Multimedia Data Management

Using Metadata to Integrate and Apply Digital Media

Chapter

1

Overview on Using Metadata to Manage Multimedia Data

Susanne Boll, Wolfgang Klas, and Amit Sheth*

Department of Database and Information Systems,
Faculty of Computer Science, University Ulm,
Oberer Eselsberg, D-89069 Ulm, Germany
{boll,klas}@informatik.uni-ulm.de

**Large Scale Distributed Information Systems Lab,*
Department of Computer Science, University of Georgia,
415 Graduate Studies Research Center
Athens GA 30602-7404, USA
amit@om.cs.uga.edu

Abstract

This chapter introduces the reader to the key role that metadata can play for managing various digital media, often also called new media, such as image, audio, and video in information systems. We present different views on how metadata can be modeled, classified, extracted, managed, and applied, to support a convenient handling of digital media. We discuss various issues on the role of metadata from different viewpoints: the application scenario perspective, the media processing perspective, and the metadata type and classification perspective. We also look at the world of standards for handling metadata.

Finally, the chapter gives an outline of this book and provides a coherent view over all the chapters. Each of the chapters and its main issue in the field is introduced in brief with respect to the concepts and issues discussed in this chapter.

1.1 The Need for Metadata in Multimedia Systems

Real applications in industry face the serious problem of how to handle large amounts of a variety of digital media. Therefore, there is a growing demand for database and information systems support in the area of modeling, management, and processing digital media. There is a need to explicitly capture a fair amount of content-information as well as application-specific semantics by means of a variety of *metadata*, e.g., multimedia indexes, attribute-based annotations, and intentional descriptions, to allow appropriate access to, selection of, and processing of digital media.

One of the key problems to be solved is the development of metadata, that is, the generation, structuring, representation, management, and proper utilization of data or information about data. While the issue of metadata has received a fair share of attention in conjunction with structured data (e.g., biomolecular data) and text documents, most of the current practices in the context of digital media and multimedia data management are still quite ad hoc.

Metadata plays a far more important role in managing multimedia data than does the management of traditional (well-)structured data or information retrieval techniques applied to text-only data. Some of the reasons for this are:

Different query paradigm. The exact-match paradigm for querying is no longer suitable or adequate for querying or retrieving various types of digital media.

Inadequate processing techniques. Content-based processing techniques are too hard to analyze, and very large data sets are often limited or inadequate.

Lacking efficiency. When a content-based search is possible, it cannot be used very frequently (e.g., for every query) for performance reasons and because of varying application- and modality-specific search criteria.

Semantics of multimedia data. Derived data and interpreted data (which may be considered a part of the metadata) as well as context and semantics (which may be easier to base on metadata rather than raw data) are of greater value when dealing with multimedia data or new/digital media (especially audio-visual data).

Let us discuss a couple of items listed above in a little more detail. Various digital media or components of multimedia data involve very large raw data volume. This has consequences on effective management and retrieval of the digital media. Content-based retrieval on raw data means that the query capabilities are limited to the number of available matching algorithms. Performance is lacking when queries are executed on large data sets. Indirect retrieval and processing, however, that use abstract information or metadata of the digital media seem to be a promising approach to enhance querying and processing and to improve response time as metadata will be of much less volume than the digital media themselves.

Semantics of multimedia data like images, video, and audio is implicit to the raw media data. By analyzing and processing, semantics can be made explicit to some extent on different abstraction levels, from feature values to knowledge-based concepts. Metadata describing this semantics explicitly may still not be sufficient for exact-match querying. For example, color distribution feature values of an image for red, black, and yellow still do not allow the conclusion that the image shows a sunset. The user's term *bar* may be ambiguous as it may match the concepts *barrier* or *pub* or *legal profession* dependent on the user's idea. A different querying paradigm is needed to allow a proper mapping of the user's ideas to the explicitly available semantics of a media object and, finally, to the raw data.

Metadata that is derived or extracted from digital media have the additional advantages of being more amenable to traditional data retrieval and manipulation techniques than the raw digital media. Examples include text annotations of images or other media or attribute-based metadata that can be managed by a structured database. These can then be easily applied or analyzed using text processing or information retrieval techniques and database management systems, respectively. For fuller exploitation of data and to provide information system interoperability it is necessary to support correlation of data. It is often more convenient and useful to identify and specify correlation at the metadata level rather than between raw data items.

In many situations metadata, as it is traditionally defined, is not sufficient for managing digital media and multimedia data. In the case of structured databases, the norm is to use schema descriptions and associated information (such as database statistics) as metadata. In the case of unstructured textual data and information retrieval, metadata is generally limited to indexes and textual descriptions of data. Richer forms of information, as conveyed by context, ontologies, and semantics, are very important here [8]. Metadata in such cases provide a suitable basis for building the higher forms of information.

With this book we try to look at the problems of *generating/extracting*, *structuring*, *representing*, *storing/organizing*, and *using* metadata for various digital media and multimedia data as well as their applications. We look at metadata for images, audio and speech, video, structured documents, and geo-spatial data and finally at the role of metadata in multimedia documents stored in digital libraries and in mixed-media-based information access.

