



HAL
open science

D1.6 Intermediate validation reports for pilots TONALITIES, TUNES, MUSICBO and CHILD 1st version

Christophe Guillotel-Nothmann, Marco Gurrieri, Chukwudi Festus, Adam
Filaber, Peter Kranenburg, Van, Eleonora Marzi, Arianna Graciotti

► To cite this version:

Christophe Guillotel-Nothmann, Marco Gurrieri, Chukwudi Festus, Adam Filaber, Peter Kranenburg, Van, et al.. D1.6 Intermediate validation reports for pilots TONALITIES, TUNES, MUSICBO and CHILD 1st version. IReMus. 2023. halshs-04079571

HAL Id: halshs-04079571

<https://shs.hal.science/halshs-04079571v1>

Submitted on 24 Apr 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - ShareAlike 4.0 International License



*Polifonia: a digital harmoniser for musical heritage knowledge,
H2020*

**D1.6 Intermediate validation reports for pilots TONALITIES, TUNES, MUSICBO and
CHILD 1st version**

Deliverable information	
WP	WP1 Pilots and Web portal
Document dissemination level	PU Public
Deliverable type	R Document, report
Lead beneficiary	CNRS
Contributors	
Document status	Final version
Document version	V1.0
Date	28/02/2023
Authors	Chukwudi Festus Uwasomba, OU Adam Filaber, CNRS-IReMus Christophe Guilloitel-Nothmann, CNRS-IReMus [task leader] Marco Gurrieri, CNRS-IReMus Peter van Kranenburg, KNAW Eleonora Marzi, UNIBO Arianna Graciotti, UNIBO





INTENTIONALLY BLANK PAGE



Project information

Project Start Date: 1st of January 2021

Project Duration: 40 months

Project website: <http://www.polifonia-project.eu>

Project contacts

Project Coordinator

Valentina Presutti

ALMA MATER STUDIORUM - UNIVERSITÀ DI BOLOGNA

Department of Language, Literature and Modern Cultures (LILEC)

E-mail: valentina.presutti@unibo.it

Project Manager

Marta Clementi

ALMA MATER STUDIORUM - UNIVERSITÀ DI BOLOGNA

Research division

E-mail: marta.clementi3@unibo.it

POLIFONIA consortium

No.	Short name	Institution name	Country
1	UNIBO	ALMA MATER STUDIORUM - UNIVERSITÀ DI BOLOGNA	Italy
2	OU	THE OPEN UNIVERSITY	United Kingdom
3	KCL	KING'S COLLEGE LONDON	United Kingdom
4	NUI GALWAY	NATIONAL UNIVERSITY OF IRELAND GALWAY	Ireland
5	MiC	MINISTERO DELLA CULTURA	Italy
6	CNRS	CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE CNRS	France
	SORBONNE	SORBONNE UNIVERSITE (LinkedTP)	France
7	CNAM	CONSERVATOIRE NATIONAL DES ARTS ET METIERS	France
8	NISV	STICHTING NEDERLANDS INSTITUUT VOORBEELD EN GELUID	Netherlands
9	KNAW	KONINKLIJKE NEDERLANDSE AKADEMIE VAN WETENSCHAPPEN	Netherlands



10	DP	DIGITAL PATHS SRL	Italy
----	----	-------------------	-------



Project Summary

European musical heritage is a dynamic historical flow of experiences, leaving heterogeneous traces that are difficult to capture, connect, access, interpret, and valorise. Computing technologies have the potential to shed a light on this wealth of resources by extracting, materialising and linking new knowledge from heterogeneous sources, hence revealing facts and experiences from hidden voices of the past. *Polifonia* makes this happen by building novel ways of inspecting, representing, and interacting with digital content. Memory institutions, scholars, and citizens will be able to navigate, explore, and discover multiple perspectives and stories about European Musical Heritage.

Polifonia focuses on European Musical Heritage, intended as musical contents and artefacts – or music objects – (tunes, scores, melodies, notations, etc.) along with relevant knowledge about them such as: their links to tangible objects (theatres, conservatoires, churches, etc.), their cultural and historical contexts, opinions and stories told by people having diverse social and artistic roles (scholars, writers, students, intellectuals, musicians, politicians, journalists, etc.), and facts expressed in different styles and disciplines (memoire, reportage, news, biographies, reviews), different languages (English, Italian, French, Spanish, and German), and across centuries.

The overall goal of the project is to realise an ecosystem of computational methods and tools supporting discovery, extraction, encoding, interlinking, classification, exploration of, and access to, musical heritage knowledge on the Web. An equally important objective is to demonstrate that these tools improve the state of the art of Social Science and Humanities (SSH) methodologies. Hence their development is guided by, and continuously intertwined with, experiments and validations performed in real-world settings, identified by musical heritage stakeholders (both belonging to the Consortium and external supporters) such as cultural institutes and collection owners, historians of music, anthropologists and ethnomusicologists, linguists, etc.



Executive summary

Polifonia is driven by ten pilot studies, which provide both the validation criteria and the input requirements to the other research and development work packages. The pilots are heterogeneous in terms of knowledge domains, e.g., bells heritage, popular Irish music, history of music in Bologna, music influence on children. They involve interdisciplinary teams that bring different experience and methodological practices. This deliverable describes the efforts made so far, in connection with WP5, for the validation of four pilot studies dedicated to the study of musical cultural heritage: Tonalties, Tunes, MusicBo, and Child. A validation framework, consisting of “core criteria”, shared by all the pilots, and “pilot-specific validation methods”, is applied by the pilots and adapted to their specific needs. The results obtained are subjected to a critical analysis. Finally, lessons are drawn from this first validation step for the final validation (M40). Six short appendices make this report self-contained, including a glossary of terms, a description of the four pilots, and a description of the *Polifonia* work packages’ organisation.

Document history

Version	Release date	Summary of changes	Author(s) -Institution
V0.1	25/10/2022	First draft released	CNRS
V0.2	15/01/2023	Protocols defined by pilot leaders	CNRS, KNAW, UNIBO, OU
V0.3	16/01/2023	First version sent to internal reviewers	CNRS, KNAW, UNIBO, OU
V0.4	14/02/2023	Revised version according to the remarks and suggestions of the internal reviewers	CNRS, KNAW, UNIBO, OU
V0.5	23/02/2023	Final version sent to the coordinator	CNRS, KNAW, UNIBO, OU
V1.0	28/02/2023	Submission to EU	UNIBO

