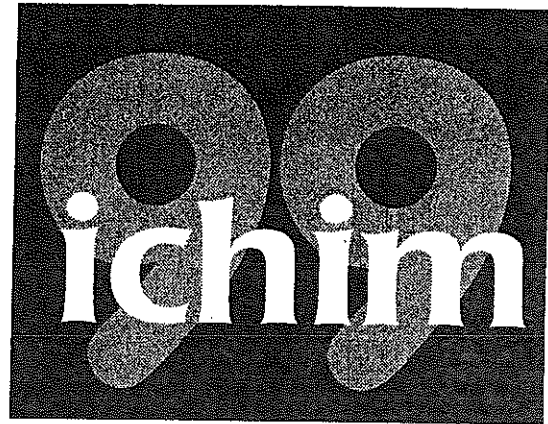


A2-31
1999

6



**Cultural Heritage
Informatics:
selected papers
from ichim99**

IST. EL. INF.
BIBLIOTECA
ARCHIVIO

A2-31
1999

**Edited by
David Bearman and
Jennifer Trant**

<i>A Nomadic Information System for Adaptive Exhibition Guidance</i> _____	103
Reinhard Oppermann and Marcus Specht, German National Research Center for Information Technology, Institute for Applied Information Technology, Germany	
<i>Hypermedia in Archaeological Exhibitions: different kinds of interactivity and visualization</i> _____	111
Martin Kriszat, University of Hamburg, Germany	
<i>Virtual Reality and the Future of Publishing Archaeological Excavations: the multimedia publication of the prehistoric settlements on Tsoungiza at Ancient Nemea</i> _____	125
Mary K. Dabney, James C. Wright, Bryn Mawr College, and Donald H. Sanders, Learning Sites, Inc., USA	
<i>Museum Informatics and Information Infrastructures Supporting Collaboration across Intra-Museum Boundaries</i> _____	133
Paul F. Marty, Spurlock Museum, USA	
<i>Emerging Tools & Techniques Of Digital Media: history log and multiple futures</i> _____	139
Slavko Milekic, The University of the Arts, USA	
<i>Technologies and methods for interactive exhibit design: From wireless object & body tracking to wearable computers</i> _____	147
Flavia Sparacino, Kent Larson, Ron MacNeil, Glorianna Davenport, Alex Pentland, Massachusetts Institute of Technology, USA	
 Access	
<i>Electronic Esperanto: The Role of the Object Oriented CIDOC Reference Model</i> _____	157
Martin Doerr, Foundation for Research and Technology Hellas, Greece and Nicholas Crofts, Direction des Systèmes d'Information, Geneva, Switzerland	
<i>Finding Museum Information in the Internet Commons: a report on the CIMI Dublin Core Metadata Testbed Project</i> _____	175
John Perkins, Consortium for the Computer Interchange of Museum Information (CIMI), Canada and Angela T. Spinazze, ATSPIN Consulting, USA	
<i>Aquarelle – Networked Cultural Information</i> _____	179
G. L. Mallen and M. J. Stapleton, System Simulation, Ltd., UK	
<i>From Images To Images: the ALADIN project</i> _____	185
Katell Briatte and Marie Despres-Lonnet, Université Charles de Gaulle, France	
<i>'Guiding' the World to Canada's Museums: the yellow brick road or the straight and narrow?</i> _____	199
Kati (Ecaterina) Geber, Canadian Heritage Information Network (CHIN), Canada	

Using & Re-using Archive Information for Multimedia Applications: the virtual museum of Italian computer science history

Franca Garzotto and Paolo Paolini, Politecnico di Milano, Giovanni Mainetto, CNUCE-CNR, Serena Pisani and Pasquale Savino IEI-CNR, Italy