For each medium, we address a variety of issues and try to give answers to questions like:

- What are the characteristics of the media type considered, e.g., text, image, voice, and video, and the application domains that provide the context for metadata-based exploitation of the digital media?
- What are typical types and examples of metadata in the context of a media type, e.g., abstractions from raw video data?
- What are the content, the reference terms, the ontology of the metadata?
- What are the strategies and techniques used for manually and/or automatically generating/extracting and maintaining metadata?
- How is metadata structured, and what kind of language is used to describe the structure?
- What is the relationship between data and metadata (e.g., size) and does this induce any implications in handling metadata?
- How is metadata stored and organized, i.e., which techniques are used to represent metadata in, for example, databases?
- How does metadata facilitate information discovery and retrieval as well as other processing of digital media?
- Are there metadata standards related to the digital media or application domain of concern and what role do standards play?

Next, we present three perspectives that govern the management of digital media and multimedia data. These perspectives evolve from or are supported by the chapters of this book, and many of the examples given are based on these chapters. The issues outlined above are addressed in the context of these perspectives, i.e., particular application scenarios, media processing techniques, and the characteristics of metadata.

1.2 The Application Scenario Perspective

We briefly discuss in the following the application areas covered by the subsequent chapters in this book with respect to metadata issues.

Image retrieval, navigation, and browsing in collections of images. More and more application areas, such as medicine, maintain large collections of digital images. Efficient mechanisms to efficiently browse and navigate through the collections, however, are still lacking. Semantic content-based image browsing and navigation are needed instead of searching and viewing directory trees for image files such as `apwr38.gif`. An important issue is to extract images according to the user's (semantic) association and impression of an image, e.g., sunset at the sea, and not only according to mere (syntactic or structural) image features such as the color or texture. For the retrieval of images, a suitable definition for semantic equality and similarity of images is also needed. To be able to support the mapping from the users' ideas down to the raw image data, a model describing the association between users' concepts and image characteristics and semantics is needed (see for example Chapter 7).

Advances in techniques for obtaining images of the body's interior have greatly improved medical diagnosis. New imaging methods include various X-ray systems, computerized tomography, and magnetic resonance. The introduction of computerized tomography (CT) was a major advance in visualizing almost all parts of the body, particularly useful in diagnosing tumors and other space-occupying lesions. These new techniques lead to accumulation of masses of digital medical images stored in medical image archives.

For satisfying diagnosis, however, it is not sufficient to store and access a patient's CT images with the patient's record-id. Rather, suitable querying mechanisms are needed for a useful employment of the images in medical diagnosis (see Chapter 6). The questions of a surgeon to a medical image archival may be: How does my patient's tumor look compared to similar cases of brain tumors? What is the normal growth rate of a special type of brain tumor? Does the spatial growth of a brain tumor decrease with a certain drug therapy?

The images themselves do not give hints about whether they show a brain tumor or where it is located in the body. Therefore the knowledge of the spatial content of the images and the evolutionary behavior of the spatial content for a medical image (e.g., for a brain tumor) must be used or made available when processing a surgeon's queries. The result of a query should then be a collection of images that have similar spatial characteristics compared to a given image or a sequence of images showing the growth of a brain tumor over a year's time.

Searching image collections can be employed to find a starting point in a "web" of images from which the user may want to start a navigation through images and related information.

After having selected a particular image of interest, navigating through an image collection can take place, e.g., by choosing a particular part of the image that is currently being displayed. This selection can lead to associated data such as a set of related images or some other related textual information. One may also navigate in the (hyper-)textual information and may come back to the image collection via special links/hooks in the text and find an image associated with the respective textual information.

For example, an image of a person comprises various regions having semantic content, e.g., the various subregions that correspond to the eyes, lips, and nose. When viewing a media object, the related information can be investigated for learning, e.g. which person can be seen on the image. Additionally an information *location* associated with a person's image can lead to an associated building and room of the location and then "finally" to the image of the person's office (see for example Chapter 5).

Navigating image collections might also involve navigating three-dimensional (3-D) representations, e.g., of the body's interior. A sequence of CT images can be the basis of a computed 3-D graphics representation of the brain. A surgeon may navigate through this representation of the brain. She/he may select a particular volume of interest, the thalamus, and enter it, viewing it at a higher resolution to see whether there is a growth inside. The surgeon may also select a part of the 3-D representation inside a thalamus that allows him/her to view photographs of patients with similar growth, etc.

This kind of support for image retrieval, navigation, and browsing requires a lot of semantic knowledge which can be represented by means of metadata and which can be used by the retrieval, navigation, and browsing algorithms.

Video. In many domains video clips are archived digitally, e.g., in news agencies. Besides the archiving of the digitized video clips, an important issue is to browse through a collection of videos and select them either entirely or partially. A difference between searching, browsing, and navigating in videos in contrast to images is the temporal aspect. The abstract information that is added to video to support retrieval can change within the video depending on the part of the video. Furthermore, querying against a video database results in a sequence of video clips, each of which is time-dependent.

For example, nowadays we have to watch a provider's news and cannot eliminate those news items we are not interested in. Personalized news (for an example see Chapter 10), cut to special personal interest, will make a news watcher independent of the news and of the time the news is actually on air. According to a user profile, videos are searched, and those parts of the present news items are selected that fit a questioner's need. With semantic knowledge about the structure of news, newly assembled and temporally arranged news items can be composed to form a personalized news extract. The interesting issue is how to define such a user profile and how thousands of news items of a news provider can be attached metadata that in combination with a user profile allow for a satisfying mapping between the two and the successful reassembling of the personalized news.

