Federal Archives, Germany:
Proposal for a pilot project supported by the Andrew W. Mellon Foundation

International Standards for Online Finding Aids in German Archives

1. Conditions and expectations of the project

The German archival institutions are experiencing in the last years an increasing demand from scholars from other countries especially the United States which reflect the globalization of research also in the field of humanities as it could be witnessed already before in natural sciences. At the same time archives in Germany have started to use the data in their large archival systems to produce online finding aids to allow researchers to better prepare their visits to the repositories in advance from at home. With the increasing quantities of online-finding aids it is time now to join the German efforts more closely with the international developments and to allow cross boarder research in integrated gateways and services like the Research Libraries Group (RLG).

The Federal Archives “Bundesarchiv” is the archives of the federal government of Germany and its functional predecessors. It was founded in 1952 as a federal agency and reports to the Federal Commissioner for Cultural and Media Affairs in the Federal Chancellery. According to the Federal Archives Law its tasks consist in preserving the public records of the federal government and to provide access to the material and exploit it for scientific and educational purposes. T he Bundesarchiv is by far the largest archival institution in Germany with more than 800 employees and nine regional sites. It is most experienced with internet applications and disposes of a well staffed IT-department. The leading role of the Bundesarchiv in this field is widely accepted considering the fact that it was asked by the conference of the representatives of state archives, which meets twice a year, to prepare together with the archives school Marburg a grant proposal to be submitted to the German science foundation (Deutsche Forschungsgemeinschaft - DFG) for the creation of a gateway of German archives.

After several projects testing the usefulness of EAD (Encoded Archival Description) for German archival working methods the Bundesarchiv has adopted an EAD strategy last year and has started to produce and keep online-finding-aids compatible to EAD. Two software tools are used in the federal archives that can produce EAD-finding aids. Besides the Federal Archives further institutions like the professional training institution Archivschule Marburg or the state archives of North-Rhein Westfalia are thinking about the usefulness of EAD. Plans for a German or a European gateway to archives cannot be separated from reflections about the application of EAD in Europe, where it is already applied to a large extent in the UK and France, and its use for further exchange and collaboration. Considering problems that libraries experience because of different standards with national background traditions and the difficulties to agree to a common world wide standard it seems to be extremely useful just in this moment, when the reflections inside the archival profession start about using standards for internet application and data exchange to promote the far developed standards family

around archival description and internet presentation like EAD, EAC (Encoded Archival Context) and METS (Metadata Encoding and Transmission Standard) in Germany before own standards are developed. The use of common standards would allow to join the efforts in their further development and to circumvent the later costly work on exchange interfaces and data migration.

The projects and experiences with EAD since the late 90s have shown, that in spite of some smaller inconveniences EAD is appropriate for German online-finding aids and that it can be adopted to the German standards. Now it is time to develop methods that present added values for the daily archival work, so that a critical mass of EAD-finding aids is produced and with it the enhanced functionalities become more obvious. The knowledge about the standards should be spread more widely inside the archival community and tools are needed for easier access to the most relevant information. The use of the international standard EAD offers several benefits for German archives. It would ease the integration of German archival descriptions into international networks necessary for the increasingly globalizing historical research and serves as an additional access layer on top of the existing systems.

2. The professional background

Germany has a very stable tradition of standardized archival description since more than a hundred years assured by a central training institution which had to be passed by every archivist applying for a position in public archives services. This standardized techniques include the establishment of holding guides for repositories representing with the tectonics the complete structure of their holdings and including standardized elements for the description of the fonds. These guides were published in print and since some years they are mostly accessible online either in structured form or as database queries. The description of records groups or fonds in the guide refers to the corresponding finding aids. They represent the next, more detailed level with standardized title pages, an introduction, the structured lists of items and a list of index terms with optional annexes. The strategy for online presentation in most cases lays the priority in presenting 100% of the holdings guide online and than as many finding aids of a repository as possible.

