Faculdade de Engenharia da Universidade do Porto

Mestrado em Gestão de Informação

Metadata-oriented Multimedia Information Retrieval
(Pesquisa de Informação Multimédia Orientada por Meta-informação)

Karen An Koen Goethals

Licenciada em Engenharia Electrotécnica
('Burgelijk Electrotechnisch Ingenieur')

pela Faculdade de Engenharia da Universidade de Gent
('Faculteit Toegepaste Wetenschappen Universiteit Gent')

Dissertação submetida para satisfação parcial dos
Requisitos do grau de mestre em

Gestão de Informação

Dissertação realizada sob a supervisão da

Professora Doutora Maria Cristina de Carvalho Alves Ribeiro

do Departamento de Engenharia Electrotécnica e de Computadores

da Faculdade de Engenharia da Universidade do Porto

Porto, Janeiro de 2001
I. Abstract

With the advent of the information superhighways, a vast amount of data is currently available on the Internet. The advances in the methods for processing and coding image, video and audio have led to a wide spectrum of non-traditional forms of data being available across the network. Text, image, video and audio data, all perhaps stored in multiple heterogeneous formats, form the core of what is known today as multimedia data.

The large amounts of multimedia data that are being collected in multimedia archives have originated the need for means to search the archives. The diversity in the multimedia data and in the applications that manipulate it also requires the development of standard formats for information exchange.

It is recognised that, unlike text, for which effective content-based retrieval methods are currently used, it is not easy to automatically index or search image and video items based on their content. The main difficulties have to do with the complexity of extracting useful features from such objects. It is therefore necessary to make use of any available descriptions associated with the objects, concerning either the content itself or the context of the objects' production and use.

Metadata is getting a fair share of attention in current research on solutions to the querying of multimedia objects. The number of metadata-related projects demonstrate this trend. Because of the importance of this subject, MPEG has dedicated a group, MPEG-7, to work on a standard for a 'Multimedia Content Description Interface'.

This dissertation proposes the integration of descriptive metadata and content metadata in a description scheme for generic multimedia items and demonstrates its use in information storage and retrieval. It is argued that the efficiency of retrieval and the usefulness of its results can improve with this approach.

A description scheme is proposed for the integration of metadata and it is tested on the implementation of a prototype information storage and retrieval system based on a relational database. The use of the description scheme is demonstrated with a data set of a photographic archive (part of the MPEG-7 Content Set), documents of a historical personal archive and a video of the MPEG-7 Content Set. The prototype serves as a workbench for the evaluation of the advantages of the combination of content and context metadata in multimedia documents.

In order to address the problems posed by information exchange between heterogeneous systems, some exploration of the possibilities available for converting the results of queries in the prototype to XML is presented.
II. Resumo

Com o aparecimento das auto-estradas da informação, grandes volumes de dados estão actualmente disponíveis na Internet. Os avanços registados no processamento e codificação de imagem, vídeo e áudio têm feito proliferar na rede dados em diversos formatos não tradicionais. Dados com texto, imagem, vídeo e áudio, usando possivelmente múltiplos formatos, constituem o que se designa actualmente como dados multimédia.

Com as grandes quantidades de dados que estão a ser coleccionadas em arquivos multimédia, tornam-se necessárias formas de os pesquisar. A diversidade nos dados multimédia e nas aplicações que os manipulam também faz aparecer a necessidade de formatos normalizados para a troca de informação.

É reconhecido que, ao contrário do texto em que estão já disponíveis métodos eficientes de pesquisa baseada no conteúdo, não é fácil indexar ou pesquisar automaticamente itens de imagem ou vídeo com base no seu conteúdo. As dificuldades principais têm a ver com a complexidade associada à extracção de parâmetros úteis para caracterizá-los a partir do seu conteúdo. Torna-se por isso necessário usar todas as descrições disponíveis dos objectos, sejam estas acerca do conteúdo propriamente dito ou do contexto de produção e uso do objecto.

A meta-informação tem sido alvo de atenção na actual procura de soluções para a interrogação em objectos multimedia. O número de projectos relacionados com meta-informação demonstra esta tendência. A importância da descrição de objectos audiovisuais com vista à pesquisa levou o MPEG a dedicar um grupo, o MPEG-7, ao trabalho de desenvolvimento de uma norma designada “Interface para a Descrição de Conteúdos Multimédia”

A presente dissertação propõe a integração de meta-informação descritiva e de conteúdo num esquema de descrição para itens multimédia genéricos e demonstra o seu uso no armazenamento e pesquisa de informação. Defende-se que a eficiência da pesquisa e a utilidade dos seus resultados podem melhorar com esta abordagem.

Propõe-se um esquema de descrição para a integração de meta-informação e testa-se o mesmo na implementação de um protótipo de sistema de armazenamento e pesquisa de informação baseado numa base de dados relacional. O uso do esquema de descrição é demonstrado com um conjunto de dados de um arquivo fotográfico (que faz parte do conjunto de dados de teste do MPEG-7), um conjunto de documentos de um arquivo histórico pessoal e um vídeo também pertencente ao conjunto de dados do MPEG-7. O protótipo serve como banco de ensaios para a avaliação das vantagens da combinação de meta-informação de conteúdo e de contexto em documentos multimedia.

Para lidar com os problemas postos pela troca de informação entre sistemas heterogéneos, foram exploradas algumas hipóteses de conversão dos resultados de interrogações efectuadas no protótipo para XML.
III. Résumé

Avec l'arrivée de la transmission de l'information plus efficace, une vaste quantité de données est actuellement disponible sur l'Internet. Les avances dans les méthodes pour traiter et coder l'image, le vidéo et l'acoustique ont mené à une gamme étendue de formes des données non traditionnelles disponible à travers le réseau. Des données de texte, d'image, de vidéo et des données sonores - tous peut-être enregistrés dans des formats différents - forment traditionnellement le cœur de ce qui est connu aujourd'hui comme données de multimédia.

Les grandes quantités de données de multimédia qui sont rassemblées en archives de multimédia ont lancé le besoin de moyens de rechercher les archives. La diversité dans les données de multimédia et dans les applications qui la manipulent également exige le développement des formats standards pour l'échange de l'information.

Il est clair qu'il n'est pas facile de classer ou rechercher des éléments d'image et de vidéo basés sur leur contenu, bien que ces méthodes sont actuellement utilisées pour du texte. Les principales difficultés se rapportent à la complexité d'extraire les dispositifs utiles à partir de tels objets. Il est donc nécessaire de se servir de toutes les descriptions disponibles associées aux objets, au sujet le contenu lui-même ou le contexte de la production et de l'utilisation de l'objet. Metadata obtient une partie équitable de l'attention dans la recherche actuelle sur des solutions à la question des objets de multimédia. Le nombre de projets metadata-connexes démontre cette tendance. En raison de l'importance de ce sujet, MPEG a consacré un groupe, MPEG-7, pour standardiser un 'Multimedia Content Description Interface'.

Cette dissertation propose l'intégration du metadata descriptif et du metadata de contenu dans un arrangement de description pour les éléments génériques de multimédia et son utilisation est démontrée par sauver et rechercher l'information. Il est prouvé que l'efficacité de la recherche de l'information et l'utilité de ses résultats s'améliorent avec cette approche.

Un arrangement de description est présenté pour l'intégration du metadata et cela est testé par l'exécution d'un prototype pour sauver et retrouver de l'information, basé sur une base de données relationnelle. L'utilisation de l' arrangement de description est démontrée avec des données d' archives photographiques (qui fait partie du MPEG-7 Content Set), des documents d' archives personnelles historiques et d'un vidéo du MPEG-7 Content Set. Le prototype sert à l'évaluation des avantages de la combinaison du metadata de contenu et de contexte dans des documents de multimédia.

A cause des problèmes qui se posent par échange de l'information entre les systèmes hétérogènes, quelques possibilités sont explorées pour convertir les résultats des requêtes dans le prototype en XML.
IV. Samenvatting

De komst van de informatiesnelweg zorgt voor een omvangrijke hoeveelheid data op het Internet. De vooruitgang geboekt in methoden voor het verwerken en coderen van beelden, video en audio, hebben geled tot een diverse waaier aan niet-traditionele datavormen beschikbaar via het netwerk. Data van tekst, beelden, video en audio - elk eventueel in verschillende formaten opgeslagen - vormen traditioneel de ‘kern’ van wat vandaag de naam ‘multimedia’ draagt.

De grote hoeveelheden multimedia data die verzameld worden in multimedia archieven, liggen aan de oorsprong van de vraag naar middelen om deze archieven te doorzoeken. De diversiteit in de multimedia data en in toepassingen die deze data manipuleren, vragen ook om de ontwikkeling van standaard formaten voor de uitwisseling van informatie.

Het is duidelijk dat het niet eenvoudig is om beelden en video gebaseerd op hun inhoud te indexeren of te doorzoeken, hoewel er dergelijke methodes reeds in gebruik zijn voor tekst. De grootste moeilijkheden houden verband met het afleiden van nuttige kenmerken uit de multimedia objecten. Daarom is het noodzakelijk om gebruik te maken van beschikbare beschrijvingen geassocieerd met de objecten, die ofwel de inhoud zelf behandelen ofwel de context van de productie en het gebruik van het object.

In het hedendaags onderzoek krijgt meta-informatie een groot stuk van de aandacht als mogelijke oplossing om multimedia documenten te zoeken. Een bewijs van deze trend zijn het toenemend aantal metadata projecten. Omwille van het belang van dit onderwerp, heeft MPEG een groep, zijnde MPEG-7, speciaal gewijd aan het werk voor het standariseren van een ‘Multimedia Content Description Interface’.

Deze thesis stelt de integratie voor van beschrijvende metadata en inhoudelijke metadata in een ‘description scheme’ (beschrijvend schema) voor algemene (multimedia) items en het gebruik van het schema wordt gedomineerd in het opslaan en terugvinden van informatie. Er wordt aangetoond dat de efficiëntie van het terugvinden van informatie en de bruikbaarheid van de resultaten via deze aanpak verbeterd.

Er wordt een schema voorgesteld voor de integratie van meta-informatie en dit wordt getest in de uitvoering van een prototype systeem voor het opslaan en terugvinden van informatie, gebaseerd op een relationele database. Het gebruik van het schema wordt gedomineerd met een dataset van een fotografisch archief (dat deel uitmaakt van de MPEG-7 Content Set), een historisch persoonlijk archief en een video van de MPEG-7 Content Set. Het prototype is bedoeld om de voordelen van een gezamenlijke voorstelling van inhoud en context metadata in multimedia documenten te evalueren.

Omwille van problemen die zich stellen bij de informatie-uitwisseling tussen heterogene systemen, worden tevens enkele mogelijkheden onderzocht om zoek-resultaten in het prototype om te zetten in XML.
V. Acknowledgements

I would like to thank

Prof. Cristina Ribeiro, for her help and recommendations for my work

Prof. Gabriel David, for giving me the opportunity to be part of the MetaMedia project

All the persons who made in one way or another a contribution to my work, especially Suzana

An, Liesbet, Marijke and Sofie for their back-up from Belgium

Ricardo for his never-ending patience

And my parents for their never-ending support.
VI. Table of Contents

I. Abstract ........................................................................................................ III
II. Resumo .......................................................................................................... V
III. Resumé ....................................................................................................... VII
IV. Samenvatting .............................................................................................. IX
V. Acknowledgements ..................................................................................... XI
VI. Table of Contents ..................................................................................... XIII
VII. Table of Figures ....................................................................................... XVI
VIII. Table of Tables ...................................................................................... XIX
IX. Abbreviations ............................................................................................ XXI
X. Keywords .................................................................................................. XXIII
XI. Glossary .................................................................................................... XXV

1. INTRODUCTION ......................................................................................... 1
   Organisation of the dissertation ................................................................. 4

2. METADATA .................................................................................................... 5
   2.1 Introduction .............................................................................................. 5
   2.2 Metadata on the Internet ....................................................................... 6
   2.3 Metadata Standards .............................................................................. 7
   2.4 ISAD(G) .................................................................................................. 7
   2.5 Dublin Core ............................................................................................ 8
   2.6 Resource Description Framework ......................................................... 10
      2.6.1 Introduction ...................................................................................... 10
      2.6.2 RDF Data Model ........................................................................... 11
      2.6.3 RDF Syntax .................................................................................... 12
      2.6.4 RDF Schema ................................................................................ 12
   2.7 MPEG-7 .................................................................................................. 13
      2.7.1 Introduction ...................................................................................... 13
      2.7.2 Goals .............................................................................................. 14
      2.7.3 Applications ................................................................................... 16

3. MARKUP LANGUAGES ................................................................................ 17
   3.1 Introduction ............................................................................................. 17
   3.2 Standard Generalised Markup Language (SGML) ................................... 17
   3.3 HyperText Markup Language (HTML) ..................................................... 18
      3.3.1 XHTML .......................................................................................... 18
   3.4 Extensible Markup Language (XML) ......................................................... 19
      3.4.1 Origin and goals ............................................................................ 19
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.4.2</td>
<td>Semi-structured</td>
<td>20</td>
</tr>
<tr>
<td>3.4.3</td>
<td>Syntax</td>
<td>20</td>
</tr>
<tr>
<td>3.4.4</td>
<td>Document Type Declaration (DTD)</td>
<td>21</td>
</tr>
<tr>
<td>3.4.5</td>
<td>Schemas</td>
<td>22</td>
</tr>
<tr>
<td>3.4.6</td>
<td>Extensible Stylesheet Language (XSL)</td>
<td>23</td>
</tr>
<tr>
<td>3.4.7</td>
<td>Namespaces in XML</td>
<td>23</td>
</tr>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>25</td>
</tr>
<tr>
<td>4.2</td>
<td>Relational Model</td>
<td>25</td>
</tr>
<tr>
<td>4.3</td>
<td>Architecture of a DBMS</td>
<td>26</td>
</tr>
<tr>
<td>4.4</td>
<td>Multimedia DBMS</td>
<td>28</td>
</tr>
<tr>
<td>4.4.1</td>
<td>Multimedia-specific aspects</td>
<td>28</td>
</tr>
<tr>
<td>4.4.2</td>
<td>Multimedia in traditional DBMS</td>
<td>28</td>
</tr>
<tr>
<td>4.5</td>
<td>Future</td>
<td>30</td>
</tr>
<tr>
<td>4.6</td>
<td>Exporting relational data in XML</td>
<td>31</td>
</tr>
<tr>
<td>4.6.1</td>
<td>Requirements</td>
<td>31</td>
</tr>
<tr>
<td>4.6.2</td>
<td>Possible approaches</td>
<td>31</td>
</tr>
<tr>
<td>4.6.3</td>
<td>SilkRoute</td>
<td>32</td>
</tr>
<tr>
<td>4.6.4</td>
<td>XML constructor</td>
<td>33</td>
</tr>
<tr>
<td>4.6.5</td>
<td>Integrating keyword search</td>
<td>33</td>
</tr>
<tr>
<td>4.6.6</td>
<td>Oracle (8i)</td>
<td>34</td>
</tr>
<tr>
<td>4.6.6.1</td>
<td>Generate XML</td>
<td>35</td>
</tr>
<tr>
<td>4.6.6.2</td>
<td>Exchange XML</td>
<td>36</td>
</tr>
<tr>
<td>4.6.6.3</td>
<td>Store XML</td>
<td>36</td>
</tr>
<tr>
<td>5.1</td>
<td>Classification</td>
<td>39</td>
</tr>
<tr>
<td>5.2</td>
<td>Problems in Multimedia Information Retrieval</td>
<td>40</td>
</tr>
<tr>
<td>5.3</td>
<td>Research</td>
<td>42</td>
</tr>
<tr>
<td>5.4</td>
<td>Examples of multimedia information retrieval systems</td>
<td>42</td>
</tr>
<tr>
<td>5.4.1</td>
<td>Search-engines on the Web</td>
<td>42</td>
</tr>
<tr>
<td>5.4.2</td>
<td>Describing and retrieving photos using RDF and HTTP</td>
<td>43</td>
</tr>
<tr>
<td>5.4.3</td>
<td>Columbia's content based visual query project</td>
<td>43</td>
</tr>
<tr>
<td>5.4.4</td>
<td>Query by Image Content (QBIC)</td>
<td>45</td>
</tr>
<tr>
<td>5.4.5</td>
<td>Content-base search in Oracle 8i</td>
<td>46</td>
</tr>
<tr>
<td>5.5</td>
<td>Conclusions</td>
<td>46</td>
</tr>
<tr>
<td>6.1</td>
<td>Introduction</td>
<td>49</td>
</tr>
<tr>
<td>6.2</td>
<td>System Of Description – Package</td>
<td>51</td>
</tr>
<tr>
<td>6.2.1</td>
<td>Description Unit</td>
<td>51</td>
</tr>
<tr>
<td>6.2.2</td>
<td>Segment</td>
<td>51</td>
</tr>
<tr>
<td>6.2.3</td>
<td>Hierarchy</td>
<td>53</td>
</tr>
<tr>
<td>6.3</td>
<td>Storage – Package</td>
<td>55</td>
</tr>
<tr>
<td>6.4</td>
<td>About – Package</td>
<td>56</td>
</tr>
</tbody>
</table>
# Table of Contents

6.5 Creator – Package .............................................................................57  
6.6 Owner – Package .............................................................................58  
6.7 Multiple Language Support .................................................................60  
6.8 Relation with MPEG-7 Multimedia Description Schemes ....................62  
6.8.1 Content Organisation ......................................................................66  
6.8.1.1 Collections .............................................................................66  
6.8.1.2 Models ..................................................................................67  
6.8.2 Content Description .......................................................................67  
6.8.2.1 Structural Aspects .................................................................67  
6.8.2.2 Conceptual Aspects ...............................................................69  
6.8.3 Content Management ....................................................................70  
6.8.3.1 Media Description .................................................................71  
6.8.3.2 Creation Information ..............................................................72  
6.8.3.3 Usage Information .................................................................73  
6.8.4 Navigation and Access .................................................................73  
6.8.5 User Interaction ............................................................................74  
6.8.6 Basic Elements and Schema Tools ................................................75  
6.8.6.1 Basic Elements ....................................................................75  
6.8.6.2 Schema Tools .......................................................................75  

7. ARCH-METAMEDIA ..........................................................................77  
7.1 Relational translation for the prototype ..............................................77  
7.2 Population of the prototype ...............................................................83  
7.2.1 Archive of Rafa de Queiroz ..........................................................83  
7.2.2 Video ("Science Eye": Bridge Construction) of the MPEG-7 Content Set ........................................................................87  
7.3 Queries in Arch-metamedia ...............................................................90  
7.3.1 Hierarchy ....................................................................................90  
7.3.2 Segment ......................................................................................90  
7.3.3 About Package ...........................................................................91  
7.3.4 Owner & Creator .......................................................................91  
7.3.5 Combinations ............................................................................91  
7.4 Comparison ......................................................................................92  
7.5 Conversion to XML in Arch-metamedia ............................................93  
7.6 Conclusion .......................................................................................94  

8. CONCLUSION ..................................................................................95  
Future work ............................................................................................95  

9. REFERENCES ....................................................................................97
VII. Table of Figures

Figure 1. Graphical representation of the tasks in the MetaMedia project ...........................................4
Figure 2. A simple example of the RDF data model ..............................................................................11
Figure 3. A simple example of the RDF data model, using the DC vocabulary .................................13
Figure 4. Scope of the MPEG-7 standardisation .....................................................................................14
Figure 5. A UML based representation of possible relations between Ds and DSs .........................15
Figure 6. Major components of a DBMS ...............................................................................................26
Figure 7. XML, portal to a world of data .................................................................................................35
Figure 8. Generating XML from query results with Oracle’s XML SQL Utility ...............................35
Figure 9. Package diagram of the MetaMedia Description Scheme ....................................................50
Figure 10. UML presentation of the SystemOfDescription – Package .................................................52
Figure 11. Example for a film of the hierarchy in SystemOfDescription – Package .......................54
Figure 12. Implementation of hierarchy in the SystemOfDescription – Package .............................55
Figure 13. UML presentation of the Storage – Package .........................................................................56
Figure 14. UML presentation of the About – Package ...........................................................................57
Figure 15. UML presentation of the Creator – Package .........................................................................58
Figure 16. UML presentation of the Owner – Package ...........................................................................60
Figure 17. UML presentation of the multiple language support in the MetaMedia DS ....................58
Figure 18. Overview of the MultiMedia Description Schemes (MPEG-7) ..........................................62
Figure 19. Relational presentation of the Description Unit .....................................................................77
Figure 20. Relational presentation of the Segment ................................................................................78
Figure 21. Relational presentation of the Hierarchy ...............................................................................79
Figure 22. Relational presentation of the About – Package ...................................................................79
Figure 23. Relational presentation of the Storage – Package ..................................................................80
Figure 24. Relational presentation of the Creator – Package ..................................................................80
Figure 25. Relational presentation of the Owner – Package I .................................................................81
Figure 26. Relational presentation of the Owner – Package II ...............................................................81
Figure 27. Relational presentation of the ‘Utilities – Package’ ...............................................................82
Figure 28. Levels of Scheme 0 in Arch-metamedia .................................................................................85
Figure 29. Levels of Scheme 1 in Arch-metamedia .................................................................................85
Figure 30. Hierarchy of segments in video ‘Science Eye’ .......................................................................87
Figure 31. Hierarchy of the MPEG-7 Content Set (Scheme 2) ..............................................................89
Figure 32. Conversions in the MetaMedia project .................................................................................94
VIII. Table of Tables

Table 1. The Dublin Core Metadata Element Set.........................................................9
Table 2. Attributes for the Dublin Core Metadata Element Set......................................10
Table 3. The RDF data model of Figure 3 translated into XML (using namespaces)............13
Table 4. A simple XML document................................................................................20
Table 5. DTD for Table 4.............................................................................................21
Table 6. SQL syntax......................................................................................................26
Table 7. Examples of hierarchical video DS and levels......................................................53
Table 8. Mapping of the Archive of Eça de Queirós into the MetaMedia DS.......................84
Table 9. Example of information available of the MPEG-7 video......................................87
Table 10. Mapping of the ‘Science Eye’ video description into the MetaMedia DS.............88
IX. Abbreviations

ANSI : American National Standard Institute
API : Application Programming Interface
AV : Audio-visual
BLOB : Binary Large Object
CBR : Content Based Retrieval
CSS : Cascading Style Sheets
D : Descriptor
DB : Database
DBMS : Database Management System
DC : Dublin Core
DDL : Description Definition Language
DS : Description Scheme
DTD : Document Type Declaration/Definition
DOM : Document Object Model
HTML : HyperText Markup Language
HTTP : HyperText Transfer Protocol
ISAD(G) : General International Standard Archival Description
ISO : International Standardisation Organisation
MPEG : Moving Pictures Expert Group
OCLC : Online Computer Library Center
OODBMS : Object-Oriented Database Management System
PICS : Platform for Internet Content Selection
RDBMS : Relational Database Management System
RDF : Resource Description Framework
RXL : Relation to XML Transformation Language
SGML : Standard Generalised Markup Language
SQL : Standard Query Language
UML : Unified Modelling Language
URI : Uniform Resource Identifier
WWW : World-Wide Web
W3C : World-Wide Web Consortium
XMI : XML Metadata Interchange Language
XML : eXtensible Markup Language
X. Keywords

Metadata
Description Scheme
MPEG-7
Multimedia Documents
Multimedia Database
Content-based Information Retrieval
Mark-up Languages
XML
Relational Database
XI. **Glossary**

**P420**: Proposal for Video Description Scheme

**P422**: Proposal for Text and Still Image Description Scheme

**Data**: *(MPEG)* Data is audio-visual information that will be described using MPEG-7, regardless of storage, coding, display, transmission, medium, or technology.

**Feature**: *(MPEG)* A Feature is a distinctive characteristic of the data which signifies something to somebody.

**Descriptor (D)**: *(MPEG)* A Descriptor is a representation of a Feature. A Descriptor defines the syntax and the semantics of the Feature representation.

**Descriptor Value**: *(MPEG)* A Descriptor Value is an instantiation of a Descriptor for a given data set (or subset thereof).

**Description Scheme (DS)**: *(MPEG)* A Description Scheme specifies the structure and semantics of the relationships between its components, which may be both Descriptors and Description Schemes.

**Description**: *(MPEG)* A Description consists of a DS (structure) and the set of Descriptor Values (instantiations) that describe the Data.

**Description Definition Language (DDL)**: *(MPEG)* The Description Definition Language is a language that allows the creation of new Description Schemes and, possibly, Descriptors. It also allows the extension and modification of existing Description Schemes.

**Multimedia Description Scheme (MMDS)**: *(MPEG)* Multimedia Description Schemes refer to specific metadata structures being developed for MPEG-7 that are used to describe and annotate audio-visual data. In the MPEG-7 standard, the Multimedia Description Schemes are categorized as pertaining to visual, audio, or generic description. In general, the Multimedia Description Schemes (DSs) can contain MPEG-7 Descriptors and other DSs and can be extended for domain-specific applications.
**Content (Information):** *(MetaMedia)* information that describes the item itself and the relations with its (content) segments.

**Context (Information):** *(MetaMedia)* information that can not be extracted from the item, like information concerning the creation, storage, owner and relation with other items.

**Description Unit:** *(MetaMedia)* metadata for the item

**Document:** *(MetaMedia)* any item to be described and individually retrievable.

**BLOB:** a large object whose value is composed of unstructured binary ('raw') data.

**CLOB:** a large object whose value is composed of character data that corresponds to the database character set (defined for the Oracle 8i database).
1. Introduction

How do we find a photograph from a large archive that contains thousands or millions of pictures? How does a CNN video journalist find a specific clip from the myriad of videotapes, ranging from historical to contemporary, from sports to humanities? How do people organise and search the content of personal videotapes of family events, travel scenes, or social gatherings?

The era of ‘the information explosion’ has brought about the wide dissemination and use of visual information, particularly digital images and video, which we are also seeing in combination with text, audio and graphics.

A multimedia database system deals with the storage, manipulation and retrieval of all types of digitally presentable information objects such as text, still images, video and sound. Providing mechanisms that allow the user to retrieve desired multimedia information is an important issue in multimedia databases. Information about the content of the multimedia objects is contained within the multimedia objects and is usually not encoded into attributes. So, a picture of a football-game hides this content in the image itself, because there is no associated attribute like an event for example that clarifies for a search engine that this image is about a football-game.

Because content equality is not well defined, special techniques are needed for the retrieval of multimedia objects with content similar to that specified in the user’s query. In text databases, information retrieval techniques allow one to retrieve a document, if the document’s keywords are close to those specified in the query. But a multimedia document is a structured collection of attributes, text, image, video and audio date.

Multimedia document retrieval should be possible through the structure, attributes, and media content of the multimedia document. This poses several interesting challenges, due to the heterogeneity of data, the fuzziness of information, the loss of information in the creation of indexes, and the need of an interactive refinement of the query result.

A multimedia information retrieval system differs from a traditional information retrieval system in 2 main aspects:
- The structure of multimedia objects is more complex than the structure of typical textual data, handled by traditional information retrieval systems, and
- It is difficult to define similarities, since spatial and temporal aspects have to be taken in account.

Both aspects make it difficult to formulate an appropriate query and to search images and videos.

Content-based querying is the searching in multimedia databases based on the intrinsic properties of the multimedia, rather than on coded alphanumeric descriptors. Although the expectations are high for the new generation multimedia database systems, concerning automated content-based retrieval, research will still have to reveal if those expectations about content-based querying are going to be redeemed.

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

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 common practice 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\(^1\), and semantics, are very important here. Metadata in such cases provides a suitable basis for building the higher forms of information.

The intent of the MetaMedia Project ("Metadata for Multimedia Archives" [1]) has been to bring together telecommunications, information systems and archival researchers around a common set of goals:
- To participate in the MPEG-7 standardisation effort;
- To study the characteristics of a language suitable to specify description schemes for a wide range of kinds of documents;
- To define some particular description schemes for photography, video and historic archives;
- To develop an information system able to store and search such descriptions;
- To demonstrate the concept of automatic production of metadata.

\(^1\) An ontology is a hierarchical taxonomy of terms describing certain knowledge topics.
1. Introduction

The project team is composed of researchers from universities (Ienesc-Porto and the University of Minho) and from archives (Portuguese Photographic Centre, National Archives), already involved in projects in related areas, and has the support of an external consultant for the archival matters.

The participation in the MPEG-7 standardisation resulted in 2 proposals for MPEG-7 Description Schemes, namely a 'Video Description Scheme' (P420) and a 'Text and Still Image Description Scheme' (P422). The latter description scheme was used to develop a prototype, called Arch-MetaMedia, that is a relational database containing images of the Portuguese Photo Center, and which make part of the MPEG-7 Content Set. The implementation of the description scheme in this prototype and its implications for metadata concerning content and context are described in [3].

The continuation of this work, subject of this dissertation, contained three main aspects, namely unifying the previous Description Schemes, extending the data in the prototype with other domains and media, and perform test queries on data from different media and of different nature.

So the work started by equalising the Video Description Scheme and Text & Still Image Description Scheme, what resulted in the (unified) MetaMedia Description Scheme. This Description Scheme, defined in UML (see appendix A), was translated in relational schemes, which were then implemented in the Arch-MetaMedia prototype, so that the prototype would be able to support a wider range of multimedia data. To obtain more diversity in the database, the prototype was then enriched with some documents of the Archive of Eça de Queirós of the Fundação Eça de Queirós (under the custody of the Arquivo Distrital de Porto) and with a video (V10) of the MPEG-7 Content Set. With the resulting prototype, the advantages of metadata about content and context to conduct queries, are demonstrated.

The MetaMedia team of the University of Minho translated the MetaMedia DS in a XML presentation by using an intermediate language namely the XML Metadata Interchange Language (XMI). Their work also focused on the conversion from XML to the relational schemes, and the other way around. For the latter conversion the existing possibilities will be explored here.
Organisation of the dissertation

Chapter 2 presents the state of development of standards for the description of metadata. In particular, the norms of Dublin Core, Resource Description Framework and MPEG-7 are explored.

In chapter 3 markup languages (SGML, HTML and XML) are presented, as they play an important role in metadata initiatives.