A similar application scenario can be derived from the demand of a critic who only wants to watch those parts of a film that suffice to write a quick review of the film or the special demand of a sport enthusiast who has only time to see a sequence of all field goals of a certain football game of a certain team in order to be able to talk about the game the next day or the post-game analysis of football teams to support planning of strategies and analyze performance (for an example see Chapter 9).

Audio and speech. Radio stations collect many if not all of their important and informative programs such as radio news in archives. Often it is of interest to reuse or to refer to parts of such programs in other radio broadcasts. However, to efficiently retrieve parts of radio programs it is necessary to have the right metadata generated from and associated with the audio recordings (see for example Chapter 8). This asks for retrieval of audio that contains spoken text. One important issue here is the detection of text in the audio, i.e., speech recognition. Here problems that arise in speech recognition because of different pronounciation of words by different speakers and language peculiarities must be overcome. Another important issue is the mapping between a high-level vague query, like a textual or a query containing spoken text, to the metadata attached with the audio recordings. This calls for an organization of the metadata to support efficient query evaluation and for a query evaluation model that determines those recordings in an archive that are relevant to a user's query.

Structured document management. As the publishing paradigm is shifting from popular desktop publishing to database-driven publishing, processing of structured documents becomes more and more important. Interesting issues are the description of document structure and layout, structure and content-oriented retrieval of components of documents, full-text retrieval, presentation of document content on various output channels like print media, CD-ROMs, WWW, etc. Particular document information models like SGML (Standard Generalized Markup Language) and HyTime (Hypermedia- /Time-Based Structuring Language) introduce a lot of descriptive information, i.e., metadata, on the structure and content of documents (see for example Chapter 2). Such metadata can be used during processing for improving system performance, e.g., database configuration, and document type-specific query optimization, or for providing new functionality, e.g., higher expressiveness of query statements, integrating information retrieval techniques with database functionality, and providing query templates based on document structure or layout. Metadata can be used at various system layers, e.g., at the specification layer by means of document type definitions, for the internal representation of documents and its components including storage models and indexing, for maintaining histories of processing a document, to support declarative access and query processing, etc.

Metadata about structure can be used by the author of a news article to retrieve interesting parts of documents in a huge document archive. For a well-targeted query the document structure available via metadata can be exploited. Not only can all the documents be retrieved by the author that contain the name "Helmut Kohl" but also all documents that contain the name in their heading as this is a known structural element in the documents. Efficient retrieval is achieved by exploiting document structure as the metadata can be used for indexing, and that is essential for short query response time. A typesetter of a newspaper's title page can make use of the metadata to properly lay out the article, that is, to process the document like "Place the *title* in 18pt Helvetica at the top of the page, align the first two *paragraphs* beneath the *headline*, and let the remaining *paragraphs* follow on the next page."

Geographic and environmental information systems. Geographic and environmental information systems are used by various parties who have very special information needs. Such systems have to provide an integrated view on individual geographic and environmental data sets. Obviously, one key problem is the provision of descriptive information on the content provided to the end users and for the information system itself in order to facilitate transparent integrated access to different information sources. The problems faced in this application domain are related to some extent to the integration of heterogeneous databases (see for example Chapter 4). Approaches taken in this field deal with a significant amount of metadata for global query decomposition, global transaction management, schema integration, and management of federated information systems. Other important issues addressed are exercise of control over the degree of uncertainty and accuracy of the data. An important milestone achieved in this application domain is the availability of national and international standards for metadata frameworks (see for example Chapter 3).

Digital libraries. Digital libraries offer a wide range of services and collections of digital documents and constitute a challenging application area for the development and implementation of metadata frameworks. All digital library projects currently under development have to solve metadata issues. Most of the approaches have oriented their metadata framework toward the description of collections of digital material ranging from full text to spatially referenced data sets to multimedia material like video and audio. This allows development of a semantically rich library catalog which in turn provides for the development of rich functions for searching, retrieving, evaluating, and obtaining information. For example, the approach taken in the Stanford Digital Library project focuses on the establishment of an infrastructure for various interoperable services. Metadata therefore is used to describe services and to provide a general interoperability framework. The Alexandria Digital Library project follows the traditional library paradigm, employing many of the well-known traditional forms of metadata like thesauri and subject headings. Metadata turns out to be one of the most critical and essential ingredient for communicating information in a well-performing digital library (see Chapter 12).

Mixing various media to support information access. Access to multimedia information most often is single-media-oriented, i.e., the media type of a query is the same as of the data, or the retrieved information is of a single media type. Mixed-media access is an approach which allows for using a variety of media types in queries independent of the media type of the data (see for example Chapter 11). In other words, a single query can be used to retrieve information from data which consists of different media, or queries using different media types can be used to more precisely describe the information needs. Using word spotting techniques one can build up a time index of keywords on recorded speech (audio) or a location index of keywords on images of text. Using speaker segmentation techniques one can partition audio segments corresponding to different speakers. If additional information is available on speakers, one may be able to also identify speakers. Semantically rich information available in speech but not in text is prosodic information occurring in emphatic speech, i.e., changes in pitch, amplitude, and timing, which allows recognition of segments of speech intended to be important, at least from the speakers viewpoint. Having available these kinds of metadata one may be able to pose semantically rich queries based on a combination of the different aspects and media.

