MeTaMaF: Metadata Tagging and Mapping Framework for Managing Multimedia Content

M. Anwar Hossain, Md. Anisur Rahman, Iluju Kiringa and Abdulmotaleb El Saddik
Multimedia Communications Research Laboratory (MCRLab)
School of Information Technology and Engineering (SITE)
University of Ottawa, Canada
{anwar, anis}@mcrlab.uottawa.ca, kiringa@site.uottawa.ca, abed@mcrlab.uottawa.ca

Abstract

Metadata comes into forefront as a savior of multimedia search and management complexity. However, the existence of the diverse set of metadata standards and the different vocabularies used by these standards has made that task especially challenging in recent days. In this paper, we propose a framework for managing multimedia content by effectively dealing with the heterogeneous multimedia metadata in a transparent fashion. The cornerstone of our approach is to leverage the existing metadata standards and provide schema-level mapping among them. The proposal also allows adding new tags/vocabularies with the given metadata vocabularies, defining equivalency relationship among those vocabularies, and separately managing the metadata vocabularies outside of the actual media files. Our approach naturally extends the search perimeter to cover different media types thereby producing more relevant search results. We have developed a prototype of the framework and evaluated its performance in terms of some basic query execution times.

1. Introduction

The huge growth of audio, video, image, text and other multimedia contents in personal collections and industry brings difficulty to manage media data effectively. The difficulty arises due to unsuitability of exact-match technique in querying large volume of multimedia data, complexity of content-based processing techniques, and inflexibility of retrieving the semantics and context from raw data [8]. Consequently, the problem with large collection of multimedia objects appears when the users try to retrieve specific objects based on a certain criteria. Therefore, solving problems such as searching for a certain song by a certain singer in a directory full of mp3 files is not doable. The use of metadata seems promising for solving such problems.

Metadata is a very generic term and is becoming part of day to day vocabulary. It is defined as the data about data. It is associated with data (association can be done by embedding metadata in the media data or storing the metadata independently and linking them separately with data) so that consequent searches would be done on the metadata rather than the multimedia itself. Due to smaller size and structured representation, metadata is useful in optimizing queries, improving query-response time, explaining answers, and mediating among information sources etc. It has become obvious that the role of metadata in managing multimedia data is more demanding than the management of traditional structured and textual data.

Although, there has been a universal agreement on the importance of metadata for managing multimedia data, no consensus for a common metadata standard has yet been made. Different communities and standard bodies came up with different metadata standards (e.g. EXIF [2], ID3 [3], ASF [10], DublinCore [4], MPEG-7 [5, 6] etc.) for identifying and managing multimedia contents. Some proposed metadata to be embedded within the media object (e.g. EXIF in jpeg, ID3 in mp3 etc.), while others advocate for separately maintained metadata (e.g. DublinCore for document, MPEG7 for multimedia etc.). These standards do not follow a common vocabulary and structure, which makes the management of multimedia content very challenging.

In order to overcome the challenges in dealing with several multimedia metadata standards, we propose a multimedia management framework that we call MeTaMaF. Furthermore, we make the following contributions in this paper:

- First, we present a mapping mechanism that logically relates vocabularies of different multimedia metadata standards. The mapping is performed both at the schema and content level. The schema-level mapping is realized by establishing semantic correspondences among the tags of the metadata schema (e.g. EXIF, ID3, MPEG-7). The content-level mapping is realized by consulting the similar terms of the tag values.
- Second, our framework allows adding new tags with the existing tag sets. This helps media owner/collector to keep additional information about any media for better identification. For example, a user may choose to add a new tag 'classification' and use it for the purpose of grouping his/her multimedia collection.
- Third, we present an algorithm for translating user queries according to the schema and content level mappings. The process hides the heterogeneity of the metadata standards from the user.
- Finally, we implemented the framework as desktop and web-based prototype. We tested the prototype and received encouraging results of sample query execution times.

The rest of this paper is organized as follows. Section 2 briefly comments on related research. Section 3 portraits a motivating scenario. Section 4 summarizes the objectives of the proposed framework. Section 5 describes the process of building the metadata repository from multimedia collection. The metadata mapping mechanism is presented in Section 6. In Section 7, we describe the query translation method. Section 8 provides implementation details along the test results. The last section draws the conclusion and states some possible future scopes.