Chapter 4 explains relational database issues, in particular those specific for multimedia data. The possibilities of exporting the query results of the relational database to XML are also discussed, for reasons of easy information exchange.

The state of the art in multimedia information retrieval is explored in chapter 5.

The unified MetaMedia Description Scheme is described in chapter 6, where a comparison with the current status of the MPEG-7 Multimedia Description Schemes is also presented.

The development of the prototype, Arch-MetaMedia, is the subject of chapter 7. The prototype is the practical implementation of the Description Scheme discussed in chapter 6. Here are also the results of the queries performed in the database presented.

Finally, in chapter 8 the conclusions of this work can be found, together with future work.
2. Metadata

2.1 Introduction

Searching today is essentially a process of matching the query terms to words in a document. So, in most methods, if the terms don’t match, then the document will not be retrieved, no matter how much it is about the subject of the query. In this context, there is a role reserved for metadata.

*Metadata* is information on the organisation of data, the various data domains and the relationships between them, or in short, metadata is ‘data about data’ [4]. Annotations, a set of keywords and metadata on history, age and quality of data are some very specific, well-known types of metadata. Other forms of metadata are information about authenticity, availability and accessibility, digital signatures, copyright, reproduction, etc. Metadata can be created at the time of creation of an object, either by or under the auspices of the creator. It can also be added later as part of the traditional cataloguing process.

Two categories of metadata are *descriptive metadata*, which is metadata that is external to the meaning of the document and pertains more how it was created, and *semantic metadata*, which characterises the subject matters that can be found within the document’s contents. Whether in the traditional context or in the Internet context, the purpose of metadata is usually to facilitate and improve the retrieval of information, but there are other purposes like for example defining the conditions of access to a document...

By attaching metadata to documents, various processing steps can conveniently use the metadata. 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 that cannot be answered by processing raw material only. This is usually the case when metadata stems from manual augmentation of media.
2.2 Metadata on the Internet

With the widespread use of electronic communication, the concern with information retrieval passed from the specialised community of the documentation sciences to the global scientific and technical community, and even to the lay person. Available online resources grow rapidly, therefore making information retrieval a vital task and spreading the research on metadata. The purposes of metadata on the Web can be found in areas of cataloguing, content rating, intellectual property rights, digital signatures, privacy levels, applications in electronic commerce.

Three methods of transmitting metadata about a document exist on the WWW [5]:

- The metadata can be contained within the document itself (in the form of HTML META tags or XML tags, see next chapter);
- It can be passed along in the HTTP header when the client and server communicate to transfer the document; or
- It can exist outside the document altogether.

The history of metadata [6] at the W3C\(^2\) began in 1995 with PICS, the Platform for Internet Content Selection. PICS is a platform for communicating ratings at Web pages from a server to clients. These ratings, or rating labels, contain information about the content of Web pages: for example, whether a particular page contains sex or violence. Instead of being a fixed set of criteria, PICS introduced a general mechanism for creating rating systems. Different organisations could rate content based on their own objectives and values, and users could set their browser to filter out any Web pages not matching their own criteria. PICS is a restricted metadata framework. It allows certain things to be expressed very precisely about Web pages; in particular.

Inspired by PICS, a general metadata framework was developed under the name Resource Description Framework, RDF, (which will be discussed further in 2.5). Once the Web has been sufficiently ‘populated’ with rich metadata, it is expected that the Web of today, (a vast amount of unstructured mass of information), may be transformed in the future into something more manageable- and thus something far more useful.

\(^2\) W3C was created to develop common protocols that enhance the interoperability and promote the evolution of the WWW. It is an industry consortium jointly run by the MIT Laboratory for Computer Science in the US, the National Institute for Research in Computer Science and Control in France, and Keio University in Japan. Services provided by the consortium include: a repository of information about the WWW for developers and users, reference code implementations to embody and promote standards, and various prototypes and sample applications to demonstrate use of new technology.
2.3 Metadata standards

The value of metadata is limited if there is no common agreement on what elements to use or what their content should be. Metadata cannot fully serve its purpose unless it is subjected to a certain amount of standardisation.

The effective use of metadata among applications requires common conventions about semantics, syntax and structure. Individual resource description communities define the semantics, or meaning, of metadata that address their particular needs. Syntax, the systematic arrangement of data elements for machine processing, facilitates the exchange and use of metadata among multiple applications. Structure can be thought of as a formal constraint on the syntax for the consistent representation of semantics [7].

Challenges of defining metadata are:

- To be able to retrieve items on a specified topic regardless of the media type, therefore metadata should not be specialised by media.
- Functionality and simplicity need to be balanced, because people won’t be willing to generate complex metadata descriptions for simple items.
- Interoperability is a requirement to ensure common grounds for multimedia descriptions, although different specifications may exist for different purposes.
- The rapid growth of information needs requires extensibility.

2.4 ISAD(G)

The General International Standard Archival Description, ISAD(G) [2], was developed by the ICA Ad Hoc Commission on Descriptive Standards (ICA/DDS), which became a permanent committee at the ICA International Congress on Archives in Beijing, China, in 1996. This standard provides general guidance for the preparation of archival descriptions, where the purpose of archival description is to identify and explain the context, and content, of archival material in order to promote its accessibility.

The ISAD standard organises archives in a multilevel way, where the following multilevel description rules govern:

- Description from the general to the specific
- Information relevant to the level of description
- Linking of descriptions
- Non-repetition of information
ISAD identifies or defines 26 elements that may be combined to constitute the description of an archival entity. The rules are organised into 7 areas of descriptive information:

1. **Identity Statement Area**: where essential information is conveyed to identify the unit of description. (Elements: reference code; title; date(s); level of description; extent and medium of the unit of description (quantity, bulk, or size))

2. **Context Area**: where information is conveyed about the origin and custody of the unit of description. (Elements: name of creator(s); administrative/biographic history; archival history; immediate source of acquisition or transfer)

3. **Content and Structure Area**: where information is conveyed about the subject matter and arrangement of the unit of description. (Elements: scope and content; appraisal, destruction and scheduling information; accruals; system of arrangement)

4. **Conditions of Access and Use Area**: where information is conveyed about the availability of the unit of description. (Elements: conditions governing access; conditions governing reproduction; language/scripts of material; physical characteristics and technical requirements; finding aids)

5. **Allied Materials Area**: where information is conveyed about materials having an important relationship to the unit of description. (Elements: existence and location of originals; existence and location of copies; related units of description; publication note)

6. **Notes Area**: where specialised information and information that cannot be accommodated in any of the other areas may be conveyed. (Elements: note)

7. **Description Control Area**: where information is conveyed on how, when and by whom the archival description was prepared. (Elements: archivist's note; rules or conventions; date(s) of descriptions)

### 2.5 Dublin Core

The Dublin Core (DC) grew out of a workshop sponsored by the OCLC\(^3\) and the National Centre for Supercomputing Applications in 1995. It was a response to the need to improve retrieval of information resources, especially on the WWW [8]. DC is intended to be a basic collection of metadata elements. The Core was originally conceived to be used by content creators, but interest has become widespread among specialised resource description groups, such as museums and libraries.

---

\(^3\) The Online Computer Library Centre is a non-profit, membership, computer library service and research organisation whose computer network and services link more than 25000 libraries in 63 countries and territories. (http://www.oclc.org)
The DC Metadata Element Set, Version 1.1 – Reference Description [9] contains 15 metadata elements (see Table 1), which can be divided in 3 groups, namely content, intellectual property and instantiation.

Table 1: The Dublin Core Metadata Element Set [9].

<table>
<thead>
<tr>
<th>Element</th>
<th>Group</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td>Content</td>
<td>A name given to the resource.</td>
</tr>
<tr>
<td>Creator</td>
<td>Intellectual property</td>
<td>An entity primarily responsible for making the content of the resource.</td>
</tr>
<tr>
<td>Subject</td>
<td>Content</td>
<td>The topic of the content of the resource.</td>
</tr>
<tr>
<td>Description</td>
<td>Content</td>
<td>An account of the content of the resource.</td>
</tr>
<tr>
<td>Publisher</td>
<td>Intellectual property</td>
<td>An entity responsible for making the resource available.</td>
</tr>
<tr>
<td>Contributor</td>
<td>Intellectual property</td>
<td>An entity responsible for making contributions to the content of the resource.</td>
</tr>
<tr>
<td>Date</td>
<td>Instantiation</td>
<td>A date associated with an event in the life cycle of the resource.</td>
</tr>
<tr>
<td>Type</td>
<td>Instantiation</td>
<td>The nature or genre of the content of the resource.</td>
</tr>
<tr>
<td>Format</td>
<td>Instantiation</td>
<td>The physical or digital manifestation of the resource.</td>
</tr>
<tr>
<td>Identifier</td>
<td>Instantiation</td>
<td>An unambiguous reference to the resource within a given context.</td>
</tr>
<tr>
<td>Source</td>
<td>Content</td>
<td>A reference to a resource from which the present resource is derived.</td>
</tr>
<tr>
<td>Language</td>
<td>Content</td>
<td>A language of the intellectual content of the resource.</td>
</tr>
<tr>
<td>Relation</td>
<td>Content</td>
<td>A reference to a related resource.</td>
</tr>
<tr>
<td>Coverage</td>
<td>Content</td>
<td>The extent or scope of the content of the resource.</td>
</tr>
<tr>
<td>Rights</td>
<td>Intellectual property</td>
<td>Information about rights held in and over the resource.</td>
</tr>
</tbody>
</table>

A DC Metadata Initiative Working Draft [9a] proposes a list of resource types. Recommended best practice is to select a value from a controlled vocabulary. The following list is offered as default list of resource types: collection, dataset, event, image, interactive resource, model, party, physical object, place, service, software, sound, text. Each DC element is defined using a set of 10 attributes, see Table 2.
Table 2: Attributes for the Dublin Core Metadata Element Set [9a].

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Default Value</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>-</td>
<td>The label assigned to the data element.</td>
</tr>
<tr>
<td>Identifier</td>
<td>-</td>
<td>The unique identifier assigned to the data element.</td>
</tr>
<tr>
<td>Version</td>
<td>1.1</td>
<td>The version of the data element.</td>
</tr>
<tr>
<td>Registration authority</td>
<td>Dublin Core Metadata Initiative</td>
<td>The entity authorised to register the data element.</td>
</tr>
<tr>
<td>Language</td>
<td>EN</td>
<td>The language in which the data element is specified.</td>
</tr>
<tr>
<td>Definition</td>
<td>-</td>
<td>A statement that clearly represents the concept and essential nature of the data element.</td>
</tr>
<tr>
<td>Obligation</td>
<td>Optional</td>
<td>Indicates if the data element is required to always or sometimes be present (contain a value).</td>
</tr>
<tr>
<td>Datatype</td>
<td>Character string</td>
<td>Indicates the type of data that can be represented in the value of the data element.</td>
</tr>
<tr>
<td>Maximum occurrence</td>
<td>Unlimited</td>
<td>Indicates any limit to the repeatability of the data element.</td>
</tr>
<tr>
<td>Comment</td>
<td>-</td>
<td>A remark concerning the application of the data element.</td>
</tr>
</tbody>
</table>

There is a continuing tension between the original design for simplicity and the desire of specialised groups to increase the level of detail to meet their needs. This is called the minimalist / structuralist issue [8]. In the minimalist point of view, the simplicity is important both for creation of metadata and for the use of metadata by tools. The goal of semantic interoperability across communities can only be achieved if there is a simple core of elements that are understood to mean the same thing in every case. Specialised groups may need to develop their own list to extend the Core.

2.6 Resource Description Framework

2.6.1 Introduction

The Resource Description Framework (RDF) [10], developed under the auspices of the W3C, is an infrastructure that enables the encoding, exchange, and reuse of structured metadata [7]. This infrastructure enables metadata interoperability through the design of mechanisms that support common conventions of semantics, syntax and structure. RDF is the result of a number of metadata communities bringing together their needs to provide a robust and flexible architecture for supporting metadata on the Web.
The broad goal of RDF [6] is to define a mechanism for describing resources that makes no assumption about a particular application domain, nor defines the semantics of any application domain. RDF metadata can be used in a variety of application areas, for example:

- In resource discovery to provide better search engine capabilities;
- In cataloguing for describing the content and content relationships available at a particular Web site, page, or digital library;
- By intelligent software agents to facilitate knowledge sharing and exchange;
- In content rating;
- In describing collections of pages that represent a single logical ‘document’;
- For describing intellectual property rights of Web pages.

2.6.2 RDF Data Model

RDF data basically consists of nodes and attached attribute/value pairs. Figure 2 presents a simple example.

![Diagram]

Figure 2: A simple example of the RDF data model [10a].

The figure represents the data model corresponding to the statement “the author of thesis ‘Metadata’ is Karen Goethals”. ‘Thesis Metadata’ is a single resource of the data model, where a resource is defined as any object that is uniquely identifiable by an Uniform Resource Identifier (URI). The properties associated with resources are identified by property-types, and those property-types have corresponding values, which are either atomic (text strings, numbers, etc.) or other resources or metadata instances. So in the above example, there is a property-type “author” and a corresponding value of “Karen Goethals”.

To distinguish characteristics of the data model, the RDF Model and Syntax Specification [10a] represents the relationships among resources, property-types, and values in a directed labelled graph. In this case, resources are identified as nodes, property-types are defined as directed label arcs, and string values are quoted.
2.6.3 RDF Syntax

The essence of RDF is the model of nodes, attributes, and their values. A syntax [10a] representing this model is required to store instances of this model into machine-readable files and to communicate these instances among applications. XML (see 3.4) has been adopted for this syntax.

RDF is designed to have the following characteristics:

- Independence,
- Interchange,
- Scalability,
- PropertyTypes are Resources,
- Values can be Resources,
- Properties can be Resources.

The reason why XML alone is not enough, is that XML falls apart on the scalability design goal. Problems of XML are that the order in which elements appear in an XML document is significant and often very meaningful, and (that) XML allows complicated constructions which are hard to handle (even in moderate amounts). RDF properties are simple 3-part records (Resource, PropertyType, Value), so they are easy to handle and look things up, even in large numbers. On the other hand, XML is a necessary part of the solution to RDF's interchange design goal.

2.6.4 RDF Schema

RDF supports the use of conventions that will facilitate modular interoperability among separate metadata element sets. RDF provides a means for publishing both human-readable and machine-processable vocabularies⁴. RDF in itself does not contain any predefined vocabularies for authoring metadata. Anyone can design a new vocabulary, the only requirement for using it is that a designating URI is included in the metadata instances using this vocabulary. A central registry is not required. This permits communities to declare vocabularies that may be reused, extended or refined to address applications or domain specific descriptive requirements.

RDF Schemas [10b] are used to declare vocabularies. They define the valid properties in a given RDF description, as well as any characteristics or restrictions of the property-type values themselves. The XML Namespace mechanism (see 3.4.6.5) serves to identify RDF schemas. XML namespaces provide a method for unambiguously identifying the semantics and
conventions governing the particular use of property-types by uniquely identifying the
governing authority of the vocabulary. So for the example of Figure 2, we can replace the
property-type "author" by the "creator" element defined by the Dublin Core. (see Figure 3)

\[ \text{Dissertation A} \xrightarrow{\text{DC: Creator}} \text{"Karen Goethals"} \]

Figure 3: A simple example of the RDF data model, using the DC vocabulary.

The corresponding syntactic way of expressing this statement using XML namespaces to
identify the use of the DC Schema is shown in Table 3.

\begin{table}
\centering
\begin{tabular}{|l|}
\hline
< ? xml : namespace ns = "http://www.w3.org/RDF/RDF/" prefix = "RDF" ?>
< ? xml : namespace ns = "http://purl.oclc.org/DC" prefix = "DC" ?>

< RDF : RDF>
  < RDF : Description RDF : HREF = "http://www.fe.up.pt/~mgi99015" >
    < DC : Creator > Karen Goethals < /DC : Creator >
  < /RDF : Description >
< /RDF : RDF>
\hline
\end{tabular}
\end{table}

Table 3: The RDF data model of Figure 3 translated into XML, (using Namespaces).

2.7 MPEG-7

2.7.1 Introduction

With the availability of more and more audio-visual information, there is an increasing
difficulty to find desired items. While solutions exist that allow searching for textual
information, searching information on audio-visual content however, is very difficult, as no
generally recognised or standardised description of this material exist. This need originates the

\(^4\) Vocabularies are the set of properties, or metadata elements, defined by resource description communities.

MPEG-7 is concerned with the standardisation of a ‘Multimedia Content Description Interface’, so MPEG-7 will be a standardised description of various types of multimedia information [11a]. This description will be associated with the content itself, to allow fast and efficient searching for material that is of interest to the user. The standard does not comprise the (automatic) extraction of features. Nor does it specify the search engine (or any other program) that can make use of the description. (see Figure 4)

![Diagram of multimedia standardisation](image)

**Figure 4: Scope of the MPEG-7 standardisation [11a].**

The difference with other norms of the MPEG family is that MPEG-7 intends to provide a representation of *information about the content*, while MPEG-1, 2, and 4 give a representation of *the content itself\(^6\).*

### 2.7.2 Goals

MPEG-7's goal [11a] is to describe multimedia content, while managing the data flexibly and globalising the data resources. (MPEG-7 data may be physically located with the associated AV material, in the same data stream or on the same storage system, but the descriptions could also live somewhere else on the globe.) Therefore, the goal of the standard is the development of a


\(^6\) MPEG-1 and 2 in a waveform or sample-based, frame-based presentation; MPEG-4 in an object-based presentation.
A *Descriptor* defines the syntax and the semantics of the feature representation, where a *feature* is a distinctive characteristic of the data signifying something to somebody (e.g. colour histogram, text of the title, etc.). *Description Schemes* are models of the multimedia objects and of the universes that they represent, (e.g. the data model of the description). They specify the types of the Descriptors that can be used in a given description, and the relationships between these Descriptors and between other Description Schemes. (e.g. a movie, temporally structured as scenes and shots, including some textual descriptors at the scene level, and colour, motion and some audio descriptors at the shot level.) The *Description Definition Language* is a language that allows the creation of new Description Schemes and, possibly, Descriptors. It also allows the extension and modification of existing Description Schemes. DDL is not a modelling language such as the Unified Modelling Language (UML), but it is rather to be used to represent the result of modelling, i.e. DSs and Ds. MPEG-7 has built its DDL on W3C’s XML Schema Language (see 3.4.5).
2.7.3 Applications

MPEG-7 will allow different granularity in its descriptions, offering the possibility to have different levels of discrimination. The level of abstraction is related to the way the features can be extracted: many low-level features can be extracted in fully automatic ways, whereas high-level features need (much) more human interaction. Because the descriptive features must be meaningful in the context of the application, they will be different for different user domains and different applications.

Applications for MPEG-7 [11c] can be categorised in 2 domains, namely ‘pull’ and ‘push’ applications. Pull, or retrieval, applications involve databases, audio-visual archives, and the Web-based Internet paradigm (a client requests material from a server). These applications remain for many the primary applications for MPEG-7, although the proposed multimedia content descriptors are acknowledged to serve much more than search applications. Push applications on the other hand are more related to broadcasting and the emerging webcasting. This means that the push applications need selection and filtering, (instead of indexing and retrieval). MPEG-7 application domains can be found in journalism, cultural services, geographical information systems, biomedical applications, film, video and radio archives, to name just a few.
3. Markup Languages

3.1 Introduction

Markup is defined [4] as extra textual syntax that can be used to describe formatting actions, structure information, text semantics, attributes, display, presentation of information,... A markup language provides the 'lexical' and 'syntactic' structure of a document. The lexical structure of a document defines [12] how the components of information it contains are recognised (or what characters appear in a document and how those characters can be used to make up words). The syntactic structure defines interrelationships between these components (or how these words can be put together to form a structure). The marks used in a markup language are called 'tags'. Usually there is an initial and ending tag surrounding the marked item, to avoid ambiguity.

The beginning of markup languages was marked in the late 1960's, when Charles Goldfarb coined the term 'markup language' while working on GML (Generalised Markup Language), an IBM research project, that automated documents in general and led to the development of SGML in the 1970's [13].

Presentation is an important aspect of the markup languages. As use of markup became more widespread in the 1970's and 1980's, people realised that the best way to use markup was by a formal separation of structure and presentation. The standard metalanguage for markup is SGML.

3.2 Standard Generalised Markup Language (SGML)

The Standard Generalised Markup Language (SGML) [14] is a metalanguage for tagging text. SGML provides the rules for defining a markup language based on tags [15]. By defining the role of each piece of text in a formal model, users of programs based on the SGML can check that each element of text is used in the correct place. The difference with other markup languages is that SGML not simply indicates where a change of appearance occurs, or where a new element starts. SGML wants to clearly identify the boundaries of every part of a document,
without being told where every part starts or ends. Therefore, rules are provided that allow the computer to recognise where the various elements of a text entity start and end. When these rules are carefully applied, the amount of coding that needs to be entered by a human operator can be reduced to a bare minimum. An SGML document consists of a description of the structure of the document and the text itself marked with tags that describe the structure.

The lexical structure of a document is defined in an SGML Declaration, while the syntactic structure is defined by a Document Type Declaration (DTD). Tags are denoted by angle brackets. An ending tag is specified by adding a slash before the tag name, for example:

\[<\text{tag}>\text{marked text}</\text{tag}>\]

### 3.3 HyperText Markup Language (HTML)

The HyperText Markup Language (HTML) [16] is an instance of SGML. HTML was created in 1992 and has evolved during the past years. It is a simple language, well suited for hypertext\(^7\), multimedia, and the display of small and simple documents.

Most documents on the Web are currently stored and transmitted in HTML. Other media, like images or audio can be embedded within HTML documents. There exists also a field for metadata, which might be used for instance by search-engines. In 1997 the Cascade Style Sheets (CSS) were introduced, to fix the presentation style of a document. Style sheets separate information about presentation from document content, which in turn simplifies Web site maintenance, promotes Web accessibility, and makes the Web faster.

#### 3.3.1 XHTML

XHTML 1.0 is W3C’s recommendation [16a] for the latest version of HTML (following from earlier work on HTML 4.01, 4.0, 3.2 and 2.0). XHTML 1.0 is the first major change to HTML since HTML 4.0 was released in 1997. It brings the rigor of XML (see 3.4) to Web pages and is the keystone in W3C’s work to create standards that provide richer Web pages on an ever increasing range of browser platforms including cell phones, televisions, cars, wallet sized wireless communicators, kiosks and desktops.

XHTML is modular making it easy to combine with markup tags for things like vector graphics, multimedia, electronic commerce and more. XHTML 1.0 is specified in 3 ‘flavours’. You specify which of these variants you are using by inserting a line at the beginning of the

\[^7\) Hypertext is high-level interactive navigational structure allowing the user to browse text non-sequentially on a computer screen. It consists basically of nodes which are correlated by directed links in a graph structure.\]
document. Each variant has its own Document Type Declaration (see 3.4.4) which sets out the rules and regulations for using HTML in a succinct and definitive manner.

- XHTML Transitional: this 'flavour' is to take advantage of XHTML features including style sheets, but nonetheless to make small adjustments to markup for the benefit of those viewing Web pages with older browsers which can't understand style sheets.
- XHTML Strict: for a really clean structural markup, free of any tags associated with layout. (To get the desired font, colour and layout effects, it should be used together with W3C's Cascading Style Sheet language [17]).
- XHTML Frameset: with this 'flavour' you can use HTML frames to partition the browser window into 2 or more frames.

3.4 Extensible Markup Language (XML)

3.4.1 Origin and goals

XML stands for eXtensible Markup Language [18] and was developed by a XML Working Group formed under the auspices of the W3C in 1996, and which was chaired by Jon Bosak of Sun Microsystems. While SGML was created for general document structuring and HTML as an application of SGML for Web documents, XML is intended to be a simplification of SGML for general Web use. The goal of XML is to enable the delivery of self-describing data structures of arbitrary depth and complexity to applications that require such structures.

XML is not a markup language, as HTML is, but a metalanguage that is capable of containing markup languages in the same way as SGML. XML allows a human-readable semantic markup, which is also machine-readable. XML is a data representation with the characteristics of a document. Therefore, one can first parse a XML file to extract the data and then process it (machine-centric use), or one can present the file as a document by using style sheets (human-centric use). This helps us to do both things in the same application at the same time.

XML and HTML were designed with different goals: XML was designed to describe data and focus on what data is, while HTML was designed to display data and focus on how data looks. XML is not a replacement for HTML. In the future development of the Web it is most likely that XML will be used to structure and describe Web data, while HTML will be used to format and display the same data.
So how can XML be used? First, XML can keep data separated from your HTML: data can be stored in a separate XML file, and formatted and displayed by HTML. This way changes in the underlying data won’t have any influence on the HTML code. Second, XML can also store data inside HTML documents (as ‘data islands’). Third, XML can be used to exchange data between computer systems and databases that contain data in incompatible formats. Fourth, XML can be used to store data in files or databases.

3.4.2 Semi-structured

XML is nothing else than a particular standard syntax for semi-structured data exchange. Semi-structured data refers to data with some of the following characteristics [19]:

- The schema is not given in advance and may be implicit in the data.
- The schema is relatively large (compared to the size of the data) and may be changing frequently.
- The schema is descriptive rather than ‘prescriptive’, i.e. it describes the current state of the data, but violations of the schema are still tolerated.
- The data is not strongly typed, i.e. for different objects, the values of the same attribute may be of different types.

3.4.3 Syntax

Table 4: A simple XML document

```xml
< ? xml version = "1.0" ? >
< book type = "dissertation" >
  < title > Metadata-oriented Multimedia Information Retrieval </title >
  < author >
  < firstname > Karen </ firstname >
  < lastname > Goethals </ lastname >
  </ author >
</ book >
```

A simple XML document is presented in Table 4. The first line is the XML declaration, which should always be included. All XML documents must contain a single tag pair to define the root element. All other elements must be nested within the root element. All elements can have sub (children) elements. Sub elements must be in pairs and correctly nested within their parent
3. Markup Languages

In the example of Table 4, <book> is the root element and contains 2 child elements, namely <title> and <author>. <first name> and <last name> are 'sub children' of <author>. XML elements can have attributes in name/value pairs just like in HTML. 'Type' is an attribute of the root element <book>. Attribute values must always be quoted in XML.

Other differences compared to HTML are that all XML elements must have a closing tag, XML tags are case sensitive (a capital letter is different from a small letter) and XML elements must be properly nested. A 'well formed' XML document is a document that conforms to the XML syntax rules.

3.4.4 Document Type Declaration (DTD)

XML allows an application independent way of sharing data. With a Document Type Declaration (DTD), independent groups of people can agree to use a common DTD for interchanging data.

Declarations allow a document to communicate meta-information to the parser\textsuperscript{8} about its content. Meta-information includes the allowed sequence and nesting of tags, attribute values and their types and defaults, the names of external files that may be referenced and whether or not they contain XML, the formats of some external (non-XML) data that may be referenced, and the entities that may be encountered. [20]

A DTD for Table 4 might look like the DTD in Table 5.

\begin{center}
\textit{Table 3: DTD for the example of Table 4.}
\end{center}

\begin{verbatim}
<!ELEMENT book (title+, author*)>
<!ATTLIST book type (dissertation | article) dissertation>
<!ELEMENT title (#PCDATA)>
<!ELEMENT author (firstname?, lastname)>
<!ELEMENT firstname (#PCDATA)>
<!ELEMENT lastname (#PCDATA)>
\end{verbatim}

There are 4 kinds of declarations in XML:

- Element type declarations identify the names of elements and the nature of their content.

The first line of Example 2 defines an element 'book' with minimum 1 'title' element (denoted by the '+'), and zero or more 'author' elements (denoted by the '*').

---

\textsuperscript{8} A XML parser is used to read and update (create and manipulate) a XML document.
The title element contains PCDATA, which stands for parsed character data. PCDATA is
text that will be parsed by a parser. Tags inside the text will be treated as markup and
etties will be expanded. The author element can have 2 sub-elements: an optional
‘firstname’ element (‘?’ denotes zero or one occurrence) and one ‘lastname’ element. Both
sub-elements contain PCDATA. (Another option is CDATA or character data: the text will
not be parsed by a parser.)

- Attribute list declarations identify which elements may have attributes, what attributes they
  may have, what values the attributes may hold, and what value is default.
  So the second line of Table 5 defines an attribute called ‘type’ for the element ‘book’. The
  value for ‘type’ can only take 2 values: ‘dissertation’ or ‘article’. The former is the default
  value. Instead of working with default values, one can also work with implied, required or
  fixed attribute values.

- With an entity declaration (< ! ENTITY ... >), a name can be associated with some
  other fragment of content.

- Notation declarations identify specific types of external binary data. This information is
  passed to the processing applications, which may make whatever use of it as it wishes.
  (e.g. < ! NOTATION GIF87A SYSTEM "GIF" >, which means that an image with the
  name ‘GIF87A’ and of format ‘gif’ can be found be the system.)

A ‘valid’ XML document is a ‘well formed’ XML document which conforms to the rules of a
DTD, otherwise it is a ‘non-valid’ XML document (that might be ‘well formed’ though).

3.4.5 Schemas

A schema [21] is a model for describing the structure of information. The term has its origin in
the database world, but in the context of XML, a schema describes a ‘model’ for a whole class
of documents. The model describes the possible arrangements of tags and text in a valid
document. Or in other words: a schema is an agreement on a common vocabulary for a
particular application that involves exchanging documents. In schemas, models are described in
terms of constraints. Content model constraints [22a] describe the order and sequence of
elements, and datatype constraints [22b] describe valid units of data. The purpose of a schema is
to allow machine validation of document structure.

DTDs are the schema mechanism for SGML, but they have some limitations:

- They are written in a different (non-XML) syntax.
- They have no support for namespaces (see 3.4.7).
- They only offer extremely limited data typing.
They have a complex and fragile extension mechanism based on string substitution.

XML schemas try to overcome these limitations and attempt to be more expressive than DTDs. New features offered by XML schemas [22] are:

- Richer datatypes
- User defined types (called Archetypes)
- Attribute grouping
- Refinable archetypes, or "inheritance"
- Namespace support

Although XML schemas are expected to replace the DTDs in the future, in the short term DTDs still have a number of advantages:

- Widespread tool support: all SGML tools and many XML tools can process DTDs.
- Widespread deployment: a large number of document types are already defined using DTDs (e.g. HTML, TEI\(^9\)).
- Widespread expertise and many years of practical applications.

