NATIONAL ENDOWMENT FOR THE HUMANITIES

Narrative Section of a Successful Application

The attached document contains the grant narrative and selected portions of a previously funded grant application. It is not intended to serve as a model, but to give you a sense of how a successful application may be crafted. Every successful application is different, and each applicant is urged to prepare a proposal that reflects its unique project and aspirations. Program guidelines also change and the samples may not match exactly what is now required. Please use the current set of application instructions to prepare your application.

Prospective applicants should consult the current Office of Digital Humanities program application guidelines at https://www.neh.gov/grants/odh/digitalhumanities-advancement-grants for instructions.

Applicants are also strongly encouraged to consult with the NEH Office of Digital Humanities staff well before a grant deadline.

Note: The attachment only contains the grant narrative and selected portions, not the entire funded application. In addition, certain portions may have been redacted to protect the privacy interests of an individual and/or to protect confidential commercial and financial information and/or to protect copyrighted materials.

Project Title: The Klezmer Archive Project

Institution: Klezmer Institute

Project Director: Christina Crowder

Grant Program: Digital Humanities Advancement Grants, Level II

LIST OF PARTICIPANTS

Project Team

*Project Director: Christina Crowder - Executive Director at Klezmer Institute

*Grant Administrator/Project Coordinator: Clara Byom, M.M. - Development Director at Klezmer Institute

*Eléonore Biezunski, Ph.D. - Sound Archivist at YIVO Institute for Jewish Research

*Dan Kunda Thagard - Independent Software Developer

*Yonatan Malin, Ph.D. - Associate Professor in the College of Music and Program in Jewish Studies at the University of Colorado - Boulder

Andrew Parisi, Ph.D. - Independent Machine Learning Engineer

*Max Rothman - Independent Software Engineer

*Matthew Stein - Designer at Enigmida and Independent Computational Musicologist and Software Engineer

Schyler VerSteeg - Product and UX Researcher at CVS Health

Project Advisory Board

Members of the Advisory Board will be available to the project team to answer questions and provide advice, participate in 1-2 yearly meetings with the Advisory Board & Members of the Klezmer Archive Team, participate in a user research call with others in my area of expertise, advocate for the project, and share potential opportunities with the team.

Jeanette Casey, M.M., M.L.S. - Music Librarian at Mills Music Library at UW-Madison Libraries

Jean Delahousse - Independent Semantic Web Consultant and Analyst

Emese Ilyefalvi, Ph.D. - Assistant Professor of Folkloristics at Eötvös Loránd University, Budapest, Hungary and research fellow at MTA–ELTE Lendület Historical Folkloristics Research Group

Mark Kligman, Ph.D. - Professor of Musicology and Ethnomusicology at University of California - Los Angeles

Judith Pinnolis, M.M, M.S. - Associate Director, Instruction and Engagement at Berklee College of Music/ The Boston Conservatory at Berklee

*Pete Rushefsky - Executive Director at Center for Traditional Music and Dance

*Amanda (Miryem-Khaye) Siegel - Research Librarian at New York Public Library

Lyudmila Sholokhova, Ph.D. - Curator of the Dorot Jewish Collection at The New York Public Library

David Lewis - Researcher at University of Oxford e-Research Centre in Oxford and Lecturer in Computer Science at Goldsmiths, University of London

TBD - An expert in Digital Text Studies, particularly right to left script, to be identified during the grant period

*denotes individuals who are klezmer musicians

Consultants

Michel Biezunski, Ph.D. - Founder and CEO of Infoloom
Jean Delahousse - Semantic Web and Knowledge Graph Consultant, DOREMUS project
MEI/MIR Specialists TBD
Designer & Front End Engineer TBD (may be one more individuals)
Graphic Designer for Outreach TBD

The Klezmer Archive Project - Narrative
Project Overview
Klezmer, the instrumental music of the Ashkenazic Jews of Eastern Europe, was and continues to be a transnational music based in oral tradition. The Klezmer Archive (KA) project is creating a universally accessible, useful digital archival tool for interaction, discovery, and research on available information about klezmer music and its network of contemporary and historical people. During the grant period, the team will prototype and test tools and concepts through an iterative process involving community feedback, consultation with specialists, and experiments. By the end of the Level II grant period, the Klezmer Archive team will have produced prototypes of core components of the resource; a basic front end developer tool kit to be used for the web interface; a white paper for public release and/or publication; conference presentations; and, public announcements (including dev blog posts).

