



The DFG Viewer for Interoperability in Germany

Ralf Goebel

Deutsche Forschungsgemeinschaft, DFG, Germany,
ralf.goebel@dfg.de

Sebastian Meyer

Saxon State Library Dresden, Germany,
sebastian.meyer@slub-dresden.de

Abstract

This article deals with the DFG Viewer for Interoperability, a free and open source web-based viewer for digitised books, and assesses its relevance for interoperability in Germany. First the specific situation in Germany is described, including the important role of the Deutsche Forschungsgemeinschaft (German Research Foundation). The article then moves on to the overall concept of the viewer and its technical background. It introduces the data formats and standards used, it briefly illustrates how the viewer works and includes a few examples.

Key Words: interoperability; metadata standards; METS; MODS; OAI-PMH; open source

The Deutsche Forschungsgemeinschaft (German Research Foundation)

As the largest third-party funding body in Germany, the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation¹) also funds the establishment of digital collections at German libraries. The DFG itself is not a library institution nor is it exclusively responsible for libraries. Rather, it is the central self-governing research funding organisation of German science

and research. The DFG is financed by the federal and state governments, but acts independently of politics and independently of political guidelines. The DFG's core business is the funding of research projects in all disciplines, from the humanities and social sciences to the life and engineering sciences. In addition, the DFG also funds projects at major research libraries and archives.²

The Virtual National Catalogue for Digitised Books

In Germany there is no national catalogue of the academic state and university libraries — neither for printed nor for digital books. This has much to do with the political system in Germany and the autonomy of the German states in the areas of education and science. Today, there are six large union catalogues in Germany for the collections of the academic state and university libraries. The metadata of newly catalogued digital content are integrated in one of these six regional union catalogues.

For printed books, a 'virtual' national view of German library holdings was created by the Karlsruhe Virtual Catalogue³ many years ago. The Virtual Catalogue offers a meta search across the six regional union catalogues. The current version of the Virtual Catalogue is already able to identify digitally available titles, especially digital content from the six regional union catalogues.

DFG also funded a portal called the Central Index of Digitised Imprints.⁴ The idea was to offer a uniform, comprehensive view of public domain digitised books. In the early stages of the project, the libraries had to deliver their metadata to the portal. This procedure unfortunately was not crowned with success. For that reason the portal had to be restructured. Presently, the portal is on track to develop an innovative harvesting-based method of data collecting from the different German libraries, using the Open Archives Initiative Protocol for Metadata Harvesting.

The Very Best Design for a Viewer

In recent years, all libraries developed their own viewers for displaying digital material. The wheel was reinvented each time with great ambition. Where should the buttons for paging forward and backward be positioned? What background colour should be used? I am sure most readers can recall similar discussions in their institutions and countries. You will not find a single

presentation of digitised content at one library that is identical with that of another library in Germany. Yet we all are talking about international and networked information systems, not local special solutions. For this reason, a uniform web presentation style was developed, the so-called 'DFG Viewer'.⁵ All DFG funded digitisation projects are now asked to support that style. The viewer was originally designed for displaying digitised books. It can, however, also be used for all other materials, such as medieval handwritings.

It is not the DFG's intention to replace local solutions, but rather to supplement these with a nationally uniform one. The DFG viewer should always come into play when digital material is retrieved from national catalogues. With the click of the mouse the user can then be further guided to the local presentation environment of the libraries and use additional functionalities that are available there. Therefore I like to speak of the viewer as a minimum presentation standard on a national level.

Cross-Institutional Presentation

Libraries are well aware of the fundamental change in the structure of information provision that is taking place. Some libraries react to this situation with increased individualisation, profiling and the development of unique features. Unifying the existing range of viewers for digital books is impossible against this background. For this reason, a few libraries have, with support from the DFG, developed a format for data exchange, which has made it possible to display decentrally stored digital images in a centrally hosted presentation environment. The reference application was installed at the Saxon State Library of Dresden. It facilitates a uniform display of digital material in the DFG Viewer, independent of where the digital content was produced or where it is stored. The book that you page through in the DFG Viewer in Dresden can be physically stored on a server in Munich, Wolfenbüttel, Dresden, Halle or someplace else in the world. The presentation environment is always the same. How the Viewer functions is described in the DFG's 'Practical Guidelines on Digitisation'.⁶

(Ralf Goebel)

Technical Background

Before a national presentation interface for digitised books could even be considered, the wide range of data formats used in Germany had to be stan-

standardised and a technology decision had to be made on data transfer. Two considerations were important here: First, the intent was not to create a unique solution for Germany, but instead, international standards already in use should serve as a model. Second, the technical hurdles for libraries were to be kept as low as possible in order to increase the acceptance of the standards and to promote their adoption.