### 3.4.6 Extensible Stylesheet Language (XSL)

In order to display XML documents, it is necessary to have a mechanism to describe how the document should be displayed. The eXtensible Stylesheet Language (XSL) [23] is a language for transforming and formatting XML. XSL is the stylesheet language of XML.

One of the design goals for XSL was to make it possible to transform data from one format to another on a server, returning readable data to all kinds of browsers.

### 3.4.7 Namespaces in XML

XML namespaces [24] is a proposal for a simple method to be used for qualifying names used in XML documents by associating them with schemas, identified by Uniform Resource Identifier.

\(^9\) TEI, or Text Encoding Initiative, is an international project to develop guidelines for the encoding of textual material in electronic form for research purposes. (http://www.hcu.ox.ac.uk/TEI/)
A document instance of a XML application may contain markup defined in multiple schemas, which may have been authored independently. A reason to work this way is that it is beneficial to reuse parts from existing, well-defined schemas, and it has the advantage of allowing search engines or other tools to operate over a range of documents that vary in many respects but use common names for common element types. Therefore, document constructs should have universal names. The combination of the universally-managed URI namespace and the local schema namespace produces names that are guaranteed universally unique.
4. Relational Databases

4.1 Introduction

The term database refers to a collection of data that is managed by a database management system, also called a DBMS, or DB system [25]. The first commercial DBMSs appeared in the late 1960's. These evolved from file systems, that stored data over a long period of time, and allowed the storage of large amounts of data. (However, file systems do not generally guarantee that data cannot be lost if it is backed up, and they don't support efficient access to data items whose location in a particular file is not known.) The first important applications of DBMSs were the ones where data was composed of many small items, and many queries or modifications were made, like in banking and airline reservations systems.

The early DBMSs encouraged the user to visualise data much as it was stored. These DB systems used several different data models for describing the structure in a DB, the most common among them were the 'hierarchical' or tree-based model and the graph-based 'network' model. The latter was standardised in the late 1960s through a report of the Committee on Data Systems and Languages. A problem with these early models and systems was that they did not support high-level query languages.

4.2 Relational model

A paper written by Codd [26] in 1970 changed DB systems significantly. He proposed that DB systems should present the user with a view of data organised as tables, called relations. Behind the scenes, there might be a complex data structure that allowed rapid response to a variety of queries. Unlike earlier DB systems, the user of a relational system would not be concerned with the storage structure. A RDBMS, as defined by Codd, has 3 major parts:

- Data that is presented as tables
- Operators for manipulating tables
- Integrity rules on tables
Queries could be expressed in a very high-level language, which greatly increased the efficiency of DB programmers. SQL has become the most popular language for RDBMS [27], as well as an ANSI and ISO standard. It provides operators for manipulating tables. The SQL `select` statement queries tables. An example of the syntax of the select command is presented in Table 6.

‘Table-1’, ‘table-2’ and any others are combined into one temporary table. The attribute list specifies which columns should be retained in the temporary table. The predicate expression specifies which rows should be retained.

Table 6: SQL syntax [27].

```
SELECT attribute-list
FROM table-1, table-2,...
WHERE predicate-is-true;
```

4.3 Architecture of a DBMS

![Diagram showing the major components of a DBMS](image)

Figure 6: Major components of a DBMS [25].
Figure 6 shows the major components of a DBMS [27]. The disk-shaped component indicates the place where data is stored. Also metadata is stored in this component. (For a relational DBMS this includes for example the names of the relations.) Often a DBMS maintains indexes for the data. An index is a data structure that helps us to find data items quickly, given a part of their value.

The task of a storage manager is to obtain requested information from the data storage and to modify the information there when requested to by the levels of systems above it. The query processor handles not only queries, but also requests for modification of the data or the metadata. Its job is to find the best way to carry out a requested operation and to issue commands to the storage manager that will carry them out.

The transaction manager component is responsible for the integrity of the system. The ‘proper’ execution of transactions requires the following properties:

- **Atomicity**: either all of a transaction is executed or none of it is.
- **Consistency**: the database has a notion of a ‘consistent state’, for example, an appropriate consistency condition for an airline database is that no seat is assigned to two different customers.
- **Isolation**: when 2 or more transactions run concurrently, their effects must be isolated from one another.
- **Durability**: if a transaction has completed its work, its effect should not get lost in case the system fails, even if it fails immediately after the transaction completes.

The transaction manager interacts with the query manager, since it must know what data is being operated upon by the current queries, and it may need to delay certain queries or operations so that these conflicts do not occur. It interacts with the storage manager because schemes for protection of data usually involve storing a ‘log’ of changes to the data. By properly ordering operations, the log will contain a record of changes so that after a system failure even those changes that never reached the disk can be re-executed.

There are 3 types of input to the DBMS: queries, which are questions about the data; modifications, which are operations to modify the data; and schema modifications, which are commands issued by authorised personnel (DB administrator), who are allowed to change the schema of the DB or create a new DB.

It is common to use a client-server architecture for a DB system, in which requests by one process (the client) are sent to another process (the server) for execution. Therefore, the work of the components of Figure 6 is divided into a server process and one or more client processes.

In the simplest client/server architecture, the entire DBMS is a server, except for the query interfaces that interact with the user and send queries or other commands to the server.
Relational systems generally use the SQL language for representing requests from the client to the server. The DB server then sends the answer, in the form of a table or relation, back to the client.

4.4 Multimedia DBMS

4.4.1 Multimedia-specific aspects

A multimedia DBMS provides a suitable environment for using and managing multimedia DB information [28]. Therefore, it must not only support the various multimedia data types, but also provide facilities for traditional DBMS functions like DB definitions and creation, data retrieval, data access and organization, data independence, privacy, integrity control, version control, and concurrency support.

Multimedia document management and processing is one of the application domains for multimedia DBs. Multimedia document management is used in more specific application domains such as CAM (Computer Aided Manufacturing), education, geographic information systems, etc.

The specific nature of multimedia data also makes it important to support new special functions, like the management of huge volumes of multimedia data, effective storage management, information retrieval and handling of spatial and temporal data objects.

4.4.2 Multimedia in traditional DBMS

The integration of multimedia data in a traditional DBMS is not an easy task. Multimedia data is inherently different from conventional data. The main difference is that information about the content of multimedia is not encoded into attributes provided by the data schema (structured data). Text, image, video and audio data are typically unstructured. While a file of a person for example, can be stored in a traditional DBMS by declaring attributes like name, address, birth… and the information of each person can be attached to those attributes, this is not the case for images, where you can’t ‘split’ the information in pieces and code it into attributes without loss of information. Therefore, specific methods to represent content features and semantic features of multimedia data are needed.
Many current multimedia DBMSs are relational or object-oriented DBs with schema designed to represent the various levels of information and which have had their functionality extended to improve their ability to handle large media files.

Some conventional data models offer as special types long fields and BLOBs (binary large objects) to support multimedia data. However, these types reduce the view of multimedia data to single large un-interpreted data values, which is not adequate for the rich semantics of multimedia data. In particular, time dependent operations cannot be modelled adequately [28].

Relational DB systems based on SQL are a well understood technology. Therefore, they seem attractive to use for the storage and retrieval of multimedia data. But for the relational data model, the storage of multimedia documents requires complex mapping of hierarchical structures and sequential structures to relational tables.

A promising technology with respect to the modelling requirements of multimedia data is the object-oriented one [4]. The richness of the data model provided by OODBMSs makes them more suitable than RDBMSs for modelling both multimedia data types and their semantic relationships.

However, the performance of OODBMS in terms of storage techniques, query processing, and transaction management is not comparable to that of RDBMSs. Another drawback of OODBMSs is that they are highly non-standard.

For all the above reasons, a lot of effort has been devoted to the extension of the relational model with capabilities for modelling complex objects, typical of the object-oriented context. The goal of the so-called object-relational technology is to extend the relational model with the ability of representing complex data types by maintaining, at the same time, the performance and simplicity of relational DBMSs and related query languages.

Possible techniques for modelling multimedia data using traditional DB system technology are:

- A DB that contains only references to files which contains the original multimedia data. In addition to that, the DB may contain derived data modelled as additional attributes in the DB. The DB system can only provide services making use of metadata and not of the original multimedia data.

- A DB that stores multimedia data in attributes of types long field or BLOB. In this case, the DB system provides persistency for the data, buffering when accessing the data, and multi-user support, recovery and authorisation concepts on the granularity of long fields. The data stored is still un-interpreted and the functions available to operate on the data are generic.
Some DB systems that allow functions to be used to process data stored in the DB. This technique is useful to reuse external algorithms, programs and tools in the context of multimedia presentations.

Conventional DB approaches cannot resolve all issues for managing multimedia data. For instance, with approaches employing external references, consistency cannot be guaranteed by the system, because the DB system can’t exert any control over the referenced files. Also, approaches using un-interpreted storage for multimedia data, but not providing mechanisms for defining abstract data types, can’t provide a de-coupling of the multimedia data semantics from the applications. Thus each application has to re-implement semantics on its own.

Multimedia DB systems are still, to a large extent, at a research stage. Many are tied quite closely to a particular architecture or application, and have not yet reached the level of maturity or generality of more traditional DBMSs.

4.5 Future

Originally DBMSs were large, expensive software systems running on large computers. Today, a gigabyte fits on a single disk, and it is quite feasible to run a DBMS on a personal computer. Thus, DB systems based on the relational model have become available for even very small machines, and they are appearing as a common tool for computer applications. On the other hand, corporate DBs often occupy hundreds of gigabytes. And, as storage becomes cheaper, people find new reasons to store greater amounts of data, like images and video that take huge amounts of space. Handling such large DBs required several technological advances (like secondary storage devices and parallel computing).

Trends for (future) DB systems are (the introduction of) new technologies, like object-oriented programming, constraints and triggers, multimedia data, the WWW; and new applications such as warehousing of data or information integration. These trends change the nature of conventional DBMSs.
4.6 Exporting relational data into XML

4.6.1 Requirements

XML is rapidly emerging as a standard for exchanging data on the WWW. Despite the excitement surrounding XML, it is important to note that most operational (business) data, even for new web-based applications, continues to be stored in relational DB systems. This is unlikely to change in the foreseeable future because of the reliability, scalability, tools, and performance associated with relational DB systems. Consequently, if XML is to fulfil its potential, some mechanism is needed to publish relational data in the form of XML documents (also called XML views).

There are 2 main requirements [29] for publishing relational data as XML documents. The first is the need for a language to specify the conversion from relational data to XML documents. (This language specification describes how to structure and tag data from one or more tables as a hierarchical XML document.) The second is the need for an implementation to efficiently carry out the conversion.

Relational data is tabular, flat, normalised, and its schema is proprietary, which makes it unsuitable for direct exchange. In contrast, XML data is nested and ‘un-normalised’, and its DTD is public. Thus, the mapping from the relational data to the XML views is often complex, so a conversion tool must be general enough to express complex mappings. Existing systems are not really general [30], because they map each relational DB schema into a fixed, canonical DTD. The conversion tools must also be dynamic, i.e. only the fragment of the XML document needed by the application should be materialised (or the XML view must be ‘virtual’). To be efficient, the tools must exploit fully the underlying RDBMS query engine whenever data items in the XML view need to be materialised, (because query processors for native XML data are still immature and do not have the performance of highly optimised RDBMS engines).

4.6.2 Possible approaches

To export relational data into XML, there exist some possibilities [30]:

Currently, the interface to a relational DB is mostly a HTML form, which constructs a relational query from user inputs. Typically, such interfaces render the results of a query in HTML, but they could generate XML instead. Forms interfaces are appropriate for casual users, but
inappropriate for data exchange between applications, because they limit the application to only those queries that are predetermined by the form interface.

In another alternative, the data provider can either pre-compute the materialised view or compute it on demand whenever an application requests it. This is feasible when the XML view is small and the application needs to load the entire XML in memory, e.g. using the DOM interface. However, pre-computed views are not dynamic, i.e. their data can become stale, and are not acceptable when data freshness is critical.

Another approach is to use a native XML database engine, which can store XML data and process queries in some XML query language, but this means of course that the relational technology is not used anymore. It is unlikely that XML engines could replace completely high-performance relational databases, but a high-performance XML engine might be appropriate for data exchange. It is however expected that XML query languages won't attain performance comparable to that of SQL engines on relational data anytime soon. Moreover, this approach incurs a high space cost by materialising the data completely in XML.

In the next sections there are some example (projects) discussed that search for means to translate relational data into XML.

### 4.6.3 SilkRoute

In SilkRoute [30] data is exported into XML in 2 steps. First, a XML view of the relational database is defined using a declarative query language, called RXL (Relation to XML Transformation Language). The resulting XML view is virtual. Second, some other application formulates a query over the virtual view, extracting some piece of XML data. For this purpose XML-QL, an existing XML query language [31], is used. Only the result of that XML-QL query is materialised.

The core of SilkRoute is RXL, a declarative data transformation language that transforms flat relations to XML data. On the relational side, RXL has the full power of SQL queries and can express joins, selection conditions, aggregates, and nested queries. On the XML side, RXL has the full power of XML-QL, and can generate XML data with complex structure with arbitrary levels of nesting. Typical RXL queries are long and complex, because they express general transformations from the relational store to the XML view. (RXL has a block structure to help users organise, structure and maintain large queries.)
Once the virtual XML view is defined, SilkRoute accepts XML-QL queries and composes them automatically with the RXL query. The result of the composition is another RXL query, which extracts only that fragment of the relational data that the user requested. When a RXL query is evaluated, most of the processing is done by the underlying relational engine. To do this, the RXL query is split into a collection of SQL queries, each of which produces a set of tuples. The SQL queries are sent to RDBMS, and their flat, sorted results are merged in a single pass to construct the nested XML output.

The benefits of SilkRoute are:

- Only the relational data requested by a user query is materialised,
- The data is always produced on demand,
- And the relational engine performs most of the computation efficiently.

Problems on the other hand, are to find general techniques for translating of RXL into efficient SQL and minimisation of composed RXL views.

4.6.4 XML constructor

A key requirement for converting relational data to XML documents is a language to specify the conversion. While one approach is to invent a new language specifically for this purpose (like in SilkRoute), the approach followed in [29] is to harness and extend the power of SQL to specify the conversion of relational data to XML documents. The SQL language extension is called the XML constructor. Nested SQL statements are used to specify nesting, and SQL functions are used to specify XML element construction.

The extensions can be easily added to existing relational systems without departing from existing SQL semantics. As a result of extending SQL in this manner, standard APIs can be used to query and retrieve XML documents. This allows existing tools and applications to easily integrate relational data and XML documents.

4.6.5 Integrating keyword search

A varying or unknown structure of a XML document can make formulating queries a difficult task. In [32] an extension to XML query languages is proposed that enables keyword search at the granularity of XML elements, which helps novice users formulate queries, and also yields new optimisation opportunities for the query processor.
Relational databases can be used here in different ways, but 2 scenarios are considered. In the first scenario the whole XML data is replicated (or initially stored) in the relational DB. This scenario provides the best performance, and shows how a RDBMS can be used as a data warehouse for XML data. But it is not always possible or cost-effective to build a data warehouse. Therefore, the second approach carries query processing out in a distributed way. In this scenario, XML documents are stored by individual data sources. A RDBMS is used to store indexes which can be used to execute keyword searches and to find all relevant XML data sources for a query.

The classic structure for keyword search is the inverted file, that in its simplest form contains records of the following form:

\[
\text{< word , document >},
\]

meaning that ‘word’ can be found in the document. For XML data, this structure needs to be extended in the following ways:

- Information in the granularity of elements needs to be kept rather than documents.
- One has to record whether a keyword is the name of a tag, the name of an attribute, a word in the value of an attribute, or a word in the element’s data content.
- The depth at which the word appears for the first time within an element, needs to be recorded in order to execute ‘contains’ predicates that limit the depth.

As a result, the proposed structure for the records of inverted files is:

\[
\text{< elID , word , depth , location >}
\]

where the ‘word’, ‘depth’ and ‘location’ have the natural meaning.

The drawback of the approach is that the extended inverted files can become very large: it is not unusual that they are a factor 10 or even more larger than the original data.

4.6.6 Oracle (8i)

Developers are looking for a technology which makes it dramatically easier to acquire, integrate, and exchange information to make their growing databases work even harder on their behalf in the Internet environment [33]. XML is gaining wide industry support as well from vendors like Oracle, IBM, Sun, Microsoft, Netscape, SAP and others, as a platform- and application-neutral format for exchanging information.
4. Relational Databases

Figure 7: XML, Portal to a World of Data [33]

4.6.6.1 Generate XML

Oracle 8i has a XML SQL Utility [34] for Java consisting of a set of Java classes, that pass a query to the database and generate a XML document from the result, and write XML data to a database table.

Figure 8: Generating XML from query results with Oracle's XML SQL Utility [34]

As shown in Figure 8, the XML SQL Utility defines a fixed, canonical mapping of the relational data into a XML document, by mapping each relation and attribute name to a XML tag and tuples as nested elements. The structure of the resulting XML document is based on the internal structure of the database schema that returns the query results. The Utility can generate either a string representation of the XML document, or an in-memory XML DOM tree of elements.
In [30] the comment is made that such a view could be kept virtual, but the approach is not general enough to support mapping into an arbitrary XML format.

4.6.6.2 Exchange XML

Exchanging XML among applications that do not share a common DTD requires intermediate processing and transformation of the XML. The main limitation of DTDs is that they do not provide type information for the data, because everything in a DTD is of type string [34]. So when building a DTD from a database, type information gets lost. Applications using DTDs must assign types to data elements based on context of data or its associated tags.

4.6.6.3 Store XML

Depending on the structure of the XML document, one of the following 3 approaches can be chosen to store XML in Oracle 8i [34]:

1. Storing a XML document as a single, intact object with its tags in a CLOB or BLOB
   Storing an intact XML document in a CLOB or BLOB is a good strategy if the XML document contains static content that will be updated by replacing the entire document. (For example: articles, advertisements, legal contracts...) Documents of this nature are known as document-centric and are delivered from the database as a whole. Document-centric documents are characterised by irregular structure, larger grained data, and lots of mixed content.

2. Storing the XML document as data and distributing it untagged across object-relational tables.
   If the XML document has a well-defined structure and contains data that is updateable or used in other ways, the document is data-centric. Typically, the XML document contains elements or attributes that have complex structures. (For example: sales orders, airline flight schedules...)
   There are 2 options for storing and preserving the structure of the XML data in an object-relational form:
   - Store the attributes of the elements in a relational table and define object views to capture the structure of the XML elements.
   - Store the structured XML elements in an object table.
5 Multimedia Information Retrieval

5.1 Classification

Current practice in multimedia information retrieval can be divided in two categories:

- The first category is text-oriented. The simplest method to query multimedia data is to define keywords (metadata) associated with the data, (also known as manual indexing). With the spreading of markup languages, people began to realise that annotations provide an easy way to add metadata to a multimedia object. The annotated metadata delivers text about the multimedia object, and queries to retrieve multimedia data are performed on the associated text. With the existing algorithms the texts can be queried without much problems.

- The second category are visual retrieval systems, based on the analysis of content. Content-based querying (CBQ) [35] as it is usually called, is the searching in multimedia DBs based on the intrinsic properties of the images, rather than on coded alphanumeric descriptors. The goal is to index the multimedia data using features (metadata) derived directly from the data. Several features such as colour, texture and shape are used for images. The features are extracted by automatic analysis (e.g. image processing and image understanding for images).

Both categories don't exclude one another. Because of the problems involved with textual metadata (like the introduction of data being time-consuming,...), it is interesting to use the techniques of the content analysis systems to create automatically certain metadata (concerning content).

The MPEG-7 initiative for example works on Description Schemes to describe multimedia objects, but at the same time they promote initiatives for automatic extraction of features that could populate the Description Scheme with data.

The visual retrieval systems can be further divided into two categories [36]:

- Query-by-example systems work under the realisation that since the 'correct' match must lie within the DB, one can begin the search with a member of the DB itself, with the hope that one can guide the user towards the image that he likes over a succession of query examples.
Examples of systems that use query-by-example are QBIC [37], PhotoBook [38], VisualSEEk [39], Virage [40].

In query-by-example, one can use space-partitioning schemes to pre-compute hierarchical groupings, which can speed up the DB search. While the search speeds up, the groupings are static and need re-computation every time a new document is inserted in the DB.

- **Visual sketches**, like in QBIC, VisualSEEk... Certain sketch-based query systems [41] compute the correlation between the sketch and the edge map of each of the images in the DB, while others [42] minimise an energy function to achieve a match. In another system [43], a distance is computed between the wavelet signatures of the sketch and each of the images in the DB.

These 2 different ways of visually searching images in DBs may also be accompanied by learning and user feedback.

### 5.2 Problems in Multimedia Information Retrieval

Problems associated with the text-oriented approach, is that although the method is simple, it is insufficient since it is manual and time consuming, and the resulting indexes are highly subjective and tied to a limited vocabulary. But keywording enables very fast retrieval of data, and standard indexing approaches can be used since the keywords (strings) are a data type supported by every DBMS. For specialised applications this might be enough. However, for the classification of video, this may generally be overwhelming.

At the Bristol meeting of MPEG [44] a number of presentations were given on work related to MPEG-7. Conclusions derived from the presentations were that a real market need exists for MPEG-7 (i.e. the provision of a standardised content description methodology), and that there appears to be little knowledge available about how to combine high-level descriptive frameworks with low level feature extraction.

One of the major issues in searching for an image in a DB is how 2 images can be judged by the system as being similar. The concept of image similarity has to be defined taking into account the kind of images and possibly the application. Two kinds of similarity are usually distinguished [35]:

- **Syntactic similarity**, which is based on the basic perception of raw physical characteristics (visual or syntactic features).
- **Semantic similarity**, which involves some degree of understanding of the higher level structures that such characteristics can represent.
Ideally, the semantic content of an image, rather than its syntactic features, should be the intrinsic property of the image to be used in queries. But automatic inference of the semantic content is hard to achieve, and needs a specific, application dependent training of the extraction process.

Since it is difficult to identify the set of features that are most appropriate in any given kind of images (medical, satellite, aerial, natural scenes, etc.), several research groups focus on specific types of images. Therefore, we find in the proposed content-based information retrieval systems, domain-dependent systems and domain-independent ones. While the first systems lack flexibility and openness, they are more effective and more efficient than the latter ones, since they take advantage of a context (meta-knowledge) which is very helpful for the identification of similarity criteria and therefore the best visual features for the indexing process. The reason is that in constrained domains, it is easier to model the users’ needs and to restrict the automated analysis of the images, such as a finite set of objects.

Because multimedia involves different media and formats, the interface used to enter the query into the system becomes an issue. Since different multimedia data types may require different query interfaces, the problems to consider include how to integrate the various interfaces in an integrated multimedia DB system.

Video is simply a sequence of images, so image techniques may also apply. However, video also has the temporal dimension, therefore we may want to find specific scenes, scene changes, commercials, people eating, running or riding bicycles for example. Some systems are capable of retrieving simple movements, for example videos of skiing can be retrieved, because the movement is easy to draw (a line going down), and the system compares the sketch with movements in the stored videos.

Video shots generally contain a large number of objects, each of which is described by a complex multi-dimensional feature vector. The complexity arises partly due to the problem of describing temporal and spatial information.

One has to keep in mind that the visual paradigm works best when there are only a few dominant objects in the video with simply segmented backgrounds. It will not work well if the user is interested in video sequences that are simple to describe, but are hard to sketch out, like for example a video shot of a group of soldiers marching. It will also not work well when the user is interested in a particular semantic class of shots: he might be interested in retrieving that news segment containing the anchor person, when the news anchor is talking about Bosnia.
5.3 Research

In content-based retrieval, the *semantic* information is the purpose and the core of the query. The user asks the system for the images containing the objects of his interest. Image’s content is not only identified by data, but also by context (domain-specific knowledge) and the observer. Capturing semantic information into a general image retrieval engine requires solving problems such as:

- Feature definition and extraction,
- Segmentation, and
- Object and context identification.

Among the multimedia data, retrieving images by content is still a quite challenging problem, which is currently a hot topic in information retrieval and DB communities. Recent researches have mostly focused on [35]:

- The improvement of the analysis and the representation of image characteristics such as shape and texture;
- Usage of automatic segmentation in order to derive syntactic characteristics (colour, texture, shape), perceived as significant in images or image regions;
- The techniques and data structures for multi-dimensional indexing;
- The development of interactive tools allowing direct expression of content-based queries;
- The identification of performance evaluation criteria;
- Benchmarking.

5.4 Examples of multimedia information retrieval systems

5.4.1 Search-engines on the Web

An example of a search-engine on the Web that does queries on multimedia objects (images, video and audio separately) is Altavista [46]. When for example a query is performed on ‘CD’ by somebody who wants to search for an image of a CD, the person will retrieve a lot of images of covers of CDs with popstars. Only downwards in the ranking, an image of a CD appears.

Search-engines on the Web perform text-oriented queries, where mostly the query is based on the name of the image, which provides a very limited resource for queries. A more elaborate technique, especially for use on the Web, is to use the text that makes part of an associated link.[47]
5.4.2 Describing and retrieving photos using RDF and HTTP.

The project "Describing and retrieving photos using RDF and HTTP" [48] is a concrete example of a RDF schema and a working system around it, which might help explaining the potential of metadata on the Web.

Using metadata will automatically provide a non-visual description of the photos, hence contributing to accessibility. The metadata is separated into 3 different schemas:

- Dublin Core schema (see 2.5)
- Technical schema: captures technical data about the photo and the camera.
- Content schema: is used to categorise the subject of the photo by means of a controlled vocabulary.

The data-entry program is written in Java, but the user-interface is in fact generated at run-time directly from a machine-readable version of the schemas. So the program does not need to be changed when the RDF schema changes. In this project the interface for the multimedia object is not a big problem, since all the images are captured in JPEG format.

The metadata is entered manually, but the systems assumes that there are only slight differences between the different images (series of pictures are entered), so the fields for the metadata show the previous values and only the metadata that changes has to be modified. The idea exists to write a script that could convert the technical data included in a digital camera to the technical schema for the pictures.

5.4.3 Columbia’s content-based visual query project

Columbia’s content-based visual query project [45] uses the following research strategies:

- Create a visual feature library by automatic image analysis
  Although today’s computer vision systems cannot recognise high-level objects in unconstrained images, it appears that low-level visual features can be used to partially characterise image content. These features also provide a potential basis for abstraction of the image semantic content.

- Explore the synergy between compression and functionality
  Today’s compression standards (such as JPEG, MPEG-1, MPEG-2) are not suited to represent images and video in a way that is amenable to dynamic feature extraction. However, recent trends in compression, such as MPEG-4 and object-based video, have shown interest and promise in this direction. The goal is to develop a system in which the video objects are extracted, then encoded, transmitted, manipulated, and indexed flexibly with efficient adaptation to users’ preference and system conditions.
• Learn from users and domain ontologies\(^{10}\)
  
  What is needed, is to solve the problem of decoding semantic content in images, and user-interaction and domain-knowledge have to be further explored. These systems learn from the users’ input as to how the low-level visual features are to be used in the matching of images at the semantic level.

• Integrate visual and other multimedia features
  
The association of visual features with other multimedia features, such as text, speech, and audio, will enhance image retrieval performance. Video often has text transcripts and audio that may also be analysed, indexed, and searched. Also, images on the WWW typically have text associated with them.

Some prototype systems and on-line demos are available on [49]:

• **In Lumine**: automatic image classification using integrated text-visual features.

• **Visual Apprentice**: a user-machine co-operative system for learning visual object filters.

• **FaceTrack**: face detection, tracking and summarisation from compressed video.

• **WebClip**: network-based compressed video editor.

• **MetaSEEK**: an image meta search engine using learning from user preference.

• **WebSEEK**: a large Web image / video cataloguing system.

WebSEEK [49] is a content-based image and video catalogue and search tool for the WWW. WebSEEK collects images and videos using several autonomous Web agents which automatically analyse, index, and assign the images and videos to subject classes. The system is novel in that it utilises text and visual information synergistically to provide for cataloguing and searching for the images and videos. The complete system possesses several powerful functionalities, namely searching using image content-based techniques, query modification using content-based relevance feedback, automated collection of visual information, compact presentation of images and videos for displaying query results, image and video subject search and navigation, text-based searching, and search results list manipulations such as intersection, subtraction and concatenation.

• **VisualSEEK / SaFe**: joint region and spatial feature image searching.

SaFe is a general system for spatial and feature image search [49]. It provides a framework for searching and comparing images by the spatial arrangement of objects. In a SaFe query, objects or regions are assigned by the user. These are given properties of spatial location,
size and visual features, such as colour. The SaFe system finds the images that best match the query, based on syntactic features. SaFe uses fully automatic tools for region / feature extraction and indexing. SaFe also resolves spatial relationships, which allows the user to position objects relative to each other in a query.

- **VideoQ**: an automatic content-based search engine for video objects.

Features of VideoQ are [36]:

- Automatic video object segmentation and tracking;
- A rich visual feature library including colour, texture, shape, motion;
- Query with multiple objects;
- Spatio-temporal constraints on the query;
- Interactive querying and browsing over the WWW;
- Compressed-domain video manipulation.

VideoQ uses extensive content analysis to obtain accurate video object information [36]. Global motion of the background scene is estimated to classify the video shots as well as to obtain the local object motion. The query server contains several feature DBs, one for each of the individual features that the system indexes on. The video shot DB is stored as compressed MPEG streams. Matched video clips are dynamically 'cut' out from the MPEG stream containing the clip without full decoding of the whole stream.

The fundamental paradigm under which VideoQ operates is the visual one. This implies that the query is formulated exclusively in terms of elements having visual attributes alone, so there is no text-oriented query. The features that are stored in the DB are generated from an automatic analysis of the video stream. There is no information present in the query loop that emanates from the captions, textual annotations or the audio stream. Many retrieval systems, such as Photobook [38], VisualSEEk [39] and Virage [40], share this paradigm, but only support still image retrieval.

