MEI conversion for the Joseph Haydn Werkverzeichnis Online

Joshua Neumann

Akademie der Wissenschaften und der Literatur | Mainz

Akademie der Wissenschaften und der Literatur Mainz

CENTRE for DIGITAL MUSIC DOCUMENTATION

ABSTRACT

The work of the Joseph Haydn-Institut · Köln (JHI) has been well-known in eighteenth century studies and beyond since publication of the first four bands of the Joseph Haydn Werke (JHW) in 1958. As of early 2022, 132 print-only bands have been published, with only the 133rd–a supplemental addition–remaining in progress. No plans for a printed thematic catalogue or a digital publication of JHW exist, creating an analogue-digital dichotomy unique to this project. The next phase of work thus entails further development of the Joseph Haydn online portal in partnership with the Centre for Digital Music Documentation (CDMD) at the Akademie der Wissenschaften und der Literatur Mainz (ADWMainz). A works catalogue with complete metadata and polyphonic incipits is the current focus of the JHI-CDMD collaboration. The digital Joseph Haydn Werkverzeichnis (JHWV) builds upon recent examples of digital works catalogues (Bruckner, Gluck, Delius, and Nielsen) in terms of content engagement, data management, and structure while also accommodating extensibility.

Naturally for Gesamtausgabe or Werkverzeichnis undertakings, sources occupy a central role, and with them come massive amounts of data and metadata that require compilation and organization throughout the project lifecycle. Over the past decade, JHI has digitised the card-catalogue records for nearly 3,000 works attributed to or associated with Joseph Haydn at one point or another by converting each into an xml document. Given when this process started, and that large-scale xml projects in music documentation were not yet commonplace, JHI devised an idiomatic schema that set and achieved the goal of verbatim data preservation. Inasmuch as this system served immediate needs, it presents limitations in accessibility, shareability, and research-ability—all desirable or essential qualities for the Haydn portal. Moreover, it creates a distinct set of challenges for creating a digital Werkverzeichnis within the portal, the most pressing, time-intensive, and labor-intensive of which is transforming XML data into a standardized format and structure enabled to create the necessary qualities.
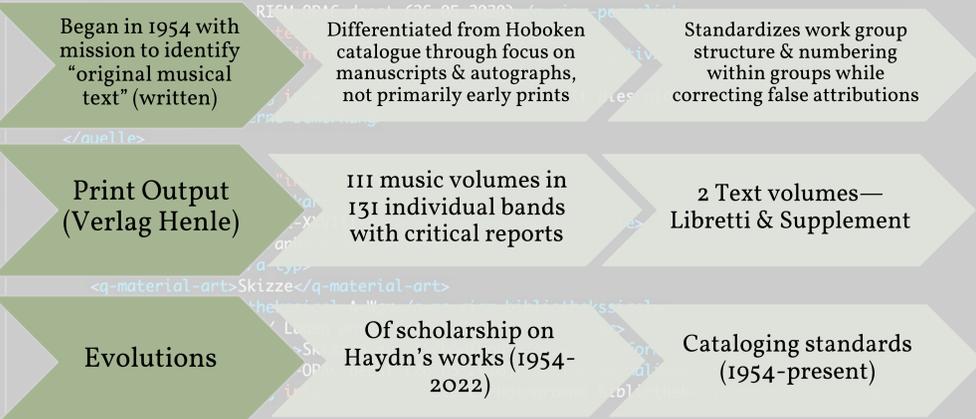
JHWV-Online GOALS

1) Mirror best practices from current online works catalogue projects
    1) Linking data between entities represented (people, places, objects) (Gluck)
    2) Full polyphonic incipits (Bruckner)
        1) Incipits as follow-on project
    3) Navigation and Methodological overview (Nielsen, Delius)
    4) Incipit searchability (Gluck)
    5) MEI-FRBR coding (Bruckner, Nielsen, Delius)
2) Embodies FAIR and CARE principles
    1) Multi-lingual web interface
    2) Open-access MEI code
    3) Accessible content editing interface
    4) Expanded use of <persName role=""> as possible
3) Set a new, technologically adaptive standard for digital works catalogues projects

JHI/JHW BACKGROUND/OVERVIEW

Began in 1954 with mission to identify "original musical text" (written)	Differentiated from Hoboken catalogue through focus on manuscripts & autographs, not primarily early prints	Standardizes work group structure & numbering within groups while correcting false attributions
Print Output (Verlag Henle)	111 music volumes in 131 individual bands with critical reports	2 Text volumes— Libretti & Supplement
Evolutions	Of scholarship on Haydn's works (1954-2022)	Cataloging standards (1954-present)

JHI Digitisation

Aim: verbatim digital preservation of information on analogue cards used to construct *Joseph Haydn Werke*

