

Klezmer Institute Digital Humanities Projects: Klezmer Archive Project & the Kiselgof-Makonovetsky Digital Manuscript Project

Christina Crowder, Klezmer Institute

Presented at the POLIN Conference, What's New, What's Next! — October 5, 2022

Panel: New Interactive Internet Approaches to Studying Eastern European Jewish Music

Abstract: In this paper we argue that both klezmer music and the Yiddish language are not simply historical artifacts, but the loci of thriving contemporary interest communities and that creating digital structures tailored to the needs of these communities is a clear imperative that matches the need to design for academic best practice in the digital humanities space. The Klezmer Archive Project and the Kiselgof-Makonovetsky Digital Manuscript Project (KMDMP) present a unique opportunity to create digital spaces that support culture bearers, and to develop high-quality, authoritative resources for newcomers to the field. Both of these projects have been the unlikely beneficiaries of Coronavirus-related lockdowns in unexpected ways. The Klezmer Archive functions as a distributed team, and much of the communal work, exchange, learning, and events in the KMDMP take place in zoom rooms.

Klezmer Institute initiated two interconnected digital humanities projects in 2020.¹ The first is the Klezmer Archive,² a new kind of digital humanities archival tool for musics of oral tradition. The second is the Kiselgof-Makonovetsky Digital Manuscript Project, which is pioneering new methods of public scholarship for musical manuscripts and translation. If we posit that both klezmer music and the Yiddish language are not simply historical artifacts, but the loci of thriving contemporary interest communities, then creating digital structures tailored to the needs of these communities is a clear imperative—even as we design for academic best practice in the digital humanities space. These

¹ This conference paper is the result in large part of co-writing with Clara Byom, Development Director of the Klezmer Institute. Find out more about these projects at <https://klezmerinstitute.org>

² The Klemzer Institute is supported by the National Endowment for the Humanities through a Phase I Digital Humanities Advancement Grant for the Klezmer Archive Project in 2021-2022.

projects present a unique opportunity to create digital spaces that support culture bearers, and to develop high-quality, authoritative resources for newcomers to the field.

Klezmer Archive Project

The Klezmer Archive project aims to create a universally accessible, useful resource for interaction, discovery, and research on all available information about klezmer music. This project is adapting and applying methodology from computational musicology and library sciences to create a tool to facilitate study of the klezmer corpus in a deeper, more systematic manner and on a more comprehensive scale than previously possible. The project team is confident that this archival tool can be built using existing software technology, but it is already known that the planned archive tool will reach beyond the current state of the art in digital humanities platforms.

The project has been awarded a National Endowment for the Humanities Phase 1 Digital Humanities Advancement Grant for 2021-2022 and is approaching the midway point in the current phase.

As many of you know, Klezmer is the instrumental music of Ashkenazic Jews of Eastern Europe, and it was and continues to be a transnational music based in oral tradition. For decades members of the klezmer community have dreamt of a centralized repository for klezmer tunes and their historical/ethnographic context, but creating such a resource within current archival structures leaves out a critical source of knowledge—klezmer culture bearers. These individuals have a deep understanding of repertoire, history, and folklore that is highly valued within the international

klezmer community, but it is only available to the whole community when it is collected and organized.

This resource will bridge the gap between oral history and archives by being a space where culture bearers and community members can engage with archival items and one another as musicians do—by drawing connections between tunes, discussing genre classifications, searching for unnamed melodies, and comparing recorded versions of the same tune—while maintaining quality standards on par with institutional collections. The project seeks to create a novel resource that combines attributes of both institutional archive structures and community-based collections and that integrates computational music analysis tools to facilitate further study.