Enhancing the Humanities
The Klezmer Archive project speaks to recent trends in the humanities that regard works of art and music not as self-standing aesthetic objects, but as entities that participate in networks of human relations. Unlike traditional archives, this will be a digital native archival resource that will not hold physical objects. Instead, the project uses archival principles to document a conceptual abstraction of music, its manifestations (sheet music, cds, YouTube videos, etc.) with reference to physical holdings via DOI, events, movement of tunes, the networks of people associated with its past, present, and future, and facilitate user-generated commentary and connections. The archival tool will be an authoritative resource with rich descriptive information, will facilitate search and analysis within and between musical items, and will highlight connections between historical and contemporary music makers. To facilitate this, the KA team will build the resource by combining knowledge graphs with features of knowledge engines, music information retrieval (MIR), pre-existing and novel tools for search, and front-end design based on the principles of generous interfaces. A Level II Digital Humanities Advancement Grant will enable the project team to develop proof of concept experiments for elements identified as essential to the core functions of the tool.

The resource will be an inclusive, welcoming space for anyone interested in the content regardless of experience. Anticipated users include scholars, independent researchers, teachers, musicians, dancers, community leaders, and individuals interested in eastern European Jewish culture. While klezmer is the passion that unites the members of this project team, cultural practitioners and researchers throughout the world experience similar problems in capturing, organizing, and accessing data for music of oral tradition and folklore. Projects in the humanities will be able to use and adapt all open-source tools and conceptual frameworks developed by the KA team, who will build upon numerous existing open-source tools and schema (i.e. MEI), and will contribute back to those ecosystems. Researchers working in the fields of folklore, music, language, ritualistics, folk dance, and any other field that documents the transmission of cultural knowledge will find tools and concepts useful for their projects.

As the humanities have embraced crowdsourcing models, it is increasingly apparent that collaboration between scholars and practitioners can be especially effective. The Klezmer Archive will continue to use the musical corpus of the [Kiselgof-Makonovetsky Digital Manuscript Project](#) (KMDMP), a community-driven klezmer manuscript digitization project, as its primary case study. The project invites participants to transcribe and translate the music and notes contained in ~850 high-resolution scans of hand-written manuscripts and catalog into digital formats for further study and performance. Datafication and technology are often perceived as enemies of oral transmission, but our experience leads us to believe this is not the case. Non-hierarchical learning environments fostered in KMDMP bridge levels of experience, increase access, and serve as a model for the community values the Klezmer Archive project wishes to embrace. The Klezmer Archive will provide models and tools to flatten the distance

between researchers and the people they study, enhancing research and serving the community simultaneously.

The tool will move beyond traditional archival descriptive data to open the "black box" of music by facilitating search within musical items. The team has adopted the open source Music Encoding Initiative (MEI) schema, the most widely used format for digitally encoding music in academic and archival projects similar to the Klezmer Archive. It supports features the project requires, such as representing ambiguous marks in handwritten scores and supporting editorial annotations. High quality MEI tooling already exists, including Verovio, a software library for engraving MEI into SVG graphics. There is room for improvement on some tooling—such as editing interfaces and software libraries for manipulating MEI data—that KA technologists are eager to contribute to. Integration of MIR tools will enable further study—from simple comparison of tune variations to complex analyses of schemas over the entire corpus. For musicians and musicologists, the ability to search across both metadata and within the music itself will facilitate a specifically musical type of thought, capturing semantic aspects of music otherwise invisible and making them available for advanced study.

Documentation of klezmer music is complicated in ways shared with other transnational and diaspora musics of oral tradition: a) multiple alphabets and languages; b) no singular composer, inconsistent naming conventions, and multiple versions of tunes; c) overlapping genre classifications. Revealing the embedded networks between musical items and the people with which they are associated (Born 2005) requires both a robust metadata ontology, and a data framework that can surface musical relationships as non-hierarchical tune families and define human relationships such as mentor, interlocutor, and sibling, etc. Working from the understanding that "the nature of the art object is a function of the social-relational matrix in which it is embedded," (Gell 1998, p. 7) the ability to view musical items as repertoire associated with particular individuals, intellectual movements, or as characteristic of specific regions can inform the study of socio-political history.

