Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Cultural Heritage 7 (2006) 92–97
Journal of Cultural Heritage
http://france.elsevier.com/direct/CULHER/

Original article

The digital opera house: an architecture for multimedia databases

G. Haus, L.A. Ludovico *

Laboratorio di Informatica Musicale (LIM), Dipartimento di Informatica e Comunicazione, Università degli Studi di Milano, 39, via Comelico, 20135 Milan, Italy

Received 5 October 2005; accepted 20 February 2006

Abstract

This paper deals with the problem of preserving, organizing and retrieving information for a typical opera house environment. On the one hand, the live fruition of a music work in a theatre is an experience very difficult to be recreated in a different context or handed down to posterity; but, on the other hand, opera houses are important centers for cultural preservation and diffusion, and their work cannot get lost immediately after performances. The processes in such an environment are very heterogeneous and complex, including not only the economic management and the logistic activities which take place in the offices, but also on-stage artistic production and craft-made activities in workshops. Probably, these latter activities are the most interesting from the point of view of cultural heritage. This paper provides a classification of the heterogeneous data to put in relationship in order to obtain a thorough and effective database. The ultimate purpose is highlighting which information should be captured, structured, and retrieved in order to transform musical performances in cultural heritage for posterity.

Keywords: Music databases; Cultural heritage; Digital knowledge; Information retrieval; Multimedia

1. Introduction

The ever-growing possibilities and needs to organize information in large databases have recently involved theatrical environments [2,3]. In the case of theatres and opera houses, databases could be employed not only to store administrative and operational data, but also to preserve and transmit the artistic activity itself.

As technology evolves and catches on in the most disparate fields, an increasing number of cultural institutions take into consideration the need to control and manage their vast amount of data and digital media assets. Technically speaking, a digital asset is any form of media that has been converted into a binary source. The term "asset" is used to indicate that such files have some sort of intrinsic value that makes it worthwhile to manage them. Thus, the locution "digital asset management" (or simply DAM) refers to the practice and domain of organizing digital files. DAM is a field related to content management, and often is considered as a superset of this subject.

Our paper deals with the design and implementation of an efficient and effective Digital Asset Management for theatrical environments. A well-designed and structured DAM, by preserving and organizing information, allows its users to save time and money. For instance, long times due to researches into traditional archives, retrieving the collocation of physical objects, composing elaborated document layouts[1] are virtually cancelled.

Before analyzing the subject thoroughly, some words should be spent about the enormous importance of opera houses' activity from the point of view of cultural heritage. Music, in its various forms and representations, constitutes a patrimony of mankind. Such heterogeneous world is made not only of chords and rests (scores), but also of sounds and voices (recordings), of images (photo and video captures), of craft-made objects (costumes, accessories, backdrops and stage tools), of men's work (hairstyle and make-up).

A well-structured database for an opera house is intrinsically different from databases appointed to other purposes. Materials of interest for opera houses include both data and meta-

* Corresponding author.
E-mail addresses: haus@dico.unimi.it (G. Haus), ludovico@dico.unimi.it (L.A. Ludovico).

[1] A trivial example is constituted by the editorial activity to compose evening's programs, which include articles, reviews, photographs, paintings and so on.

1296-2074/$ - see front matter © 2006 Elsevier SAS. All rights reserved.
doi:10.1016/j.culher.2006.02.007

data, and data themselves can be extremely heterogeneous. In fact, such databases have to store a number of document types, e.g. text, audio, still images and video; and multimedia contents, of course, require a multimedia database. As an immediate consequence, the database itself will need a large amount of disk space and a very smart way of managing related information. In other words, both hardware and software requirements will be demanding.

In this context, a particular importance is acquired by two activities: i) database planning, and ii) related applications design. In fact, even a very powerful calculation system could not be sufficient to handle the computational burden, in particular when data and metadata are not organized in a proper way. Thus, database structuring requires an accurate analysis, and related applications to retrieve information should be implemented very carefully.

