United, Linked, Connected – A Data Model for the Inventory of the Former Detmold Court Theatre (1825–1875), or: How Library Inventory History Can also Be Told

Kristina Richts-Matthaei
Paderborn University, Germany
kristina.richts@uni-paderborn.de

Irmlind Capelle
Paderborn University, Germany
irmlind.capelle@uni-paderborn.de

Abstract

Library forms of cataloging may differ greatly from the cataloging requirements of musicological research projects: They are often not detailed enough and do not take a close enough look at aspects of content relevant to research, such as handwritten entries in materials, etc. Library catalog entries of individual documents stand on their own for historical reasons, but usually do not reflect relationships to other surviving materials. This observation was the starting point for the Detmold Court Theatre Project, a six-year research project (September 2014 – January 2021) that looked at the interconnectedness of different surviving materials of a 19th century theatre company that existed from 1825 to 1875. This project dealt with a very detailed form of inventory indexing in order to hand over and make accessible the formerly related materials in their entirety. This form of indexing was called 'contextual deep indexing'. This special form of indexing took into account not only the pure performance materials but also the surviving theatre files, such as fee books, revenue and expense documents, stock lists, director's books, role and costume books, etc. All information was recorded based on autopsy (in the case of musical records on the basis of already existing RISM records from the 1980s). It was the first attempt to carry out such a form of indexing on the basis of the MEI and TEI encoding standards for a large repertory. For this purpose, a data model was needed that focuses on the linking of MEI and TEI data and enables the linking of different surviving library holdings, with a focus on a FRBR-based indexing of performance materials and associated performers as well as the structure of the theatre. Using a custom ODD-based schema, separate records were created for all works, expressions, and manifestations (in this case preserving the unity of materials kept under a common signature) and given unique identifiers so that they can now all be referenced individually.

The paper summarizes the results of this pilot project. It addresses the particularities and requirements of an inventory development that does not focus on individual objects, but on the relationship between different objects (and subjects). It presents a document-oriented (not object-oriented) data model that uses library materials to revive an entire network of a long-gone organization.

Introduction

The field of music research in the digital age is currently subject to substantial change. Libraries provide more and more of the materials preserved by them in digital form. They combine the digitized files with metadata sets or provide links to other catalog entries such as the RISM OPAC with its detailed descriptions of music manuscripts. In the best case, music researchers nowadays can get digitized material and bibliographic descriptions with just a few mouse clicks. With the current developments in the Digital Humanities, the potential for new research methods is also growing. This development is reflected primarily in editorial projects, such as Beethoven's Werkstatt,[1] and catalog projects, such as RISM[2] or Bach digital.[3] It has been accelerated by the

1 http://beethovens-werkstatt.de (accessed January 12, 2022).

2 http://www.rism.info (accessed January 12, 2022).

3 https://www.bach-digital.de (accessed January 12, 2022).

data format of the Music Encoding Initiative (MEI),[4] which offers promising opportunities for indexing metadata of musical holdings and the long term preservation of these data. One of the distinguishing features of the MEI XML schema recommendation is that it accommodates very comprehensive metadata in the file header. The abundance of data ranges from formal components like title information or names of persons, times and places of origin to comprehensive information about instrumentation, watermarks, the conservation status of historical musical sources, and much more. This is an essential prerequisite for the indexing of data of musicological research projects in an adequate depth. The schema also allows multiple degrees of integration of the linkage to digitized data and/or transcriptions. Extensive linkage and mapping possibilities create a connection to existing library data and thus enhance an extensive use of the indexed data. MEI also supports the encoding of multiple incipits, which provides significant opportunities for using mark-up that go far beyond the capabilities of other formats like Plaine & Easie,[5] for example by supporting multipart encodings, text underlays, and instrumentations.

The Detmold Court Theatre Project[6] funded by the German Research Foundation (DFG) was the first research project that used MEI as a data container for cataloging and for archival purposes at the same time. The most important aim of this project was to demonstrate the potentials of the MEI standard by a new method of cataloging on the basis of the holdings of the former Court Theatre, which existed from 1825 to 1875, and to develop an MEI and TEI based model of contextual deep indexing of musical collections. The remarkable thing about this collection, which is stored in the Regional Library of Detmold,[7] is that the material components of the Court Theatre have been retained in an astonishing variety and completeness. Additionally, almost all contextual archival materials (mostly written file material, such as revenues and expenditures, fee booklets, inventories, role and costume books, daily reports, theatre journals, etc.), which have been preserved in the Regional Library of Detmold as well. In the context of cultural history, these combined collections prove to be extraordinarily informative for cultural research.

