

# Integrated Resource Discovery and Access of Manuscript Materials: the User Perspective

by JAN ROEGIERS

## INTRODUCTION

Recent developments in ICT have given hope to users of manuscript materials that some of their old problems will now be solved. Their primary question is possibly to be understood by the librarians and archivists who, more or less jealously, keep the treasures they are interested in. Strangely enough, the world of users is very often not so familiar to them. The worlds of librarians and archivists often differ more by the methods they use, than by the material they manage. To know which manuscript materials are kept by who is not always simple. National traditions, the fortuities of history and legal regulations have produced intricate situations that cry for better cooperation between those two worlds. According to the manuals used in the education of librarians or archivists, the definition of 'archives' is clear and unambiguous, but if you compare the manuals used in different countries, you observe fundamental differences, even contradictions. The best-known example is what is mostly called in English, *private papers*, in German *Nachlässe*, in Italian *spogli*, but in French, Dutch and other languages you read *archives privées* or something similar. In fact these collections of letters, personal notes and other documents received or written by a single person and kept by him/her, are not considered archives in the true sense by most nineteenth-century archivists and in the Anglo-Saxon tradition, where 'archives' is synonymous with 'public papers' or better 'public records'. Consequently, you rarely have to look for private papers in British archival institutions, but in libraries. In most cases the content of the collection is described according to the rules of manuscript cataloguing, whereas in those countries where private papers form an important part of archival collections, they are described according to archival standards.

## DESCRIPTION METHODS: BOTTOM-UP OR TOP-DOWN

The main difference between the description methods used by archivists on the one side and librarians on the other is the existence or non-existence of a hierarchical relationship to other items or other descriptions. Librarians got their inspiration for manuscript cataloguing from the description of printed material. Their main problem was the lack of standardisation of the manuscript material itself, for which they had to establish themselves a standardised terminology, for instance by attributing titles to works without

any title page, to whole volumes without any label or inscription, or by attributing a uniform title to a very well-known text that appears under a different title. But they always describe manuscripts as individual single units, only linked to each other by the fact that they belong to the same owner and appear in the same catalogue, composed according to the same rules. This is a bottom-up approach, where the catalogue as a final result is the sum of non-related items.

Archivists, at least in our days, work top-down. They first describe the archive formed by a single institution, group, family or person, as a whole. This is the true unit of description. If possible, this description on the macro level is enriched by a further analysis of the archive, describing the series and finally the individual documents that compose the archive. In this way every element is described within its context, reflecting the original function of the document. The reader uses interrelated descriptive information, presented according to hierarchical relationships. Nowadays ISAD(G) offers a full set of standard descriptors that make it possible to computerise the archival description and to create the hierarchical relationship that is needed (ISAD(G), 2000) [1]. Programs based upon EAD, a XML-DTD, make it fully possible to obtain from the so created databases automated inventories or lists that reflect the approach of a professional archivist, taking into consideration the hierarchical relationship of descriptions. Especially American professionals are aware of the fact that this system is not only useful for archivists, but also for manuscript librarians keeping ensembles of interrelated documents, in fact keeping archives or fragments of what originally were archives. The EAD approach has also found its way into the world of museums, where it proved to be the right answer to specific needs.

The top-down approach of the archivists is the result of an originally painful experience. In the seventies and eighties archivists could only envy their colleagues, the librarians, who successfully set up large automated catalogues of their complete holdings of printed material and who started linking the databases they had created. The first attempts to come to automated access for archival collections mostly failed or were, at least, disappointing. As the main reason archivists originally saw the lack of standardisation in their terminology and procedures, a consequence of the lack of uniformity in the material they wanted to make accessible. Once the problems of standardisation solved, the problem remained. Only in the early nineties they realized that the main problem was the inability of the existing programmes, derived from library software, to deal with the problem of interrelation of the items described. Only when they exchanged the bottom-up philosophy of library catalogues for a consequential application of their own top-down approach, reliable archival software became a reality.

