A Study on Audiovisual Metadata

Duan Minglian, Yao Xingxing, Zhang Jiuzhen

Department of Information Management, Peking University, Beijing, China 100871

Tel: 86-10-62883694; 86-10-51605026; 86-10-62757981

Mail: duanminglian@china.com, xingxing_yao@digiark.net,
zhangjiuzhen@163.com

Abstract: Audiovisual resources are expanding rapidly these days. Large amounts of these complicated materials could lead to serious problems in resource description, representation and organization. Supported by the National Social Science Foundation of China, the project entitled *Innovative Study on Audiovisual Metadata and its Retrieval* has been conducted since 2002. During the preparation phase, the status of audiovisual information resources and their characteristics were analyzed; general users' and managers' needs were examined, and existing metadata methods and projects on audiovisual resources were investigated and explored. Based on those constructive and helpful findings, the *Audiovisual Metadata Set* was developed. This paper introduces the principles of the Set's design and provides an overview of its content. The work is still in progress and the discussions presented here are preliminary in nature. A great many tasks are foreseen and will need to be undertaken, to establish this Set.

Keywords: Audiovisual resources; Metadata; Audiovisual metadata; DC Metadata

1 Introduction

The digital era brings all kinds of "born digital" products, among which audiovisual resources are continually increasing. Videos, tape recordings and static pictures are familiar to almost all of us. Videos today usually synthesize text, audio and moving pictures together. Television stations, publishing companies and other organizations produce huge numbers of videos each year. Efficient representation and organization of these complicated resources is necessary for end users and managers. One basic pre-requisite for managing audiovisual resources is the development of a metadata set suitable to this special area with users having different levels of needs.

Supported by the National Social Science Foundation of China, the project entitled *Innovative Study on Audiovisual Metadata and its Retrieval* has been in progress since March 2002, at the Department of Information Management, Peking University. This paper is based on this project and focuses specifically on video metadata.

2 Study design

Since the object of this project is to develop a metadata set for audiovisual resources and apply it in practical system, we began the project with three preparation studies:

First, the status of audiovisual information resources and their characteristics were collected by a web survey and then analyzed to substantiate our objects.

Second, general users' and managers' needs were examined carefully through a questionnaire survey.

Third, existing metadata methods and projects on audiovisual resources were investigated and explored to get an overview.

Constructive and helpful findings were

gained through the preparation phase and sections 3, 4 and 5 detail the main findings in turn.

Then the *Audiovisual Metadata Set* was produced and modified, based on the Dublin Core Metadata Standard (DC). Section 6 provides an overview of this set and its content focuses on video resources. Our *Audiovisual Metadata Set* also includes another two special sections for tape recordings and static pictures, which are omitted here for length limitation. The full analysis of this material can be found in our project report.

3 Characteristics of Audiovisual resources

Different media types mean different characteristics of information resources in format, storage environment, and reading equipment, to name a few. The media types of audiovisual resources include text, image, audio and moving picture. Multimedia attributes and audiovisual resources have the following characteristics:

3.1 Large Amounts, Rapid Increase

In recent years, the publishing volume of audiovisual resources in China has greatly increased, its market proportion showing an obvious rise. "In 2001, the production of videotape in China fell to 648784(1017 different kinds), from comparison of the 905150(2473 kinds) in 1999. The amount of DVD-V's rose to 2,531,762(328 kinds) in 2001, from the 41,946(32 kinds) in 1999. The same thing happened with VCD, they increased from 634,190,900,000_7,252 kinds_in 1999 to 1,407,874,610,000_10,100 kinds_in 2001."[1] The increasing amount of recordings from TV stations is also surprising. For example, the annual incremental amount of news videos at Shandong Province TV Station is 1100 videotapes, and that of its comprehensive programs is 4200 videotapes.

3.2 Various Types, Different Formats

The storage medium of audiovisual resources is quite complicated. DVD, LD, VCD, videotapes of many specifications, CD, recordings, audiotapes, and magnetic tapes are the most commonly used storage mediums. Meanwhile DVW, MII, DVC, SX, DS, digital streaming tapes and hard disk are also in the market. The content on these storage medias can be made into different compression formats according to its purpose, such as TIFF, GIF, JFIF, and PICT.

3.3 Widely Distributed Collecting Organizations, Large Holdings

With the rapid increase in production and user's demand, audiovisual resources are being collected by many organizations. According to statistics from the project *How Much Information* which was performed by the School of Information Management and Systems at UC Berkeley, the stock volume of films in the world in 2001 was 368,530, which is 781 PB[*] in bytes, the amount of pictures is 900 billion, the equivalent of 4500 PB. That of magnetic tapes is 1000 PB; analog videotapes 6400 PB, and with copies included the total may amount to 62,351 PB. Digital tapes' stock is 1250 PB; CDs and DVDs exceed 1000 PB[2].

