Multimedia Databases: An Approach to Design

Valerie Hobbs & Diarmuid Pigott
School of Information Technology
Murdoch University
Murdoch, WA 6150
vhobbs@murdoch.edu.au, diarmuid@civilisation.org

Abstract

In this paper we describe a methodology and diagramming notation for designing a multimedia database. Multimedia databases have novel requirements compared with conventional text and numeric databases, in particular the need to consider the most appropriate physical formats of the digital artefacts and the possibility of reuse of artefacts for different semantic purposes. However the need to be able to model a problem at a high level of abstraction, and to proceed from a conceptual to physical solution, remain essential. The approach presented here ensures a parallel focus on representing both the meaning of the media artefact, and its appropriate technical specification. We present several new concepts - the media entity, the media relationship and the ERD+ diagram - that allow us to incorporate media artefacts into a conventional data modelling technique. The database design and documentation produced allow considerations of specific implementation to be deferred as long as possible.

INTRODUCTION

Multimedia databases are databases that include images, sound, video, rich text and other examples of 'rich' media instead of, or as well as, the 'conventional' data types of text and numbers. These media types have special requirements for capture, storage, retrieval and display. Unlike text databases, where we can create a conceptual model more or less in ignorance of the product it will be implemented on, the extent to which DBMS products enable direct support for rich media types varies. It is a temptation, therefore, to undertake multimedia design with a particular DBMS in mind. This however contradicts what we generally require of a modelling solution, which is to explore the requirements of the solution independent of any particular implementation.

In this paper we describe a database design methodology that we developed while teaching a multimedia databases unit this semester. We felt that there were fundamental issues relating to the 'meaning' of media that should be given equal importance to those of technical details of implementation. However, when it came to multimedia database design, we could find no existing design methodologies to follow that would allow us to do this. In consequence, we found ourselves exploring the unique requirements for a design methodology for multimedia databases, and in the process developing our own methodology.

The conventional database design methodology presented in most textbooks for text-based applications is usually summarised as a linear progression from conceptual through logical to physical design (see, for example, Connolly & Begg 1998). Although some iteration between stages occurs, the progression is essentially one way, and in theory at least, all physical considerations can be deferred until the logical design is complete and the DBMS is selected.

When we create a multimedia database, on the other hand, we are forced to confront issues of physical implementation at an early stage. Digital artefacts such as sound files and images exist - they have a physical persistence and reality that text does not. Choosing to include the name 'Mary' in a database involves typing the word in. Choosing to include a picture of Mary involves making decisions about how Mary will be photographed, what digital format the image should be in, and so on. And how we catalog and store the resultant image of Mary has obvious implications for how it can be searched and retrieved.

We will explore these issues more fully using an example. We will make use of the same example throughout this paper to demonstrate our design methodology.

An example multimedia database - the Earth Sounds database

Some natural materials, such as sand or rocks, are known to make sounds such as whistling, ringing or booming. These phenomena have been reported from various locations (for example, Ringing Rocks Park in the US, Kaui Singing Sands in Hawaii), and descriptions and possible explanations have been written up in popular and scholarly articles. Suppose now that we wish to create a multimedia database about these sound-making natural phenomena.

Consider the possible ways in which we can use media in this database. We can include photographs of the sites, recordings of the sounds, rich text documents of the articles that have been written, and so on. Now consider the sources of multimedia artefact - the actual sound, image, or text files - that we can draw on to populate the database. We may have non-digital versions of the information, such as paper documents or photographs, which we must convert to digital form. We may have digital artefacts that have already been created by someone, possibly for a completely different purpose. We may even decide to travel to the location of some of these natural phenomena, and capture our own sound recordings or images expressly to include in our database.

When we assess our collection of media artefacts, we find they cover a range of distinctiveness and potential usage. We may have managed to capture the only recording of a reported sound at a particular location. We may have several photographs of one site, taken from different viewpoints. We may have a monograph that describes several different phenomena in the same article. And we may have a recording of whistling sands that has no recorded location, which could illustrate the general phenomenon of 'whistling sands' at various sites.

This brief description highlights several issues that are relevant to multimedia database design. We summarise them here and discuss them more fully in Hobbs & Pigott (in prep).

A media artefact is a physical object with persistence, which can be reused. This has fundamental implications for the way that media artefacts are stored and used compared with conventional text databases. The concept of 'reuse' of text is meaningless. Text has no independent existence in the database - using a word involves typing it in, not locating it from a stored collection of words. If the same word is used elsewhere for a slightly different meaning, the word is entered again. We cannot make any distinction between the word itself and how it is used.