2. Contents of an opera house database

Which kind of information should find place inside the database of an opera house? As told before, we can roughly consider two different categories of information: i) data about theatre management (e.g. personal data, salaries, addresses), and ii) data and metadata about its artistic production (e.g. playbills, recordings, photos).

Often, the two different ambits are tightly related: for instance, precise information about the presence of an artist in a certain number of performances could be used not only to reconstruct the cast of the plays, but also to calculate his/her salary. Many other examples could be cited: the relationship among the costs of a particular staging and the run of the play, the availability of recordings to be marketed, and so on.

Only the second category of data and metadata is peculiar in a theatrical environment, so this paper will focus this subject, omitting comments about the organization of a traditional archive of personal data. However, even ignoring one of the two facets, the matter is still complex and challenging. For example, the contents to be stored in the database can be very heterogeneous, ranging from textual information to symbolic representation of music, from still images to video, from audio to physical objects. A more structured vision of the matter will be presented in the following sections.

We can individuate three different phases that are consecutive and logically interdependent: information capturing, information structuring, and information retrieval. In this context, we prefer the term “information” to “data” as it can embrace also the concept of metadata. Literally meaning “data about data”, metadata is information that describes another set of data. In the paper, we will underline that database entries are constituted not only by digital objects, but also by some additional information. Metadata are fundamental in order to classify and retrieve such digital objects. Besides, metadata allow one to relate objects to other objects inside the database. As a consequence, our proposal for an effective database will support both data and metadata.

3. A night at the opera

The purpose of this section is introducing some important concepts typical of theatre’s jargon.

A standard opera house usually runs different kinds of show: mainly operas, ballets, symphonic concerts, chamber concerts, and recitals, but also other initiatives such as lessons, conferences, and presentations.

By adopting the top-down approach shown in Fig. 1, the first key concept is the idea of *base version*, which represents the music work as conceived by its author(s). This definition is meaningful for operas, ballets, and concerts, and not for other kinds of manifestation such as conferences.

Among the data related to base version, we can cite the original title of the composition, the name and the role of its author(s), the date and place of its first representation, and finally the instrumental and vocal ensemble. In this context, the ensemble is a mere list of instrumental and vocal parts, often with a quantitative indication. For instance, by examining the original score we learn that S. Prokofiev’s *Peter and the Wolf* is written for flute, oboe, clarinet in A, bassoon, three horns, trumpet, trombone, timpani, triangle, tambourine, cymbals, castanets, snare drum, bass drum, and strings. On the contrary, the cast—intended as the names of the performers—cannot be specified in the base version. Similarly, for an opera or a ballet the ensemble information includes also the name of the main characters, but once again not the personal data about the interpreters.

The second key concept is the one of *staging*. By such term, we indicate a set of performances characterized by the same music program, belonging to the same season and having common features about production, staging, and cast. Under our hypotheses, *Le nozze di Figaro* by W.A. Mozart and L. da Ponte represents a “base version”, whereas its production for Teatro alla Scala’s 2005/06 opera season will constitute a “staging”.

It is worthy of noting that stagings—if compared with base version—are characterized by a superset of common features. In fact, all the common aspects of a base version are inherited by stagings, which allows one to put in correspondence different stagings of the same music work. But choosing a particular staging makes some other peculiar elements emerge, such as the same conductor, orchestra, director, stage designer, cos-

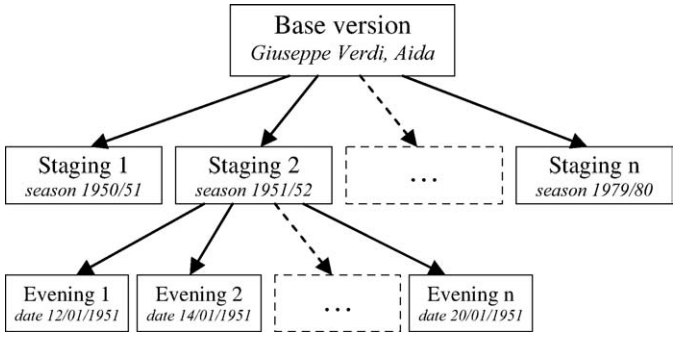

Fig. 1. Hierarchy among base version, staging and evening concepts. In italic, an example about different revivals of the same opera.

tume designer for all the performances. Among the common features of a staging, we cannot count the name of the interpreters (singers, dancers…), as the general rule in opera houses is having two or more casts that alternate on the stage in different evenings.