Table of contents

PROJECT INFORMATION	3
PROJECT CONTACTS	3
POLIFONIA CONSORTIUM.....	3
EXECUTIVE SUMMARY	6
DOCUMENT HISTORY.....	6
INTRODUCTION.....	9
TONALITIES	11
PRESENTATION OF THE VALIDATION PROTOCOL	11
1. Core validation criteria	11
a) XD and UX Design.....	11
b) Ecosystem and Rulebook	15
c) Web portal.....	16
d) Survey.....	16
2. Pilot-specific methods.....	18
a) Input validation	19
Quantitative validation of input data	19
Validation of input data quality by experts	19
b) Validation of tools/means/software.....	20
Validation of interfaces by users/experts.....	22
Validation of performance, requests/information passed between front and back end.....	25
c) Output/goal validation	26
Validation of results by ground truth.....	26
Validation of objectives through experts.....	26
APPLICATION OF THE VALIDATION PROTOCOL AND RESULTS OBTAINED	30
a) Input validation	30
Quantitative validation of input data	30
Validation of input data quality by experts	30
b) Validation of tools/means/software.....	31
Validation of interfaces by users/experts.....	31
Validation of performance, requests/information passed between front and back end.....	32
c) Output/goal validation:	32
Validation of results by ground truth.....	32
Validation of objectives through experts.....	32
LESSONS LEARNED AND FUTURE PERSPECTIVES.....	33
TUNES	34
PRESENTATION OF THE VALIDATION PROTOCOL	34
1. Core validation criteria	34
a) XD and UX Design.....	34
b) Ecosystem and Rulebook	34
2. Pilot-specific methods.....	35
APPLICATION OF THE VALIDATION PROTOCOL AND RESULTS OBTAINED	36
LESSONS LEARNED AND FUTURE PERSPECTIVES.....	39
MUSICBO.....	39
PRESENTATION OF THE VALIDATION PROTOCOL	39
1. Core validation criteria	39
a) XD and UX Design.....	39
b) Ecosystem and Rulebook	40
2. Pilot-specific methods.....	41
APPLICATION OF THE VALIDATION PROTOCOL AND RESULTS OBTAINED	47
LESSONS LEARNED AND FUTURE PERSPECTIVES.....	47



CHILD	48
PRESENTATION OF THE VALIDATION PROTOCOL	48
1. <i>Core validation criteria</i>	48
2. <i>Pilot-specific methods</i>	50
APPLICATION OF THE VALIDATION PROTOCOL AND RESULTS OBTAINED	52
LESSONS LEARNED AND FUTURE PERSPECTIVES	58
FAIR PROTOCOL	58
CONCLUSION.....	66
REFERENCES	68
APPENDICES	70
APPENDIX 1: GLOSSARY	73
APPENDIX 2: LIST OF THE PILOTS	74
APPENDIX 3: WP1 PARTICIPANTS, EFFORTS, OBJECTIVES, TASKS AND DELIVERABLES.....	75
APPENDIX 4: WORK PACKAGES IN <i>POLIFONIA</i> PROJECT	77
APPENDIX 5: STORY (PERSONA–QC–SCENARIO) TEMPLATE	80
APPENDIX 6: SURVEY TEMPLATE.....	81



Introduction

Polifonia is driven by 10 pilot studies¹ that interact with and address the preservation, management, and study of musical heritage (MH)². Focused on real-world music cases – such as Italian bell heritage, organs in the Netherlands, access to musical heritage for deaf people, and the identification of recurring melodic patterns – these pilots provide a means to validate *Polifonia*'s computational solutions; to strengthen the project's outreach; and to integrate new scenarios, early-adopters and stakeholders. They are developed by interdisciplinary teams of researchers from the humanities and social sciences, and from computer science.

This deliverable focuses on the mid-term validation of the four pilots that deal with the study of musical heritage: Tonalities³, Tunes⁴, MusicBo⁵, and Child⁶. It is a first step in the pilot validation process which will culminate in the D1.8 deliverable at the end of the project at M40⁷. The main objectives at this stage are to a) present the individual pilot's validation protocols; b) apply these protocols and present the results obtained so far; and, c) identify the lessons learned from the mid-term validation and address the avenues to be developed, strengthened, and reconsidered in view of the final validation.

The validation protocols were designed on the basis of a shared validation framework⁸ developed in connection with WP5 and presented in detail in the second version of *Polifonia*'s socio-technical Roadmap at M18 (D1.2)⁹. This validation framework aims to ensure that:

- O1: The pilot objectives reflect concrete needs expressed by the target communities and contribute to the state of the art;
- O2: the methods and approaches are sound;
- O3: the data on which the pilots are based are of sufficient quality to achieve the desired objectives;
- O4: the objectives and results expressed are achieved;
- O5: the resources, methods, and tools provided are made explicit and integrated into the *Polifonia* Ecosystem, interconnected through the PON and conveyed through the Web portal.

The Validation framework is modular in nature. It includes core criteria with which all pilots must comply and specific validation methods for individual pilots or pilot groups. In table 1 the pilots considered in this deliverable (Tonalities, Tunes, MusicBo and Child) are identified in red.

¹ The pilots are listed and described in Appendix 2.

² A glossary of abbreviations, acronyms and key concepts is provided in Appendix 1 with definitions where necessary.

³ <https://polifonia-project.eu/pilots/tonalities/>.

⁴ <https://polifonia-project.eu/pilots/tunes/>.

⁵ <https://polifonia-project.eu/pilots/musicbo/>.

⁶ <https://polifonia-project.eu/pilots/child/>.

⁷ D1.8: Final ten-pilots validation report and lessons learned.

⁸ https://github.com/polifonia-project/socio-technical_roadmap/tree/main/validation_framework.

⁹ Christophe Guillotel-Nothmann *et al.* *D1.2 Roadmap and pilot requirements 2nd version*, European Commission, Grant Agreement number: 101004746 — *Polifonia* — H2020-SC6-TRANSFORMATIONS-2018-2019-2020 / H2020-SC6- TRANSFORMATIONS-2020, Research Report, 2022.



Core validation criteria	XD and UX Design										✓
	Ecosystem + Rulebook										
	Web Portal										
	Survey (with success criteria, KPIs and validation setup)										
Pilot specific methods	Input validation	Validation of mechanically produced input data by ground truth									
		Quantitative validation of input data		✓					✓		
		Validation of input data quality by experts		✓		✓		✓	✓	✓	✓
	Validation of tools/means/software	Validation of interfaces by users/experts	✓	✓	✓	✓	✓	✓	✓	✓	✓
		Performance validation				✓					✓
		Validation of performance, requests/information passed between front and back end	✓			✓			✓		
	Output/goal validation	Compliance of the data produced with standards		✓							
		Validation of results by ground truth					✓		✓	✓	
		Validation of results by experts					✓			✓	
		Validation from a sub-corpus/case study and transfer to a larger corpus		✓							✓
		Validation of objectives through experts	✓	✓					✓	✓	
			ACCESS	BELLS	ORGANS	FACETS	INTERLINK	MEETUPS	TONALITIES	TUNES	MUSICBO

Table 1. Validation framework, core validation criteria and pilot-specific methods

The compliance of the pilots to a set of common core tools, principles, and guidelines grants a methodological and conceptual coherence to the whole while leaving enough space for individual pilot-specific validation protocols: Applying a mix of XD¹⁰ and UX Design helps to unify the design process of the pilots, ensures that their objectives are in line with real needs, and enables to monitor the extent to which the requests and needs are actually implemented. The conformity of the pilots to the Ecosystem and its Rulebook¹¹ ensures that all components are correctly referenced and documented. As a place for disseminating, enhancing and interconnecting the pilots, the Web portal¹² helps to harmonize the way Polifonia’s resources can be presented, explored, and searched by both amateur and expert users. Finally, the Polifonia Survey¹³ provides meticulous documentation on the choices, orientations, resources, and objectives of each pilot and is the main source of information for the pilot’s individual validation strategies.

The pilot specific methods cover the “input data” to assess the quality of the musical data (recordings, scores, historical texts, etc.) and to identify what changes have been made to the dataset to suit the pilot’s requirements. The validation of tools/means/software aims at verifying both the robustness and the ergonomics of the technical solutions implemented. Finally, the output/goal validation covers the validation of both the objectives and the results of the pilot.