## THE WORLD OF THE READER

The third world is the world of the reader, in most cases not educated as an archivist or a librarian and nevertheless in quest of information in our treasuries. The questions of the users of manuscript material sometimes differ from those most librarians are familiar with. Often readers seem more interested in form than in content. They are interested not in the writer, but in the scribe; not in the text, but in writing errors; not in the book, but in the binding; not in how it is now, but in how it was centuries ago; not in how it is kept now, but in the previous owners. Traditional finding tools offer valuable information about these aspects and most readers are familiar with these instruments, the printed or manuscript catalogues, lists and inventories of the collections kept by libraries and archives. The number of the existing tools is considerable. Our Leuven University library keeps some 6,000 volumes of printed catalogues of manuscript collections and the collection of printed inventories of only Belgian archives numbers almost 2,000 items. Many of these instruments are rare to find and, if our collection of manuscript catalogues is rather important, our collection of inventories of foreign archives is very modest and almost limited to the neighbouring countries. For most countries there are no instruments that present a full bibliography of the existing instruments and to learn for instance whether the private papers of someone still exist, where they are kept, if there is full description of their contents, if they are accessible or under which conditions, can take years to find out.

Catalogues and inventories, if they exist, are very often the result of highly specialised knowledge and an enormous effort. Nevertheless they are mostly used by the readers in a rather primitive way and often there is no other way. It's incredible, but there are many catalogues of manuscript collections without any index, and if there is one, it is often limited to the authors whose work is included, although the catalogue itself mentions also e.g. the provenances, for many readers a more important fact. Similarly archival lists or inventories often present no index at all and if there is one, it is mostly limited to the names of persons and places mentioned. If there is one, readers mostly restrict themselves to a consultation of the index and rarely read the introduction where the arrangement of the list is explained or where they can learn about the origin, scope and limits of the collection. In many cases, however, there is no other way, even for the most experienced specialist, than reading the whole catalogue or inventory to find out if it describes something useful. Especially occasional users of archival tools are not familiar with the methods of description archivists use, although their way to present interrelated information asks for some specific effort - often remunerated by a rich catch.

The number of existing instruments is impressive and the quality of many of them is, according to their own standards, very high. But how to use them? Their mere number often presents such an inconvenience to the reader as making it impossible to start his research. Instead of running from one library to the other, glancing through some thousands of volumes to find out if they contain a copy of the text he is studying or a

letter from the hero of his tale, he wants to dispose of instruments that will enable him to browse through endless amounts of descriptions, without leaving his desk. In the meanwhile, instead of going through the existing catalogues, he writes a standard letter to some hundreds of libraries and archives, asking if their collection might keep some documents he could use. Conclusion of this description of the reader's world: the reader wants more, not only in quantity and quality of description, but also in quality of retrieval and access. This only seems possible by new tools that take into account the many questions of the reader, and the many problems that derive from the material he is looking for. He doesn't look for a book that was printed in some hundreds or thousands of copies and of which he wants to know where the nearest copy is to be found, nor for a photocopy of an article that appeared in a journal, present in so many libraries. His first question is if there is anything that could respond to his needs and where he could find this rare bird, unique by definition. And finally, he has the same wish as all other readers: *tout, et tout de suite* ('everything, and immediately'). Until recently, this seemed impossible for manuscript material. Nowadays the reader, familiar with ICT, keenly awaits original solutions that will realise his old dreams.

I am convinced that the only way to proceed, to obtain rapid and lasting results, is to adopt the top-down approach of the archivists.

#### THE TOP-DOWN APPROACH

The bottom-up method of the librarians would mean that you could give an answer to all questions by introducing as many descriptions of individual items as possible into a common database, or by creating links uniting and covering several databases with individual descriptions. But what is an individual item? I can imagine how librarians, keepers of manuscripts, could describe all codices in their collections. But could they also describe all individual letters, kept among nineteenth and twentieth century private papers? And are they aware of the millions of letters and other individual documents, kept in an average archival depository? Or are they going to select between documents worth description and others that are not? This is why I don't believe in the final success of famous enterprises as the Dutch CEN project (*Catalogus Epistularum Neerlandicarum*), aiming at a full description of all individual letters by Dutch-speaking authors since 1600, regardless of their present depository [2]. It is a typical project drafted by librarians who aren't aware of the treasures kept by archives and who forget that the normal place to find letters is not a library, where they are mostly kept as individual 'autographs' and sometimes without specification of the addressee, but an archival depository, where letters appear in their full context and where you are aware that the addressee is often as important as the writer of the letter, sometimes even more. A bottom-up approach is only valid for very important and highly valuable materials, such as illuminated manuscripts and other medieval codices, for letters by Erasmus and

Thomas More, or for the private papers of individuals from the category of Goethe, Einstein, Proust, James Joyce, Picasso or Manzoni, people of whom you know that every scrap of paper will be published and commented on for centuries. It seems very typical that the bottom-up strategy is used by the MASTER project, short for 'Manuscript Access through Standards for Electronic Records', a project funded by the European Commission with De Montfort University of Leicester as its most active partner. Their goal is a single online catalogue of medieval manuscripts in European libraries.