The collection of performance materials from the days of the Detmold Court Theatre consists of 268 individual materials related to opera and incidental music, the most of them comprising a score, parts, scripts, libretti, role books, text books, director's books, and soufflé books. These materials were all cataloged with MEI based on autopsy to illustrate the advantages of more in-depth cataloging of materials in comparison with traditional catalog records, where the focus is always on a single document, and mostly on the physical properties of the document rather than the content. But what does an indexing actually require that does not prioritize the individual object but the relationship between different objects (here: documents) and subjects?

1 A Data Model for Contextual Deep Indexing

1.1 What Means 'Contextual Deep Indexing'?

The holdings of the former Detmold Court Theatre are nowadays kept in a very classical, library-specific manner: Each material is stored as an independent collection unit under a shelfmark that conveys to those familiar with the collection that it belongs to the Court Theatre collection. In the boxes typical for this collection, all the handwritten and printed performance materials used by the former theatre company are found in close proximity to each other. Taken together, they tell their stories [2] – each material separately, but also all the materials as a whole. Stories about plays that were performed, stories about how these plays were performed, and – related to this – stories about people who were significantly involved in these performances. But: Their history remains hidden in the materials if it is not coherently reconstructed.

This history of the former Court Theatre and its employees and participants can of course be compiled by looking through the materials in the reading room, as it has been done for decades, but this approach would be rather tedious, time-consuming, and also not very practicable when dealing with certain questions, since there are no less than 268 materials. Establishing (content-related) cross-links between the materials in this

4 https://music-encoding.org (accessed January 12, 2022).

5 In librarian practice, the Plaine & Easie Code is used for the encoding of musical incipits.

6 "Detmolder Hoftheater. Entwicklung eines MEI- und TEI-basierten Modells kontextueller Tiefenerschließung von Musikalienbeständen am Beispiel des Detmolder Hoftheaters im 19. Jahrhundert (1825–1875)", https://hoftheater-detmold.de (accessed January 12, 2022).

7 http://www.llb-detmold.de (accessed January 12, 2022).

way is therefore really out of the question and could only be done on the basis of a selection. In addition, the other surviving sources, such as file materials or director's and costume books, would have to be examined and added to the stock of performance materials in order to obtain further valuable information about the performances.

The Detmold Court Theatre project therefore dealt with a completely new form of inventory cataloging to show how combined library and musicological indexing could look like in the long term. Based on various factors, all the surviving materials of the former Detmold Court Theatre were subjected to a kind of 'overall view', i.e., each was looked through individually, re-recorded, and related to each other on the basis of certain recurring parameters such as personal names, work names, role names, or (performance) dates. However, this approach first required the creation of a digital environment in which the materials were recorded and described in a way that enables cross-connections between them. This required individually and unambiguously referenceable objects, a profound data model, and a common presentation interface that brings out the connections between the individual objects.

According to our understanding, 'contextual deep indexing' thus means: to record the information of the individual sources with content mark-ups so deeply that they can be linked and related to other objects.

1.2 The Data Model

In the field of digital editions, it has become standard practice to make contextual material available as full text and to link it via markup (see the presentation of Freischütz Digital,[8] which is exemplary for the field of music theatre). However, in the field of indexing collections, proprietary databases or library standards are still used to a great extent and cannot be made accessible on the web without further ado, and above all, different types of sources are recorded separately in their own systems. This is evident, for example, in the current discussion in the library sector about the recording of ephemera [4, 5]. Especially in the field of the indexing of playbills, which is important for theatre research (see, for example, the Düsseldorf Playbill Database developed at the University and State Library of Düsseldorf,[9] the Weimar Playbill Database[10] or the Bremen Playbill Database[11]), this has led to numerous isolated solutions, which makes a comprehensive search – e.g., for actor names – impossible.

