FWO DMP Template - Flemish Standard Data Management Plan

Version KU Leuven

Project supervisors (from application round 2018 onwards) and fellows (from application round 2020 onwards) will, upon being awarded their project or fellowship, be invited to develop their answers to the data management related questions into a DMP. The FWO expects a **completed DMP no later than 6 months after the official start date** of the project or fellowship. The DMP should not be submitted to FWO but to the research co-ordination office of the host institute; FWO may request the DMP in a random check.

At the end of the project, the **final version of the DMP** has to be added to the final report of the project; this should be submitted to FWO by the supervisor-spokesperson through FWO's e-portal. This DMP may of course have been updated since its first version. The DMP is an element in the final evaluation of the project by the relevant expert panel. Both the DMP submitted within the first 6 months after the start date and the final DMP may use this template.

The DMP template used by the Research Foundation Flanders (FWO) corresponds with the Flemish Standard Data Management Plan. This Flemish Standard DMP was developed by the Flemish Research Data Network (FRDN) Task Force DMP which comprises representatives of all Flemish funders and research institutions. This is a standardized DMP template based on the previous FWO template that contains the core requirements for data management planning. To increase understanding and facilitate completion of the DMP, a standardized **glossary** of definitions and abbreviations is available via the following link.

1. General Project Information	
Name Grant Holder & ORCID	**David Burn** (0000-0002-9137-9591)
Contributor name(s) (+ ORCID) & roles	**Toon Van Waterschoot** (0000-0002-6323-7350) – KU Leuven ESAT-STADIUS : co-supervisor WP2 **Mike Kestemont** (0000-0003-3590-693X) – Universiteit Antwerpen : co-supervisor WP3 **Bart Demuyt** (0000-0003-2902-552X) – Alamire Foundation : co-supervisor WP1 **Bram Caers** (0000-0001-7482-3198) - Leiden University : co-supervisor WP4
Project number [1] & title	S005525N **NEW PERSPECTIVES ON MEDIEVAL AND RENAISSANCE COURTLY SONG**
Funder(s) GrantID [2]	S005525N
Affiliation(s)	☒ KU Leuven ☒ Universiteit Antwerpen ☐ Universiteit Gent ☐ Universiteit Hasselt ☐ Vrije Universiteit Brussel ☒ Other: Alamire Foundation Leiden University ROR identifier KU Leuven: 05f950310 ROR identifier Universiteit Antwerpen: 008x57b05 ROR identifier Alamire Foundation: 00vrf8x48 ROR identifier Leiden University: 027bh9e22

[1] "Project number" refers to the institutional project number. This question is optional. Applicants can only provide one project number.

[2] Funder(s) GrantID refers to the number of the DMP at the funder(s), here one can specify multiple GrantIDs if multiple funding sources were used.

Please provide a short project description	The Franco-Flemish polyphony preserved in the recently discovered Leuven Chansonnier belongs to European heritage that stands artistically at the same level as the Flemish Primitives' paintings. However, the performance of the music contained in the fifteenth-century manuscript is jeopardized by the more than five centuries that separate medieval composers and today's musicians. From the initial results of the research conducted to bridge this gap, it appears important to transcend the case of the Leuven Chansonnier and to integrate it into its broader musical, literary, and historical context through multidisciplinary research. At the same time, the the research that this entails pushes the limits of traditional analysis. The response to this situation lies in data production through innovative technology and AI-driven processes. The methodology developed for this purpose is expected to prove applicable in other domains of the humanities. Starting from the case of the Leuven Chansonnier, "New Perspectives on Medieval and Renaissance Courtly Song" addresses a multifaceted set of questions. Specifically, it involves (1) multidisciplinary research into the broader societal and cultural context within which the Leuven Chansonnier and the extensive corpus of sources related to it were produced and circulated; (2) innovative research into the architecture and acoustics of the spaces in which the source corpus in general, and the Leuven Chansonnier in particular, were performed; and the transposition of the resulting data and insights into a contemporary performance context; (3) the integration of the achieved results and their dissemination both through academic and practice-oriented research and via international, national, regional, and local valorization platforms that aim to provide an authentic heritage experience to a wide audience.

List and describe all datasets or research materials that you plan to generate/collect or reuse during your research project. For each dataset or data type (observational, experimental etc.), provide a short name & description (sufficient for yourself to know what data it is about), indicate whether the data are newly generated/collected or reused, digital or physical, also indicate the type of the data (the kind of content), its technical format (file extension), and an estimate of the upper limit of the volume of the data.