Began before MEI emerged as standard format it has become

Schema development: 2011
Schema revision: 2017

Resulted in an idiosyncratic xml schema for a unique use case

Each file contains work & source metadata

Contents Overview

≈ 3200 files with unique identifiers for individual works ascribed to J. Haydn, in 38 "work" groups/trunks
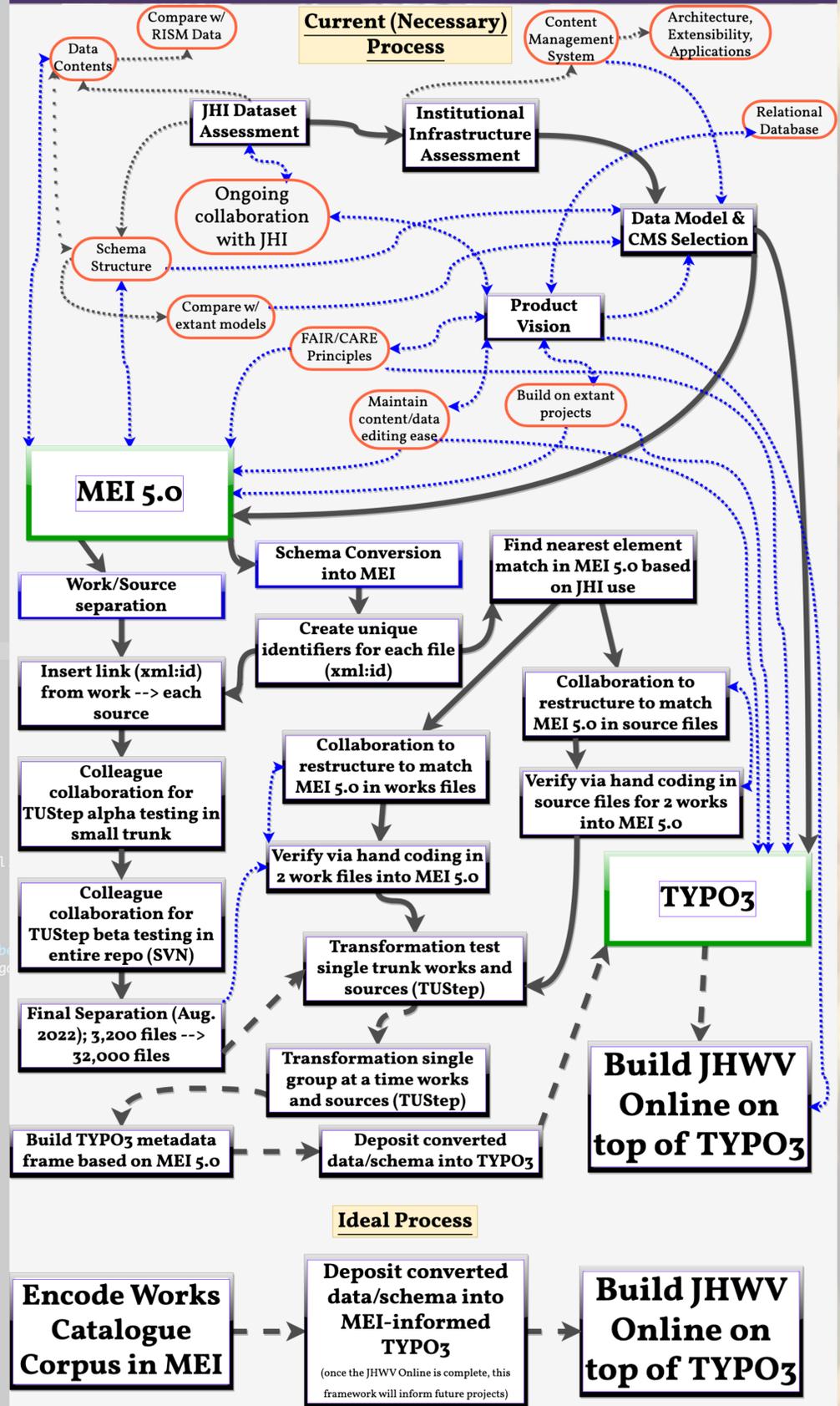
>29,200 source references

Exclusively metadata

≈180 unique element names ≈ 60/works, 120/sources

≈1400 entries with some level of challenged/uncertain attribution

WORKFLOW

Current (Necessary) Process

Compare w/ RISM Data · Data Contents · Content Management System · Architecture, Extensibility, Applications · Relational Database

JHI Dataset Assessment · Institutional Infrastructure Assessment · Ongoing collaboration with JHI · Data Model & CMS Selection · Schema Structure · Compare w/ extant models · Product Vision · FAIR/CARE Principles · Maintain content/data editing ease · Build on extant projects

MEI 5.0

Work/Source separation → Insert link (xml:id) from work --> each source → Colleague collaboration for TUStep alpha testing in small trunk → Colleague collaboration for TUStep beta testing in entire repo (SVN) → Final Separation (Aug. 2022); 3,200 files --> 32,000 files → Build TYPO3 metadata frame based on MEI 5.0

Schema Conversion into MEI → Create unique identifiers for each file (xml:id) → Collaboration to restructure to match MEI 5.0 in works files → Verify via hand coding in 2 work files into MEI 5.0 → Transformation test single trunk works and sources (TUStep) → Transformation single group at a time works and sources (TUStep) → Deposit converted data/schema into TYPO3

Find nearest element match in MEI 5.0 based on JHI use → Collaboration to restructure to match MEI 5.0 in source files → Verify via hand coding in source files for 2 works into MEI 5.0

TYPO3

Build JHWV Online on top of TYPO3

Ideal Process

Encode Works Catalogue Corpus in MEI → **Deposit converted data/schema into MEI-informed TYPO3** (once the JHWV Online is complete, this framework will inform future projects) → **Build JHWV Online on top of TYPO3**

CONCLUSIONS

- **Digitally based works catalogues are increasingly prevalent**, leveraging affordances of computer-aided research for searching, linking, and analysing metadata and content
- **Music Encoding Initiative is the standard for open access encoding** as it enables utilization of FRBR data modeling and easy adherence to FAIR and CARE principles
- **Early adoption of MEI affords comprehensive data structuring** for work and source metadata, musical content/incipits (in MEI or Plaine and Easie code), and interoperability for analytical extension; i.e. – it frees users to focus on content and ethical praxis instead of having to create something entirely new
- **Following "Ideal Process" enables faster work and easier to edit work on both technical and content levels - Save Time; do it MEI Way**

ACKNOWLEDGEMENTS & CONTACT