Abstract

The quality of content is a key attribute for assessing the global quality of a museum application. Unfortunately, producing good content, especially in multimedia digital form, is expensive and time-consuming. One way to reduce the costs without sacrificing quality is to exploit the concept of information reuse. The idea is to use (portions of) the same multimedia material in different applications, possibly adapting it for different contexts, for different categories of users, and for different delivery channels (e.g., on-line and off-line). Information reuse does not come free. To be effective, it requires a well-organized environment in which information can be easily stored, inspected, retrieved, and adapted for different purposes. This paper describes the approach adopted in the project "The Virtual Museum of Italian Computer Science History", funded by the Italian National Council of Research (CNR). In this project, all the digital material (documents, images, video interviews, etc.) is stored in a digital archive based on a multimedia database with a WWW front-end. The archive is designed for specialists only: members of the editorial board of the project; researchers in the history of science; application developers (who are looking for interesting content to include in their CD-ROMs or Web sites). Each research group involved in the project extracted and adapted from the digital archive the multimedia material needed to build a different hypermedia application in two "versions" - WWW and CD-ROM. These applications, both on-line and off-line, strongly reuse (portions of) the digital archive content, but organize and present it with a totally different style, to address the needs of non-specialists (e.g., people who have some interest, or curiosity, on the history of Italian computer science).

Introduction

Producing good-quality multimedia material is expensive and time-consuming; thus, a challenging issue is to support information "reuse" (Garzotto 1996), i.e., multiple use of the same multimedia material in different applications, in different contexts of use, and for different categories of users. In principle, developing new applications by "sharing" multimedia contents requires less time for their development and costs much less than developing their content totally from scratch.

Unfortunately, information reuse does not come free. To be effective, it requires development of a well-organized *editorial environment* in which multimedia information can be easily classified, stored, inspected, and retrieved for different purposes. It also requires some kind of *integration* between the editorial environment and the application development environment, as well as systematic *design methods* so that it is easier for developers to re-use the same base of information for building different applications. This paper will describe the technical approach adopted in the project "The Virtual Museum of Italian Computer Science History" to address the above issues.

From a cultural perspective, this project originated from the idea that, although computer science is a young, ever evolving discipline where everything becomes obsolete in a few years, its history is part of the memory of our society, and many lessons can be learned from it. We should therefore avoid the risk of forgetting the events of its evolution and the testimony of the people that made this history, which are not purely anecdotal, but which have intrinsic cultural and social value (De Marco 1999).

In acknowledging this, CNR - the Italian Research Council - launched in June 1996 the project "Museo Virtuale della Storia dell'Informatica in Italia" (Virtual Museum of Italian Computer Science History). The project's purpose is to exploit information technology in order to preserve and to disseminate the cultural heritage concerning the history of computer science in Italy, which would otherwise be lost in a few years, without any possibility of recovery. The project focuses on the "infancy" of Italian computer science, between the mid-fifties and the end of the seventies, when computer science was in its infancy and where Italy played a key role.

Cultural Heritage Informatics

digital archive is therefore conceived of as a multi-author activity;

- to support scalability, in order to accomplish the progressive growth, with time, of the multimedia content. The acquisition of the archive material is incremental; the archive is continuously enriched with new multimedia content as the collection activity proceeds (and new material is discovered). Populating the archive may, in principle, continue indefinitely, extending, in the future, the period of interest, from the early years of Italian computer science to present day;
- to make the material easily available for developers of multimedia applications working in this field and which are targeted for a non-specialistic audience;
- to keep, at the same time, complete independence from the particular applications that can be developed.

From a technological perspective, the best candidate to support the above requirements is to develop an on-line multimedia digital archive, based on standard data base technology and accessible via any standard WWW browser. The actual content is stored in a centralized data base (currently located at CNR in Pisa) developed with a commercial object-relational DBMS (Illustra/Informix). On

top of it, we have developed a WWW front end, which supports both remote update of the archive (by the various members of the editorial team of the project) and remote access, by the project members and any user who is interested in the archive content. The use of standard data base technology supports the needs of scalability, multi-user access and update, and security.

