

From object-centric towards content-centric digital
collections integrating MLA resources

Author:
Maja Sojat-Bikic

CIDOC06
GOTHENBURG
S W E D E N

1. Introduction

Museums, libraries and archives, as holders of heritage resources and users of the benefits of digital technology, enable the general public a high quality access to the heritage and at the same time preserve it for the future. Digital heritage resources are a fundamental value of the civil present and future, both in the dissemination of knowledge and in commercial use, and constitute a sound basis for the development of the digital contents industry. The EU has launched a number of projects coordinating the digitalisation of heritage contents in European countries. Representatives of European Union members adopted the Lund Principles (2001) for the digitisation of cultural and scientific contents. In 2002 a network was set up for European digitalisation projects – MINERVA (Ministerial Network for Valorising Activities in Digitisation), which carries out the plan of digitisation pursuant to the Lund principles and adopts guidelines for technology and contents.¹ Thanks to digitisation processes that have become a very popular way of linking users and information, digital culture and the digital lifestyle have taken off.

Every digitisation project is unique. Success will depend on strategic planning and a judicious choice of contents and the tools with which to realise these contents. The strength of it will stem from the size of the space and range of services that digital on- and off-line media can provide. But the strength also conceals weaknesses, because the space and range of the services of digital media can considerably hamper the ability of users to cope and may reduce the satisfaction in the consumption of the digital product. Digital information can easily get lost in digital noise. For the digital heritage to become information, the basic purpose of the information has to be taken into account – it must correspond to the task, the interests and the queries of a given user at a given moment (Bearman, 2004: 5).

Digitisation is not and must not be an end in itself. Successful digital heritage projects start with a definition of sets of functions and sets of users (Kilbride, 2004: 1). Digital collections set up relations with the real world and a real group of users, i.e., with their expectations and behaviour. Since objects from the real world that are stored in museums, libraries and archives are as a whole

inaccessible to the end user, digital collections enable an expansion of the group of users and the enlargement of the usefulness of the analogue heritage source.

Technical standards and guidelines for the digitisation process are well and completely defined (MINERVA, 2004). The guidelines for contents are much harder to define. Content shaping leads to usable and accessible digital collections that are suitable for the anticipated groups of users.

This paper will present one digital collection of contents created from MLA resources: a museum collection of pictures of an artist, a library collection of historic periodicals the articles of which were illustrated with these pictures, and archival collections – phonographic and cinematic archives, as added value. A hybrid approach to the shaping of the thematic collection has been employed.

2. The digital collection

Digitisation of a real-world object creates an electronic version of it, a digital surrogate or digital object, which is used in computer applications. Digital surrogates are digital versions of expensive, fragile, rare and inaccessible objects from the real-life space (Kilbride, 2004: 4). They are attractive, and not just to experts, because they reduce the danger of damage and dilapidation of the real objects, and enable research to be done with desktop comfort, even to the most general public, because they generate knowledge and develop understanding, viewpoints, value, pleasure, inspiration and creativity.

There are three different types of digital objects, according to the three different types of media – text, image and time-based media (sound and video) (Eadie 2005: 4). Every digital object has

data that describe it, i.e., metadata. Metadata consist of structured descriptive and administrative information about the object (identifier, type, subtype, format, analogue source, owner of analogue source, relations with other objects, rights and use restrictions and so on). Thousands

of objects are created by the digitisation process. For the sake of storage, administration and access, they have to be organised into collections (NISO Framework Advisory Group, 2004: 3).²

A digital collection is a set of logically and physically organised digital objects with a systematic access to objects. The model of the collection (in terms of hierarchy, relationship and so on) is chosen according to three basic criteria:

1. how the analogue objects are naturally organised;
2. what the purpose of their digital surrogates is;
3. what the expectations and experience of the users who use the digital resources are.

A digital collection is described by data at the level of the collection itself (collection metadata), composed of a textual description of the collection, explanations about how and why the material was selected, identifications of sources and heritage institutions for the sake of authenticating the collection, and the status of intellectual property rights (NISO Framework Advisory Group, 2004: 5).

When a collection has been once digitalised and catalogued, information can be displayed in a coherent and intelligible way, the user thus being able to move around the collection and access information effectively. What distinguishes digital from analogue collections is the user interface via which objects from the collection are accessed. A digital collection, then, can be considered from two aspects: that of content and that of the user.