Dataset Name	Description	New or Reused	Digital or Physical	Digital Data Type	Digital Data Format	Digital Data Volume (MB, GB, TB)	Physical Volume
		☐ Generate new data ☐ Reuse existing data	☐ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:		☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
Dataset WP1.1 (Music): Digitization of extant sources concordant with the Leuven Chansonnier	Image data will be captured and processed from relevant sources in Denmark (Copenhagen), France (Dijon, Paris), Germany (Berlin, Ulm), Italy: Florence, Pavia, Perugia), Portugal (Porto), South Africa (Cape Town), Spain (Segovia) and later on published on IDEM.	Generate new data	Digital	Observational data: Images	.raw; .tif; .jpeg2000	2,4 TB (estimate for 15 sources; raw, tiff, jpeg200)	
1. Research Data Summary							
metadata of extant	printed) catalogues, inventories and			Textual		sources)	

sources concordant with the Leuven Chansonnier	platforms, to complement the digital images of the sources, in the data management system CollectiveAccess in partnership with LIBIS (KU Leuven) and by a carefully developped datamodel.						
Dataset WP1.3 (Music): listing of musical and poetic variants across the sources	In this dataset, musical and poetic variants of the concordant songs across the sources will be documented by examining the available digital representations of the sources and by making up a variant table. In this process, text documents (.docx), PDF's (.pdf) and musical transcription files (Sibelius 7; .si7) will be created.	Generate new data	Digital	Observational data: Textual	.docx; .pdf; si7	< 1 GB	
Dataset WP2.1 (Space): 3D models of spaces historically linked to the corpus of courtly song	The raw visual capture data has the proprietary format of the capturing soft- and hardware (Leica) that we use in the project. As intermediary/processed files we use open standard file types .E57 and .las, or the file types .ptx and .pts that are becoming de facto standards when describing 3D visual content in ascii format. As to the file types that we use to integrate the 3D visual models with other media types: .PLY and .OBJ are becoming de facto standard 3D image formats. The acoustic data will be stored in the lossless and open PCM format (WAV). Data will initially be captured in 192 kHz, 24-bit WAV format, but will be downsampled to 48 kHz for long-term storage. 3D acoustic models derived from these data as well as the corresponding metadata will be stored in one of the formats	Generate new data	Digital	Observational data: model	.ptx; .pts; .PLY; OBJ; .wav	1 TB (estimate of 5 space)	

	specified in standard AES69-2022. More specifically, the FIR and SingleRoomSRIR as well as their extensions described in the standard will be used.						
Dataset WP2.2 (Space): Audio recordings of music performances	Audio recordings of music performances: in 96kHz, 24-bit WAV format	Generate new data	Digital	Experimental data: sound	.wav	13,60 GB (estimate 10 performances, 8 sound channels, with duration of 10 minutes)	
Dataset WP2.3 (Space): 3D video recordings of music performances	The raw capture data has the proprietary format of the 3D video capturing soft- and hardware (Evercoast). As there is no (not yet a) universal file type standard for 3D video, we use the proprietary .ecm file type (ecm = EverCoast Media), alongside open formats such as an .obj series of images.	Generate new data	Digital	Observational data: model	.ecm	96 TB (estimate 10 performances, 8 3D video channels , with duration of 10 minutes)	
Dataset WP3.1 (Text): metadata of late medieval (Dutch-language) songs	Data will be gathered from database Dutch Song Database. This material allows us to extract formal features and chart their chronological, geographical and intra-textual distribution such as stanza forms, metre (cf. Haverals et al. 2019), rhyme words, metaphors and motifs.	Reuse existing data	Digital	Compiled/ aggregated data: textual	xml, xlsx	< 1GB	
Dataset WP3.2 (Text): inventory of formulaic expressions that function as important intertextual connections	This inventory is important to evaluate the performance of computational algorithms from text reuse detection. Hereby, intertextual connections included in Brinkman's apparatus of his Gruuthuse edition (2015) can be used to build a gold standard.	Generate new data	Digital	Observational data: textual	xml, xlsx	< 1GB	
Dataset WP3.3	Data will be gathered from databases	Reuse existing	Digital	Compiled/	xml, xlsx	< 1GB	

