

PHOTOGRAPH CONSERVATION THESAURUS

Catarina Mateus
Photograph Conservator

Abstract

As a dissertation project for the MA in Preventive Conservation at Northumbria University, U.K., Catarina Mateus is developing an online photograph conservation thesaurus. It consists on designing and implementing an online database resource that can help conservators interpret and use accurate vocabulary on a daily basis. This web tool proposes that the major concept terms related to photograph conservation should be organized through hierarchical facets. The main aim is to standardize the language applied by the community of conservators. In turn, this will avoid ambiguity and help pave the way to enhancing knowledge about the following relationships: materials and applications, deterioration causes and effects, treatments and results.

This paper introduces the major concepts of a thesaurus; explains the methods used on developing the database and presents the structure and its contents.

This project is based on two models, the Getty Art & Architecture Thesaurus Online and the SIPA - Architecture and Urbanism Thesaurus. Also, it has the technical support of a Portuguese public institution, IHRU, in order to make it accessible and feasible through an online information system.

Key Words: Photograph Conservation Thesaurus (PCT); SIPA

INTRODUCTION

The idea of creating a Photograph Conservation Thesaurus (PCT) is not new; it came to light while working at the DGEMN¹ - Portuguese National Monuments Heritage Archive – from 1997 to 2006, through an outsourcing photograph conservation company, Luis Pavão Lda.

I worked with Joao Vieira², who is recently responsible for creating a digital information system named SIPA³ that manages documentation (340.000 photographs, 200.000 drawings, 12.500.000 text pages) related with Portuguese Architecture, Engineering, and Urbanism. This institution has started since 2005 a thesaurus in architecture and urbanism. This experience helped me understand how important it is to use controlled vocabularies when describing and cataloguing any type of archive documents. The relevance of creating a thesaurus related with conservation had always been discussed amongst colleagues. This would not only benefit the conservation department, in helping to stabilize and control newly formed documentation (surveys, diagnosis, reports, etc.) between conservators within the institution, but also, if created properly and according to international standards, it could become a useful tool for Portuguese users, as well as the international conservation community. In spite of all this, other priorities within the institution have put off the project.

In 2007, I decided to begin the project voluntarily as my final dissertation project for the MA in Preventive Conservation. Implementing the PCT implies not only work and commitment but also some technical resources and some guidance through information sciences. These resources were provided by IHRU that kindly approved the idea and offered the digital support necessary to make this project a reality, as well as specialized assistance in the vocabulary construction.

THE PURPOSE OF A CONTROLLED VOCABULARY

To explain the purpose of a controlled vocabulary it is necessary to introduce the concept of an **information retrieval system**. It consists in a “*group of activities and components designed to facilitate access to the subject matter of documents*”⁴. One of the components that provide this type of access is a **controlled vocabulary** created to enhance access to a database, allowing for faster and more precise searching of a specific document by researchers, avoiding ambiguity and confusion in interpreting information and eventually hamper access to materials in databases. A **thesaurus**, like a glossary or a dictionary, has controlled vocabulary, although the main difference being that it focuses on a specific theme and it is organized in a structure that shows important relationships between terms. In other words, it is “*a compilation of terms representing single concepts (...) [and] displays relationships among terms by creating what is called a semantic network. (...)*”⁵.

There are several reasons for choosing a thesaurus as an alternative to a glossary, mainly it is because a thesaurus offers the possibility of not only controlling but, also organizing vocabulary and particular concepts by representing previously established relationships between them, allowing an enhanced and comprehensive understanding of the subject. In other words, it provides not only a term but also integrates the term in its context of application. A multilingual thesaurus also allows for translation of other languages, which offers the possibility of other countries to propose national terminology that is based on the previously established concepts. The database can be complemented with introducing photographs to illustrate the term and bibliography or citations that could be useful for better understanding of the term.