The content aspect defines the purpose of the collection, i.e., the task that it performs in the digital heritage space and the objective that is achieved through its use. The content of the digital collection is chosen according to criteria that give a good characterisation of the cultural and historical material and its information value. Selection criteria according to which the collection is created can be, for example: a single donation; a certain type of material; a geographical area; a period of time; a given person; an object, or combinations. Those digital collections that come

into being from a compilation of materials from different heritage collections and different heritage institutions are more complicated. In this case the choice of content is mostly based on a given topic.

A user interface defines the manner in which the collection can be accessed – browsing and searching. The user interface itself has its own digital objects (menus, icons, input boxes, messages and so on). These are *artificial* digital objects as distinct from the *natural* digital objects that are surrogates of objects from the real world (the problem space) because of which the digital collection was created in the first place. A well designed and pleasant user interface means ease of use, greater user satisfaction and greater market effect. An interface encourages the user to take part by choice of direction of movement. The ease with which a user can find the desired content is one of the most important determinants of the usability of a digital collection. Naturalness, simplicity, intuitiveness, consistency, predictability, intelligibility, reliability, responsiveness, changeability of communication context are additional quality determinants (CIDOC Multimedia Working Group, 1997: 3). The user interface of a digital collection that brings specialised collections together has to be able to present heterogeneous resources in a coherent manner.

3. User oriented approach to digital collections

A user approach to the modelling of a digital collection raises one inevitable question: why has digitisation been carried out? The following answers are among those possible (the list is not exhaustive nor is it put in order of priorities):

- to enable heritage resources to be accessed;
- to preserve the originals;
- to provide added value;
- to integrate specialised collections;
- to offer the digital collection to the market;
- to find new users;
- to support lifelong learning.

The more a digital collection is oriented towards the needs of the users, the more attractive it is going to be. How long will it be topical and usable? Why was it created, what results will it achieve, how will its performance be quantified, will it be sold, who are the users? How can the investor be convinced of the value of the project? All these are issues that arise even before the digital collection is modelled. Modelling starts with a vision of the thematic concept, i.e., with planning an approach to the original material (Jeffrey, 2002: 16). A thematic concept is a coherent story (a storyboard) or a conceptual outline of the digital collection. It derives from several smaller ideas and from development of the basic topic and sub-topics. After the thematic concept has been defined, the phase of digital modelling is addressed. By being shaped in the digital medium, a story acquires its audiovisual expression. The best result will be given by modelling that is deeply linked with the thematic concept.

Attention to the basic principles of the user approach, based on the experience of numerous research and development teams (Macleod, 2004: 2) will facilitate the modelling of a digital collection:

1. **KNOWING THE USER/S:** The term user means the end user. Who are these users, what are their needs and expectations? Most often in fact there are several different user categories.
2. **VISION OF THE END PRODUCT:** Defining a clear vision of the final product to be produced from the digital collection. What is its purpose and benefit? Certainly the trap of production just because it is technically feasible needs avoiding.
3. **OBJECTIVES:** Define the priority objectives that will support a vision of the product and the needs of the users.
4. **USE OF GUIDELINES AND STANDARDS.** Guidelines relate to the planning of the digitisation project, the selection of original materials, the preparation of materials for the digitisation process, the preservation of digital surrogates, the metadata that describe them, and intellectual property rights. Technical standards in the area of digitisation relate to images, sound, video, 3D, metadata and taxonomy.³

5. ITERATION: The route from creative idea to usable product almost always involves iteration because the first design is hardly ever free of faults. The earlier the iterative procedures start, the more painless are the changes. Iteration is connected with parallel evaluation.
6. EVALUATION: There are many methods for evaluation of the design. Most often they depend on the judgement of experts, theoretical analysis and user testing with the application of the user interface scenario.

The user approach to modelling digital collections also implies the modelling of interactivity and navigation through the collection. Dynamic interaction of user with heritage contents can enrich user experience and contribute to learning, research and entertainment (Smith, 2000: 5). Thus from the very beginning a scenario of user actions is formulated and analysed which will enable this objective to be attained. The scenario defines the inputs and outputs, the objects and actions (de Haan, 2003: 4).

The user approach to modelling a digital collection is at the same time a selective approach to contents that will form part of the collection via digitisation. In the global information society, content is the key issue. Technological advances enable ever better processing of information and an increasingly effective communication infrastructure. Greater attention in the research has to be devoted to the digital contents so that they can be more effectively produced, have attractive functionality, be reliably exchanged, used and reused in various digital ways. Thematic collections will more easily satisfy the real needs of real users. Certainly one needs to ask: which digital projects have the greatest chance of success? Are they those that integrate MLA resources? From which aspects should one approach the creation of a digital collection of MLA contents: the museum, the library, the archive, or the hybrid? Here one comes upon the problem of the division among the heritage sectors, different standards of cataloguing, intellectual property rights, trust and so on. All cultural contents are acceptable, but digital contents are built on

available analogue resources. Quality, rarity and uniqueness of resources certainly have a part to play in the selection process. New combinations of contents are created by an integrated approach to heterogeneous collections in traditional and digital forms.