Recently the number of organizations in China that have begun to collect audiovisual resources has increased. Not only do TV stations and film archives have large holdings, but many libraries collections have also been increasing.

3.4 Special Characteristics, Serious Information Organization Problems

Compared to other kinds of information resources, creators and physical characteristics for audiovisual resources are very complicated, especially for film and video recordings.

---

[*] Petabyte(PB): 1,000,000,000,000,000 bytes OR 1015 bytes. All printed material is about 200 Petabytes

Video recordings incorporate such information as title, creator (such as actor, singer, player, speaker, director, etc.), and place of manufacture, name of manufacturer, date of manufacture, playing time, and system.

Video recordings from TV stations are even more complicated. These materials have not only titles of columns, but also sub-columns. Moreover, some sub-columns have a special title for each program on a certain date. For example, China Central TV has a column named "Oriental Space Time", which has three sub-columns respectively named "Oriental People", "Live Broadcast" and "Stories of Grassroots". Below titles of every sub-column, another unique title is given for every program. As for physical characteristics, homemade videotapes at TV stations are more complicated than those for market. They have such information as playing time, start point, end point, system, aspect ratio, sound characteristics, color, projection speed, playing speed and so on. As for dates, there are copyright validity dates, first playing dates, last playing dates, license time and manufacture dates.

While audiovisual information resources are increasing sharply, the general metadata standards available are not appropriate for today's audiovisual information resources. Cataloging rules and metadata standards for audiovisual resources are urgently needed now. This problem is a big obstacle to search and retrieval. It also restricts the film and TV industry in China. Leading all these problems is the concern of how to describe and present audiovisual information resources so as to enhance their use efficiency, which is a hot topic for information workers and researchers at home and abroad.

4 Users' Needs

User-centered ideology is popular in current research. In our project, gathering users' needs is also of great consequence.

We issued 600 questionnaires (300 organizations' questionnaires were sent to main TV stations, libraries and information centers and so on; 300 individual users' questionnaires were issued randomly), and received 150 responses. 33 organizations' feedbacks were less than anticipated, and were from organizations including libraries, information centers and TV stations. 117 responses were from individuals. Of the organizations, 13 (39.4%) were from libraries and information centers; 20 (60.6%) were from TV stations of the national and provincial level. Of the 117 individual users' responses, 60 (51.3%) were from common users; and 57 (48.7%) were from professional users. [1] Based on the survey, we finished the report entitled *Survey and Analysis on Audiovisual Information Users in China*. In that report, users' background, current conditions, information search behavior and retrieval needs, as well as information management and service of surveyed organizations, were respectively investigated and analyzed. This paper summarizes users' needs as the following:

a. Most users are young to middle-aged people, with higher level of educational background;

b. Users' needs vary distinctively. The objective resources and users' work or hobbies are closely related. General users mainly satisfy their personal interests and hobbies with audiovisual information resources; while professional users' retrieval mainly focuses on their work, which requires more accurate and deeper descriptions.

c. To satisfy special demands of different users, different organizations

provide entirely different services. Collectors, such as the Audio-visual Education Library, the AV department of information centers, the archival units of TV stations, and film archive centers, are mainly aimed at professional users while libraries aim at general users. The former have special management needs for audiovisual resources.

5 Review of existing metadata and related projects in China

Audiovisual resources have attracted increasing attention since the seventies. The IFLA set up a Round Table on Audiovisual Material in 1973, (which was renamed the Audiovisual and Multimedia Section in 1999) and many studies on audiovisual resources and related areas have taken place in many countries. Some overseas influential projects, such as the Informedia Project[3], the Audio-Visual Prototyping Project[4], Open Video Project[5], Categories for the Description of Works of Art[6], the Picture Australia [7] and the Multimedia International Digital Libraries, to name a few, were investigated and summarized in other papers. This section gives a brief introduction of the current research in China.

Collection organizations in China are mainly archival units of TV stations, AV departments of information centers and libraries (especially large public libraries and some academic libraries). Important research on audiovisual information resources is being done by these organizations too. In academic libraries, these collections are mainly distributed in music libraries, art libraries, as well as broadcast and TV libraries.

