Best Practice Recommendations for Work Catalogues based on MEI

Authors: Kristin Herold, Kristina Richts

July 2018

Preface	2
Introduction	2
Current state of work catalogues	2
Vision: A meta search for work catalogues	3
Systematic elaboration and realization	3
Data structure	3
Authority Files	5
Persons	5
Institutions	6
Geographical Names	6
Roles	6
Instrumentation	7
Proposal of core information in an MEI work catalogue	7
Person (Composer+, Librettist*, Lyricist*, other Persons*)	7
Identifiers (+)	8
Title (+)	8
Classification (+)	9
Instrumentation (*)	9
Date (*)	10
Example Coding for capturing the core information in MEI	11
Enhancement	12
Uncertainties	12
Refinement of instrumentation	12
Harmonization of events	13
Conclusion	14
Appendix	16
Sample for work catalogues	16
Overview: categories considered within work catalogues or complete edition projects	18

Preface

The best practice recommendations for work catalogues given within this document summarise the results of the discussions in working group 1 (work catalogues) during the MEI Metadata Workshop 2017 (November 7th to 9th) at the Academy of Sciences and Literature, Mainz. Many thanks go the participants of this working group: Hyacinthe Belliot (CESR Tours), Vincent Besson (CESR Tours), Margrethe Bue (Norwegian National Library), Gabriele Buschmeier (Gluck Gesamtausgabe), Katrin Eich (Brahms Ausgabe), Franz Kelnreiter (Mozarteum), Clemens Gubsch (Austrian Academy of Sciences), Claudia Heine (Strauss Gesamtausgabe), Johannes Kepper (Beethovens Werkstatt), Klaus Rettinghaus (Bach-Archiv Leipzig), Yuliya Shein (Gluck Gesamtausgabe), Christine Siegert (Beethoven Haus Bonn), Joachim Veit (Weber Gesamtausgabe) and Frank Ziegler (Weber Gesamtausgabe).

Introduction

Current state of work catalogues

One of the most common use cases for metadata in music are work catalogues which have a long tradition in music cataloging and editing. Many digital projects actually use MEI for this purpose already, often relying on the MerMEId software. It turned out that the different projects currently use varying data models, but it is possible to map the current contents to FRBRized structures.

The goal of the working group was to determine a core set of information that all work datasets have in common – regardless of the project they come from. To identify this core set, the working group agreed on the following steps:

- 1. examination of
 - a. existing uses of MEI for work catalogues or suitability of MEI for "non-MEI" work catalogues,
 - b. existing data models based on other formats,
 - c. traditional (printed) work catalogues,
- 2. comparison of the results of the preceding examinations and identification of commonalities and differences,
- 3. drafting of a proposal for best-practice recommendations on how to model work catalogues in digital form and
- 4. creation of a sample MEI file implementing the proposed model.

The working group also addressed the use of relevant authority files for work-related metadata.

¹ See the examination results of five work catalogues in the <u>appendix</u>. There is no claim for completeness.

Vision: A meta search for work catalogues

The objective of the above-mentioned effort is to reach a common standard for data acquisition in work catalogues created with MEI. In the longterm, a uniform documentation of work-related metadata is particularly advantageous as information can easily be compared, queried or reused by third parties. Based on this, one prospective main goal could be the development of a meta search for work catalogues inspired by the model of correspSearch². In order to implement such a search engine it is necessary to write mappings for existing work catalogues and to adapt (different) structures. This assumes that all of the concerned parties are willing to adopt and implement the best practice recommendations that have been coordinated by the MEI community. It should be clarified if there are legal concerns regarding data interchange in general.

Systematic elaboration and realization

Data structure

The analysis of existing work catalogues showed that there are some discrepancies in the ways of recording work-related metadata, but mainly on a structural level because the projects use different data models. On a content level the different projects agreed on a core set of information for the description of musical works, which includes

- information on ways of identifying a work (work numbers etc.),
- title information,
- information on the persons involved (composers, librettists etc.),
- information concerning the genesis of a musical work (including the date and context of creation),
- information on relations to other works of the same composer as well as
- a classification of the work.³

Some other kinds of metadata that were also mentioned don't concern the "work" directly but rather expressions of the work (musical scores, parts, libretti etc.), for example

- incipits,
- information on cast lists and instrumentation,
- information on tempo, key and metric specifications or
- information on the specific language which is used in individual expressions of the work etc.⁴

In order to develop a clear structure and to facilitate the exchange of work-related data between different projects, it is recommended not to confuse the level of work with the level of expressions of the musical work or rather information on physical sources, i.e. a specific manuscript including sketches of the composer.