(Text): smaller metadata subset of French song lyrics	as Catalogue de la Chanson Française à la Renaissance as proof-of-concept. Using machine translation, we can map the Dutch-language and French-language lyrics into the same space to retrieve textual echoes.	data		aggregated data: textual			
Dataset WP4.1 (Networks): metadata on multilingual sources containing song (1350 and 1550, in Low Countries)	Data will be gathered from various (online or printed) catalogues, inventories and platforms of different language text traditions. Sources include open-source data collections such as the Liederenbank (https://www.liederenbank.nl/) or the *Bibliotheca Neerlandica Manuscripta et Impressa* (BNM-i, https://bnm-i.huygens.knaw.nl/)	Reuse existing data	Digital	Compiled/ aggregated data: : textual	.csv	< 1 GB	
Dataset WP4.2 (Networks): additional data on multilingual text witnesses	Data generated from secondary literary review or own research of the text witness on the level of the multilinguality. Additional data will form an extra layer on top of existing data collection in WP4-1, providing more depth than is available in current reference works.	Generate new data	Digital	Observational data: textual	.csv	< 1 GB	
Dataset WP4.3 (Networks): parallel text tradition mapping of monolingual sources	restricted to songs and texts of particular interest; gathered from various (online) catalogues, inventories and platforms, with a specific focus on Middle Dutch material	Reuse existing data	Digital	Compiled/ aggregated data: : textual	.csv	< 1 GB	

GUIDANCE: *The data description forms the basis of your entire DMP, so make sure it is detailed and complete. It includes digital and physical data and encompasses the whole spectrum ranging from raw data to processed and analysed data including analysis scripts and code. Physical data are all materials that need proper management because they are valuable, difficult to replace and/or ethical issues are associated. Materials that are not considered data in an RDM context include your own manuscripts, theses and presentations; documentation is an integral part of your datasets and should described under documentation/metadata.* *RDM Guidance on data*	
If you reuse existing data, please specify the source, preferably by using a persistent identifier (e.g. DOI, Handle, URL etc.) per dataset or data type.	***Datasets WP1.2&1.3 (Music):*** idemdatabase.org diamm.ac.uk https://www.liederenbank.nl/index.php?lan=en ***Datasets WP3.1&3.3 (Text):*** https://www.liederenbank.nl/index.php?lan=en http://ricercar-old.cesr.univ-tours.fr/3-programmes/basechanson/index.htm ***Datasets WP4.1&4.3 (Networks):*** Sources include, but are not restricted to: https://www.liederenbank.nl/; https://bnm-i.huygens.knaw.nl/; https://portail.biblissima.fr/en; https://jonas.irht.cnrs.fr/index.php; https://handschriftencensus.de/; https://www.arlima.net/
Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? If so, refer to specific datasets or data types when appropriate and provide the relevant ethical approval number.	☐ Yes, human subject data; provide SMEC or EC approval number: ☐ Yes, animal data; provide ECD reference number: ☐ Yes, dual use; provide approval number: ☒ No Additional information:

Will you process personal data[3]? If so, please refer to specific datasets or data types when appropriate and provide the KU Leuven or UZ Leuven privacy register number (G or S number).	☐ Yes (provide PRET G-number or EC S-number below) ☒ No Additional information:
Does your work have potential for commercial valorization (e.g. tech transfer, for example spin-offs, commercial exploitation, …)? If so, please comment per dataset or data type where appropriate.	☒ Yes ☐ No If yes, please comment: ***Dataset WP1.1 (Music):*** The Mirador-viewer is implemented in the IDEM database so that data is interoperable in function of efficient data-connection. The Mirador viewer is a IIIF compliant viewer application running on the LIBIS KU Leuven service infrastructure. Between the holding institutions and the Alamire Foundation (KU Leuven) an agreement shall be concluded authorizing the ADL to digitise the necessary sources for further integration of these images in the Integrated Database for Early Music (IDEM) and to use them for related research and valorisation purposes. (Standard ADL agreement) ***Dataset WP2.2 (Space):*** a future integration of audio files into the IDEM data model is planned, where also the necessary agreements (e.g. with performing ensembles) will be part of.