The KA library team has conceptualized four "layers" of metadata to organize information in meaningful ways: 1) catalog; 2) domain-specific; 3) semantic; and 4) commentary (See Appendix 5). By mid-2023, the KA team will have a working draft ontology that will be thoroughly tested in Phase II. Designed to facilitate crosswalking metadata from other sources, the ontology is particularly innovative in metadata layers 2 and 3, where klezmer-specific metadata will be articulated for the first time, and the ontological conundra of "fuzzy" genre boundaries, usefully organizing conflicting or contradictory information, and human relationships will be addressed. While klezmer-specific metadata will be of minimal use to other projects, the concepts will be applicable and the KA team will develop documentation to facilitate adaptations into other subject areas. When successfully paired with front-end design inspired by emerging ideas around generous interfaces for search, these layers of metadata will facilitate discovery in ways never previously possible and will provide a framework for collections that depend on domain-specific metadata to facilitate useful search.

During Phase I, the team learned of a type of tool called knowledge engines that structure data more flexibly than is possible with relational databases, RDF, or property graphs. To our knowledge, these tools have rarely—if ever—been used in the humanities, which for the most part relies on RDF and related technologies. The KA team may be one of the first teams with the technical expertise and commitment to implementing knowledge engines for humanities work. Knowledge engines model complex real-world data similarly to how humans store knowledge—as a densely interconnected network of contextualized relationships. The data model envisioned will allow researchers to store, view, and evaluate conflicting or mutually exclusive data in ways that are not well-supported through existing technologies. In 2022 the team identified [TypeDB](#) as the most mature and well-supported open-source knowledge engine available and data ingestion into TypeDB will begin in early 2023 (see Appendix 6). Implementation of this advanced technology will not limit potential interoperability with

existing collections and institutions. The team has identified several types of interoperability that will guide decisions around tools and processes: MEI and XML/other encoding formats, relational databases and knowledge engines, standard library and archival ontologies and the KA ontology. To ensure interoperability, the team will identify specialists and have independent conversations with librarians, archivists, and technologists throughout the granting period.

The Klezmer Archive Project will serve the humanities broadly by creating resources and tools for documenting, preserving, and researching the artistic expression of cultures and the people connected to them. Integrating musical search and analytical tools will make a corpus of non-classical western music available to a wide audience of researchers in musicology, corpus studies, and ethnomusicology. Deploying the power of knowledge engine technology together with robust, domain-specific ontologies will give humanists more flexibility in the use of cutting-edge digital tools to assist in the collection, organization, and evaluation of cultural information. As the first team pursuing this application of open-source first order logic systems in the humanities, the KA project is committed to reducing barriers to entry for future projects. The Klezmer Archive will allow humanists, performers, and archivists to work together to tell the story of communities, arts, and cultures in a more comprehensive way than has been previously possible.

Environmental scan

A number of tools, models, standards, and concepts will be used for experimentation, development, and continued exploration. The [Music Encoding Initiative](#) (MEI) is an open source community-based effort that has developed "a core set of rules for recording physical and intellectual characteristics of music notation documents expressed as an [XML] schema." The [Corpus Musicae Ottomanicae (CMO)](#) project uses MEI to encode complex, non-western musical manuscripts and is of great interest to the team. The open-source [Verovio](#) engraving library can be used to render MEI for display. Both would be interoperable and can be used with [Music21](#) and other common computational musicology tools. We have referred to [Kranenberg, et al](#) (2007) for an overview of the challenges of using music information retrieval tools in orally-transmitted music collections. [MEI Friend](#), a "last mile" editor for MEI music encodings intended to assist in cleaning up encodings generated via optical music recognition or conversion from other formats, and the [MELD Framework: Music Encoding in Linked Data](#), that digitally combines notated music with contextual and interpretive knowledge in the Semantic Web will be explored in early 2023.