In contrast, a media artefact is an entire object, with its own structure and properties, and may be of considerable size. Reuse is more likely to involve referencing the same physical location than copying the artefact to a new location. More importantly, however, are the consequences for the *ways* in which we can reuse artefacts. The meaning that we attach to the contents of an artefact may differ according to circumstances. For example, the same image of a site may be used to illustrate the site itself, or the phenomenon associated with the site. Thus the physical media artefact, and what it means or represents in any particular situation, become dissociated. This has obvious consequences for how we describe the *content* of an artefact, discussed next.

A media artefact has infinite semantic content - and no absolute semantic content. Since media artefacts may be reused for different purposes, it follows that the information content of a media artefact - WHAT it represents - is not fixed. The same artefact may appear in many different databases, or even fulfil different purposes in the same database. A media artefact has, in effect, both infinite semantic content and no absolute semantic content. The implication is that the meaning of the artefact in any particular situation is decided on a case by case basis, for purely local purposes. However, it is essential that this meaning is documented with the artefact, as a set of keywords or other description.

A media artefact may be a representation of something that exists in the 'real world'. What this implies is that the process of capture or conversion to digital form inevitably forms part of the way in which the real world object is represented. This is fundamentally different from representing a problem as a text description. With text, the only issue is to define an adequate abstraction of the problem. The abstraction, if adequate, is then as good as the 'real thing'. With media artefacts, there are twin issues of both defining the abstraction and then representing it in physical form. The act of creating or collecting digital artefacts means that choices have to be made as to media type, format, quality, number of samples, views, and so on, with inevitable compromises in the digital representation of the object compared with the original.

A media artefact has information about it that is always true. Because the media artefact is a persistent object, whose technical specifications and provenance details are known, it carries this information with it wherever it goes. However the artefact is used, for whatever semantic purpose, it is still true that (for example) it is in a particular format and was digitised on a particular date from an original photograph taken by some photographer.

A database that includes media artefacts will normally be based around some conceptual model in which the artefacts play a part. This means that we need to be able to model the media component of our database as part of the overall problem domain, integrating it with the text based component.

Considering these issues with those we discussed earlier, we can now summarise our requirements for a design methodology as follows:

- It should provide a database solution, but not be tied to any specific DBMS implementation of media types or storage options
- It should progress from the conceptual to the physical, and be able to handle imperfectly known information at higher levels
- It should allow precise specification of the technical requirements for the media artefacts
- It should allow precise specification of the semantic usage of the media artefacts, independently of their physical implementation
- It should provide a modelling notation that summarises both technical specification and semantic usage of the media artefacts
- It should record the logic of decision-making by making the steps of the analysis explicit

For the particular purposes of the unit we developed this semester, we can add

- it should (as far as possible) build on techniques that the students were already familiar with - in our case entity-relationship modelling and relational database design.

A METHODOLOGY FOR MULTIMEDIA DATABASE DESIGN

The methodology that we present here is a process of integrating media modelling with conventional data modelling, and adding, in parallel, progressively more technical specification and more semantic specification for the required media artefacts. The result of the design process is a set of tables that incorporates both the data records and media artefacts. The final design is implemented in a generic manner in a relational structure, and specific hardware/software requirements for storage, retrieval and display are documented. This allows considerations of specific implementation to be deferred as long as possible.

The steps (summarised in Figure 1) are:

1. Data modelling
   - Create conventional data model - ERD and data dictionary
2. Media modelling
   - Identify media entities
   - Document media requirements
   - Document processing requirements
   - Create media artefact specification
   - Define media relationship quality
   - Draw ERD+
   - Create thesaurus of keywords
3. Integrate data and media modelling
   - Convert ERD+ to final table design
4. Finalise documentation
   - Create media dictionary
5. Create database
   - Create tables from final ERD
   - Locate or create media artefacts, and determine location of storage
   - Create artefact records, including annotation and keywording

We now describe these steps in more detail.

STEP 1: DATA MODELLING

The first step in the design methodology is to draw an initial entity-relationship diagram that is based on abstract entities only. The premise for this is that even a multimedia database, unless completely context-free, needs some conceptual structure to describe its domain of interest and thus facilitate data management. It is worth noting that the extent to which the data model is driven by the pre-existence of media artefacts, and vice versa, depends on the nature and purpose of the database. If the artefacts themselves are the reason the database is being constructed, the abstraction of the problem may be largely driven by the artefacts; or alternatively the requirement for artefacts of particular types may be determined mainly by the requirements of the problem domain.