THE RELEVANCE OF A THESAURUS IN THE AREA OF PHOTOGRAPH CONSERVATION

*“Unambiguous terminology and precision in the use of words are important aspects of examination practice. In order to communicate observations and conclusions to others and allow to comparison with other objects, there must be agreement on what the words mean”*⁶.

Barbara Appelbaum⁷

The Portuguese reality is similar to other countries where photograph conservation represents a small area of activity: due to the nonspecific language used about the subject, terms are adapted from other areas of conservation; in other situations, the language used is in English, Italian or French⁸ because there is no research dedicated to translating and adapting the foreign vocabulary. The conservation community acknowledges the fact that there is much work to be done as far as preparing conservators for photograph conservation at university level or other academic programs. In a recent article, Dusan Stulik⁹ mentioned that “*Photograph Conservation as a field of practice and study is relatively well developed in Western Europe, but it is either in its infancy or nonexistent in these other regions of Europe*” [central, eastern and southern Europe]¹⁰.

Another aspect regards understanding the accessed information in English by the conservation community, compared to the lack of knowledge about the information and experiences that are available in other languages. This way, it is impossible to exchange, learn and disseminate experiences of other professionals whose contribution can be just as useful as the ones from Anglo-Saxon countries. Universal access and an increase in information availability would be two of the consequences of permitting this information to go on the web.

After researching for thesauri dedicated to cultural heritage, none were found related to photograph conservation, however, the amount of glossaries dedicated to Photography were substantial. When observed closely, the majority of their contents are concerned with describing photography as a technique or as a method of “capturing” an image, most of them do not update or provide significant information and many lack in detail. Once again, none were found dedicated to photography as an art medium to preserve.

For these reasons, it is my belief that the Photograph Conservation Thesaurus will contribute to disseminating information as well as serve as a tool for conservators on a daily basis.

GUIDELINES AND MODELS FOR CREATING THE THESAURUS

If the main goal of thesauri are to regulate and control the vocabulary used for indexation it is expected that thesauri should be regulated under national and international standards.

In creating a credible and validated model of a thesaurus, international and national standards were used:

a) International Standards - ISO¹¹

ISO 2788:1986 - Documentation: Guidelines for the establishment and development of monolingual thesauri

ISO 5964: 1986 - Guidelines for the establishment and development of multilingual thesauri

b) American Standards - ANSI¹²

ANSI/NISO Z39.19 - Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies¹³

c) Portuguese Standards IPQ¹⁴

NP 4036: 1992 - Documentação: Tesouros monolíngues: directivas para a sua Construção e desenvolvimento - Guidelines for the establishment and development of monolingual thesauri

These standards are very similar in content but there are some slight differences that usually are related with each country's specifications. All were consulted and they are the basis for creating the regulations for the PCT.

The Models

There are many different thesaurus dedicated to heritage. Two were chosen to be the models for creating the PCT:

a) The Art & Architecture Thesaurus (ATT)¹⁵ provided by The Getty is a source of standard terminology for the use in description, cataloguing and documentation of cultural heritage collections. It currently contains *circa* 131.000 terms and other information about concepts. The terms in AAT are used to describe art, architecture, decorative arts and culture and archival materials, ranging from the time of antiquity to the present. It is built according to ISO and NISO standards in regard to the codes of practice for thesaurus construction.

AAT can be used as an assistant to database search engines, as well as manual indexing systems, helping the user forwarding his research between the terms and concepts. It has a program called "The Candidate Term Program" that allows candidates to input information and expand the coverage of specialized subject areas.

b) The Architecture and Urbanism Thesaurus (AUT) provided by IHRU, integrates *circa* 6700 scientific, technical and technological terms related with Portuguese architecture and urbanism from the 20th and 21st centuries. Its purpose is to help with classifying information produced by the Institution, especially a controlled vocabulary for enhancing information retrieval and promote the interoperability between digital systems. This system is used in different departments and databases within the institution, from the archive and library departments, to the National Heritage Inventory or architectural projects. It is available through an ERDMS¹⁶ named ARIS¹⁷, especially developed by a Portuguese company. The software controls production and updating for each record (or term); manages the produced information; associates terms to other digital resources

(for example between the Thesaurus and Library database), creating the interoperability between digital resources created by different services and public bodies and also private Portuguese archives and collections.