The following sections present the individual validation protocol for the four pilots, give a critical presentation of the results obtained, and draw lessons learned from these first validation attempts. The intermediate validation supposes that the pilots identify and describe all the elements and steps of their validation protocol. However, not all of these elements have yet to be fully defined and systematically put into practice at this stage. The second deliverable dedicated to the validation at the end of the project will have the threefold objective of a) updating the validation protocols; b)

¹⁰ Cfr. Eva Blomqvist *et al.* “Experimenting with eXtreme Design”, in *Proceedings of EKAW 2010*, ed. by P. Cimiano, H.S. Pinto, Springer, 2010, pp. 120-134; Eva Blomqvist *et al.* “Engineering Ontologies with Patterns - The eXtreme Design Methodology”, in *Ontology Engineering with Ontology Design Patterns, vol. 25. Studies on the Semantic Web*, ed. by P. Hitzler, A. Gangemi, K. Janowicz, A. Krisnadhi, V. Presutti, IOS Press, 2016, pp. 23-50.
¹¹ <https://polifonia-project.github.io/ecosystem/>.
¹² https://github.com/polifonia-project/web_portal.
¹³ <https://github.com/polifonia-project/survey>.



proposing their exhaustive application; and c) monitoring the resolution of problems, the optimization of performances, and the improvement of results since the present intermediate validation.

A global overview of the methods, results, and lessons learned is proposed in the conclusion of the deliverable.

Tonalities

Presentation of the validation protocol

Tonalities operates within the framework of the core validation criteria by sharing the project's methodological approaches (XD and UX Design), adhering to the Rulebook, making its datasets referenced into the Ecosystem, contributing to the Web Portal, and participating in the WP1 Survey.

Further, pilot specific validation methods have been identified according to Tonalities' needs and characteristics: a) validation of interfaces by users/experts, b) validation of performance, requests/information passed between front and back end, c) validation of results by ground truth, and d) validation of objectives by experts. The validation of the interface is crucial since Tonalities' analytical workflow is based on a collaborative annotation interface for music analysis¹⁴. The following sections describe how the Validation Framework is applied to Tonalities, the results obtained, and the lessons learned.

1. Core validation criteria

Following the common core tools, principles and guidelines for validation adopted, Tonalities' validation protocol is grounded on four criteria:

a) XD and UX Design

Tonalities' methodological approach involves a series of competency questions based on three Stories:

- 1) Conflicting Theoretical Interpretations¹⁵;
- 2) Create Relevant Corpus¹⁶;
- 3) Conflicting Analytical Annotations¹⁷.

These competency questions are based on the Sethus Persona¹⁸ and focus on the peculiarities of the research planning, bringing to light crucial needs and research strategies. The UX Design methodology, developed in WP2 and WP5, has ameliorated Tonalities' collaborative annotation interface¹⁹ by preparing mock-ups²⁰ and interface prototypes (*via* Balsamiq Wireframes) in order to make it accessible, intuitive, clear, speedy, practical, and agreeable. The first elements considered

¹⁴ <https://github.com/sherlock-iremus/sherlock-tonalities>.

¹⁵ https://github.com/polifonia-project/stories/blob/main/Sethus:%20Music%20Theorist/Sethus%231_ConflictingTheoreticalInterpretations.md.

¹⁶ https://github.com/polifonia-project/stories/blob/main/Sethus:%20Music%20Theorist/Sethus%232_CreateRelevantCorpus.md.

¹⁷ https://github.com/polifonia-project/stories/blob/main/Sethus:%20Music%20Theorist/Sethus%233_ConflictingAnalyticalAnnotations.md.

¹⁸ <https://github.com/polifonia-project/stories/tree/main/Sethus:%20Music%20Theorist>.

¹⁹ <https://data-iremus.huma-num.fr/tonalities/>.

²⁰ https://github.com/polifonia-project/tonalities_pilot/blob/main/tonalitiesMockUp_1.bmpr.

in the mock-up design process were the interface's main features (register protocol, introduction to the interface, documentation such as guides or scenarios, etc.) as well as the operating fields (list of works, analytical tools, etc.) (figure 1a).

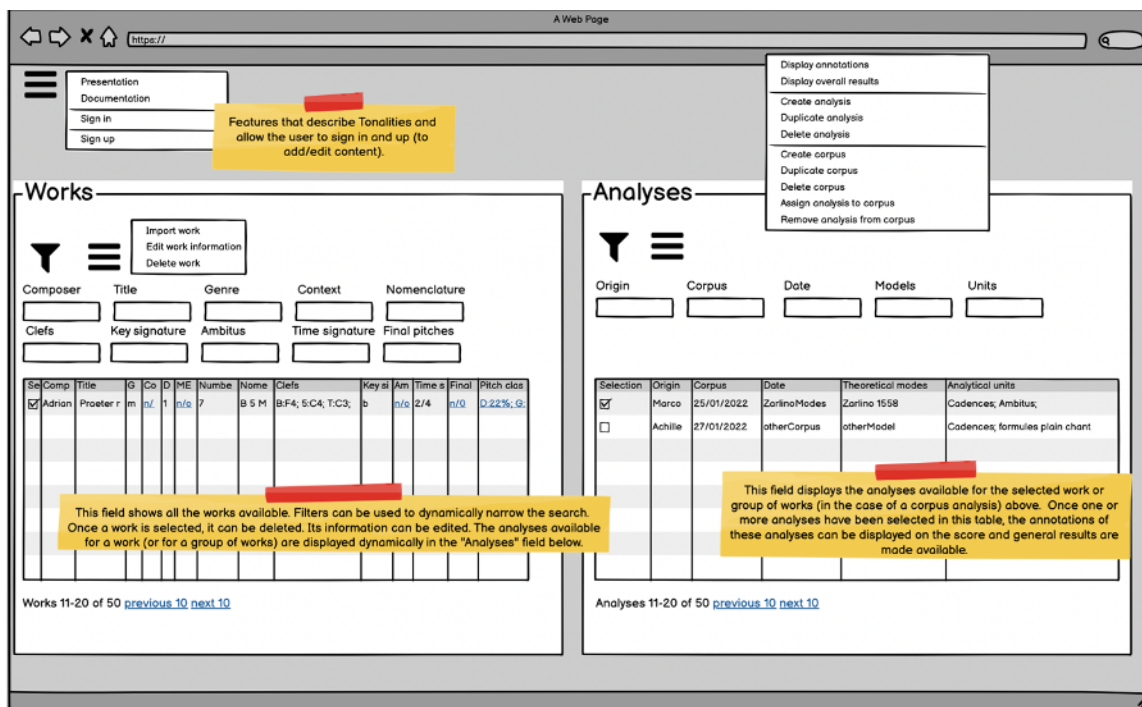


Figure 1a. Detail of the mock-up on the features and on the working fields

Several additional mock-ups have been designed to define in detail the analytical needs and tools; for example, they can associate analytical concepts to a note (figure 1b) or select multiple elements on the score to identify a cadence (figure 1c).