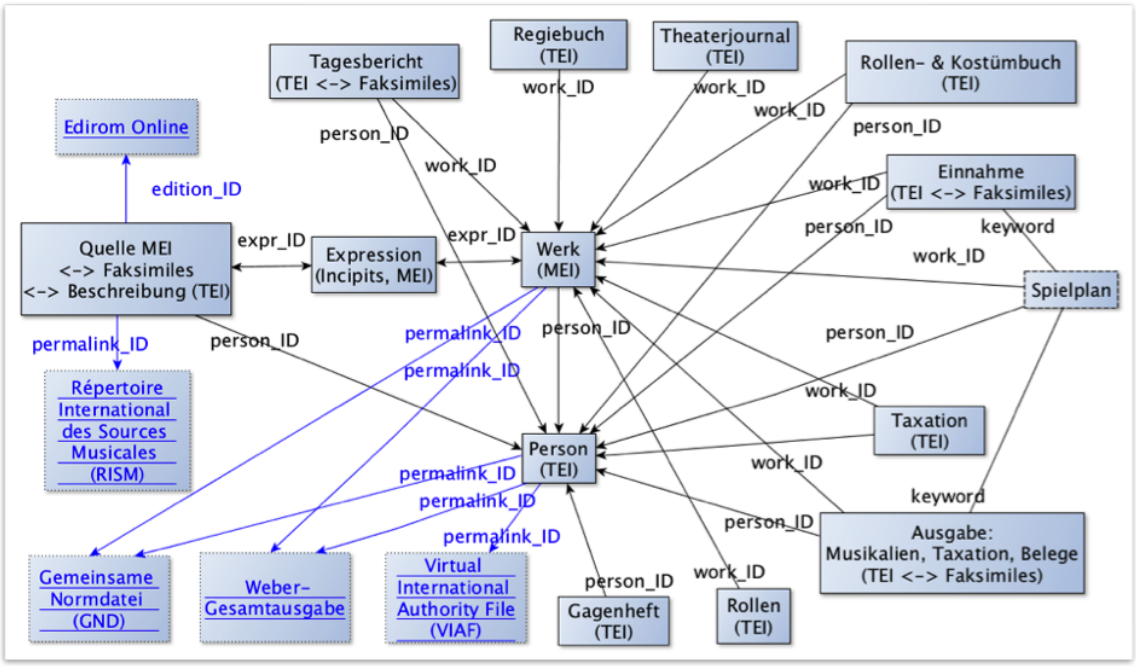

Figure 1: Data model of the Theatre Tool.

8 https://freischuetz-digital.de/ (accessed January 12, 2022).

9 http://digital.ub.uni-duesseldorf.de/theaterzettel (accessed January 12, 2022).

10 http://www.theaterzettel-weimar.de (accessed January 12, 2022).

11 https://brema.suub.uni-bremen.de/suubtheater (accessed January 12, 2022).

The Court Theatre Project has developed a data model for the contextual deep indexing of the various materials that have been preserved.[12] This data model (see Figure 1) forms the basis for the "Theatre Tool" software that was specifically developed for the presentation of the data.[13] Starting from the performance materials, which were initially split into works, expressions, and manifestations according to the FRBR principles,[14] references to all the persons and works mentioned in other materials were systematically entered. As a result, one can now see which person played which roles in which works, which people were part of the theatre company in which years, what the playbills looked like in the individual years, at which venues which plays were performed, etc. In contrast to the full FRBR model, via which relationships to persons and objects can of course also be mapped, only the entities of FRBR group 1 (work, expression, manifestation, item) were used in this data model. The individual persons were referred to with the help of identifiers; more detailed descriptions of the persons were collected in separate data records, with reference to already existing authority data records (if available). The roles or functions performed were recorded in MEI with the help of the `@role` attribute, which was semantically sufficient in this case.

By using the FRBR model for works and expressions of the work, the data collected meet both library and academic requirements. The indexing of the sources is carried out according to FRBR on three different levels: The work files record the basic data, if applicable with the date of first performance and a standardized indication of classification. The source files (corresponding to the FRBR entity manifestation) describe the available sources, which in this case are combined into a 'componentGroup', since the performance materials form a unit.