1.3 The Media Processing Perspective

Another interesting viewpoint on metadata is the kind of media processing taking place in the application frameworks. In principle, we can distinguish three types of media processing:

Generation of metadata. This kind of processing serves the purpose of generating metadata on a particular medium. It can be structured along the following lines.

- Analysis of raw material

  In many cases media objects are analyzed and metadata is generated according to the focus of the analysis. This is done very often off-line, i.e., after having recorded and stored the medium and prior to any processing which makes use of the metadata. The reason for this is the tremendous costs involved in analyzing media. Examples of this type of processing are the recognition of document structure in semi- or non-structured text, the extraction of feature values from images, segmentation of audio by speakers, recognition of scene cuts in videos, or content-based recognition of advertisement in video.

- Semi-automatic augmentation

  In contrast to the analysis of raw material, semi-automatic augmentation of media results in additional meta-information which cannot be derived from the raw material as such. It often requires comprehensive background knowledge of the context for providing additional information. Examples are the diagnostic findings of a doctor related to a computer tomography image, which are based on the doctor's experience and state of the art in medicine.

- Processing with implicit metadata generation

  The previous two types of processing are techniques for explicitly generating metadata. Alternatively, metadata can be generated implicitly when creating the raw media data. For example, a digital camera can implicitly deliver time and date for pictures or videos taken. Similarly, an SGML editor generates metadata according to the document type definition when the document is edited. In addition, processing can be guided by a *metadata creation scheme* which determines in detail, the type and organization of metadata and relationships between individual metadata items. For example, editing structured documents by using SGML editors automatically leads to metadata according to the SGML Document Type Definition (DTD) defined and used for the document. The DTD represents the *metadata creation scheme*. A similar example is the usage of specific WWW page editors which generate a lot of additional information on, for example, the page and its author according to HTML and some convention used by search engines. With the evolving new standard markup language of the Web, XML, the tendency for processing structured documents and, as a consequence, the implicit generation of metadata begins to take shape throughout the World Wide Web.

In specific application areas, many tools are already available to support generation of metadata. For example, in the field of geographic information systems a list of tools is given in [5].

Usage of metadata. By attaching metadata to media, various processing steps can conveniently use the metadata. Here, the most prominent types of processing are querying, retrieval, navigation, and browsing. These types of processing make use of metadata mostly in terms of indexing. That is, metadata is used to effectively discriminate data in order to provide for efficient processing. Furthermore, metadata can be used to allow for additional types of queries which cannot be answered by processing raw material only. This is usually the case when metadata stems from manual augmentation of media such as in the example of a doctor's diagnosis.

Maintenance of metadata. Another type of media processing affects existing, already generated and derived metadata. This usually requires updating metadata, which in turn requires explicit knowledge about the structure and semantics of metadata.

Metadata has to be updated according to changes in raw data. This may cause a new full cycle of metadata generation in order to substitute the old existing metadata on a medium. For example, if images stored in an image archive are manipulated, the feature values extracted when an image was registered in the archive may have to be recomputed. An interesting question is whether it is possible to update metadata incrementally, e.g., by performing only incremental analysis related to the modified items. For example, if the color of images changes, only color-related feature values have to be updated, but other types of analysis like the spatial placement of objects are not needed. However, incremental adding of metadata may be a problem as is known from indexing in information retrieval.

Furthermore, metadata can be updated directly without any modification of the raw material. A simple example is the correction of metadata according to the changes of semantic knowledge used for constructing the metadata, e.g., new medical findings which lead to a revision of the previous diagnosis.

1.4 The Metadata Characteristics Perspective

1.4.1 Types of Metadata

Metadata itself can be characterized and classified according to its type. In the application scenarios outlined above we can identify kinds of metadata along three dimensions: media type-specific, media processing-specific, and content-specific metadata.

Media type-specific metadata. Media types induce specific kinds of metadata, e.g., features like texture of images, frequencies in audio, font size of text, motion in video. Continuous media like audio and video have associated with them a whole lot of metadata attributes which relate to the notion of time, which static media like images and text do not have. The more specific a media type, the more specific the associated metadata attributes. For example, for speech – in comparison to audio – one can identify additional metadata items, like specific keywords present in spoken text, and speaker characteristics, like female or male, child or adult (see for example Chapter 11). In the case of video, we can have metadata that describes implementation specific properties like play-out rate, cinematographic properties like camera motion and lighting, or a time-based semantic space derived from the content of the video (see for example Chapter 10). It can be observed that media-specific metadata properties can be further differentiated along the dimensions of media processing and media content.

Media processing-specific metadata. Metadata can describe functions designed to process specific media. For example, search and retrieval functionality provided in a digital library involves metadata which relates to media objects, their contents and types, but additional metadata is intended to assist search and retrieval functions (see Chapter 12). Functions to transfer media objects over networks may also have attached metadata directing the way of transferring the media. Looking at videos (including audio) which communicate, e.g., television news, we can identify story styles, determining how media material is composed overall and how a mediator (an anchor person or reporter) interacts with the visual and aural constituents of the news message. Having such a story style explicitly identified and having it available as metadata allows further interpretation and processing of the news, e.g., a personalized filtering, diversification, and delivery of information (see Chapter 10). Another important type of media processing-specific metadata is information related to media processing performance which can be used to measure and consequently achieve desirable system performance. Similarly, meta-information about the interoperability of system components is essential to deliver the proper application functionality (see Chapter 4). In general, metadata attached to media processing functionality may be dependent on or independent of the content and type of media.