At the end of this stage we have a conventional ERD and associated data dictionary, with no media artefacts. Figure 2 shows the ERD for the Earth Sounds database.

STEP 2: MEDIA MODELLING

The purpose of the media modelling stage is to identify and document the media requirements of the database - what media artefacts will be used, and the role they play in the problem domain. The first stage specifies the broad media requirements of the database, summarised in what we term **media entities** (explained below). This is followed by two parallel activities: defining the technical specification of the media artefacts required, and defining the semantic role of the artefacts (see Figures 3a and 3b). The completed media analysis is documented in the form of additions to the original ERD, which we call the **ERD+**.

Identify media entities
We now introduce the concept of the **media entity** as a device for representing the semantic roles played by media artefacts in the database.

We define the **media entity** as representing the set of all media artefacts that are related to a particular entity for a particular purpose. We identify potential media entities by considering both the *media metatype* that is required (sound, image, video, etc) and the *role* the particular media artefact is to play in the database. For example, in the Earth Sounds database, we wish to include recordings of sounds and photographs of the sites, and we identify the media entities Site Recording and Site Photo.

An entity may have one, several or no associated media entities. An entity may also have more than one media entity of the same media metatype: in the Earth Sounds database, we have Site Photo, Site Map and Phenomenon Diagram, all images, but used to represent different things. A media entity is thus somewhat similar to the domain of an attribute, in that it represents not merely a data type, but a pool of allowable values from within a type.

We add the media entities to the ERD as a double oval symbol connected to the entity.

Document the media requirements
We document the broad technical requirements for the media artefacts that are represented by the media entity. This includes what media metatype is required, what quality is required and why (for example, we might specify that the site photos in Earth Sounds should be high quality images), and indicative (candidate) formats that could be used. This becomes the basis for the media artefact specification, described below.

Document the processing requirements
We also need to note the requirements for searching and retrieval on the media entities, eg. whether they will be retrieved in their entirety, or whether their internal structure needs to be queried. We can include here indicative queries that the database must be able to answer.

Create media artefact specification
We now define the technical specification for the required media artefacts associated with each media entity. We add the media artefact specification to the diagram in the form of a long rectangle connected to the media entity (Figure 3a). The media artefact specification has two parts - known and deferred. As much information as is known already, or which can be specified precisely, is listed in the top half of the rectangle. This specification is based on the appropriate formats for the artefact, as determined by its role and expected usage.

If the artefacts exist already (eg because they are the main focus of the database, or will be collected by the designer) then their specification will be completely known. In other circumstances we may not have complete control over the circumstances of the media artefact's creation, or must collect them from third sources and may not wish to compromise their accuracy by reconversion. We therefore state in the bottom half of the rectangle what information is deferred (ie not able to be specified at this stage).

Define media relationship quality for each media entity
The other, parallel activity in the media modelling step is specifying the semantic role of the media artefacts. We have already specified the various roles that media artefacts are to play in the database, by identifying media entities. We now need to describe these roles more precisely. We find that different roles have consequences for cardinality (how many media artefacts are linked to a particular entity record), and for whether artefacts are replaceable.

We introduce the term **media relationship quality** to describe the role the artefact plays in relation to its associated entity, and define it in terms of *substitutability* and *specificity*.

We can say that an artefact is *substitutable* if another artefact could fulfil the same purpose for the entity just as well. *Not substitutable* means that no other artefact will 'do' for that purpose. An artefact is *specific* to the entity it is related to if it holds that particular meaning (as defined through the media entity) for only one instance of the entity. *Not specific* means the artefact could describe another entity instance (of that entity type) just as well with the same meaning.

Based on the criteria of specificity and substitutability we can distinguish three types of relationship. We call these types diagnostic, illustrative and symbolic. We describe these types fully in Hobbs & Pigott (in prep), and summarise them here.

In a **diagnostic** relationship, artefacts are both *specific* and *not substitutable*. In effect the artefact is used to represent a fact about the entity. In the Earth Sounds database, we identify the artefacts represented by media entities WrittenAccount and SiteSeismics as diagnostic.

In the **illustrative** type of relationship, a media artefact may be used as an illustration of some aspect of the entity. Other similar artefacts could be substituted, or several could be used, but the artefact would not be reusable by other entities in the same entity type: in other words the artefacts in an illustrative relationship are *specific* but *substitutable*. In Earth Sounds, the artefacts represented by SitePhoto, SiteMap and SiteRecording are all considered as being illustrative.