The Archive Data Schema

The archive contains various classes of information: information about personages who, with various roles, contributed to the development of computer science in Italy, their biographies, audio and video interviews and photographs; about public and private institutions that had (and partially still have) a relevant role in the history of Italian computer science; about software systems (as operating systems of the first computers) and hardware systems (as electronic computers and calculators) with historical and technical cards, photographs, audio and video recordings; and, finally, about events that have marked the fundamental steps in the development of computer science in Italy. The data base schema describes these various classes of information and models the relationships between them. Figure 1 shows a simplified version of the schema to provide an intuitive idea of the data organization. As a matter of fact, the multimedia digital archive is much more complicated

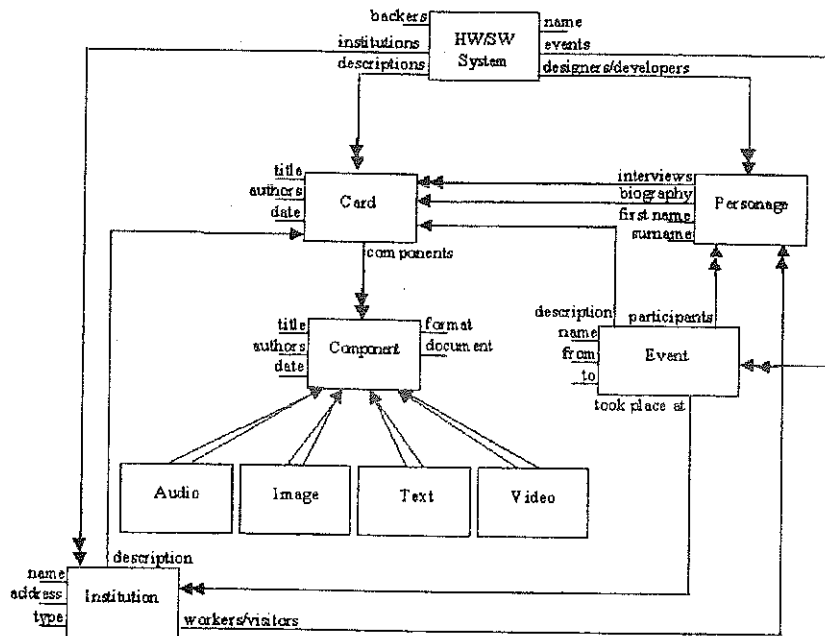


Figure 1. The Multimedia Digital Archive Schema

Bibliography, Documents with historical soundness, Institutions, Events, Images, Personages, Cards, Hardware Systems, Software Systems, Textual Documents, Functional Units and Video Documents) that can be used as a starting point; successive steps are performed by following the links starting from these categories, and correspond to the database schema relationships. Multimedia presentation tools allow one to visualize the information dynamically extracted from the digital archive. This dynamism is a particularly interesting aspect since new information inserted in the archive will be automatically visualized without further efforts.

If a user is, for example, interested in a given personage, he can access the Personage category and choose the person about which he is searching for information. The information presented depends on what is available about the personage; for example, a guided tour presenting the chosen personage can present his/her photograph first, then his/her biography (see figures 3 and 4). Moreover, the user can see interviews of the personage (as figure 4 shows), access the events to which he/she has taken part, the institutions in which he/she has worked or visited, and the software and hardware systems on which he/she has worked.

At any time, the user can change category and see, for example, the events that have marked important steps in the history of Italian computer science.

The hypermedia applications

The hypermedia applications developed in the project have different purposes with respect to the multimedia archive. The latter is designed for pro-

fessional use. Its intended users are members of the editorial board of the project, who need to insert new material or update the existing one; or researchers in the history of science, who are looking for material which has some interest for their studies; or application developers, who need interesting content to include in their CD-ROMs or Web sites. As such, the digital archive stores *all* the material collected during the project, in a number of different formats and versions. The content is organized using technical criteria, to optimize storage allocation. The interface tries to maximize the efficiency of a specialist's access, based on queries and goal-oriented exploration.

In contrast, the hypermedia applications, both online and off-line, are mainly designed for non-professional use. Their main requirement is to promote some interest in the early years of Italian computer science, to stimulate reflection, especially by the young generation, on what happened in industry and academy in those years and how it relates to the present. Intended users are "casual" visitors who are mainly motivated by curiosity or people who have some interest on the history of Italian computer science, either for education purposes or just for personal culture.

