Places	14,43%
Activities/Scenes	10.70%
Buildings	9,35%
Linguistic messages	6,67%
Animals	2.34%
Concepts	2,17%
Geographical zones	0,71%
Medium	0,47%
Type of shot	0,17%

Figure 9: Elements taken into account during image descriptions

Image descriptions have then permitted us to supplement image-sorting results. The words used by a given group have been compared with the captions given to the piles during test #1 and with the content of each pile. This gave us the opportunity to better understand how the different groups had constituted their stacks. In other words, what have been the determining elements for putting an image into a specific class?

Our findings confirm the fact that children have less abstraction abilities than adults do. Indeed, most of them based their classification on singular elements to be seen on the images when other groups created piles with a greater distance to the visible objects. For example all but one child put the "cow" and the "horse" pictures in a specific stack while adults associated them to a more generic class labelled "agriculture" or "farming".

Propositions

The various experiments we have carried out enabled us to precisely analyze how subjects perceived and understood an iconographic document. As a result, we propose a theoretical model of image reading based on the concepts of archetype and scenotype. This model will serve as a basis for the definition of auxiliary tools to analyze and index iconographic documents stored in historical databases in the perspective of their opening to the general public.

Archetype

Our investigation showed that some pictures which are strongly monoreferential act as archetypes of an object, a building, an activity, etc. For example, the picture of the Dunkerque church has been associated with other churches by everyone and all descriptions contain the term "church". Thus, this image could exemplify the "generic" church for our panel.

This led us to define the concept of archetypal images. Archetypal images or archetypes are exemplary images that illustrate a discrete item in a very consensual way. This illustrative function encourages us to use these images as informative documents to be linked to keywords in our thesaurus, in order to explain-by-example the specialized terminology. For example, when a user interested in architecture finds the term "roodscreen" in the thesaurus, he/she will simultaneouly see the iconographic document indexers as the most representative of this architectural element. The archetype will then have the same function as the informative note in a traditional thesaurus.

Now we are studying how hypertext links could relate archetypes to other documents depicting the same element. Images could also be associated to the archetype and thus inherit its general properties.

Scenotype

Most of the documents in the archive show not a single element such as a church, a cow or a car, but a plurality of objects, persons, buildings, etc. Our experiments about image sorting, description and retrieval showed that the different items which are taken into account by users can be distributed

into four categories: places, actors, objects and activities. These categories remind us of the well-known aristotelician categories of nominative (who?), accusative (what?), locative (where?) and ablative (how?). This observation led us to introduce, at the same time as archetypal images, the concept of archetypal scenes or scenotypes. A scenotype is an association of different elements (places, actors, objects or activities) which illustrate a scene in a very consensual way. In other words, a scenotype is an exemplary scene that shares the most characteristics of a topic and no or only few characteristics of other topics.

From then on, it becomes possible to explain how images have been split into classes. Indeed consensual and "nomadic" images can be distributed on a relevance scale according to their relative distance from the scenotype.

When studying the different clusters produced by the co-occurrence analysis, we can say that the more consensual images, i. e. those which are both strongly and most often linked to others, represent scenotypes. For example, image #14 is a scenotype of harvest at the beginning of the century. All elements (place, tools, activity and actors) combine to build a coherent scene. Contrarily, in image #112, the place, tools and activity suggest a blacksmith's shop and would have induced the classification of the picture in a "craft industry" class, had not the age of the actors conflicted with this scenotype.

When considering the case of images that pose interpretation problems, we can note that subjects are not at a loss as long as the document depicts enough elements to be associated to a scenotype. In this way, the apparent paradox we have mentionned about sorting and interpretation is resolved. Picture #89 contained enough elements to be associated to the scenotype "country scene", even if the central element was unidentified. Conversely, no element in picture #149 could make think of something known. Moreover, elements somehow contradicted each other (inside/outside).

Elements for designing a new descriptive note

We will use the "date" and "geographical location" fields to generate the two first interaction spaces of our interface. As far as subject indexing is concerned, we propose to conceive a new descriptive note based on the concept of scenotype. The note will include four sections corresponding to the four categories: "Actors", "Objects", "Places" and "Activities". The minimal information that must be given by indexers is the presence of at least one element for each category, by simply checking a box. This obligatory information should reduce the bias of the point of view and guarantee a minimal indexing for each image. For example, the mere fact of checking a box to inform of the presence of actors in the picture will allow to retrieve documents showing people, even if the indexer was unable to mention their name or function and, because of that, could have considered their presence as insignificant.

Elements for working out an iconographic thesaurus

We have proposed to structure the indexing vocabulary into a thesaurus which could be used by indexers to describe the documents and by users to browse the archive by means of the tree of concepts, which is the third interaction space in our overview interface.