In the **symbolic** relationship, we acknowledge that the artefact, plus its meaning, could be shared by several instances of an entity. We also acknowledge that there is nothing particular about the artefact itself: another one could be used to represent the meaning just as well. Thus the symbolic relationship is both *not specific* and *substitutable*. The symbolic type of relationship occurs when the artefact is used in a metaphorical way, not necessarily associated with its original or most obvious meaning. In the Earth Sounds database, PhenomenonType Photo is designated as symbolic. Here we wish to use a single image (or images) to provide an illustration of the general type of phenomenon (ie pertaining to rocks, sand, etc), to be shared by several records of phenomena of that type.

We identify the media relationship quality of each media entity in the database by adding the appropriate initial to the line (Figure 3b).

It is important to remember that the definition of an artefact as having a particular MRQ applies only to the particular role it is playing - there is nothing inherent in the artefact itself. Thus, we could envisage the situation where a media artefact was used in one part of a database in a D relationship, and reused in another as S. However, the *role* the artefact played would be different in each place, and would be defined through separate media entities.

Draw the ERD+

We have now extended our original data model, as shown in the ERD, with media entities, media artefact specifications, and media relationship quality indicators. We call this final diagram the ERD+. The ERD+ provides a diagrammatic summary of both the technical and semantic requirements for media artefacts, and their incorporation into the conventional data model. The ERD+ for the Earth Sounds database is shown in Figure 4.

Create a thesaurus of keywords

We have referred throughout this paper to the semantic content of media artefacts - we now need to consider how this is to be described, so that artefacts can be retrieved directly through their meaning as well as through association with data records. We assume that the media artefacts we use will always need to be described by keywords and a text description, regardless of whether content-based searching (for example, IBM's Query By Image Content™) is also available.

To enable precision of data entry and retrieval, some mechanism is required to ensure a limited domain of words that can be used for the purpose. There are various ways in which such words lists can be assembled, but the systematic approach requires the construction of a thesaurus. Thesaurus construction is basic to all cataloguing and keywording systems, and we follow the conventional principles here. In this step we define a thesaurus of terms appropriate to the theme of our database.

Included in the thesaurus' structure is a hierarchical structure with broader terms and narrower terms which enables searches of varying generality and specificity, while providing cued lists of acceptable terms for searching. The mechanism of preferred terms and deprecated terms, permits a system of guidance where incorrect terms are used.

STEP 3: INTEGRATE DATA AND MEDIA MODELLING

This step involves converting the ERD+ into a set of relational tables that can be used as the basis for the multimedia database. The ERD+ resolves into tables in the following way:

- All conventional (data) entities are converted to tables.
- Two 'artefact tables' are created to hold the information about artefacts derived from the media entity and media artefact specification.
- Three 'media relationship quality' tables are created to link artefact records with their associated entities, in the manner appropriate to their roles.
- A table is created to hold the set of keyword terms defined in the thesaurus.

Entity tables

Each conventional entity becomes a table, and the usual rules for representing related entities through the use of foreign key fields apply.

Artefact tables

The information from the media entity and media artefact specification is represented in two tables, one to hold the technical information that is fixed when the artefact is created (Artefact-Technical), and the other to hold the details of the role (or roles) the artefact plays (Artefact-Role).

The **Artefact-Technical** table holds technical information for all artefacts in the database. It includes an ID field, a pointer mechanism for locating the artefact for the database (disk address or URL), and a set of fields that contain the technical annotative information about each artefact. These consist of at least information pertaining to the initial process of creating the artefact, information pertaining to conversion to the current format, including references to external documentation held in the media dictionary, and copyright information. We anticipate that eventually we will make use of the proposed MPEG-7 standard (ISO, 1998) in defining the technical information required. Artefact-Technical also contains a field to record media metatype, to facilitate searching on different media types.

Each artefact used in the database is a record in the Artefact-Technical table. The information in this table (apart from location) would generally not change, and can be exported with the artefact if necessary to be used in different databases.

The **Artefact-Role** table holds semantic information about the role (or roles) the artefact plays in the database. The table has an ID field to identify the record, a description field, which contains a text description of the contents of the artefact, a keywords field, which holds a subset of the keywords defined in the thesaurus, and a field for the name of the media entity the artefact is used for. The Artefact-Role table also contains a foreign key field to link the artefact role to the appropriate record in the Artefact-Technical table.

Storing artefact information in two tables rather than one enables us to separate the fixed (technical) information about the artefact itself from the shifting semantic information about the artefact's content. The relationship between Artefact-Technical and Artefact-Role is one-to-many - thus we can use the same physical artefact (as denoted in the Technical table) to be reused in the database by playing more than one role (as denoted in the Role table).