Some libraries in China (for example, Peking University Library and Shanghai Library) use CNMARC to represent audiovisual information resources. The National Library of China and the Guangzhou Library of Guangdong Province had even developed special management systems for audiovisual information by themselves. As far as we know, their metadata framework is very simple (it includes some basic information such as title, creator, publishing information, price, and type.) and as such it cannot sufficiently describe complicated audiovisual resources.

Many TV stations have also developed their own audiovisual information management systems, but the lack of a general and standardized AV metadata schema has become a big problem in this area.

This being the case, the Ministry of Broadcast, Film and Television in China began research on AV metadata in 2001. On August 18, 2003, the Cataloging Specification for Broadcast and Television AV Materials was issued. Cataloging objects were divided into four levels: level of program, level of segment, level of scene, level of shot. With reference to DC metadata, 15 elements were set up for level of program; 12 for level of segment; 6 respectively for level of scene and level of shot.

At the same time, some projects were in progress in Taiwan, China. The Art University of Taibei, The Information Science Institute and the Computer Center of the Central Academe carried out a project named "the Taiwan electronic AV digital museum in social science and humanities". In the early stages, the project focused on professional 16mm films and BETACAM videotapes, which was aimed at offering a free and globally shared electronic AV database. First, they will develop a metadata system and chose ECHO_European Chronicles On-line_as the standard and revised it according to certain conditions. Now they

have brought forward "the General view on Digital AV Material Elements".

In general, audiovisual resources metadata research in China is still in its early stages. Besides the lack of uniform cataloging standards, subject indexing rules and classification rules, the lack of a standardized metadata elements set is the chief problem. Our project smoothly marched into the next phase: developing the set of audiovisual metadata.

6 Our Study on Audiovisual Metadata Set

6.1 Principles of Design

Considering the current state of the audiovisual metadata field, our design principles were established first as following:

a. User's Needs Principle. To satisfy both the audience and the manager's (librarians, administrators) needs, the general user and the professional's needs is our ultimate goal.

b. Simplicity and Operability Principle. This principle relates to the practicality and ease of use that is one of the most attractive features of DC. Compared with complex MARC metadata, DC's simplicity makes it more practical and easy to use.

c. Breadth and Depth Principle. While multiform audiovisual resources take on varied characters, general user and professional's needs also vary greatly. Thus different breadth and depth of representation is required according to the different materials. As for metadata elements and qualifiers, there should be optional combinations of them, which are suitable for both macroscopical audiovisual resources (such as videotapes, recording tapes and whole compact disk) and microcosmic ones (for example, segments or scenes of a video). This principle makes it possible for

efficient retrieval, and helps users choose and utilize it, by revealing the multi-layers of material..

d. Open and Extensibility Principle. So far one metadata set can only represent limited audiovisual resources. Some special and professional needs have to be considered and met separately, by choosing given elements and qualifiers or supplying new ones ipso facto. Hereby our set should be open and extensible to permit this kind of expansion.

e. Interoperability Principle. "Metadata's interoperability embodies on its support to interoperability among distinct architectural systems." [8] Our set should be expedient for its own application systems and for other organizations' employments. Adopting DC[9] in both its semantic and metadata structure makes this possible to a certain extent.

6.2 Description layers of audiovisual resources

Description of audiovisual resources is divided into three layers: the collective layer, the individual layer and the analytic layer (illustrated in **Figure 1**). Just as traditional series/sub series, monograph and piece-analytic, a series of videotapes is represented as a collection layer; one separate tape (story unit) of the series is represented as individual layers and scenes of that tape are analytic layers.

Videotapes that are published differently have different description layers. In the case of TV programs, the analytic layer can be a certain segment of one program from a TV play series, or a singing performance of a music evening.

6.3 Audiovisual Metadata Set

Referenced to DC elements and qualifiers, our *Audiovisual Metadata Set* presents 15 elements and 84 qualifiers altogether. The elements are delineated by 9 attributes

according to the ISO/IEC 11179 specification and Standardization of Data Elements; those attributes are: Name, Identifier, Version, Language, Definition, Obligation, Datatype, Maximum Occurrence and Comment. Since the full delineation of the specification is too large to present here, we will give an overview by listing 15 elements and the qualifiers used for videos as following:

Elements: Title, Creator, Subject and Keywords, Description, Publisher, Contributor, Date, Resource Type, Format, Resource Identifier, Source, Language, Relation, Rights Management and Physical Description. These elements are the same for static pictures, audio information and videos. Compared to DC elements, our set adds "Physical Description" while it takes "Coverage" out, keeping the numbers of elements equal.