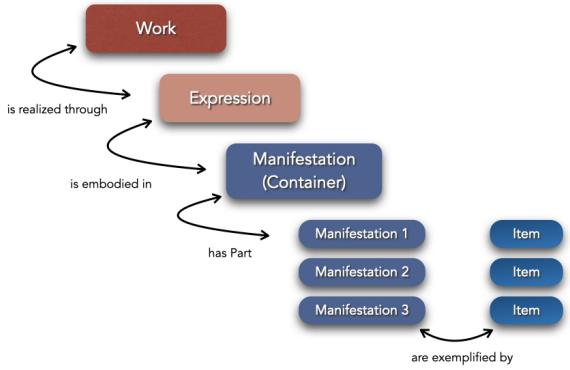

Figure 2: FRBR-based indexing of the musical records.

The link between work and source is always the expression file, because the respective performance material of the theatre in Detmold is as a unit an expression of the work, while that of another theatre would be another expression. An adaptation of an opera for wind ensemble, for example, would also be another expression of the same work. The relationships between the files are described with relations as specified by FRBR: `hasRealization`, `isEmbodimentOf`, `hasEmbodiment`, `isPartof`, etc. (see Figure 2).

In addition to these files, which are necessary for source indexing, files for persons and *dramatis personae* were created. By means of a unique ID for each file, clear identification is possible in case of any recurrence of a work, role, or person name. While these files are typical catalog indexes which, at least in the case of the source files, go far beyond the usual library indexing, the extensively handed down contextual material is indexed partly as registers, but mainly in full text. Both forms of indexing are based on the XML standards TEI and MEI, so that all data can be marked up (see Figure 3).

In addition, authority files (GND, VIAF, GeoNames) are used for persons, works, and, if applicable, places, so that external information can be integrated. However, since many persons and works are not well known or cannot be clearly identified and thus cannot be clearly assigned to an authority data ID, the use of project-specific IDs remains necessary.

12 https://hoftheater-detmold.de/47-2/das-modell/ (accessed January 12, 2022).

13 https://github.com/Detmolder-Hoftheater/TheaterTool (accessed January 12, 2022).

14 https://www.ifla.org/publications/functional-requirements-for-bibliographic-records (accessed January 12, 2022).

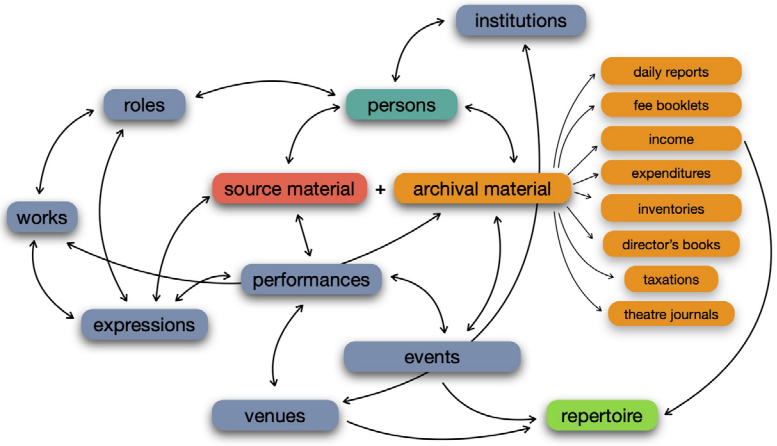

Figure 3: Data relations between the different entities of the indexed materials.

Linking the data across all source boundaries results in various connections in terms of content: Thus, for the persons of the Detmold Court Theatre, on the one hand, the data on salary and possible special allowances, length of employment, and additional occupations in the theatre business can be retrieved, and on the other hand, the works and even the roles in which these persons were employed. For the performance materials, information on dating and scribes is linked from the files, and the entries in the costume and director's books provide the first clues to the presentation of individual works on stage.