Digital multimedia technology facilitates a new approach to digital contents. It breaks up the linearity of traditional ways of creating and dissemination contents. The object of a digital collection can be shown in different media (text, image, sound, video, animation and so on) and it can be reached via interaction and navigation. The status of the public is changed. It no longer consists of an amorphous readership, but of individual users with specific characteristics. The user acquires the right to the autonomous approach to contents (Pascon, 1997: 6).

4. Case study: an example of an end-user digital collection on a painter, a city and a period

Zagreb City Museum holds a collection of paintings by a well-known Croatian artist, Otto Antonini (1892-1959) who was also editor and illustrator of a popular magazine called *Svijet* (*World*), and most of the pictures in the collection were first published in the magazine. *Svijet* used to be published on Saturdays, coming out in Zagreb from 1926 to 1936, and it was much more than a local review. It wrote about the life of the city and the country, world events, exotic regions, culture, fashion (the germs of local glamour), sport, technical novelties, lifestyle and so on.

The initiating idea of the project for a thematic digital collection about the painter, the city and the period was to bring together specialised collections of images, newspaper and archive sources and to create a collection of contents for the scholarly and the general public. In this manner, the picture, as museum object, is located in its historical context, and obtains additional value from other heritage sources.

The project comprehended:

1. Modelling the contents (planning the information structure) – choice and evaluation of contents and analysis of their short-term and long-term use;
2. Designing the navigation (the routes through the information structure);
3. Designing the presentation of the contents (the presentational aspect);
4. Designing functionality (functions available to the users).

Each digitisation project must be planned respecting selection criteria that will properly characterise the material and its information value. During the selection of contents from *Svijet*, the following criteria were respected:

- *Svijet* is an important source of information about various aspects of life in the 1920s and 1930s;
- *Svijet* constantly kept up with Zagreb and the Zagreb middle classes for a long enough period of time;
- It stopped coming out 70 years ago (intellectual property rights expired);
- Availability of material in the library of Zagreb City Museum;
- All the numbers available;
- The originals in a good state of preservation.

Then the objectives of the project that digitisation had to satisfy were set:

- Enabling access to the collection of images and the historical periodicals;
- To convey part of the history of Croatian painting and journalism to a broad circle of current and future readers;
- To link a certain amount of original journalistic information about Zagreb and the world into a coherent story about the painter, the city and the period (snapshots of the social life of the 1920s and 1930s);
- To provide added value – expand the historical context for the requirements of the end user;
- To turn the story into a digital collection;

- To achieve digital collection functionality;
- To enable searching according to various criteria – dates, topics, persons, corporate entities, localities, objects and events.

Criteria for the selection of materials were set:

- Topics related to Zagreb life;
- Good coverage of the original sections of the periodical – architecture, business, science, culture, social life, fashion, sport, humour, popular reading matter such as fiction and so on.

The project boundaries were set:

- The volume of information to be limited to the volume of portable digital medium (CD ROM);
- Use of existing software tools;
- Keeping down costs.

The technical decisions were made:

- Segmentation of the periodical into articles;
- Separate scanning of images;
- Input of articles and images into a relational database
- Combination of retyping and OCR-ing of articles, depending on quality of printing and layout of text, and the consequent amount of interventions if OCR quality was not satisfactory.

The 524 articles (see Fig. 1) selected are connected in their contents mainly with Zagreb, as well as with everyday activities of ordinary life – travel, sport, fashion, beauty queen contents, season changes, festive occasion customs and atmosphere. Unfortunately, because of the many thousands of articles, it was what had to be lost that needed choosing.