Content-specific metadata. Media objects may have associated metadata which is solely derived from the content represented by the media objects. Such information is independent of media type and media processing. It reflects the semantics of the media object in some given context. Different media objects may have identical content-specific metadata. For example, an animation of the launch of a rocket, a video and a photography on the launch of a space shuttle, and a document describing the functionality of rocket engines may share the same content-specific metadata related to the functionality of rocket engines. From this example one can also observe that the organization of metadata may be very critical. In the case of the video in the previous example, the metadata related to the functionality of rocket engines may be just a small piece of metadata associated with that video. In the case of the document this metadata could be the only one associated with the document. Obviously, granularity of metadata units and their organization in terms of, e.g., hierarchies, is an interesting issue.

Besides the characterization of metadata along the three dimensions, one can identify very specific, well-known types of metadata.

Annotations. Media objects may be annotated, which, in principle, constitutes metadata to be used further on. Annotations to a media object may be of any kind of media, although most traditional applications use only textual annotations. Annotations may relate to the media contents, the media format, processing type, temporal and spatial space associated with media objects, etc. Most often, annotations are generated as the result of an intellectual task, but they may also be generated automatically by a media analysis step.

Set of keywords. Keywords are a traditional way of providing metadata on a subject and are well known in the context of indexing. The issue with keywords is the existence of a common metric or ontology accepted by the various parties involved in the media processing. It is mandatory to apply this metric or ontology both to the identification and assignment of keywords and to the usage of keywords for processing, e.g., searching and retrieval.

Metadata on history, age, and quality of data. Metadata can include information about the history, age, and quality of media objects and/or associated metadata. The distinction between age and quality of media objects and age, history, and quality of associated metadata, i.e., a case of having metadata on metadata, is an interesting issue, but is not further discussed here.

In principle, metadata can be original information which actually existed before any media processing toward the generation of any further metadata takes place – one could also call it defacto metadata – or it can be derived through some processing. This distinction may have some influence on the further processing, e.g., some device like X-ray equipment delivers a media object in some specific format which cannot be influenced by anybody. Information about this format constitutes original metadata. If the media object is converted into some other format like DICOM or JPEG due to any subsequent processing, information about the new representation constitutes derived metadata. Medical doctors may be very sensitive to the qualitative differences of the X-ray images and may want to have access to the processing history, including the history of metadata.

We have now taken a closer look on the different types of metadata we have identified in the application scenarios. In the following we will sketch a classification of metadata.

1.4.2 Metadata Classification

Classifying metadata to get suitable abstractions aids in exploiting metadata. One of the interesting classifications appears in [2]. Another classification used in the InfoQuilt project [10] appears in [7] and is adapted by some of the chapters in this book. A summary of this classification as adapted from [8] is presented next.

Content-independent metadata. This type of metadata captures information that does not depend on the content of the document with which it is associated. Examples of this type of metadata are *location, modification-date* of a document and *type-of-sensor* used to record a photographic image. There is no information content captured by these metadata, but these might still be useful for retrieval of documents from their actual physical locations and for checking whether the information is current or not.

Content-dependent metadata. This type of metadata depends on the content of the document it is associated with. Examples of content dependent metadata are *size* of a document, *max-colors*, *number-of-rows*, and *number-of-columns* of an image. Content-dependent metadata can be further sub-divided as follows:

- **Direct content-based metadata.** This type of metadata is based directly on the contents of a document. A popular example of this is full-text indices based on the text of the documents. *Inverted tree* and *document vectors* are examples of this type of metadata.
- **Content-descriptive metadata.** This type of metadata describes the contents of a document without direct utilization of those contents. This type of metadata often involves use of knowledge or human perception/cognition. An example is denoting the fragrance of an image containing a flower. Another example of this type of metadata is textual annotations describing the contents of an image. This type of metadata comes in two flavors:
  - **Domain-independent metadata.** These metadata capture information present in the document independent of the application or subject domain of the information. Examples of these are the *C/C++ parse trees* and *HTML/SGML document type definitions.*
  - **Domain-specific metadata.** Metadata of this type is described in a manner specific to the application or subject domain of information. Issues of vocabulary become very important in this case as the terms have to be chosen in a domain-specific manner. Examples of such metadata are *relief, land-cover* from the GIS domain and *area, population* from the Census domain. In the case of structured data, the database schema is an example of such metadata. Another interesting example is *domain-specific*

ontologies, terms from which may be used as vocabulary to construct metadata specific to that domain.

1.5 The World of Metadata Standards

Standards are an important means to achieve common representation schemes and interoperability of systems, and hence can play a pivotal role in exploiting metadata. There are very many activities going on this area including

- the development of a metadata taxonomy to help structure the discourse on metadata,
- the development of ontologies related to metadata attributes, and description of data elements and domains in terms of naming, typing, classification, and semantics,
- the definition of a meta-model registry structure to achieve mappings among different meta-models, and
- the definition of generic functionality for tools for the development and operation of metadata bases.