### 5.4.4 Query By Image Content (QBIC)

The QBIC system of IBM supports content-based retrieval of multimedia data. Instead of exact matches, searches are approximate and driven by similarity. The procedure (query-by-example) uses a combination of browsing and searching. Basically, the DB system does not retrieve images based on explicitly defined characteristics, but assists the user in narrowing the search space. Developing a QBIC application consists of 3 logical steps: DB population, feature calculation and image query [28].
The QBIC system provides a method for computing (syntactic) features. This way, text based searching can be extended by allowing retrieval on image parameters. This retrieval method can be extended to different object types. It is not a method of extracting conceptual or symbolic information, or a replacement for text based approaches. QBIC is not restricted to images. While QBIC [37] is visual, it is not exclusively so, as the images have been manually annotated allowing for keyword searches on the DB.

5.4.5 Content-based search in Oracle 8i

With Visual Information Retrieval of Oracle8i [50], the two general approaches (namely, using information manually entered or included in the table design, and predetermined classification schemes; and, using automated image feature extraction and object recognition to classify image content) are combined in designing a table to accommodate images. Traditional text columns are used to describe the semantic significance of the image (for example, that the automobile in the picture won a specific award), and to use the Visual Information Retrieval type for the image to permit content-based queries based on intrinsic attributes of the image (for example, how closely its colour and shape match a picture of a specific automobile).

A content-based retrieval system processes the information contained in image data and creates an abstraction of its content in terms of visual attributes. Any query operations deal solely with this abstraction rather than with the image itself. Thus, every image inserted into the DB is analysed, and a compact representation of its content is stored in a feature vector, or signature. The signature contains information about the following visual attributes (or syntactic features):

- Global colour: represents the distribution of colours within the entire image.
- Local colour: represents colour distributions and where they occur in an image.
- Texture: represents the low-level patterns and textures within the image, such as graininess or smoothness.
- Structure: represents the shapes that appear in the image, as determined by shape-characterisation techniques, such as edge detection.
- Facial: represents unique characteristics of human faces.

5.5 Conclusions

Although there exists a tendency for researchers to concentrate on content analysis of multimedia objects, the text-oriented approaches for multimedia querying will not lose their relevance. Currently some syntactic features can be extracted automatically, which is not the case for
semantic features. However, there is often a strong correlation between syntactic and semantic aspects, which makes it possible to handle the syntactic features rather than the semantic ones (which are more complex). With the discovery of such correlations, it might be possible to extract semantic features in the future.

Many content-based image search systems have been developed for various applications. There has been substantial progress in developing powerful tools which allow users to specify image queries by giving examples, drawing sketches, selecting visual features (e.g. colour and texture), and arranging spatial structure of feature. Using these approaches, the greatest success is achieved in specific domains such as remote sensing and medical applications [45].

On the other hand, for the temporal aspects there are still many questions to solve. The notion of analysing video and detecting actions is far from reality at this point in time. However, as the importance of multimedia DBs grows, so will the demand for the ability to intelligently query them.

As long as the possibilities of content analysis are not completely defined, the text-oriented approaches are a solution for querying multimedia data. As automatic analysis techniques evolve, they can be used to introduce automatically metadata, this way saving time and reducing errors.

With the new ‘synthetic films’, segmentation will be done automatically, since the film is composed of different objects put together by computer. Systems for analysing pictures will also take advantage of this segmentation, which means that generating metadata for these kinds of films is easier to achieve.
6. The Metamedia Description Scheme

6.1 Introduction

The MetaMedia Description Scheme is developed in the context of the MetaMedia project. The main goal is to build a multimedia database that fulfils the requirements of archival storage and search. Therefore, descriptors and a description scheme are defined in the context of the MPEG-7 initiative (see 2.7), thus the terminology used is coming from MPEG-7. The approach taken is the text-oriented one, but MetaMedia leaves the possibilities open for feature extraction and considers also possible ways for data interchange.

The goal of the Description Scheme is to have a metadata model for multimedia objects that comprises a structural description and a set of descriptors for textual documents, images and video. Therefore, a Description Scheme aims at transposing to the metadata the essential features of a category of objects, in order to incorporate into the manipulation and archive tools the means to produce descriptions according to a standard, and to enable search utilities to retrieve objects based on their characteristics and content features. The multimedia objects that can be described with the Description Scheme may be individual items as well as collections of items.

Two aspects of an object are distinguished here, the context and the content, and both will have associated metadata. Metadata concerning the context includes features that have to do with the object’s creator and owner, the support used for its storage, the relationships with other objects, the conditions under which it has been used. Metadata concerning the content includes summaries and keywords, events or entities the object is about, and in general any features of the object’s content.

Content information indexing and retrieval is expected to become current technology, see also chapter 5. However, the descriptors associated with structural and context information will not lose their relevance. Multimedia archives will become larger and archival techniques will have an impact on their physical and logical organisation. It is the intention of this Description
Scheme to balance context and content in the descriptions, to obtain an integration between traditional queries (by name, title...) and queries related with content analysis. Whereas the 'content queries' talked about in Internet-related projects in the previous chapter were more restricted to queries on colour, motion, etc. (syntactic features), the content descriptors included in the Description Scheme have also values that can't be retrieved (so far and in the near future) by automatic extraction. This is the case with events, for example which need interpretation of the data, rather than analysis of elementary features. But with the huge amount of information on the Internet it is hard to add semantic data that can not be automatically extracted, compared to a well organised archive.

The Description Scheme proposes the classification of materials on levels that correspond to the granularity of description, and gives the ability to query documents.

With MPEG-7 in mind, the description scheme, the content descriptors and the content hierarchies were developed. The ISAD standard (see 2.4) provided the inspiration for the contextual descriptors and the hierarchical description. ISAD was interpreted by considering creator, owner, location, etc. as entities in the model and not just as descriptors.

The Description Scheme is defined in UML (see Appendix A), because this language admits the natural expression of hierarchies and inheritance. Moreover, UML is a good basis for translating into relational schemes and XML.

The proposed Description Scheme includes sub-models, namely Creator, Owner, Storage and About (see Figure 9) [51,52].

![Diagram]

Figure 9: Package Diagram of the MetaMedia DS.

---

11 Sub-chapter 6.1 (Introduction) till 6.7 (Multiple Language Support) contain several references to P420 [51] and P422 [52], but for reasons of readability these references will not be repeated in the text.
6.2 System of Description - Package

System of Description is the central package that organises the description of multimedia material. It contains the entities that structure the whole description. It is also the package that establishes the main relationships between the materials. A hierarchical structure of containment is adopted, whereby each object is described making use of the description of its container object, if any. The levels in the hierarchy are left open, to be established by the archives manager according to the nature of the materials.

The main entity in this package is the Description Unit. The Description Unit is the place to identify and place general descriptors for the unit being described. A unit is either a single atomic item or a collection of items with common descriptive features. The description for a collection captures features in common for a set of objects which typically have been put together at the time of creation. For a Description Unit there is always an associated Description Level, and possibly a set of related units. Scheme and Description Level are entities with a different status: they are meant to control the hierarchical structure, in a way acceptable for the archives.

Description Units and Segments are instantiated by the archivist who describes materials, while the archives manager is responsible for defining Schemes and Description Levels, tailoring the structure of the archive its contents.

6.2.1 Description Unit

The Description Unit contains a structural description of a multimedia document. It supports context information, which makes it suitable for archive organisations (large quantities of documents). Moreover, a top-down philosophy is used, which builds on the archive’s experience. The Description Unit entity groups together the core of descriptors for each object.

6.2.2 Segment

The Segment contains content descriptions of a multimedia document. So here descriptions associated with the segmentation of the (AV) content are handled. Moreover, in Segment, MPEG-4 and advanced index mechanisms (feature extraction tools) can be integrated.

Each Description Unit can have various Segments, and each Segment belongs to at least one Description Unit. In case of time-dependent media like video, the association TimePeriod clarifies which time-period of the Description Unit the Segment is describing. The Segment is in
6.2 System of Description - Package

System of Description is the central package that organises the description of multimedia material. It contains the entities that structure the whole description. It is also the package that establishes the main relationships between the materials. A hierarchical structure of containment is adopted, whereby each object is described making use of the description of its container object, if any. The levels in the hierarchy are left open, to be established by the archives manager according to the nature of the materials.

The main entity in this package is the Description Unit. The Description Unit is the place to identify and place general descriptors for the unit being described. A unit is either a single atomic item or a collection of items with common descriptive features. The description for a collection captures features in common for a set of objects which typically have been put together at the time of creation. For a Description Unit there is always an associated Description Level, and possibly a set of related units. Scheme and Description Level are entities with a different status: they are meant to control the hierarchical structure, in a way acceptable for the archives.

Description Units and Segments are instantiated by the archivist who describes materials, while the archives manager is responsible for defining Schemes and Description Levels, tailoring the structure of the archive its contents.

6.2.1 Description Unit

The Description Unit contains a structural description of a multimedia document. It supports context information, which makes it suitable for archive organisations (large quantities of documents). Moreover, a top-down philosophy is used, which builds on the archive's experience. The Description Unit entity groups together the core of descriptors for each object.

6.2.2 Segment

The Segment contains content descriptions of a multimedia document. So here descriptions associated with the segmentation of the (AV) content are handled. Moreover, in Segment, MPEG-4 and advanced index mechanisms (feature extraction tools) can be integrated.

Each Description Unit can have various Segments, and each Segment belongs to at least one Description Unit. In case of time-dependent media like video, the association TimePeriod clarifies which time-period of the Description Unit the Segment is describing. The Segment is in
Figure 10: UML presentation of the SystemOfDescription Package.
a way a lower level description unit, covering content: it has no context information besides that induced by the Description Unit it is attached to.

For a Segment, there are new types of information that can be attached. The Segment in the description scheme is adapted to be able to support different kinds of media, so information for each of them can be gathered in different objects like Text, Image and Video.

The description of Video can be more detailed through the objects VideoObjectItem and AudioObjectItem. Descriptors may now have a time-dependent nature. Those descriptors can be found in TimeDependentDescriptor and the associated values are in Sample.

6.2.3 Hierarchy

The main ideas behind the proposed Description Schemes are:

- The archive is organised as a containment hierarchy with a top level and a varying number of sub-levels, depending on the actual material and how far the description work has gone;
- The set of attributes to describe each level is the same, with the advantage of making it easy to add or remove a level in a hierarchy;
- The actual description must follow the rule of generality, asserting each element as high in the hierarchy as appropriate and avoiding redundancies at lower levels where only more specific information, relevant for the respective level, should be explicitly added; generic information can be obtained by searching higher in the hierarchy;
- There are separate representations for the structural and external characteristics and for the detailed description of the document's content;
- The creation context is important and induces the overall organisation of the archive.

To be able to deal with a wide range of situations where more or different levels are appropriate, a fixed hierarchy was avoided by adopting a recursively aggregated class. Each concrete scheme one may add to the class Scheme, has a definite set of levels connected by the meta-part-of association. It acts as a pattern for the allowable levels in the instances of the Description Unit class and its part-of association.

By use of the hierarchy, one can choose how far the level of description will go. If the amount of available documents is too big to handle, then documents can be gathered and the common descriptors can be defined. If later more distinct descriptors are needed for the separate documents, they can always be added at a lower level. The descriptors defined in the higher level don't have to be repeated, because they are inherited. For example, a television broadcaster...
might not have the means to describe all the video archive. Therefore, the documentaries might be gathered for instance as ‘series’, and descriptors like owner, copyright, storage could be specified for the collection. If later more detailed descriptions are required for some documentaries, they can be added at a level lower than the series of which they make part, while they inherit the descriptors that were defined in the concerned series.

The descriptions of a (set of) video stream(s), for instance, can be organised according to a well-defined temporal hierarchy. This hierarchy reflects several possible divisions of the video content being described in time intervals. The higher in the hierarchy, the larger the time interval considered. Some of the possible hierarchies (examples in Table 7, Figure 11 and Figure 12), are well known and already in use in different areas. This fact could be used to make a small set of schemes based on these well-known video organisations.

Table 7: Examples of hierarchical video description schemes and levels.

<table>
<thead>
<tr>
<th>Level Number</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scheme number</td>
<td>News</td>
<td>News Series</td>
<td>Program</td>
<td>Story/Item</td>
<td>Story Detail</td>
</tr>
<tr>
<td>2</td>
<td>Film</td>
<td>Film</td>
<td>Sequence</td>
<td>Shot</td>
<td></td>
</tr>
</tbody>
</table>

![Diagram](image_url)

Figure 11: Example of the Hierarchy in SystemOfDescription Package for a film.
6.3 Storage – Package

The Storage Package contains descriptors that point to the actual storage of the documents. The Support class is meant to clarify the physical aspects of the storage of the multimedia document. This class originates from P420, but the Section association (with attributes Starttime and Stoptime) is left out, to unify the Video Description Scheme with Description Schemes for other media. Through the Support class information about the format of the AV support described can be captured.
The association Storage (between the Description Unit and Support) adds information about the physical location of the Description Unit. It is a many-to-many relationship because the Description Unit may be stored in several folders and a single CD-ROM (Support) can contain many photos (Description Units).

![UML diagram of Storage package]

Figure 13: UML presentation of the Storage Package.

6.4 About – Package

The About Package (former name: Event Package) includes descriptions of the entities and events related with the Description Unit. Therefore, this Package adds semantics to the analysis of the document with an obvious interest for conducting searches.

The Event class is meant to store information about real world events such as a World Cup or a wedding. Many documents may be related with a single event, and a single document may be related with several events.

With the class Entity Type a classification of the events can be made (while the Event class gives more detailed information). This classification is structured as a hierarchy (where the highest level is the most generic level).
If for example the Beatles or the Queen are in an image then this information may be asserted through the association to the corresponding World Entity. (Many documents may be related with a 'World Entity', and one document may refer several World Entities.) The class World Entity is intended for entities that are related with the content of the Description Units. Entities related with ownership and creation of the Description Unit, are handled in the Owner Package and Creator Package respectively.

![Diagram](image)

*Figure 14: UML presentation of the About Package.*

### 6.5 Creator – Package

(In P420 this Package is called Productor Package)

Each Description Unit may have several Creators.

In the Creation class the date and place of creation of the Description Unit are recorded, together with the Role of the Creator with respect to the Description Unit (in P422 this Role is implemented in a separate class, while P420 doesn’t contain a Role for the Productor). Examples of Roles are the author of a photograph or the author of the negative.

Each Creator (who can have a different identification along time) has a history and may belong to another creator (e.g. the news department in a broadcaster), which is clarified by the hierarchy of CreatorDescription. Different identifications may refer to the same creator entity.

Therefore the association Main Identification is used to make clear if the considered Creator (and his identification) is the main identification of a given entity, described in the corresponding CreatorDescription.
Each creator has a geographical location associated through the association with the class *Location* (former name *Local*).

![Diagram](image)

*Figure 15: UML presentation of the Creator Package.*

### 6.6 Owner – Package

Each *Description Unit* describes material that may be owned by some (one or more) entity. The class *Owner* represents those entities.
For each set of documents, we can describe information on the custodial history (history of the custody of the document by several entities), where the original is, if we hold a copy, and where the copies are kept.

For the sake of avoiding redundant information, the class Original in P420 and P422 (which indicated what is the original of the Description Unit that can be owned by someone) is replaced by the attribute Original in the Copy class (the attribute indicates if the concerned 'copy' is the original). Furthermore, the Copy class has the location of the copies that a certain owner has.

![UML Diagram]

Figure 16: UML presentation of the Owner Package.
Since the content can have several owners in different time instants, it is important to register the custodial history of the content being described through the *Custody* class. Each *Description Unit* may require several instances of *Custody*, each describing it as kept by a certain *Owner* over a specified period of time. On the other hand, an *Owner* may have several instances of *Custody* for several *Description Units*.

It is important to register the materials that are intimately related with the material being described, but belong to another owner: for instance, a series that is part in one archive and part in another. That can be interesting from a search perspective and can be done through the *Associated Material* class.

Another association is the *Related Description Unit* where it is possible to register any cross-relations between *Description Units* in the same archives.

Each owner entity is associated with a geographical location through the association with class *Location* (former name *Local*).

### 6.7 Multiple Language Support

![UML Diagram](image)

*Figure 17: UML presentation of the multiple language support in the MetaMedia DS.*

Support of multiple language multimedia description is a requirement to make the world-wide flow of information real and effective.
As there are several descriptors that are language-dependent, it is necessary that each one can be translated to any desired language. This is achieved by migrating each text attribute in an affected class to an auxiliary class connected through a *Language* association. For each main class, the auxiliary class has as many instances as required languages.

The multiple language support can be implemented in all the packages, but to keep the Description Scheme clear, the multiple language support was left out in the diagrams of the different packages. To demonstrate the implementation of the multiple language support, the *Description Unit* will serve here as a model.

The *Reference Code* in the *Description Unit*, for example, is language independent, thus remains an attribute of the *Description Unit* class. Keywords on the other hand are language dependent, so a class *Keywords* associated with the *Description Unit* and with a *Language* will be created (see example in Figure 17, where Keywords and Summary are language dependent descriptors). Thus, each instance of the *Keywords* class contains keywords in a certain language.
6.8 Relation with MPEG-7 Multimedia Description Schemes\textsuperscript{12}

"Multimedia Description Schemes provide a way to describe in XML the important concepts related to audio-visual data in order to facilitate the searching, indexing and filtering of audio-visual data."

The MPEG-7 Multimedia Description Schemes (also called MMDS) are organised into the following areas according to their functionality:

- Basic elements
- Schema tools
- Content description
- Content management
- Content organisation
- Navigation and access
- User interaction

Figure 18 gives an overview of the structure.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{mdbg.png}
\caption{Overview of the Multimedia Description Schemes [11d].}
\end{figure}

\textsuperscript{12} The subchapter 6.8 contains a lot of references of [11d], which are not always repeated for reasons of readability, but the cited text can be recognised by the "\ldots".
From the overview in Figure 18, we can see that the MMDS are presented on the level of the archives, a level higher than the one of the MetaMedia DS, which is presented on the level of the items, meaning that everything is built around the object.

Making a rough comparison of the MMDS overview, with the Package diagram of the MetaMedia DS, we find that the About Package can be categorised as Content Description, and the Creator, Owner and Storage Package get the label Content Management. The SystemOfDescription Package balances between Content Management, Content Description and Content Organisation.

"Because the descriptive features must be meaningful in the context of the application, they will be different for different user domains and different applications. This implies that the same material can be described using different types of features, tuned to the area of application. (The level of abstraction is related to the way the features can be extracted: many low-level features can be extracted in fully automatic ways, whereas high level features need (much) more human interaction.)"

Since the MetaMedia DS was designed especially for multimedia archives, issues related with the context (and not only content) were thought of as being important. But the MMDS don’t make this clear distinction between context and content.

In the MetaMedia DS the Description Unit, Creator, Owner and Storage are considered as context information, while Segment and About treat the content information. If we make this distinction in the MMDS, then Content Management treats context and Content Description treats content, while Content Organisation, User Interaction and Navigation & Access concern both context and content.

"The idea for the MMDS is that, next to having a description of the content, it is also required to include other types of information about the multimedia data, like:

- **The form** - An example of the form is the coding scheme used (e.g. JPEG, MPEG-2), or the overall data size. This information helps determining whether the material can be ‘read’ by the user.

In the MetaMedia DS the form can be handled by ‘format’ and ‘size’ in the Description Unit, and extra information concerning analog or digital support can be recorded in the AnalogSupport and DigitalSupport classes.

- "Conditions for accessing the material" - This includes links to a registry with intellectual property rights information, and price."
This kind of information is not worked out very well in the MetaMedia DS, since the DS is aimed for archives, and not specific for content creation (especially in a business context), where those concepts are more important.

- "**Classification** - This includes parental rating, and content classification into a number of pre-defined categories."

Archives can have a very hierarchical structure, so the MetaMedia DS includes the means to construct this hierarchy in the DS. But the attention goes here to the structure itself and not to a classification by content.

- "**Links to other relevant material** - The information may help the user speeding up the search."

In the MetaMedia DS, there are links to other associated material. But here the distinction is made between related material in the same archive, and related material owned by an other entity.

- "**The context** - In the case of recorded non-fiction content, it is very important to know the occasion of the recording (e.g. Olympic Games 1996, final of 200 meter hurdles, men)"

The About Package of the MetaMedia DS records events associated with a Description Unit, so the context can be searched in the class Event.

"In many cases, it will be desirable to use textual information for the descriptions. Care will be taken, however, that the usefulness of the descriptions is as independent from the language area as possible."

In the MetaMedia DS multiple language support is provided (for selected descriptors) to make a world-wide flow of information possible.

"A very clear example where text comes in handy is in giving names of authors, film, places. Therefore, MPEG-7 Description tools will allow to create descriptions of content that may include:

- Information describing the creation and production processes of the content (director, title, short feature movie)"

  (MetaMedia DS: Description Unit, Creator Package)

- "Information related to the usage of the content (copyright pointers, usage history, broadcast schedule)"

  (Information given less relevance in MetaMedia DS; attribute copyright in Description Unit)
• "Information of the storage features of the content (storage format, encoding)"
  (Storage Package in MetaMedia DS)
• "Structural information on spatial, temporal or spatio-temporal components of the content
  (scene cuts, segmentation in regions, region motion tracking)"
  (Segment of SystemOfDescription Package in MetaMedia DS)
• "Information about low level features in the content (colours, textures, sound timbres,
  melody description)"
  (in Segment-part of the MetaMedia DS)
• "Conceptual information of the reality captured by the content (objects and events,
  interactions among objects)"
  (MetaMedia DS: About Package)

For the organisation of collections, the MMDS can count on the Collection DS and the Model
DS (both in Content Organisation). "The Collection DS organises collections of audio-visual
content, segments, events, and/or objects. This allows each collection to be described as a whole
based on the common properties. The Model DS provide tools for modelling the attributes and
features of audio-visual content."

The hierarchy of the Description Unit permits in the MetaMedia DS organisation of collections
by context. The MetaMedia DS does not include organisation of collections by content,
although a hierarchy on segment-level and on events is included, but this is rather seen as a
contextual organisation.

The classes DescriptionLevel and Scheme (in SystemOfDescription) also organise the archive
(context), but this 'organisation' is to be fixed by the archive manager, and can only be used,
not changed by the cataloguers. (So with the configuration of the archive a certain amount of
Schemes and Levels will be defined.) Archivists can refine the level of description by using the
hierarchy of the Description Unit (refinement of context description), or by using the hierarchy
of the Segment (refinement of the content description). With these hierarchies, one can record
generic information, or record information as detailed as desired, or feasible. In the MMDS, the
level of description for content can be defined with Structural Aspects, "which describes the
audio-visual content from the viewpoint of real-world semantics and conceptual notions. (The
Structure DSs are organised around a Segment DS that represents the spatial, temporal or
spatio-temporal structure of the audio-visual content.)"
"Multimedia Description Schemes provide a standard set of data types and description structures needed for generating XML descriptions of audio-visual content. The descriptions are human-readable and can be searched, transmitted and filtered in applications that deal with audio-visual data.

Many Description Schemes (DSs) are computed automatically by processing the audio-visual data directly, such as by using feature extraction algorithms or content analysis methods. Other descriptions are generated manually by human annotation or by processing external meta-data related to the audio-visual material.”

So far, there were no automatic extraction algorithms used to populate the MetaMedia DS, but this matter will receive attention in future work.

“In MPEG-7, the DSs are categorised as pertaining specifically to the audio or visual domain, or pertaining generically to the description of multimedia.”

With exception of the Segment-part, all attributes in the MetaMedia DS are generically. In the Segment, a distinction is made for Texts, Images and Video. Since no audio was used yet, this part has not been implemented separately. Both Images and Video contain visual information, but they are separated on the temporal aspect, while the MMDS don’t make the separation on this level. Moreover, in MetaMedia it is believed that it can be useful to have the Text class, because some documents have specific attributes related with the text and the importance of those attributes might for example help describing mark-up documents.

6.8.1 Content organisation

"MPEG-7 provides DSs for organising and modelling collections of audio-visual content, segments, events, and/or objects, and describing their common properties. The collections can be further described using different models and statistics in order to characterise the attributes of the collection members.”

6.8.1.1 Collections

"The Collection Structure DS describes collections of audio-visual content or pieces of audiovisual material such as temporal segments of video. The Collection Structure DS groups the audio-visual content, segments, events, or objects into collection clusters and specifies properties that are common to the elements. The CollectionStructure DS describes also statistics and models of the attribute values of the elements, such as a mean colour histogram for
a collection of images. The CollectionStructure DS also describes relationships among collection clusters."

Through the hierarchy of Segment in the MetaMedia DS, the Description Unit can be further divided but then based on (differences in) content. However, there are no clusters made with properties common to the elements. Also the hierarchy available for the EventType was not intended for clusters, but to be able to make the descriptions of an event from the generic to the more specific. Thus, the clusters are not predefined, but might be composed with queries on the concerned hierarchies and properties.

6.8.1.2 Models

"The Model DSs provide tools for modelling the attributes and features of audio-visual content. The Probability Model DS provides fundamental DSs for specifying different statistical functions and probabilistic structures.

The Analytic Model DSs describe collections of examples of audio-visual data or clusters of Descriptors that provide models for particular semantic classes."

So far, tools for the modelling of the audio-visual content attributes and features are not implemented in the MetaMedia DS.

6.8.2 Content description

"MPEG-7 provides also DSs for content description. These elements describe the Structure (regions, video frames, and audio segments) and Semantics (objects, events, abstract notions)."

6.8.2.1 Structural aspects

"The core element of this part of the description is the Segment DS, that addresses the description of the physical and logical aspects of audio-visual content.

A segment represents a section of an audio-visual content item. It has five subclasses: AudioVisual Segment DS, Audio Segment DS, Still Region DS, Moving Region DS and Video Segment DS. Therefore, it may have both spatial and temporal properties.

The Segment DS is abstract (in the sense of object-oriented programming) and cannot be instantiated on its own: it is used to define the common properties of its subclasses. Any segment may be described by creation information, usage information, media information and
textual annotation. Moreover, a segment can be decomposed into sub-segments through the **Segment Decomposition DS**.

A segment is not necessarily connected, but may be composed of several non-connected components. Connectivity refers here to both spatial and temporal domains.”

“The **Segment DS** is recursive, i.e., it may be subdivided into sub-segments, and thus may form a hierarchy (tree). The resulting segment tree is used to define the media source, the temporal and / or spatial structure of the audio-visual content.”

In the **MetaMedia DS** the **Segment** class can be **Text, Image** and **Video**. Because the archive is, so far, only intended for texts, images and video, and not specific for audio-documents, no special subclass for audio was implemented. (But the DS can be easily extended with such a subclass for audio.) **Segment** defines the common properties of its subclasses **Text, Image** and **Video**, and is also an abstract class. Opposed to the MMDS, information about creation, usage, etc. is not connected to the **Segment** but that kind of information is linked at the **Description Unit**.

The temporal, spatial and spatio-temporal decomposition in the **MetaMedia DS** takes place in the **Segment** class: spatial concepts will be handled rather in the subclass **Image**, while (spatio-)temporal aspects will be handled in the subclass **Video**. So, if a **Segment** is for instance a video containing some text, then the specific characteristics of the text will be handled in the subclass **Text**. Specific frames of the video or regions of a shot will be treated in the **Image** subclass, while the temporal characteristics will be left for the **Video** subclass.

The **Video** subclass shows correspondence with the **VideoSegment DS**. The **Video** may be a **VideoObjectItem**, and **TimeDependent-Descriptors** may be associated with the **Video**.

Although there is no specific Audio subclass, the **Video** class includes **AudioObjectItems**. The characteristics of **Audio-ObjectItem** can be compared with the **AudioSegment DS**, but then with the audio seen as part of video.

The **Image** subclass can be compared with **StillRegion**.

The **MetaMedia DS** has no specific descriptors like the **Moving-RegionDS**, but items with ‘moving regions characteristics’ can be handled through the **Video** subclass.

The **MetaMedia DS** has a subclass **Text** (subclass of **Segment**), at the same level as **Image** and **Video**, which does not appear in the MMDS of MPEG.

Like in the MMDS, in Segment the hierarchy is based on a spatial or temporal division, and the sub-components don’t need to be connected (spatial or temporal).
"A segment may also be decomposed into various media sources such as various audio tracks or various viewpoints from several cameras. The hierarchical decomposition is useful to design efficient search strategies (global search to local search). It also allows the description to be scalable: a segment may be described by its direct set of Descriptors and DSs, but it may also be described by the union of the Descriptors and DSs that are related to its sub-segments."

The viewpoint of several cameras can be registered in the MetaMedia DS through the TimedependentDescriptor class (and Value). With the hierarchy of Segment, the description in a Segment can be refined by the descriptions in its subsegments, to any level of detail. So the description at the group level has its descriptors, and the description at the finer levels has new ones.

6.8.2.2 Conceptual aspects

"For applications where the structure is of no real use, but where the user is mainly interested in the semantic of the content, an alternative approach is provided by the Semantic DS. In this approach, the emphasis is not on segments but on Events, Objects, States and Concepts. The Event DS describes a meaningful temporal localisation. The event can be either concrete or abstract. A concrete event can have an instance in the real world or the media (e.g. wedding). An abstract event is created from a concrete event by abstraction, that is, by substituting variables in a concrete event for properties."

In the MetaMedia DS, the About Package serves as Semantic DS. Through the class Event semantic information can be registered about real world events, described in the MMDS as a meaningful temporal localisation. It is also possible in the MetaMedia DS to work with abstract and concrete events, and this by using the hierarchy of the EventType class.

"The Object DS describes semantically a physical or an abstract object. A physical object exists in the media or the real world as an entity with spatial and temporal extent. An abstract object is the resulting entity of applying abstraction to a physical object. As in the case of Event, abstraction is the process of replacing attributes of an entity by a variable. Essentially, this creates a template of the entity in question."
The *About Package* in the *MetaMedia DS* has a class *WorldEntity* that registers entities (called objects in MMDS) appearing in a document. There is no hierarchy attached to the *WorldEntity* class, so the objects can not be distinguished in abstract and concrete. However, it is possible to link more than one *WorldEntity* to the *Description Unit*. *WorldEntities* might be objects as well as persons in *MetaMedia*, where persons in the MMDS would get a subdivision in *AgentObject*.