2. Related Work

MeTaMaF includes concepts from several research directions; metadata-based multimedia management being the most prominent one. We will provide some of the most relevant works in this regard. However, our discussions will also cover some variant of metadata-based multimedia management approach such as annotations and free tagging schemes.

An early work suggesting the potential on using metadata for multimedia management is presented in [8]. The authors stated different aspects of metadata including its modeling, extraction, management and application for processing heterogeneous media. They also discussed about the various roles of metadata and available metadata standards, all these remains as a motivation for subsequent metadata-centric multimedia management activities. Some of these activities are VideoAnywhere [7], InfoHarness [15], UniMedia [17] etc.

MADCOW [9] is a multimedia digital annotation system, which allows users to create, modify, save, search and filter private and public annotations related to the content of web pages. The participating users can share their thoughts, proposals and comments with each other. This approach has the potential of creating a huge knowledge base consisting of annotated data about multimedia web resources. However, the annotation system may generate very subjective data about the resource, which may not have been intended by the original author. Although, the MADCOW system may not be suitable to manage a whole collection of multimedia data following a wide spectrum of metadata standard, it can be used in the context of MeTaMaF as an added value to enhance the metadata tagging process.

FLICKR [11], a popular image tagging and sharing service, is used by web users to upload images and then describe the contents with ad-hoc "tags". It uses a free-text keyword system (aka folksonomy) in which the user community defines the vocabulary through use. This introduction of user-defined tags and shared collection of photos in Flickr greatly empowers the community collective intelligence [12]. Application developers may leverage its tagging services through published API. Although FLICKR deals with images only, it can be extended to tag other multimedia contents. Furthermore, defining tags in FLICKR is very similar to defining keywords or categories in general. This approach is easy to follow but lacks the expressiveness that may be defined in a structured metadata-based approach consisting of key-value pairs.

The commercial system Singingfish [13] provides audio/visual multimedia search services to internet users. It combines metadata with additional rule-based classification system. It employs spiders to extract content metadata and uses other knowledge base technologies for searching and relevance ranking [14]. In addition to the spiders, audio and video content can voluntarily be contributed to the Singingfish repository to be processed and included in its search domain. Singingfish also uses MPEG-7 description schemes to characterize streaming media and facilitates searching those media [1] over the Internet. Unlike MeTaMaF, the Singingfish approach might seem suitable for a traditional search engine; it may not be used to manage multimedia collections in personal and organization level, due to various customized requirements that influence such media management process.

3. Motivating Scenario

The problem of managing large volume of multimedia data can be well illustrated in the following scenario. We assume that we have a large collection of image files in JPEG format (conforms to EXIF [2] standard) and audio files in MP3 format (conforms to ID3 [3] standard). Figure 1 shows the partial schema of EXIF and ID3. In response to a user query for media files containing the images of 'George Michael' and audio performed by 'George Michael', we expect to retrieve all types of media that satisfies the search criteria. A simple file name based search cannot retrieve expected search results. However, assuming that the media files contain the information we are looking for are present in some embedded metadata format and we extracted those metadata and maintained in separate repositories, our query can be translated to find the results by searching the value of relevant metadata tags from those repositories.

EXIF (ImageDescription, DateTime, Artist, …)
ID3 (Title, Year, Performer, …)

Figure 1. EXIF and ID3 metadata schema

For example, the images of George Michael may be found by searching the 'ImageDescription' tag of EXIF and the audio performed by George Michael may separately be found by searching the 'Performer' tag of ID3 as shown in figure 2.

ImageDescription	Date/Time	Artist	…
George Michael	1990	John	

(a) EXIF instance

Title	Year	Performer	…
Careless whisper	1990	George Michael	

(b) ID3 instance

Figure 2. Multimedia metadata instances

It is obvious that due to different structure and vocabulary of the metadata standards, searching for the media files to match the search keyword cannot be performed in a homogeneous manner. For example, in the above scenario, even though we are interested in looking for media files related to 'George Michael', we need to search the information in two different tags in two different standards, although those tags contain semantically similar information. This motivated us to work on our proposed framework that provides a unified mechanism to manage heterogeneous multimedia metadata for searching multimedia objects. In the following we gradually describe the framework.