This last consideration brings us to introduce the final key concept: the *evening*, where not only the music program but also the performers are fixed. An evening represents the implementation of a base version (or many base versions, or even a selection from different base versions), according to a particular staging, in a particular date and place and with a definite cast of performers. According to this interpretation, evenings are in biunique relationship with fliers or playbills, which are the single-page leaflets advertising the event.

Thanks to our top-down approach, it is evident that moving from base version to evening concepts we are providing more and more punctual and detailed information, whereas the levels of abstraction and aggregation lowers: the base version of G. Verdi's *Otello* was performed thousands of times at Teatro alla Scala, while the 7th December 2001 performance is a single evening; on the other hand, each *Otello* evening is characterized by very detailed information, but all the evenings belonging to the same base version share few common features.

A particular situation happens when a performance can be subdivided in a number of blocks, each presenting completeness and independence from others; in this case, we can assume that each block belongs to a different base version. A typical example is represented by those vocal recitals where arias from different operas are sung, as each aria belongs to the base version of the corresponding opera. This situation, namely many base versions involved in a single evening, can occurs in composite[2] concerts, ballets, and operas.

4. Data capturing

The first step to build the database is capturing the information of some interest for our purposes. Identifying all the possible types of heterogeneous material is not a trivial process. As Teatro alla Scala experience taught us [5,6], an opera house is a very rich and complex informative environment, whose materials and documents include:

- scores and symbolic representations of music;
- audio recordings;
- video recordings;
- fliers, playbills and posters;
- photos;
- sketches;
- fashion plates;
- costumes and related accessories;
- stage tools and equipment;
- stage maps;
- other textual documents, such as bibliography, discography, libretto, short descriptions and reviews of music works.

This list does not claim completeness, but it is sufficient to illustrate the heterogeneity of data and metadata the database will store.

First, the original form of the material often has to be converted. Obviously, in a multimedia database only digital objects and data can be entered. Some contents from the aforementioned list are already digital encodings: for instance, digital photos, digital recordings, and computer-edited texts. Besides, some contents undergo an analog-to-digital conversion which prevents (or should prevent, under particular conditions) informative loss; it is the case of image scanning, analog audio and analog video digitalization, typing of digital documents (texts and scores) from a hard copy. Finally, there are physical objects that can be captured only from certain points of view and a limited number of times. In this case, we know that the digital copy does not allow one to catch and appreciate all the facets of the digitalized object. Let us cite the examples of a stage tool or a costume (Fig. 2): many digital photos can be taken and stored to show their features, but the whole photographic sequence is not sufficient to provide a three-dimensional view of the original object. For physical objects, a virtually exhaustive set of digital descriptions (both textual and visual) should be chosen; but such choice would represent—in any case—an arbitrary operation which introduces informative loss. For instance, three photos of a costume can show its characteristics better than a single shot, as well as 100 photos give us a far better three-dimensional view of the subject, but neither 1000 photos can make up for the sensation of touching the costume's fabric.

5. A proposal for database structuring

When designing the database, a problem soon arises: what are the most appropriate relationships among the aforementioned digital contents listed in Section 4 and the concepts of base version, staging and evening listed in Section 3? A possible answer is provided by the entity-relationship diagram shown in Fig. 3. This layout represents just an example among many combinations of different constraints and requirements. Some aspects of our proposal—for instance the type of materials involved and their relationships—are a direct consequence of LaScalaDAM project [4], which took place at Teatro alla Scala of Milan and was coordinated by our laboratory[3].

[2] In this context, "composite" means "made of pieces by different authors" or "made of distinct pieces by the same author". For example, we can define composite an evening when G. Puccini's triptych is performed: the single base versions are respectively *Il tabarro*, *Suor Angelica*, and *Gianni Schicchi*. On the contrary, according to our definition the performance of C. Orff's *Carmina Burana* is not composite: even if the work could be subdivided into more atomic parts ("O Fortuna", "Fortuna Plango Vulnera", "Veris Leta Facies"…), the author conceived it as a single scenic cantata.