Connecting the encoded music to manuscript images provides essential context. KMDMP manuscript pages have marginalia in different handwritings and inks and all of this information tells a part of the story that cannot be represented in the current digital encoding process. In Phase II the team will consider how the [Web Annotation Specification](#), [TEI](#), and [IIIF](#) standards can inform the connection and display of encoded music and source images. The [Transcription for Paleographical and Editorial Notation](#) (TPEN) tool, which allows users to transcribe text and have it displayed on the image, is a notable example in this domain. Representing uncertain, partial, and incomplete date information will be a critical concern, so the team will explore ways to model known and unknown time using standards and models found in the [OWL Time Ontology](#), [CIDOC Conceptual Reference Model](#), and [Extended Date/Time Format](#) (Library of Congress).

Experimentation will continue with ways to incorporate knowledge engine features into the Klezmer Archive. [TypeDB](#) is an open source product similar to a knowledge engine that can be used as a foundation for the project. [CYC](#) is a closed source machine reasoning platform that will be used as inspiration. Early research on generous interfaces and discovery has surfaced a number of projects and people of interest, such as the [LD4 Discovery Affinity Group](#), [FRBRVis](#), [Dario Rodighiero](#), [Gissoo Doroudian/Bona Fide](#), [Matthiu Jacomy](#)/[Gephi](#), and [Huda Khan](#).

Throughout the research phase, the team explored non-music areas of research in the

humanities for concepts and models. [Egil Bakka's](#) (2019) work to connect classification and typology in the field of folk dance to theoretical concepts from the fields of archeology and biology has inspired the team to find solutions to what has been identified as the "fuzzy tune and genre problem." Learning about the [Humanities Networked Infrastructure](#) (HuNI) project validated ideas about the value of representing complexity, contestation, and connection. When considering metadata ontologies, the team worked most with [DoReMus](#), a project that sought to improve music description and make three French music collections inter-searchable via Linked Open Data. DoReMus harmonized [CIDOC CRM](#) and [FRBRoo](#), which was extended with classes and properties specific to musical data, and a set of shared multilingual vocabularies. The KA team will continue to explore this ontology, as well as other standard metadata systems such as [MARC,](#) [The Music Ontology,](#) and [DublinCore](#) to ensure interoperability.

In addition to its novel technical goals, the Klezmer Archive seeks to become a central hub for its community of musicians, audience, researchers, and educators in the same way that resources like [The Session](#), [Hymnary](#), and [the Country Dance and Song Society](#) have become the connective tissue of their respective communities. The project team will set up meetings and workshops with individuals from many of these identified projects/programs and draw upon their experience to inform the project design process.

History of the project

The project team came together over lunch in December 2019 after the Yiddish New York festival. The meeting revealed that everyone there had already spent considerable energy thinking about how to make a digital archive for klezmer, and that the nascent team included a mix of people with programming, archival, and research expertise—as well as deep domain expertise as klezmer musicians themselves. By another stroke of good fortune, a user experience researcher joined the team early on, helping to foreground the needs of the community of users from the very beginning. Genesis of the Klezmer Archive Project coincided with development of the [Kiselgof-Makonovetsky Digital Manuscript Projects](#) (KMDMP). As KMDMP progressed, the idea to use its corpus as the test case for the more ambitious digital archive project began to emerge.

In 2020, the project team and advisory board participated in an exercise to explore basic questions of how the archive would solve the needs of its intended users, leading to a foundational set of nine principles which continue to guide the project. The team also conducted comparative analysis of similar projects and technologies, held group transcription observation sessions, and met with Michel Biezunski of Infoloom to learn about his topic mapping program. The project was awarded a Phase I NEH Digital Humanities Advancement Grant for 2021-2022 (HAA-277220-21).

The project team began 2021 by pursuing research in the four main interest areas in the DHAG work plan (Technology, User Experience, Fuzzy Tune/Genre Boundaries, and Editorial/Data Ingestion) and working in four overlapping groups (tech, library/archive, user experience, corpus studies). The team's User Experience expert led a series of interviews with potential users of the Klezmer Archive to learn about how potential users think about and relate to klezmer, and what features they would like to see in the resource. The library team researched ontologies used for organizing data in music library and archive systems and experimented in Infoloom's Networker, leading the team to develop an adaptation of DoReMus concepts. The library team observed ongoing crowdsourced digitization of KMDMP materials, and engaged with developing metadata vocabulary lists, recognizing corpus-specific issues within musical notation and text, and identifying participants to invite for further data processing and ingestion experiments. The tech team researched data modeling, leading the team to learn about Knowledge Engines and receive consultation from Andrew Parisi. The tech team developed and deployed a preliminary search tool for use on the KMDMP corpus in mid-2022. Team members attended conferences and held meetings with individuals working on related projects to learn

about current work in global musicology, generous interfaces, knowledge graphs, ontology, models for validating different kinds of data, and the potential use of entailment to make inferences on data sets. In May of 2022 the team launched a dev blog to provide updates and the team held public community meetings in June of each year.