4. System Objectives

MeTaMaF is a framework for managing heterogeneous multimedia content. This section summarizes the objectives of this framework that are described in a very independent manner to allow reusing them by other research entities and/or using them as an evaluation checkmark of any multimedia content management system.

- Tag framework/standard independent

The system should not be limited to a specific standard; the user should be able to specify his/her tags freely.

- Automatic metadata extraction if applicable

To simplify the process of managing multimedia files, the system should try to extract the metadata out of the multimedia whenever possible.

- Tag equivalency relationship

As seen in all the reviewed metadata frameworks, some of the tags are common between different tagging scheme and the only difference is the naming convention (e.g. author is equivalent to writer). However, care should be given to the semantic interpretation of the tag names when such relationship is defined.

- Simple query tool

Users should be provided with such tool so that they can easily query the information of interest.

- Simplified management of multimedia files

This can be done by using any of the well established standards for file management while maintaining the relation to the metadata.

- Search natural metadata management environment

Multimedia metadata should be managed (indexed and searched) in a search natural environment such as RDBMS. Because managing metadata in a formatted file environment (e.g. XML) can have a direct impact on the performance as well as increased cost of maintaining the system. However, XML should be used for exchanging metadata information among the participating entities.

- Multi-user environment

The system should be designed for multi-users as that environment is the primary target of such systems (e.g. multimedia document management system in large organization).

- Web-enabled

As the world converges faster to web based applications because of its advantages (e.g. light weight, run anywhere…), having a web front is becoming the default.

- Platform independent

This is the current trend in the software technologies, most of the new applications should run anywhere with minor modifications.

5. Metadata Collection

MeTaMaF's management functionality depends on metadata that are either embedded within the media files or are present as external XML descriptions. It reuses several open libraries to extract the embedded metadata (e.g. EXIF, ID3, ASF etc.) already present in the multimedia files and stores them in relational repository that acts as an abstraction of the actual multimedia content. In addition, in the case where a media file has external metadata descriptor as XML document, it parses that descriptor to identify the [tag, value] pair and stores them along with the extracted embedded metadata in the repository. This repository constitutes the collection of metadata which are leveraged for any search and management operation.

However, there are cases when the extracted metadata contains either null values or incomplete information, which may be edited and enriched by the user. Although the incompleteness of metadata information seems natural as the content owners do not take the measures to tag the high-level content (e.g. author, title etc.) when it is created, the low level extracted information (e.g. color, size, format etc.) can be used for specialized query operation.

The extracted and parsed metadata can further be augmented to populate empty or incomplete tag values or by adding new tags through the Tag Editor. Users normally use their domain specific knowledge and experience to edit the tag values. Tag editing can be performed after metadata extraction or at a later time (especially when the extraction is performed on large collection in a batch process).

6. Metadata Mapping

The cornerstone of MeTaMaF is the mapping among heterogeneous metadata vocabularies. To this effort, we analyzed the different multimedia metadata tag structures according to their inherent semantics based on common sense and subsequently established the correspondence among related tags. That is, we build equivalency relationships between tags that are semantically similar. We define this equivalency relationship as follows.

Definition 1. *When tag (x) is configured to have tag (y) as an equivalent then all query statements executed on tag(x) will be expanded to cover tag(y) and vice versa.*

We build tag correspondences among the metadata vocabularies that are partially shown in Table 1 and briefly described in the following. According to the EXIF standard of Digital Still Camera (DSC) image, the artist tag refers to the creator/author who is primarily responsible for the resource. The artist tag is semantically similar to the performer tag of ID3, author tag of ASF standard, and creator tag of DublinCore, all of which inherently advertise the ownership of the object. In case of MPEG-7 the ownership of a media may be expressed using similar tag such as creator.