The need for this kind of resource is clear for many kinds of folkloric musics, but particularly so for Klezmer. Documentating klezmer is inherently complicated because it is a transnational orally transmitted music. Multiple languages, alphabets, and transliterations create complex metadata issues that require significant work to accommodate. Like any folkloric music, the lack of singular composers, inconsistent naming conventions, multiple versions of the same tune, and overlapping genre classifications quickly push against the limits of static archival taxonomies. To address these issues, the team is investigating ways to structure corpus-specific metadata vocabularies and how to build tools for curated user contributions within a flexible architecture. Semantic metadata will show relationships between items, linking multiple recordings, tune variations, and shared melodic material and any other user-identified relationship artifact-to-artifact.

Allowing scholars and practitioners to contribute both factual data and more subjective classifications and commentary will create a space for equally valid, but sometimes competing

narratives. This approach more closely represents the way that culture bearers understand this music and better reflects the collective knowledge of the community. To both accommodate and elevate the complexity of an artifact's web of connections rather than presenting a singular narrative is a fundamentally different task than traditional archive structures take on, but doing so creates a more complete and inclusive history.

For example, the team is investigating ways to display music as an image of the original manuscript, digital notation, and pitch map all together. Another visual format that has been discussed is how to implement a “talmudic” approach for the visual display of information. In this view you would have a central text (in our case an artifact of music) surrounded by specific location information, connections to other texts (music items), and commentary that relates to the central item that can co-exist even when it presents different interpretations. Applying a well-considered, weighted rank to different commenters would allow the tool to feature the commentary and additional knowledge added by culture bearers and scholars about specific tunes with a high visual ranking while still allowing community members to contribute to the text record about the item.

What the Archive team is working on in Phase 1

The work over the course of the two year NEH grant period is to develop a plan for the archive tool—what is referred to as the Minimum Viable Product, or MVP. This plan will reflect all of the features that must be included in order to fulfill the mission of the archive tool (but no more than that). In concert with identifying the features of the MVP, a proposed data structure will be developed to support these features (and others, such as incorporating recordings, to be added later), along with an elucidation of the technical (i.e programming) requirements needed to execute each of the

features. The grant outcome will be a robust plan describing what the archive tool needs to include, how the tool will be structured, and what technical parts and pieces need to be adapted to build it.

The project team includes experts in the fields of archives and libraries, corpus studies, user experience, system architecture, and computational musicology, along with the project management team. The “secret sauce” of the team—in all seriousness—is that all but one of the core team are practicing klezmer musicians. This fact eliminates any distance between the people who are making the tool and those who will use it.³

Over the course of the grant period, the Archive Project team is meeting with experts in a number of related fields, attending conferences, and reading widely. Project management is adapted from the Agile project management system, with an emphasis on an iterative process that responds quickly as new information is acquired with the intention of bringing this kind of industry awareness along with the project as it evolves. As a very brief overview, the Klezmer Archive Project has four work streams.

Tech: Investigating data structures based on prior art, comparative analysis, research and meetings with specialists. The team will begin to develop tools to help scale the data ingestion process and experiment with different technologies that could be used to build the Archive.

Fuzzy Tune Boundary Research: The semantic structure of the klezmer corpus is fuzzier than that of the musical genres for which similar work has most commonly been done in the past (e.g. classical, jazz, and pop). To understand how to structure klezmer tune data, the team

³ It is also a lucky detail that the team also happens to make a great klezmer ensemble, with three violins, flute, clarinet, tsimbl, and accordion.

must first understand how klezmer practitioners and researchers conceptualize that information so the tool matches the range of their mental models. This work is being done through interviews and topic mapping experiments.

Editor Experience/Data Ingestion: The editing and data ingestion stream is involved with collecting volunteer-donated scores and text translation; developing editorial policies for music and text from manuscript sources; and learning from the volunteer experience in the Kiselgof and other projects.

End-User Experience: Informed by comparative analysis and observations from the “fuzzy” and “editorial” work streams, the team will identify the most important features for the MVP, develop models, and begin wireframe prototyping of the archival component of the Klezmer Archive.