[3] See Glossary Flemish Standard Data Management Plan

Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material/Data transfer agreements, research collaboration agreements)? If so, please explain to what data they relate and what restrictions are in place.	☒ Yes ☐ No If yes, please explain: ***Dataset WP1.1 (Music):*** The terms and conditions of IDEM (https://www.idemdatabase.org/terms) specify the constraints regarding use of the images: the generated images are stored and made accessible through the Mirador viewer of the Integrated Database for Early Music. There will be no limitation towards the consultation of the images without charge. However, download of the images is subject to the approval of the holding institution. Any constraints on the use of the data are defined in consultation with the holding institution of the original source. (e.g. copyright mention in case of publication) ***Dataset WP2.2 (Space):*** accessibility and download possibilities and/or restrictions regarding audio files will also become part of IDEM.
Are there any other legal issues, such as intellectual property rights and ownership, to be managed related to the data you (re)use? If so, please explain to what data they relate and which restrictions will be asserted.	☐ Yes ☒ No If yes, please explain:

2. Documentation and Metadata

Clearly describe what approach will be followed to capture the accompanying information necessary to keep **data understandable and usable**, for yourself and others, now and in the future (e.g. in terms of documentation levels and types required, procedures used, Electronic Lab Notebooks, README.txt files, Codebook.tsv etc. where this information is recorded). *RDM guidance on documentation and metadata*.	***Dataset WP1.1 (Music):*** During ingest images are processed by means of a precise method of filenaming so that individual files can be easily identified after ingest. In case of clear separate production units inside sources, a specific table of content is provided to aid in navigating the source. This happens by means of a metadata-file (.xlsx) and uses a Dublin core metadata standard. ***Dataset WP1.2 (Music):*** Data is gathered directly in the developed datamodel in CollectiveAccess manualy of by means of data import. Sufficient datafields are provided to give the necessary documentation and notes. ***Dataset WP1.3 (Music):*** A clear folder structure is sufficient since the dataset consists of rather individual work documents. ***Dataset WP2.1 (Space):*** For each room that is modelled the raw measurement data is collected, accompanied by a text file with a clear description of its representation, and an image thumb file of the room. The text data includes, but is not limited to, location, time, historical context, possible links to the musical corpus studied in this project, and a description of the room type according to standard AES69-2022. Specifically, for the acoustic measurements, the equipment used to generate and record the sound signals will be detailed in a rich text document in which pictures of the experimental setup will be included in as well. Moreover, all source and microphone positions used for the sound generation and recording will be accurately measured and their 3-D coordinates will be saved in a CSV file. For sources and microphones that follow a dynamic spatial trajectory during the recordings, the 3-D coordinates of a set of positions along the trajectory will be saved in a CSV file along with the corresponding time stamps. These CSV files will also be published along with the data.

	Datasets WP2.2&2.3 (Space): The descriptive part of the metadata for the audio file type and 3D video file type of recordings of music performances contains linked information: in a first phase the recordings will be documented using text and pictures. Descriptive text will include data such as (and not limited to): performing ensemble, performers (number of, name, ...), director, composition, composer, location, type of microphones used, microphone positions, etc.. Pictures will further illustrate this. In a second phase, with the integration into IDEM, metadata distilled from this information will ensure full interoperability between the audio file type and 3D video file type , and also to other data types within IDEM. ***Datasets WP3.1-3.3 (Text):*** Any data, software and models will be accompanied by a technical README (in markdown format) in the respective repositories (Zenodo and GitHub). On Zenodo, the distributed datasets will be assigned a DOI to increase the discoverability of the datasets. This file will detail the origin, current version and rationale of the relevant files. ***Datasets WP4.1-4.3 (Networks):*** Data will be made available to the highest degree possible through existing (metadata) platforms, channeling back refined information into the source reference works that we used as a starting point. As a complete collection, the gathered data will be made accessible through DANS https://dans.knaw.nl/nl/; or through Zenodo.

Will a metadata standard be used to make it easier to **find and reuse the data**? If so, please specify which metadata standard will be used. If not, please specify which metadata will be created to make the data easier to find and reuse. *REPOSITORIES COULD ASK TO DELIVER METADATA IN A CERTAIN FORMAT, WITH SPECIFIED ONTOLOGIES AND VOCABULARIES, I.E. STANDARD LISTS WITH UNIQUE IDENTIFIERS.*	***Dataset WP1.1 (Music):*** ☒ Yes ☐ No The mapping template for ingesting basic metadata related to the digital object into the long-term preservation (LTP) system uses the Dublin Core (DC) standard. Both DC elements and DC terms are available; The digital representations (JPEG2000) stored in Rosetta are represented as IIIF manifest files. These manifest files describe the structure of the object and associated metadata according to the specifications of the International Image Interoperability Framework (IIIF). This allows the digital objects to be visualized easily and anywhere in the world in any IIIF compliant viewer application. Mirador was chosen as viewer application and customized for the IDEM database in order to easily place images from different sources next to each other for comparative research; ***Dataset WP1.2&3 (Music):*** ☒ Yes ☐ No The specialized datamodel for descriptive metadata in CollectiveAccess can be linked to Dublin Core metadata standards, amongst others. Connections to other widely known databases as Répertoire International des Sources Musicales (RISM), and CantusDatabase contribute to the findability and reusability of the data. For authority records, connection is made with standards as VIAF. ***Datasets WP2.1-2.3 (Space):*** ☐ Yes ☒ No