Data Format

Eventually the 'Metadata Encoding and Transmission Standard' (METS) and the 'Metadata Object Description Schema' (MODS) were selected as data formats. Both of them are maintained by the Library of Congress and are based on the extensible mark-up language of the World Wide Web Consortium. Therefore they are widely used and they can easily be implemented with the aid of familiar technologies. The formats are primarily designed for describing objects in digital libraries and thus are very well suited to our purpose.

An application profile⁷ supplements the official specifications and indicates which of the format's data fields are mandatory or optional in the DFG Viewer. A valid METS file can be generated with very little effort, since there are very few mandatory fields in the application profile. Nevertheless, there are still enough optional possibilities available to represent digitised books of nearly any level of complexity.

The viewer does not support every METS and MODS element defined in the official specifications, but each field not mentioned in the application profile is simply ignored. So if a library needs to use these fields — for example in the context of another application — it may do so without having to worry about the DFG Viewer.

Some Deeper Insights into METS ...

To illustrate this, I would like to give you some insight into the formats and their use in the context of the DFG Viewer. The METS file holds information on the physical and logical structure of the digitised book as well as details about the image files representing the book's pages. The descriptive and administrative metadata are encoded in MODS and embedded in the METS file (Figure 1).

Fig. 1: The METS file.



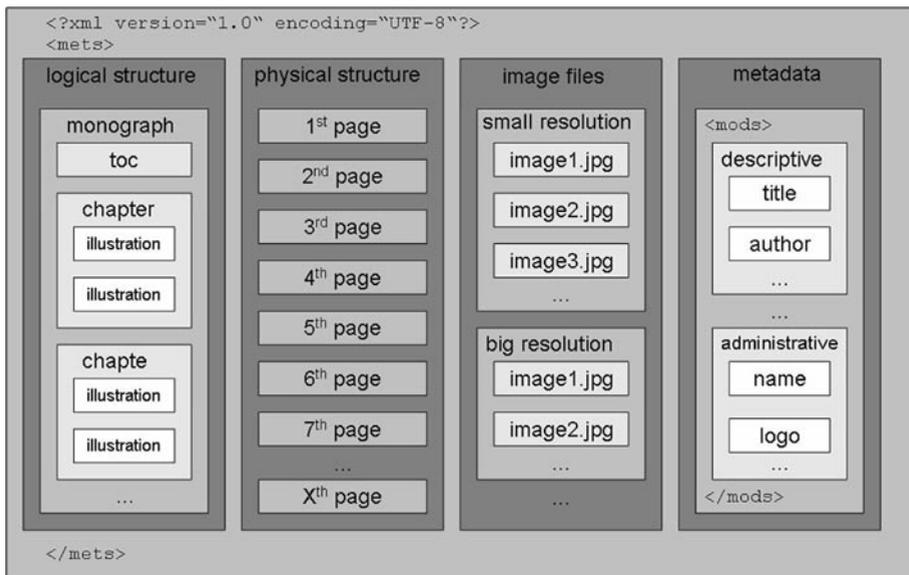
The physical structure represents the sequence of the pages in the book. The DFG Viewer uses this to obtain information on the number of pages, their order and pagination as well as persistent identifiers on the page level. The logical structure, on the other hand, comprises the chapters, tables of contents and other bibliographical structures with their corresponding labels from which the DFG Viewer generates navigation for the document. It also includes information on the document type and again any persistent identifiers of the structural units. Both structural representations are linked to one another in such a way that the respective physical pages are assigned to each logical structural unit. The information on image files includes details on the storage location, their file format and their context of use. The storage location must be an HTTP URL, since the images are always stored at the contributing library and are only retrieved by the DFG Viewer as needed. The context of use describes up to three different image resolutions and one additional file format which the DFG Viewer offers in the form of zoom levels or as downloadable file respectively. These details are also linked to the physical structure, allowing all corresponding images to be linked to each physical page.

The physical sequence of the pages and one image per page suffice as minimum requirements. This gives the DFG Viewer enough information to facilitate browsing through the digitised book. Only if the METS file includes optional information on the logical structure or on the various image variants does the DFG Viewer also display the document's navigation or buttons for zooming and downloading. Thus, each library which carries out digitisation can decide for itself on the complexity of its METS file and which functions of the DFG Viewer it would like to utilise.