In the MMDS these objects and events are also used to describe segments by defining relationships between them: for example, an object football can be kicked towards the object goal by an (agent) football-player, and this in the event ‘scoring’. The problem in *MetaMedia* is that the *events* and *worldentities* are linked with the *Description Unit*, and there are no descriptors defined for the relationships (between *Events* and *Worldentities*). Thus, to be able to describe the scene of the above example, the ‘relationship’ between *Events* and *Worldentities* can be described by using the *Segment Description*.

"The *State DS* identifies semantic properties of the entity at a given time, in a given spatial location, or in a given media location."

"Finally, the *Concept DS* describes abstract elements that are not created by abstraction from concrete objects and events. Concepts such as freedom or mystery are typical examples."

The *MetaMedia DS* information about the state of an entity and a concept of a document has to be handled in attributes like ‘Description’. For more specific searches it might however be worthwhile considering to add in the *About Package* a State class, linked to the *WorldEntity* Class, and a Concept class, linked to the *Description Unit*.

6.8.3 Content Management

"The Content Management description tools allow the description of the life cycle of the content, from content to consumption.

The content described by MPEG-7 descriptions can be available in different modalities, formats, coding schemes, and there can be several instances.

The only part of the description that depends on the storage media or the encoding format is the *MediaInformation*. The remaining part of the MPEG-7 description does not depend on the various profiles or instances and, as a result, can be used to describe jointly all possible copies of the content."
6.8.3.1 Media Description

"The Media Description describes the storage media such as the compression, coding and storage format of the audio-visual data. The Media Information DSs identify the master media, which is the original source from which different instances of the audio-visual content are produced. The instances of the audio-visual content are referred to as Media Profiles, which are versions of the master obtained perhaps by using different encodings, or storage and delivery formats. Each Media Profile is described individually in terms of the encoding parameters, storage media information and location."

"The MediaProfile D is composed of:

- **MediaFormat D**: contains description tools that are specific to the coding format of the media profile.
- **MediaInstance D**: contains the description tools that identify and locate the different media instances (copies) available of a media profile.
- **MediaTranscodingHints D**: contains description tools that specify transcoding hints of the media being described. The purpose of this D is to improve quality and reduce complexity for transcoding applications.
- **MediaQuality D**: represents quality rating information of an audio or visual content. It can be used to represent both subjective quality ratings and objective quality ratings."

The Storage Package of the MetaMedia DS includes data similar to the Media Information DS. Instead of Media Profiles, the term Support is used. Documents available in different formats, will imply as much instances of Support as there are formats (just like the Media Profiles). With this approach however, one can not specify what is the master media, because there is only one Description Unit, with different Supports, but the identification of Original and Copy happens in the Owner Package which is linked to the Description Unit. To identify master media and copy, there are 2 possible solutions:

- The information is put in an attribute of the Support class (e.g. ‘observation’).
- One can consider each format as a different Description Unit (can be linked in the hierarchy), each with one Support type, and the identification of the master media takes place in the Copy class of the Owner Package.

In the Storage Package the physical location is specified in the Storage class (cfr. MediaInstance D). The DSs derived from the Media Profile DS can be related to the Support class, like the AnalogSupport and DigitalSupport subclasses.
6.8.3.2 Creation Information

"The creation and production information description tools describe author-generated information about the generation/production process of the AV content. This information cannot usually be extracted from the content itself. This information is related to the material but it is not explicitly depicted in the actual content."

"The Creation D contains the description tools related to the creation of the content, including places, dates, actions, materials, staff (technical and artistic) and organisations involved."

In the MetaMedia DS this information can be found in the Creator Package. The Creation class contains where and when the content was created, while the 'by whom' is handled in the Creator class. Information about the history and the location of the creator are associated with the latter class.

"The Classification D contains the description tools that allow classifying the AV content. The Classification D is used for the description of the classification of the AV content. It allows searching and filtering based on user preferences regarding user-oriented classifications (e.g., language, style, genre, etc.) and service-oriented classifications (e.g., purpose, parental guidance, market segmentation, media review, etc.)."

In the Description Unit of the MetaMedia DS attributes like language, style and genre are present, but there are no specific filters implemented based on user preferences.

"The Related Material D contains the description tools related to additional information about the AV content available in other materials."

Similar content in the same archive can be found through the Owner Package of the Multimedia DS. The following possibilities exist:
- Related material of a Document can be found through the Related Description Unit, which registers any cross-relationships between Description Units.
- The Associated Material class registers the materials that are intimately related with the materials being described, but belong to another owner.
- Through the Copy class, copies of a Description Unit belonging to a certain owner can be found.
6.8.3.3 Usage Information

"The content usage information description tools describe information about the usage process of the AV content.

The UsageInformation DS description will incorporate new descriptions each time the content is used (e.g., UsageRecord DS, Income in Financial D), or when there are new ways to access to the content (e.g., Availability D)."

"The Rights D contains the description tools related to the rights holders of the annotated content (IPR) and the Access Rights.

The Financial D contains information related to the costs generated and income produced by AV content.

The Availability D contains the description tools related to the availability for use of the content.

The UsageRecord DS contains the description tools related to the past use of the content."

The Usage Information in the MetaMedia DS is rather limited, because that kind of information is more important for content creation purposes than for archives.

There exists no specific class for rights, but information like copyrights can be found in the Description Unit. Information about the Owner and Custody can be found in the Owner Package.

6.8.4 Navigation and access

"MPEG-7 provides also DSs for facilitating browsing and retrieval of audio-visual content by defining summaries, partitions and decompositions, and variations of the audio-visual material.

- **Summaries**: provide compact summaries of the audio-visual content to enable discovery, browsing, navigation, visualisation and sonification of audio-visual content. The Summary DSs involve two types of navigation modes: hierarchical and sequential. In the hierarchical mode, the information is organised into successive levels, each describing the audio-visual content at a different level of detail. In general, the levels closer to the root of the hierarchy provide more coarse summaries, and levels further from the root provide more detailed summaries. The sequential summary provides a sequence of images or video frames, possibly synchronised with audio, which may compose a slide-show or audio-visual skim."
The Description Unit contains a Summary component, but this is a summary at the document level. A summary of arbitrary audio-visual items is not provided in the MetaMedia scheme, but it would imply that a Summary component was added to the Segment, although the Segment Description might be considered as useful in this case.

- **Partitions and Decompositions**: describe different decompositions of the audio-visual signals in space, time and frequency. The partitions and decompositions can be used to describe different views of the audio-visual data, which is important for multi-resolution access and progressive retrieval."

This has no correspondence in the MetaMedia DS.

- **Variations**: provide information about different variations of audio-visual programs, such as summaries and abstracts; scaled, compressed and low-resolution versions; and versions with different languages and modalities – audio, video, image, text, and so forth.”

This has no correspondence in the MetaMedia DS.

### 6.8.5 User Interaction

"MPEG-7 provides also DSs for describing User Interaction. The UserPreference DSs describe preferences of users pertaining to the consumption of the audio-visual content. This allows the MPEG-7 audio-visual content descriptions to be matched to the UserPreference DS descriptions in order to select and personalise audio-visual content for more efficient and effective access, presentation and consumption."

The MetaMedia DS does not include specific elements for user interaction. However, it is foreseen that the interaction of the users with the developed system will be studied in the future, so there might come some adaptations to promote user interaction.
6.8.6 Basic Elements and Schema Tools

6.8.6.1 Basic Elements

"The MPEG-7 Multimedia DS specification has defined a number of basic elements that are used repeatedly as fundamental constructs throughout the definition of the MPEG-7 DSs. Many of the basic elements provide specific data-types and mathematical structures, such as vectors, matrices, and histograms, which are important for audio-visual content description. Also included as basic elements are constructs for linking media files and localising segments, regions, and so forth. Many of the basic elements address specific needs of audio-visual content description, such as the description of time, places, persons, individuals, groups, organisations, and other textual annotation."

In the MetaMedia DS the basic elements are left for specification in the language of implementation.

6.8.6.2 Schema Tools

"The MPEG-7 Multimedia DS specification has also defined a number of Schema Tools that facilitate the creation and packaging of MPEG-7 descriptions. The Schema Tools consist of MPEG-7 Root Elements, Top-Level Elements and Package Tools. The Root Elements, which are the starting elements of MPEG-7 descriptions, allow the creation of full MPEG-7 XML documents and MPEG-7 description fragments. The Top-Level Elements, which follow the MPEG-7 Root Element in an MPEG-7 description, organise the MPEG-7 DSs for specific content-oriented description tasks, such as for describing image, video, audio, or audio-visual content; collections; users; or world semantics. The Package Tools are designed for grouping or associating related DS components of a description into folders or packages. The packages are useful for organising and conveying the structure and types of MPEG-7 description information to search engines in querying and to aid users to overcome unfamiliarity with MPEG-7 description information in browsing."

In the MetaMedia DS the Schema Tools depend on the implementation. (Oracle for example has tools for the translation to XML documents.)

The MetaMedia DS is defined in UML, and therefore the language is used to group the Descriptors into Packages (SystemOfDescription, Storage, About, Creator, Owner).
7. Arch-metamedia

7.1 Relational translation for the prototype

The start-point of the Arch-metamedia prototype was a prototype that was first developed as a prototype of a descriptive archive system according to the ISAD(G) standard [2]. In Arch-metamedia the ‘Description Scheme for Still Images and Text’ was implemented in Oracle Designer, as described in [3], and the database was populated with images of the Portuguese Photo Center. (These images make also part of the MPEG-7 Content Set.)

With the intention of having a prototype for more diverse multimedia documents, Archmetamedia had to be updated with the Multimedia Description Scheme, discussed in the previous chapter. The MetaMedia DS defined in UML is the central piece where relational and markup articulate, so using the UML presentation, the MetaMedia DS was translated into relational schemes and put into Oracle Designer. The translation of the UML presentation of the DS into relational schemes can be seen in Figure 19 to 27.

![Diagram](attachment:image.png)

*Figure 19: Relational presentation of the Description Unit (SystemOfDescription – Package)*
7. Arch-metamedia

Figure 20: Relational presentation of the Segment (SystemOfDescription – Package)

The combination of Figures 19, 20 and 21 represent the relational presentation of the SystemOfDescription Package (cfr. Figure 10). Figure 19 shows that the Theme, Summary, Keywords and Description, all 'part of' the Description Unit, are implemented for multiple language use. The same counts for Segment Description in Figure 20.

To be able to implement more elaborate schemes (meaning that a level can have more than one level higher in the hierarchy), the Upperlevel class exists: this class clarifies for a certain level which levels can be ‘parent’ or upperlevel. (see Figure 21)
Figure 21: Relational presentation of the Hierarchy (SystemOfDescription – Package)

Figure 22: Relational presentation of the About – Package.
Figure 23: Relational presentation of the Storage-Package.

Figure 24: Relational presentation of the Creator – Package.
Figure 25: Relational Presentation of the Owner – Package (I).

Figure 26: Relational presentation of the Owner – Package (II.)
Figure 22 shows the relational presentation of the About Package (cfr. Figure 14). Event Description, Event-type Description, Entity Description and Entity Role Description are separate classes due to the multiple language implementation.

The Storage Package (cfr. Figure 13) in relational form is shown in Figure 23. The multiple language descriptors here are Storage Description and Support Observation.

Figure 24 shows the Creator Package (cfr. Figure 15). Creation Role is to identify the reason why the document was created, while the purpose of Creator Type is to make a classification of the kinds of creators (persons or organisations). Both are language dependent.

Figure 25 and 26 represent the relational presentation of the Owner Package (cfr. Figure 16). The role of an owner, the description of a custody and associated material, and the observation of a copy are all implemented as language dependent descriptors.