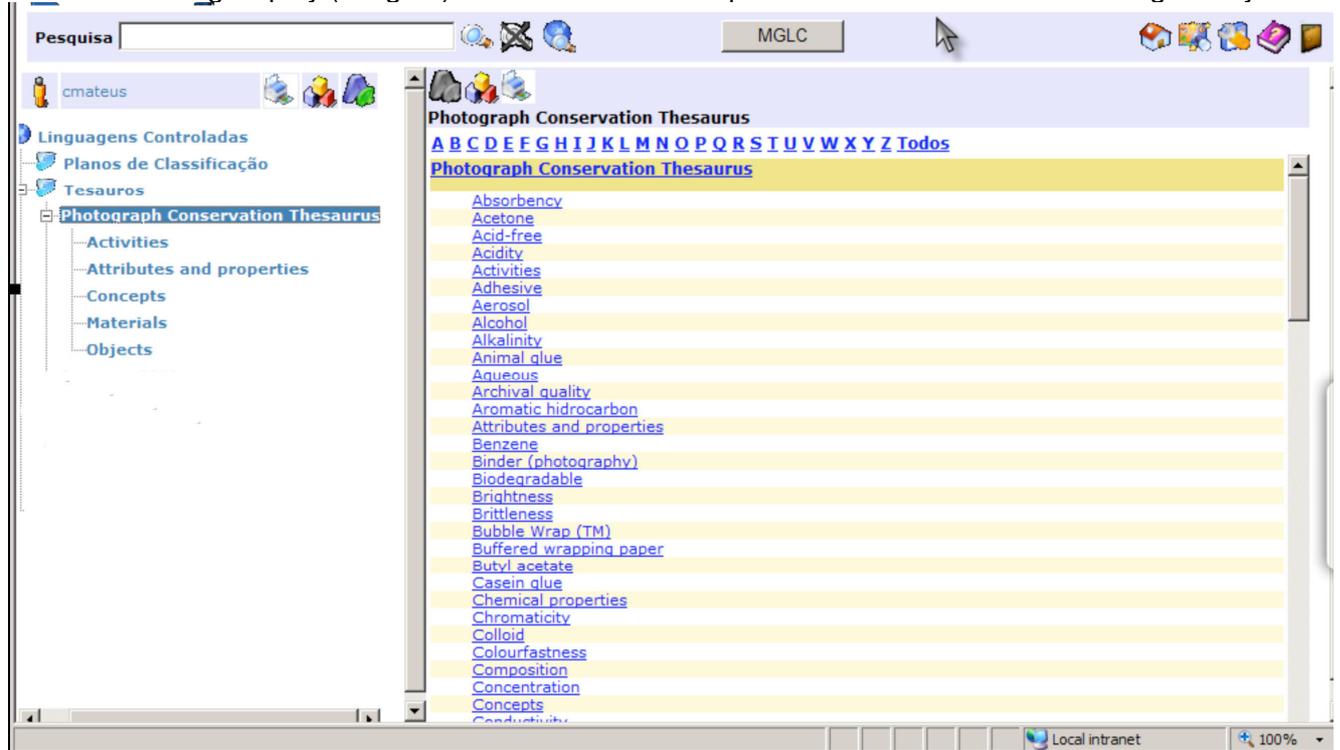
This article's purpose is not to explore the differences and similarities between the two models, only refer them as references and sources used for the PCT.