...and MODS

The aforementioned information applied only to the METS standard. The *metadata* are described in MODS. These MODS data are, however, embedded in the METS file.

Fig. 2: MODS metadata embedded in the METS file.



The metadata comprise descriptive details. These include bibliographic information on the document, such as title, author and place of publication. Descriptive metadata can exist not only for the entire document, but also for each logical structure. For example, an illustration is to be associated with

a certain artist or a dedication is to be addressed towards a certain recipient. Thus, each logical structural unit can have its own MODS section with descriptive metadata. For the DFG Viewer, however, only the title, author, place of publication and year of publication are necessary.

The administrative metadata provide information on the originator of the digital material and the library-related context. The DFG Viewer evaluates information on name, logo, contact options and homepage of the originator and displays links to the catalogue entry and local presentation of the book at the corresponding library.

Data Transmission

While METS and MODS represent a suitable exchange format, they do not answer the question how data transmission takes place. An international standard had to be selected for this purpose as well. In addition to being easy to implement, the standard was to be able to transmit XML data without first requiring any data conversion. The standard chosen was the 'Protocol for Metadata Harvesting' of the Open Archives Initiative (OAI-PMH). Although the specification is based on the use of Dublin Core as metadata format, any other XML-based format can also be used.

An OAI-PMH interface allows specific datasets to be called up from a repository — exactly what is required by the DFG Viewer. Furthermore, the interface offers options for delivering a table of contents of all datasets contained in the repository and the sorting of the datasets in individual sets that, for example, may represent collections.

The OAI-PMH specification is also an international standard that is used by multi-national projects, such as Europeana. Therefore the German libraries did not need to implement a new interface especially for the DFG Viewer, but rather, they could use this in the context of many other projects as well.

Summing up

With METS/MODS and OAI-PMH, the standardisation of data format and transmission mentioned at the beginning has been attained not only on a national level, but for international connectivity as well. Thanks to the standardising force of the DFG as the largest provider of funding for digitisation

at German libraries, its standards are being adopted rapidly, and this goes for the DFG Viewer as well.

The viewer itself is operated as a web service at the Saxon State Library in Dresden. It queries the METS data via the OAI protocol for metadata harvesting and then interprets this information. Depending on the complexity of the METS data, it offers the user more or fewer options. These range from simply browsing the document to displaying citable identifiers on the page level, various zoom functions, a download option and extensive navigation functions on the content level.⁸

The DFG Viewer is, however, only a presentation interface, not a directory system or catalogue. Therefore, it does not possess its own data management functionality. Each time a record is retrieved, the Viewer must be explicitly told which digitised book is to be displayed and where it is located. Thus the DFG Viewer is best suited to be used in combination with catalogue and directory systems. These systems offer the user the opportunity to perform searches and can provide the DFG Viewer with all information necessary for displaying a digitised book.

Using the DFG Viewer

The DFG Viewer is a free web service and can be used by anyone without having to worry about license fees or local presentation software. It is implemented as an extension to the free Content Management System TYPO3 and is also available for download under the conditions of the General Public License. If desired, libraries can even operate their own instance of the Viewer. An integrated METS/MODS validator based on the application profile also makes it a practical tool for developers who wish to create METS files suitable for the DFG Viewer. Especially small libraries which cannot afford to host and maintain their own viewer for digitised books, can use the DFG Viewer as their primary presentation system. They just need to have a simple web server which provides the METS files and images.

The format and transmission standards have not yet spread throughout Germany, but continue to grow in acceptance. Because of the open architecture and the use of international standards, the DFG Viewer can also be useful in many other contexts. Readers of *LIBER Quarterly* are welcome to try the DFG Viewer and the authors would welcome your feedback.

What's Coming Next?

The next steps in the development of the DFG Viewer will be an improvement in the support of digitised manuscripts and the possibility to display full texts. Additional XML data formats are required for this purpose. These formats can be embedded in METS in the same way as MODS.