Figure 1. Description layers of audiovisual resources
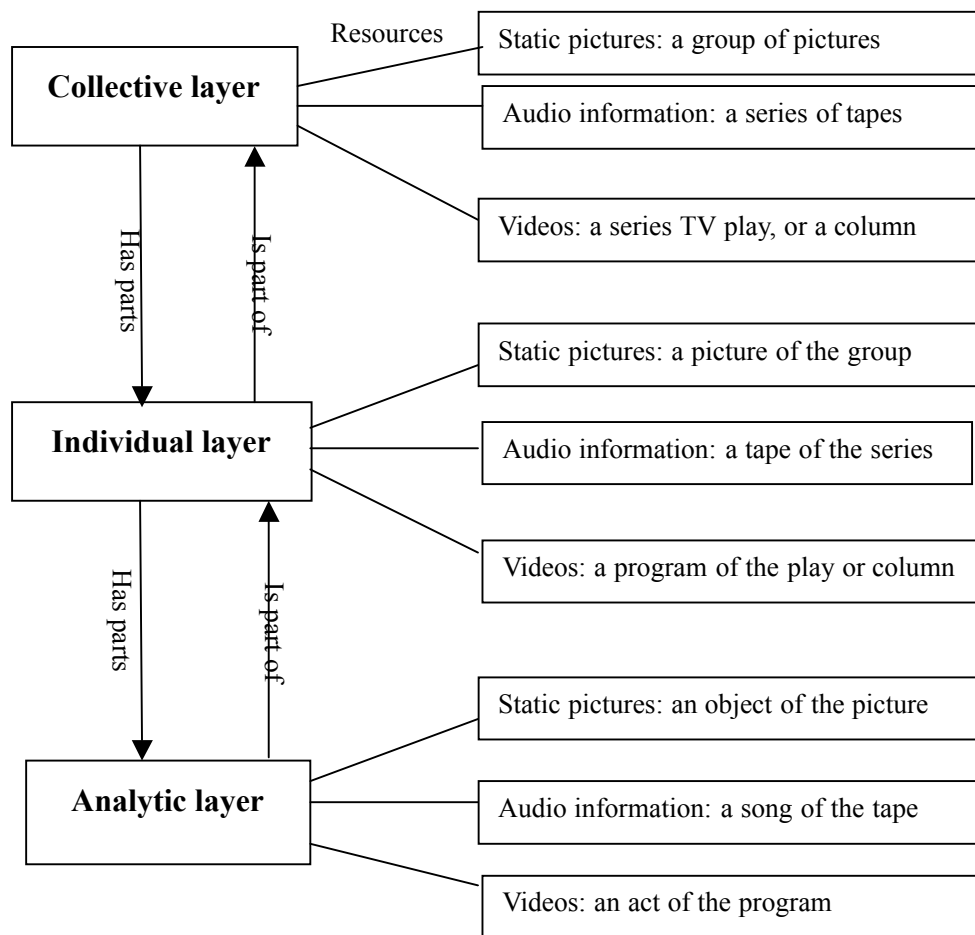
Qualifiers are designed for different audiovisual materials, and are classified as element-qualifiers and code-system-qualifiers. There are some qualifiers with the same definitions and usages as DC Qualifiers, such as "Alternative" for the element "Title", "Coverage" for the element "Format", and so on. But more than 60 percent of Qualifiers were added or modified to suit different kinds of audiovisual materials. We'll detail those for videos below (the entire definition and usage can be found in our full project report. Qualifiers without note are all

element-qualifiers):

- Qualifiers for the element "**Description**" include: notes, abstracts, audience, awards, tableOfContents, version, colorMode, shootingPlace, cameraMotion, sceneRange, cameraAngle, placeofCollection, and holdingInstitution.
- Qualifiers for the element "**Date**" include: created, issued, modified, valid, published, manufactured, copyright, firstBroadcasted, broadcasted, shot.
- Qualifiers for the element "**Language**" contain: track, subtitle and ISO639-2 (Code System Qualifier).

- Qualifiers for the "**Rights Management**" element are: secretLevel, owner, kind, statement, user, authorizedScope, deadline, usage, and times.
- And Qualifiers for the element "**Physical Description**" consist of: extentOfItem, playingTime, startPoint, endPoint, system, aspectRatio, specialProjectionCharacteristics, soundCharacteristics, trackConfiguration, color, projectionSpeed, and size.

The following are two simple examples that illustrate the usage of the Audiovisual Metadata Set.

Example 1: An Individual Layer Record of a Video recording

Title: _____

Creator [personal]: ___, ___

Creator [personal]: ___, ___

Creator [personal]: ___, ___

Type: _____

Publisher [placeOfManufacture]: ___

Publisher [manufacturerName]: _____

Date [manufactured]: 1997-10-01

Physical Description [extentOfItem]: 1 ____

Physical Description [playingTime]: 94min.