The screenshot shows the TONALITIES software interface. At the top, there's a browser address bar with 'https://'. Below it, a navigation bar has 'Theoretical model', 'Analysis', and 'Annotation' tabs. On the left, a list of composers is shown: Zarlino (1558), Praetorius (1619), Powers (1981), PDQ Bach (1757), Meeùs (2003), and Cathé (2012). The main area displays musical notation for six voices (C, S, A, T, Q, B) with lyrics. Annotations are placed over notes, such as 'C_A', 'C_P', 'C_F' for Soprano and 'B_A', 'B_P', 'B_F' for Tenor. A yellow callout box explains: 'Case 1: (individual score elements): the user clicks on an individual note and associates it with an analytical class. Once the score is displayed, the user selects a theoretical model ("Theoretical models" menu). The analytical concepts associated with the model are then displayed in the upper right-hand field.' Another yellow callout box explains: 'An analytical concept is associated with a note by clicking on the note and then selecting the concept in the upper right window. The analytical concept is displayed as an abbreviation in the score. Details about the concept are displayed and can be edited in the field at the bottom right. An annotation already made can be edited/deleted by clicking either on the note or on the annotation text.' On the right, a 'Theoretical concepts' menu is open, showing a tree structure: Cadence (Cantus, Antepenultimate, Penultimate, Finalis), Altus (Antepenultimate, Penultimate, Finalis), Tenor (Antepenultimate, Penultimate, Finalis), and Basus (Antepenultimate, Penultimate, Finalis). At the bottom, there are icons for Harmonic patterns, Imitations, Roots, and Harmonic patterns.

Figure 1b. Annotation case 1: selection of individual score elements

The screenshot shows the TONALITIES software interface. At the top, there's a browser address bar with 'https://'. Below it, a navigation bar has 'Theoretical model', 'Analysis', and 'Annotations' tabs. The main area displays musical notation for six voices (C, S, A, T, Q, B) with lyrics. A dashed rectangular selection box is drawn around a section of the score, covering parts of the Soprano, Alto, and Tenor staves. A yellow callout box explains: 'An analytical concept is associated with a score section either by selecting individual notes within this section (for example with shift + click) or by selecting two elements that correspond each a) to the lowest offset/highest voice and n) highest offset/lowest voice (for example cmd + click). A dotted frame is added to the score to identify x/y coordinates. Selected notes within this section have a distinct color.' On the right, a 'Theoretical concepts' menu is open, showing a tree structure: Cadence (Harmonic patterns, Imitations, Roots), and Basus (Harmonic patterns).

Figure 1c. Another annotation case (score selection of several elements)

Analytical annotations at all levels – both macrostructural and microstructural – can be the subject of free comments which in turn may lead to additional comments. The resulting argumentative chains and networks of claims are all directly or indirectly based on the application of explicit

theoretical models. The ergonomics with which the comments can be produced and visualized through the interface will be evaluated based on user feedback in the coming iterations of the interface validation.

A final aspect regulated through the mock-ups has been the part of the interface that will permit to compare the analytical results and visualise the schematic overviews of given analytical elements (figure 1d).

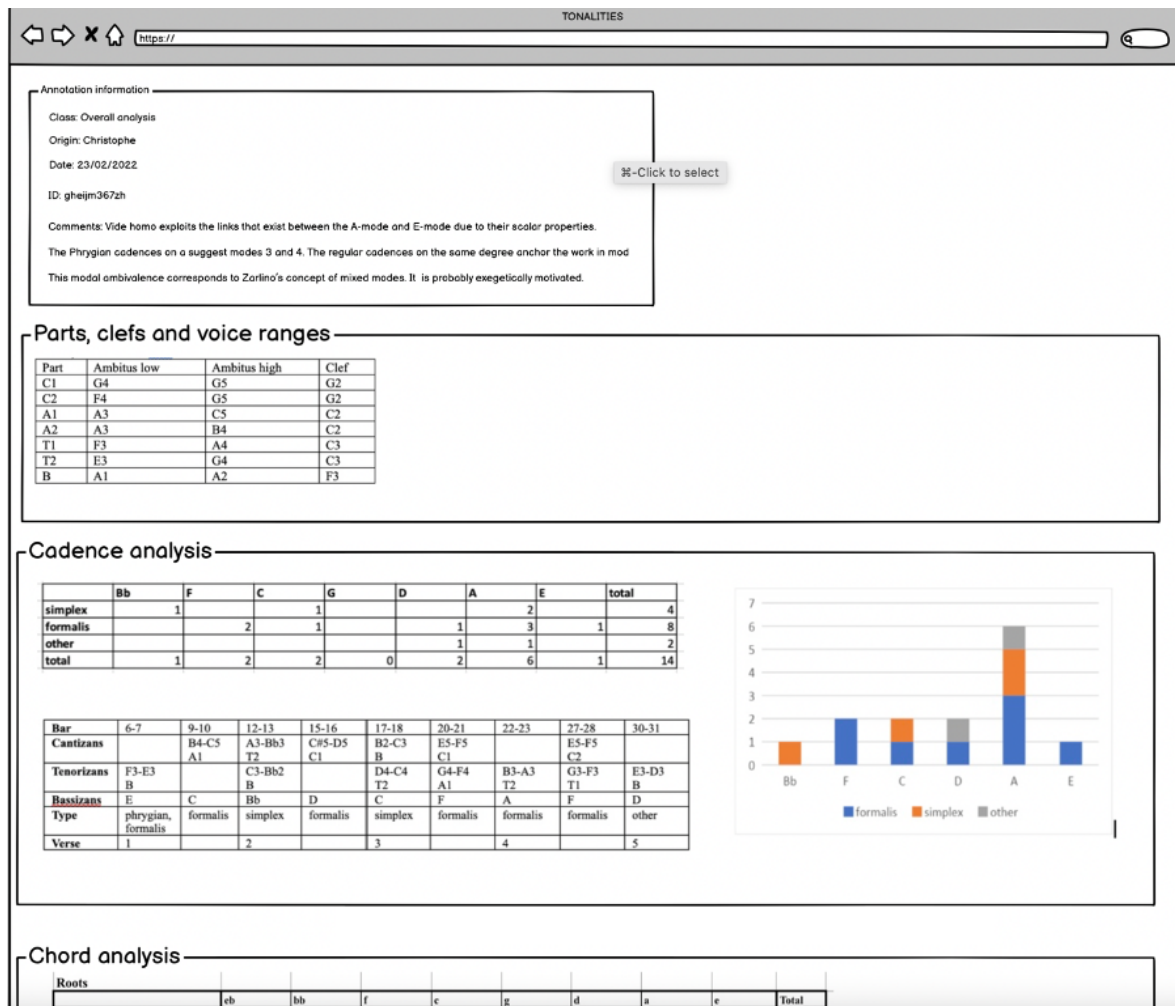


Figure 1d. Analytical results and schematic overviews

A rating system (from 1 up to 5) of IT technical complexities and of musicological priorities has been applied²¹ to establish an efficient action plan (Diagram 1). For example, creating a score and an analysis library has been classified as a high musicological priority and received a low rating (1) in terms of IT complexity. This element was thus quickly developed and integrated into the interface.

²¹ https://github.com/polifonia-project/tonalities_pilot/blob/main/prioritiesAnnotationTool.png.