Further information can be found on the web pages of the Library of Congress⁹, the Open Archives Initiative¹⁰ and, of course, the homepage of the DFG Viewer.¹¹

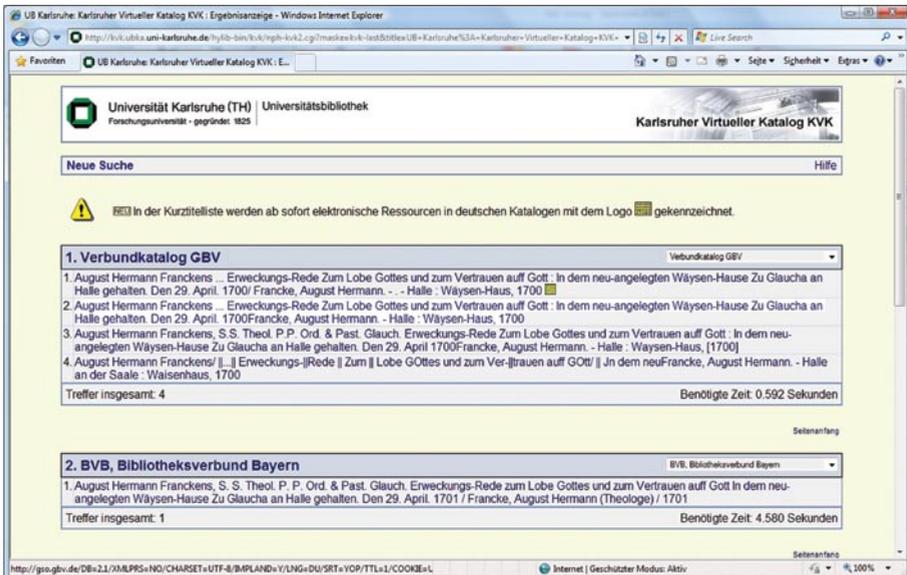
(Sebastian Meyer)

Example: from the Karlsruhe Virtual Catalogue to the Local Presentation

Step 1

Figure 3 is the result of a meta search across the German Union Catalogues. One entry is marked with a yellow symbol. This indicates an item that is accessible in a digital format. With a click on the entry the Northern Germany Union Catalogue is displayed (step 2).

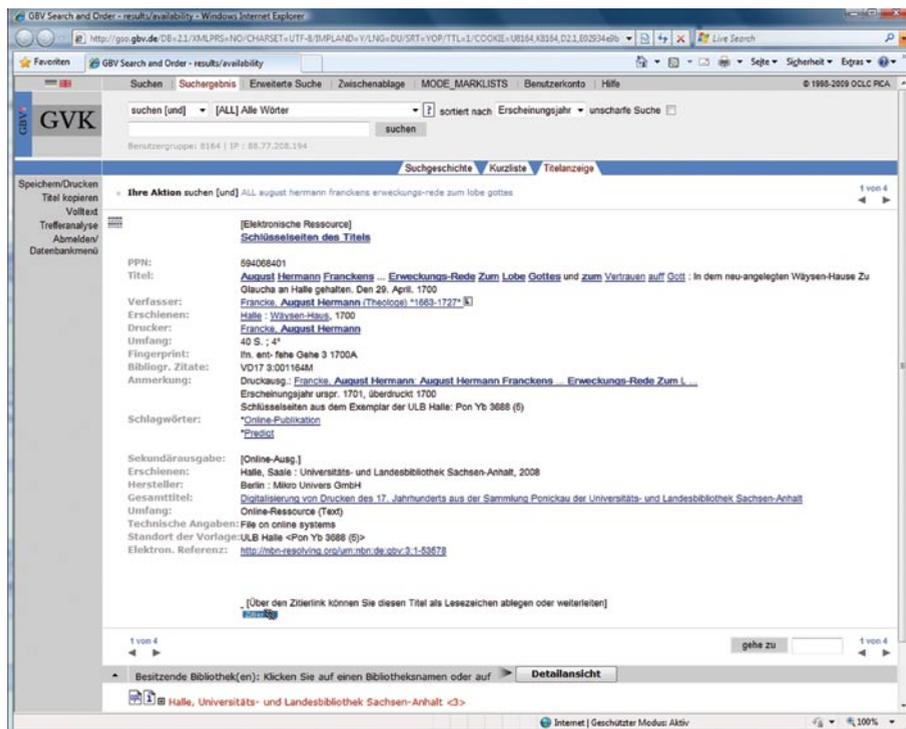
Fig. 3: Step 1.



Step 2

In Figure 4 the Northern Germany Union Catalogue (GVK) is displayed. One further click on the persistent URL below¹² shows the digitised book embedded in the environment of the viewer (step 3).

Fig. 4: Step 2.