[3] LIM (Laboratorio di Informatica Musicale)—DI*Co* (Dipartimento di Informatica e Comunicazione), Università degli Studi di Milano—Web site: http://www.lim.dico.unimi.it.

Fig. 2. Three photos to capture the same subject: a sword and its sheath.

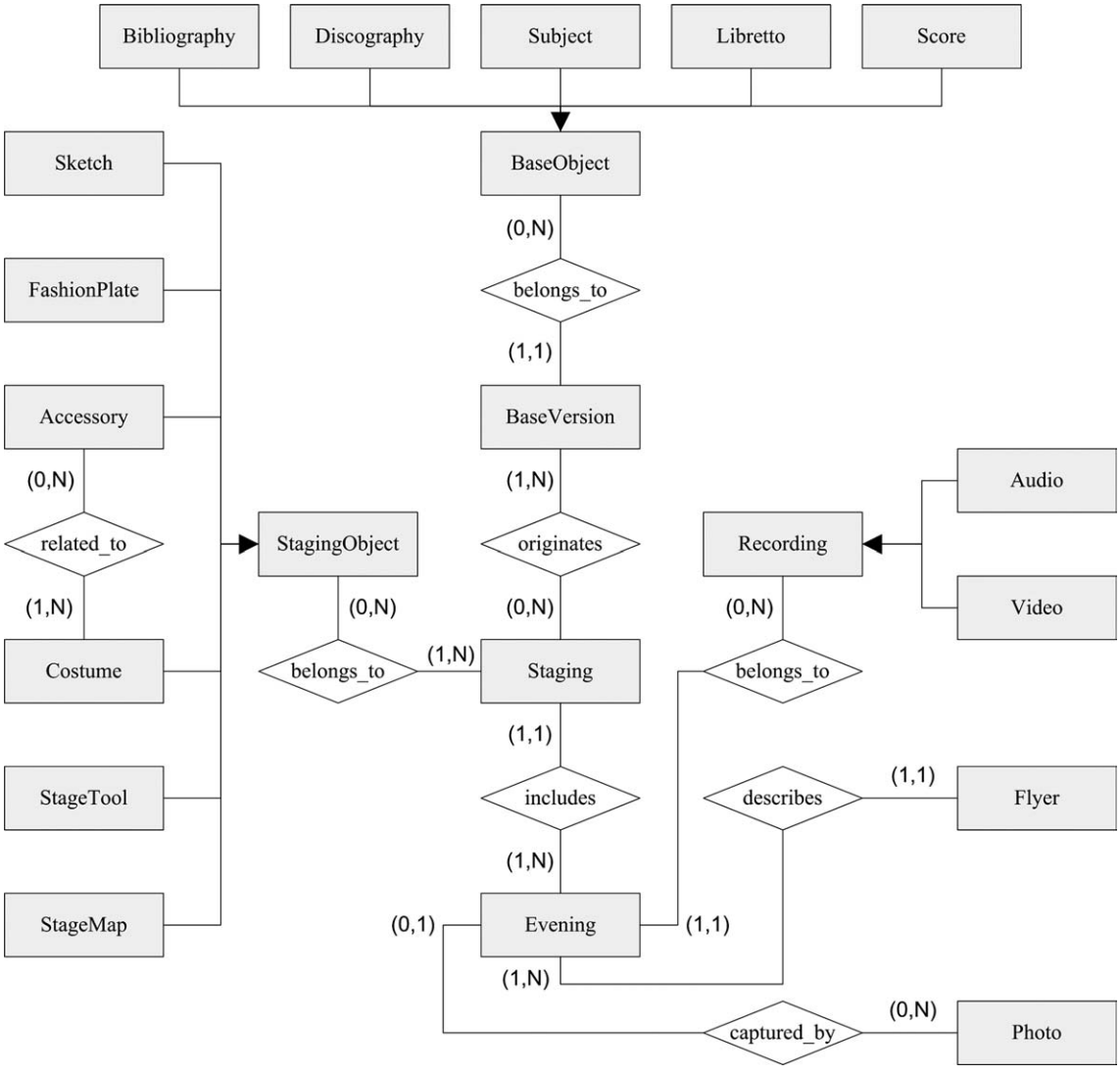

Fig. 3. The entity-relationship diagram for the proposed multimedia database.