Table 1. Metadata tag correspondences

Mapping properties	Mapping constructs
Ownership	{EXIF.artist}={ID3.performer}= {ASF.author}={DC.creator}= {MPEG7.creator.role}
Identification	{EXIF.imageDescription}={ID3.title} ={ASF.title}={DC.title}= {MPEG7.title/ MPEG7.audioSegment.songTitle/ MPEG7.videoSegment.seriesTitle/ MPEG7.main.tilte/ MPEG7.alternative.title/ ..}
Creation time	{EXIF.Date/Time}={ID3.year}= {ASF.creationDate}={DC.date}= {MPEG7.creationDate}
Copyright information	{EXIF.copyrightHolder}= {ID3.copyrightMessage}= {ASF.copyright}={DC.rights}= {MPEG7.rightsHolder}
Media type/group	{ID3.genre}={ASF.genre}= {MPEG7.classification.genre}

However, the expressivity of MPEG-7 allows more fine grained details of the creator such as creator role, agent data type etc. Therefore, the artist/performer/author tags are relevant with MPEG-7.creator.role tag.

Similar procedure is followed to map the other relevant tags. For example, the identification of an object in EXIF standard term may be discovered from the imageDescription tag as it refers to the title of an image. For ID3, DublinCore and ASF, the identification is expressed with the title tag. However, in MPEG-7 title may be defined to different parts of a multimedia object. Such as, a video may have a title whereas the audio inside a video may have another title. Hence the title information may enumerate as MPEG7.main.title, MPEG7.alternate.title and so on. All such variations are carefully considered in MeTaMaF.

Although we provide initial tag correspondences laid in Table 1 as a starting point in MeTaMaF, new equivalency relationships may be established at run time. This is required if the user adds new tags to the existing tag collections.

The mapping process followed by MeTaMaF is based on the understanding of semantics of the tags. However, other schema matching and/or model management approaches such as Similarity Flooding (SF) algorithm [18], CUPID [19], or [20] could have been utilized to generate a tentative mapping among the different multimedia metadata schemas and later refined it for our purpose.

6.1. Content Mapping

MeTaMaF is designed to support two types of content mapping in addition to the tag level mapping. The first type of content mapping, unlike tag mapping, is performed on the tag values at the time of executing a query. This approach mainly focuses on the synonym of a query term and extends the search perimeter to cover the synonyms. To find the synonym of a term, dictionary-like approach such as the WordNet [21, 22], is adopted in our framework.

The second type has been introduced due to the presence of multiple instances of a word, which may not be captured by the WordNet approach. This situation mainly arises in the case of nouns. For example the term "George Michael" may be mapped to "G. Michael" or "Michael, G" and so on. This situation is handled by allowing the user to manage a content collection we refer as *Similar Value Collection (SVC)* that is gradually enriched by the domain specific knowledge. Such a collection is consulted by the query engine to translate the query conditions.

7. Query Translation

One of the key functionalities of MeTaMaF is to provide multimedia search facility through which a user may retrieve different types of multimedia content with a single query. Central to this process is the translation of a query *Q* to the vocabulary of heterogeneous multimedia metadata schema by using the equivalence of tags of those schemas. The query vocabulary in this case refers to the tags and tag values used in the condition(s) of the query. The query result is the list of actual multimedia content satisfying the query condition.

Consider a query that retrieves the multimedia content with the condition *NAME = "George Michael"* AND *YEAR = "1990"*. Here *NAME* is an abstract tag in the user interface layer, which is equivalent to *identification* properties in Table 1. Similarly *YEAR* refers to the *Creation time* properties in Table 1.

In order to translate the abstract query conditions to search-specific conditions we replace the conditions with several disjunctions based on the tag-equivalence relationship. Such as, *NAME = "George Michael"* and *YEAR = "1990"* is translated to [*imageDescription = "George Michael"* OR *title = "George Michael"*] AND [*Date/Time = "1990"* OR year = *"1990"* OR *creationDate = "1990"* OR *date = "1990"*].