In the legacy finding aids the core data for each descriptive unit consist mainly of a title, further descriptions, running dates, and a reference code together with necessary remarks. Online finding aids contain further data for instance supplementary information on different levels about processing, appraisal decisions, history of the records creator and information about what had happened to the described records before their transfer. The finding aids had been prepared manually for printing until the first computerized work possibilities were offered to repositories. The first ADP-applications were developed in the 80s. As the data for the descriptive units were captured before on card files now these data were entered into databases. The databases were sorted according to the numbers of the arrangement scheme, entered with the descriptive units and the lists were than exported to a text processor to produce the finding after adding title page, introduction and index terms. The techniques of databases were used

to elaborate highly sophisticated relations between different parts of the description as well as the arrangement scheme.

The use of databases for the descriptive units assured the formal consistency of the descriptive data of single units. However the weak point was the need to edit the finding aid in a different system after export of the data to a word processor, because of the need to keep track of any changes in the database for further uses. Some attempts were made to offer the databases for direct search by users. But than the context was missing and that made the search very inefficient. Therefore parallel to the development of EAD in the US tools were developed to generate online-finding aids with import of data from the databases. A consortium of Bundesarchiv, Archivschule Marburg and the archives directorate of the Land Baden-Württemberg, named PARSIFAL, developed from 1995 on MIDOSAonline as a converter of database data into HTML-finding aids. This effort was financed by the German science foundation DFG with 80.000$. When EAD version 1 was issued in 1998 reflections started about whether a conversion of German finding aids to EAD might be possible. To get better information and experiences about a possible adaptation of EAD in Germany a joint working group was installed in 2000, funded collaboratively by CLIR (Council on Library and Information Resources) and DFG with members of the EAD working group of SAA (Society of American Archivists) and the same number of German archivists, which discussed differences and communalities between German and American archival descriptive practices and the possibilities to implement EAD for German finding aids. Knowledge about description practices in American archives was a key to understanding EAD better as a way to standardize the structure of the complete texts of finding aids, irrespective of how they were produced. EAD offered the possibility to circumvent the weak point of the database applications by opening new conceptual ways for the production of finding aids, especially with the use of XML. Again the perspective on the whole finding aid was possible.

In response to the knowledge gained from the joint working group the tool MidosaXML was developed, now completely based on XML, with the capability to import data from databases in different formats. MidosaXML, conceived as finding aids editor, looks like a XML-editor, but works on the level of the finding aid, not of the DTD and therefore needs no knowledge on XML or EAD, but can produce EAD finding aids as well as HTML presentations or print out manuscripts from the same data. Even if it works with XML it offers several functionalities known from databases like sorting or searching in different categories, from word processing like formatting texts, and allows for instance manual construction of arrangement schemes with drag and drop functions.

Besides the practical use of MidosaXML as an independent tool its development served as a pilot for the applicability of XML-working methods for the combination of data from database and text processing into the finding aid structure. It furthermore proved that the conversion of German descriptions to EAD does not entail information losses. Parallel to this development the Bundesarchiv organized another project funded by the DFG to test the applicability of EAD for the international exchange of data from its large descriptive database BASYS (Bundesarchiv-System). The outcomes BASYS-Fox (Finding aids Online with XML) show now that also in this case EAD is applicable and useful. Other state

archives in Germany on the level of the Länder have just started to think about conversion of their data into EAD. Especially when cross boarder collaboration with neighboring countries is planned, EAD becomes inevitable. The next step now is to join the finding aids to the description of fonds in the holdings guides and make them online accessible including the navigation between both levels. EAC seems to be very much adapted to this intention, because its elements include all information needed for the guides, including an authority control for the names of the records creators, which was not available until now.