In late 2022 the project team will focus on ingesting KMDMP data in a format accessible for testing early prototypes for data ingestion. The tech team will continue to refine and iterate on the search tool based on feedback from the community. The library team aims to have 80 percent of a draft ontology completed by the end of 2022. In early 2023 the team will produce a white paper for public release and/or publication, a plan for transcribing and ingesting the entirety of the KMDMP collection, an architecture plan for the software of the MVP, and appropriate public announcements and presentations introducing the project to the relevant fields as outlined in the Phase I work plan. The team is on track to meet these goals.

Activities and project team

The project team will convene regularly for monthly all-hands meetings and three in-person retreats. Additional meetings will be scheduled with specific user groups, the public, advisory board, and domain experts identified by ongoing research and networking. The complete work plan can be found in Attachment 3. The work is divided into four streams: User Experience/Outreach; Ontology/Knowledge Engine; Music Information Retrieval & Notated Score Ingestion/Editorial Processing; and Data Ingestion/Editorial Processing. Each feature will be developed in an iterative approach with tight feedback loops of research, defining goals, design planning, and tech prototyping.

The team anticipates developing and testing feature prototypes based on an initial prioritization:

P1- Music search, data modeling/inference;
P2- Score ingestion, metadata ingestion, metadata search;
P3- Artifact view/presenting scores.

Ultimately, prototype success will be determined by rigorous testing and iteration. To evaluate the project's progress and success overall, the team will invite domain experts (e.g. archivists/librarians, musicians, musicologists, historians) and the Advisory Board for probing, high-level conversations. As tools are shared with users, the team will seek feedback via conversations, surveys, and UX research.

The Klezmer Archive project is uniquely capable of executing on this plan because the assembled team is composed of experienced software industry, archival, academic, and musical experts wholly focused on carrying it through to completion. This project is extremely ambitious in scope, presenting both technical and operational challenges to successfully design complicated, sensitive processes. Specific risks include:

1. UX problems, such as showing sufficient context for laypeople to understand contradictory assertions, may be more difficult to solve than anticipated. The team includes individuals with industry UX experience to manage the workflow and deliver results on time and on budget.
2. Technical problems that are harder than expected to solve. The team includes software industry professionals with a strong track record of delivering working products, but to mitigate this risk, the team will continue involving academic and industry contacts to navigate complications as they arise.
3. Inability to ingest data at the necessary rate and quality. The team recognizes that the data ingestion process must be treated like a product in its own right with research, trials, and iterations on a small scale before ramping up ingestion.

Final products and dissemination

During the two years of funding for a Level II Digital Humanities Advancement Grant, the project team will iterate on prototypes and concepts that will accomplish the Klezmer Archive mission. Results will include: prototypes of core components of the resource; a basic front end developer tool kit to be used for the web interface; a white paper for public release and/or publication; and, public announcements, conference presentations, the dev blog, the Klezmer Institute mailing list, and social media.

Papers and presentations will be accessible to the public via the Klezmer Institute website and authors of conference papers and blog posts will be named as contributors. The team anticipates presenting at or attending a combination of the following conferences: LD4, Digital Humanities Summer Institute, The Knowledge Graph Conference, International Society for Music Information Retrieval Conference, Graphen & Netzwerke Conference, Music Encoding Conference, Music Library Association, and International Society for Ethnology and Folklore.

Everything built throughout the entire lifecycle of the project—from software to content—will be freely available to technologists, musicians, and scholars alike, and anyone with the relevant knowledge will be encouraged to contribute. Contributors to code development will be credited in the GitHub repository and public announcements where applicable. Outside consultants will be identified as co-authors of text and/or code contributors. Volunteer participants in the KMDMP contribute score notations under a contributor license agreement, and all digital notation outputs from the project are licensed through a [Creative Commons CC-BY](#) 4.0 license with attribution to the Vernadsky National Library of Ukraine as the holder of the source materials. Annotated digital images of the physical sources that comprise the KMDMP corpus will be shared via IIIF with a [Rights Statement](#) determined in collaboration with the Vernadsky National Library of Ukraine once hostilities with the Russian Federation cease. At later stages of the project, cooperative arrangements will be sought with archives and libraries with relevant holdings, and appropriate Rights Statements will be agreed with each cooperating institution.