² CorrespSearch is a web service that offers a meta search in scholarly editions of letters; https://correspsearch.net/

³ An overview of the project specific criteria can be found in the appendix.

⁴ ibidem.

Therefore, it is recommended to rely on the model of the Functional Requirements for Bibliographic Records (FRBR)⁵ that has been implemented in MEI because it provides very detailed options for a specified storage of metadata.⁶ The model addresses the following four main levels:

- 1. work,
- 2. expression,
- 3. manifestation and
- 4. item.

In MEI core information concerning the four FRBR levels *work*, *expression*, *manifestation* and *item* is encoded within the following elements:

- WORK: <workDesc> <work>
- EXPRESSION: <workDesc> <work> <expressionList> <expression>
- MANIFESTATION: <fileDesc> <sourceDesc> <source>
- ITEM: <fileDesc> <sourceDesc> <source> <itemList> <item>

Due to historical backgrounds in the development of MEI the structure is like written above; but following a consistent implementation of the FRBR model it would be better to have <sourceDesc> and <workDesc> on the same level.

Following the FRBR model *work* and *expression* related metadata in MEI are recorded with the elements *<work>* and *<expression>*. Both act on an abstract level, whereas the levels of manifestation and item describe the physical objects. As information on expressions of a work is included in the MEI element *<work>* (*<workDesc> <work> <expressionList> <expression>*), a direct connection of the two levels of information is given. Information encoded on the level "work" is subsequently inherited.

⁵ https://www.ifla.org/publications/functional-requirements-for-bibliographic-records

⁶ The use of the FRBR model in projects working with MEI has been carried out very successful in the Detmold Court Theatre Project (http://www.hoftheater-detmold.de), and in the Catalogue of Carl Nielsen's Works (http://www.kb.dk/dcm/cnw/navigation.xq), created with the MEI Metadata Editor MerMEId, developed by the Danish Centre for Music Editing (DCM) of the Royal Danish Library in Copenhagen et al.

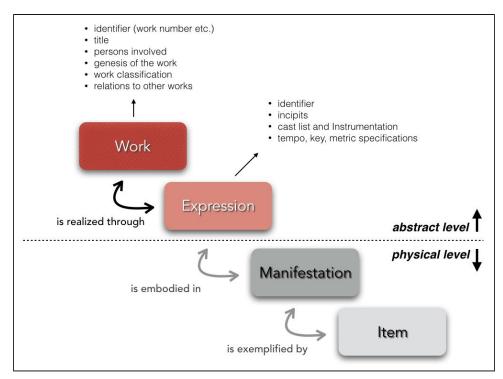


Figure 1: Four levels of description in MEI

Authority Files

To ensure a uniform data acquisition and to increase the connectivity to other projects it is highly recommended to work with authority files systematically. There are accepted authority files for:

- persons
- institutions
- geographical names
- roles
- instrumentation
- etc.

Persons

Regarding the encoding of personal names it has proved positive to link this data to one of the following authority files:

- Gemeinsame Normdatei (GND)⁷
- Virtual International Authority File (VIAF)8
- International Standard Name Identifier (ISNI)⁹

⁷ http://www.dnb.de/DE/Standardisierung/GND/gnd_node.html

⁸ https://viaf.org

⁹ http://www.isni.org

It is highly recommended to enrich each personal name with the attributes @auth, @authURI and @codedval. The @codedval attribute should contain the identifier of the person within the linked authority file. See the following example:

Institutions

Regarding the encoding of institutional names it is recommended to enrich the <corpName>-Element with the attributes @auth, @authURI and @codedval. An adequate authority file for institutional names can be the Gemeinsame Normdatei of the German National Library.