	Since no formally acknowledged metadata standard specific to our discipline exists, no metadata standard will be used. A customised data model was developed by LIBIS (KU Leuven). The data model meets the research questions and the metadata needs recorded by the Alamire Foundation (KU Leuven). Special attention is given to the interrelations between the different data types: the integration of also audio files (audiorecordings of performances) and 3D video files (recordings of the same performances) into the IDEM data model; the addition of also the 3D models of spaces to the LIBIS data model. ***Datasets WP3.1-3.3 (Text):*** ☒ Yes ☐ No The project will use publicly documented open standards, such as UTF-8 encoded text files, in particular XML for hierarchical data modelling need and tabular spreadsheet formats (XLSX). An earlier paper that relied on this approach can be found here: https://zenodo.org/records/14732466. This repository illustrates how similar research have been shared before using FAIR principles. Any data, software and models will be accompanied by a technical README in the respective repositories (Zenodo and GitHub). On Zenodo, the distributed datasets will be assigned a DOI to increase the discoverability of the datasets. This file will detail the origin, current version and software requirements. ***Datasets WP4.1-4.3 (Networks):*** ☒ Yes ☐ No We will cater our metadata to the standard used by the destination platform. Following the FAIR principles, we will opt for a widely shared standard such as MARC21, which can be catered to specific needs, for example in capturing the bibliographic irregularities of Medieval Manuscript material. See for example: https://rbms.info/dcrm/amremm/

3. Data Storage & Back-up during the Research Project

Where will the data be stored? *Consult the interactive KU Leuven storage guide to find the most suitable storage solution for your data.*	☐ OneDrive (KU Leuven) ☐ Sharepoint online ☐ Sharepoint on-premis ☐ Large Volume Storage ☐ Digital Vault ***Dataset WP1.1 (Music):*** ☒ Other: Image data will be stored in a long-term preservation repository Teneo (previously named Rosetta) with unlimited storage capacity. It archives and provides persistent access to born-digital or digitised data objects and datasets that need to be kept for the long term, and includes metadata, usage policies and access tools. ***Dataset WP1.2 (Music):*** ☒ Other: Descriptive metadata will be stored in the specialized datamodel developed in the Data Management System (DMS) of CollectiveAccess. ***Dataset WP1.3 (Music):*** ☒ OneDrive (KU Leuven) Research documents will be stored on a personal OneDrive.

	Datasets WP2.1-2.3 (Space): ☒ Other: In a first phase, the raw recorded data will be stored locally in the data center of the Library of Voices (Alamire Foundation).New data generated from these files during post-production and integration of media will also be stored on the same server. A selection of these files will in a second phase, with the integration into the IDEM data model, be stored and managed in IDEM. A master copy of all measurement data and corresponding metadata will also be stored on ESAT servers. Copies can be made and kept on personal devices. ***Datasets WP3.1-3.3 (Text):*** ☒ Other: Initially, all data will be created and stored on the project collaborators' personal laptops that are backed up on a regular (weekly) basis by MAC OS X's Time Machine application (to external hard disks in separate physical locations). On a 6-monthly basis (roughly), we aim to release, incrementally expand and update datasets on Zenodo. ***Datasets WP4.1-4.3 (Networks):*** ☒ Sharepoint on-premis Leiden University will establish a 'Research Drive' which is a closed data collection with restricted access for researchers involved in the research project, for the duration of the project.