To ensure accessibility to individuals with disabilities, the project will consult with one or more front-end experts during this granting period, aiming to build for [WCAG 3.0 compatibility](#) from the start. The team will also consult the W3C [Web Accessibility Initiative page](#) and the U.S. Department of Justice's [Guidance on Web Accessibility and the ADA](#).

After the grant period, the team expects to launch an initial version of the archive tool to the public with the entirety of the KMDMP manuscripts incorporated. Later stages of the project will address questions of how to engage institutions with klezmer music collections, and how to incorporate and process digital material shared with the Klezmer Archive project.

EXPANDED WORK PLAN
Start Date: September 1, 2023 | End Date: August 31, 2025

Workstream Key:
UE - User Experience/Outreach/Design (VerSteeg, Crowder, Byom)
KE - Knowledge Engine/Ontology (Crowder, Byom, Biezunski, Parisi)
MIR - Music Information Retrieval & Notated Score Ingestion/Editorial Processing (Kunda Thagard, Rothman, Stein, Crowder, Byom, Malin)
DI - Metadata Ingestion/Editorial Processing (Kunda Thagard, Rothman, Stein, Crowder, Byom, Biezunski, Parisi)

Each workstream will involve various team members throughout the grant period depending on the tool being developed and tested. Although the workstream teams listed above will be most involved in the work, the team will also identify consultants to assist with specific topics. The team will hire a front end engineer and designer (one or more people) to develop a tool kit that can be deployed by the programming team. The team anticipates developing and testing feature prototypes based on an initial prioritization: **P1:** Music search; data modeling/inference; **P2:** Score ingestion; metadata ingestion; metadata search; **P3:** Artifact view/presenting scores. Each feature will be developed in an iterative approach with tight feedback loops of research, defining goals, design planning, and tech prototyping. The schedule will be flexible and each step will inform the step that follows.

(UX) User Experience/Outreach/Design - Research to inform prototyping and testing, user interface choices (including generous interfaces), and further exploration of target user journeys based on analysis from the research conducted in Phase I. This workstream will also lead the communication and collaboration with the front end engineer and designer.
(KE) Ontology/Knowledge Engine - Research and experimentation on data modeling, ontology, typology, and schema for structuring metadata that are flexible enough to evolve and adapt for future needs within this domain, and to connect to other domains/systems. Experimentation with existing linked data tools, and prototypes/processes developed in the MIR and DI streams. Research on crosswalking metadata and interoperability with existing collections.
(MIR) Music Information Retrieval & Notated Score Ingestion/Editorial Processing - Development of editorial and technical processes for ingesting scores. This stream will rely on KMDMP digitized scores as testing materials.
(DI) Data Ingestion/Editorial Processing - Development of editorial and technical processes for ingesting metadata. This stream will use KMDMP metadata and other case studies to test ontology concepts on anticipated media formats.

Ongoing throughout the grant period
(All) - Monthly All-Hands Meetings which function in quarterly cycles and fall into three categories: Planning and Prioritization, Design Discussions, and Open/TBD.
(UX & DI) - Research and Discovery Phase for Generous Interfaces and Search
(UX) - Identify, supervise, and collaborate with front end engineer and designer
(KE) - Use Knowledge Graph to test and iterate on Ontology & Schema

September - November 2023
(UX) - Continued open-ended research on user needs, Further exploration of target users/journeys
(KE) - Conversion of Ontology into Schema
(MIR) - Iteration on musical search tool, Clean KMDMP Case Study Scores, Modify XML to MEI 5.0.0-dev Conversion Tool, Adopt IIIF Image and Presentation APIs for KMDMP Case Study

Images
(DI) - Experimenting and Iterating on Metadata Input Process and Tools for Internal Use
(including editing and viewing interface with TypeDB)