TEATRO ALLO SCALO

Rappresentazioni
N. 52, 56, 58

CONCERTI SINFONICI 2003/2004

TEATRO PELEGRIN PAJUELO
Madrid

Abbonamento
Stagione Sinfonica

DOMENICA 29 FEBBRAIO 2004 - ORE 20 - TURNO A
MERCOLEDI 3 MARZO 2004 - ORE 20 - TURNO B
VENERDI 5 MARZO 2004 - ORE 20 - TURNO C

CORO E ORCHESTRA
DELL'UNIVERSITA' DEGLI STUDI
DI MILANO

Direttore
GOFFREDO HAUS

WOLFGANG AMADEUS MOZART
SINFONIA N. 35 IN RE MAGG. K 385 "HAFFNER"
Allegro con spirito
Andante
Menuetto
Presto

CARL ORFF
CARMINA BURANA
Cantiones profanae per soli, coro e orchestra
Fortuna Imperatrix Mundi
I. Primo Vere
Uf dem Anger
II. In Taberna
III. Cour d'amour
Blanziflor et Helena
Fortuna Imperatrix Mundi

Maestro del Coro
Luca A. Ludovico
Coro di voci bianche dell'Università degli Studi di Milano diretto dal M° Adriano Baratè

IL CONCERTO SARÀ TRASMESSO IN DIRETTA STEREOFONICA DA RAI RADIOTRE DOMENICA 29 FEBBRAIO

PREZZI (Tasse comprese)
Platea € 50,00 - € 40,00 - € 30,00 - € 25,00 - € 20,00 - Gallerie € 30,00 - € 25,00 - € 20,00 - € 15,00 - € 10,00 - € 5,00
Sui prezzi dei biglietti dei posti prenotati o acquistati tramite il sistema RossBITS si applica la maggiorazione del 20% come servizio di prenotazione.
Sui prezzi dei biglietti acquistati direttamente agli sportelli della Biglietteria Centrale a partire da un mese prima e fino al giorno precedente la recita si applica la maggiorazione del 10% come servizio di prevendita.
A termine di legge è vietato, durante lo spettacolo effettuare anche parzialmente riprese filmate o registrazioni e scattare fotografie in sala o nei ridotti. Il Teatro aprirà 45 minuti prima dell'inizio dello spettacolo.

Impaginazione e stampa LEVA S.p.A. Arti Grafiche - Viale Edison, 605 - Sesto S. Giovanni - Milano

Fig. 4. An example of "composite" flier: the same playbill describes three evenings.

Let us briefly comment on some peculiar aspects of our proposal, that is the result of our most recent efforts and an evolution of [1].

In Fig. 3, the skeleton of database structure is evident: the framework is composed by the three key concepts listed before (namely BaseVersion, Staging, and Evening) together with their relationships.

Music score as well as a number of textual documents (e.g. librettos) and data (e.g. authors, title, publication date) can be associated to the base version. Each sub-entity presents its own metadata, and these peculiarities will be considered when implementing the database, but from a theoretical point of view they all can be grouped under the concept of BaseObject. Some minor problems can arise: for example, various autographic score versions or different titles for the same work. Fortunately, those problems can be easily solved thanks to a 1:N instead of a 1:1 relationship.

According to our proposal, sketches, fashion plates, costumes, stage maps and equipment are naturally related not to a base version, but to a staging. In fact, the choice of the stage material to use depends on the dictates of the director, of the costume designer, of the scene painter and so on[4]. Once again, even if the aforementioned objects are usually related to a particular staging, we cannot ignore common events as staging revivals or "contaminations" among different works whose production employs the same equipment. This is the reason why in Fig. 3 the theoretical 1:N cardinality between Staging and StagingObject entities becomes a M:N relationship. In other words, our database structure supports the following situations: for a given staging many objects can be used, as well as the same object can be employed in different productions.