Media relationship quality tables

Each of the types of relationship quality (Diagnostic, Illustrative or Symbolic) between each media entity and its associated entity translates into different expected cardinalities between the entity and the physical artefact. The media relationship quality (MRQ) tables are three tables, one for each quality type, that define the cardinality of the relationship between the conventional entities and the media artefact in its role.

When the quality of the relationship between entity and media entity is Diagnostic, we said that the artefact was specific and non-substitutable. We therefore expect there to be a single media artefact associated with the entity for the purpose described by the media entity - the expected cardinality is 1:1. We represent this by an intersection table between the table derived from the original entity and the Artefact-Role table, which we call the MRQ-D table.

When the relationship quality between entity and media entity is Illustrative, we said that the artefact was specific but substitutable. We therefore expect there to be potentially many media artefacts associated with a particular entity: the expected cardinality is 1:N. We represent this by the MRQ-I table.

Finally, when the relationship quality between entity and media entity is Symbolic, we said that the artefact was both non-specific and substitutable. Thus we expect the relationship between entity and artefact to be M:N. We represent this in the MRQ-S table.

(We note that the intersection table is actually required only in the case of the Symbolic relationship: the D relationship could be implemented by the foreign key of the Artefact-Role in the related entity table, and the I relationship by the foreign key of the entity table in the Artefact-Role table. However, we choose to use three intersection tables for consistency and increased flexibility in dealing with unusual cases.)

The generalised table structure is shown in Figure 5. For clarity, only one entity is shown, related to each of the MRQ tables. This would be equivalent to the entity having one or more media entities of types D, I and S. In general, if an entity has an associated media entity (or entities) of type D it will have an entry in the MRQ-D table, if it has a media entity of type I it will have an entry in the MRQ-I table, and so on.

Thesaurus table
The thesaurus is implemented as a single table, with each keyword term as primary key. Other fields include semantic note (whether the term is broad or narrow, or deprecated), referred term (the related broader term, if appropriate) and scope (the category of the keyword). The thesaurus table is used as a source of the keywords that are used in the Artefact-Role table.

The schema for the final database is shown in Figure 6.

STEP 4: FINALISE DOCUMENTATION - CREATE MEDIA DICTIONARY

A multimedia database will require many things that are not standard with a conventional database. We also need to record the implementation information that is implied from the media artefact specification. We call this part of the documentation the Media Dictionary to distinguish it from the table descriptions stored in the Data Dictionary. The media dictionary contains at least the following information:

- Formats used in the database, and the artefacts that use them. (This includes all possibilities of 'deferred' details: eg. if the image format is deferred, but is one of three possibilities, JPEG, BMP, GIF, then all three formats are included in the Media Dictionary.)
- Hardware required to run the system, including specific hardware for individual formats.
- Special software required to run the media artefact (such as plug-ins)
- We also include in the Media Dictionary any additional hardware/software implications for the required retrieval that were documented in general terms in the media modelling stage.

STEP 5: CREATE THE DATABASE

The final stage in the methodology is to create the prototype database itself. This involves:
- Creating the tables from the final ERD+, as discussed
- Locating or creating all media artefacts, if not already done
- Entering the records for each table, adding technical annotation for each artefact in the Artefact-Technical table, and keyword information for each artefact-role in the Artefact-Role table.
- Determining the most appropriate location of artefact storage, and including this in the location field in Artefact-Technical.

At this stage we have a database structure that can handle retrieval of artefacts based on various criteria (discussed below), assuming the host operating environment can display the artefact to the user. If more sophisticated searching is required then we may need to modify the database structure to accommodate specific requirements of the software (for example, interfacing to a GIS package or an integrated analysis package).

DISCUSSION

The design methodology presented here fulfils the criteria we specified at the start of this paper. It allows us to model the requirements for both technical specification and semantic usage of media artefacts, and to proceed from a high level of abstraction to a physical design. The ERD+ diagram provides a convenient notation for summarising both technical and semantic requirements.

The methodology allows us to deal with the issue of physical reuse of media artefacts, and to separate the cases where reuse is for a different purpose in the database (what we call 'contingent' reuse), and the where reuse is deliberate, with both the media artefact and its meaning being shared. Contingent reuse of artefacts is taken care of by the fact that the same artefact can appear more than once in the Artefact-Role table, and semantic reuse by the S type of media relationship quality.