December 2023 - February 2024
(All) - In-Person Retreat (June)
(UX) - User Research on IIIF experience, Continued open-ended research on user needs,
Further exploration of target users/journeys, Research on Music Search Display
(KE) - Continue the conversion of Ontology into Schema
(MIR) - Iteration on musical search tool, Ingest KMDMP Case Study Scores, Convert Score from
XML to MEI Format, Begin Score Editorial Process in MEI, KMDMP Community Participants
Test and Use IIIF API Region Selection with URL Linking
(DI) - Iterating on Metadata Input Process and Tools for Internal Use

March - May 2024
(UX) - User Research on IIIF experience continues, Research User-Directed UI
Design/Experience, Research on Music Search Display
(KE) - Use Knowledge Graph to test and iterate on Ontology & Schema
(MIR) - Continue Score Editorial Process, Iterate on Score Ingestion & Editorial Process
(DI) - Iterating on Metadata Input Process and Tools for Internal Use

June - August 2024
(All) - In-Person Retreat (October/November)
(UX) - Research User-Directed UI Design/Experience, Research & Experiments on Music
Search Display
(KE) - Use Knowledge Graph to test and iterate on Ontology & Schema
(MIR) - Iterate on Score Ingestion and Editorial Process
(DI) - Iterating on Metadata Input Process and Tools for Internal Use

September - November 2024
(UX) - Research User-Directed UI Design/Experience, Research & Experiments on Music
Search Display
(KE) - Focused experiments on Knowledge Engine Special Features (n-ary relations, inference,
attribution, validation)
(MIR) - Iterate on Score Ingestion and Editorial Process, Working towards improved
Domain-Specific Music Search, Building Generous Interfaces for browsing our collection built
on top of musical search
(DI) - Iterating on Metadata Input Process and Tools for Public Use

December 2024 - February 2025
(All) - In-Person Meeting (June)
(UX) - Research User-Directed UI Design/Experience, Research and Experiments on Music
Search Display
(KE) - Focused experiments on Knowledge Engine Special Features
(MIR) - Iterate on Score Ingestion and Editorial Process, Iteration on Domain-Specific Music
Search, Experiments on Music Search Display
(DI) - Iterating on Metadata Input Process and Tools for Public Use

March - May 2025
(UX) - Research User-Directed UI Design/Experience
(KE) - Focused experiments on Knowledge Engine Special Features
(MIR) - Iterate on Score Ingestion and Editorial Process, Iteration on Domain-Specific Music Search, Experiments on Music Search Display
(DI) - Iterating on Metadata Input Process and Tools for Public Use

June - August 2025
(All) - In-Person Retreat (October/November)
(UX) - Research User-Directed UI Design/Experience
(KE) - Focused experiments on Knowledge Engine Special Features
(MIR) - Iterate on Score Ingestion and Editorial Process, Iteration on Domain-Specific Music Search, Experiments on Music Search Display
(DI) - Iterating on Metadata Input Process and Tools for Public Use

Anticipated Conferences/Workshops:
LD4, Digital Humanities Summer Institute, The Knowledge Graph Conference, International Society for Music Information Retrieval Conference, Graphen & Netzwerke Conference, Music Encoding Conference, Music Library Association, and International Society for Ethnology and Folklore
Anticipated Individual Meetings: Meeting and conference calls with individuals from projects identified in the Environmental Scan and through continued research will be scheduled on an ad hoc basis as the project develops. To ensure interoperability, the team will have conversations with librarians, archivists, and technologists at institutions with significant klezmer music holdings. Meetings with consultants and advisors including Michel Biezunski, Jean Delahousse, Grace Van't Hof, MEI/MIR specialists David Lewis and Andrew Hankinson, and front end engineers/designers will be held throughout the grant period.