In addition to the above translation, the query conditions may further be extended for content-level mapping as described in Section 6.1. For example, the query condition *NAME = "George Michael"* may be translated further as [*imageDescription = "George Michael"* OR *imageDescription = "G. Michael" OR title = "George Michael"* OR *title = "G. Michael"*]

7.1. Query Translation Method

The exact process of query translation is formulated in the algorithm presented in Figure 3. Some auxiliary functions are used in this algorithm. The function *extract_conditions(Q)* returns the set of conditions present in query Q. The *extract_tag(C)*, *extract_operator(C),* and *extract_val(C)* functions return the tag, operator and the values respectively from the condition C. The *extract_equivalent_tags(t)* function returns all equivalent tags of the given tag *t*. The *extract_similar_content(val)* returns all equivalent terms (i.e. synonyms and similar terms) of *val* consulting the *WordNet* and *SVC*. Finally, *replace_condition(Q,C,C')* returns a query by replacing condition C in Q by a new condition C'.

```
QueryTranslation (Q)
Input: A query Q from user.
Output: Translated query Q'
begin
   Q'= Q;
   SofC = extract_conditions(Q);
   for each C ∈ SofC do
       t  =   extract_tag (C);
       op= extract_operator (C);
       val= extract_val (C);
       SofET = extract_equivalent_tags(t) ;
       SofEV=extract_similar_content(val) ;
       C' = C;
       for each eqT ∈ SofET do
           for each eqV ∈ SofEV do
                C' = C' or (eqT op val);
           end for
        end for
        Q'= replace_condition(Q',C,C') ;
    end for
end
```

Figure 3. Query translation algorithm

8. Implementation

We implemented the functionalities of the MeTaMaF framework described in this paper. Figure 4 shows the overall architecture of the framework. The user or the media collector uses the user interface to upload media files into media repository. The upload operation also extracts the embedded metadata from the media files and stores them in relational database according to the data model supported by this framework. This data model consists of six major entities. The *FileSystem* entity represents a file system node (e.g. FTP server) to which multimedia files can be uploaded and later be retrieved

The *Multimedia* entity refers to an abstraction of actual multimedia instance containing the file location. The *Tag* entity contains metadata tags of different standards. The *MultimediaTag* entity contains the list of tags and their values that corresponds to each multimedia object. It therefore represents the connection between the *Multimedia* and the *Tag* entities. The *EquivalentTag* entity holds the association among the semantically related tags as partially stated in Table 1. Ideally a tag may be equivalent to many other tags. The *SimilarValueCollection* entity maintains the list of domain specific data values that correspond to same term or concept as described in Section 6.1.

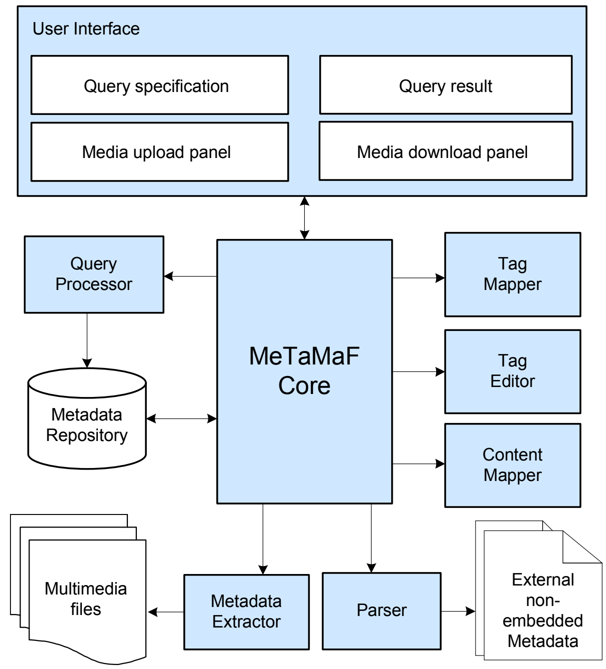

Figure 4. Multimedia metadata instances

Once the media files are uploaded into the system, the user can search media files through the search interface by specifying some query conditions. The specified query is then translated according to the algorithm presented in Figure 3 and subsequently the final results are delivered to the user.

The prototype is built on J2EE architecture [23] (currently known as Java EE) that follows EJB3.0 specifications and is tested on JBoss 4.03 application server. Java (J2SE 5.0) is used as a programming language. The persistence layer is implemented using MySQL [24] database. The overall system is tested on IBM computers with Pentium 4 3.60GHz CPU, 1.00GB RAM and MS Windows XP operating system.