The database structure described is able to handle various types of queries involving media artefacts using normal SQL. Artefacts can be searched and retrieved on their links to data records, their own keyword set, the role they play and their media metatype. Queries on artefacts can be integrated with queries on data records. We note several general types of queries that are supported:

- 'Find all media artefacts [optionally: associated with this data record]'
- 'Find all media artefacts of this type (=media metatype) [associated with this data record]'
- 'Find all media artefacts with this content (=keyword) [associated with this data record]'
- 'Find all media artefacts associated with this role (=media entity) [associated with this data record]'

We have used the design methodology described here throughout our multimedia databases unit this semester, in which the students each designed and built a multimedia database on a theme of their own choosing. We found the media entity, media relationship quality and ERD+ concepts particularly useful as a means of getting the students to focus on *why* they wanted to include particular types of media artefacts in their database, and in helping them model at an appropriate level of detail. We have found the methodology a useful pedagogical tool, and believe it to have wider application.

REFERENCES

Connolly, T., and Begg, C. *Database Systems: A Practical Approach to Design, Implementation and Management.* Second Edition. Addison-Wesley, 1998.

Hobbs & Pigott (in prep). What is the meaning of media? Paper to be submitted to 2000 IRMA International Conference.

International Organisation for Standardisation, 1998. MPEG-7 Requirements Document v.7, ISO/IEC JTC1/SC29/WG11: Coding Of Moving Pictures and Audio. http://sound.media.mit.edu/mpeg4/audio/public/mpeg7/w2461.html