EAC presents a very interesting development especially for German archives. In the older archival tradition no need for authority control was felt, because a different form of a name might just bear informational value which was looked for. In archives new information drawn from something nobody saw or realized before is more valuable than the well known facts. However with the new technologies authority control may open the ways to new methods of research, that help to find a path through the material. A seamless combination of different research strategies including hyperlink navigation and key term search supported by authority control may help users to redefine their questions and their orientation so that they can investigate more efficiently the holdings in shorter time and with more security that they did not leave out any relevant material. Besides authority control the possibility of managing descriptions of records creators with EAC seems to offer furthermore a chance to actualize the replacement and enhancement of what was tried in the 70s with former competencies card files and was stopped because of the inefficiency of the analogue method. These files were intended to capture descriptions of changing competencies and functions in administrative agencies. EAC finally seems to offer the chance to revitalize these efforts. It can handle this information much more smoothly and use it for building networks together with the description of the fonds. This capacity of describing biographies or competencies, together with the authority control of changing personal or agency nominations and government structures combined with the description of the fonds makes it very attractive. These two aspects, the authority control for research and the description of changing competencies offer the possibility to use EAC for encoding the holdings guides. A pilot would be useful to demonstrate its usability and the linking capacities to EAD descriptions. This might than be enhanced by the links to images of the fonds described.

In contrast to database applications for descriptions and the early production of online-finding aids digitization of archival material has been regarded by German archives with a lot of hesitation considering the high costs of storage, retrieval, and securing longevity. Furthermore the archives were afraid of the danger of creating bone yards of information in the internet, which would not offer all its richness because of the loss of structure and interrelatedness. They furthermore worried about the usefulness of large amounts of digitized images considering the fact that their users are eager to see what other did not yet see before them. Therefore several strategies to use digitization are discussed. They cover digitization on demand, images for learning tools and digital expositions. All archival reproductions especially from German holdings need to be presented inside their physical and logical context to be understandable. Therefore images need to be linked to the archival description and tools are crucial to handle these links. It might be imagined that digital reproductions are delivered to the users

together with an extract from the corresponding finding aid or that images in an exposition refer to the description as explanation of their background. E-learning tools including images from the archival material as examples might be built up the way that future users can try out, what would be their best research method, how long they would need for their work and how many items they should order per day. To make imaging more attractive for German archives tools are needed, that allow to construct the different instances of combined information for the different purposes at the desk of the archivists. Therefore metadata of digital reproductions, including the technical metadata, the addresses, the legend, and other information must be managed in a way that links them up to the descriptions. METS seems to be a structure form for metadata of images that can fulfill these expectations when it is built into tools for the archivists desk.

Joining archival descriptive data in EAD-format, EAC-authority files and the metadata for images seems to be a promising way to deliver complete archival information via the internet and with internet technologies. In this combination it might be used for digitization on demand, for cultural expositions or for tutorials for scholars and students including e-learning tools for the preparation of a visit to archives. Therefore the following step should be the study of METS to test its applicability for the management of digitized images from the records of the respective fonds to make them available together with their context of creation and the archival description. With these experiences the traditional hesitation towards digitization may be overcome.

Of special interest in this context is the project just begun by the University of California, San Diego Libraries, the New York University Libraries, and the Five Colleges Libraries to support development of software applications for archival processing and the management of archival information under the name of “The Archivist’s Toolkit”. Especially the incorporation of METS within a descriptive module and the integration of components for digital and media asset tracking and management are of high interest because there is not yet any tool for these purposes available. The intention of this project to develop reusable tools based on open source software and especially intended for small and medium archives seems to be very promising also for the perspective from abroad. For the developments ahead an open source toolkit might be of great use as the archival community in German explores how to improve its systems. Conversely the long German experiences with the use of databases for description as well as with the XML-based tool MidosaXML might be of use for the developers of the toolkit. It seems therefore to be useful for both sides to establish contacts with visits and exchange of experiences to learn from the project and to avoid double expenditures.