Geographical Names

Geographical names can be refined with links to databases that hold geographical names, for example GeoNames¹⁰. Similar to personal or institutional names those details should be encoded by the use of the attributes @auth, @authURI and @codedval.

```
<geogName auth="GeoNames"
    authURI="http://www.geonames.org/" codedval="2643743">London</geogName>
```

Roles

Regarding the encoding of personal names it is highly recommended to specify them with information on specific roles or functions these persons had in a process, i.e. the creation of a work. Currently, roles can be recorded best using the MARC Code List for Relators¹¹.

```
<respStmt>
     <resp auth="MARC List of Relator Terms"
          authURI="http://id.loc.gov/vocabulary/relators/" codedval="cmp"/>
          </respStmt>
```

¹⁰ http://www.geonames.org/

¹¹ https://www.loc.gov/marc/relators/

Instrumentation

As instruments can have many different names a recording of instrumentation should always include values that accurately identify them. This can be done using the MARC Instruments and Voices Code List¹². Another recommended list is provided by the ufficio ricerca fondi musicali¹³.

```
<perfResList>
    <head>Instrumentation</head>
    <perfRes authURI="https://www.loc.gov/standards/valuelist/marcmusperf.html"
        count="2" codedval="wb">Ob</perfRes>
    <perfRes authURI="https://www.loc.gov/standards/valuelist/marcmusperf.html"
        count="2" codedval="bb">Tr</perfRes>
    <perfRes authURI="https://www.loc.gov/standards/valuelist/marcmusperf.html"
        count="2" codedval="ba">Hr</perfRes>
    <perfRes authURI="https://www.loc.gov/standards/valuelist/marcmusperf.html"
        count="2" codedval="sa">Vl</perfRes>
    <perfRes authURI="https://www.loc.gov/standards/valuelist/marcmusperf.html"
        count="1" codedval="sb">Va</perfRes>
    <perfRes authURI="https://www.loc.gov/standards/valuelist/marcmusperf.html"
        count="1" codedval="sb">Va</perfRes>
    </perfRes authURI="https://www.loc.gov/standards/valuelist/marcmusperf.html"
        count="1" codedval="sd">Bc</perfRes>
    </perfResList>
```

Proposal of core information in an MEI work catalogue

The following core information data fields were agreed on at the Metadata meeting mentioned above to be the minimum set of information, that is needed to identify a work without any doubt and which should be available or able to be obtained for every musical work.

```
(explanation: * means 0 or more; + means 1 or more)
```

```
Person (Composer+, Librettist*, Lyricist*, other Persons*)
```

If possible, enrich each personal name with the attributes @role, @authURI and @codedval. For roles of persons see the MARC Code List for Relators¹⁴.

¹² https://www.loc.gov/standards/valuelist/marcmusperf.html

¹³ http://www.urfm.braidense.it/risorse/medium_query_en.php

¹⁴ http://www.loc.gov/marc/relators/relaterm.html

```
codedval="118629662">Kind, Friedrich</persName>
    </respStmt>
    </titleStmt>
    </work>
```

Identifiers (+)

Use **Identifiers** for <frbr:work/> and <frbr:expression/> separately. The working group agreed on the following recommendation: every work and every expression must have an xml:id, works starting with a "w", expression starting with an "e".

```
<work xml:id="w_H65753"/>
```

Additional identifiers from authority files or other work-related directories, that also reference the work should be coded by the use of a separate <identifier>-element. If possible, the different identifiers should be specified with information about the source the identifier comes from. This is to be done best by the use of the @auth and @authURI attributes. A work can have several identifiers from different directory systems. See the following examples:

```
<work xml:id="w_H65753">
    <identifier auth="Jähnsverzeichnis" label="JV">277</identifier>
    </work>
```

```
<work>
    <identifier auth="GND" authURI="http://d-nb.info/gnd/">99541727X</identifier>
    </work>
```

Title (+)

Use **title** for <frbr:work/> and <frbr:expression/> separately. Use "Einheitssachtitel" (type="uniform") for name string.

```
<work authURI="http://viaf.org/viaf/" codedval="185515787">
  <titleStmt>
```

```
<title type="uniform">Der Freischütz</title>
  </titleStmt>
  </work>
```

Classification (+)

Use **classification** for <frbr:work/> and <frbr:expression/> separately. Use the authority files of the Library of Congress. A recommended vocabulary for musical genre and form is the genre and form term list of the Library of Congress Music Genre/Form Project Group¹⁵. Another recommended authority file is the MARC 21 Bibliographic 008 Music¹⁶. We should encourage the LOC that it is needed to work on this list and include further classification categories, but use it as a starting point.