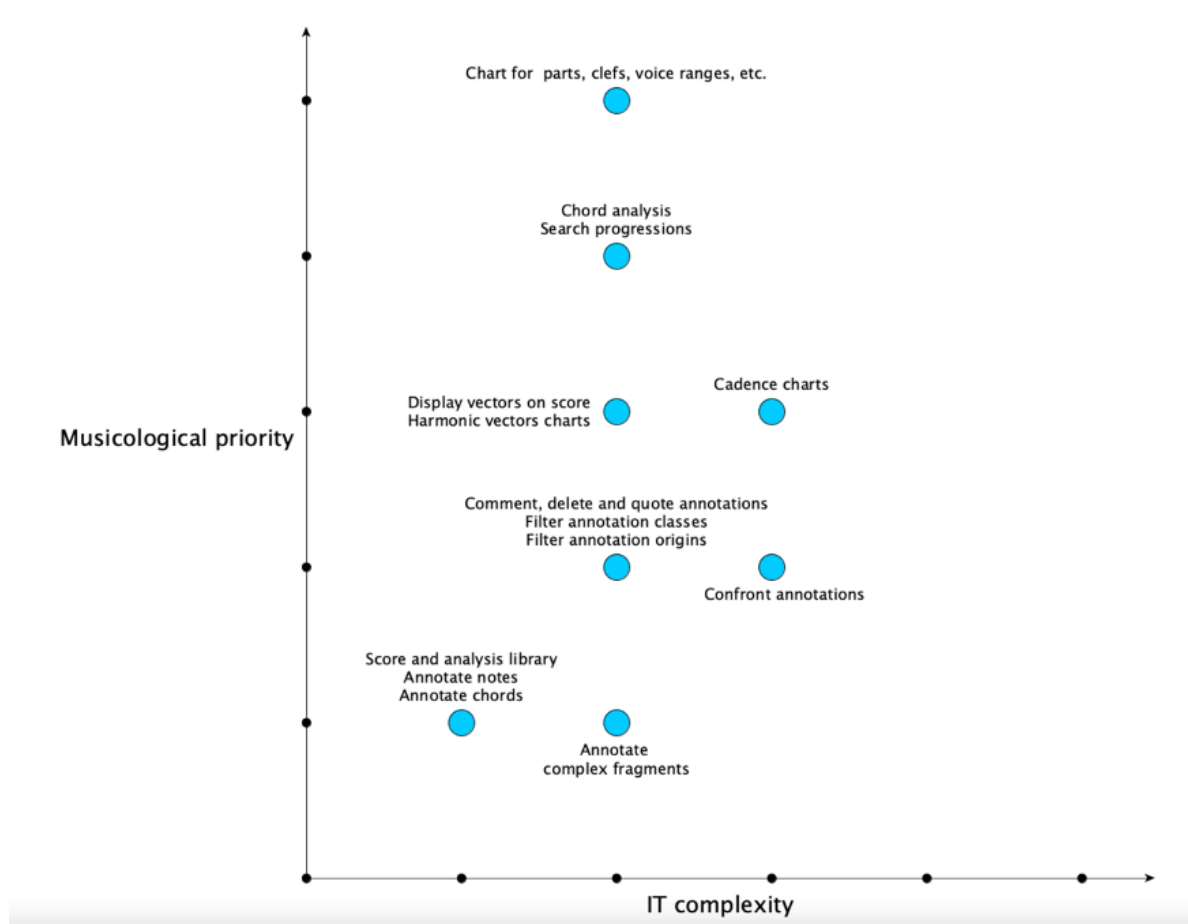


Diagram 1. Musicological priorities and IT technical complexity levels

Another UX Design criterium applied by Tonalities is the personalisation approach²², thanks to which the end user will be able to customize some aspects of his/her annotation process. For example, Tonalities already enables the end user to add a personal commentary to the annotations, provide a tool for labelling a simple selection, and suggest a personal analytical scenario not yet available in the interface: a new scenario could be, for example, to set up a series of new analytical tools for the application of a roman numeral analysis.

b) Ecosystem and Rulebook

Among the components made available in the Polifonia Ecosystem²³, Tonalities contributed to MusoW by providing detailed data about some web musical resources such as *The Josquin Research Project*, the *CRIM Project*, the *Marenzio Online Digital Edition*, etc²⁴. All these projects are important for Tonalities' analytical purposes, as they produced qualitative and addressable symbolic musical data in MEI format²⁵. They also constitute a valuable source of MEI scores in a quantitative viewpoint, meeting one of the main goals of the Tonalities pilot, that is to promote

²² It is a procedure that allows UX designers to customize user journeys and experiences according to the choices and requirements of customers. It is based on a deep understanding of the end user's needs and allows to personalise the content of a webservice.

²³ <https://polifonia-project.github.io/ecosystem/>.

²⁴ For a complete list of all the projects taken into account in Tonalities see table 5.

²⁵ <https://music-encoding.org>.



collaborative work on online large corpora. All components produced by Tonalities are compliant with the Rulebook and are referenced in the Ecosystem. These components include:

- Persona/Stories: Sethus, and three Stories related to;
- Software: Tonalities' Collaborative Annotation Interface for Music Analysis, once finalised, will be referenced in the Ecosystem;
- Data: the Tonalities ontology²⁶, a Satisfaction questionnaire²⁷, and a questionnaire for validating the objectives.

c) Web portal

Tonalities' main contributions to the Web portal can be summarized as follows: a) disseminating actions have been entered in Polifonia's Zenodo documentation list, which will be used to create a section of the Web portal dedicated to the scientific activity; b) three different sets of Competency Questions have been produced by Tonalities, contributing to the exploratory data analysis performed to understand the scope of pilots data, the distribution of entities and relations between entities, and the information patterns and the cross-connections between Polifonia pilots; and c) providing a first release of data *via* Melody²⁸ for communicating results and in progress outcomes of the pilot.

d) Survey

The Survey²⁹ has been the object of several implementations and iterations, as shown in table 2.

Iterations	Implementations	Timeline
1 st iteration		7 May 2021
2 nd iteration	First update with integration of the suggestions from the 1 st iteration	14 May 2021
	Minor format update, reorganised standards in the technical part, merged with DMP Annex 1	17 June 2021
3 rd iteration	Validation protocols and general update	15 November 2021

²⁶ <https://github.com/polifonia-project/modal-tonal-ontology>.

²⁷ [https://github.com/polifonia-project/tonalities_pilot/blob/main/Tonalities'%20Interface%20Satisfaction%20Questionnaire%20\(last%20version\).pdf](https://github.com/polifonia-project/tonalities_pilot/blob/main/Tonalities'%20Interface%20Satisfaction%20Questionnaire%20(last%20version).pdf).

²⁸ <https://projects.dharc.unibo.it/melody/tonalities/tonalities>.

²⁹ <https://github.com/polifonia-project/survey>.



4 th iteration	Objective and validation questions, and stakeholders	13 June 2022
	License added	18 June 2022
5 th iteration	Questions updates according to Polifonia's 6 th meeting	10 November 2022

Table 2. Survey's implementations and iterations

Tonalities' contributions to the different iterations helped to centralize information relating to the pilot's success criteria, validation setup, validation methods, KPIs and stakeholders competent for the validation (musicologists, music analysts, music students, and music lovers) (see table 3). The regular update of the survey allows to monitor the evolution, progress and application of Tonalities' validation setup. It also helps to compare and adjust Tonalities' validation protocol with other pilots.