How will the data be backed up? *WHAT STORAGE AND BACKUP PROCEDURES WILL BE IN PLACE TO PREVENT DATA LOSS?*	☐ Standard back-up provided by KU Leuven ICTS for my storage solution ☐ Personal back-ups I make (specify) ☐ Other (specify) ***Dataset WP1.1 (Music):*** ☒ Standard back-up provided by KU Leuven ICTS for my storage solution Backup of image data will be guaranteed in Teneo by LIBIS, using services provided by ICTS, KU Leuven's Department of ICT ***Dataset WP1.2 (Music):*** ☒ Standard back-up provided by KU Leuven ICTS for my storage solution Backup of descriptive metadata will be guaranteed in CollectiveAccess by LIBIS, using services provided by ICTS, KU Leuven's Department of ICT. ***Dataset WP1.3 (Music):*** ☒ Standard back-up provided by KU Leuven ICTS for my storage solution ***Datasets WP2.1-2.3 (Space):*** ☒ Standard back-up provided by KU Leuven ICTS for my storage solution

	Datasets WP3.1-3.3 (Text): ☒ Personal back-ups (specify) Project collaborators' personal laptops will be backed up on a regular (weekly) basis by MAC OS X's Time Machine application (to external hard disks in separate physical locations). ***Datasets WP4.1-4.3 (Networks):*** ☒ Other (specify) Data gathered within the Leiden branch of the project will be backed up locally. These drives are centrally hosted and regular backups are guaranteed. https://digitalscholarship.nl/rds/service/surf-researchdrive/
Is there currently sufficient storage & backup capacity during the project? If yes, specify concisely. If no or insufficient storage or backup capacities are available, then explain how this will be taken care of.	☒ Yes ☐ No If no, please specify:

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons? *CLEARLY DESCRIBE THE MEASURES (IN TERMS OF PHYSICAL SECURITY, NETWORK SECURITY, AND SECURITY OF COMPUTER SYSTEMS AND FILES) THAT WILL BE TAKEN TO ENSURE THAT STORED AND TRANSFERRED DATA ARE SAFE.* *Guidance on security for research data*	***Datasets WP1.1-1.3 (Music):*** The security of the data is guaranteed by LIBIS and by using services provided by ICTS, KU Leuven's Department of ICT. Both Teneo and CollectiveAccess have individual user logins for the authorized researchers and personel. ***Datasets WP2.1-2.3 (Space):*** The security of the data locally stored in the Library of Voices is guaranteed by ICTS. The data stored on the ESAT servers has access regulated by an access control list (ACL) that grants read-write access to the project owner read-only access to specific users. The ACL is managed by the project owner. Client computers can access the data using: SMB2 (or higher) from specific IP ranges NFSv4 from specific (IT managed) systems. ***Datasets WP3.1-3.3 (Text):*** Most of the data which will be used and created is free of copyright and is not privacy-sensitive, so that data access should not be restricted. For data backup policy, see below. ***Datasets WP4.1-4.3 (Networks):*** By using a shared Research Drive within the Leiden research line of the project, which guarantees restricted access for project members only.

What are the expected costs for data storage and backup during the research project? How will these costs be covered?	***Dataset WP1.1 (Music):*** The storage costs of Teneo by LIBIS are 1.700 EUR/TB/year, for the 2,5 TB over four years, this will come down to 16.320 EUR. Costs will be covered with funding from the KU Leuven and/or Alamire Foundation (structural and project resources). ***Dataset WP1.2 (Music):*** The recurrent storage (50 GB limit) and maintenance costs of the whole metadata model in CollectiveAccess are 2171,10 EUR/year. This is sufficient for storage of this new dataset. Costs will be covered with funding from the KU Leuven and/or Alamire Foundation (structural and project resources). ***Datasets WP2.1-2.3 (Space):*** Alamire Foundation will provide local data storage at the Library of Voices (estimated cost of 6.000 €) with mirroring at ICTS KU Leuven (estimated cost 150 €/TB/year)..Costs will be covered with funding from the KU Leuven and/or Alamire Foundation (structural and project resources). ESAT servers: There will be no additional costs for the storage and back up, as the data will be stored at the servers of the research unit, which are available for the researchers belonging to this unit. ***Datasets WP3.1-3.3 (Text):*** No significant foreseeable costs to be reported here: our use of the Faculty's server, Zenodo, Github and Transkribus is free. Hard drives will be bought of the project budget for small equipment. ***Datasets WP4.1-4.3 (Networks):*** The costs will be covered by Leiden Faculty of Humanities, partly through the overhead provided by FWO.