Although we took a lot of ideas from existing sources (Garnier,1984; Groupe image,1993, Orbach Natanson & Alexander,1995), the specificity of the LIB.R.I.S. collection required the use of a great number of specialized terms that were not to be found in the existing iconographic thesauruses. Moreover, we wanted the thesaurus to be organized in four branches, in accordance with the four categories which make up the scenotype: actors, places, objects, activities. Our thesaurus is currently under construction. Its design raises interesting issues regarding the status of a given term. Indeed, the same element should appear in several branches according to its function in the scenotype. For example, a building could be considered as an object if it was the central topic of the picture (architectural testimony), or as a location if it was part of the scenery.

Conclusion

As suggested by A.E. Cawkell, "At present the only practical way of identifying or indexing and retrieving images by humans, is by the language of descriptive words" (Cawkell,1992). But words misrepresent images and the sooner the user gets in touch with the pictures themselves, the better. For this reason, we suggest to go directly from images to images in two ways: First by the mean of an overview interface consistently showing an "image" of the database and of its content. Then by organizing the archive in broad sets, constituted according to scenotypal proximity, and from which the user will visually select relevant images.

A working prototype based on these features is now under development. Further tests will be undertaken with the same panel, in order to validate our hypothesis.

References

Ahlberg, C., Shneiderman, B. "Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays" *CHI'94*, 1994.

Bates, M., J, Wilde, D. N., & Siegfried, S. "Document Familiarity, Relevance and Bradford's Law: The Getty Online Searching Project Report" N. 5. *Information Processing & Management, 32*, 697-707, 1996.

Bovey, J. D. "A graphical retrieval system" in *Journal of Information Science*, 179,188, 1993.

Brochu, D., & Davallon, J. "L'organisation du musée confronté à l'arrivée de nouvelles technologies" in *Médiation sociale, systèmes d'information et réseaux de communication*, (pp. 389-399). Metz: SFSIC, 1998.

Brown, P., & Hidderley, R. "The democratic indexing of images" in *The New Review of Hypermedia and Multimedia*, 107-120, 1996.

Callon, M., Courtial, J.P., Penan, H. *La Scientométrie*. Paris: Presses Universitaires de France, 1993.

Cawkell, A. E. "Selected aspects of image processing and management: review and future prospects" in *Journal of Information Science, 18*, 179-192, 1992.

Chang, S.-K. e. a. "The design of pictorial databases upon the theory of symbolic projections" in *Lecture notes on Computer Science*, 303-323, 1990.

Després-Lonnet, M., Aubry, M., & Briatte, K. "The LIB.R.I.S. project: from heritage conservation to user-oriented concerns" in *VINE* (107 "Digital images in libraries), 8-15, 1998.

Després-Lonnet, M., & Briatte, K. "Can one give online public access to scientific digitised resources?" in *Digital Resources in the Humanities*, DRH'98 Glasgow: HATII, University of Glasgow, 1998.

Dickson, J. "An analysis of user errors in Searching an Online Catalog" in. *Cataloging and Classification Quarterly, 4*(3), 19-38, 1984.

Gardin, J.-C. c *Une archéologie théorique*. Paris: Hachette, 1979.

Garnier, F. *Thesaurus iconographique: système descriptif des représentations*. Genève: Le Léopard d'Or, 1984.

Gervereau, L. *Voir, Comprendre, Analyser les images*. Paris: Editions La découverte, 1997.

Groupe image *Thésaurus des images médiévales: pour la constitution de bases de données iconographiques*. Paris: Centre de Recherches Historiques, 1993.

Hudrisier, H. *L'iconothèque: documentation audiovisuelle et banques d'images*. Paris: La Documentation Française, 1982.

Laulan, A.M. "Problèmes généraux de la relation temps/espace/image/discours" in BPI, C.G.P. (Eds.), *L'image fixe: espace de l'image et temps du discours* (pp. 10-19). Paris: La Documentation Française, 1983.

Orbach Natanson, B., & Alexander, A. "Thesaurus for Graphic Materials I: Subject Terms". *http://lcweb.loc.gov/lexico/tgm/*, 1995.

Robertson, G. "Information visualization using 3D interactive animation" in *Communications of the ACM, 34*(4), 56-71, 1993.

Shneiderman, B. *Designing the user interface. Strategies for effective human-computer interaction* (3rd edition). Reading, MA: Addison-Wesley, 1992.

Taylor "Authority files in Online catalogs: an investigation." in *Cataloging and Classification Quarterly, 4*(3), 1-17, 1984.

Wiberley, S. E. J. "Names in space and time: the indexing vocabulary of the humanities" in *Library Quarterly, 58*, 1-28, 1988.