The top-down approach means that the first task of librarians and archivists would consist in establishing an easily accessible database, presenting an overview of all existing repositories where manuscript material is to be found. What the user really needs is *something about everything*, not *everything about something*. I cannot enough insist on this principle, so often neglected by perfectionist librarians and archivists. The first need of the reader is to be aware of the mere existence of resources that could be useful to him. A very short and simple description is often enough. To put it in good scholastic Latin: *melius est esse quam non esse*, or: it is better to be than not to be. And in our days, as we see every day in our contacts with students, not to be on the Internet equals not to be at all.

It makes sense to plan the general overview as a national task for every European country and to link these databases at the European level in a later stage. Austria has set a good example with its *Handschriftbestände in Österreich*, set up by the Kommission für Schrift- und Buchwesen des Mittelalters of the Academy of Sciences (KSBM), the UK with the *National Register of Archives* (NRA), maintained by the Historical Manuscripts Commission (HMC). In fact, similar projects are set up in several countries. In my own country, I can mention the Archiefbank Vlaanderen, meant to become a national register of private archives and shaped after the British model. In many countries it could make sense to think of two databases, one for libraries and another for archives. Nevertheless the distinction becomes more and more artificial, especially since the creation in many countries of specialised documentation centres that collect printed and manuscript material, archives as well as books, periodicals, posters or photographs, all related to a specific topic. You also have to take into account that other types of institutions very often keep important collections of manuscript material, such as museums. Of course it becomes also possible to create a national or regional instrument that enables cross-searching in different databases, such as is done with the Dutch Cultuurwijzer, where the holdings of archives, libraries, museums and other institutions can be searched simultaneously. A very successful example of such a tool is the Online Archive of California (OAC) that provides access to materials such as manuscripts, photographs and works of art held in libraries, museums, archives and other institutions across California.

Apart of general information on the library or other institution, the database should comprehend as soon as possible an overview of all existing collections, holdings, *fonds*,

*Bestände*, or how they might be called, presenting a description on macro-level and references to the existing catalogues, lists, or inventories, published material as well as unpublished. How can one expect a scholar who has never visited the Brussels Royal Library and who discovers that some ten thousand manuscripts of this library are described in a published catalogue in thirteen volumes [3], without any kind of index, to know that he could find in the Library itself a full index in the form of a card system? It might be worth travelling to Brussels, or writing a simple request for information, instead of going through thirteen volumes. The presentation of individual institutions should imply a standardised typology of the depositories themselves, making it possible to link similar institutions as is already done by the Deutsche Archivschule in Marburg which can show you the way to University Archives all over Germany.

## INTEGRATED DISCOVERY & ACCESS

In a further stage the many already existing tools should be fully integrated into the databases. The splendid catalogues, lists and inventories could be integrated in many ways. It would be a great loss of information if they should be replaced by uniformly standardised descriptions, made according to some MARC standard and considerably poorer than the products of the great scholars who composed the catalogues we are familiar with. The most attractive of all solutions, for the experienced user as well as for the beginner, is a scan of the full text, enriched by tagging according to something as XML, for instance by using the EAD-DTD. This would enable systematic searches through different catalogues, and at the same time keep all less standardised information. Something similar is set up by the British Library for the automation of the 70 volumes of its manuscripts catalogue, describing a million of manuscripts. France has been inspired by this example for the automation of the more than hundred volumes of the *Catalogue général des manuscrits des bibliothèques de France* (Creff, 2001). I hope that this work in progress will result in a database as rich as the printed edition. It is likely that the major libraries in Europe dispose of the means for a similar conversion project. The great question is who will take care of the conversion of the other, smaller libraries, very often great in their collection, but poor in their present financial and personal means.

Unfortunately much manuscript material, kept by important as well as by smaller libraries, archives, documentation centres and other institutions, such as museums, has up to now not been described at all. As I have already mentioned, it is much more important for the reader to know that a collection exists and to have a general idea of its content, than to dispose of a description in full detail of every single item. In a second stage short descriptions, according to a generally accepted standard, could enrich the macro-description at low cost. It also seems preferable to use the same standard to encode the essential data contained in old manuscript lists or in rather administrative

inventories and acquisition lists. In a top-down philosophy it remains possible at any moment to enrich the existent description with more details, e.g. by analysing the letters or other individual documents united in an item originally described as a whole, or by describing the single illustrations and by adding information on other aspects than the content itself.