8.1. Example Interface

Among the several user interfaces we provide a sample desktop client view in Figure 5, which shows that the client has used the metadata extraction service for extracting metadata from an mp3 file.

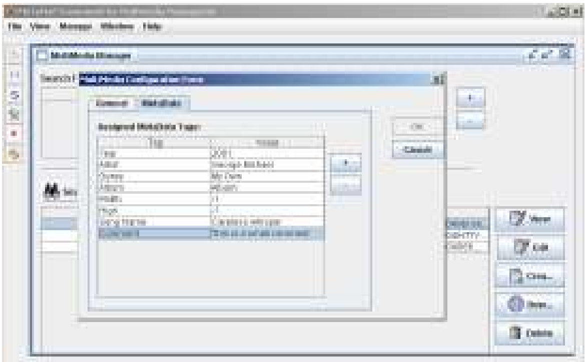

Figure 5. Multimedia metadata instances

8.2. Limitation

Metadata-based information management approach requires human to tag the content, which is inherently subjective in nature and sometimes impractical in the case of large multimedia collections if done manually. To partially address this issue, we adopted the established mechanisms to extract embedded metadata that are already present within the media files, although this process depends on the availability of such information. In reality, low-level system data (e.g. height, width, color, resolution etc.) are automatically generated by the equipment. However, the multimedia files are not tagged properly with semantic information by the media curator and hence the automatic metadata extraction produces very few meaningful data. Given such circumstances, the test result we produce may not be fully representative of the real life environment. However we would like to emphasize that the approach presented here is worth experimenting.

Another drawback of our work is that we did not perform any usability test with the prototype. Such test would provide valuable insight about the system's ease of use, navigability, user friendliness and other interesting parameters.

8.3. Result Analysis

We tested our prototype and evaluated it in terms of execution time of search operations that are performed on the web. The test result is presented in Figure 6. The time in the figure represents sum of query translation time and query execution time. The test includes queries that contain upto four search conditions. However, the system does not pose any limitation on the number of search conditions to be specified. We observed that the query time slightly increases with the increase of search conditions. The test was performed using a media repository containing metadata of 1000 media files of different formats. The query translation in this test includes the tag equivalency relationship and does not consider content mapping.

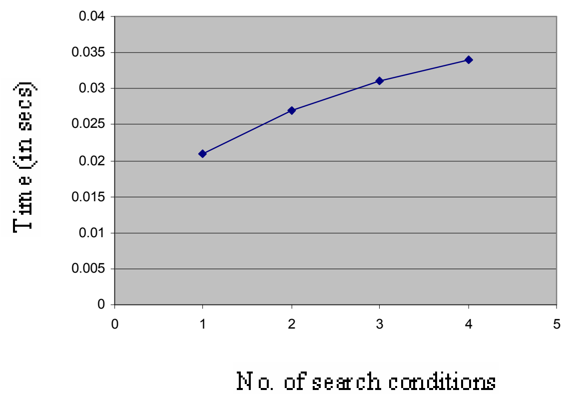

Figure 6. Query translation and execution time

9. Conclusion

This paper presents the design and implementation of a metadata based multimedia management framework. It uses several existing metadata tagging standards covering a wide range of applicability ranging from application specific (e.g. ID3, EXIF etc.) to application independent (e.g. MPEG-7). Our focus was to manage these diverse metadata formats in order to facilitate searching and retrieving multimedia content in a wider context. This has been realized by generating correspondences among the tags of different metadata schemas. This approach enables the use of metadata formats that are already in extensive use within existing media archives as well as leveraging newly emerging metadata formats.

We plan to extend our framework by adding new extraction plugins for various other types of multimedia formats such as .wrl, .x3d, .mov, qt, .mp4 etc. and augment the embedded metadata information through techniques like data mining and semantic based crawling. We aim to build J2ME client that would provide the search and retrieval services to nomadic users. The objectives of the multimedia content management system presented in this study have a great range of coverage and should be refined more and eventually it might become the benchmark for multimedia content management.

10. Acknowledgements

The authors acknowledge the financial assistance of Natural Sciences and Engineering Research Council of Canada (NSERC) and of Ontario Ministry of Energy, Science and Technology (PREA).