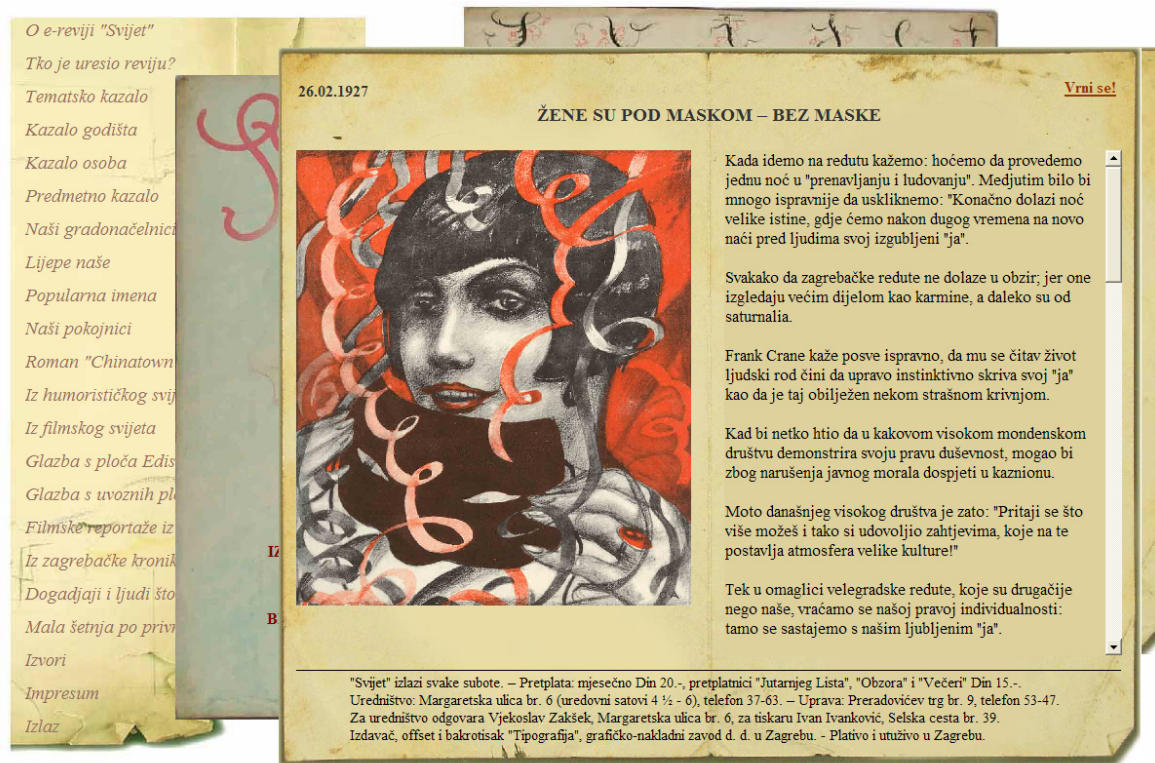


Fig. 1 An article on women in disguise

Sound and video, additional sections (*Our Mayors*, *Our Departed*, *Our Lovelies*, *Popular Names*, *From the World of Humour*, *From Zagreb Chronicles*, *From the World Chronicles*, *A Stroll around Business Zagreb* and so on) constitute the added value of the digital surrogate. Thus for example the section *Our Lovelies* shows a collection of glamorous covers that communicate the ideal of feminine beauty and the fashion departures of the 1920s and 1930s, while *Popular Names* gives the most popular male and female names (see Fig. 2).

Svijet published the scores and words of the hit tunes, and the added value here consists of the same songs from old gramophone records. Excerpts from news footages of the 30s are thematically linked to selected articles. To ignore Antonini's advertising illustrations, a source of sociological, economic, political and cultural information, would have been to impoverish the context of the time. And to ignore humour would be the same – a joke too is a document of a time.



Fig. 2 Top twelve female and male names of 30s

At the beginning of every digitisation effort the question arises as to whom this is all for. The anticipated group of users of this digital collection covers the general readership, and special groups such as journalists, historians, students, schoolchildren, their teachers and so on. The collection can serve as a readily available reference library and as an educational aid in schools.

Conclusion

The diversity of heritage material and the great sets of data require a content-oriented approach to the shaping of digital heritage collections. The technical guidelines and standards for digitisation are well defined. The key issue is the content, its effective production, attractive functionality and market value. Digital contents are constructed on available heritage resources for real users in real life situations. The attributes of the resources, such as quality, rarity, uniqueness, cultural importance for the local community and so on play key roles. The digital heritage space is a surrogate for the analogue heritage space, and is created primarily out of democratic principles,

enabling the people to access the heritage, while at the same time preserving the integrity of the original. The creation of digital collections from digital contents opens up new opportunities in learning and research, for experts and students, lifelong learners, lovers of culture and history and other groups of users. Digitisation has many benefits, among which is certainly the large number of users that can access digital resources. But users should not be taken for granted. Evaluation of the contents to be transposed into digital resources and an early understanding of the reasons for such a migration are of critical importance for the success of any digitisation process.

Digital technology enables the integration of different heritage resources from different heritage communities. There is more benefit from digital multimedia technology precisely because of this ability to integrate than from the technological capacities per se. Collaboration in the digital future, partnership in the heritage sector (archives, libraries, museums, history and culture societies, radio, TV, those who possess digital contents) will result in content enrichment and new user opportunities.