Targeted User Group Interviews: (Dates TBD and participants to be added)
Librarian and Archivist Group - Judith Pinnolis, Amanda Miryem-Khaye Siegel, Jeanette Casey
Jewish Music Researchers and Community Leaders - Pete Rushefsky, Mark Kligman
Folklore/Folk Music Digital Archivists - Emese Ilyefalvi, Peter Van Kranenberg
Text Specialists - Asya Vaisman Schulman, Helen Beer

Risks: The Klezmer Archive project is uniquely capable of executing on this plan because the assembled team is composed of experienced software industry, archival, academic, and musical experts wholly focused on carrying it through to completion. This project is extremely ambitious in scope, presenting both technical and operational challenges to successfully design complicated, sensitive processes. Specific risks include:
1. UX problems, such as showing sufficient context for laypeople to understand contradictory assertions, may be more difficult to solve than anticipated. The team includes individuals with industry UX experience to manage the workflow and deliver results on time and on budget.
2. Technical problems that are harder than expected to solve. The team includes software industry professionals with a strong track record of delivering working products, but to mitigate this risk, the team will continue involving academic and industry contacts to navigate complications as they arise.
3. Inability to ingest data at the necessary rate and quality. The team recognizes that the data ingestion process must be treated like a product in its own right with research, trials, and iterations on a small scale before ramping up ingestion.

DATA MANAGEMENT PLAN

Types of Data Collected: The Project will collect five types of data: User Research Data, Code, Encoded Music Data, Manuscript Image Data, and Metadata.

Data Storage and Preservation of Access (by data type):
User Research Data- The raw data will be kept confidential due to legal and ethical concerns, but high-level summaries may be shared in presentations or publications. Presentations and publications will be made available under a Creative Commons license.
Code- All source code developed by the project will be published under open source licenses, under the klezmer-archive Github Organization.
- Reusable libraries under a permissive open source to be determined (e.g. MIT, BSD)
- Application code will be open sourced either under the same license as libraries, or under a copyleft license like GPL
- Contributions to other open source projects that are not maintained by the Klezmer Archive Project will be made under the licenses determined by maintainers of those projects

The Klezmer Archive Project team will maintain the publicly hosted open source projects and accept code contributions from the community. The project will ensure that we have the necessary rights to make community contributions equally open source, for example via a contributor agreement. The project will provide documents of how to contribute to each code project (documenting how to get started, workflow, coding conventions, architectural goals, and licensing information). All code contributed to projects maintained by the Klezmer Archive Project will pass through code review and automated testing before being merged into the main branch of the project
Encoded music- Digitally-encoded music will be crowdsourced from volunteers. Volunteers submit encoded music (MusicXML and native file formats) under a Creative Commons License, with the ownership of the encoded document provided to the Klezmer Archive Project. The encoded score will go through automated and manual transformations to convert the document into the necessary format, fill in missing metadata, ensure the document conforms to the style guide, make corrections, and editorial annotation. The volunteer digitizer is cited as the encoder, and the Klezmer Archive Project editorial team cited as the editor. Digitally-encoded music notation encoded by the Klezmer Archive project or given to and edited by the Klezmer Archive Project will be made available under a Creative Commons license, in standard music encoding formats such as MEI, MusicXML, and Kern. All digital-encoded music will also be accessible through the application we are building, where it will be searchable, browsable, and downloadable using a user-friendly interface.
Manuscript image data- Images of source manuscripts will be made available under a Creative Commons license. The images will also be served by a IIIF-compliant image server, and we will fully implement the IIIF Image and Presentation APIs to facilitate easy access and reuse of the manuscript image data.
Metadata- Klezmer Project will maintain a knowledge graph consisting of data derived from metadata attached to encoded music artifacts, data made available by host institutions, information referenced in academic literature, and comments and connections made by users of

the Klezmer Archive Project's application. This data will be hosted using TypeDB, an open-source database by Vaticle. The Klezmer Archive Project will expose an API for querying the data in the knowledge graph. Data will be available in one or more RDF translations, either live or via a regular export.

Period of Data Retention: Data shared with and created by the Klezmer Archive Project will be shared as soon as possible with the public on the Klezmer Archive Github organization, the Klezmer Institute website, or some other easily-accessible manner based on project needs, e.g. google drive or a custom website. Public access will be retained for the duration of the Klezmer Archive Project's existence. Data will be preserved for a minimum of 5 years beyond the life of the project.

Data That Will Be Retained and Not Shared:
- User data from our application (logins, passwords, etc.) will be maintained securely by the Klezmer Archive team and not published.
- Analytics data from our applications, useful for understanding user preferences and use-cases of our application will be securely maintained by the project team and not published.
- Logs and tracing data generated by our applications, useful for the project team for debugging, troubleshooting, and analytics, will be securely maintained by the project team and not published.