On the thin line between enriching metadata and keeping usability

Bach digital is one of the oldest musicological projects in the web. Started originally on 28th July 2000,[1] it has been reborn in 2008, growing and improving itself ever since. It's goal is to provide extensive information on every musical work by a member of the Bach family, focussing on all known sources of these works. By now more than 2000 sources are fully digitized and viewable online in high quality.

Bach digital's sister project *Bach Repertorium*, which started in 1993, switched for cataloguing the works by Bach family members to the *MerMEId* database in 2015.[2] The first volume prepared completely with MerMEId was published last year in print,[3] another volume is expected to be released this year with two more catalogues coming up.

However, the metadata in *Bach digital* is not stored in MEI nor based on MEI, and perspectively will not be in the future.[4] It started from a relational (Microsoft Access) database, that has been converted to XML – and it still is somewhat relational. But we learned a lot from MEI. In the last ten years we constantly adapted and improved our data model for different reasons: cleaning up, including more information, restructuring and overall basically making it easier to re-use the data we provide. Also in ten years the underlying operating system changed four times, and eight different major versions of the application software (based on MyCoRe)[5] were used.

As musicological research still unveils new insights, we try to keep not only the systems running but also all the data up-to-date. All the changes that happed over the last years within *Bach digital* will be presented and explained, what drove our decisions.

A big step forward was the recent release of the integrated person database with more than 6000 entries so far, giving basic information on life and profession and links back to works and sources for composers and scribes. In principle it combines and enhances the indexes of the *Neue Bach Ausgabe*, the released volumes from the *Bach Repertorium*, and some other publications like *Die Bach-Quellen in Wien und Alt-Österreich*. Also it allows us to present the huge collection of handwriting samples compiled by the late Yoshitake Kobayashi and stored at the bach archive.

It is planned to enrich the metadata on musical works with musical incipits. For the latest volumes from the *Bach Reperorium* project these have been set with Sibelius, which will be batch converted to MEI and then uploaded to the database. Within the dataset the music is rendered with *Verovio*.[6] Some test examples are already online.[7]

---

1   IBM press release from 24 July 2000, https://www-03.ibm.com/press/us/en/pressrelease/1612.wss.

2   See talk "MerMEId in practice" by Wolfram Albrecht Enßlin and Klaus Rettinghaus at MEC2016, Montreal, Canada.

3   *Johann Ernst Bach : thematisch-systematisches Verzeichnis der musikalischen Werke (BR-JEB) / bearbeitet von Klaus Rettinghaus*, Stuttgart 2018.

4   The edition project *Corpus Musicae Ottomanicae* (CMO) is like *Bach digital* also a MyCoRe based application and it is using the MEI schema. It is thought about cooperating in some areas to take advantage of synergy effects. See https://corpus-musicae-ottomanicae.de/.

5   See http://mycore.org/.

6   By now the JavaScript toolkit of *Verovio* is used. For future versions of *Bach digital* it is planned to integrate the Java toolkit into the application to render the incipits server-side.

7   E.g. see https://www.bach-digital.de/receive/BachDigitalWork_work_00009090.

Some of the other topics that shall be presented and discussed further are:

- librettos

- bibliographic records

- improving and deeper integration of provenances

- JSON-LD and semantic web

- providing edited music

- overall development and deployment