11. References

[1] E., Rehm, "Representing internet streaming media metadata using MPEG-7 multimedia description schemes," *Proc. ACM Multimedia Conf.* (MM2000), Los Angeles, CA, Oct.–Nov. 2000.

[2] Exchangeable Image File Format, http://www.exif.org/specifications.html

[3] MP3 metadata standard, http://www.id3.org/index.html

[4] The Dublin Core Metadata Initiative, http://dublincore.org/

[5] MPEG-7 Overview, http://www.chiariglione.org/ MPEG/standards/mpeg-7/mpeg-7.htm

[6] Manjunath, B. S., P. Salembier, and T. Sikora, eds., *Introduction to MPEG-7: Multimedia Content Description Language*, John Wiley & Sons, New York, 2002.

[7] A. Sheth, C. Bertram, and K. Shah, "VideoAnywhere: A System for Searching and Managing Distributed Video Assets," *SIGMOD Record,* vol. 28, no. 1, pp. 104-109, 1999.

[8] S. Boll, W. Klas, and A. Sheth, "Overview on Using Metadata to manage Multimedia Data", Multimedia Data Management, 1998: 1-24.

[9] P. Bottoni et al., "MADCOW: a multimedia digital annotation system," *Proc. Working Conf. Advanced Visual interfaces* (Gallipoli, Italy, May 25 - 28, 2004). ACM Press, New York, NY, pp. 55-62.

[10] Advanced Systems Format (ASF) Specification, available at: http://www.microsoft.com/windows/ windowsmedia/forpros/format/asfspec.aspx

[11] Flickr, http://www.flickr.com

[12] A. Weiss, "The power of collective intelligence," *NetWorker*. New York, vol. 9, no. 3, pp. 16-23, Sept. 2005.

[13] http://search.singingfish.com/sfw/home.jsp

[14] M. Fritz, "Singingfish: Advancing the Art of Multimedia Search," *EContent*, April 2003, http://www.econtentmag.com/Articles/ArticleReader.aspx?ArticleID=4291

[15] I. Shah and A. Sheth, "INFOHARNESS: managing distributed, heterogeneous information," *IEEE Internet Computing*, vol. 3, no. 6, pp. 18-28, Nov/Dec. 1999.

[16] Y. Rui et al., "Relevance Feedback: A Power Tool in Interactive Content-Based Image Retrieval," *IEEE Trans. on Circuits and Systems for Video Technology* , Special Issue on Segmentation, Description, and Retrieval of Video Content, vol. 8, no. 5, pp. 644-655, Sept, 1998.

[17] H.H. Kim, S.S. Park, and W. Kim, "A Framework for the Integration of Multimedia Data," *Journal of Object Technology*, vol. 4, no. 5, pp. 27-35, July-August 2005.

[18] S. Melnik, E. Rahm, and P.A. Bernstein, "Rondo: a programming platform for generic model management," *Proc. ACM SIGMOD int'l Conf. Management of Data* (San Diego, California, June 09-12, 2003), ACM Press, New York, NY, 193-204.

[19] J. Madhavan, P.A. Bernstein, and E. Rahm, "Generic Schema Matching with Cupid," *Proc. 27th int'l Conf. Very Large Data Bases* (September 11 - 14, 2001). P. M. Apers, P. Atzeni, S. Ceri, S. Paraboschi, K. Ramamohanarao, and R. T. Snodgrass, Eds. Very Large Data Bases. Morgan Kaufmann Publishers, San Francisco, CA, pp. 49-58.

[20] P.A. Bernstein, "Applying Model Management to Classical Meta Data Problems," *Proc. CIDR 2003*, pp. 209-220.

[21] G.A. Miller, "WordNet: a lexical database for English," *Commun. ACM,* vol. 38, no. 11, pp. 39-41, Nov. 1995.

[22] JWordNet, http://jwn.sourceforge.net/

[23] I. Singh, B. Stearns, and M. Johnson, Designing Enterprise Applications with the J2EE Platform, 2nd edition. Addison-Wesley, 2002.

[24] MySQL database management systems, http://www.mysql.com/