Kiselgof-Makonovetsky Digital Manuscript Project

The klezmerim interviewed by Kiselgof were professionals, with an eclectic collection of different tunes in their repertoire that let them earn a living, generation after generation, adapting to the new tastes and audiences. Their voices faded through numerous social cataclysms, but luckily, the archive survived. While these materials are very important for research, their biggest impact is going to be when they return to the fingers of the musicians and the feet of the dancers, when they continue their journey through human hearts. Luckily, we discovered this time capsule at a post-revival time, when there is a generation of new klezmerim more than capable of processing and reviving this repertoire. The communities of Yiddish Summer Weimar, Yiddish New York, Klez Canada, etc. are the real heirs to those who contributed to the archive in the first place. Why not pass them their inheritance, and let them develop it further - while helping to create a curated academic edition?

— Anna Rogers (née Gladkova)

The ***Kiselgof-Makonovetsky Digital Manuscript Project (KMDMP)*** connects musicians, Yiddishists, and scholars today with klezmer musicians from more than 100 years ago through an international digital humanities project that is innovating new forms of public scholarship through accessible online platforms. The project community is digitally notating and translating handwritten music and text collected during the An-ski Ethnographic Expeditions of 1912-1914 and the Makonovetsky Wedding Manuscript, making them accessible to scholars and musicians for further study, computational analysis, and performance.

The corpus comprises 850 high-resolution photographs of handwritten music manuscripts and catalogue pages that are held in the Institute of Manuscripts at the Vernadsky National Library of Ukraine. The material was acquired in Kyiv by Anna Rogers in 2017 through a collaboration with Mariko Mishiro of Tokyo University of Arts; Chitoshi Hinoue at Kyoto University of Art and Design; Pete Rushefsky at the Center for Traditional Music and Dance in New York; and, Dr. Walter Zev Feldman. Klezmer Institute was invited to create a project to work with the material on the condition that all of the material should be made available to the public right from the very beginning, and that digitization, editing and curation should follow as an open source project using Creative Commons licenses.

Klezmer Institute has been guided by this mission from the beginning, and the project design centers around digitally notating the approximately 1,400 melodies, and transliterating and translating all Yiddish, Russian, and other language text, catalogue information, and annotations found in the materials. It is unfolding in three phases. The launch of the project in November 2020 marked the completion of the first phase (Dissemination), in which the scans of the materials were prepared and disseminated. The project is currently in Phase Two (Digitize & Transcribe) in which

scholars, musicians, linguists, and other interested individuals are invited to participate in the digitization, transcribing, and transliterating of the materials. Depending on the acquisition of more materials, Phase Two could last through 2025. Phase Three (Analyze, Perform, & Transform) began in Summer of 2021 when the project was the feature topic at the Yiddish Summer Weimar Festival and will be ongoing, as participants are encouraged to record/perform the pieces they are digitizing. Phase Three also encompasses development of editorial teams that will create critical editions and player folios for publication.

Given the imperative to release all of the materials right from the beginning, it was essential that the project platform be accessible, inexpensive (free), and flexible. It is organized on two web pages (one public and one private for registered participants), a Facebook group, a Google email group, and the KMDMP Commons Google folder. Project materials include a public folder containing PDFs with the original manuscripts (access and preservation resolution); tracking sheets to coordinate digitization and translation/transliteration; guides and other resources; and, a public folder containing the PDFs of digitized tunes. Participant registration and multi-format score uploads are coordinated through the automation platform Integromat.

The significance of the materials in this collection cannot be overstated—both for klezmer music practitioners and for the study of Ashkenazic Jewish history and culture. Public release of these manuscripts increases the easily accessible European klezmer repertoire more than fivefold, and gives an unprecedented view into the repertoire and musical trends encountered by working musicians in the time period. The intentional choice to release the collection to the public without prior curatorial or editorial intervention allows participants to discover details about the individual musicians and ethnographers who contributed to the collection on their own and in community, and

to take in the full range of genres and tunes in the corpus. Through this volunteer-driven, collaborative project, researchers and performers are developing a more nuanced and comprehensive understanding of the European klezmer repertoire and its context. Moving forward, genre-based study of this corpus will illuminate social trends in the small towns visited by the An-sky team. For musicologists, the digitized corpus will facilitate many kinds of hitherto-impossible computational analyses, such as geographical distribution of musical motifs, data-informed genre distinctions, and many others.

