

## Manuscript Cataloguing in a Comparative Perspective: State of the Art, Common Challenges, Future Directions

7-10 May 2018 at the Centre for the Study of Manuscript Cultures in Hamburg

10 May, 11.30 – 12.30 Discussion and Conclusions

Claudia Fabian, Bayerische Staatsbibliothek, Leiterin der Abteilung Handschriften und Alte Drucke

Ladies and gentlemen, dear colleagues,

this is now the fourth day of your meeting, and I admire all of you present here to sum up and finalize the state of art of manuscript cataloguing in a comparative perspective. I will do this from the point of view of a librarian responsible for infrastructure, the creation of manuscript catalogues and the building of manuscript databases. Through the various presentations you have become aware about the common challenges, and why they are common.

It is first because of the objects – the manuscripts:

All of them are unique, truly individual, some of them of outstanding value. They present a broad variety: in time, in place, in language, in materiality, in tradition. They can be considered under a number of different aspects. They have a long history and itinerary which has not yet come to its end. They are traditionally considered by an international learnt audience and research on them continues. Although – or because - they are unique, they are prone to all kinds of contextualisation. They teach about culture in its broadest and widest sense, they are part of various traditions, one of them the tradition of collecting, another central and recurrent one: the tradition of texts.

But (2) today, in the library or other memory institutions, we do not only have to preserve, care for and consider the unique original items. We also have to deal with photo archives of manuscripts, microforms, digital copies – the VHMM did remind us of the importance of this approach. They easily able to develop a life of their own besides the original and which allows for a variety of new approaches.

And (3), both libraries and research projects, we have today a rich tradition of description of some manuscripts (2000 years, we heard, for Yao manuscripts) – and a lively concern for those items still hidden, not or under-described, important information resources threatened by oblivion, loss or destruction although unique testimonies for cultural heritage. The existing catalogues often are catalogues in print or even manuscript format: they claim for conversion and reconsideration as they contain a wealth of scholarly concern and insights to be translated into the digital environment, to remain known and to build on. We also have numerous databases for or concerned with. Most of them are outside the

mainstream librarian or information technology approach. We see in various projects how these databases or information tools become obsolete, need migration. We have excellent tools and wonderful ideas about approaching the manuscripts with refined technological means: only to mention electronic editions, online bibliographies, digital annotation tools, automated recognition of characters, of images, and comparison of images ... but we are aware that these e-forms of publication and access are not yet future proved. Longterm archiving might work, but longterm access is another challenge.

A first conclusion of all this should be positive: manuscripts remain on the agenda, they have an ongoing value and interest for research. The care which has to be taken is not finished with any published volume or accomplished database, not with a digital copy or a printed facsimile of the manuscript. All these elements are steps, essential steps in an ongoing research process making manuscripts and manuscript culture relevant for today's and tomorrow's challenges and questions. Again, I think VHMM did rightly stress our common responsibility in this, but also illustrates this ongoing concern: only think how the Benedictines cared for manuscripts since their foundation.

And there are even better news today: The challenges we face in research projects, in research institutes, in libraries are today common challenges – maybe thanks to the digital environment and infrastructure which crosses traditional institutional borders. So we can face them with strong partners to find best future directions and walk on firm paths.

Description and access, a key business of the library, owner and doorkeeper of manuscripts, throughout centuries, is still, again and with new accents, top on the agenda of the research library. It is no longer mainly focused on how best and fastest to deal with the increasing modern print collection or electronic publications. The new international cataloguing framework, called „Resource description and access“ (RDA), takes already within its name account of ALL resources. It is now time to consider especially those which so far – for a number of reasons - remained outside the so-called „bibliographic universe“. Manuscripts are one of the immediate challenges for RDA. There is an expert group in the German speaking RDA community defining rules for manuscript cataloguing, taking into account what exists in order to align and standardize existing practice, a second, more advanced group does the same for private archival material, among them letters. RDA has a much broader approach than the MARC records, is fit for linked open data and integration of controlled vocabularies. It allows different methods of describing – among them even unstructured text. Aligning the manuscripts to RDA allows to integrate them into a broad, widely shared and thus sustainable culture and universe of description and access.

RDA includes and aligns standardized vocabularies and authority files. Authority records play a future proved and increasing role for access and interoperability of information in the semantic web. From very early onwards manuscript culture was somehow involved in what later became authority records – and we have seen part of this learnt tradition and practice of indices, clavis, census work very active in the 19th century and till today. In the late 1960ies IFLA published an authority list for titles of religious works. In my first years in the

library in the late 1980ies I inherited a list of medieval authors from the manuscripts department (later: Personennamen des Mittelalters, PMA), which became one of the first name authority files within GND (Gemeinsame Normdatei), the comprehensive authority file for German speaking countries). Today GND offers a huge number of authority records and develops them in an international context and – this is important and relatively new – reaches out to a research oriented environment. But for an authority record it does not matter whether it comes from GND, CERL Thesaurus, Getty Thesauri, LoC or any other accepted authority file. The importance is not the standardized name form or the variants, there may be several acceptable names. The important thing is the record number – the authority number -, clearly designating and defining the entity (and thus differentiating it from other similar entities). The definition of entities with a link to an authority record is a vital step in making manuscript descriptions accessible and interoperable in the internet world or semantic universe. And please note: databases do not only profit from authority work. They have a lot to contribute to it, more than standardized data, they can offer research proved information, quality authority data. There is room for development and cooperation in various research projects.