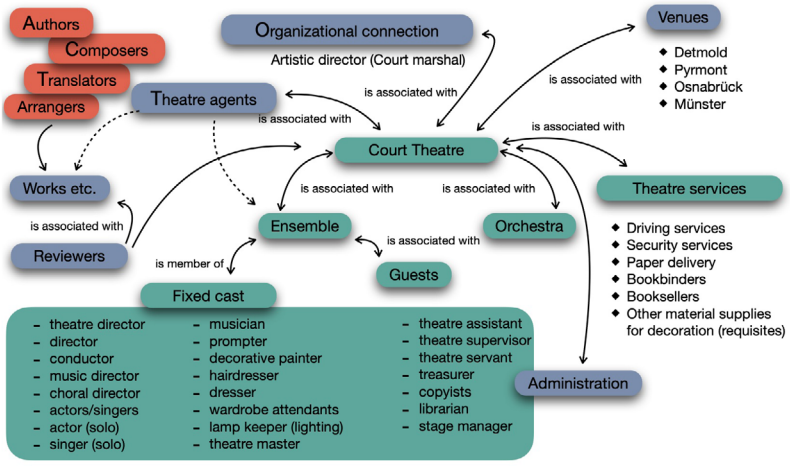

Figure 4: Personal relations in the Court Theatre Network.

In summary, it can be said that the theatre network (see Figure 4) is a network which focuses on the individuals and relates them to their activities that can be traced back in the surviving sources. This sets it apart from other, event-based[15] or performance-based[16] models that are thought of in terms of the individual event or the individual performance. Individual events only appear in the theatre programs from 1820 to 1847, which could be reconstructed on the basis of the theatre's surviving revenues and are now generated by the Theatre Tool on the basis of TEI indexing.

15 https://performance.slub-dresden.de (accessed January 12, 2022).

16 A discussion on presentation models of performance-based informations took place at the DHd Conference 2020 "Spielräume: Digital Humanities zwischen Modellierung und Interpretation", March 2–6, 2020, in the panel of Kathrin Dennerlein "Datamodelling Drama and (Musical)theater", see [3].

2 Presentation of the Data/Visualization

All data collected in the project context are visualized with the help of the software Theatre Tool and presented in the Detmold Court Theatre Portal.[17] This is based on XQuery and JavaScript. Within the portal, information is compiled and displayed in bundles with the help of references that query the assigned IDs of individual person, work, and role data sets. For example, users can search specifically for individual works or select them from the repertoire list and receive a work record as a result, which provides references to those areas of the portal in which the work occurs. For example, Figure 5 shows the references for the work *La Dame Blanche*, including links to the theatre program, income, expenditure, theatre journals, director's books, taxations, daily reports as well as role and costume books.

If facsimiles are available for a work or if performance materials have been indexed, the user will see this to the right of the work information. General information on the material as well as descriptions of the individual sources are displayed under the respective library signature. The expression file connecting work and manifestation is not visible, but contains the incipit encodings from which the incipit displayed in the portal is generated.