In the original prototype the concept of Location was more elaborated (with geographical information), and this implementation was kept in the current prototype (as can be seen in Figure 27).

```
<table>
<thead>
<tr>
<th>COUNTRY</th>
<th>LANGUAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td># CODE</td>
<td># CODE</td>
</tr>
<tr>
<td>NAME</td>
<td>NAME</td>
</tr>
<tr>
<td>has</td>
<td>describes</td>
</tr>
<tr>
<td></td>
<td>describes</td>
</tr>
<tr>
<td>LOC_NAME</td>
<td>LOC_NAME</td>
</tr>
<tr>
<td># LAN_CODE</td>
<td># LAN_CODE</td>
</tr>
<tr>
<td># LOC_CODE</td>
<td># LOC_CODE</td>
</tr>
<tr>
<td></td>
<td>belongs to</td>
</tr>
<tr>
<td></td>
<td>describes</td>
</tr>
<tr>
<td></td>
<td>describes</td>
</tr>
<tr>
<td>DISTRICT</td>
<td>LOCATION</td>
</tr>
<tr>
<td># CODE</td>
<td># CODE</td>
</tr>
<tr>
<td>NAME</td>
<td># CODE</td>
</tr>
<tr>
<td># COU_CODE</td>
<td># ENCODE</td>
</tr>
<tr>
<td></td>
<td>STARTDATE</td>
</tr>
<tr>
<td></td>
<td>TOW_CODE</td>
</tr>
<tr>
<td></td>
<td>VIL_CODE</td>
</tr>
<tr>
<td>has</td>
<td>has</td>
</tr>
<tr>
<td></td>
<td>has</td>
</tr>
<tr>
<td></td>
<td>has</td>
</tr>
<tr>
<td></td>
<td>has</td>
</tr>
<tr>
<td>TOWN</td>
<td>VILLAGE</td>
</tr>
<tr>
<td># CODE</td>
<td># CODE</td>
</tr>
<tr>
<td>NAME</td>
<td>NAME</td>
</tr>
<tr>
<td># DIS_CODE</td>
<td># TOW_CODE</td>
</tr>
<tr>
<td># ZC_SUB_ZIP_CODE</td>
<td># ZC_SUB_ZIP_CODE</td>
</tr>
<tr>
<td># ZC_ZIP_CODE</td>
<td># ZC_ZIP_CODE</td>
</tr>
</tbody>
</table>

Figure 27: Relational Presentation of the 'Utilities - Package'
Since the description scheme in the database changed, the data, already present from the previous scheme, had to be transferred into the new scheme. The most important differences here were:

- The different implementation of multiple language support.
- Information about creators and owners were gathered in ‘Agent’, with associated ‘Actions’; the information present in the classes Agent and Actions were transferred to the appropriate descriptors in the Creator and Owner Packages.

The difficulties encountered with the transfer to the new scheme were mainly due to the care of not loosing the data present in the database (of the Portuguese Photo Center). Since the introduction of data is very time-consuming, the loss of data had to be avoided. Therefore, the old and new scheme were thoroughly compared to define the necessary ‘strategy’ of copying, looking at the old and new constraints (‘null’, foreign keys...).

The current implementation resembles a relational presentation that is closer to the UML presentation of the Description Scheme than before.

7.2 Population of the prototype

To obtain more diversity in the documents in the prototype, Arch-metamedia was enriched with documents of the Archive of Eça de Queirós and a video (“Science Eye” –V10) of the MPEG-7 Content Set.

7.2.1 Archive of Eça de Queirós

The data of the Archive of Eça de Queirós, which has descriptions according to the ISAD norm, was delivered in MS Access, and thus had to be converted to fit in the tables of the MetaMedia Description Scheme. The mapping of the information in Access to the relational scheme in Oracle is summarised in the Table 8. (The first column presents the name of the column in MS Access, the second and third column give the name of the table and the attribute (column) respectively in Oracle to which it was mapped.)
Table 8: Mapping of Archive of Eça de Queirós (MS Access) into the MetaMedia Description Scheme

<table>
<thead>
<tr>
<th>MS Access</th>
<th>Oracle Table</th>
<th>Oracle Attribute</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Cod núcleo documental</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Título</td>
<td>Description Unit</td>
<td>Main_title</td>
</tr>
<tr>
<td>Referencia</td>
<td>Description Unit</td>
<td>Reference_code</td>
</tr>
<tr>
<td>Nível de descrição</td>
<td>Level</td>
<td>Description</td>
</tr>
<tr>
<td>Dimensão</td>
<td>Description Unit</td>
<td>Record_nature</td>
</tr>
<tr>
<td>História administrativa_biográfica</td>
<td>Creator_description</td>
<td>History</td>
</tr>
<tr>
<td>Organização e ordenação</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Data de produção inicial</td>
<td>Description Unit</td>
<td>Earliestdate</td>
</tr>
<tr>
<td>Data de produção final</td>
<td>Description Unit</td>
<td>Latestdate</td>
</tr>
<tr>
<td>Tipologia</td>
<td>Description Unit</td>
<td>Du_type</td>
</tr>
<tr>
<td>Unidades de descrição relacionadas</td>
<td>Related_du</td>
<td>Du_reference_code2 + Relation</td>
</tr>
<tr>
<td>Ambito e conteúdo</td>
<td>Summary</td>
<td>Summary</td>
</tr>
<tr>
<td>Características físicas</td>
<td>Description Unit</td>
<td>Preservation</td>
</tr>
<tr>
<td>Localização/cota</td>
<td>Storage</td>
<td>Location</td>
</tr>
<tr>
<td>Documento catalogado</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Dimensões (Cxl)</td>
<td>Description Unit</td>
<td>Du_size</td>
</tr>
<tr>
<td>Local de produção</td>
<td>Creation</td>
<td>Place</td>
</tr>
<tr>
<td>Autor</td>
<td>Creator</td>
<td>Name</td>
</tr>
<tr>
<td>Destinatário</td>
<td>Creator</td>
<td>Name</td>
</tr>
<tr>
<td>Suporte</td>
<td>Support</td>
<td>Medium</td>
</tr>
<tr>
<td>Idioma</td>
<td>Description Unit</td>
<td>Language</td>
</tr>
<tr>
<td>Cota antiga</td>
<td>Storage_description</td>
<td>Description</td>
</tr>
<tr>
<td>Tipo de letra</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Colocação antiga</td>
<td>Storage_description</td>
<td>Description</td>
</tr>
<tr>
<td>Estado de conservação</td>
<td>Description Unit</td>
<td>Preservation</td>
</tr>
<tr>
<td>OBS</td>
<td>Description Unit</td>
<td>Note</td>
</tr>
</tbody>
</table>

Some remarks about the introduction of the data in the MetaMedia Scheme:

- The names of the authors were put in the Creator table, with as associated creator-type ‘author’ (‘autor’).
- The names of the addressees were also put in the table of Creators, mentioning in the associated creation-role that it concerned an ‘addressee’ (‘destinatário’).
- For letter-type (‘tipo de letra’) there exists right now no appropriate field in the MetaMedia Scheme. However, one could easily extend the Description Unit with an attribute called ‘lettertype’, since other document-specific attributes like negative-number (for pictures), total-time-duration (for video) already exist in the Description Unit.

On the other hand, for some documents, especially those in markup languages, the lay-out is important and (sometimes) treated separately (as is the case with stylesheets). For such documents it would be recommendable to have a class ‘Layout’ linked to the Text class,
where all the information about the layout (like letter-type, font, colour) of the document could be gathered.

The Scheme ('Scheme1') or Hierarchy of the Archive of Eça de Queirós has the levels shown in Figure 29. It has more levels, and the 'upperlevels' are more diverse, compared to the Scheme ('Scheme0') of the images from the Portuguese Photo Center (Figure 28).

---

**Figure 28:** Levels of Scheme0 in Arch-metamedia.

**Figure 29:** Levels of Scheme1 in Arch-metamedia.
With the introduction of the data of the archive of Eça de Queirós, a small demonstration for an
archivist revealed some problems about the use of the Creator and Owner Package. An archivist
has a clear definition of every term in mind, for example who can be classified as ‘owner’ and
‘creator’. But in the MetaMedia Schema, all entities related with the context of a document have
to be described in the Creator and Owner Package. If we look for example at the addressees of
some letters in the Archive, then the names of the addressees are put in the Creator class and the
creation-role is set to ‘addressee’ (‘destinatário’). Thus, in a first categorisation, the entities are
assigned to Owner or Creator, and afterwards these categories are refined through creation-role,
creator-type and owner-role. This caused initially some confusion for the archivist for whom an
‘addressee’ is not a ‘creator’… But if the Description Scheme didn’t work with this higher
classification of context entities into Creator or Owner, every possible associated entity would
need to be implemented as a separate class or package. This way, one would probably loose the
more generic view on the Scheme.

Besides the problem of conversion between descriptions, the Archive of Eça de Queirós showed
also a lack of systematic organisation in the imported data. The data set was delivered in
electronic format, but not in a systematic machine-readable organisation.
Some information was implicit in the data. For example, the Archive of Eça de Queirós has a
strong hierarchy. Some subsections have a person’s name as title. It appears that those names
are actually the ‘creator’ (‘produtor’) of the Description Units of lower levels in the concerned
subsection. This was nowhere explicitly mentioned, so for the computer-scientist this
information would get lost, while for the archivist, if the information is also put in the Creator
Package, data is unnecessarily repeated. To avoid problems like this (for computer-scientists) in
the future, it would be useful if rules like above would be clarified by the archivists.
Thus, we are handling here with different views on the same material. For a human reader, the
creator in the title makes no confusion. For an automatic system that has an established place for
the creator, it is necessary to put it in that specific place.

The manual entering of data is of course subject to mistakes. The data in the archive of Eça de
Queirós was no exception on this ‘rule’. Some mistakes encountered were that the same
reference code appeared twice and that some upper reference codes (the reference codes of the
level above) were missing. Some other mistakes might be left unnoticed. The use of a
hierarchical composed reference code that is automatically produced (by use of the level and the
upper reference code), might resolve problems like this. (The Archive of Eça de Queirós uses a
hierarchical composed code, but the human interference caused here some irregularities.)
7.2.2 Video ("Science Eye": Bridge construction – V10) of MPEG-7 Content Set

The video ("Science Eye": Bridge construction) of the MPEG-7 Content Set was chosen because of the accompanied description of the 'shots', which could serve to populate and to demonstrate the use of the segment-part of the Arch-metamedia database. The difficulty to find descriptions like the one used, calls for automation, to be able to add easier the necessary descriptions. Of course, there is also the problem of handling the amount of information associated with a video.

![Diagram](image)

*Figure 30: Hierarchy of segments in video 'Science Eye'*

The video contains a hierarchy on the Segment level, which is presented in Figure 30. A concrete example of the division into segment levels, might give here a better view. The document level contains the complete video, and handles about a special steel used for the building of the longest bridge in the world. One of the scene-groups in this document is about the National Research Institute for Metals. A scene in this scene-group contains an introduction of Dr. Takahashi. In a last division that scene contains several cuts, among which a cut of the National Research Institute for Metals in Tukuba City, a cut with a zooming-in to Dr. Takahashi, and a cut where Dr. Takahashi explains about display monitors. The specific information that was attached about these described divisions, is put in Table 9, to have an idea what kind of information was available.

*Table 9: Example of information available of the MPEG-7 video.*

<table>
<thead>
<tr>
<th>Division</th>
<th>Begin</th>
<th>End</th>
<th>(Shot-)Abstract</th>
<th>Effect</th>
<th>Camera</th>
</tr>
</thead>
<tbody>
<tr>
<td>Document</td>
<td>00:00:00:00</td>
<td>00:09:15</td>
<td>Ultra-steel suspends the longest bridge in the world</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The mapping of the data (in the description) into the MetaMedia Scheme is shown in Table 10.

<table>
<thead>
<tr>
<th>MPEG-7 Description</th>
<th>Oracle Table</th>
<th>Oracle Attribute</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘program’, ‘scene-(group)’, ‘cut’</td>
<td>Segment</td>
<td>Seg_level</td>
</tr>
<tr>
<td>begin</td>
<td>Timeperiod</td>
<td>Starttime</td>
</tr>
<tr>
<td>end</td>
<td>Timeperiod</td>
<td>Stopotime</td>
</tr>
<tr>
<td>program title</td>
<td>Description_unit</td>
<td>Main_title</td>
</tr>
<tr>
<td>(shot-)abstract</td>
<td>Segment_description</td>
<td>Description</td>
</tr>
<tr>
<td>camera</td>
<td>Sample</td>
<td>Value</td>
</tr>
<tr>
<td>voice-abstract</td>
<td>Audioobjectitem</td>
<td>Description</td>
</tr>
<tr>
<td>Research engineer</td>
<td>Entity_description</td>
<td>Description (name)</td>
</tr>
<tr>
<td></td>
<td>Entity_role_description</td>
<td>Description (‘Research Engineer’)</td>
</tr>
<tr>
<td>Reporter</td>
<td>Entity_description</td>
<td>Description (name)</td>
</tr>
<tr>
<td></td>
<td>Entity_role_description</td>
<td>Description (‘Reporter’)</td>
</tr>
<tr>
<td>Creator</td>
<td>Name</td>
<td>(name)</td>
</tr>
<tr>
<td>Creator_type</td>
<td>Description</td>
<td>(‘Reporter’)</td>
</tr>
</tbody>
</table>

‘Camera motion’ is a Timedependent-descriptor (according to MPEG-7), and the associated values (like ‘pan’, ‘air-shot’..) are gathered in ‘value’ of the Sample class.
'Effects' were handled according to their nature:

- 'SUPER' was indicated in Segment as part of the 'cut' to which it belonged (a level lower), and got a code in the Text table (since it concerned a string)
- 'DVE' and 'DISSOLVE' were also indicated in Segment as part of the 'cut' to which they belonged, but they got a code in the Video table.

Instead of inserting video as a 'stand-alone' document, the data in the accompanied license document MPEG98/N2466 [11e] was used, to insert information about the other documents of the MPEG-7 content-set, in the following way:

- The hierarchy used for the MPEG-7 content set is shown in Figure 31.

![Diagram](image)

*Figure 31: Hierarchy of the MPEG-7 Content Set (Scheme 2)*

- 'Category' was put in genre (Description Unit), 'type' was inserted in du_type (Description Unit) and 'short description' found its place in description (Description).
- The content of 'source' (who is responsible for the making of the concerned item) was put in Creators name, while in Creation the associated Creation-role was put to 'source'.
- The owner is set to MPEG-7.
- For copyright information (Description Unit) the name of the license document is mentioned, and for the video inserted in the database, the license document is also put in Related Description Units.
- The medium (Support) is for all of these documents 'CD-ROM', while in location (Storage) the number of the CD-ROM of the MPEG-7 Content Set appears (for example, for the used video, this is '26', because the video is recorded on the 26th CD-ROM of the MPEG-7 Content Set).
- Since the images of the Portuguese Photo Center are also part of the MPEG-7 Content Set, the reference code of the 'fonds'-level of these images was linked to the reference code of the corresponding reference code on the 'document (collection)' level.
7.3 Queries in Arch-metamedia

The queries performed in the Arch-metamedia are all expressed in SQL and text-based. By using a Description Scheme for multimedia documents one can perform queries on each of the attributes, both context or content. For example, the documents created before 1500 can be searched, see Appendix B or the documents with ‘boat’ (‘barco’) as keyword can be queried. Moreover, the results of these queries can also be adapted: one can opt to retrieve the documents themselves (Appendix I), or specific information like where it is stored (Appendix B), who owns it... (To get these results, the foreign keys are used, and by defining appropriate views, these queries can be seriously simplified.) Queries can deliver results independent of the medium: because the queries are performed on the data (text) present in the Description Scheme, the results might be videos as well as images. Unfortunately this is hard to demonstrate with the current data in the database, since the video and the images have practically no subjects in common.

7.3.1 Hierarchy

The structure of archives is very important, therefore the hierarchy can not be neglected. With an Oracle Report, this hierarchy can be easily visualised as can be seen in Appendix C and D, for the hierarchy of the Archive of Eça de Queirós and the hierarchy in the segments of the “Science Eye” video respectively.

7.3.2 Segment

By using the classes related with the Segment, more documents can be retrieved than a search on keywords would deliver. An example is shown in Appendix E, where a query on ‘Cementite’ is performed. Both the Segment Description and the Audio-objectitem (which contains the voice-abstracts in text-format) are queried here. This way more parts of the video are retrieved for being relative to the subject, since the parts where a person is speaking about a certain subject can also be of interest, and not just those parts that have in their description the concerned word.
7.3.3 About Package

It is apparent how in none of the original data-sets used to populate the database, anything was mentioned about events or entities appearing in the document. Therefore these fields were populated with data that could be extracted from the other data (or by looking at the document itself), creating this way not a very diverse population. Nevertheless, the About Package is considered as very important for conducting searches on content, since it concerns matters that people easily remember: a king (entity) on an image or a video about a marriage (event) for example.

The classes Event-type and Entity-role are intended to create a controlled vocabulary for more easily querying. (An important consideration here is that SQL only returns exact matches.) Event-type (with its hierarchy) intends to give a more general description, while Event is more specific oriented, for example: sale (‘venda’) for Event-type, has the related Events sale of real estate (‘venda de imobilário’) and transactions (‘transações’), see Appendix F.

7.3.4 Owner & Creator

In the packages Owner and Creator there are some obvious context queries that can be performed, like who is the author (creator-type) of a document, or to which archive (owner-role) belongs a document.

On the other hand, with the class Creation Role (‘Crea_role’) one can query also on content, since this class is intended to contain the reasons of creation for a document. For example, the documents can be searched that where destined (creation role = ‘Destinatário’) for Eça de Queirós (Creator = ‘Eça de Queirós), see Appendix G.

7.3.5 Combinations

If one wants to perform a specific search, it can be useful to combine fields for querying. For example if we know that Eça de Queirós was addressee of a document, and the document had passport (‘passaporte’) in its title, see Appendix G. Or when we are looking for deeds (du_type = ‘escrituras’) about the sale of real estate (event = ‘venda imobilario’), see Appendix H. By combining queries, one can narrow the search-results, for example, if we know there is an image about a party (event-type = ‘festa’), and there was a table shown (entity = ‘mesa’), see Appendix I.
On the other hand, if users are not performing specific searches, they can retrieve more results by querying different fields. For example, somebody who wants to retrieve documents about transactions ('transacções') can query the documents with 'transactions' appearing in the fields of title, keyword, description, summary, event or entity. (see Appendix J) Another example is one of the previous queries, that searched for 'Cementite' in the fields of segment_description and audio-objectitem (Appendix E).

Searches (on content) can be performed both on the Description Unit 'level' and the Segment, but this is not demonstrated with an example due to the diversity of the documents present on both 'levels'. It is for example possible to combine queries for words in the fields of segment-description and summary together.

To get a ranking in the results, it would be necessary to give a weight to each field (so the ones with 'transactions' appearing in 'event' could be ranked higher than those who have the word in the title for instance.) Of course, the question is popping up here how the weights would be assigned to the fields... Querying documents by their descriptions means that the associated weights don't have to be assigned by the number of occurrences of a certain word, a technique which is currently used a lot, but which is not always a good representative for assigning a weight of relevance.

7.4 Comparison

The method used, namely the use of a scheme for metadata to archive, is also demonstrated in 'Describing and retrieving photos using RDF and HTTP' ([48] see 5.4.2), but that (W3C) project worked only with images and used a schema based on RDF. Because they concentrate on photos their 'Technical Schema' and 'Content Schema' contains more specific attributes than the MetaMedia Description Scheme. The latter can be used for any multimedia document, but nevertheless most of the information of their RDF Schema could be transferred without much problems to the MetaMedia Scheme. The elements of the 'Dublin Core Schema' are all contained in the MetaMedia Description Scheme, mostly in a more elaborate manner.

In the 'suggestions for extensions' the authors suggest that there could be an additional schema for further describing portrait and group photos that would allow listing the people in the picture. In the MetaMedia Scheme this can be done through the Entity class in the About Package. On the other hand, if the authors think about topic-specific schemes, this raises again the minimalist/structuralist issue (see 2.4): more specific schemes and useful for small communities, or more general schemes and common applicable? (A detailed scheme allows for
very specific searches, but for a user not familiar with the used scheme or for a user searching without that specific target in mind, it appears easier to search in a more general scheme.)

Because sometimes a detail of a photo is interesting as a picture on its own, another suggestion is that there could be a system to address a region of a photo and attach metadata to it. This 'system' is implemented in the MetaMedia Scheme by means of the hierarchy of Segment.

Looking at a system implemented on the Internet, namely WebSEEK (see 5.4.3), it appears more 'limited' than the searches we can perform with the MetaMedia Scheme: in WebSEEK searches are performed by selecting 'categories' or subject classes. But by combining the SAFE system (see 5.4.3) with WebSEEK one can perform searches on content (syntactic features), like a colour appearing in a certain region. So far, this is not implemented in the Arch-metamedia prototype, but by using Oracle's Visual Information Retrieval Package [50], this shortcoming might be overcome in the future. The classes to be used for this kind of queries are the Timedependent-descriptors and their associated Samples. (Right now, there is only one Timedependent-descriptor modelled in the prototype, namely 'camera motion', of the 'Science Eye' video.)

7.5 Conversion to XML in Arch-metamedia

The MetaMedia Description Scheme defined in UML and implemented in Rational Rose, was used by the MetaMedia team of the University of Minho, to translate into XML by using the intermediate XML language.

The conversion of the data in the relational prototype to data in XML format is a task that is being worked on by the MetaMedia team of the University of Minho.

A program (written in Perl) takes care of the generation of a DTD out of the structure of the database. The approach therefore followed (discussed in [54]) is the approach taken by the XML Authority in providing automated schema creation from relational data sources. It provides mechanical mappings of key database structure concepts. Another program (written in Java) puts the data of the database in XML format.
Arch-MetaMedia is a prototype based on the MetaMedia Description Scheme, that can be used for multimedia documents. The used test-data are images of the Portuguese Photo Center and a video of NHK, both part of the MPEG-7 Content Set, and the Eça de Queirós Historic Archive, property of the Fundação de Eça de Queirós.

The presented examples of queries show how the metadata can be used to retrieve other metadata or the document itself. The MetaMedia Description Scheme contains both context and content metadata, thereby also allowing searches both on content and context.

The implementation of Arch-MetaMedia demonstrates the usefulness and the need of a metadata description scheme to search and retrieve multimedia documents. The documents can be retrieved by their associated descriptions, and independently of the number of occurrences of a certain term.

The MetaMedia Description Scheme itself is still open for extension. With the population of more data, more needs for specific attributes will be discovered, so the Description Scheme might be adapted. Some possible extensions have already been noted, (for example a ‘Layout’ class linked to the Text class.)

Closer contact with archivist is needed to discover rules of implicit data, which is invisible to the computer-system. Otherwise non-professional users might ‘loose’ information.
8. Conclusion

Multimedia searching is a relatively new field, and has still many research issues. It requires close interaction between multiple technical disciplines and application users. Researchers have made great progress in recent years, but a few critical issues have still not been addressed adequately. Today's content-based image search systems allow for image queries based on image examples, feature specification, and text-based search. But the linkage between low-level visual features and high-level semantics is still weak.

By the use of a Description Scheme that combines context and content metadata, this dissertation shows that context metadata and content metadata complement each other for the retrieval of multimedia documents. The defined Description Scheme was tested in the implementation of a prototype database system for information storage and retrieval.

The queries performed on the developed prototype demonstrate that the use of metadata, and more specific the proposed Description Scheme, makes querying multimedia documents an easier task, since one is not limited to merely searching keywords or subject-classes. By using a Description Scheme the multimedia documents can be retrieved on their metadata descriptions, be it context or content.

For commercialisation of the system, the standardisation of a description scheme and the automatic introduction of the metadata are requirements for easy and widespread dissemination. Especially for use on the Web, the introduction of metadata is an overwhelming job if not done automatically. For users unfamiliar with the underlying schema, the composition of the queries should happen with some guidance. Moreover, the interface has to be adapted to be able to show the different formats that might be the result of a query.

Future work

The use of metadata automatically provides a non-visual description of documents, hence contributing to their accessibility. Initiatives like the MPEG-7 Multimedia Content Description Interface are important for the global applicability (and recognition) of the Description Schemes.
Although recent research focuses more on content metadata, the value of context metadata will not lose its value, especially not for archives. With the awareness of the value of intellectual properties for example, attributes like creator or copyright will be an essential part of the information associated with a document.

For a wide-scale integration of Metadata Schemes, the development of automatic extraction tools for the attributes will be necessary. The continuation of the MetaMedia project – MetaMedia II – has as an objective to integrate feature extraction tools in the developed prototype. One of the possibilities that could be explored is the interMedia package of Oracle [48], because it can for instance extract automatically data from a video and store the metadata information in attributes of the Oracle InterMedia Video object. So if the data from that video object could be transferred into the appropriate tables of the MetaMedia Description Scheme, those data would not have to be entered manually anymore. Also, with the evolutions in technology, more and more metadata will get produced automatically, like in digital cameras for example. Therefore, ways should be explored to transfer the data automatically into the Description Scheme. With the amount of data available, it would relieve the job of cataloguing if the description of multimedia objects was generated at the time of creation.

MetaMedia II will also research further the possible exchanges between relational data and markup, more specific XML: so far, relational databases have the highest efficiency in querying data, while XML is getting the universal standard (on the Web) for exchange of data.

The evolutions concerning the MPEG-7 Descriptors will be followed, to update the MetaMedia Description Scheme with Descriptors that are not yet integrated, but are considered as important by MPEG-7.

MetaMedia II will devote also activities to the development of tools for a multimedia archival environment: the prototype will serve to develop an appropriate interface, which makes it easier for archivist to visualise and enter the data, and which allows users to perform queries. It is clear that the wide divulgation of a system depends on the user-friendliness of the system, so this aspect cannot be neglected.