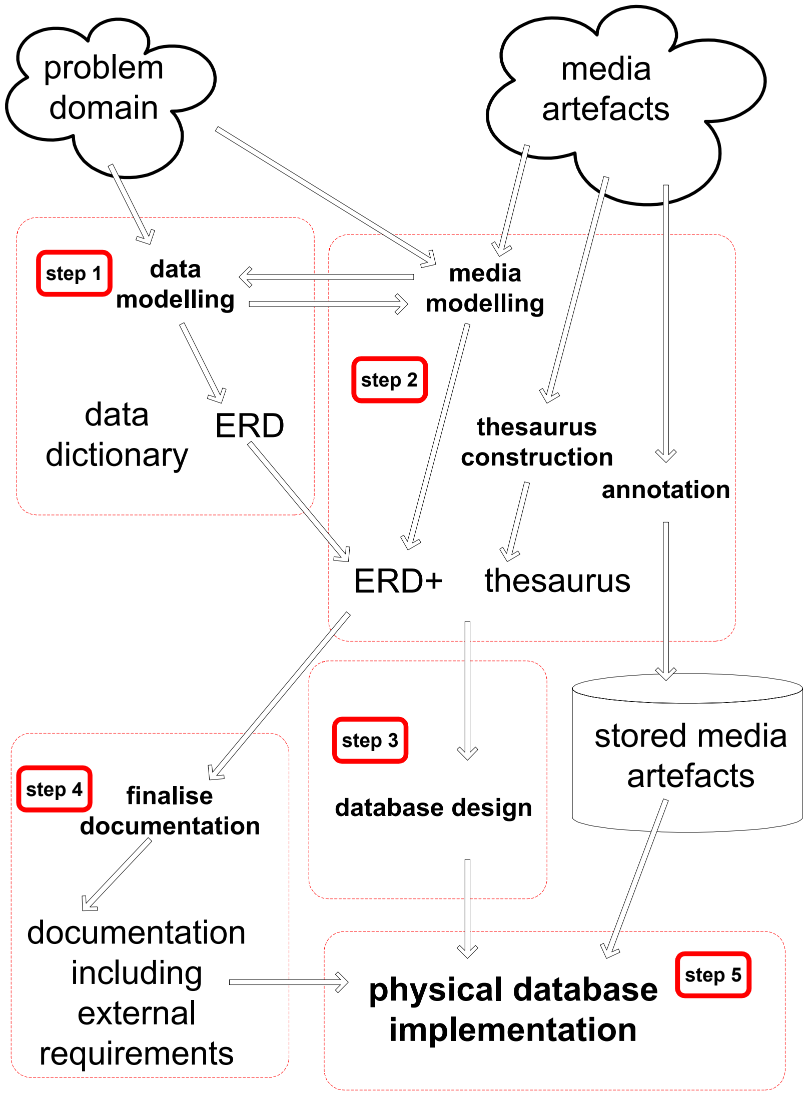

Figure 1. The multimedia database design methodology

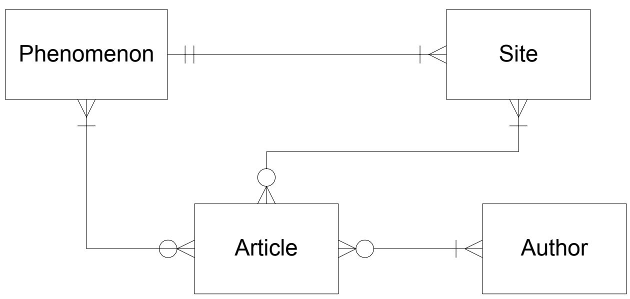

Figure 2. The ERD (showing data only) for the Earth Sounds database

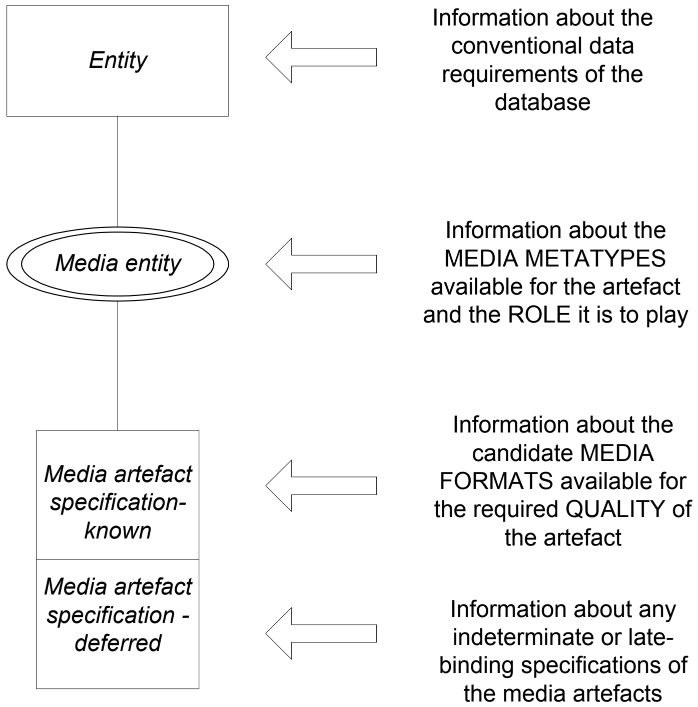

Figure 3a. Defining the technical specification for the media artefacts.