Instrumentation (*)

Information on the **instrumentation** is captured on the level of <frbr:expression/>. Use the *MARC Instruments and Voices Code List*¹⁷ and see also the *Unimarc codes for medium of performance*¹⁸. On expression level information on instrumentation should only be encoded in a standardized form. Please take care not to mix it with transcriptions of instrumentation labelings given in single sources which should be recorded on manifestation level.

¹⁵ https://www.loc.gov/catdir/cpso/lcmlalist.pdf

¹⁶ http://loc.gov/marc/bibliographic/bd008m.html

¹⁷ https://www.loc.gov/standards/valuelist/marcmusperf.html

¹⁸ http://www.urfm.braidense.it/risorse/medium query en.php

Date (*)

Date of creation

Capture the date or the time span of creation with <frbr:expression/>, not with <frbr:work/>. If more detailed information regarding date of creation of different expressions, e.g. a complete score and an additional piano score, use <frbr:expression/> to capture each single expression.

```
<expression>
  <creation>
      <date notbefore="1791-12-05" notafter="1799-10-05" cert="low"/>
      </creation>
  </expression>
```

Date of first performance

Besides the date of creation of a work, the date of the first performance is very often a crucial information help to identify a work. We recommend, but it is of course not mandatory to capture the information about a first performance as an event as followed:

See the example coding to prove the usability and implementation of the core set and their adressings to their corresponding FRBR entities implemented in MEI¹⁹

```
<workDesc>
 <work xml:id="w_Z100075"> <!-- identifier -->
    <identifier auth="GND"</pre>
                authURI="http://d-nb.info/gnd/">300264135</identifier>
    <titleStmt> <!-- title -->
      <title type="uniform">Die Soldaten</title>
      <respStmt> <!-- composer -->
        <persName role="marc:cmp" authURI="http://zimmermann-gesamtausgabe.de/"</pre>
                  codedval="Z400002">Bernd Alois Zimmermann
      </respStmt>
    </titleStmt>
    <history> <!-- dates -->
      <creation><!-- date of creation -->
        <date notBefore="1957" notAfter="1965"></date>
      <eventList><!-- date of first performance -->
        <event type="firstPerformance">
          <date isodate="1965-02-15"/>
          <geogName type="place" authURI="http://www.geonames.org/"</pre>
                    codedval="2886242">Köln</geogName>
          <corpName auth="GND" authURI="http://d-nb.info/gnd/"</pre>
                    codedval="2057209-8">Oper Köln</corpName>
        </event>
      </eventList>
    </history>
    <perfMedium> <!-- instrumentation -->
      <perfResList/>
    </perfMedium>
    <expressionList>
      <expression xml:id="e Z100075">
        <classification>
          <termList>
            <term> <!-- classification -->
              <identifier
         authURI="http://loc.gov/marc/bibliographic/bd008m.html">op</identifier>
          </termList>
        </classification>
      </expression>
    </expressionList>
  </work>
</workDesc>
```

¹⁹ The example (with minor additions/modifications) is taken from a poster, which was presented at the Music Encoding Conference 2018 in Washington by Kristin Herold for the Bernd Alois Zimmermann-Gesamtausgabe

Enhancement

Uncertainties

Dealing with uncertainties is often an important topic for researchers, who want to note down if i.e. names, dates, places or research results in general cannot be proven precisely. Therefore, many of the MEI elements offer a @cert attribute, which "signifies the degree of certainty or precision associated with a feature"²⁰. Such uncertainties cannot totally be handled by library catalogues, that often only contain uncertainties concerning datings, for example "ca. 1910".