PHOTOGRAPH CONSERVATION THESAURUS (PCT)

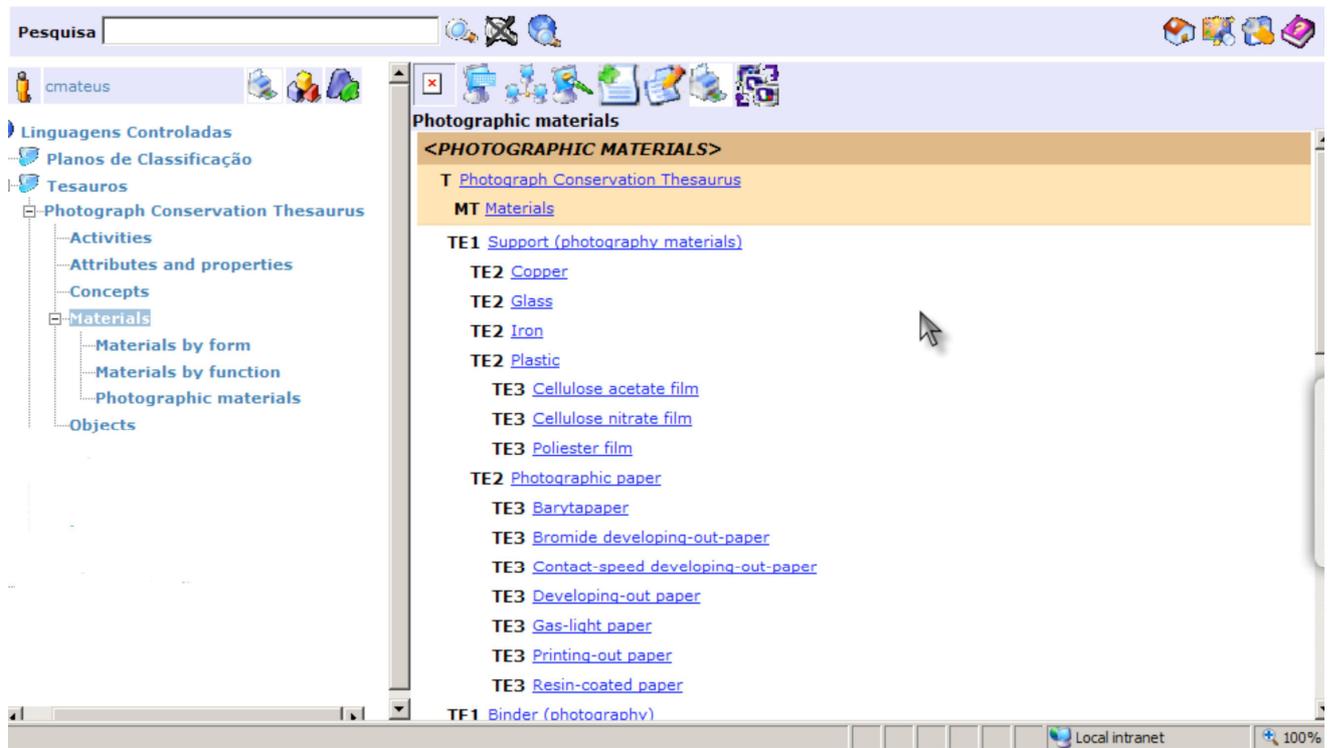
Alphabetical and Classified Presentation

When accessing the PCT, the researcher will find two different ways of presentation: an Alphabetical and Classified presentation of terms.

The following display (image 1) shows the terms in alphabetical order. It is similar to a glossary.



The Classified presentation (image 2) has a tree form and shows the relationships between terms and its hierarchies or subdivisions.



In practice, the indexer just needs to insert the data related to the term once in the database (using the classified component). The system will organize itself automatically according to alphabetical order.

Facets and Types of Hierarchies

The hierarchical structure of the thesaurus is divided into six top levels named Facets, each one subdivided in Broader and Narrower terms. The Facets chosen for the PCT are based on the models above, with the conviction that they are the most appropriate to describe the majority of terms related to conservation and preservation. Not all Facets presented in the models were included in the PCT because some were not considered adequate or significant for the current database.