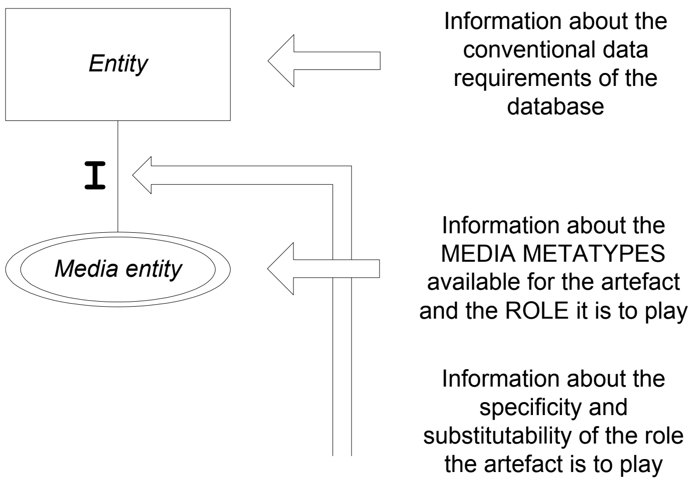

Figure 3b. Defining the semantic role of the media artefacts

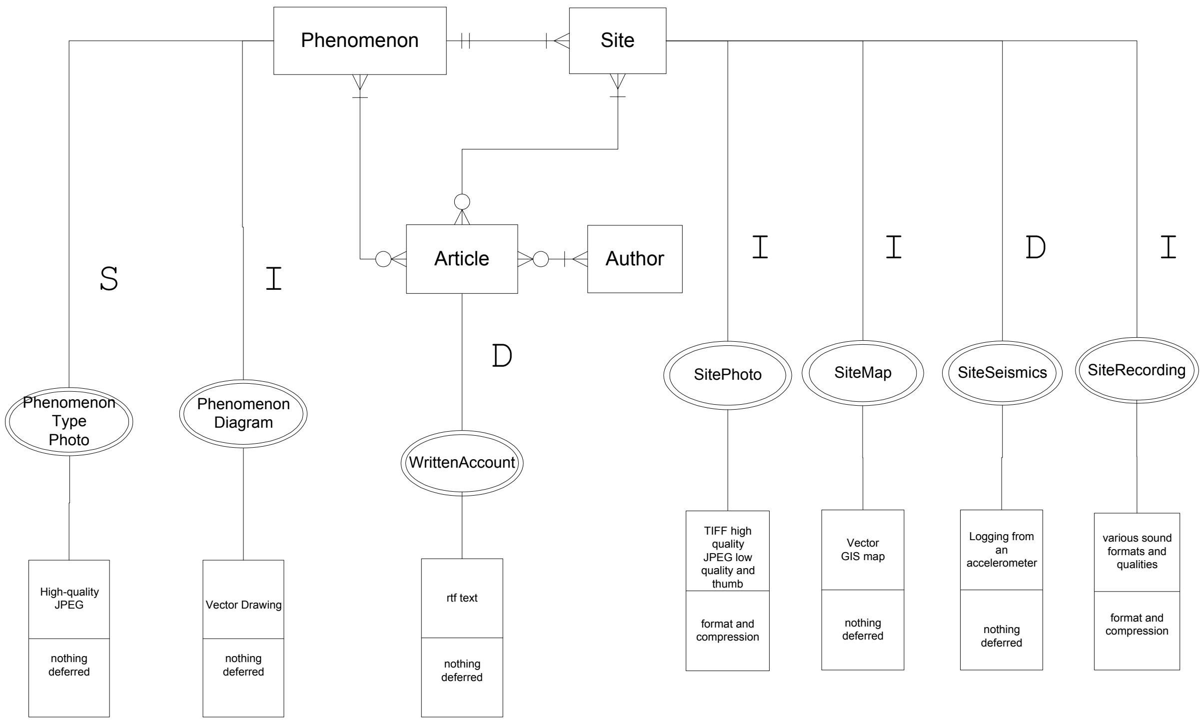

Figure 4. The completed ERD+ for the Earth Sounds database.

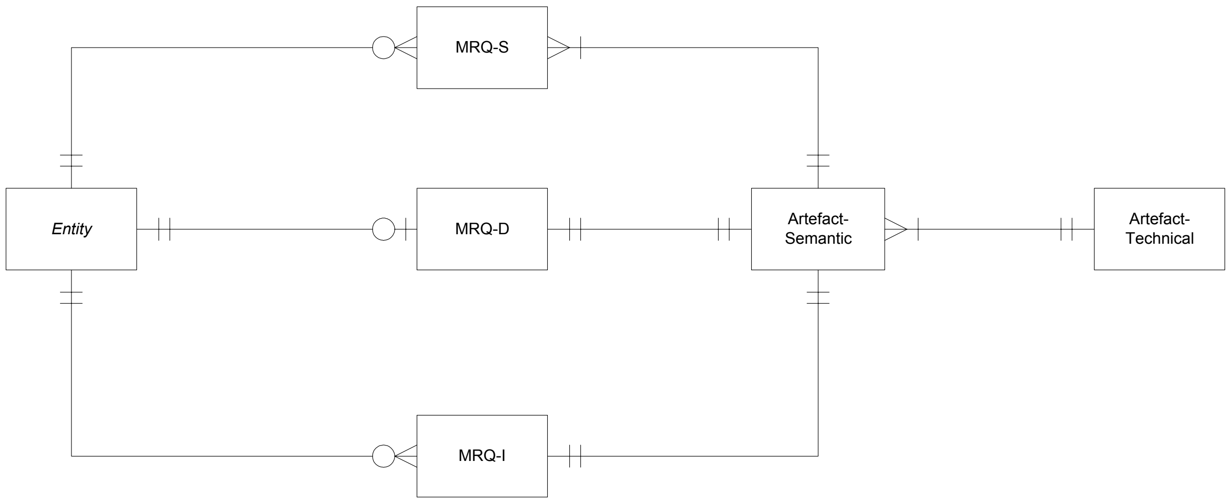

Figure 5. Generalised table design after conversion from ERD+.

ARTEFACT-TECHNICAL (ArtefactTechID, Location, CreationDate, CreationEquipment, CreatorName, CreationSettings, LocaleDescription, ConversionSettings, FormatInfo, CODEC, CopyrightOwner, CopyrightDescription, Provenance)

ARTEFACT-ROLE (ArtefactRoleID, Description, Keywords, MediaEntity, **ArtefactTechID**)

MRQ-D (**EntityID, ArtefactRoleID**)
MRQ-I (**ArtefactRoleID, EntityID**)
MRQ-S (**ArtefactRoleID, EntityID**)

Figure 6. Final database schema. Artefact-Technical fields are indicative.