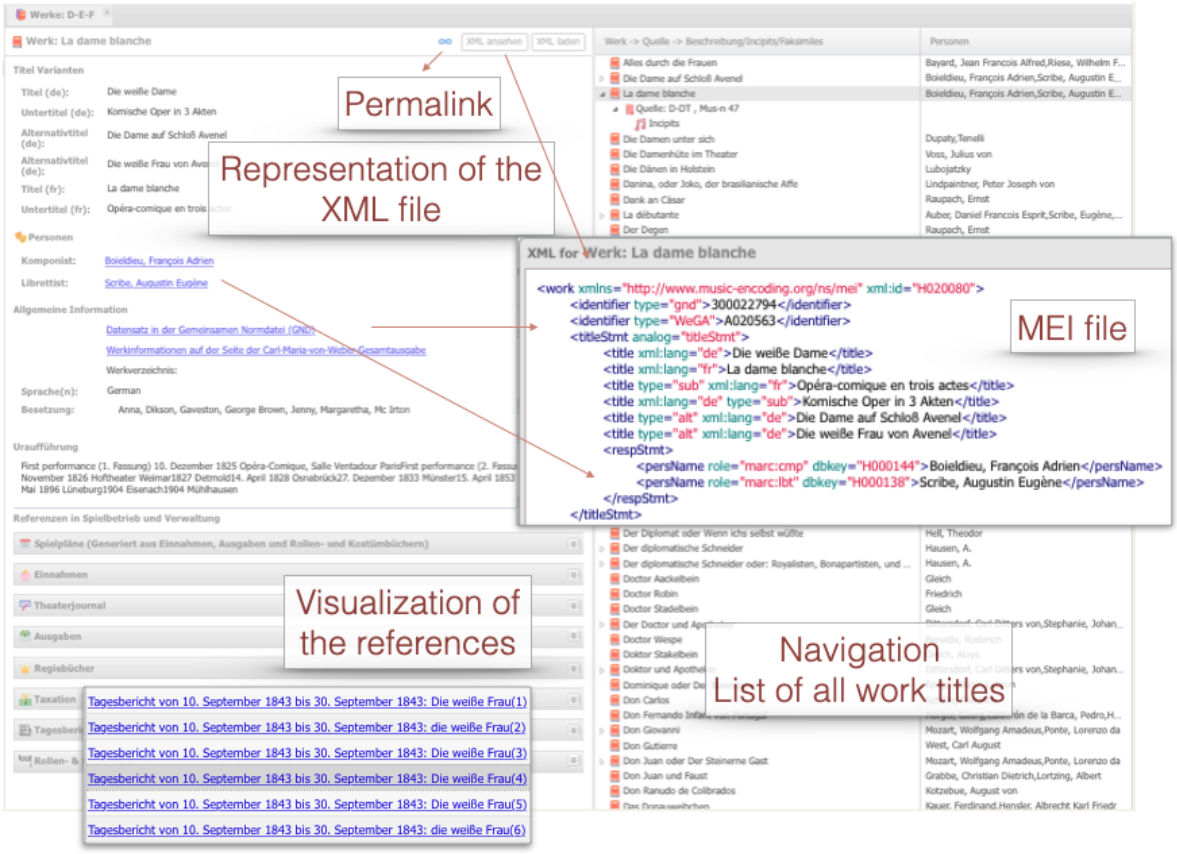

Figure 5: Visualization of references to the work *La dame blanche*.

The facsimiles are displayed with the help of a Leaflet plug-in,[18] a library for displaying interactive maps on the web. So far, the software offers a simple search for persons, roles, and works, created with a Fuse.js plugin, a fuzzy-based library.[19] As recommended in web presentations, the contents can be downloaded as XML files to enable further work with the data (search, query in a larger context, etc.). Of course, the data can also be used as examples for indexing in other projects. The works, sources, persons, and roles can be directly referenced by other projects with the help of permalinks.

17 https://dev.hoftheater-detmold.de/index.html (accessed January 12, 2022).

18 https://leafletjs.com (accessed January 12, 2022).

19 https://fusejs.io (accessed January 12, 2022).

Since the Detmold Court Theatre Project has so far mainly indexed materials on music theatre, some music-specific applications have been integrated into the software. For example, the beginnings of individual musical numbers are reproduced with note incipits to make them quickly comparable. In order to provide the musicologist with information on the original score arrangement, clef, spelling of the instruments, etc., not only a voice or piano score is reproduced – as is traditionally the case –, but the first measures and the vocal part are reproduced in full score. The encoding of the musical incipits is done with MEI, the presentation with a Verovio[20] plug-in (see Figure 6).

Another special feature of the project is the exemplary deep indexing of some selected performance materials: For these, the facsimiles of the sources are also made available, in a format that allows access to the exact measure. Only this form of indexing makes it possible to indicate interventions in the musical text not on the basis of materiality (e.g., deletion on p. 4v to 5r penultimate measure) but on the basis of content (e.g., deletion of mm. 17–20 in no. 1) and thus in a comprehensible way for the user. The software Edirom is used for an automatic identification of measures, and the Theatre Tool is linked to Edirom Online[21] for the presentation. Although Edirom was developed for the preparation of musical material, it can also be used to map text sources, e.g., by scenes or even lines.