Refinement of instrumentation

Looking at the current encoding possibilities for the indication of cast and instrumentation in MEI it became apparent that it is not possible yet to encode information as detailed as needed. A big advantage of digital work catalogues in comparison to printed ones is the possibility to semantically search the performance medium.

Working on that topic in the context of the Bernd Alois Zimmermann-Gesamtausgabe it became clear that the transformation of prose text of the instrumentation suggested by the MEI Guidelines does not affect the needs of precision for work catalogue projects yet as there is no fully developed possibility to encode alternative instrumentations yet. Regarding this topic the Zimmermann Project developed the following proposal for an enhancement of the perResList>-element that should be discussed within the MEI community. It is proposed to add an element alternationGrp> capturing one or more alternative instrumentations using an alternate>-element, prefixed with the namespace <baselines > bazga:> to mark the proposed element.

```
<perfMedium>
  <perfResList>
    <perfResList>
      <bazga:alternationGrp>
        <bazga:alternate type="primary">
          <perfRes count="2"</pre>
                authURI="http://www.loc.gov/standards/valueList/marmusperf.html"
                codedval="wa">Fl</perfRes>
        </bazga:alternate>
        <bazga:alternate type="secondary">
          <perfRes count="2"</pre>
                authURI="http://www.loc.gov/standards/valueList/marmusperf.html"
                codedval="we">Picc</perfRes>
        </bazga:alternate>
      </bazga:alternationGrp>
    </perfResList>
    <perfResList>
      <bazga:alternationGrp>
```

²⁰ MEI Guidelines, Chapter 11: Editorial Markup, see: http://music-encoding.org/guidelines/v3/content/edittrans.html#edittransAbbrev

```
<bazga:alternate type="primary">
          <perfRes count="2"</pre>
                authURI="http://www.loc.gov/standards/valueList/marmusperf.html"
                codedval="wb">Ob</perfRes>
        </bazga:alternate>
        <bazga:alternate type="secondary">
          <perfRes count="1"</pre>
                authURI="http://www.loc.gov/standards/valueList/marmusperf.html"
                codedval="wb" n="1">Ob</perfRes>
          <perfRes count="1"</pre>
                authURI="http://www.loc.gov/standards/valueList/marmusperf.html"
                codedval="ba" n="2">EHr</perfRes>
        </bazga:alternate>
      </bazga:alternationGrp>
    </perfResList>
 </perfResList>
</perfMedium>
```

Harmonization of events

The analysis of current encoding practices has shown that there are some types of events occuring in almost all of the existent work catalogues, like dates of first performances or productions (recording, broadcast). The recommendation is to record these events following a common scheme. It is conceivable to offer a set of templates for the most common event types. A growing set of sample sets, also contributed by the whole MEI community could be of great help to harmonize the data structure of work catalogues.

To emphasize this idea, an example for a sample set is given as followed: Dealing with the event of the first performance of works by Zimmermann the need for sample sets arrived quite quickly, since the information about first performances of his works shows the feasibility for a very structured way of encoding instead of using prose descriptions. Two kinds of first performances occured within Zimmermann works. The first one is the "traditional" first performance of a musical piece while the second one (radio production) takes into consideration the recording and first live broadcasting of a production. The information is very consistent for both of them, so a sample encoding for such kind of information is possible and shown below.

Conclusion

The recommendations mentioned above are the result of the preliminary considerations that have been elaborated by the working group on work catalogues created with MEI by the end of 2017. One further step is to synchronize the results within the group and to convey them into the MEI community. The recommendations should be made publicly accessible on the MEI website.

In the long term, it would be excellent if a tool like the metadata editor MerMEId would support all of the recommendations to promote a common creation of work catalogue data across all projects.

While drafting the recommendations, the following three questions arose:

- Are these recommendations of any help or interest for libraries? How can we enhance the collaboration between institutions providing MEI data and libraries and the mapping and transformation of data. And which information of library catalogues is not addressable (yet) in MEI?
- The recommendations of core elements are quite well developed, but is this really all core information we need or want to see in a work catalogue? Or did we already capture too many elements and labeled them as core?
- In addition to core elements in an MEI work catalogue, much more information is needed (see Appendix for a sample list of different projects and their collected

topics). How can we deal with this issue? If we realize, that some information like dedicatee or incipits are part of work catalogues by several projects, but not all of them (=core elements). We suggest to develop sample sets comparable to the ideas in the chapter `Harmonization of events`. It could be a common task for the MEI community to enhance these recommendations. This would a) facilitate the start for new projects to create their work catalogues in MEI and b) increase the interoperability of data for existing projects by adapting these recommendations.