A full scan of the document itself can of course be very useful for the reader, but is not what he generally expects. It makes sense of course, out of reasons of conservation, for extremely precious or fragile materials, for documents that are frequently used, for documents that are of more interest for readers abroad than for a native public or for documents with a high pedagogical value. It is unthinkable to scan completely the vast collections of post-medieval times. Readers often regret the enormous sums absorbed by prestige projects for reproduction or scanning that merely serve the image of the institution and its keeper, whereas large parts of the holdings have never been described or remain inaccessible for various reasons.

Manuscript material often asks for descriptors that are very unfamiliar to librarians, used to printed material. Some examples. First there is the problem of accessibility, often depending on a special permission, especially for more recent documents. It is very important to the reader to be aware of these restrictions. The more he gets easy access to faraway collections, the more it becomes necessary to avoid disappointments, by warning him before he takes the plane. The same is more and more true for copyright problems, where legal regulations become more and more strict and lawyers more and more inventive to enrich themselves. There are also other unexpected elements that can interest a reader of unique material. When I work at the Handschriftensammlung of the Österreichische Nationalbibliothek, I find in every volume a slip bearing the names of the earlier readers who used it. Of course, it is always flattering to know yourself the successor of a famous scholar, but often it is also very useful to know that you are or possibly are not aware of the earlier and especially the more recent use of the document. We have to avoid cases such as the one of the unfortunate scholar who recently presented a manuscript for publication to me, a fully annotated edition of a text from the Vatican Library of which the deciphering and transcription had taken almost a month, and whom I had to tell that an edition of it had been published a year before. Libraries and archives that keep unique material have to establish good contacts with their readers, to be informed about work in progress based upon their collections, and to use this information when necessary. The use of a unique document creates a link between the reader and the document, which has to be recorded, and, as far as our susceptibility for privacy permits, could be communicated to other readers.

A question I have avoided until now is in how far an integrated form of access for printed and manuscript material meets the expectations of the user. Let us say first that integrated access of manuscript material, kept by several institutions, is one of the reader's main wishes. The present depository of a document is often so fortuitous and

unexpected, that integration of the existing or new finding tools is an absolute demand. The first requirement for such integration is the use of good authority files for persons and corporate bodies. Once more archivists have been the first to realise that this is not only a question of avoiding homonyms by the use of standard names, but also includes a set of descriptive elements that have to be presented according to a specific standard. Because of the importance of access points of archival retrieval, the Committee on Descriptive Standards (CDS) of the International Council on Archives (ICA) developed a separate standard, ISAAR(CPF), short for 'International Standard Archival Authority Record for Corporate Bodies, Persons and Families'. [4]. This example has inspired the initiators of the LEAF project, 'Linking and Exploring Authority Files', supported by the European Commission and aiming at a 'pan- European Central Name Authority File', meant among others to serve for MALVINE. It is evident that it doesn't make sense to use different authority files for printed and manuscript material. If LEAF really becomes a common European tool, it will enable integrated access of materials of different types and the readers will surely appreciate the improved extended possibility for retrieval. Libraries that would link their catalogue of printed material to LEAF and enrich LEAF by their own authority files would be of a great help for their own and other readers.

## CONCLUSION

MALVINE is meant as the answer of the library world to the need of the users for integrated discovery and access of modern manuscript material (Weber, 2002). It is described as "an electronic network of European institutions, independent of heterogeneous technical solutions, to enhance access to disparate holdings of modern manuscripts and letters, kept and catalogued in European libraries, archives, documentation centres and museums". This multilingual metadata based search engine for a specialised sector has to provide harmonised access to a large number of European collections. The list of participants in this stage of the project is already impressive. MALVINE is inspired by the archival EAD standard. The ambitions and the expectations are high and the first results, visible since July 2003, are promising. In my opinion the real success of the enterprise will depend on the ability to adopt a top-down strategy. Otherwise the content of MALVINE will remain an arbitrary accumulation of knowledge already available on paper. Another question that is not fully clear to me is the assignment of tasks between MASTER, meant for medieval manuscripts, and MALVINE, for post-medieval or modern material. There is a large category possible in between! The success of MALVINE and similar projects will also depend on the ability of librarians, archivists and others who keep the keys of the treasuries, in explaining the top-down approach to their readers. Readers are used to the bottom-up approach of library catalogues and only a few of them who are experienced in intensive use of archives, are familiar with the other way. Interfaces will have to focus on this pedagogical aspect. If they succeed, the readers will finally have the same access to

manuscripts as they now enjoy for printed materials. Readers will no more have the impression of belonging to an underdeveloped third world. The three worlds of librarians, archivists and readers will meet.