1.5.1 Metadata Standards in the Context of Multimedia Systems

Some prominent standards related to metadata management and relevant in the context of multimedia systems are summarized next.

- The ISO 11179 standard [9] addresses the specification and standardization of registration of data elements.
- The Metadata Coalition Interchange Standard is a joint effort of industry vendors and users addressing a variety of problems and issues regarding the exchange, sharing, and management of metadata. This joint effort resulted in the Metadata Interchange Specification V 1.1 [11].
- The Meta Content Format (MCF) [14] addresses the abstraction, standardization and representation of the structures used for organizing information. In order to adequately describe information organization structures,

MCF allows objects representing entities such as people, organizations and projects to be first class citizens on the same level as files, folders and Web pages. MCF is a general purpose structure description language. In addition to the syntax and semantics, it also provides a standard vocabulary for describing common objects such as people, organizations, meetings, etc. MCF is a lingua-franca schema for integrating different information sources. It does not use or provide a standard schema. Rather it provides a framework for determining which data from one schema can be automatically and dynamically converted into data in a very different schema.

- Dublin Core Metadata Element Set [3] is a consensus which represents a simple resource description record that has the potential to provide a foundation for electronic bibliographic description in the context of on-line libraries that may improve structured access to information on the Internet and promote interoperability among disparate description models.

- Standards developed by the U.S. Federal Geographic Data Committee (FGDC), e.g., Content Standards for Digital Geospatial Metadata [4] which provides a common set of terminology and definitions for the documentation of geospatial data. The standard establishes names of data elements and groups of data elements to be used for these purposes, the definitions of these data elements and groups, and information about the values that are to be provided for the data elements.

- HL7 is a standard developed by the Health Level-7 group [15] (an ANSI accredited standards developer) whose members are hospitals, professional societies, health care industry including almost all of the major health care systems vendors, and individuals worldwide. HL7 includes specifications for medical records and information management.

- UDK, a proposal for a European Environmental Catalogue, constitutes a meta-information system and navigation tool set for collections of environmental data [12, 13]. It is already in place in Austria and Germany and under evaluation in many other European countries.

- Product Data Exchange using STEP (PDES) is an American National Standard (ANS), developed by the U.S. PRO IGES/PDES Organization (IPO), which is accredited by the American National Standards Institute (ANSI). PDES [1] is an adoption of the International Standard for the Exchange of Product Model Data (STEP). STEP has been approved by more than 20 countries worldwide, including all major U.S. trading partners.

1.5.2 Related Standardization Bodies

Major standardization activities are driven by various groups including national and international standardization authorities and a coalition of industry and user groups. Some interesting bodies relevant in this context are

- The NASA/Science Office of Standards and Technology [16] aims at the evolution and adoption of data systems standards for advancing the communication of scientific knowledge about Earth, the solar system, and our universe. It participates in and coordinates standardization activities related to data archiving and exchange.
- The Metadata Coalition [17] is a joint effort of industry vendors and users, which came up with the Metadata Metadata Interchange Specification V 1.1.
- Global Change Data and Information System (GCDIS) [18], while not a primary standardization body, constitutes a collection of distributed information systems operated by United States Federal Government agencies involved in global change research. GCDIS is designed to provide comprehensive global change related data and information to scientists and researchers, policy makers, educators, industry, and the public at large. GCDIS data and information span the world and are broadly multidisciplinary. GCDIS coordinates the efforts of the participating agencies to develop the GCDIS and, hence, deals to a significant extend with metadata issues.
- X3L8 is a technical committee of Accredited Standards Committee X3 [19], which is accredited by ANSI. X3L8 establishes standards for specifying and standardizing data. The focus of the work is on establishing ways to describe data to facilitate human use and to enable intelligent computer processing. Data is described through use of metadata (data about data). Metadata issues covered by the committee include naming, identification, definitions, classification, and registration. The standards developed by the committee are used in many areas, such as Electronic Data Interchange (EDI), data administration, information management, application development for information systems, and data access/interchange via the World Wide Web (WWW) and National Information Infrastructure (NII). X3L8 is involved in the development of ISO/IEC 11179 and of the X3L8 Metamodel for the Management of Sharable Data.
- X3T2 is a technical committee of Accredited Standards Committee X3 [19], accredited by ANSI. It focuses on Data Interchange and is concerned

generally with defining the data that is interchanged between a variety of entities ranging from subroutines and functions to remote applications. Obviously, there are many metadata issues involved in this.

1.5.3 Sample Standardization Projects

A partial list of interesting big projects which apply standards for the handling of metadata include

- the ESA Prototype International Directory (ESA-PID) [20], the European Space Agency's Prototype International Directory, which is an information source for rapid and efficient identification, location and overview descriptions of data sets of interest to the earth and space science research community including collections of earth observation images,
- the Environmental Data Registry (EDR) [21] of the U.S. Environmental Protection Agency (EPA), which provides information about data elements used in selected EPA systems,
- the Basic Semantic Repository (BSR) [22], a joint ISO/UN-ECE initiative, which tries to apply some kind of stable semantic reference point as a standard means of naming data elements,
- the Data Documentation Initiative (DDI) [23], developing a Document Type Definition (DTD) for an international codebook standard using the Standard Generalized Markup Language SGML, and
- the Government Information Locator (GILS) [24] identifying and describing information resources throughout the U.S. Federal Government, and providing assistance in obtaining the information.