Facets	Examples of Broader terms
Attributes and Properties	Physical properties Chemical properties Mechanical properties Optical properties Physicochemical properties Conservation Condition
Agents	People related with conservation activities People related with photograph activities
Activities	Photographic Processes Preservation activities Conservation activities Processes of Deterioration Processes of registration and documentation Methodologies and Techniques
Materials	Materials by form Materials by type Materials by function Photographic Materials Conservation Materials Preservation Materials
Objects	Documentation and Information Photographs by type Photographs by form Instruments for conservation and preservation Equipment for conservation and preservation
Concepts	General concepts

Table 1: Facets and Broader terms provided by the PCT

Hierarchical Relationships

All terms representing a single concept have been collected together and one has been established as the descriptor. The descriptors can be graphically displayed through three types of relationships:

Hierarchical relationship – the most important relationship. It is used to show the whole/part or genus/species relationships between terms (Broader Term BT/ Narrower Term NT).

Equivalent relationship – introduces synonyms for a specific concept, including variant spellings. It is designated as USE FOR (UF). For example, Mineral spirit is an equivalent term to White spirit.

The **Hierarchic** and **Equivalent** relationships imply that the terms must be reciprocal, i.e., that the database must show both views.

Associative relationship – Correspond to those terms related with the concept, which means within the context of a particular thesaurus but are not in the same hierarchy (designated as Related Term – RT). This association does not need to be reciprocal, i.e., the term may not be displayed in the other related hierarchy.

Fields

The indexing database requires the submission of minimum specific fields for each term. The fields are:

- a) ID - identifies the record uniquely (IDs are generated by the system)
- b) Parent (the hierarchical relationship. It can be broader/narrower between terms)
- c) Preferred term (the word or group of words used to refer a specific concept)
- d) Language (Portuguese and English)
- e) Contributor of the term (automatically generated by the system)

The other fields optionally submitted are:

- a) Scope Note (Term concept)
- b) Sources for the scope note
- c) Detailed Note
- d) Citations
- e) Non-preferred term (Use For)
- f) Illustration (Image that illustrates the concept)
- g) Translation to other languages
- h) Related Term
- i) Guide Terms - are not indexing terms, but sometimes are needed to provide structure to the hierarchy, forming levels below which records may be positioned. To identify the guide terms angle brackets are used.

The database fields and correspondent abbreviators used in the PCT are the same ones recommended for ISO standards. But the digital platform is currently built according to Portuguese standards (NP), which means that the terms are translated in Portuguese.

In the following table correspondences were made between the abbreviators, their meaning and the Portuguese translation¹⁸:

ISO 2788:1986 Abbreviator	Meaning	Correspondence to ARIS platform
ID	Identification number	ID
SN	Scope Note	NE
USE	Use	USE
UF	Use For	UP
TT	Facet - Top Term	MT
BT	Broader Term	TG
NT	Narrower Term	TE
RT	Related Term	TR

Table 2. Universal abbreviations and correspondance with the ARIS platform

METHODOLOGY

Gathering terms

Gathering terms is a time consuming task and it must have a set of rules in order to make the process easier. Authors are unanimous in this matter: most of them suggest guidelines to achieve good results.

Before building a completely new thesaurus, (...) it is advisable to review existing thesauri to see if one will meet local needs. Sometimes an existing thesaurus can be modified to make it usable [for a specific purpose]¹⁹. In this case, despite the non-existence of a specific thesaurus in this area, some terms may be gathered from other databases related to cultural property.

Some recommendations indicate that an empirical approach is more reliable than a deductive approach, i.e., instead of creating a vocabulary from top to bottom (from Broader Terms to Narrower terms) it is more efficient to collect the terms from different sources and then a category or hierarchy of terms is formed only if there is a need for it.

This inductive method has two important principles: the literary warrant and the user warrant. The first one is based on the theory²⁰ of Wyndham Hulme²¹ that when applied to a thesaurus, indicates that a term is justified only if it occurs within the literature of a subject with some degree of frequency. The literature used to collect terms must be up-to-date, and it should start with the latest publications.