Step 3

In Figure 5 the digitised book is presented in the environment of the viewer. The viewer is hosted at the Saxon State Library of Dresden, the image is located on a server at the Halle University and State Library. With a last click the local library viewer is displayed (step 4).

Step 4

Figure 6 shows the local library viewer.

Fig. 5: Step 3.

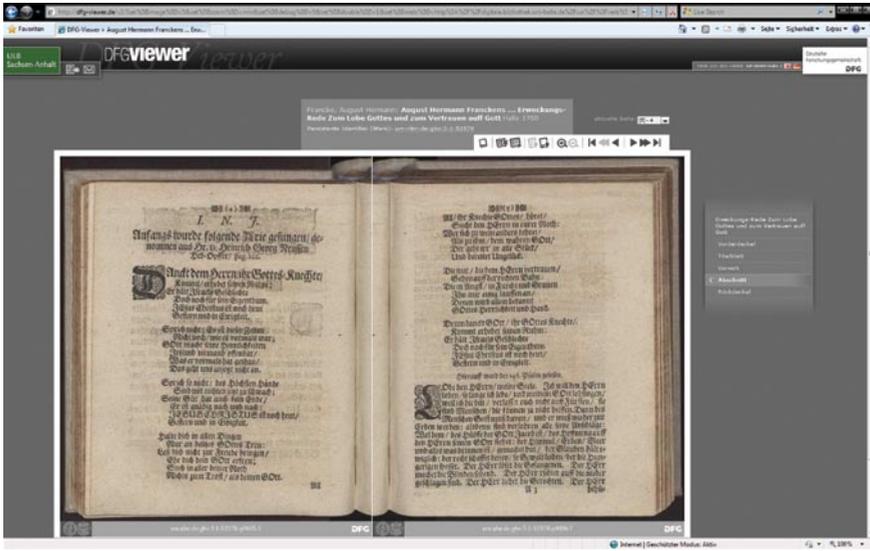
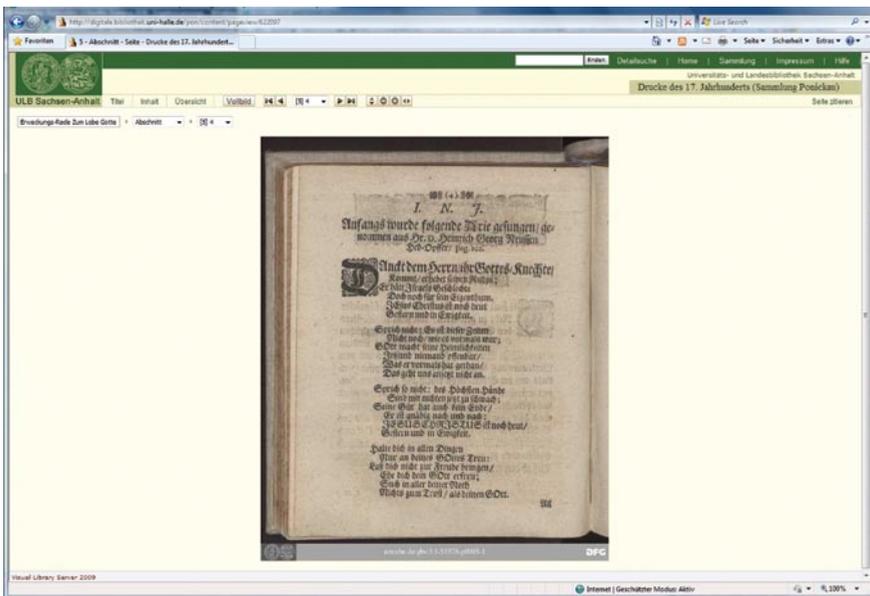


Fig. 6: Step 4.



(Sebastian Meyer)

Notes

¹ <http://www.dfg.de>

² <http://www.dfg.de/lis>

³ <http://www.ubka.uni-karlsruhe.de/kvk.html>

⁴ <http://www.zvdd.de>

⁵ <http://www.dfg-viewer.de>

⁶ <http://www.dfg.de/lis/en> ('Publications')

⁷ <http://dfg-viewer.de/en/profile-of-the-metadata/>

⁸ Examples: <http://dfg-viewer.de/en/demonstrator/>

⁹ <http://www.loc.gov/standards/>

¹⁰ <http://www.openarchives.org/>

¹¹ <http://dfg-viewer.de/en/>

¹² <http://nbn-resolving.org/urn:nbn:de:gbv:3:1-53578>