## NOTES

1. ISAD(G) is a result of the Committee on Descriptive Standards (CDS) of the International Council on Archives (ICA). A second edition was published in 1999 and is available in English at [http://www.ica.org/biblio/cds/isad\\_g\\_2e.pdf](http://www.ica.org/biblio/cds/isad_g_2e.pdf) and French at [http://www.ica.org/biblio/cds/isad\\_g\\_2f.pdf](http://www.ica.org/biblio/cds/isad_g_2f.pdf)
2. The CEN catalogue can be consulted via the catalogues of the Royal Library at the Hague, on the spot or by registered readers. The project can also be regarded as merely a catalogue of the holdings of a (large) group of libraries; in this sense it is rather a success, but can be misunderstood by readers who see it as a database of all relevant materials.
3. Joseph van den Gheyn a.o., *Catalogue des manuscrits de la Bibliothèque royale de Belgique*, 13 vols., Brussels 1901-1948.
4. More information and the full text of the latest version can be found on the website of ICA. A group of archivists, who met for the first time in Toronto in March 2001, created a more complete high-level model for the recording and exchange of information about the creators of archival materials, termed "Encoded Archival Context" (EAC), compatible with the ISAAR(CPF) standard.

## REFERENCES

1. Creff, Jean-Arthur. « Quelle informatisation pour le catalogue général des manuscrits des bibliothèques publiques de France? ». *Gazette du livre médiéval*, 39 (automne 2001), 41-45.
2. Dongelmans, B.P.M., A.M.T. Leerintveld. *Digital access to Book Trade Archives*, Leiden : Academic Press Leiden, 2002. VIII, 84 p.
3. Weber, Jutta, "MALVINE, LEAF and Kalliope: Some co-operation models". In: *Digital access to Book Trade Archives*, Leiden : Academic Press Leiden, 2002, p. 49-68.

*Integrated Resource Discovery and Access of Manuscript Materials: the User Perspective*

WEB SITES REFERRED TO IN THE TEXT

- Archiefbank Vlaanderen. <http://www.archiefbank.be/>
- Archivschule Marburg. <http://www.uni-marburg.de/archivschule/>
- British Library Manuscripts Catalogue. <http://molcat.bl.uk/>
- Cultuurwijzer. <http://www.cultuurwijzer.nl/>
- EAC - Encoded Archival Context. <http://www.library.yale.edu/eac/>
- EAD - Encoded Archival Description. <http://www.loc.gov/ead/>
- HMC - Historical Manuscripts Commission. <http://www.hmc.gov.uk/>
- ICA - International Council on Archives. <http://www.ica.org/>
- ISAAR(CPF) - International Standard Archival Authority Record for Corporate Bodies, Persons and Families, 1995. [http://www.ica.org/biblio/cds/isaar\\_eng.html](http://www.ica.org/biblio/cds/isaar_eng.html)
- ISAD(G) : General International Standard Archival Description : adopted by the Committee on Descriptive Standards, Stockholm, Sweden, 19-22 September 1999. Ottawa : International Council on Archives, 2000. 2<sup>nd</sup> ed. [http://www.ica.org/biblio/cds/isad\\_g\\_2e.pdf](http://www.ica.org/biblio/cds/isad_g_2e.pdf)
- KSBM - Kommission für Schrift- und Buchwesen des Mittelalters of the Academy of Sciences. <http://www.oeaw.ac.at/ksbm/>
- LEAF - Linking and Exploring Authority Files. <http://www.crxnet.com/leaf>
- MALVINE - - Manuscripts and Letters via Integrated Networks in Europe. <http://www.malvine.org/>
- MASTER - Manuscript Access through Standards for Electronic Records. <http://www.cta.dmu.ac.uk/projects/master/>
- NRA - National Register of Archives. <http://www.hmc.gov.uk/nra/>