Part of these users may need to be stimulated to use the system. In general, they do not require the availability of the whole body of information stored in the archive. Textual descriptions of "concepts" (biographies, events, machines, etc.) can be short for these users (the archive instead also stores long detailed presentations about each subject). They are particularly attracted by visual content (images) and above all, by live, interactive material such as video interviews (where more curious and anecdotal information can be discovered). In terms of

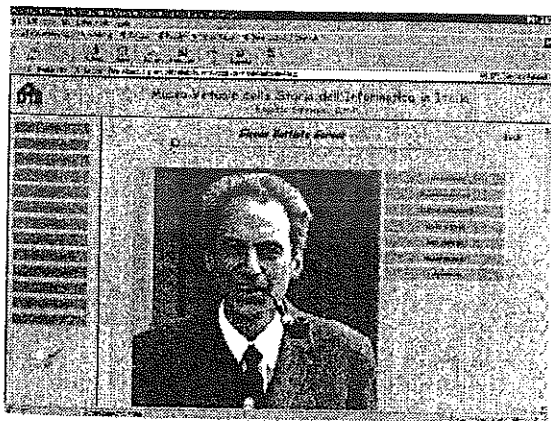


Figure 3. A personage in the digital archive of the Virtual Museum of Italian Computer Science History

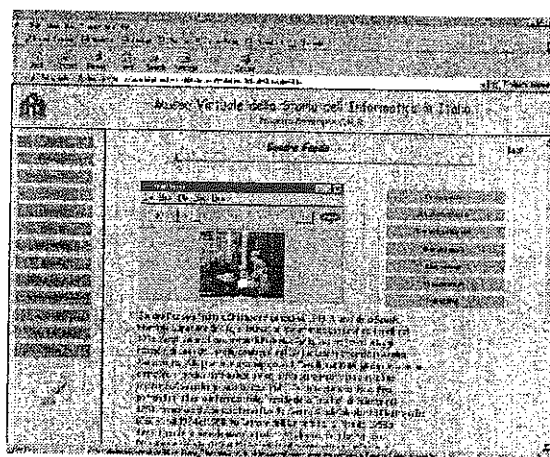


Figure 4. Another personage in the digital archive of the Virtual Museum of Italian Computer Science History

Cultural Heritage Informatics

- component Long Bio: includes a Long text Description of the person

Computer System

- component Description: includes Computer Code, Date, Use (commercial or scientific), Computer Generation, Short Presentation, one or more Images
- component SW (software): includes a Short Description of the software developed for the system
- component Functional Unit (FU): includes Computer Code, Functional Unit Name (e.g., Arithmetic Unit), one or more Images, Short Description
- component Functional Unit (FU includes Computer Code, Functional Sub-Unit Name (e.g., Arithmetic Unit), one or more Images, Short Description

Q&A (question and answer)

- component Q&A: Interviewed Person Name, Institution Name, Question - transcription, Answer - transcription, Video Fragment

The conceptual design of the applications includes the definition of "collections" (Garzotto 1994a) that represent sets of entities, or sets of other collections. Members of a collection can be homogenous or heterogeneous. They can be grouped together for different reasons: because they have the same nature (e.g., "all interviews"); because they are mutually related by some semantic relationship (e.g., all places where a given person worked); because they are all useful for a given user task (e.g., have a quick overview of the most important information on a given theme"). Collections have a distinguished element called "center" which provides an overview of the collection content and the entry points (links) to access collection members.

The main collections are listed in the following table, where we indicate the types of objects grouped by each collection. All collections are homogeneous, i.e., they group elements of a given type, except those named "Guided Tour on Theme X_i" (on the bottom of the table). These are heterogeneous and groups information objects of different types that are related to a given theme (e.g., "The arrival of the first computer from the U.S.").

