# Lute Music Project

Project with Tim Crawford and David Lewis  
Goldsmiths

## Background

Around £5K is available for a three-month period (May, June, July 2023) with a spending and payment cutoff at the end of July. The aim is to review the ECOLM and related music resources and produce a report and recommendations for a low-cost sustainability plan permitting incremental improvement. Some consideration should be taken for interoperability with the contemporary F-TEMPO search project. (These terms are explained below.)

## ECOLM – Electronic Corpus of Lute Music

ECOLM (1999-2002) produced a queryable database of lute encodings with metadata in a SQL database with a web interface. It is still running thanks to the efforts of a single unnamed IT person at Goldsmiths. Encodings use TabCode format (not widely used elsewhere) and are manually reviewed and annotated for quality control.

ECOLM II (2002-2006) extended the ECOLM data and performed research work on the basis of it. As far as visible outcomes are concerned, ECOLM and ECOLM II may be treated as the same project, resulting in a database of around 2000 encodings.

ECOLM III (2012) attempted to crowd-source further high-quality encodings by starting from OMR (optical music recognition) scans and enlisting experts and enthusiasts to make corrections. A pleasant editing tool was developed and there was significant initial interest, but it tailed off probably because the edits were not incorporated into any actual query system the users could use – they ended up in a separate database from the existing ECOLM. Failing the “cardinal rule of crowdsourcing” – ensure users can actually benefit from the edits they make. ECOLM III had successful aspects and perhaps with more time might have corrected course, but its instigators appear to see it as a failure.

## F-TEMPO – Full-Text Search of Early Music Prints Online

Search from printed music, currently in test using vocal pieces. Not directly applicable to lute music – can accept MEI encodings but requires monophonic indexes for search and needs to know the tuning to interpret tablature.

## New Third-Party Contributions

Several collections of lute music have been made available by various curators:

John Robinson, Lute Society publications. Around 7000 encodings consisting of printed music, prose, and semi-structured metadata. Intended for players rather than scholars.

Sarge Gerbode, lutemusic.org. Around 20,000 encodings in playing editions with semi-structured metadata. Some duplicates and adaptations. Intended for players rather than scholars.

Jan Burgers, Pierre Phalèse database. Phalèse was a 16th-century publisher; Burgers has about 1000 encodings in tab form of as-yet unknown quality, with publication metadata.

Peter Steur, mss.slweiss.de. Metadata catalogue. Around 68,000 listings of which 62,000 have incipits (opening ideas) encoded. Originally curated with the late Markus Lutz.

## Miscellaneous Notes and Queries

* Gerbode has an overall index spreadsheet (out of date? According to Tim) and Alexander Kerelis has made a simple but elegant search front-end based on it (<https://keliris.dev/lute-search>). It works by parsing the xlsx to JSON using a Rust crate (Calamine) then loading to Meilisearch (open source with paid hosting option)
* Where did the crowd-sourced data from ECOLM III ultimately go? Did it eventually end up in the ECOLM I/II database?
  + “It has a separate set of tables in our MySQL database, and no public visibility. I believe it’s probably still split by system rather than combining all the lines into single pieces.” – David
  + “The only data dump was to the Arts and Humanities Data Service, for sustainability (no comment). At the time, this was not a thing we thought about. Obviously, there’s no reason not to share the data other than that the schema is entirely bespoke. Happy to put a dump on a git repo if that helps get things started.” – David
* David mentions “a Github repo” for the mss.slweiss.de content. The page links to a Git index at <http://slweiss.de/gitlist/mss.git/commits/master> but it doesn’t load. Is there an actual Github repo or another mirror?
  + “There is the git repo (but not GitHub, as you say) you found. The online ‘database’ (or rather website) is generated by php scripts (maybe others, too) from cvs files in the sub-directory ‘trunk/mss’ in the repo - AFAIK there is no SQL d/b or DBMS involved - Markus did mention to me in an email back in 2018 that he had tried MySQL or similar but had given up on that. However, the cvs files on which the website depends are generated in some way from files uploaded by Peter or Markus - what happens to these uploads I don’t know.” – Tim
  + Actual Git repo can be cloned from <http://slweiss.de/git/mss.git/>
* Learn more about what RISM does and how!
  + It’s a catalogue of sheet music available *somewhere* – with scholarly metadata (full librarian metadata)
  + Does not generally contain the actual music (although it can) – its original purpose is to tell you where to find it
  + Develops or sponsors the Verovio MEI renderer and Muscat cataloguing application
  + “RISM is an old cataloguing project (used to be book based) of historical sources of notated music … now built on the library standard MARC... librarians use Muscat, RISM’s tool. That generates MARC for local library catalogues, **and** contributes to RISM as a union catalogue. So no, there’s no music transcription, but there are records to which transcriptions could be attached (in a LOD sort of way..)” – David
  + “There are well over 1m music incipits, though, as David says, not every source/item includes them” – plus notes on limitations – Tim
* How essential is musical content to ECOLM?
  + “Totally essential, as the general motivation is to provide as much as possible in the way of full-text encodings with their own metadata, which needs to be reconciled with [mss.slweiss.de](http://mss.slweiss.de) , which is a broader resource” – Tim

## What do users want from these resources?

* Performance resources – tablature principally. These are relatively well-served already in the sense that most of the existing databases are for performers.
* Historical background and observations, as in the Lute Society notes.
* “Discovery” – what does this desire consist of? Content search? Similarity? Metadata such as composer, publication date, variations… performance difficulty? Links to recordings?
* Musicological analysis. Common themes. Styleometry (what does this cover?)
* Curated metadata about tab quality etc

## Things Mentioned and Not To Forget About

* Verovio – general MEI renderer, supports lute tab I think?
* TabCode – format devised by Tim himself which “nobody else uses”
* Luteconv – open source converter between tab formats, “works beautifully”
* Tunings: Tablature editions (almost?) always assume players know how to tune the instrument. You have the upper six playing strings, and then bass strings which may further depend on the key. The upper strings have two dominant tunings styles before and after some (unspecified) cutoff date – before, guitar-like; after, usually a minor chord.

## My deliverables?

* Summary of situation as it stands
* Survey of options
* Recommendations

## Desirable Qualities of a Solution

* Transparency and replicability by third parties. What happens if everyone involved with the project leaves? Can it be rebuilt from public data?
  + Note: not calling for a guarantee of organisational continuity. If nobody out there is interested at all, it might still disappear.
* Accessibility and user value. It must support the things that users want, straightforwardly.
* Ease of contribution. It should be possible for people other than the maintainers to contribute material or corrections and to see and use their own contributions in the context of the system.
  + With the usual questions about quality control and provenance.
* Incremental improvement. It should be possible to start with “whatever we have” and update it as resources permit.

## Not Goals

* Essentially anything better served by a “performance interface”
* In the short term it is most important to have a process for managing and querying data (contributed in potentially many forms) rather than to have a single interface for contributing and managing that data. Lesson of ECOLM III