9. References


[3] GAIO, Susana Maria Silva Santos – Meta-informação no armazenamento e recuperação de documentos multimédia. Dissertation for Master Course in Information Management accompanied by Prof. Dr. Maria Cristina Ribeiro and presented in the Faculty of Engineering of the University of Porto in 2000.


http://www.w3.org/RDF/Overview.html

[10a] LASSILA, O; SWICK, R. - RDF Model and Syntax Specification.
W3C Proposed Recommendation, February 1999

W3C Candidate Recommendation, March 2000.


Doc. ISO/MPEG N2466, MPEG, Atlantic City, October 1998


http://www.oasis-open.org/cover/general.html

SGML Centre, 1992.
http://www.personal.u-net.com/~sgml/sgml.htm

[16] W3C – HyperText Markup Language
http://www.w3.org/MarkUp

[16a] XHTML 1.0: The Extensible HyperText Markup Language.
http://www.w3.org/TR/2000/REC-xhtml1-20000126

[17] W3C – Cascading Style Sheet Language
http://www.w3.org/Style/CSS/
http://www.w3.org/TR/2000/REC-xml-20001006

http://www-rodi.inria.fr/Fmbrepubs_dana.html

http://www.xml.com/pub/98/10/guide0.html


[22] *W3C XML Schemas*
http://www.w3.org/XML/Schema

W3C Candidate Recommendation, October 2000.
http://www.w3.org/TR/2000/CR-xmlschema-1-20001024/

W3C Candidate Recommendation, October 2000.

[23] *W3C Extensible Stylesheet Language*
http://www.w3.org/Style/XSL/Overview.html

[24a] CLARK, J. -- *XSL Transformations (XSLT) version 1.0*.
http://www.w3.org/TR/1999/REC-xslt-19991116

http://www.w3.org/TR/REC-xml-names


http://www.es.colorado.edu/~getrich/Classes/csci5817/Term_Papers/veeramani/ora.html


Appendix A
**UML terminology**

**Package**

![Package diagram]

Packages serve to partition the logical model of a system. They are clusters of highly related classes that are themselves cohesive, but are loosely coupled relative to other such clusters. A package can be referenced by its name.

**Class**

![Class diagram]

A class captures the common structure and common behaviour of a set of objects. A class is an abstraction of real-world items. A class can be referenced by its name, and can define several attributes and operations.

**Association**

![Association diagram]

An association represents a semantic connection between classes. The cardinality of an association can be 0..0, 0..1, 0..n, 1..1, 1..n or n, on both ends of the association.
Generalisation

A generalisation relationship between classes shows that the subclass (A) shares the structure or behaviour defined in one or more superclasses (B), where B contains the generic elements, and A the more specific elements. ('is-a' relationship) The declaration of the specific elements is completely consistent with the declaration of the generic elements, whereby more additional information can be added.

Aggregation

An aggregation shows a whole and part relationship, where A is a part component and B is the aggregate class.
Appendix B
Query on latest date

Reference Code: PT/FEQ/FAMEQ/B-A-005-d1
Main Title: Autos de confirmação
Earliest Date: 02 NOV 1448 Latest Date: 02 NOV 1448
Storage Location: 11-2 Medium: pergaminho

Reference Code: PT/FEQ/FAMEQ/B-A-025-d1
Main Title: Escrituras de arrendamento (família)
Earliest Date: 06 JAN 1456 Latest Date: 06 JAN 1456
Storage Location: 11-4 Medium: pergaminho

Reference Code: PT/FEQ/FAMEQ/B-B-029-d1
Main Title: Escrituras de doação
Earliest Date: 12 JUL 1480 Latest Date: 12 JUL 1480
Storage Location: 11-19 Medium: pergaminho

latest date
## Query on hierarchy of Archive de Eça

<table>
<thead>
<tr>
<th>Level Description</th>
<th>Level Code</th>
<th>Reference Code</th>
<th>Upper Code</th>
<th>Main Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fonds</td>
<td>1</td>
<td>PT/FEQ</td>
<td>PT/FEQ</td>
<td>Arquivo Pessoal de Eça de Queirós</td>
</tr>
<tr>
<td>Sections</td>
<td>3</td>
<td>PT/FEQ/APEQ/A</td>
<td>PT/FEQ/APEQ/A</td>
<td>Actividade privada (Eça)</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/APEQ/A-004-004</td>
<td>PT/FEQ/APEQ/A-004</td>
<td>Certidões (1892-1900)</td>
</tr>
<tr>
<td>Documents</td>
<td>8</td>
<td>PT/FEQ/APEQ/A-004-d1</td>
<td>PT/FEQ/APEQ/A-004-d1</td>
<td>Certidões (1892-1900)</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/APEQ/A-007-007</td>
<td>PT/FEQ/APEQ/A-007-007</td>
<td>Correspondência recebida (1869-1900)</td>
</tr>
<tr>
<td>SubSeries</td>
<td>6</td>
<td>PT/FEQ/APEQ/A-007-007-007</td>
<td>PT/FEQ/APEQ/A-007-007-007</td>
<td>Correspondência recebida de Família Eça de Queirós</td>
</tr>
<tr>
<td>Documents</td>
<td>6</td>
<td>PT/FEQ/APEQ/A-007-007-d1</td>
<td>PT/FEQ/APEQ/A-007-007-d1</td>
<td>Correspondência recebida (1869-1900)</td>
</tr>
<tr>
<td>Documents</td>
<td>8</td>
<td>PT/FEQ/APEQ/A-007-007-d2</td>
<td>PT/FEQ/APEQ/A-007-007-d2</td>
<td>Correspondência recebida (1869-1900)</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/APEQ/A-009-009-009</td>
<td>PT/FEQ/APEQ/A-009-009-009</td>
<td>Diplomas de Eça</td>
</tr>
<tr>
<td>Documents</td>
<td>8</td>
<td>PT/FEQ/APEQ/A-009-009-d2</td>
<td>PT/FEQ/APEQ/A-009-009-d2</td>
<td>Diplomas de Eça</td>
</tr>
<tr>
<td>Documents</td>
<td>8</td>
<td>PT/FEQ/APEQ/A-009-009-d2</td>
<td>PT/FEQ/APEQ/A-009-009-d2</td>
<td>Diplomas de Eça</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/APEQ/A-011-011</td>
<td>PT/FEQ/APEQ/A-011-011</td>
<td>Documentos diversos (Eça de Queirós)</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/APEQ/A-012-012</td>
<td>PT/FEQ/APEQ/A-012-012</td>
<td>Documentos pessoais de Eça de Queirós</td>
</tr>
<tr>
<td>Composed</td>
<td>7</td>
<td>PT/FEQ/APEQ/A-012-012</td>
<td>PT/FEQ/APEQ/A-012-012</td>
<td>Documentos pessoais de Eça de Queirós</td>
</tr>
<tr>
<td>Documents</td>
<td>8</td>
<td>PT/FEQ/APEQ/A-012-012-d1</td>
<td>PT/FEQ/APEQ/A-012-012-d1</td>
<td>Documentos pessoais de Eça de Queirós</td>
</tr>
<tr>
<td>Composed</td>
<td>7</td>
<td>PT/FEQ/APEQ/A-012-012-d2</td>
<td>PT/FEQ/APEQ/A-012-012-d2</td>
<td>Documentos pessoais de Eça de Queirós</td>
</tr>
<tr>
<td>Documents</td>
<td>8</td>
<td>PT/FEQ/APEQ/A-014-014</td>
<td>PT/FEQ/APEQ/A-014-014</td>
<td>Passaportes (Eça)</td>
</tr>
<tr>
<td>SubSections</td>
<td>4</td>
<td>PT/FEQ/APEQ/A-A</td>
<td>PT/FEQ/APEQ/A-A</td>
<td>Gestão patrimonial (Eça)</td>
</tr>
<tr>
<td>Composed</td>
<td>7</td>
<td>PT/FEQ/APEQ/A-A-007-007</td>
<td>PT/FEQ/APEQ/A-A-007-007</td>
<td>Correspondência recebida (patrimonial)</td>
</tr>
<tr>
<td>Composed</td>
<td>7</td>
<td>PT/FEQ/APEQ/A-A-007-007-007</td>
<td>PT/FEQ/APEQ/A-A-007-007-007</td>
<td>Correspondência recebida (patrimonial)</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/APEQ/A-A-008</td>
<td>PT/FEQ/APEQ/A-A-008</td>
<td>Declarações de impostos</td>
</tr>
<tr>
<td>Documents</td>
<td>5</td>
<td>PT/FEQ/APEQ/A-A-017</td>
<td>PT/FEQ/APEQ/A-A-017</td>
<td>Processos particulares relativos a seguros</td>
</tr>
</tbody>
</table>

**Hierarchy**
<table>
<thead>
<tr>
<th>Level Description</th>
<th>Level Code</th>
<th>Reference Code</th>
<th>Upper Code</th>
<th>Main Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>Composed Documents</td>
<td>7</td>
<td>PT/FEQ/APEQ/A-A-017-p1</td>
<td>PT/FEQ/APEQ/A-A-017</td>
<td>Processos particulares relativos a seguros</td>
</tr>
<tr>
<td>Documents Series</td>
<td>5</td>
<td>PT/FEQ/APEQ/A-A-018</td>
<td>PT/FEQ/APEQ/A-A</td>
<td>Recibos de letras</td>
</tr>
<tr>
<td>SubSections 4</td>
<td>4</td>
<td>PT/FEQ/APEQ/A-B</td>
<td>PT/FEQ/APEQ/A</td>
<td>Colecção/objectos (Eça)</td>
</tr>
<tr>
<td>SubSections 4</td>
<td>4</td>
<td>PT/FEQ/APEQ/A-C</td>
<td>PT/FEQ/APEQ/A</td>
<td>Colecção/periódicos (Eça)</td>
</tr>
<tr>
<td>Sections 3</td>
<td>3</td>
<td>PT/FEQ/APEQ/B</td>
<td>PT/FEQ/APEQ/B</td>
<td>Actividade pública (Eça)</td>
</tr>
<tr>
<td>SubSections 4</td>
<td>4</td>
<td>PT/FEQ/APEQ/B-D</td>
<td>PT/FEQ/APEQ/B-D</td>
<td>Consulado de Bristol</td>
</tr>
<tr>
<td>Series 5</td>
<td>5</td>
<td>PT/FEQ/APEQ/B-D-013</td>
<td>PT/FEQ/APEQ/B-D</td>
<td>Escrituras de arrendamento (Bristol)</td>
</tr>
<tr>
<td>Series 5</td>
<td>5</td>
<td>PT/FEQ/APEQ/B-D-014</td>
<td>PT/FEQ/APEQ/B-D</td>
<td>Passaportes (1880-1887)</td>
</tr>
<tr>
<td>Documents Series</td>
<td>8</td>
<td>PT/FEQ/APEQ/B-D-014-d2</td>
<td>PT/FEQ/APEQ/B-D-014</td>
<td>Passaportes (1880-1887)</td>
</tr>
<tr>
<td>Documents Series</td>
<td>8</td>
<td>PT/FEQ/APEQ/B-D-014-d1</td>
<td>PT/FEQ/APEQ/B-D-014</td>
<td>Passaportes (1880-1887)</td>
</tr>
<tr>
<td>Series 5</td>
<td>5</td>
<td>PT/FEQ/APEQ/B-D-010</td>
<td>PT/FEQ/APEQ/B-D</td>
<td>Documentos de despesa (Bristol)</td>
</tr>
<tr>
<td>SubSections 4</td>
<td>4</td>
<td>PT/FEQ/APEQ/B-E</td>
<td>PT/FEQ/APEQ/B-E</td>
<td>Consulado de Paris</td>
</tr>
<tr>
<td>Series 5</td>
<td>5</td>
<td>PT/FEQ/APEQ/B-E-007</td>
<td>PT/FEQ/APEQ/B-E</td>
<td>Correspondência recebida em Paris</td>
</tr>
<tr>
<td>SubSeries 6</td>
<td>6</td>
<td>PT/FEQ/APEQ/B-E-007-02</td>
<td>PT/FEQ/APEQ/B-E-007</td>
<td>Correspondência recebida de entidades superiores em Paris</td>
</tr>
<tr>
<td>Series 5</td>
<td>5</td>
<td>PT/FEQ/APEQ/B-E-014</td>
<td>PT/FEQ/APEQ/B-E</td>
<td>Passaportes (1890-1899)</td>
</tr>
<tr>
<td>Documents Series</td>
<td>8</td>
<td>PT/FEQ/APEQ/B-E-014-d3</td>
<td>PT/FEQ/APEQ/B-E-014</td>
<td>Passaportes (1890-1899)</td>
</tr>
<tr>
<td>Documents Series</td>
<td>8</td>
<td>PT/FEQ/APEQ/B-E-014-d1</td>
<td>PT/FEQ/APEQ/B-E-014</td>
<td>Passaportes (1890-1899)</td>
</tr>
<tr>
<td>Documents Series</td>
<td>8</td>
<td>PT/FEQ/APEQ/B-E-014-d2</td>
<td>PT/FEQ/APEQ/B-E-014</td>
<td>Passaportes (1890-1899)</td>
</tr>
<tr>
<td>Series 5</td>
<td>5</td>
<td>PT/FEQ/APEQ/B-E-014</td>
<td>PT/FEQ/APEQ/B-E</td>
<td>Processos de correspondência de Paris</td>
</tr>
<tr>
<td>Series 5</td>
<td>5</td>
<td>PT/FEQ/APEQ/B-E-016</td>
<td>PT/FEQ/APEQ/B-E</td>
<td>Processos particulares de empréstimo (Paris)</td>
</tr>
<tr>
<td>Documents Series</td>
<td>7</td>
<td>PT/FEQ/APEQ/B-E-016-p1</td>
<td>PT/FEQ/APEQ/B-E-016</td>
<td>Processos particulares de empréstimo (Paris)</td>
</tr>
<tr>
<td>Series 5</td>
<td>5</td>
<td>PT/FEQ/APEQ/B-E-005</td>
<td>PT/FEQ/APEQ/B-E</td>
<td>Conta corrente (1889-1900)</td>
</tr>
<tr>
<td>Composed Documents</td>
<td>7</td>
<td>PT/FEQ/APEQ/B-E-005-p1</td>
<td>PT/FEQ/APEQ/B-E-005</td>
<td>Conta corrente (1889-1900)</td>
</tr>
<tr>
<td>Documents Series</td>
<td>5</td>
<td>PT/FEQ/APEQ/B-E-010</td>
<td>PT/FEQ/APEQ/B-E</td>
<td>Documentos de despesa de Paris</td>
</tr>
<tr>
<td>Series 5</td>
<td>5</td>
<td>PT/FEQ/APEQ/B-E-013</td>
<td>PT/FEQ/APEQ/B-E</td>
<td>Escrituras de arrendamento (Paris)</td>
</tr>
<tr>
<td>Series 5</td>
<td>5</td>
<td>PT/FEQ/APEQ/B-E-020</td>
<td>PT/FEQ/APEQ/B-E</td>
<td>Processos de liquidação de sociedade comercial</td>
</tr>
<tr>
<td>SubSections 4</td>
<td>4</td>
<td>PT/FEQ/APEQ/B-B</td>
<td>PT/FEQ/APEQ/B-B</td>
<td>Consulado de Havana</td>
</tr>
<tr>
<td>Series 5</td>
<td>5</td>
<td>PT/FEQ/APEQ/B-B-003</td>
<td>PT/FEQ/APEQ/B-B-003</td>
<td>Autos de nomeação (Havana)</td>
</tr>
<tr>
<td>Documents Series</td>
<td>8</td>
<td>PT/FEQ/APEQ/B-B-003-d1</td>
<td>PT/FEQ/APEQ/B-B-003</td>
<td>Autos de nomeação (Havana)</td>
</tr>
<tr>
<td>Series 5</td>
<td>5</td>
<td>PT/FEQ/APEQ/B-B-007</td>
<td>PT/FEQ/APEQ/B-B-007</td>
<td>Correspondência recebida em Havana</td>
</tr>
<tr>
<td>Level Description</td>
<td>Level Code</td>
<td>Reference Code</td>
<td>Upper Code</td>
<td>Main Title</td>
</tr>
<tr>
<td>-------------------</td>
<td>------------</td>
<td>----------------</td>
<td>------------</td>
<td>------------</td>
</tr>
<tr>
<td>SubSeries 6</td>
<td>6</td>
<td>PT/FEQ/APEQ/B-B-007-02</td>
<td>PT/FEQ/APEQ/B-B-007</td>
<td>Correspondência recebida de entidades superiores em Havana</td>
</tr>
<tr>
<td>Documents 8</td>
<td>8</td>
<td>PT/FEQ/APEQ/B-B-007-02-d1</td>
<td>PT/FEQ/APEQ/B-B-007-02</td>
<td>Correspondência recebida de entidades superiores em Havana</td>
</tr>
<tr>
<td>Documents 8</td>
<td>8</td>
<td>PT/FEQ/APEQ/B-B-007-02-d2</td>
<td>PT/FEQ/APEQ/B-B-007-02</td>
<td>Correspondência recebida de entidades superiores em Havana</td>
</tr>
<tr>
<td>Documents 8</td>
<td>8</td>
<td>PT/FEQ/APEQ/B-B-007-02-d3</td>
<td>PT/FEQ/APEQ/B-B-007-02</td>
<td>Correspondência recebida de entidades superiores em Havana</td>
</tr>
<tr>
<td>Series 5</td>
<td>5</td>
<td>PT/FEQ/APEQ/B-B-014</td>
<td>PT/FEQ/APEQ/B-B</td>
<td>Passaportes (1872-1874)</td>
</tr>
<tr>
<td>Documents 8</td>
<td>8</td>
<td>PT/FEQ/APEQ/B-B-014-d1</td>
<td>PT/FEQ/APEQ/B-B-014</td>
<td>Passaportes (1872-1874)</td>
</tr>
<tr>
<td>Documents 8</td>
<td>8</td>
<td>PT/FEQ/APEQ/B-B-014-d2</td>
<td>PT/FEQ/APEQ/B-B-014</td>
<td>Passaportes (1872-1874)</td>
</tr>
<tr>
<td>Documents 8</td>
<td>8</td>
<td>PT/FEQ/APEQ/B-B-014-d3</td>
<td>PT/FEQ/APEQ/B-B-014</td>
<td>Passaportes (1872-1874)</td>
</tr>
<tr>
<td>SubSections 4</td>
<td>4</td>
<td>PT/FEQ/APEQ/B-C</td>
<td>PT/FEQ/APEQ/B</td>
<td>Consulado de Newcastle-On-Tyne</td>
</tr>
<tr>
<td>Series 5</td>
<td>5</td>
<td>PT/FEQ/APEQ/B-C-002</td>
<td>PT/FEQ/APEQ/B-C</td>
<td>Autos de aceitação diplomática</td>
</tr>
<tr>
<td>Documents 8</td>
<td>8</td>
<td>PT/FEQ/APEQ/B-C-002-d1</td>
<td>PT/FEQ/APEQ/B-C-002</td>
<td>Autos de aceitação diplomática</td>
</tr>
<tr>
<td>Series 5</td>
<td>5</td>
<td>PT/FEQ/APEQ/B-C-003</td>
<td>PT/FEQ/APEQ/B-C</td>
<td>Autos de nomeação (Newcastle-On-Tyne)</td>
</tr>
<tr>
<td>Documents 8</td>
<td>8</td>
<td>PT/FEQ/APEQ/B-C-003-d1</td>
<td>PT/FEQ/APEQ/B-C-003</td>
<td>Autos de nomeação (Newcastle-On-Tyne)</td>
</tr>
<tr>
<td>Series 5</td>
<td>5</td>
<td>PT/FEQ/APEQ/B-C-007</td>
<td>PT/FEQ/APEQ/B-C</td>
<td>Documentos de despesa de Newcastle-on-Tyne</td>
</tr>
<tr>
<td>SubSections 4</td>
<td>4</td>
<td>PT/FEQ/APEQ/B-A</td>
<td>PT/FEQ/APEQ/B</td>
<td>Administração do Concelho de Leiria</td>
</tr>
<tr>
<td>Series 5</td>
<td>5</td>
<td>PT/FEQ/APEQ/B-A-003</td>
<td>PT/FEQ/APEQ/B-A</td>
<td>Autos de nomeação (Leiria)</td>
</tr>
<tr>
<td>Documents 8</td>
<td>8</td>
<td>PT/FEQ/APEQ/B-A-003-d1</td>
<td>PT/FEQ/APEQ/B-A-003</td>
<td>Autos de nomeação (Leiria)</td>
</tr>
<tr>
<td>Sections 3</td>
<td>3</td>
<td>PT/FEQ/APEQ/C</td>
<td>PT/FEQ/APEQ</td>
<td>Actividade editorial</td>
</tr>
<tr>
<td>Series 5</td>
<td>5</td>
<td>PT/FEQ/APEQ/C-001</td>
<td>PT/FEQ/APEQ/C</td>
<td>Artigos propostos</td>
</tr>
<tr>
<td>Series 5</td>
<td>5</td>
<td>PT/FEQ/APEQ/C-006</td>
<td>PT/FEQ/APEQ/C</td>
<td>Contratos de Sociedade</td>
</tr>
<tr>
<td>Series 5</td>
<td>5</td>
<td>PT/FEQ/APEQ/C-019</td>
<td>PT/FEQ/APEQ/C</td>
<td>Revisões de provas</td>
</tr>
<tr>
<td>Documents 8</td>
<td>8</td>
<td>PT/FEQ/APEQ/C-019-d1</td>
<td>PT/FEQ/APEQ/C-019</td>
<td>Revisões de provas</td>
</tr>
<tr>
<td>Series 5</td>
<td>5</td>
<td>PT/FEQ/APEQ/C-007</td>
<td>PT/FEQ/APEQ/C</td>
<td>Correspondência recebida de editores e tradutores</td>
</tr>
<tr>
<td>SubPonds 2</td>
<td>2</td>
<td>PT/FEQ/FAMEQ(CR)</td>
<td>PT/FEQ/FAMEQ</td>
<td>Família Eça de Queirós (Condes de Resende)</td>
</tr>
<tr>
<td>Sections 3</td>
<td>3</td>
<td>PT/FEQ/FAMEQ/B</td>
<td>PT/FEQ/FAMEQ(CR)</td>
<td>Gestão patrimonial</td>
</tr>
<tr>
<td>SubSections 4</td>
<td>4</td>
<td>PT/FEQ/FAMEQ/B-A</td>
<td>PT/FEQ/FAMEQ/B</td>
<td>Posse e Administração de Propriedades</td>
</tr>
<tr>
<td>Level Description</td>
<td>Level Code</td>
<td>Reference Code</td>
<td>Upper Code</td>
<td>Main Title</td>
</tr>
<tr>
<td>-------------------</td>
<td>------------</td>
<td>----------------</td>
<td>------------</td>
<td>------------</td>
</tr>
<tr>
<td>Series 5</td>
<td>PT/FEQ/FAMEQ/B-A-003</td>
<td>PT/FEQ/FAMEQ/B-A</td>
<td>Apontamentos relativos a administração de propriedades</td>
<td></td>
</tr>
<tr>
<td>Series 5</td>
<td>PT/FEQ/FAMEQ/B-A-005</td>
<td>PT/FEQ/FAMEQ/B-A</td>
<td>Autos de confirmação</td>
<td></td>
</tr>
<tr>
<td>Documents 8</td>
<td>PT/FEQ/FAMEQ/B-A-005-d2</td>
<td>PT/FEQ/FAMEQ/B-A-005</td>
<td>Autos de confirmação</td>
<td></td>
</tr>
<tr>
<td>Documents 8</td>
<td>PT/FEQ/FAMEQ/B-A-005-d1</td>
<td>PT/FEQ/FAMEQ/B-A-005</td>
<td>Autos de confirmação</td>
<td></td>
</tr>
<tr>
<td>Series 5</td>
<td>PT/FEQ/FAMEQ/B-A-006</td>
<td>PT/FEQ/FAMEQ/B-A</td>
<td>Autos de penhora</td>
<td></td>
</tr>
<tr>
<td>Series 5</td>
<td>PT/FEQ/FAMEQ/B-A-011</td>
<td>PT/FEQ/FAMEQ/B-A</td>
<td>Certidões de sentenças e provisões</td>
<td></td>
</tr>
<tr>
<td>Series 5</td>
<td>PT/FEQ/FAMEQ/B-A-015</td>
<td>PT/FEQ/FAMEQ/B-A</td>
<td>Declarações de quitação</td>
<td></td>
</tr>
<tr>
<td>Documents 8</td>
<td>PT/FEQ/FAMEQ/B-A-025</td>
<td>PT/FEQ/FAMEQ/B-A</td>
<td>Escrituras de arrendamento (família)</td>
<td></td>
</tr>
<tr>
<td>Series 5</td>
<td>PT/FEQ/FAMEQ/B-A-034</td>
<td>PT/FEQ/FAMEQ/B-A</td>
<td>Índices descritivos de propriedades</td>
<td></td>
</tr>
<tr>
<td>Series 5</td>
<td>PT/FEQ/FAMEQ/B-A-036</td>
<td>PT/FEQ/FAMEQ/B-A</td>
<td>Inventários de bens e direitos</td>
<td></td>
</tr>
<tr>
<td>Series 5</td>
<td>PT/FEQ/FAMEQ/B-A-040</td>
<td>PT/FEQ/FAMEQ/B-A</td>
<td>Petições judiciais</td>
<td></td>
</tr>
<tr>
<td>Series 5</td>
<td>PT/FEQ/FAMEQ/B-A-049</td>
<td>PT/FEQ/FAMEQ/B-A</td>
<td>Procurações (1525-1879)</td>
<td></td>
</tr>
<tr>
<td>Series 5</td>
<td>PT/FEQ/FAMEQ/B-A-056</td>
<td>PT/FEQ/FAMEQ/B-A</td>
<td>Títulos particulares de contratos de arrendamento (1663)</td>
<td></td>
</tr>
<tr>
<td>Series 5</td>
<td>PT/FEQ/FAMEQ/B-A-044</td>
<td>PT/FEQ/FAMEQ/B-A</td>
<td>Processos judiciais</td>
<td></td>
</tr>
<tr>
<td>Series 5</td>
<td>PT/FEQ/FAMEQ/B-A-001</td>
<td>PT/FEQ/FAMEQ/B-A</td>
<td>Acções ordinárias (para execução de bens)</td>
<td></td>
</tr>
<tr>
<td>Series 5</td>
<td>PT/FEQ/FAMEQ/B-A-008</td>
<td>PT/FEQ/FAMEQ/B-A</td>
<td>Autos de provisão</td>
<td></td>
</tr>
<tr>
<td>Series 5</td>
<td>PT/FEQ/FAMEQ/B-A-014</td>
<td>PT/FEQ/FAMEQ/B-A</td>
<td>Declarações de obrigação</td>
<td></td>
</tr>
<tr>
<td>Documents 8</td>
<td>PT/FEQ/FAMEQ/B-A-027-d1</td>
<td>PT/FEQ/FAMEQ/B-A-027</td>
<td>Escrituras de composição amigável</td>
<td></td>
</tr>
<tr>
<td>Documents 8</td>
<td>PT/FEQ/FAMEQ/B-A-027-d2</td>
<td>PT/FEQ/FAMEQ/B-A-027</td>
<td>Escrituras de composição amigável</td>
<td></td>
</tr>
<tr>
<td>Series 5</td>
<td>PT/FEQ/FAMEQ/B-A-027</td>
<td>PT/FEQ/FAMEQ/B-A</td>
<td>Medidações de propriedades</td>
<td></td>
</tr>
<tr>
<td>Series 5</td>
<td>PT/FEQ/FAMEQ/B-A-038</td>
<td>PT/FEQ/FAMEQ/B-A</td>
<td>Processos particulares de administração de propriedades</td>
<td></td>
</tr>
<tr>
<td>Series 5</td>
<td>PT/FEQ/FAMEQ/B-A-042</td>
<td>PT/FEQ/FAMEQ/B-A</td>
<td>Autos de vedorias e apegações</td>
<td></td>
</tr>
<tr>
<td>Series 5</td>
<td>PT/FEQ/FAMEQ/B-A-009</td>
<td>PT/FEQ/FAMEQ/B-A</td>
<td>Pareceres jurídicos</td>
<td></td>
</tr>
<tr>
<td>Series 5</td>
<td>PT/FEQ/FAMEQ/B-A-039</td>
<td>PT/FEQ/FAMEQ/B-A</td>
<td>Sentenças cíveis relativas a posse de bens</td>
<td></td>
</tr>
<tr>
<td>SubSections 4</td>
<td>PT/FEQ/FAMEQ/B-B</td>
<td>PT/FEQ/FAMEQ/B</td>
<td>Transação e transmissão de</td>
<td></td>
</tr>
</tbody>
</table>

hierarchy
<table>
<thead>
<tr>
<th>Level Description</th>
<th>Level Code</th>
<th>Reference Code</th>
<th>Upper Code</th>
<th>Main Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/B-B-010</td>
<td>PT/FEQ/FAMEQ/B-B</td>
<td>Certidões de documentos relativos a transacções patrimoniais</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/B-B-024</td>
<td>PT/FEQ/FAMEQ/B-B</td>
<td>Escrituras ante-nupciais</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/B-B-026</td>
<td>PT/FEQ/FAMEQ/B-B</td>
<td>Escrituras de autenticação</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/B-B-029</td>
<td>PT/FEQ/FAMEQ/B-B</td>
<td>Escrituras de doação</td>
</tr>
<tr>
<td>Documents</td>
<td>8</td>
<td>PT/FEQ/FAMEQ/B-B-029-dl</td>
<td>PT/FEQ/FAMEQ/B-B-029</td>
<td>Escrituras de doação</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/B-B-037</td>
<td>PT/FEQ/FAMEQ/B-B</td>
<td>Mandados judiciais</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/B-B-041</td>
<td>PT/FEQ/FAMEQ/B-B</td>
<td>Processos de arrematação e trepasso de foros</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/B-B-046</td>
<td>PT/FEQ/FAMEQ/B-B</td>
<td>Processos particulares relativos a inventários e partilhas</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/B-B-054</td>
<td>PT/FEQ/FAMEQ/B-B</td>
<td>Testamentos</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/B-B-055</td>
<td>PT/FEQ/FAMEQ/B-B</td>
<td>Títulos particulares de contratos de compra e venda</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/B-B-057</td>
<td>PT/FEQ/FAMEQ/B-B</td>
<td>Títulos particulares de contratos de permuta</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/B-B-058</td>
<td>PT/FEQ/FAMEQ/B-B</td>
<td>Títulos particulares de repartição de bens e direitos</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/B-B-059</td>
<td>PT/FEQ/FAMEQ/B-B</td>
<td>Autos de posse</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/B-B-062</td>
<td>PT/FEQ/FAMEQ/B-B</td>
<td>Escrituras de compra e venda</td>
</tr>
<tr>
<td>Documents</td>
<td>8</td>
<td>PT/FEQ/FAMEQ/B-B-028-d3</td>
<td>PT/FEQ/FAMEQ/B-B-028</td>
<td>Escrituras de compra e venda</td>
</tr>
<tr>
<td>Documents</td>
<td>8</td>
<td>PT/FEQ/FAMEQ/B-B-028-d11</td>
<td>PT/FEQ/FAMEQ/B-B-028</td>
<td>Escrituras de compra e venda</td>
</tr>
<tr>
<td>Documents</td>
<td>8</td>
<td>PT/FEQ/FAMEQ/B-B-028-d1</td>
<td>PT/FEQ/FAMEQ/B-B-028</td>
<td>Escrituras de compra e venda</td>
</tr>
<tr>
<td>Documents</td>
<td>8</td>
<td>PT/FEQ/FAMEQ/B-B-028-d2</td>
<td>PT/FEQ/FAMEQ/B-B-028</td>
<td>Escrituras de compra e venda</td>
</tr>
<tr>
<td>Documents</td>
<td>8</td>
<td>PT/FEQ/FAMEQ/B-B-028-d4</td>
<td>PT/FEQ/FAMEQ/B-B-028</td>
<td>Escrituras de compra e venda</td>
</tr>
<tr>
<td>Documents</td>
<td>8</td>
<td>PT/FEQ/FAMEQ/B-B-028-d7</td>
<td>PT/FEQ/FAMEQ/B-B-028</td>
<td>Escrituras de compra e venda</td>
</tr>
<tr>
<td>Documents</td>
<td>8</td>
<td>PT/FEQ/FAMEQ/B-B-028-d9</td>
<td>PT/FEQ/FAMEQ/B-B-028</td>
<td>Escrituras de compra e venda</td>
</tr>
<tr>
<td>Documents</td>
<td>8</td>
<td>PT/FEQ/FAMEQ/B-B-028-d8</td>
<td>PT/FEQ/FAMEQ/B-B-028</td>
<td>Escrituras de compra e venda</td>
</tr>
<tr>
<td>Documents</td>
<td>8</td>
<td>PT/FEQ/FAMEQ/B-B-028-d5</td>
<td>PT/FEQ/FAMEQ/B-B-028</td>
<td>Escrituras de compra e venda</td>
</tr>
<tr>
<td>Documents</td>
<td>8</td>
<td>PT/FEQ/FAMEQ/B-B-028-d6</td>
<td>PT/FEQ/FAMEQ/B-B-028</td>
<td>Escrituras de compra e venda</td>
</tr>
<tr>
<td>Documents</td>
<td>8</td>
<td>PT/FEQ/FAMEQ/B-B-028-d10</td>
<td>PT/FEQ/FAMEQ/B-B-028</td>
<td>Escrituras de compra e venda</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/B-B-032</td>
<td>PT/FEQ/FAMEQ/B-B</td>
<td>Escrituras de permuta</td>
</tr>
<tr>
<td>Documents</td>
<td>8</td>
<td>PT/FEQ/FAMEQ/B-B-</td>
<td>PT/FEQ/FAMEQ/B-B</td>
<td>Escrituras de permuta</td>
</tr>
</tbody>
</table>

**Hierarchy**
### Query on hierarchy of Archive de Eça

<table>
<thead>
<tr>
<th>Level Description</th>
<th>Level Code</th>
<th>Reference Code</th>
<th>Upper Code</th>
<th>Main Title</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Series</strong></td>
<td>5</td>
<td>PT/FEQ/FAMEQ/B-B-032-d1</td>
<td>032</td>
<td>Processos de transacções patrimoniais</td>
</tr>
<tr>
<td><strong>Series</strong></td>
<td>5</td>
<td>PT/FEQ/FAMEQ/B-B-053</td>
<td>PT/FEQ/FAMEQ/B-B</td>
<td>Sentenças cíveis relativas a transmissão de bens</td>
</tr>
<tr>
<td><strong>Documents</strong></td>
<td>8</td>
<td>PT/FEQ/FAMEQ/B-B-053-d1</td>
<td>PT/FEQ/FAMEQ/B-B-053</td>
<td>Sentenças cíveis relativas a transmissão de bens</td>
</tr>
<tr>
<td><strong>Series</strong></td>
<td>5</td>
<td>PT/FEQ/FAMEQ/B-B-004</td>
<td>PT/FEQ/FAMEQ/B-B</td>
<td>Autos de arrematação</td>
</tr>
<tr>
<td><strong>SubSections</strong></td>
<td>4</td>
<td>PT/FEQ/FAMEQ/B-C-002</td>
<td>PT/FEQ/FAMEQ/B</td>
<td>Administração do espólio literário de Eça de Queirós</td>
</tr>
<tr>
<td><strong>Series</strong></td>
<td>5</td>
<td>PT/FEQ/FAMEQ/B-C-002</td>
<td>PT/FEQ/FAMEQ/B-C</td>
<td>Adaptações de obras de Eça de Queirós (teatro e rádio)</td>
</tr>
<tr>
<td><strong>Series</strong></td>
<td>5</td>
<td>PT/FEQ/FAMEQ/B-C-013</td>
<td>PT/FEQ/FAMEQ/B-C</td>
<td>Conta corrente (com editores)</td>
</tr>
<tr>
<td><strong>Series</strong></td>
<td>5</td>
<td>PT/FEQ/FAMEQ/B-C-030</td>
<td>PT/FEQ/FAMEQ/B-C</td>
<td>Pareceres jurídicos (Marcelo Caetano)</td>
</tr>
<tr>
<td><strong>Series</strong></td>
<td>5</td>
<td>PT/FEQ/FAMEQ/B-C-059</td>
<td>PT/FEQ/FAMEQ/B-C</td>
<td>Transcrições de cartas, obras literárias e documentos de Eça</td>
</tr>
<tr>
<td><strong>Sections</strong></td>
<td>3</td>
<td>PT/FEQ/FAMEQ/C-023</td>
<td>PT/FEQ/FAMEQ/C-023</td>
<td>Gestão financeira</td>
</tr>
<tr>
<td><strong>Series</strong></td>
<td>5</td>
<td>PT/FEQ/FAMEQ/C-023</td>
<td>PT/FEQ/FAMEQ/C</td>
<td>Documentos relativos ao enterro e trasladação de Eça de Queirós</td>
</tr>
<tr>
<td><strong>SubSections</strong></td>
<td>4</td>
<td>PT/FEQ/FAMEQ/C-A-016</td>
<td>PT/FEQ/FAMEQ/C-A</td>
<td>Conta corrente (1919)</td>
</tr>
<tr>
<td><strong>Series</strong></td>
<td>5</td>
<td>PT/FEQ/FAMEQ/C-A-016</td>
<td>PT/FEQ/FAMEQ/C-A</td>
<td>Declarações para pagamento de aluguel</td>
</tr>
<tr>
<td><strong>Series</strong></td>
<td>5</td>
<td>PT/FEQ/FAMEQ/C-A-017</td>
<td>PT/FEQ/FAMEQ/C-A</td>
<td>Diário</td>
</tr>
<tr>
<td><strong>Series</strong></td>
<td>5</td>
<td>PT/FEQ/FAMEQ/C-A-017</td>
<td>PT/FEQ/FAMEQ/C-A</td>
<td>Recibo</td>
</tr>
<tr>
<td><strong>Series</strong></td>
<td>5</td>
<td>PT/FEQ/FAMEQ/C-A-018</td>
<td>PT/FEQ/FAMEQ/C-A</td>
<td>Documentos de despesa (1850c.-1945c.)</td>
</tr>
<tr>
<td><strong>SubSections</strong></td>
<td>4</td>
<td>PT/FEQ/FAMEQ/C-B-012</td>
<td>PT/FEQ/FAMEQ/C-B</td>
<td>Empréstimo de Capitais</td>
</tr>
<tr>
<td><strong>Series</strong></td>
<td>5</td>
<td>PT/FEQ/FAMEQ/C-B-012</td>
<td>PT/FEQ/FAMEQ/C-B</td>
<td>Certificados de hipotecas</td>
</tr>
<tr>
<td><strong>Series</strong></td>
<td>5</td>
<td>PT/FEQ/FAMEQ/C-B-030</td>
<td>PT/FEQ/FAMEQ/C-B</td>
<td>Escrituras de empréstimo</td>
</tr>
<tr>
<td><strong>Series</strong></td>
<td>5</td>
<td>PT/FEQ/FAMEQ/C-B-030</td>
<td>PT/FEQ/FAMEQ/C-B</td>
<td>Escrituras de hipotecas</td>
</tr>
<tr>
<td><strong>Series</strong></td>
<td>5</td>
<td>PT/FEQ/FAMEQ/C-B-031</td>
<td>PT/FEQ/FAMEQ/C-B</td>
<td>Escrituras de quitação</td>
</tr>
<tr>
<td><strong>Series</strong></td>
<td>5</td>
<td>PT/FEQ/FAMEQ/C-B-031</td>
<td>PT/FEQ/FAMEQ/C-B</td>
<td>Escrituras de quitação</td>
</tr>
<tr>
<td><strong>Series</strong></td>
<td>5</td>
<td>PT/FEQ/FAMEQ/C-B-049</td>
<td>PT/FEQ/FAMEQ/C-B</td>
<td>Processos relativos a pedidos de empréstimo</td>
</tr>
<tr>
<td><strong>Series</strong></td>
<td>5</td>
<td>PT/FEQ/FAMEQ/C-B-051</td>
<td>PT/FEQ/FAMEQ/C-B</td>
<td>Registos de receita e despesa</td>
</tr>
<tr>
<td><strong>Sections</strong></td>
<td>3</td>
<td>PT/FEQ/FAMEQ/D</td>
<td>PT/FEQ/FAMEQ/D-CR</td>
<td>Actividades individuais</td>
</tr>
<tr>
<td><strong>SubSections</strong></td>
<td>4</td>
<td>PT/FEQ/FAMEQ/D-A</td>
<td>PT/FEQ/FAMEQ/D-A</td>
<td>Gaspar de Figueiroa</td>
</tr>
<tr>
<td><strong>Series</strong></td>
<td>5</td>
<td>PT/FEQ/FAMEQ/D-A</td>
<td>PT/FEQ/FAMEQ/D-A</td>
<td>Autos de nomeação (Gaspar de Figueiroa)</td>
</tr>
<tr>
<td><strong>SubSections</strong></td>
<td>4</td>
<td>PT/FEQ/FAMEQ/D-B</td>
<td>PT/FEQ/FAMEQ/D-B</td>
<td>Duarte Carneiro Rangel</td>
</tr>
<tr>
<td><strong>Series</strong></td>
<td>5</td>
<td>PT/FEQ/FAMEQ/D-B</td>
<td>PT/FEQ/FAMEQ/D-B</td>
<td>Dispensas eclesiásticas</td>
</tr>
<tr>
<td>Level Description</td>
<td>Level Code</td>
<td>Reference Code</td>
<td>Upper Code</td>
<td>Main Title</td>
</tr>
<tr>
<td>-------------------</td>
<td>------------</td>
<td>----------------</td>
<td>------------</td>
<td>------------</td>
</tr>
<tr>
<td>SubSections</td>
<td>4</td>
<td>PT/FEQ/FAMEQ/D-E</td>
<td>PT/FEQ/FAMEQ/D</td>
<td>Tomé da Silva Baldaia</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/D-E-064</td>
<td>PT/FEQ/FAMEQ/D-E</td>
<td>Certificados de Tomé da Silva Baldaia</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/D-E-047</td>
<td>PT/FEQ/FAMEQ/D-E</td>
<td>Processos de petição de Tomé da Silva Baldaia</td>
</tr>
<tr>
<td>Composed Documents</td>
<td>7</td>
<td>PT/FEQ/FAMEQ/D-E-047-d1</td>
<td>PT/FEQ/FAMEQ/D-E-047</td>
<td>Processos de petição de Tomé da Silva Baldaia</td>
</tr>
<tr>
<td>SubSections</td>
<td>4</td>
<td>PT/FEQ/FAMEQ/D-D</td>
<td>PT/FEQ/FAMEQ/D</td>
<td>Francisco Carneiro de Figueiroa</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/D-D-045</td>
<td>PT/FEQ/FAMEQ/D-D</td>
<td>Processos de nomeação</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/D-D-064</td>
<td>PT/FEQ/FAMEQ/D-D</td>
<td>Certificados</td>
</tr>
<tr>
<td>SubSections</td>
<td>4</td>
<td>PT/FEQ/FAMEQ/D-F</td>
<td>PT/FEQ/FAMEQ/D-F</td>
<td>João de Figueiroa Pinto</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/D-F-061</td>
<td>PT/FEQ/FAMEQ/D-F</td>
<td>Autos de legitimação</td>
</tr>
<tr>
<td>Documents</td>
<td>8</td>
<td>PT/FEQ/FAMEQ/D-F-061-d1</td>
<td>PT/FEQ/FAMEQ/D-F-061</td>
<td>Autos de legitimação</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/D-F-064</td>
<td>PT/FEQ/FAMEQ/D-F</td>
<td>Certificados de João de Figueiroa Pinto</td>
</tr>
<tr>
<td>SubSections</td>
<td>4</td>
<td>PT/FEQ/FAMEQ/D-G</td>
<td>PT/FEQ/FAMEQ/D-G</td>
<td>Francisco Carneiro de Figueiroa (2)</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/D-G-065</td>
<td>PT/FEQ/FAMEQ/D-G</td>
<td>Correspondência recebida por Francisco Carneiro de Figueiroa (2)</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/D-I-073</td>
<td>PT/FEQ/FAMEQ/D-I</td>
<td>Certidões (1817)</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/D-I-071</td>
<td>PT/FEQ/FAMEQ/D-I</td>
<td>Inquéritos judiciais</td>
</tr>
<tr>
<td>SubSections</td>
<td>4</td>
<td>PT/FEQ/FAMEQ/D-K</td>
<td>PT/FEQ/FAMEQ/D-K</td>
<td>Documentos de despesa (1890)</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/D-L-018</td>
<td>PT/FEQ/FAMEQ/D-L</td>
<td>Emília Eça de Queirós</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/D-L-070</td>