Endnotes

1 MINERVA has developed a considerable publishing activity, and publications are available on-line in several languages (<http://www.minervaeurope.org>). One of the publications, Charter of Parma, 2003, continues and backs up the Lund Principles concerning the digitisation of the heritage, and says that digital processing of material is the most essential step that European heritage establishments should take with the objectives of preserving and valorising the heritage and cultural diversity, facilitating access to the heritage, and encouraging the production of digital contents and services (MINERVA, 2003).

2 The National Information Standards Organization (NISO) Framework Advisory Group has published A Framework of Guidance for Building Good Digital Collections, giving a review of activities and a framework for action while creating and managing digital collections.

3 The problem with most guidelines and standards is their sheer numerousness, and for this reason it might be difficult to cope. Useful sources are MINERVA's Good Practice Handbook (MINERVA Working Group 6, 2003) and The NINCH Guide to Good Practice in the Digital Representation and Management of Cultural Heritage Materials (NINCH, 2003). Useful instructions can also be found at the VNET5 portal (<http://www.vnet5.org>)

founded by the European Commission with the objective of providing guidelines for the development of electronic publishing projects meant for the end users of digital contents.

References

- Bearman, D. (2004) Sustaining Culture, Enhancing Life: User Requirements for Cultural Heritage, at http://eu2004.digitaliseringerfgoed.info/sites/cultuurtechnologie/contents/i000133/eu2004_users.doc, accessed 23 December 2005
- CIDOC Multimedia Working Group (1997) Multimedia Evaluation Criteria, at <http://www.archimuse.com/papers/cidoc/cidoc.mmwg.eval.crit.html>, accessed 4 March 2006
- de Haan, G. (2003) Design for Global Access to Cultural Heritage. Greenwich UK: The 2nd BCS HCI Group Workshop on Culture and Human Computer Interface, at <http://home.tiscali.nl/gdehaan/articles/imass-cult-and-hci.html>, accessed 5 March 2006
- Eadie, M. (2005) The Digitisation Process: An Introduction to Some Key Themes, at <http://ahds.ac.uk/creating/information-papers/digitisation-process/index.htm>, accessed 4 March 2006
- Jeffrey, R. (ed.) (2002) Producing Online Heritage Projects. Hull: Canadian Heritage Information Network.

Kilbride, W. (2004) Why bother with digitisation? Users and using digital resources, at <http://ahds.ac.uk/creating/information-papers/why-bother-digitising>, accessed 22 December 2005

Lund Principles (2001) at http://www.cordis.lu/ist/digicult/lund_p_browse.htm, accessed 2 January 2006

Macleod, M. (2004) "User-Centred Product Creation in Electronic Publishing – Good Practice Models" in: A Guide to Good Practice in Collaborative Working Methods and New Media Tools Creation, at <http://ahds.ac.uk/creating/guides/new-media-tools/mcleod.htm>, accessed 14 February 2006

MINERVA (2003) Charter of Parma. Rome: Ministerial Network for Valorising Activities in Digitisation.

MINERVA (2004) Technical Guidelines for Digital Cultural Content Creation Programmes. Rome: Ministerial Network for Valorising Activities in Digitisation.

MINERVA Working Group 6 (2003) Good Practice Handbook, version 1.2, at http://www.minervaeurope.org/structure/workinggroups/goodpract/document/bestpracticehandbook1_2.pdf, accessed 2 January 2006

NINCH (2003) The NINCH Guide to Good Practice in the Digital Representation and Management of Cultural Heritage Materials. Glasgow: Humanities Advanced Technology and Information Institute, University of Glasgow & National Initiative for a Networked Cultural Heritage, at <http://www.ninch.org/guide.pdf>, accessed 21 December 2005

NISO Framework Advisory Group (2004) A Framework of Guidance for Building Good Digital Collections. 2nd edition, Bethesda, MD: National Information Standards Organization.

Pascon, J. L. (1997) "Developing a National Strategy for Multimedia Cultural Heritage", in: Proceedings of the Museums and the Web International Conference. Los Angeles, at <http://www.archimuse.com/mw97/speak/pascon.htm>, accessed 2 January 2006

Smith, B. (2000) "Digital Heritage and Cultural Content in the New Information Society Technologies Programme", Cultivate Interactive, issue 1, at <http://www.cultivate-int.org/issue1/ist>, accessed 20 December 2005

are used here as examples to show the wide scope of the research domain. They share characteristics with each other but also with different fields of expertise to the extent that they belong to different vernacular space time. This in turn means to highlight the fact that not only.