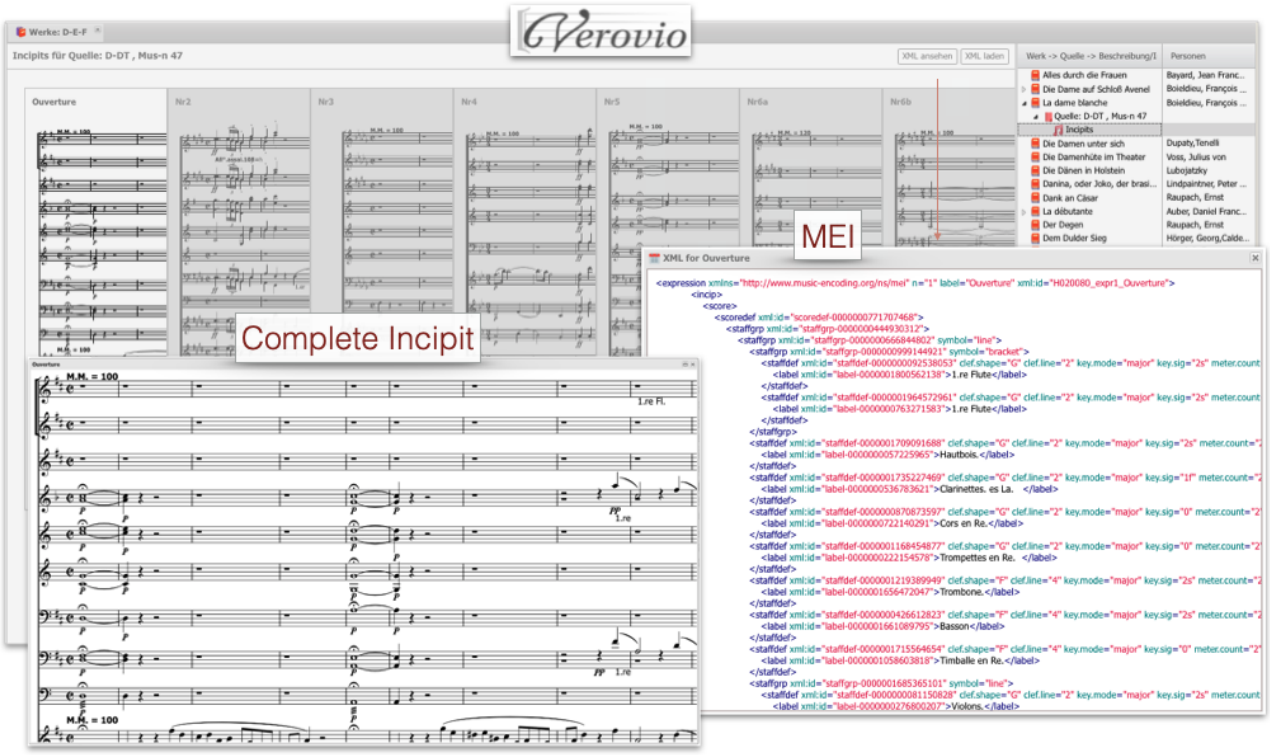

Figure 6: Visualization of musical incipits.

The Theatre Tool is designed to represent these complex text and data structures but can easily be adapted to other requirements: For example, the material recorded in the project is predominantly handwritten material, which is why the fourth level envisaged by FRBR, the copy (item), is not taken into account according to the 'manifestation singleton' rule.[22] Of course, this level could also be represented. Since the main interest of the indexing lies in the working methods and personnel of the Detmold Court Theatre Society, the locations mentioned in the sources are marked, but there are no independent files for them (with the possibility of references) and so far no search option.

20 https://www.verovio.org/ (accessed January 12, 2022).

21 https://github.com/Edirom/Edirom-Online (accessed January 12, 2022).

22 Manifestation singletons are at the same time source and item, because there is only one copy of them, e.g., autographs.

With the increasing digitization of library collections, the digitized files could be integrated into the Theatre Tool via IIIF, which would solve a number of legal problems. How these can be synchronized with the sources indexed in terms of content still needs to be examined.

3 Contextual Deep Indexing: Needs and Perspectives

Based on project experience, it can be said that there is a need for further coordination between science and libraries, but this is of great interest to both sides. It is clear that such a form of contextual deep indexing as was carried out in the project described here as well as the work on composer work indexes, for example, can only be undertaken by specialists. Nevertheless, there is great interest in making this detailed information on individual sources accessible via the owning libraries as well. The use of standards and authority files as tested in the Court Theatre Project is a first step in this direction, but more thought must certainly be given to interfaces for data exchange. There is a great need for interfaces for (automated) data exchange, especially on the part of research projects. The connection to library workflows is still completely lacking here. The heterogeneous library and catalog landscape found in Germany, for example, does not simplify things, but instead complicates them. In the course of the project, we also noticed that it is not so easy to obtain appropriate work authority files, not to mention data sets for individual expressions of a work. It was a particular concern of ours to link all the works in the repertoire of the Detmold Court Theatre with authority files. In reality, however, we had to learn that only a small part of our work datasets could be linked with good conscience to already provided data sets, and that (in the case of Germany) it is also not so easy to contribute data to the Integrated Authority File (GND)[23] as a research project because there are no workflows for automated processes for this so far. So once the technical hurdles have been overcome, we hope that we will still be able to contribute a great deal to the expansion of the GND Werknormdaten after the end of the project.