The most convincing and driving force for integrating manuscripts in common access and discovery systems is the digitization of the originals. The unique objects hidden so far in treasure rooms, remote collections, anyway: a fixed place, known only to learnt experts, willing to study huge volumes and to travel, are thus available for everybody on the internet. The digitizing institution profits from a streamlined workflow from the catalogue record to a digital workflow tool, the digital copy is automatically provided with some basic metadata from this record, and the link automatically integrated into the record. The catalogue record (best in the form of linked open data) with the link to the digital copy can then travel into all kind of systems: digital portals, union catalogues etc. Nothing distinguishes the digital copy of a manuscript from the digital copy of a printed book, an image, an object etc. – if not the metadata. The digital copies of manuscripts are in deep need of metadata, as a manuscript even less explains itself than a printed book. Thus digitisation of manuscripts puts pressure on their description and on all kinds of interoperability issues.

How nice would it be to link from the digital manuscript to the digital special catalogue – but is there only one???? Inserting links, keeping links up-to-date, keeping servers alive is a question of continuous care and effort – the same is true for the services expected from a digital surrogate, the linking must be automatically done and easily accessible.

IIIF is a new, future oriented step for digitized manuscripts. It is an interoperability framework – allowing digital images to leave their silos. For IIIF the image is first. You must make it understandable and interoperable. For this you need metadata – cataloguing – which now can come second. IIIF thus emulates the „normal“ approach to an original: you have a manuscript – you describe (annotate) it ... You can even compare it to what you know, have already seen. Digital images, interoperable digital images provide what research needs and demands: evidence and the checking of evidence.

In the last years a new communication grew between manuscript oriented research projects, infrastructure providers and conservation institutes (museums, archives, libraries) and these four days have been a living testimony of this spirit of cooperation. We are better placed today, as we have tools available that make this effort practically feasible and the results visible, thus more convincing than a mere theory. Today, libraries do not only provide raw material or technical infrastructure. They are partners in the common concern for interoperability and sustainability of research results in manuscript studies. We can put our instruments and knowledge together for better visibility, perception and impression – within the wide area of digital humanities. The sharing, the cooperation, the dialogue works among disciplines, but also among institutional borders. For sure, this is an effort, this is a cost, but it will have a profit – for the objects and the outcome of research. We live in happy times: DFG is about to start a digitization campaign, building on the masterplan for digitizing medieval manuscripts in German collections published in 2015 – and here cooperation between libraries and research projects is a key issue.

So what to do?

To profit from this kairós, we have to consider our various projects and achievements under the aspect of interoperability, of milestones in building a network, of bridging between one database and another, of enriching information, of reaching out to other, related information tools. This maybe a bit painful, as long as we regard our projects and our manuscripts in a holistic view which we might see threatened, dissolved into pieces. But deconstruction is not the aim of the operation, although modularisation is an issue. The aim is to recognize the numbers of entities we are dealing with, to assess their value in a networked environment and to make these entities reach out of their environment into a network, to contextualize in various contexts, cultural, linguistic, collection, provenance oriented, but also to create new, interdisciplinary relationships of all kind – even beyond the traditional reach of humanities, e.g. with material sciences (and we heard quite some interesting issues about weight and measures). By doing so, the database which provides these entities gains in importance and visibility – proves what it wants to be: a specialist approach to a specialist object on which to build for further and deeper insights.

Defining entities is the key issue, followed by desambiguation and clear (interoperable) naming of the entity. The entity at the heart of most of our databases is the manuscript, the material, physical object of the real world. As we heard, there are ambitions to create an International Standard Manuscript Identifier (ISMI) to allow clear ISO-based naming of this real world manuscript for interoperability. As long as we are not yet there, standardized naming in each database, based on the holding institution (which ideally is represented by an ISIL) and the shelfmark, is a first step. The GND foresees a record for a manuscript as material object, called „Schriftdenkmal“. Our new national manuscripts portal will clearly separate between a record for the manuscript as a cultural heritage object and possibly multiple records for the description of the object. (The manuscript itself as a real world „cultural heritage object“ as called in the museum’s world must be clearly distinguished

from its digital copies (which are designed by a URL, URN, ark, DOI – and thus seem to be easily interoperable) as there may be many of them, and especially IIF encourages to build even more. It must also be clearly distinguished from descriptions, catalogue entries referring to it. There may also be many of them – and they can be quoted more and more by relatively stable database numbers.)

A lot of manuscript studies are less focussed on the material object than on the aspect of text transmission and tradition. RDA invites to integrate this work into the logical framework of the FRBR, defining a work title. In Germany, Handschriftencensus now cooperates with the national authority file to integrate their work titles for medieval German texts into GND. For sure, an authority record approach will not always allow for the kind of detail a specialized database can provide. But it can allow to take in the kind of information useful to make the specialist database interoperable.

Provenance is another important aspect of manuscript description. Again here authority files provide a vocabulary into which information of special databases can be integrated.

Persons, institutions, languages, countries, material support, time, place ... all this is well cared for in the RDA framework – and these entities or information elements are present and mostly harmonized according to internal rule in practically all manuscript databases. A number of specialist databases, for watermarks, bindings, already offer interoperable specialist information to be taken account by linking.

If you look for partners in analysing your database for alignment: the group RDA and manuscripts will listen to you, the group of Handschriftenportal, CERL with its Thesaurus.

With all this in mind, let us reconsider our effort to make best out of our manuscripts within our projects and reach out beyond the project, the special database, the particularities, find new partners, build new bridges – it will be rewarding. Days like these allow us to come to know each other and to engage in fruitful dialogues, conversations – and I hope so – cooperations and synergies.