Chapter 9
Metadata for Information Retrieval in Archives

Vicent Giménez Chornet
Universitat Politècnica de València, Spain

ABSTRACT
The traditional system for searching for documents in archives is based on correct preparation of the description instruments: guides, inventories, catalogues and index cards. The emergence of three factors in the last few decades has made the system, used by archive institutions for centuries, no longer the most appropriate for information retrieval. These three factors are information and communication technologies, and electronic indexing, preferably using appropriate thesauri.

This chapter is an attempt to analyse which technological elements and which description and indexing elements directly impact information retrieval of documents. As a result of the analysis, some requirements are proposed that must be observed in archival information systems installed in organisations with the aim of optimising effective information retrieval.

INTRODUCTION
The features inherent in archival documentation, especially the high diversity of types and the absence of a title and intellectual author, which is not the same as the producer, has made its control more difficult, even using different description models. Therefore, this has slowed down implementation of a good system of information retrieval. Organisations have for centuries used archival staff for the control and search of documents that were more or less efficient for the amount of documentation managed. The current problem, since the 20th century, is the huge amount of documentation produced by public administrations, which has not corresponded with an increase in archival staff. The impossibility of cataloguing all the documents archived because of infrastructural deficiencies slows down the process of locating documents that any user might need at any given moment.
Archivists are not uninvolved in research and in the search for solutions enabling more effective retrieval of documents produced by organisations. Often, the urgency to find documents is higher in administrative archives where, as a general rule, the main users are the document producers and they themselves require a rapid response in searching for documents in order to process or resolve current administrative situations. But a recent interest in locating documents has also arisen in historical archives, caused by the digitisation of the records. An essential requirement in this case is the document description and to be able to link digital images to a record of the document.

Due to the interest in information retrieval, researchers in archival centres and educational institutions have worked in a variety of environments to find effective solutions to the demands of today’s society. These solutions have centred on the description, the incorporation of information and communication technologies and document management tools. The questions to be answered after some years of proposals are whether these are sufficient or should new requirements be put forward? And in the archives, has the problem of information retrieval been effectively resolved?

To analyze this issue we will see what proposals exist to describe metadata archival funds, following the implementation of international standards and the problems caused by inadequate description for a good retrieval. Finally, we will propose the basic requirements to be applied in the description in the various metadata standards for archival description, for the efficient retrieval of information.

**Metadata Proposals for Information Retrieval**

Seeing the effectiveness achieved by libraries in information retrieval, one of the solutions proposed in the sphere of archives is the normalisation of the description. In this sense, as far back as 1985, the AMC format was proposed, as a part of the MARC format, for describing hand-written documentation deposited in libraries (Sahli, N., 1985). In the 1990s, the archival community was developing its own standards for description and control of its assets (Jimerson, R.C., 2002). Thus EAD was born, based on XML tags (Pitti, Daniel V., 1997), and ISAD (G), created by the International Council of Archives, 2000). Their aim was to describe documents in an individual way and also to be able to describe series or other types of document groupings. Later, the EAC-CPF (Society of American Archivist, 2008) and the ISAAR (CPF) (International Council of Archives, 2004) were created to describe the producers of these documents: organisations, individuals or families.

The standardisation effort for archival descriptions was accompanied by technological advances. The improvements in communications implemented since the 1990s, fundamentally the Internet and Intranet, technological advances, computers with more storage capacity and processor speed and progress made in software are indispensable for and inseparable from archival description standards when implementing a system for document management.

The standards have been considered not as something new or “a revolutionary transformation of archival practice but just another step in the universal trend undertaken for over a century to convert this practice into a working method based on a terminology” (Esteban Navarro, M. A., Gay Molins, P., 1998). But the standards have brought something more that had not been considered in archival tradition; the description of the context, that is, of the producers of the documents: “archival descriptive traditions have placed greater emphasis on the context and function of records and collections, in keeping with the nature of archival materials, which are the product of organic activity rooted in the context of creation” (Freund, L., Butterworth, R., 2008). The elements of these standards serve as metadata labels, both for describing contemporary admin-
Related Content

Intelligent Biometric System Using Soft Computing Tools
[www.igi-global.com/chapter/intelligent-biometric-system-using-soft/39581?camid=4v1a](www.igi-global.com/chapter/intelligent-biometric-system-using-soft/39581?camid=4v1a)

Two Rough Set-based Software Tools for Analyzing Non-Deterministic Data
[www.igi-global.com/article/two-rough-set-based-software-tools-for-analyzing-non-deterministic-data/111311?camid=4v1a](www.igi-global.com/article/two-rough-set-based-software-tools-for-analyzing-non-deterministic-data/111311?camid=4v1a)

A Hyper-Heuristic Using GRASP with Path-Relinking: A Case Study of the Nurse Rostering Problem
He Jiang, Junying Qiu and Jifeng Xuan (2013). *Interdisciplinary Advances in Information Technology Research* (pp. 89-99).
[www.igi-global.com/chapter/hyper-heuristic-using-grasp-path/74534?camid=4v1a](www.igi-global.com/chapter/hyper-heuristic-using-grasp-path/74534?camid=4v1a)

Fuzzy Logic Approach in Risk Assessment
[www.igi-global.com/chapter/fuzzy-logic-approach-in-risk-assessment/184375?camid=4v1a](www.igi-global.com/chapter/fuzzy-logic-approach-in-risk-assessment/184375?camid=4v1a)