Navigation is quite rich in the hypermedia applications. Users can choose where to start accessing

Entity Type ⇒ Collection ↓	Person	Place	System	Historical Document	Q&A	Event	Curiosity
All persons (alphabetic order)	✓						
Persons by Place	✓						
Persons by Event	✓						
Persons by system	✓						
All computer systems (alphabetic order)			✓				
Systems by generation			✓				
Systems by development place			✓				
Systems by person			✓				
Sw products			✓				
All Events (Chronological order)						✓	
Events by year						✓	
Events by place						✓	
All places (alp. order)		✓				✓	
Q&A by subject (alp. order)					✓		
Q&Q by Person					✓		
All Curiosities (alp. order)							✓
Curiosity by subject							✓
All historical documents (chronological order)				✓			
Historical document by event				✓			
Guided tour by Theme X ₁	✓	✓	✓	✓	✓	✓	✓
Guided tour by Theme X _n	✓	✓	✓	✓	✓	✓	✓

Structure Table: The structure of the collections in the hypermedia applications of the Virtual Museum of Italian Computer Science History

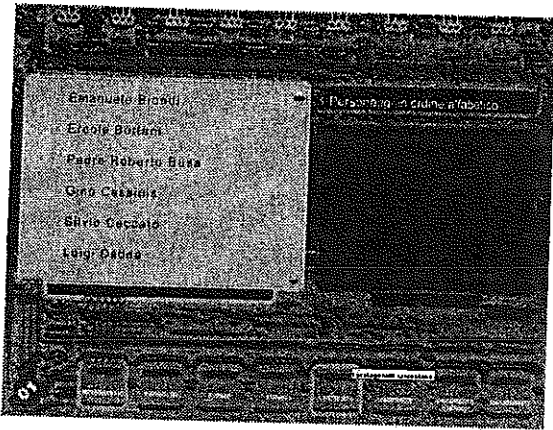


Figure 8. The entry point to the collection "Personaggi by Place (Milano)"

computer mother board. The gray background shows, in the shadow, micro circuits and board units. The idea of circuits and CPU units are also suggested by the buttons, the lines separating or connecting the various areas on the screen, the frame around the content fields. Colors are strong and contrasted, with a careful use of gray, black, white, intense red, intense yellow, and "electric" blue.

Conclusion and future work

This paper has presented the approach adopted in the Project "The Virtual Museum of Italian Computer Science History" for the reuse of multimedia documentation in the creation of interactive multimedia applications.

The experience gained in the project for the creation of the multimedia archive - together with the applications for its access and management - and for the development of different hypermedia applications both on-line and off-line (e.g., on CD-ROM) brings us to the following conclusions:

- The cost of the production, selection, organization, etc. of the multimedia material is largely dominant with respect to the cost for the development of the applications.
- The possibility of reusing, even part of the material produced, provides a significant reduction in the cost of the development of a new application. This reduction is much more significant when different versions of the same application are produced (such as the CD-ROM and WWW versions) or when updates of the application are created (e.g. when new material becomes available).

- The possibility of reuse is strongly dependent on the availability of a well-organized environment, like the digital archive developed in the project. This acts as a multimedia repository of the collected material, supporting hypermedia developers to search and to retrieve, in a very effective way, what needs to be reused for application development. The digital archive also acts as an editorial environment, which makes the process of storing new content and maintaining the overall body of knowledge sources easy, efficient, and well organized.

- A multimedia repository where data are stored in different formats, guarantees that the digitized material can be accessed in the long term. This makes *reuse over time* possible. The problem of format obsolescence is extremely important and critical, especially in the case of archives containing historical data: the solution adopted in the project is also based on the adoption of procedures for the updating of existing material, thus making possible the insertion of new data formats when necessary.

The work done so far has also evidenced the need for further research work in order to obtain an optimal solution. In particular, we are considering continuing our research activity in the following areas:

- In many cases, it is useful to enable the reuse of parts of applications, besides the multimedia contents. Research work is needed to understand how *reuse of application components (modules)* can be done, how subparts of an application can be extracted.
- Related to the previous aspect, there is the need of creating an archive of reusable application components. This also requires the definition of appropriate mechanisms for the selection of the relevant application modules (query mechanisms) together with a suitable query formulation interface.
- The final and most ambitious goal is to provide a strong integration of the application with the multimedia repository. This means that data are stored in the archive and the application acquires them only when needed, by using query by content mechanisms instead of directly accessing the objects. This approach makes possible the development of applications whose content can be updated continuously, as soon as new material becomes available.