Considering the sub-entities grouped by the concept of StagingObject, two of them present a particular relationship: Costume and Accessory. The latter category includes, for example, jewelry, wigs, and shoes. The relationship cardinality is justified by the following considerations: some costumes don't have any accessory, or perhaps accessories are stored and clas-

[4] To tell the truth, sometimes the authors annotate precise indications on their scores as regards scenery, stage layouts, movements of the artistic masses.

sified together with their costumes; on the contrary, in our opinion each accessory must be related to some costume.

By examining Fig. 3, an asymmetry becomes soon evident. BaseVersion is related to BaseObject, as well as Staging is related to StagingObject, but Evening is not related to a unique EveningObject entity. The reason resides in the different cardinalities we have to manage for the relationships involving Evening. In fact, audio and video recordings are necessarily related to one and only one evening; on the contrary, a flier is usually related to a single show, but it can describe also many evenings that share the same music program and the same performers (Fig. 4); finally, photos can be even unrelated to any evening: this is the case of pictures taken during non-musical events or depicting private moments of artists' life.

6. Conclusions

An effective and efficient multimedia database presents many advantages for opera houses. From the point of view of the theatre establishment, it makes management operations and editorial initiatives much simpler.

However, such database can assume a very important role also from the cultural heritage perspective. In fact, thanks to computer-based technologies, database contents can be easily transmitted, both in space and in time dimension.

A digital version of the original material can be handed on from generation to generation. In a certain sense, not only data and metadata, but also physical objects can be transmitted, after ad hoc digitalization campaigns.

Besides, thanks to network technologies, a worldwide access to database contents can be granted to any authorized user. On-line publication of multimedia contents requires a well-structured interface, containing effective navigation tools and appropriate forms of data aggregation. However, as a result of this effort, researchers, performers, music students, opera fans and simple keens could access such cultural treasure, anytime and anywhere.

Acknowledgments

The authors wish to acknowledge the partial support of this project by Italian MIUR (FIRB "Web-Minds" project N. RBNE01WEJT_005) and the Italian National Research Council, in the framework of the research program "Methodologies, techniques, and computer tools for the preservation, the structural organization, and the intelligent query of musical audio archives stored on heterogeneous magnetic media", Finalized Project "Cultural Heritage", Subproject 3, Topic 3.2, Subtopic 3.2.2, Target 3.2.1.

References

[1] E. Ferrari, G. Haus, The Multimedia ORDBMS Architecture of the Musical Archive at Teatro alla Scala, Proc. of the IEEE 1999 Multimedia Conference, Firenze.

[2] G. Haus, M.L. Pelegrin, Music processing technologies for rescuing music archives at teatro alla Scala and Bolshoi theatre, Journal of New Music Research, Swets and Zeilinger, Amsterdam, 2001.

[3] L. Diana, E. Ferrari, E. Haus E., Saving the Multimedia Musical Heritage of Teatro alla Scala for Querying in a Web-Oriented Environment, Proc. of the IEEE 2001 Wedelmusic Conference, Florence, 2001.

[4] G. Haus, in: Rescuing La Scala's Audio Archives, IEEE Computer, 31(3), IEEE CS Press, Washington, 1998, pp. 88–89.

[5] G. Haus, A. Paccagnini, M.L. Pelegrin, Characterization of Music Archives' Contents. A Case Study: the Archive at Teatro alla Scala, Atti del XII° Colloquio di Informatica Musicale, AIMI/Università di Udine, Gorizia, 1998.

[6] G. Haus, A. Paccagnini, M.L. Pelegrin, "Characterization of Music Archives' Contents. A Case Study: the Archive at Teatro alla Scala", Proc. of the 3rd International Congress on Science and Technology for the Safeguard of Cultural Heritage in the Mediterranean Basin, Alcalà de Henares, Spain, 2001.