The second one, the user warrant, guarantees that the chosen term is the most representative of future users of the thesaurus. Users questionnaires can be used in order to achieve best results.

When choosing a term frequently used in literature, proper names of persons, organizations, geographic places, named subjects and named events are excluded. The reference to the sources used as warrant for the term should be indicated, through a brief reference to the source or a full bibliography citation.

Characteristics of the term

Criteria should be defined in order to create terms with consistency. The established and most relevant criteria are described above:

- a) Use Roman alphabet
- b) Diacritics or special characters should be avoided except when translating to other languages
- c) Use lower case except on terms based on proper nouns or adjectives
- d) Use singular terms; the plural should be avoided and only used when the singular does not express the accurate concept
- e) Use compound terms (more than one word) only when extremely necessary to define a term
- f) Avoid all punctuation except commas, hyphens, and apostrophes
- g) Use parentheses only for the trademark symbol (*TM*)
- h) Avoid abbreviations, acronyms, and codes, unless the abbreviation is better known and more often used than the full term. Example HVAC
- i) Trademarks or brand names may be included only if the term is found in common usage to refer to a generic material, object, or process.
- j) Verbal nouns or adjectives should be used in gerund
- k) British English is chosen for the Preferred Term. It may be included variant terms (synonyms) in American English when the spelling differs. Synonyms in other languages may also be included.
- l) Translations should have literary warrant in a source that indicates that the term has the same meaning as described in the scope note.

Description of the term – scope note

The scope note is used to explain the meaning of a term and differs from a definition of a glossary or dictionary because it does not provide all the possible meanings of that term but only the single concept related to the context of the thesaurus. It is a free text field but should have the same criteria of composition:

- a) Description of the term from general to particular concepts
- b) First sentence should highlight the fundamental information, i.e., the specific meaning of the term. The following text should answer to the composition, the characteristics, the application and how it is used. The full description and examples can be introduced in the Detailed Note. Citations that are pertinent to enhance the note are introduced in the Citation field.

A detailed layout for a narrower term is presented in the image 3. It is possible to see the term, and the hierarchical relationships, the scope note, an image to illustrate the concept and other information considered necessary.

Albumen print	
ALBUMEN PRINT	
T Photograph Conservation Thesaurus	
MT Objects	
Nota Explicativa	
Nota	<p>Albumen prints (1850-1890s) are probably the most common photographic print material of the 19th century. The albumen print is a printing-out process used for contact printing with a negative, usually a glass plate collodion. A paper sheet is coated with an emulsion composed of light-sensitive salts of silver suspended in albumen (egg white). Most prints are mounted on cardboard because the paper is rather thin and curls easily. They can be recognized by the cream coloured highlight areas and glossy surface. After 1855, albumen prints were almost always toned with gold chloride, which enriched their colour to a purple-brown tone and increased their permanence.</p> <p>The first sign of deterioration of albumen prints is that the highlights become yellowish, eventually the whole print starts to fade and becomes yellow or brown, which is due to the effect of light on the albumen layer. Display, use and exposure to unfavorable circumstances, may cause serious damage within approximately forty years. Poor quality of the paper will cause staining of the albumen layer, while dryness and chemical changes can make the photograph crack and curl. Attempts to remount a curled albumen print can be very dangerous for the photograph. Albumen prints must be kept in the dark and protected by UV filters when exposed.</p>
TG1 <Photographic documents by form>	
TG2 Documents by form	
Recursos de Informação	
Referência Documental	
Sobieszek, Robert A., <i>British Masters of the Albumen Print - A Selection of Mid-Nineteenth-Century Victorian Photography</i> , International Museum of Photography at George Eastman House & The University of Chicago Press, Chicago & London, 1976	
ECPA (European Comitee for Preservation and Access) To have and to hold - Processes http://www.knaw.nl/ecpa/photo/proces05.htm	
	
DGEMN/, Foto Nº 564702	
Albumen print	

These general aspects are considered to be the most important to mention in this article. More specific methodology is available in the dissertation.