1.6 Outline of the Book

This book consists of twelve chapters. Following this introduction we first look at the domain of document management, then at the state of the art in metadata handling in geographic and environmental applications, followed by a collection of chapters addressing metadata issues in the context of various media types, and finally we address the role and potential of metadata in the context of digital libraries.

Employment of database technology to administer documents of arbitrary type leads to a new, database-oriented view of publishing in which meta information occurs in a variety of ways. The chapter *Metadata Handling in HyperStorM* by K. Böhm focuses on the organization of metadata, the role of standards, and the storage and the usage of metadata for the processing of multimedia documents. It provides a classification of metadata as it appears in the context of multimedia documents. The chapter describes information models of standards for document interoperability, i.e., SGML/HyTime and DFR, with respect to metadata issues. It also discusses the usage of metadata for document storage, that is, how information models can be reflected within the database system. It shows how the nature of multimedia documents leads to new ways of exploiting metadata.

In the chapter *Metadata in Geographic and Environmental Data Management*, O. Günther and A. Voisard give an overview of metadata schemes and implementations in geographic and environmental information systems. Case studies include the Content Standards for Digital Geospatial Metadata of the U.S. Federal Geographic Data Committee (FGDC) and the Catalogue of Data Sources (CDS) of the European Environmental Agency. Furthermore, they discuss the UDK (Environmental Data Catalogue) project, an international software engineering effort to facilitate access to environmental data from the government and other sources, in greater detail. The chapter presents the UDK data model, its implementation as a distributed information system, and its upcoming integration into the World Wide Web.

Metadata Management for Geographic Information Discovery and Exchange by P. Drew and J. Ying looks at the role of metadata in the context of gluing together and making integrated use of existing geographic information systems illustrated by the geographic information exchange facility under construction in Hong Kong. The chapter focuses on the operational management needed to capture and transfer metadata across distributed geographic information systems. It shows how the transformation infrastructure works for the purpose of discovery and retrieval of information in different systems.

Using Metadata for the Intelligent Browsing of Structured Media Objects, by W. Grosky, F. Fotouhi, and Z. Jiang, discusses the role of metadata used in techniques for browsing and querying large image databases in an easy and efficient way. It is shown how images can be combined with textual information, establishing a kind of content-based hypermedia space, and how one can impose a virtual world metaphor on the information, which help the user to navigate through the image collection. Associative retrieval also includes navigation between a 3-D graphics representation of a virtual world and particular individual images as well as sets of images.

In *Content-Based Image Retrieval Using Metadata and Relaxation Techniques* by W. W. Chu, C.-C. Hsu, I. T. Ieong, and R. K. Taira, an approach for content-based image retrieval is proposed. The approach supports semantic query operations like *nearby* or *far away*, similarity operators, and the usage of conceptual terms. Image content is reflected by three metadata layers: at the pixel level, at the semantic level including spatial, temporal, and shape features, and at the knowledge layer. These layers provide for an integration of image representation and image content and allow for querying images by features and content.

In *A Metadatabase System for Semantic Image Search by a Mathematical Model of Meaning*, Y. Kiyoki, T. Kitagawa, and T. Hayama discuss use of metadata to provide associative search of images for a set of user-given keywords. The chapter presents a new method based on metadata for selecting images. The metadata used do not depend on the characteristics of the media, but use keywords that are associated with images based on user's impression. The approach uses various existing ontologies.

An information retrieval system that allows a simultaneous search for speech documents and text documents is presented in *Metadata for Content-based Retrieval of Speech Recordings* by M. Wechsler and P. Schäuble. It presents arguments for the use of *best-match retrieval* as opposed to *exact-match retrieval* and an approach which is based on the automatic generation of metadata to support such retrieval. It also discusses metadata organization by means of a new controlled indexing vocabulary, and its use that provides for the same retrieval effectiveness as a conventional Boolean retrieval system.

The chapter by R. Jain and A. Hampapur, *Video Data Management Systems: Metadata and Architecture*, characterizes video metadata and its usage for content based processing. Through an analysis of applications and the nature of queries, it leads to a rather comprehensive set of features (metadata items) that are included in their data model for representing video. The chapter discusses challenges of integrating video into databases and the resulting requirements for a video data model. A specific data model for describing video data and an architecture for a video data management system is presented.

Information personalization is increasingly viewed as an essential component of any front-end to a large information space. Personalization can achieve both customization of the presentation of information as well as tailoring of the content itself. In the chapter *The Use of Metadata for the Rendering of Personalized Video Delivery*, W. Klippgen, T. D. C. Little, G. Ahanger, and D. Venkatesh investigate techniques for personalizing information delivery based on metadata associated with diverse information units including video. They begin with a survey of approaches to information personalization and the requirements for this task. Subsequently, they present a characterization of the use of metadata to facilitate video information personalization.

The chapter by F. Chen, M. Hearst, D. Kimber, J. Kupiec, J. Pedersen, and L. Wilcox titled *Metadata for Mixed-Media Access* discusses a new information access paradigm which allows formulating queries in different media types than the media type of the data queried. The media considered in the chapter are speech, images of text, and full-length text. It illustrates how the combination of different media types can be used for accessing information and how metadata is automatically generated and used for supporting data access.