Access to a significant portion of the field collections from the An-ski Ethnographic Expeditions (1912-1914) allows researchers to gain a more complete understanding of the informants, their repertoire, and even, in some cases, their thoughts about documenting this music for future generations. The Kiselgof repertoire comes from a geographically different region than the field work conducted by the revered Russian Jewish ethnographer Moshe Beregovski. As scholar Walter Feldman notes, "Kiselgof went to areas of Northern Ukraine and some parts of Belarus where he was from, including even the westernmost region of northwest Ukraine in Kremenets, which used to border on Austrian Galicia. Beregovski did not do fieldwork in these areas and so now we have a rather significant repertoire coming from the border between the north and south of Jewish Eastern Europe. The material augments our knowledge of European klezmer music in a very significant way." Beregovski's curated publications represent the highest level of Jewish musical scholarship in his time, and they include selected repertoire from the material now available digitally in KMDMP. The ability to examine Beregovski's volumes through the lens of the larger corpus allows us to contextualize choices that may have been made for ideological reasons, and to re-center individual informants by examining their more complete repertoires as represented in the manuscripts.

We believe that this kind of public scholarship is the only way to make these collections available to a broad audience in a timely fashion. Optical Music Recognition (OMR) for handwritten music remains in its infancy compared to OCR for handwritten text and is not available on any public platform; likewise, music notation software is relatively more complicated to use than word processors. Nevertheless, as of mid-September 2021, approximately 760 tunes have been digitized and large portions of text have been translated/transliterated—about 50 percent of the KMDMP materials. This represents literally thousands of hours of donated, volunteer labor. The community is enthusiastic and the project team is committed to continuing the work, discussions, and collaborations until the digitization and translations are complete.

KMDMP is a startup initiative of the Klezmer Institute and has yet to receive grant support beyond small individual cash donations and immeasurable hours of donated time from within the community of klezmer practitioners and scholars. This all-volunteer model has been successful for the initial digitization phase, but future editorial & curatorial efforts will require more robust funding to compensate editorial teams—members of which we hope to recruit from the KMDMP community itself.

Though it was envisioned from the beginning that the KMDMP manuscripts would serve as a test corpus for the Klezmer Archive Project, the Kiselgof project has informed the Archive project in many unexpected ways. It was known that working with incoming music and text would be helpful for defining useful metadata specific to the corpus. More surprising has been the evolving process of accommodating the foundational choice to be platform agnostic for incoming digitized music files. In an effort to minimize friction for potential score volunteers, it was decided to accept digital notation from any music notation software, leading to a decision to request three kinds of files: native software

format (.native); uncompressed music xml (.xml); and PDF (.pdf, for visual confirmation of the desired layout). Accommodating updates to incoming files was another high priority along with version control. Automating the score intake process through the Integromat platform has been its own side project for the Archive team and has been tremendously informative for understanding how to interact with and accommodate community volunteers in a low-tech, but nevertheless technical space. Understanding what works and what doesn't when it comes to engaging with project participants has been informative for the User Experience component of the Archive project, and the administrative team is actively working to sustain interest in the project now that Coronavirus restrictions are easing and something closer to normal life returns for the musicians and Yiddishists who have been the primary participants to date.

Returning to the quote at the beginning of this section, Anna Rogers is correct in identifying the communities of Yiddish Summer Weimar, Yiddish New York, and KlezKanada as the true heirs to those who shared their music and music manuscripts with the An-ski Ethnographic team. These two projects present a unique opportunity to design digital spaces that support culture bearers and to create high-quality, authoritative resources for newcomers to the field.