Survey's validation section	Tonalities pilot
In order for the pilot to be successful, what criteria would the new tools and methods need to satisfy?	<p>Tools easy to use/ be understood in a documented historical/analytical/methodological environment.</p> <p>Certain tasks (i.e. search all harmonic patterns corresponding to a specific query) carried out more quickly</p> <p>Accessibility and addition of data from various sources</p> <p>Compare analytical results and annotations efficiently</p>
What will be the validation setup?	<p>Component 1: Goals and objectives: Do the objectives of the pilot correspond to real musicological needs? Do they take into account suggestions made by musicological colleagues?</p> <p>Component 2: Analytical results: Does the theoretical model allow for the correct inference of modal attributions? - Use a test corpus (c. 150 works) and check whether the attributions made by the theorists can be inferred from the models. - Use the justification for the modal inference produced by the reasoner as metrics?</p> <p>Component 3: Data processing: Is the analytical information provided by user(s) or computed by algorithms correctly stored and retrievable? - Use a test corpus of multi-layered analyses (c. 15 works which have more than 1 analysis) and check if the information layers are stored correctly</p> <p>Component 4: User engagement / ergonomics: Does the music annotation environment meet usability requirements? Can the CQ be answered with the interface and tools provided?</p>
What are the pilot's validation methods?	<p>Component 1: Validation of objectives through experts</p> <p>Component 2: Validation of results by ground truth</p> <p>Component 3: Validation of performance, requests/information passed between front and back end based on performance measures</p> <p>Component 4: Validation of interfaces by users/experts</p>



What KPIs or metrics will be used for validating it?	<p>Component 1: Usability survey/interviews about the pilot's general aims, questions and methodological orientations. Takes into account CQ.</p> <p>Component 2: Test corpus validated against ground truth (f-score), Number of models, size of current corpus</p> <p>Component 3: Time/Accuracy/Performance of requests/information passed between front and back end</p> <p>Component 4: Usability survey/interviews for the interface starting from the mock-up</p>
What values for KPIs/metrics shall be achieved for claiming success?	<p>Component 1: 90% positive evaluation of the user survey</p> <p>Component 2: 90% correct modal-tonal identification; at least 150 works with one analysis, at least 50 works with 2 or more analyses, at least 5 theoretical models;</p> <p>Component 3: to be specified</p> <p>Component 4: 90% positive evaluation of the user survey</p>
Can you name at least three (additional) stakeholders from different institutional backgrounds and countries who could contribute to the definition/validation of the pilot?	Frans Wiering, Anne-Emmanuelle Ceulemans, Richard Freedman, Dmitri Tymoczko, Tim Crawford, Jesse Rodin, Anne Cummings, David Lewis, Daniele Sabaino, Katelijne Schiltz, Laurent Pugin

Table 3. Tonalities' answers to the Survey's validation section

2. Pilot-specific methods

The pilot specific methods developed by Tonalities are summarized in table 4:

Specific methods	Tonalities
a) Input validation	/1) Quantitative validation of input data /2) Validation of input data quality by experts (case study: ZVP)
b) Validation of tools/means/software	/1) Validation of interfaces by users/experts /2) Validation of performance, requests/information passed between front and back end
c) Output/goal validation	/1) Validation of results by ground truth /2) Validation of objectives through experts

Table 4. Tonalities: pilot-specific validation methods

Concerning the Input validation, Tonalities adopted a a/1) Quantitative validation of input data for fixing the amount of data to produce till the end of the project (M40), while a a/2) Validation of input data quality by experts has been applied following a specific protocol for a case study on Zarlino.



The second group of specific methods dedicated to the Validation of tools/means/software is articulated into two main elements: b/1) Validation of interfaces by users/experts, for which a dedicated Satisfaction Questionnaire has been submitted to a group of stakeholders; and b/2) Validation of performance, requests/information passed between front and back end. Two other methods compose the Output/goal validation: c/1) Validation of results by ground truth, c/2) Validation of objectives through experts, for which a dedicated Questionnaire will be submitted in the coming months.

a) Input validation

Quantitative validation of input data

Tonalities aims to reach at the end of the project (M40) the following quantitative results:

- to propose at least 5 theoretical models;
- to make available circa 1000 music compositions in MEI scores, with at least 450 works with one analysis, and at least 150 works with 2 or more analyses.

Validation of input data quality by experts

The choice to work on music scores encoded in MEI files allows to enrich their XML code with additional tags providing all genres of information concerning the scientific and philological work made on the musical composition. Furthermore, the MEI code assures the addressability of each component through XML IDs.

Table 5 summarizes in the first column the corpora considered until now. These corpora are based on existing research projects that have already produced a consistent number of MEI scores which are made available in free access on the websites of each project (as listed in the third and fourth columns). In certain cases, the corpus of MEI scores has been directly created within Tonalities or produced in projects promoted by the IReMus lab. The second column contains the number of MEI scores that Tonalities aims to process. The total number of MEI scores amounts to roughly 1.000 items. The fifth and sixth columns respectively list the license of each project and the license adopted by Tonalities. Concerning the licenses, a detailed database has been elaborated³⁰.

Corpus	Number of scores considered in Tonalities	Project	URL	Original license	Final license
Bach's chorals	371	JSBChorales.net: Bach Chorales	http://www.jsbchorales.net/ http://neuma.humanum.fr/home/corpus/all:composers:bach:chorals/	CC BY-SA	CC-BY-NC 4.0
CRIM	100	Citations: the Renaissance	https://crimproject.org	CC-BY-NC 4.0	CC-BY-NC 4.0

30

<https://docs.google.com/spreadsheets/d/1cJ5oipaeJCibNVV9fX4sYMeEotiU69Gfea3bloF2TQU/edit#gid=0>.

		Imitation Mass Project			
Du Chemin	20	The Lost Voices Project	http://digitalduchemin.org	CC BY-NC 3.0	CC-BY-NC 4.0
Gesualdo	50	Gesualdo Online	https://ricercar.gesualdo-online.cesr.univ-tours.fr	GNU_GPL_v3	CC-BY-NC 4.0
Josquin's secular works	68	The Josquin Research Project	https://josquin.stanford.edu	CC-BY-SA 4.0	CC-BY-SA 4.0
Marenzio	50	Marenzio Online Digital Edition	http://www.marenzio.org/index.html	?	CC-BY-NC 4.0
Measuring polyphony	10	Measuring Polyphony. Digital Encodings of Late Medieval Music	https://measuringpolyphony.org	CC BY 3.0 US	CC-BY-NC 4.0
Saint-Saëns	100	Œuvres instrumentales complètes de Camille Saint-Saëns (1835-1921)	https://www.iremus.cnrs.fr/fr/programme-de-recherche/oeuvres-instrumentales-completes-de-camille-saint-saens-1835-1921	?	CC-BY-NC 4.0
Tasso	50	Tasso in Music Project	https://www.tassomusic.org	?	CC-BY-NC 4.0
Works quoted in Zarlino	192	Cantus Index Choral Wiki Global Chant Database IMSLP Petrucci Music Library The Josquin Research Project	http://cantusindex.org https://www.cpdl.org http://www.globalchant.org/index.php https://imslp.org/wiki/Main_Page https://josquin.stanford.edu	CC-BY-SA 4.0 CPDL copyright ? CC-BY-SA 4.0 CC-BY-SA 4.0	CC-BY-SA 4.0 CC-BY-NC 4.0 CC-BY-NC 4.0 CC-BY-SA 4.0 CC-BY-SA 4.0
Total	1000				

Table 5. Corpora of MEI scores considered in Tonalities pilot

Tonalities requires its final users to adopt their own specific editorial protocol to be applied to the corpus of works they aim to analyse. Each editorial protocol must be declared and published on GitHub. An example of such a protocol is the one applied to the Zarlino corpus, currently investigated by the Tonalities research team. This protocol represents a specific methodology that other similar case studies can use as a model.