Appendix

Sample for work catalogues

The following overview presents categories recorded by five work catalogues/edition projects: the Kritische Ausgabe der Werke von Richard Strauss, the Anton Bruckner-Internetportal, the Christoph Willibald Gluck – Complete Works project, the Bernd Alois Zimmermann Complete Edition and the Carl Maria von Weber Complete Edition. The categories are not filtered yet. The six core elements have been determined on the basis of the information given by the projects.

• Kritische Ausgabe der Werke von Richard Strauss

http://www.rsi-rsqv.de/

- o title (Titel)
- dedication (Widmung)
- incipit of text (Textincipit)
- version (Fassung)
- systematology of the work (Werksystematik)
- o roles (Rollen)
- work number (Werknummer)
- person (Person)

• Anton Bruckner-Internetportal (Webarchiv)

http://www.bruckner-online.at/

- o using MerMEId for generating all information for their digital work catalogues
- usable as work and source catalogue
- coded information for works are:
 - title
 - work number
 - all involved persons (composers, dedicatees, authors, instrumentalists; all with certainty-attribute and GND link)
 - dedication
 - relations to other works
 - classification of the respective work
- musical information of the respective work like:
 - versions
 - tempo
 - key
 - incipit (text & music)
 - ensemble
 - instruments/cast in the ensemble
 - movements
- source information:
 - source description of the content (additions, connections to other sources, provenance)
 - texture of the source (dimension, list of hands)
 - RISM signature
 - classification/quality of the source
 - assignment of sources to versions
- date of composition, work genesis, performances (in general reduced of Bruckners life span)

Christoph Willibald Gluck – Sämtliche Werke

http://www.gluck-gesamtausgabe.de/

- o separation in search for work and sources
 - search for work (filter are Gluck-specific)
 - type of work (Werktyp)
 - original language (Originalsprache)
 - authenticity (Authentizität)
 - containment of time (Zeitliche Eingrenzung)
 - search for sources
 - category (Kategorie)
 - type/subtype (Typ/Untertyp)
 - form of documentation (Überlieferungsform)
 - quality (Qualität)
 - containment of time (Zeitliche Eingrenzung)
- general remarks:
 - use of "assigned", if authorship can not be proven without doubt
 - database TYPO3
- **Bernd Alois Zimmermann-Gesamtausgabe**, Work catalogue by Heribert Henrich (printed²¹), not a digital work catalogue yet
 - o title
 - dedication (Widmung)
 - date of creation (Entstehungszeit)
 - assignment (Auftrag)
 - text (Text)
 - performance medium/cast (Besetzung)
 - composition (Aufbau)
 - sources (Quellen)
 - publication (Veröffentlichung)
 - first performance (Uraufführung)
 - documents (Dokumente)
 - work genesis (Werkgeschichte)
 - literature (Literatur)
- Carl Maria von Weber Gesamtausgabe (printed)
 - o JV, WeV and opus numbers
 - Title
 - involved persons (composer, librettist etc.)
 - incipits
 - date of creation
 - classification (music/text)
 - relations to other works etc.

-

²¹ http://d-nb.info/99541727X

Overview: categories considered within work catalogues or complete edition projects

	Strauss	Bruckner	Gluck	Zimmermann	Weber
Person (Composer, Librettist, Lyricist, others)	х	х	х	x	х
Identifiers	х	х			х
Title	х	х	х	х	х
Classification		х	х		х
Instrumentation		х		х	х
Date		x	x	x	х
Dedication	х	х		х	х
Tempo		х			х
Key		х			
Incipit	х	х			х
Language			х		
First Performance		х		х	
Work genesis/history of work		х		х	
Structure of composition	х	х		х	
Relations to other works		х			х

The first six categories in bold print present the so called core elements.