5. Data Preservation after the end of the Research Project	
Which data will be retained for at least five years (or longer, in agreement with other retention policies that are applicable) after the end of the project? In case some data cannot be preserved, clearly state the reasons for this (e.g. legal or contractual restrictions, storage/budget issues, institutional policies...). *Guidance on data preservation*	☒ All data will be preserved for 10 years according to KU Leuven RDM policy ☐ All data will be preserved for 25 years according to CTC recommendations for clinical trials with medicinal products for human use and for clinical experiments on humans ☐ Certain data cannot be kept for 10 years (explain)
Where will these data be archived (stored and curated for the long-term)? *Dedicated data repositories are often the best place to preserve your data. Data not suitable for preservation in a repository can be stored using a KU Leuven storage solution, consult the interactive KU Leuven storage guide.*	☐ KU Leuven RDR ☐ Large Volume Storage (longterm for large volumes) ☐ Shared network drive (J-drive) ☐ Other (specifiy): ***Datasets WP1.1 (Music):*** ☒ Other (specifiy): Long-term preservation (LTP) repository Teneo by LIBIS ***Datasets WP1.2&1.3 (Music):*** ☒ Other (specifiy): Data management system (DMS) Collective Access by LIBIS ***Datasets WP2.1-2.3 (Space):*** Alamire Foundation will provide local data storage at the Library of Voices with mirroring at ICTS KU

	Leuven. A master copy of all measurement data and corresponding metadata will be stored on ESAT servers. Copies can be made and kept on personal devices. ***Datasets WP3.1-3.3 (Text):*** ☒ Other (specifiy): All newly created datasets will be persistently archived on Zenodo, which guarantees long-term preservation of scientific data. Upon the finalization of the project, all data will be saved in an individual folder on the Microsoft SharePoint Sites of the Department of Literature (University of Antwerp). The folder will be accessible to the PhD student, the PhD advisors, and the research manager of the Faculty of Arts (Géraldine De Visscher). This ensures that an original copy of the relevant files will remain available in the host institution. More regular backups to Github and Zenodo (on a monthly basis) when relevant are described below. ***Datasets WP4.1-4.3 (Networks):*** ☒ Other (specifiy): Possibly DANS (https://dans.knaw.nl/nl/)
What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	***Dataset WP1.1 (Music):*** The storage costs of Teneo by LIBIS are 1.700 EUR/TB/year. Costs will be covered with funding from the KU Leuven and/or Alamire Foundation (structural and project resources). ***Datasets WP1.2&1.3 (Music):*** The recurrent storage (50 GB limit) and maintenance costs of the whole metadata model in

	CollectiveAccess are 2171,10 EUR/year. This is sufficient for storage of this new dataset. Costs will be covered with funding from the KU Leuven and/or Alamire Foundation (structural and project resources). ***Datasets WP2.1-2.3 (Space):*** Alamire Foundation will provide local data storage at the Library of Voices (estimated cost of 6.000 €) with mirroring at ICTS KU Leuven (estimated cost 150 €/TB/year)..Costs will be covered with funding from the KU Leuven and/or Alamire Foundation (structural and project resources). ESAT servers: There will be no additional costs for the storage and back up, as the data will be stored at the servers of the research unit, which are available for the researchers belonging to this unit. ***Datasets WP3.1-3.3 (Text):*** No significant foreseeable costs to be reported here: our use of the Faculty's server, Zenodo, Github and Transkribus is free. Hard drives will be bought of the project budget for small equipment. ***Datasets WP4.1-4.3 (Networks):*** The costs will be covered by Leiden Faculty of Humanities, partly through the overhead provided by FWO.

6. Data Sharing and Reuse

Will the data (or part of the data) be made available for reuse after/during the project? Please explain per dataset or data type which data will be made available. *Note that 'available' does not necessarily mean that the data set becomes openly available, conditions for access and use may apply. Availability in this question thus entails both open & restricted access. For more information: https://wiki.surfnet.nl/display/standards/info-eu-repo/#infoeurepo-AccessRights*	☒ Yes, as open data ☐ Yes, as embargoed data (temporary restriction) ☐ Yes, as restricted data (upon approval, or institutional access only) ☐ No (closed access) ☐ Other, please specify: ***Dataset WP1.1 (Music):*** ☒ Yes, as open data (with user registration) ***Datasets WP1.2-1.3 (Music):*** ☒ Yes, as open data ***Datasets WP2.1-2.3 (Space):*** ☒ Yes, as open data
If access is restricted, please specify who will be able to access the data and under what conditions.	***Dataset WP1.1 (Music):*** Image data will be available after registration on IDEM website. Download restrictions depending on the conditions and requirements of the holding institutions. With an explicit agreement of the holding institution, it is possible to facilitate the download of the images from the IDEM database thus guaranteeing an open access treatment for the cleared images.