Case study on Zarlino

A specific protocol – the Zarlino Validation Protocol (ZVP)³¹ – has been adopted for a test corpus containing all the musical works mentioned by Zarlino in the fourth book of his musical treatise *Le Institutioni harmoniche* (Venice, the author, 1558). This specific protocol establishes the following parameters:

1. pitches;
2. rhythmic values;
3. distribution of the lyrics (with a particular attention to the convenable syllabification and to the prosody);
4. articulation in sections or internal parts.
5. the adoption of the clefs of the original source;
6. the use of a conversion table between modern and Early music time signatures;
7. the adoption of the *integer valor*³²;
8. the conversion of the *Mensurstrich* in normal bar lines.

The attention to the clefs employed in the original source is due to the fact that their position and typology is an important indicator of a compositions’ modal attribution by the music theorists of that period. The remaining parameters serve to provide a standard version of the score to the end users and guarantee a uniform dataset for the IT team.

The ZVP can be subject to adaptations and modifications due to the progressing workflow.

b) Validation of tools/means/software

The main tool developed in the framework of Tonalities is a collaborative annotation interface for music analysis, for which two specific methods of validation have been developed, as described here after.

Validation of interfaces by users/experts

In order to validate its interface, Tonalities gathered a focus group of three experts, asking their feedback through a specific Satisfaction Questionnaire³³ shown in table 7.

1. How satisfied/dissatisfied are you with each of the following features?						
Interface features	Very satisfied	Moderately satisfied	Neutral	Moderately dissatisfied	Very dissatisfied	Please explain why
Registration protocol						
Presentation of the scores						

³¹ https://github.com/polifonia-project/tonalities_pilot/blob/main/validation/input-data-quality/zarlino_validation_protocol.

³² The value of the notes is the same of the one in the original source, so that a *minima* in the source is transcribed as a minim in the modern score.

³³ https://github.com/polifonia-project/tonalities_pilot/tree/main/validation/interface.



Organization of the interface						
Annotation tools (generally):						
a. Select a verticality						
b. Assign a root						
c. Create a selection						
d. Create a nested selection						
e. Delete a selection						
f. Select a theoretical model						
g. Find a class/property						
h. Apply a class/property						
i. Add an annotation						
j. Delete an annotation						
Visualisation of annotations						
Adding commentaries						



Comparing annotations						
Logging out						
General questions						
1. What is your most favourite feature of the interface? A:						
2. What is your less favourite feature of the interface? A:						
3. Please name any important features that are missing in the interface. R:						
4. In what ways the interface is better than the traditional process? A:						
5. Is it faster/slower than the traditional process? A:						
6. Do you have any suggestion(s) for improving the interface? A:						
7. What about:						
a. ergonomics? A:						
b. available features? A:						
c. intuitiveness/clarity? A:						
8. How likely are you to recommend the interface to a friend or colleague?						



	Very likely	Likely	Neutral	Unlikely	Very unlikely
9. Which analytical annotation tasks/scenarios would you like to be able to carry out with this annotation tool? A:					
10. Do you have any other suggestions? A:					

Table 7. Tonalities’ Interface Satisfaction Questionnaire

To answer the questionnaire, users were asked to carry out two scenarios.³⁴ These scenarios were initially prepared as teasers of a demonstration on the interface presented at the Music Heritage Knowledge Graphs³⁵. This scenario-guided methodology has been developed in collaboration with the WP5 and consists of asking the stakeholders to concretely apply the scenario in order to evaluate how it works in the questionnaire. The questionnaire is also meant to collect feedback on the problems, on the expectations of the scientific community, on the ergonomics and design about the interface. At the end of the project (M40), Tonalities aims to obtain at least a 3.0 score for this particular validation method, using the following rating:

- 0 = Very dissatisfied
- 1 = Moderately dissatisfied
- 2 = Neutral
- 3 = Moderately satisfied
- 4 = Very satisfied.

For its future new iterations (up to 2), we are planning to involve a larger group of stakeholders (up to 15), and to question other dimensions such as the grade of difficulty for finding the features to use. We will also ask our stakeholders to propose their own scenarios.

Validation of performance, requests/information passed between front and back end

For this specific method of validation concerning the interface, Tonalities operated into two main directions:

- To ensure that all analytical annotations made through the interface have an author and a date through a SparQL query checking the conformity of the graph.
- Despite the heterogeneity of the analytical annotation anchors (note, verticality, arbitrary selection), it is necessary to ensure that a unified and fast graph traversal can be used to establish the link between an annotation and the score containing its anchor. The complexity and structure of the RDF data graph varies depending on whether an annotation

³⁴ <https://nakala.fr/10.34847/nkl.2ba4qe52>.

³⁵ <https://iswc2022.semanticweb.org/index.php/workshops/>.



is about a note, a verticality, or an arbitrary selection of musical signs annotated with a complex musicological concept characterised by an internal structure (such as a cadence). For statistical reporting purposes (e.g. to get a count of all annotations on all scores in the corpus), it is important to be able to trace back from the annotation to the score via a simple RDF path that is independent of the complexity of the annotation's anchor. Thus, we created shortcuts between each annotation and the partition containing its anchor in the form of RDF predicates.

The Tonalities interface is based on the Sherlock annotation model, developed at the IReMus to produce a homogeneous representation, at the laboratory level, of musicological knowledge units, including editors, composers, annotators, theoreticians, etc. This model is intended to differentiate and interconnect the different "roles" a same person may have: for example, Zarlino as an author of theoretical models and a composer of motets, or a collaborator of the IReMus as a score editor and an analyst. etc. The creation/update of entities is done in an automated way in Tonalities: a) by user authentication and b) by importing works and theoretical models with explicit information (on composers, editors, theorists, encoders, etc.). The exploration of these roles is not planned at the level of Tonalities, but at the level of the Sherlock project.

c) Output/goal validation

The last subgroup covers the validation of both the results and the objectives of the pilots.

Validation of results by ground truth

Tonalities aims at classifying musical works into modal-tonal categories on the basis of an in-depth understanding of their inner organisation. Our starting point are a) the collection of scores made available and b) the modelling of historical and today's music theories. The workflow that has been implemented consists of extracting from the scores the musical knowledge (cadences, *ambitus*, alterations, etc.) on which the formal definitions of the ontologies are based. This extraction is done both on the basis of manual annotations and AI. The knowledge thus obtained is instantiated in the ontologies so that a reasoner (HermiT) can classify the works according to their modes. The validation protocol plans that the modal-tonal attributions be compared to ground truth, for example to the modal attributions made by the historical theorists themselves, once these attributions are verified by expert analysts.