The last chapter, *A Framework for Meta-Information in Digital Libraries* by K. Beard and T. R. Smith, discusses the challenges of developing meta-information for digital libraries and outlines the contents of meta-information needed to support the basic search, retrieval, evaluation, and transfer functions in a digital library. The authors develop a framework for modeling meta-information in a digital library catalog. The framework provides a basis for the rational design of meta-information and for the analysis and resolution of catalog intraoperability and interoperability issues. The Alexandria Digital Library (ADL) provides specific examples of this framework.

This book represents culmination of our quest, which started with the special issue of SIGMOD record [6], in understanding the similarities and the differences in effective management of different digital media, with emphasis on the role metadata plays in achieving this objective. We hope that the effort of the distinguished authors of the following chapters leads us to good progress in our journey toward infocosm, a society in which people will have information anywhere, any time, and in many forms, for effective decision making, better learning, and more fun.

References

[1] ANS U.S. PRO/IPO. Product Data Exchange using STEP (PDES): Initial Release (Parts 1, 11, 21, 31, 41,42, 43, 44, 46, 101, 201 and 203). ANS U.S. PRO/IPO-200-1994, 1994.
URL http://uspro.scra.org/catalog/ANS.html).

[2] K. Böhm and T. Rakow. Metadata for Multimedia Documents. In [6].

[3] Dublin Core Metadata Element Set: Reference Description. OCLC Online Computer Library Center, Inc., Office of Research and Special Projects, Dublin, Ohio, USA, January 1997.
URL http://www.oclc.org:5046/research/dublin_core/.

[4] FGDC. *Content Standard for Digital Geospatial Metadata.* U.S. Government, Federal Geographic Data Committee, Washington, D. C., April 1997. URL ftp://fgdc.er.usgs.gov,
URL http://www.fgdc.gov/Metadata/metahome.html.

[5] FGDC, Federal Geographic Data Committee. *July 1997 Survey of FGDC Metadata Tools*, FGDC, USA, July 1997.
URL http://www.fgdc.gov/Metadata/toollist/metatools797.html.

[6] W. Klas and A. Sheth (eds.). Metadata for Digital Media. Special issue of SIGMOD Record, 23 (4), ACM Press, December 1994.

[7] V. Kashyap, K. Shah, and A. Sheth. Metadata for building the MultiMedia Patch Quilt. In S. Jajodia and V.S. Subrahmanian (eds.). *Multimedia Database Systems: Issues and Research Directions.* Springer-Verlag, 1995.

[8] V. Kashyap and A. Sheth. Semantic Heterogeneity in Global Information Systems: the Role of Metadata, Context and Ontologies. In M. Papazoglou and G. Schlageter (eds.). *Cooperative Information Systems: Current Trends and Directions.* Springer-Verlag, 1997, pp. 139–178.

[9] ISO/IEC. Information technology. *Specification and standardization of data elements – Part 6: Registration of data elements.* ISO/IEC 11179–6, 1997.

[10] InfoQuilt. Information Brokering for Globally Distributed Heterogenous Digital Media. Large Scale Distributed Information Systems Lab, Department of Computer Science, University of Georgia.
URL http://lsdis.cs.uga.edu/infoquilt/.

[11] Metadata Coalition ([17]). Metadata Interchange Specification. Version 1.1, August 1997. URL http://www.he.net/ metadata.

[12] T. Schütz and R. Böhm. Die Datenstrukturierung des Metainformationssystems Umwelt-Datenkatalog, In H. Kremers (ed.). *Umweltdatenbanken.* Metropolis, Marburg, 1994.

[13] T. Schütz and H. Lessing. Metainformation von Umwelt-Datenobjekten - zum Datenmodell des Umwelt-Datenkataloges Niedersachsens. In A. Jaeschke et al. (eds.). *Informatik für den Umweltschutz 7. Symposium Ulm.* Springer-Verlag, Berlin Heidelberg, 1993.

[14] Meta-Content Format. URL http://mcf.research.apple.com/wp.html.

[15] Health Level-7 (HL7). ANSI accredited standards developer. URL http://www.mcis.duke.edu/standards/HL7/hl7.htm.

[16] NASA/Science Office of Standards and Technology (NOST). National Space Science Data Center (NSSDC), Goddard Space Flight Center (GSFC), Greenbelt, MD. URL http://www.gsfc.nasa.gov/nost/.

[17] Metadata Coalition. Industry Initiative for Metadata Interchange. URL http://www.he.net/ metadata/.

[18] Global Change Data and Information System (GCDIS) Group. Federal Committee on Environment and Natural Resources (CENR), Subcommittee on Global Change Research (SGCR), Subgroup of the Data Management Working Group (DMWG). URL http://www.gcdis.usgcrp.gov/.

[19] National Committee for Information Technology Standards. URL http://www.x3.org/.

[20] Prototype International Directory. European Space Agency (ESA).

[21] Environmental Data Registry (EDR). U.S. Environmental Protection Agency (EPA), Washington. URL http://www.epa.gov/edr/.

[22] The Basic Semantic Repository. URL http://www.cs.mu.oz.au/research/icaris/bsr.html.

[23] The Data Documentation Initiative. URL http://www.icpsr.umich.edu/DDI/.

[24] Government Information Locator Service (GILS). U.S. Federal Government. URL http://www.usgs.gov/gils/index.html.