Physical Description [color]: ___

Physical Description [startPoint]: 00: 01: 00: 00

Physical Description [endPoint]: 00: 94: 00: 00

Physical Description [system]: PAL

Language [track]: chi

Description [notes]: 1997_10_____

Description [abstracts]:
_____
_____
_____
_____
_____
_____
_____

Description [shootingPlace]: _____

Date [shot]: 1997-10-01

Description [placeOfCollection]: ___
Description [holdingInstitution]: _____
Identifier [callNumber]: S000007
Rights Management [owner]: _____
Rights Management [authorizedScope]: _____
Rights Management [deadline]: 2002-11-01_0000-00-00

Example 2: An Analytic Layer Record of the Video recording

Title: _____
Creator [personal]: ___, ___
Creator [personal]: ___, ___
Creator [personal]: ___, ___
Type: _____
Physical Description [startPoint]: 00: 36: 52: 00
Physical Description [endPoint]: 00: 40: 39: 00
Description [abstracts]: _____
Description [shootingPlace]: _____
Date [shot]: 1997-10-01
Relation [isPartOf]: _____
Description [placeOfCollection]: ___
Description [holdingInstitution]: _____
Identifier [callNumber]: S000007

7 Conclusion and Future works

The *Audiovisual Metadata Set* is just the preliminary achievement of our project. The object material is fixed on three representative resources of audiovisual information: audio information, videos and static pictures. Considering the characters of different materials and the relationship between materials and their components, we defined the Elements and Qualifiers strictly according to DC; those newly added Elements and Qualifiers are pointed out distinctly to ease their application. Both professional and general users' needs can be satisfied by the flexibility of our set.

The mapping of our set to both DC and MARC (CNMARC and USMARC) is listed in our full report as a form. This lays the foundation for future interoperation among different metadata sets.

The ultra goal of our project, to satisfy both general user and professional's needs and to facilitate audiovisual resources' representation and management, requires a great deal of tasks to be undertaken. Work is still in progress and the logical framework (metadata set) that has been developed requires conversion into technical framework, which then needs to be applied in practical systems.

Some of the future works are as follows:

- To evaluate and refine the Audiovisual Metadata Set through real application in a limited scope.
- To encode the Set both with XML Document Type Definition and RDF Schema.
- To register the Set with authoritative metadata registries.
- To popularize the Set by publishing.
- To design an audiovisual resource management system and to develop the guidelines for it.
- To accelerate the efficient retrieval

mechanism of audiovisual resources.

- To realize the interoperation among different metadata sets.
- To construct a well-suited Digital video library and then offer convenient services for users.

Acknowledgements

We would like to express our gratitude to all those who have helped us in one way or another during the writing and editing stages of this paper. We are especially grateful to our project team members: Shen Zhenghua, Zhang Haoda, Gao Wei, Li Dongxiu and Wei Chengguan.

References

[1] _____.
_____.
_____ 48_1__48-52_77
[2] Lyman P, Varian H R. How much information 2003? [2003-10-06].
http://www.sims.berkeley.edu/research/projects/how-much-info-2003/
[1] _____.
_____.
_____ 48_1__48-52_77
[3] Wactlar H. Informedia II Digital Video Library: Auto Summarization and Visualization Across Multiple Video Documents and Libraries. [2003-10-16].
http://www.informedia.cs.cmu.edu/dli2/index.html
[4] Library of Congress. Digital Audio-Visual Preservation Prototyping Project. [2003-10-16].
http://lcweb.loc. gov/rr/ mopic/avprot/
[5] NISO Z39.87-2002 / AIIM 20-2002 Data Dictionary – Technical Metadata for Digital Still Images. [2003-10-16].
http://www.niso.org/tandards/resources/Z39_87_trial_use.pdf
[6] Baca M, Harpring P. Categories for the Description of Works of Art. [2003-10-06].
http://www.getty.edu/research/conducting_research/standards/cdwa/
[7] National Library of Australia. The Picture Australia. [2003-10-06]. http://www.pictureaustralia.org www.pictureaustralia.org
[8] _____ ____.
_____ _____.
http://cdls.nstl.gov.cn/cdls2w3c/2003/SpcMetadata/docs/spcmd_manual.Doc
[9] Dublin Core Metadata Element Set, Version 1.1: Reference Description. 2003-06-02.
http://dublincore.org/documents/2003/06/02/dces