PROSPECTS AND FUTURE WORK

The Photograph Conservation Thesaurus will be, in the future, available through SIPA and will use the same digital platform (applied to the Architecture Thesaurus) for indexing and displaying the information that is being produced in the dissertation. Presently, the information is going through a testing phase, i.e., only accessible in the database via intranet, which means that it does not have public access at this moment.

Realizing that the thesaurus is an extensive task that needs more than a few months to achieve; and also that generally it is prepared by an interdisciplinary group of experts; the main goal of the present project is to begin forming a structural model for a Photograph Conservation Thesaurus, creating rules and guidelines for the introduction of terminology, based on ISO standards, and also recognized sources such as the AAT Online Thesaurus.

It is not intended to be a finished product (a thesaurus, like a dictionary, is in a constant updating process), but instead, to make it available to the public, meanwhile encouraging community of users to participate and introduce their own translated language terms through a previously built-in validation system.

This project is proposed with the humble expectation that it will motivate conservators, researchers, investigators and students to contribute and develop it in the future, and turn it into a reference as a working tool.

¹ In 2007, the institution closed down and was merged in a new Public Institution named IHRU – Institute of Habitation and Urban Rehabilitation <http://www.monumentos.pt>

² João Vieira is Head of Department of Information, Library and Archives and the President of ICA - International Council of Archives

³ SIPA . Sistema de Informação para o Património Arquitectónico. (Architectural Heritage Information System).

See http://www.monumentos.pt/Monumentos/forms/002_G.aspx

⁴ Lancaster, F. W. - *Thesaurus construction and use : a condensed course*. Paris : UNESCO, 1985. (PGI-85/WS/11)

⁵ Lanzi, Elisa, *Introduction to Vocabularies, Enhancing Access to Cultural Heritage Information*, Getty Research Institute, 1998

⁶ APPLEBAUM, Barbara, *Conservation Treatment Methodology*, pg37, Butterworth- Heinemann publications, UK, 2007

⁷ Barbara Applebaum is an objects conservator in the USA and also the author of *Guide to Environmental Protection of Collections*, Madison, Conn., Sound View Press, 1991

⁸ Examples such as *Foxing*, *Trattegio* or *Craquelet* are commonly used by conservators in Portugal.

⁹ Dusan Stulik is a senior conservation scientist at Getty Conservation Institute

¹⁰ The Getty Conservation Institute Newsletter “Conservation”, pag.16, vol.22,nº3,2007

¹¹ ISO – International Organization for Standardization

¹² ANSI – American National Standards Institute

¹³ See:

http://www.niso.org/kfile_download?pt=RkGKiXzW643YeUaYUqZ1BFwDhIG4-24RJbcZBWg8uE4vWdpZsJDs4RjLz0t90_d5_ymGsj_IKVAGZww13HuDIXTWtlUyrc77y9xWOHiF7M8%3D

¹⁴ IPQ – Instituto Português da Qualidade . (Portuguese Quality Institute)

¹⁵ www.getty.edu/research/conducting_research/vocabularies/aat/

¹⁶ ERDMS – Electronic Records and Documents Management System

¹⁷ ARIS – Archives and Records Information System

¹⁸ Translation available through Portuguese Standards NP 4036

¹⁹ in LANCASTER, F. W. - *Thesaurus construction and use : a condensed course*. Pag.10, Paris, UNESCO, 1985. (PGI-85/WS/11)

²⁰ Applied to bibliographic classification says that a term is justified only if literature on this topic is known to exist.

²¹ E. Wyndham Hulme (1859-1954) proposed the concept of literary warrant as the basis of book classification and the definition of subject classes, where in classes and names of classes would be derived from existing literature rather than a preconceived philosophical order of sciences, the predominant and favored type of classification system in Hulme's day.