3. The aims of the project

The project should support the integration of German archives into the network of international standards for online finding aids and the use of the internet for the presentation of archival holdings, especially with the use of EAD, EAC, and METS

by creating a specialized information website making available translations and best practice guidelines, as well as pilot applications. Results might be a thorough documentation on the three standards for the German professional public and guidelines for best practice for the application of the standards in German context. This would facilitate the introduction of the standards and it might considerably reduce the time needed for creating awareness of the standards and their benefits and for establishing trust in the usefulness for the own practical work. On the other hand the participation of German archives in the international user community might also support the stabilization of the standards development and their maintenance. Therefore the standards shall be better available. Besides the documentation the first steps for the development of tools for easy application should show which way to go further. One result might be the paving of the way for further initiatives that aim at providing a toolset for the archivist workbench in Germany that facilitates the application of the standards and enhances at the same time the productivity of the archival processing as well as of the use of archival material.

4. Work packages

To prepare the adaptation of the three standards for German archival practice the following steps should be undertaken:

- Translation of the central documents of all three standards into German
- Application testing with a pilot project for one or two record groups including description in EAD-format, combined with the authority control of the organizational background using EAC and with searchable images of the collections.
- Evaluation of the experiences
- Drafting of best practice guidelines for the application of the three standards in the German archival context.

These steps should be organized in the following work packages:

WP1: Translations

In order to promote the acceptance of the mentioned international standards for online finding aids in German archives the following texts should be translated:

1.1 EAD
EAD-Tag Library
EAD-Applications Guidelines (new version)

1.2 EAC
EAC-Tag Library

1.3 METS
Overview and Tutorial
METS Tag-Library

The translations should be accompanied by the publication of articles about the context of the network of standards.

As a sort of kick-off meeting a colloque should be held about the functions of the standards especially METS with examples of applications in US institutions. Partners might be RLG and the University of Virginia, Columbia University and New York University. Three to four persons should be invited and asked to present their developments and experiences, especially concerning the combination of archival description and METS. Participants should be interested German archivists from the state archives who are effective as multipliers.

WP2: Development of a pilot application (testbed)

2.1 Two record groups should be selected as a test bed for a pilot application. These might be an administrative record group of an agency with often changing responsibilities or something similar and a collection of personal papers. The criteria for the selection ought to be

- that a complete archival description and finding aid is available in MidosaXML and EAD;
- that a digitization would be easily applicable, for instance from microfilm.

2.2 For these record groups structures and layouts for capturing tools for the metadata for the images (according to METS) and for EAC-Data should be developed together with a linking mechanism to combine them to the descriptive data in a flexible yet robust way.

2.3 The following step would be the development of different multidimensional access ways to the data with different yet interlinked research interfaces, between which the user can switch at will and will be supported by retracing his own search history. That would need flexible finding aids with structural navigation, full text search (“google”) and authority term access like records creators and their different nomination in different periods. Further access ways combined to the description of the images might be thought of. They should include tools like page turners and representation of bound volumes.

2.4 Evaluation process

During the development all decisions should be recorded, alternatives described and experiences and reflections laid down, so that the development of the pilot could serve for evaluation and as basis for drafting best practice guidelines with proposals for practical applications

WP3: Create a website

Create a EAD/EAC/METS website for the pilot application and as a central focal point on standards for online finding aids. The website should contain the translated documentation, further information material on the standards as well as link lists. If feasible a newsletter or a special listserve might be conceived targeting at the professional community of archivists.

WP4: Contacting the Archivist’s Toolset Project in US

Contacts to the Archivist’s Toolset Project at San Diego/New York for a feasibility study to see whether their developments might be applicable as such or in an adapted way to German Archives.

5. Expected results

1. Website for a better access to the documentation about the standards with translation of the mentioned standards, other useful information and support. It will be used for testing of the pilot application and for the documentation of the project, its development and experiences.

2. Pilot-application with evaluation and best practice guidelines for special application strategies in Germany presented in a way that its developments and findings can be reused by the archival community.

3. Feasibility Study: What can be adopted from the Archivist's Toolset project, what should be adapted and how should it be presented for use in the German archival context.