Are there any factors that restrict or prevent the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)? Please explain per dataset or data type where appropriate.	☐ Yes, privacy aspects ☐ Yes, intellectual property rights ☐ Yes, ethical aspects ☐ Yes, aspects of dual use ☐ Yes, other ☒ No If yes, please specify:

Where will the data be made available? If already known, please provide a repository per dataset or data type.	☐ KU Leuven RDR ☐ Other data repository (specify) ☐ Other (specify) ***Datasets WP1.1-1.3 (Music):*** ☒ Other (specify) Integrated Database for Early Music (IDEM), powered by the web application Omeka S by LIBIS. ***Datasets WP2.1-2.3 (Space):*** ☒ Other (specify) The Online Database for Acoustic Heritage (ODAH) will be made available in an Open Access repository. This database will contain all spatial room impulse responses and raw acoustic audio recordings of this project as well as corresponding metadata and documentation describing this data. Either on the LIBIS platform (as support for the IDEM database – to be decided by the Alamire Foundation) or on Zenodo, TBD. ***Datasets WP3.1-3.3 (Text):*** Smaller data will be publicly shared on Github during development. For more persistent archiving, the data will be mirrorred on Zenodo. ***Datasets WP4.1-4.3 (Networks):*** As the datasets will contain mostly qualitive, rather than quantitative data, tey will enrich other relevant platforms, e.g. IDEM, Dutch Song Database.

When will the data be made available?	☐ Upon publication of research results ☐ Specific date (specify) ☐ Other (specify) ***Datasets WP1.1-1.3 (Music):*** ☒ Other (specify) Image and descriptive data will gradually be made available in the course of the project ***Datasets WP2.1-2.3 (Space):*** ☒ Other (specify) During the review phase at the time of publication ***Datasets WP3.1-3.3 (Text):*** ☒ Other (specify) Minimally upon publication of the associated publications. ***Datasets WP4.1-4.3 (Networks):*** ☒ Other (specify) Upon completion of dataset/completion of project time. Intermediate sharing of information with relevant platforms is possible.

Which data usage licenses are you going to provide? If none, please explain why. *A DATA USAGE LICENSE INDICATES WHETHER THE DATA CAN BE REUSED OR NOT AND UNDER WHAT CONDITIONS. IF NO LICENCE IS GRANTED, THE DATA ARE IN A GREY ZONE AND CANNOT BE LEGALLY REUSED. DO NOTE THAT YOU MAY ONLY RELEASE DATA UNDER A LICENCE CHOSEN BY YOURSELF IF IT DOES NOT ALREADY FALL UNDER ANOTHER LICENCE THAT MIGHT PROHIBIT THAT.* *Check the RDR guidance on licences for data and software sources code or consult the License selector tool to help you choose.*	☒ CC-BY 4.0 (data) ☐ Data Transfer Agreement (restricted data) ☐ MIT licence (code) ☐ GNU GPL-3.0 (code) ☐ Other (specify)
Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, please provide it here. *INDICATE WHETHER YOU INTEND TO ADD A PERSISTENT AND UNIQUE IDENTIFIER IN ORDER TO IDENTIFY AND RETRIEVE THE DATA.*	☒ Yes, a PID will be added upon deposit in a data repository ☐ My dataset already has a PID ☐ No

What are the expected costs for data sharing? How will these costs be covered?	***Datasets WP1.1-1.3 (Music):*** Recurrent costs for Omeka S website application by LIBIS are 1437,47 EUR/year. Costs will be covered with funding from the KU Leuven and/or Alamire Foundation (structural and project resources). ***Datasets WP2.1-2.3 (Space):*** Costs will be covered with funding from the KU Leuven and/or Alamire Foundation (structural and project resources). ESAT servers: There will be no additional costs for the storage and back up, as the data will be stored at the servers of the research unit, which are available for the researchers belonging to this unit. ***Datasets WP3.1-3.3 (Text):*** No significant foreseeable costs to be reported here: our use of the Faculty's server, Zenodo, Github and Transkribus is free. Hard drives will be bought of the project budget for small equipment. ***Datasets WP4.1-4.3 (Networks):*** NA, as data sharing will happen on other relevant platforms.

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	KU Leuven (Musicology; ESAT), Alamire Foundation, Universiteit Antwerpen, Leiden University
Who will manage data storage and backup during the research project?	LIBIS and KU Leuven ICTS
Who will manage data preservation and sharing?	Alamire Foundation in collaboration with LIBIS and KU Leuven ICTS
Who will update and implement this DMP?	KU Leuven - Musicology: David Burn