<td>PT/FEQ/FAMEQ/D-L</td>
<td>Passaportes (1900)</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/D-L-073</td>
<td>PT/FEQ/FAMEQ/D-L</td>
<td>Títulos de reforma</td>
</tr>
<tr>
<td>Level Description</td>
<td>Level Code</td>
<td>Reference Code</td>
<td>Upper Code</td>
<td>Main Title</td>
</tr>
<tr>
<td>------------------</td>
<td>------------</td>
<td>----------------</td>
<td>------------</td>
<td>------------</td>
</tr>
<tr>
<td>SubSections</td>
<td>4</td>
<td>PT/FEQ/FAMEQ/D-O-065</td>
<td>PT/FEQ/FAMEQ/D-O</td>
<td>José Maria Eça de Queirós</td>
</tr>
<tr>
<td>SubSections</td>
<td>4</td>
<td>PT/FEQ/FAMEQ/D-P</td>
<td>PT/FEQ/FAMEQ/D-P</td>
<td>Documentos de avaliação pedagógica de José Maria Eça de Queirós</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/D-P-021</td>
<td>PT/FEQ/FAMEQ/D-P</td>
<td>Alberto Eça de Queirós</td>
</tr>
<tr>
<td>SubSections</td>
<td>4</td>
<td>PT/FEQ/FAMEQ/D-P-065</td>
<td>PT/FEQ/FAMEQ/D-P</td>
<td>Documentos de avaliação pedagógica de Alberto Eça de Queirós</td>
</tr>
<tr>
<td>SubSections</td>
<td>4</td>
<td>PT/FEQ/FAMEQ/D-Q-065</td>
<td>PT/FEQ/FAMEQ/D-Q</td>
<td>Matilde Eça de Queirós</td>
</tr>
<tr>
<td>SubSections</td>
<td>4</td>
<td>PT/FEQ/FAMEQ/D-C-047</td>
<td>PT/FEQ/FAMEQ/D-C</td>
<td>Francisco Neto de Piqueirao</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/D-C-063</td>
<td>PT/FEQ/FAMEQ/D-C</td>
<td>Processos de petição de Francisco Neto de Piqueirao</td>
</tr>
<tr>
<td>SubSections</td>
<td>4</td>
<td>PT/FEQ/FAMEQ/D-H-063</td>
<td>PT/FEQ/FAMEQ/D-H</td>
<td>Correspondência recebida por Francisco Neto de Piqueirao</td>
</tr>
<tr>
<td>SubSections</td>
<td>4</td>
<td>PT/FEQ/FAMEQ/D-N</td>
<td>PT/FEQ/FAMEQ/D-N</td>
<td>Certidões (1877)</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/D-N-066</td>
<td>PT/FEQ/FAMEQ/D-N</td>
<td>Maria Eça de Queirós</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/D-N-067</td>
<td>PT/FEQ/FAMEQ/D-N</td>
<td>Cursos de literatura</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/D-N-066</td>
<td>PT/FEQ/FAMEQ/D-N</td>
<td>Diplomas</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/D-N-065</td>
<td>PT/FEQ/FAMEQ/D-N</td>
<td>Correspondência relativa a gestão patrimonial</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/D-N-065-065</td>
<td>PT/FEQ/FAMEQ/D-N</td>
<td>Apontamentos e transcrições de cartas, documentos e artigos</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/D-N-065-01</td>
<td>PT/FEQ/FAMEQ/D-N</td>
<td>Diversos (miscelânea)</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/D-N-056</td>
<td>PT/FEQ/FAMEQ/D-N</td>
<td>Discursos e conferências de Maria Eça de Queirós</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/D-N-072</td>
<td>PT/FEQ/FAMEQ/D-N</td>
<td>Adaptações de obras de Eça de Queirós</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/D-N-068</td>
<td>PT/FEQ/FAMEQ/D-N</td>
<td>Documentos de produção literária de Maria Eça de Queirós</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/D-N-022</td>
<td>PT/FEQ/FAMEQ/D-N</td>
<td></td>
</tr>
<tr>
<td>Level Description</td>
<td>Level Code</td>
<td>Reference Code</td>
<td>Upper Code</td>
<td>Main Title</td>
</tr>
<tr>
<td>------------------</td>
<td>------------</td>
<td>-------------------------</td>
<td>---------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/D-N-021</td>
<td>PT/FEQ/FAMEQ/D-N</td>
<td>Documentos de avaliação pedagógica de Maria Eça de Queirós</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/D-N-074</td>
<td>PT/FEQ/FAMEQ/D-N</td>
<td>Processos particulares para deslocação ao estrangeiro</td>
</tr>
<tr>
<td>Sections</td>
<td>3</td>
<td>PT/FEQ/FAMEQ/3</td>
<td>PT/FEQ/FAMEQ (CR)</td>
<td>Coleção periódicos e recortes de Jornais</td>
</tr>
<tr>
<td>Sections</td>
<td>3</td>
<td>PT/FEQ/FAMEQ/A</td>
<td>PT/FEQ/FAMEQ (CR)</td>
<td>Constituição e regulamentação</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/A-036</td>
<td>PT/FEQ/FAMEQ/A</td>
<td>Inventários de documentos</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ/A-019</td>
<td>PT/FEQ/FAMEQ/A</td>
<td>Documentos genealógicos</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/FAMEQ-019</td>
<td>PT/FEQ/FAMEQ (CR)</td>
<td>Documentos diversos (1867-1968)</td>
</tr>
<tr>
<td>SubPonds</td>
<td>2</td>
<td>PT/FEQ/APARQ</td>
<td>PT/FEQ</td>
<td>Arquivo Pessoal de António Eça de Queirós</td>
</tr>
<tr>
<td>Sections</td>
<td>3</td>
<td>PT/FEQ/APARQ/A</td>
<td>PT/FEQ/APARQ</td>
<td>Actividade privada (António EÇQ)</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/APARQ/A-007</td>
<td>PT/FEQ/APARQ/A</td>
<td>Correspondência recebida (1909-1965)</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/APARQ/A-009</td>
<td>PT/FEQ/APARQ/A</td>
<td>Documentos de avaliação pedagógica de António Eça de Queirós</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/APARQ/A-012</td>
<td>PT/FEQ/APARQ/A</td>
<td>Documentos relativos à Legião Portuguesana</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/APARQ/A-018</td>
<td>PT/FEQ/APARQ/A</td>
<td>Dossiers relativos ao &quot;Clube de Eça&quot;</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/APARQ/A-023</td>
<td>PT/FEQ/APARQ/A</td>
<td>Processos relativos a edições de obras de Eça</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/APARQ/A-024</td>
<td>PT/FEQ/APARQ/A</td>
<td>Processos relativos a esferéisias commemorativas</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/APARQ/A-026</td>
<td>PT/FEQ/APARQ/A</td>
<td>Desenhos e esboços</td>
</tr>
<tr>
<td>SubSections</td>
<td>4</td>
<td>PT/FEQ/APARQ/A-021</td>
<td>PT/FEQ/APARQ/A</td>
<td>Serviços Aéreos Portugueses (SAP)</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/APARQ/A-021</td>
<td>PT/FEQ/APARQ/A</td>
<td>Processos de correspondência (SAP)</td>
</tr>
<tr>
<td>SubSections</td>
<td>4</td>
<td>PT/FEQ/APARQ/A-007</td>
<td>PT/FEQ/APARQ/A</td>
<td>Companhia Agrícola e Fabril da Guiné (CAGF)</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/APARQ/A-013</td>
<td>PT/FEQ/APARQ/A</td>
<td>Dossiers administrativos</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/APARQ/A-002</td>
<td>PT/FEQ/APARQ/A</td>
<td>Apontamentos auto-biográficos</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/APARQ/A-006</td>
<td>PT/FEQ/APARQ/A</td>
<td>Copiador 1</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/APARQ/A-010</td>
<td>PT/FEQ/APARQ/A</td>
<td>Documentos de despesa privados</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/APARQ/A-011</td>
<td>PT/FEQ/APARQ/A</td>
<td>Documentos de produção literária de António</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/APARQ/A-019</td>
<td>PT/FEQ/APARQ/A</td>
<td>Documentos sobre eleições</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/APARQ/A-002</td>
<td>PT/FEQ/APARQ/A</td>
<td>Processos de correspondência</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/APARQ/A-004</td>
<td>PT/FEQ/APARQ/A</td>
<td>Artigos de opinião</td>
</tr>
<tr>
<td>Sections</td>
<td>3</td>
<td>PT/FEQ/APARQ/B</td>
<td>PT/FEQ/APARQ</td>
<td>Actividade pública (António EÇQ)</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/APARQ/B-006</td>
<td>PT/FEQ/APARQ</td>
<td>Discursos e conferências de</td>
</tr>
<tr>
<td>Level Description</td>
<td>Level Code</td>
<td>Reference Code</td>
<td>Upper Code</td>
<td>Main Title</td>
</tr>
<tr>
<td>-------------------</td>
<td>------------</td>
<td>-----------------</td>
<td>------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>SubSections</td>
<td>4</td>
<td>PT/FEQ/APABQ/B-A</td>
<td>PT/FEQ/APABQ</td>
<td>António Eça de Queirós</td>
</tr>
<tr>
<td>Series</td>
<td>5</td>
<td>PT/FEQ/APABQ/B-A</td>
<td>PT/FEQ/APABQ/B</td>
<td>Secretariado Nacional de Informação (SNI)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>005</td>
<td>007</td>
<td>Compilações de recortes de jornais</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PT/FEQ/APABQ/B-A</td>
<td>PT/FEQ/APABQ/B</td>
<td>Correspondência recebida (SNI)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>015</td>
<td>017</td>
<td>Dossiers relativos a imprensa</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PT/FEQ/APABQ/B-A</td>
<td>PT/FEQ/APABQ/B</td>
<td>Processos de correspondência (oficial e particular)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>021</td>
<td>014</td>
<td>Dossiers relativos a feiras e exposições internacionais</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PT/FEQ/APABQ/B-A</td>
<td>PT/FEQ/APABQ/B</td>
<td>Dossiers relativos a viagens oficiais</td>
</tr>
<tr>
<td></td>
<td></td>
<td>017</td>
<td>017</td>
<td>Relatórios de viagens oficiais</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PT/FEQ/APABQ/B-A</td>
<td>PT/FEQ/APABQ/B</td>
<td>Emisora Nacional (EN)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>025</td>
<td>020</td>
<td>Processos de direcção</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PT/FEQ/APABQ/B-A</td>
<td>PT/FEQ/APABQ/B</td>
<td>Processos de correspondência (particular e oficial)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>021</td>
<td>016</td>
<td>Dossiers relativos a regatas</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PT/FEQ/APABQ/B-A</td>
<td>PT/FEQ/APABQ/B</td>
<td>Adaptações de obras de Eça de Queirós 1951</td>
</tr>
<tr>
<td></td>
<td></td>
<td>001</td>
<td>003</td>
<td>Apontamentos sobre a reforma da Emisora Nacional</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PT/FEQ/APABQ/B-A</td>
<td>PT/FEQ/APABQ/B</td>
<td>Copiador 2</td>
</tr>
<tr>
<td>Sections</td>
<td>3</td>
<td>PT/FEQ/APABQ/C</td>
<td>PT/FEQ/APABQ</td>
<td>Coleção/periódicos, recortes de jornais, publ. Oficiais</td>
</tr>
</tbody>
</table>
Appendix D
<table>
<thead>
<tr>
<th>Segment Code</th>
<th>Upper Code</th>
<th>Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td></td>
<td>Program</td>
<td>Building of the Akashi-channel bridge</td>
</tr>
<tr>
<td>1.01</td>
<td>100</td>
<td>Scene-group</td>
<td>Introduction of the Akashi-channel bridge</td>
</tr>
<tr>
<td>1.02</td>
<td>101</td>
<td>Scene</td>
<td>Introduction of the Akashi-channel bridge</td>
</tr>
<tr>
<td>1.03</td>
<td>102</td>
<td>Cut</td>
<td>Air-shot of the Akashi-channel bridge</td>
</tr>
<tr>
<td>1.159</td>
<td>103</td>
<td>Effect Super</td>
<td>Effect: super; string &quot;The Akashi-Channel Bridge&quot;</td>
</tr>
<tr>
<td>1.04</td>
<td>102</td>
<td>Cut</td>
<td>Fanning-shot of the bridge</td>
</tr>
<tr>
<td>1.05</td>
<td>102</td>
<td>Cut</td>
<td>Dolly-shot of cables of the Akashi-channel bridge</td>
</tr>
<tr>
<td>1.06</td>
<td>102</td>
<td>Cut</td>
<td>Tinas report</td>
</tr>
<tr>
<td>1.07</td>
<td>102</td>
<td>Effect DVE</td>
<td>Effect: DVE</td>
</tr>
<tr>
<td>1.08</td>
<td>102</td>
<td>Cut</td>
<td>Close-up of a cross section of the Akashi-channel bridges cable</td>
</tr>
<tr>
<td>1.09</td>
<td>100</td>
<td>Scene-group</td>
<td>Effect: super; string &quot;Cross Section of the Cable&quot;</td>
</tr>
<tr>
<td>1.10</td>
<td>109</td>
<td>Scene</td>
<td>Cross section of the Akashi-channel bridges cable</td>
</tr>
<tr>
<td>1.11</td>
<td>110</td>
<td>Cut</td>
<td>National Research Institute for Metals</td>
</tr>
<tr>
<td>1.12</td>
<td>111</td>
<td>Effect Super</td>
<td>National Research Institute for Metals in Tokuba city</td>
</tr>
<tr>
<td>1.13</td>
<td>110</td>
<td>Cut</td>
<td>Effect: super; string &quot;National Research Institute for Metals&quot;</td>
</tr>
<tr>
<td>1.14</td>
<td>110</td>
<td>Cut</td>
<td>Zooming-in to Dr. Toshihiko Takahashi who is an inventor of the Ultra-steel</td>
</tr>
<tr>
<td>1.15</td>
<td>109</td>
<td>Scene</td>
<td>Effect: super; string &quot;Dr. Toshihiko Takahashi&quot;</td>
</tr>
<tr>
<td>1.16</td>
<td>115</td>
<td>Cut</td>
<td>Dr. Takahashi explains display monitors.</td>
</tr>
<tr>
<td>1.17</td>
<td>115</td>
<td>Cut</td>
<td>Close-up of a display monitor</td>
</tr>
<tr>
<td>1.18</td>
<td>115</td>
<td>Cut</td>
<td>CEMENTITE makes strong steel.</td>
</tr>
<tr>
<td>1.19</td>
<td>115</td>
<td>Cut</td>
<td>Long-shot of the rolled wire</td>
</tr>
<tr>
<td>1.20</td>
<td>115</td>
<td>Effect Super</td>
<td>Close-up of the rolled wire</td>
</tr>
<tr>
<td>1.21</td>
<td>118</td>
<td>Effect Super</td>
<td>A photomicrograph of the inner structure of the wire</td>
</tr>
<tr>
<td>1.22</td>
<td>118</td>
<td>Effect Super</td>
<td>Effect: super; string &quot;Photomicrograph of The Wire&quot;</td>
</tr>
<tr>
<td>1.23</td>
<td>118</td>
<td>Effect DVE</td>
<td>Effect: super; string &quot;CEMENTITE&quot;</td>
</tr>
<tr>
<td>1.24</td>
<td>115</td>
<td>Cut</td>
<td>Animation for the explanation</td>
</tr>
<tr>
<td>1.25</td>
<td>115</td>
<td>Cut</td>
<td>A photomicrograph of broken CEMENTITE after plating.</td>
</tr>
<tr>
<td>1.26</td>
<td>115</td>
<td>Effect Super</td>
<td>Effect: super; string &quot;CEMENTITE After Plating&quot;</td>
</tr>
<tr>
<td>1.27</td>
<td>115</td>
<td>Cut</td>
<td>Two photomicrographs</td>
</tr>
<tr>
<td>1.28</td>
<td>115</td>
<td>Effect Super</td>
<td>Effect: super; string &quot;Before Plating After Plating&quot;</td>
</tr>
<tr>
<td>1.29</td>
<td>115</td>
<td>Effect DVE</td>
<td>Effect: DVE</td>
</tr>
<tr>
<td>1.30</td>
<td>109</td>
<td>Scene</td>
<td>Preventing CEMENTITE's break-down</td>
</tr>
<tr>
<td>1.31</td>
<td>122</td>
<td>Cut</td>
<td>Tina and an engineer are coming into a laboratory.</td>
</tr>
<tr>
<td>1.32</td>
<td>122</td>
<td>Cut</td>
<td>Fanning-shot of the Atom Probe</td>
</tr>
<tr>
<td>Segment Code</td>
<td>Upper Code</td>
<td>Level</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
<td>------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>167</td>
<td>124</td>
<td>Effect Super</td>
<td>Effect: super: string &quot;The Atom Probe&quot;</td>
</tr>
<tr>
<td>125</td>
<td>122</td>
<td>Cut</td>
<td>Attaching the needle</td>
</tr>
<tr>
<td>126</td>
<td>122</td>
<td>Cut</td>
<td>Panning-shot of the needle.</td>
</tr>
<tr>
<td>127</td>
<td>122</td>
<td>Cut</td>
<td>Tina looks in a display monitor of the Atom Probe.</td>
</tr>
<tr>
<td>128</td>
<td>122</td>
<td>Cut</td>
<td>Image of atoms that are jumping out from steel and colliding with the sensor.</td>
</tr>
<tr>
<td>129</td>
<td>122</td>
<td>Cut</td>
<td>Panning-shot of the Atom Probe and a computer.</td>
</tr>
<tr>
<td>130</td>
<td>122</td>
<td>Cut</td>
<td>Display monitor of the computer.</td>
</tr>
<tr>
<td>131</td>
<td>122</td>
<td>Cut</td>
<td>Image of 3D atomic structure of the wire</td>
</tr>
<tr>
<td>168</td>
<td>131</td>
<td>Effect Super</td>
<td>Effect: super: string &quot;CEMENTITE&quot;</td>
</tr>
<tr>
<td>132</td>
<td>122</td>
<td>Cut</td>
<td>Magnified 3D atomic structure of the wire</td>
</tr>
<tr>
<td>133</td>
<td>122</td>
<td>Cut</td>
<td>Panning-shot of the magnified 3D atomic structure</td>
</tr>
<tr>
<td>169</td>
<td>133</td>
<td>Effect Super</td>
<td>Effect: super: string &quot;Silicon Atoms&quot;</td>
</tr>
<tr>
<td>134</td>
<td>122</td>
<td>Cut</td>
<td>Interview of Dr. Takahashi</td>
</tr>
<tr>
<td>135</td>
<td>109</td>
<td>Scene</td>
<td>Silicon wire</td>
</tr>
<tr>
<td>136</td>
<td>135</td>
<td>Cut</td>
<td>Lumps of silicon</td>
</tr>
<tr>
<td>170</td>
<td>136</td>
<td>Effect Super</td>
<td>Effect: super: string &quot;Silicon&quot;</td>
</tr>
<tr>
<td>137</td>
<td>135</td>
<td>Cut</td>
<td>Panning-shot of lumps of silicon</td>
</tr>
<tr>
<td>138</td>
<td>135</td>
<td>Cut</td>
<td>Close-up of the silicon-blended wire</td>
</tr>
<tr>
<td>171</td>
<td>138</td>
<td>Effect Super</td>
<td>Effect: super: string &quot;Silicon-blended Wire&quot;</td>
</tr>
<tr>
<td>179</td>
<td>138</td>
<td>Effect DVE</td>
<td>Effect: DVE</td>
</tr>
<tr>
<td>139</td>
<td>109</td>
<td>Scene</td>
<td>Experiment to test strength of the wire</td>
</tr>
<tr>
<td>140</td>
<td>139</td>
<td>Cut</td>
<td>Tina tries to bent the wire by her hands.</td>
</tr>
<tr>
<td>141</td>
<td>139</td>
<td>Cut</td>
<td>Attaching the wire to the test equipment</td>
</tr>
<tr>
<td>142</td>
<td>139</td>
<td>Cut</td>
<td>Close-up of the attached wire</td>
</tr>
<tr>
<td>143</td>
<td>139</td>
<td>Cut</td>
<td>Graph on a display monitor</td>
</tr>
<tr>
<td>144</td>
<td>139</td>
<td>Cut</td>
<td>Close-up of the broken wire</td>
</tr>
<tr>
<td>172</td>
<td>144</td>
<td>Effect Super</td>
<td>Effect: super: string &quot;Conventional Wire&quot;</td>
</tr>
<tr>
<td>145</td>
<td>139</td>
<td>Cut</td>
<td>Close-up of the testing wire</td>
</tr>
<tr>
<td>173</td>
<td>145</td>
<td>Effect Super</td>
<td>Effect: super: string &quot;Silicon-blended Wire&quot;</td>
</tr>
<tr>
<td>180</td>
<td>145</td>
<td>Effect DVE</td>
<td>Effect: DVE</td>
</tr>
<tr>
<td>146</td>
<td>139</td>
<td>Cut</td>
<td>Close-up of the testing wire</td>
</tr>
<tr>
<td>174</td>
<td>146</td>
<td>Effect Super</td>
<td>Effect: super: string &quot;Silicon-blended Wire&quot;</td>
</tr>
<tr>
<td>181</td>
<td>146</td>
<td>Effect DVE</td>
<td>Effect: DVE</td>
</tr>
<tr>
<td>147</td>
<td>139</td>
<td>Cut</td>
<td>Close-up of the testing wire</td>
</tr>
<tr>
<td>148</td>
<td>100</td>
<td>Scene-group</td>
<td>Manufacturing process of the wire</td>
</tr>
<tr>
<td>149</td>
<td>148</td>
<td>Scene</td>
<td>Manufacturing process of the wire</td>
</tr>
<tr>
<td>150</td>
<td>149</td>
<td>Cut</td>
<td>Blast furnace</td>
</tr>
<tr>
<td>151</td>
<td>149</td>
<td>Cut</td>
<td>Cooling the wire</td>
</tr>
<tr>
<td>152</td>
<td>149</td>
<td>Cut</td>
<td>Ring shaped wire</td>
</tr>
<tr>
<td>153</td>
<td>149</td>
<td>Cut</td>
<td>Cooled wire</td>
</tr>
<tr>
<td>154</td>
<td>100</td>
<td>Scene-group</td>
<td>Conclusion</td>
</tr>
<tr>
<td>155</td>
<td>154</td>
<td>Scene</td>
<td>The Akashi-channel bridge</td>
</tr>
<tr>
<td>156</td>
<td>155</td>
<td>Cut</td>
<td>Air-shot of the Akashi-channel bridge</td>
</tr>
</tbody>
</table>

**segment hierarchy**
Query on the hierarchy in the segments

<table>
<thead>
<tr>
<th>Segment Code</th>
<th>Upper Code</th>
<th>Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>182</td>
<td>156</td>
<td>Effect Dissolve</td>
<td>Effect: dissolve</td>
</tr>
<tr>
<td>157</td>
<td>155</td>
<td>Cut</td>
<td>Air/Long-shot of the Akashi-channel bridge</td>
</tr>
<tr>
<td>158</td>
<td>155</td>
<td>Cut</td>
<td>Interview of Dr. Takahashi</td>
</tr>
<tr>
<td>175</td>
<td>158</td>
<td>Effect Super</td>
<td>Effect: super: string &quot;Dr. Toshihiko Takahashi, National Research Institute for Metals&quot;</td>
</tr>
</tbody>
</table>
Appendix E
Du Reference Code: MPEG/CON/V/V10
Main Title: Ultra-steel suspends the longest bridge in the world

Segment: Segment
Code: 115  Level: Scene
Starttime: 00:01:51:26  Stoptime: 00:03:20:25

Segment Description: CEMENTITE makes strong steel.
Audio-objectitem:

Du Reference Code: MPEG/CON/V/V10
Main Title: Ultra-steel suspends the longest bridge in the world

Segment: Segment
Code: 118  Level: Cut
Starttime: 00:02:04:10  Stoptime: 00:02:32:24

Segment Description: A photomicrograph of the inner structure of the wire
Steel is made by mixing iron with carbon. A tough stratum that is a combination of iron and carbon is called CEMENTITE. Stable CEMENTITE makes the wire strong.

Audio-objectitem:

Du Reference Code: MPEG/CON/V/V10
Main Title: Ultra-steel suspends the longest bridge in the world

Segment: Segment
Code: 119  Level: Cut
Starttime: 00:02:32:25  Stoptime: 00:02:53:26

Segment Description: Animation for the explanation
Audio-objectitem: CEMENTITE is weak in tolerance to heat. CEMENTITE is broken by heat in the process of plating.

Du Reference Code: MPEG/CON/V/V10
Main Title: Ultra-steel suspends the longest bridge in the world

Segment: Segment
Code: 120  Level: Cut
Starttime: 00:02:53:27  Stoptime: 00:03:00:25

Segment Description: A photomicrograph of broken CEMENTITE after plating.
Query on segment content

Du Reference Code : MPEG/CON/V/V10
Main Title : Ultra-steel suspends the longest bridge in the world

Segment
Code : 121
Level : Cut
Starttime : 00:03:00:26
Stopetime : 00:03:20:25

Segment Description : Two photomicrographs
Audio-objectitem : CEMENTITE is broken after plating. The main subject is to develop a method for protecting CEMENTITE in the process of plating.

Du Reference Code : MPEG/CON/V/V10
Main Title : Ultra-steel suspends the longest bridge in the world

Segment
Code : 122
Level : Scene
Starttime : 00:03:20:26
Stopetime : 00:05:50:11

Segment Description : Preventing CEMENTITE break-down
Audio-objectitem :

Du Reference Code : MPEG/CON/V/V10
Main Title : Ultra-steel suspends the longest bridge in the world

Segment
Code : 126
Level : Cut
Starttime : 00:03:55:27
Stopetime : 00:04:05:22

Segment Description : Panning-shot of the needle.
Audio-objectitem : The research group of Dr. Takahashi have been inquiring the atomic structure of CEMENTITE.

Du Reference Code : MPEG/CON/V/V10
Main Title : Ultra-steel suspends the longest bridge in the world

Segment
Code : 131
Level : Cut
Starttime : 00:04:42:26
Stopetime : 00:04:58:15

Segment Description : Image of 3D atomic structure of the wire
Audio-objectitem : CEMENTITE is displayed as pale green layer.

segment content
Du Reference Code : MPEG/CON/V/V10
Main Title : Ultra-steel suspends the longest bridge in the world

Segment Code : 132
Segment Level : Cut
Starttime : 00:04:58:16
Stop-time: 00:05:08:16

Segment Description : Magnified 3D atomic structure of the wire
Audio-objectitem : CEMENTITE is displayed as a mob of green points in the center part.

Du Reference Code : MPEG/CON/V/V10
Main Title : Ultra-steel suspends the longest bridge in the world

Segment Code : 133
Segment Level : Cut
Starttime : 00:05:08:17
Stop-time: 00:05:26:10

Segment Description : Panning-shot of the magnified 3D atomic structure
Audio-objectitem : Silicon atoms are displayed as yellow points enclosing CEMENTITE. The silicon is protecting CEMENTITE.
Appendix F
## Query on event-type

<table>
<thead>
<tr>
<th>Event-type Code</th>
<th>Du Reference Code</th>
<th>Main Title</th>
<th>Event Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>PT/FEQ/FAMEQ/B-B-028</td>
<td>Escrituras de compra e venda</td>
<td>Venda de imobiliario</td>
</tr>
<tr>
<td>5</td>
<td>PT/FEQ/FAMEQ/B-B-028-d1</td>
<td>Escrituras de compra e venda</td>
<td>Venda de imobiliario</td>
</tr>
<tr>
<td>5</td>
<td>PT/FEQ/FAMEQ/B-B-028-d10</td>
<td>Escrituras de compra e venda</td>
<td>Venda de imobiliario</td>
</tr>
<tr>
<td>5</td>
<td>PT/FEQ/FAMEQ/B-B-028-d11</td>
<td>Escrituras de compra e venda</td>
<td>Venda de imobiliario</td>
</tr>
<tr>
<td>5</td>
<td>PT/FEQ/FAMEQ/B-B-028-d2</td>
<td>Escrituras de compra e venda</td>
<td>Venda de imobiliario</td>
</tr>
<tr>
<td>5</td>
<td>PT/FEQ/FAMEQ/B-B-028-d3</td>
<td>Escrituras de compra e venda</td>
<td>Venda de imobiliario</td>
</tr>
<tr>
<td>5</td>
<td>PT/FEQ/FAMEQ/B-B-028-d4</td>
<td>Escrituras de compra e venda</td>
<td>Venda de imobiliario</td>
</tr>
<tr>
<td>5</td>
<td>PT/FEQ/FAMEQ/B-B-028-d5</td>
<td>Escrituras de compra e venda</td>
<td>Venda de imobiliario</td>
</tr>
<tr>
<td>5</td>
<td>PT/FEQ/FAMEQ/B-B-028-d6</td>
<td>Escrituras de compra e venda</td>
<td>Venda de imobiliario</td>
</tr>
<tr>
<td>5</td>
<td>PT/FEQ/FAMEQ/B-B-028-d7</td>
<td>Escrituras de compra e venda</td>
<td>Venda de imobiliario</td>
</tr>
<tr>
<td>5</td>
<td>PT/FEQ/FAMEQ/B-B-028-d8</td>
<td>Escrituras de compra e venda</td>
<td>Venda de imobiliario</td>
</tr>
<tr>
<td>5</td>
<td>PT/FEQ/FAMEQ/B-B-028-d9</td>
<td>Escrituras de compra e venda</td>
<td>Venda de imobiliario</td>
</tr>
<tr>
<td>5</td>
<td>PT/FEQ/FAMEQ/B-B-055</td>
<td>Títulos particulares de contratos de compra e venda</td>
<td>Transacção</td>
</tr>
</tbody>
</table>
Appendix G
<table>
<thead>
<tr>
<th>Du Reference Code</th>
<th>Main Title</th>
<th>Creation Role</th>
<th>Creator</th>
</tr>
</thead>
<tbody>
<tr>
<td>PT/FEQ/APEQ/A-014-d1</td>
<td>Passaportes (Eça)</td>
<td>Destinatário</td>
<td>Eça de Queirós</td>
</tr>
<tr>
<td>PT/FEQ/APEQ/B-B-014-d1</td>
<td>Passaportes (1872-1874)</td>
<td>Destinatário</td>
<td>Eça de Queirós</td>
</tr>
<tr>
<td>PT/FEQ/APEQ/B-B-014-d2</td>
<td>Passaportes (1872-1874)</td>
<td>Destinatário</td>
<td>Eça de Queirós</td>
</tr>
<tr>
<td>PT/FEQ/APEQ/B-B-014-d3</td>
<td>Passaportes (1872-1874)</td>
<td>Destinatário</td>
<td>Eça de Queirós</td>
</tr>
<tr>
<td>PT/FEQ/APEQ/B-B-014-d4</td>
<td>Passaportes (1872-1874)</td>
<td>Destinatário</td>
<td>Eça de Queirós</td>
</tr>
<tr>
<td>PT/FEQ/APEQ/B-D-014-d1</td>
<td>Passaportes (1880-1887)</td>
<td>Destinatário</td>
<td>Eça de Queirós</td>
</tr>
<tr>
<td>PT/FEQ/APEQ/B-D-014-d2</td>
<td>Passaportes (1880-1887)</td>
<td>Destinatário</td>
<td>Eça de Queirós</td>
</tr>
<tr>
<td>PT/FEQ/APEQ/B-E-014-d1</td>
<td>Passaportes (1890-1899)</td>
<td>Destinatário</td>
<td>Eça de Queirós</td>
</tr>
<tr>
<td>PT/FEQ/APEQ/B-E-014-d2</td>
<td>Passaportes (1890-1899)</td>
<td>Destinatário</td>
<td>Eça de Queirós</td>
</tr>
<tr>
<td>PT/FEQ/APEQ/B-E-014-d3</td>
<td>Passaportes (1890-1899)</td>
<td>Destinatário</td>
<td>Eça de Queirós</td>
</tr>
</tbody>
</table>

creationrole & title
Appendix H
<table>
<thead>
<tr>
<th>Du Reference Code</th>
<th>Main Title</th>
<th>Event</th>
<th>Du Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>PT/FEQ/FAMEQ/B-B-028-d1</td>
<td>Escrituras de compra e venda</td>
<td>Venda de imobiliario</td>
<td>Escrituras</td>
</tr>
<tr>
<td>PT/FEQ/FAMEQ/B-B-028-d2</td>
<td>Escrituras de compra e venda</td>
<td>Venda de imobiliario</td>
<td>Escrituras</td>
</tr>
<tr>
<td>PT/FEQ/FAMEQ/B-B-028-d3</td>
<td>Escrituras de compra e venda</td>
<td>Venda de imobiliario</td>
<td>Escrituras</td>
</tr>
<tr>
<td>PT/FEQ/FAMEQ/B-B-028-d4</td>
<td>Escrituras de compra e venda</td>
<td>Venda de imobiliario</td>
<td>Escrituras</td>
</tr>
<tr>
<td>PT/FEQ/FAMEQ/B-B-028-d5</td>
<td>Escrituras de compra e venda</td>
<td>Venda de imobiliario</td>
<td>Escrituras</td>
</tr>
<tr>
<td>PT/FEQ/FAMEQ/B-B-028-d6</td>
<td>Escrituras de compra e venda</td>
<td>Venda de imobiliario</td>
<td>Escrituras</td>
</tr>
<tr>
<td>PT/FEQ/FAMEQ/B-B-028-d7</td>
<td>Escrituras de compra e venda</td>
<td>Venda de imobiliario</td>
<td>Escrituras</td>
</tr>
<tr>
<td>PT/FEQ/FAMEQ/B-B-028-d8</td>
<td>Escrituras de compra e venda</td>
<td>Venda de imobiliario</td>
<td>Escrituras</td>
</tr>
<tr>
<td>PT/FEQ/FAMEQ/B-B-028-d9</td>
<td>Escrituras de compra e venda</td>
<td>Venda de imobiliario</td>
<td>Escrituras</td>
</tr>
<tr>
<td>PT/FEQ/FAMEQ/B-B-028-d10</td>
<td>Escrituras de compra e venda</td>
<td>Venda de imobiliario</td>
<td>Escrituras</td>
</tr>
<tr>
<td>PT/FEQ/FAMEQ/B-B-028-d11</td>
<td>Escrituras de compra e venda</td>
<td>Venda de imobiliario</td>
<td>Escrituras</td>
</tr>
<tr>
<td>PT/FEQ/FAMEQ/B-B-028</td>
<td>Escrituras de compra e venda</td>
<td>Venda de imobiliario</td>
<td>Escrituras</td>
</tr>
</tbody>
</table>
Appendix I
Query on entities and event-types

du Reference Code : POR CPF 1006  Main Title : APR 14

entity & event-type
Appendix J
Reference Code : PT/FEQ/APEQ/A-A-017  
**Main Title :** Processos particulares relativos a seguros  
**Keywords :** Processo  
**Summary :** Processos organizados não oficialmente, por um indivíduo, relativos a assuntos e transações específicas sobre constituição de apólices de seguros.  
**Description :**  
**Entity :**  
**Event :**

Reference Code : PT/FEQ/FAMEQ/B-A-044  
**Main Title :** Gestão patrimonial  
**Keywords :** Gestão patrimonial  
**Summary :** Congrega esta secção a documentação relativa à actividade funcional de gestão e administração do património familiar, compreendendo portanto a constituição, manutenção e transacção de propriedades e bens. O conjunto documental analisado diz majoritariamente respeito às quintas de Vila Nova e Tijheiro, situadas em Baião. Inclui igualmente numa área específica a administração do espólio literário de Eça.  
**Description :**  
**Entity :**  
**Event :**

Reference Code : PT/FEQ/FAMEQ/B-B  
**Main Title :** Processos judiciais  
**Keywords :** Processo  
**Summary :** Processos de natureza cível ou orfanológica relativos a questões de transmissão e transações patrimoniais. Podem ser completos ou compor-se de partes isoladas de processos, como articulados, réplicas, apensos,...  
**Description :**  
**Entity :**  
**Event :**

Reference Code : PT/FEQ/FAMEQ/B-B-007  
**Main Title :** Transacção e transmissão de bens  
**Keywords :**  
**Summary :** Compreende esta subsecção documentos relativos a qualquer tipo de transferência de posse de bens ou repartição dos mesmos.  
**Description :**  
**Entity :**  
**Event :**

Reference Code : PT/FEQ/FAMEQ/B-B  
**Main Title :** Transacção e transmissão de bens  
**Keywords :**  
**Summary :** Compreende esta subsecção documentos relativos a qualquer tipo de transferência de posse de bens ou repartição dos mesmos.  
**Description :**  
**Entity :**  
**Event :**

Reference Code : PT/FEQ/FAMEQ/B-B-007  
**Main Title :** Autos de posse  
**Keywords :** Auto  
**Summary :** Documentos notariais declarando a entrada em posse de um bem imóvel, por parte do destinário, após concluída uma transacção judicial ou administrativa.  
**Description :**  
**Entity :**  
**Event :**
Reference Code: PT/FEQ/FAMEQ/B-B-010  
Main Title: Certidões de documentos relativos a transacções patrimoniais.

Keywords: Certidão
Summary: Certidões em que se trasladam e certificam actos notariais relativos a transacções antigas, de natureza patrimonial (compras/vendas, arrendamentos, etc.).

Reference Code: PT/FEQ/FAMEQ/B-B-043  
Main Title: Processos de transacções patrimoniais.

Keywords: Dossier
Summary: Conjunto de documentos de diferentes tipologias agrupados e constituídos em processo, para a concretização de determinada transacção patrimonial.

Reference Code: PT/FEQ/FAMEQ/B-B-055  
Main Title: Títulos particulares de contratos de compra e venda.

Keywords: Título
Summary: Documentos particulares declarando e veiculando a realização de uma transacção entre partes.

Reference Code: PT/FEQ/FAMEQ/B-B-057  
Main Title: Títulos particulares de contratos de permuta.

Keywords: Título
Summary: Documentos particulares declarando e veiculando a realização de transacções (escambos) entre partes.