A similar problem exists with regard to the handing back of RISM data. Thus, for source indexing, we were thankfully able to make use of the LOD datasets provided by RISM, converted them to MEI and enriched them with many, many details on the individual source components. But how can we return them? And if we do return them, are the data structures on the RISM side sophisticated enough to fully present such data depths? Here, the question must be raised how cooperation can be organized in the future, and whether it would not make absolute sense, at least in the case of the work and expression files, to obtain authority files for works and expressions as a research project via RISM.

And there are other areas that should be improved in the cooperation between library and science in the future: We noticed that library and source indexes are not structured enough. Almost all information important for research (e.g., information on scribes, reflections or references in the literature on dating, references to supplements) is described, if at all, in unstructured free-text fields.

When using MARC codes, for example, the codes for professions in the field of music are not differentiated enough, and the possibilities for 'classification', which are obligatory in the library field, are either very general or more exact than can be determined by specialists. Instead of a 'classification', researchers need the original genre designations in order to be able to describe the individual terms with sufficient precision.

However, many developments can also be highlighted positively: With *musiconn.performance*, a powerful platform for searching events is being created, to which we will contribute our events from the theatre programs. With IIIF technology, numerous rights problems regarding the use of digitized material are solved, and the GND is slowly opening up with regard to the joint development of authority data.

Conclusion

The data of the Court Theatre Project are to be considered an important contribution to the research of the German theatre landscape in the 19th century: They are open and freely available, clearly structured and comprehensible, because they are available in TEI or MEI 4.0, including approximately 900 score incipits, and are made available not only in the project's web portal but also on Zenodo [1]. They are ready for evaluation, be it classical or digital, for musicological, theatrical, historical, or sociological research.

23 https://www.dnb.de/EN/Professionell/Standardisierung/GND/gnd_node.html (accessed January 12, 2022).

Works Cited

[1] Capelle, Irmlind, Kristina Richts-Matthaei, and Elena Schilke. *Detmolder Hofhteater 1825–1875*. Version 1.0 [Data set], June 29, 2021, https://doi.org/10.5281/zenodo.4903639 .

[2] Capelle, Irmlind, and Joachim Veit. "Was verbirgt sich in den Kisten? Materielle Zeugen zum Reden gebracht: Der Bestand Mus-n 245 zum *Freischütz* in der Lippischen Landesbibliothek Detmold" in *Musikwissenschaft der Vielfalt. Festschrift für Rebecca Grotjahn zum 60. Geburtstag*, ed. Joachim Iffland, Johanna Imm, Nina Jaeschke and Sarah Schauberger. München: Allitera, 2021, 42–76, urn:nbn:de:hbz:51:1-25587.

[3] Dennerlein, Katrin, Birk Weiberg, Dominik Stoltz, Katrin Bircher, Klaus Illmayer, and Gesa zur Nieden. "Datamodelling Drama and (Musical)theater" in *DHd 2020. Spielräume: Digital Humanities zwischen Modellierung und Interpretation. Book of Abstracts*, 2020, 59–62, https://doi.org/10.5281/zenodo.3666690.

[4] Pernerstorfer, Matthias J., ed. *Theater – Zettel – Sammlungen*, vol. 1: *Erschließung, Digitalisierung, Forschung*; vol. 2: *Bestände, Erschließung, Forschung* (Bibliographica 1, 2). Wien: Hollitzer Wissenschaftsverlag, 2012–2015.

[5] Wiermann, Barbara. "musiconn.performance – Musikalische Ereignisdaten im Fachinformationsdienst Musikwissenschaft" in *Kooperative Informationsinfrastrukturen als Chance und Herausforderung: Festschrift für Thomas Bürger zum 65. Geburtstag*, eds. Achim Bonte and Juliane Rehnolt. Berlin, Boston: De Gruyter Saur, 2018, 398–415, https://doi.org/10.1515/9783110587524-041.