Validation of objectives through experts

A first internal validation of the objectives has been defined within the core validation criteria, grounded on the Competency Questions. Their definition generated a progressive adaptation in Tonalities' research team through the elaboration of individual scenarios and personas. After several confrontations and internal debates, the research team converged on the Sethus' Story and Persona.³⁶ For the validation of the final objectives, Tonalities decided to involve also a group of experts and of potential end users, asking them to answer a specific Questionnaire. It is structured into 9 questions, introduced by a short presentation of each objective, and followed by decrescent rating answers.

³⁶ The different Stories related to Sethus in GitHub reflect this process of adaptation.



Questionnaire on the objectives³⁷

- Analysing large collections of digital scores in an open and interactive collaborative environment.** Tonalities aims at analysing large digital corpora of notated music at all levels of granularity and to process them in series³⁸: the end user can work on large corpora of symbolic music data (with a possible integration of on site transcriptions of ethnomusicological corpora), propose his/her analyses, save them, and make them addressable and easily quotable. The aim is to create an online workplace where the end users can add their own MEI music scores that are accessible from everywhere and that can be directly analysed on the screen. Several MEI scores are already available in the interface and come from existing web projects dedicated to music analysis like *The Lost Voices Project* or *The Josquin Research Project*. Tonalities' approach is complementary to all these projects, opening to an innovative online interaction with digital music scores.

How useful is this objective?

Very useful	Useful	Neutral	Useless	Less than useless

- Choosing a theoretical model.** Tonalities proposes to the end user the possibility to choose among several theoretical models, articulated in classes and properties that can be used to contextualise and to justify the user's analytical annotations (e.g.: this chord can be classified as incomplete according to this theoretical model). The end user can also propose to integrate new theoretical models into the interface. The possibility to apply different theoretical models, from different periods of the history of music and of different geographical origins, permits us to evaluate how the analysis changes through the lens of a given theoretical framework. For example, a 16th century madrigal composed in the Venetian area can be analysed using classes and properties modelled on Gioseffo Zarlino's *Le istituzioni harmoniche* (Venice, the author, 1558), but it can also be analysed applying a more recent theoretical viewpoint like Harold Powers' tonal types. According to the theoretical framework adopted, the analytical results may considerably differ.

How useful is this aspect?

Very useful	Useful	Neutral	Useless	Less than useless

³⁷ https://github.com/polifonia-project/tonalities_pilot/blob/main/validation/results-objectives/Questionnaire%20on%20the%20objectives.docx.

³⁸ With reference to Georges Molinié's concept of serial stylistics, serial musical analysis, as it is here intended, draws its conclusions by comparing works to which the same analytical procedure is applied by iteration. See Georges Molinié, & Alain Viala. *Approches de la réception : sémiostylistique et sociopolitique de Le Clézio*, Presses universitaires de France, 1993, pp. 43-46.



- 3. **Comparing different theoretical models.** Tonalities’ main epistemological challenge is to explore the internal properties of music in a large spectrum of repertoires (including the ethnomusicological ones), trying to make their mechanisms emerge through the lens of several theoretical viewpoints formulated along the centuries. In particular the choice to decline the term tonality in the plural form, as already done by Anthony Pople in his homonymous project³⁹, is meant to emphasise the great diversity of tonal organizations and theoretical understandings throughout history and intellectual and contexts. In this perspective, Tonalities adopts the methodological posture to engage the analysis to one or many theoretical frameworks that are made explicit. The end user is in the position to look “inside” these models, to compare them, and to evaluate their differences, changes and/or omissions.

How useful is the possibility to compare different theoretical models?

Very useful	Useful	Neutral	Useless	Less than useless

- 4. **Argued, documented and authored analyses.** The end user can formulate an argued, documented and authored analysis on the basis of the theoretical model(s) adopted. The choice of a theoretical model represents a specific viewpoint avoiding an arbitrary and not referenced approach. Each analytical observation formulated is authored and attributed to a user, whose identification is guaranteed by the registration via ORCID.

How useful is the possibility to formulate an argued, documented and authored analysis?

Very useful	Useful	Neutral	Useless	Less than useless

- 5. **Collaborative dimension.** The possibility to work in a collaborative dimension permits to analyse large corpora otherwise unlikely approachable by a single analyst. Tonalities’ interface offers, therefore, the possibility to coordinate a team of analysts coming from different parts of the world in order to apply either the same or different analytical approaches to an extended group of musical works.

How useful is the collaborative dimension of *Tonalities*?

Very useful	Useful	Neutral	Useless	Less than useless

³⁹ <https://research.hud.ac.uk/tonalities/>.



--	--	--	--	--

6. **Critical comments.** The end user can add his/her own critical comments on each level of the analytical process, from a single analytical annotation to the final conclusions eventually expressed. Comments can also be made on the analyses published by other users. This is intended to encourage an online debate able to host all the different opinions, to motivate the theoretical approaches, to convey the confrontation among several analytical viewpoints, and to make every level of complexity, ambiguity or nuances emerge. Each comment bears a date, is authored and addressable via dedicated urls.

How useful is the possibility to add critical commentaries?

Very useful	Useful	Neutral	Useless	Less than useless

7. **Publishability, quotability and reusability.** Tonalities aims to make the analyses publishable, quotable and reusable in innovative ways, giving the possibility to:

- a. export the analytical annotation,
- b. export the critical commentaries in text format files,
- c. export comparative diagrams and schemas.

Furthermore, Tonalities guarantees the addressability of the analytical annotations and of the critical comments produced.

How useful is this aspect?

Very useful	Useful	Neutral	Useless	Less than useless

8. **Suggestions.** Do you have any other objective to suggest?

A:

9. Please describe a user scenario relevant to your research work based on the (proposed and/or added) objectives.

At the end of the project (M40) Tonalities aims to obtain at least a score of 3.0 for this particular validation method, using the following rating:



- 0 = Useless.
- 1 = Almost useless.
- 2 = Neutral.
- 3 = Useful.
- 4 = Very useful.

Application of the validation protocol and results obtained

a) Input validation

Quantitative validation of input data

Corpus	Number of MEI scores already available
Bach's chorals	371
Works quoted in Zarlino	40
Total	411

Table 8: Corpora and MEI scores already available

Table 8 summarizes the number of MEI scores already available (411). This represents circa 40% of the final amount of scores Tonalities aims to reach at the end of the project (circa 1000 MEI scores).

Validation of input data quality by experts

The ZVP applied to the input datasets caused some sporadic problems of redundancy in the encoded MEI scores originating from The Josquin Research Project that were reused in the Zarlino corpus. They are likely due to the collaborative work carried out in some of the scores: some music symbols have been double encoded and superposed, which obfuscates the visualisation of the scores in the Tonalities Interface. This problem will necessitate a specific quality control protocol on the MEI scores produced in the framework of the Josquin Research Project.

Furthermore, the ZVP has been applied with specific attention to the Zarlino corpus. The input data provided to Tonalities guaranteed a starting point for the pilot's analytical approach and specific research purposes. Based on a small initial corpus of seven MEI scores already available in the interface, the ZVP permitted to:

1. make approximately 50 manual edits⁴⁰;
2. adopt the clefs of the original source in 100% of the cases;
3. apply the conversion table for the time signature to the 100% of the scores;
4. adopt the integer valor in two cases;
5. convert the original Mensurstrich into normal bar lines in one case.

⁴⁰ Correct 4 wrong pitches; Redefine 13 rhythmic values; Redistribute lyrics or to adopt the convenable syllabification linked to the prosody in 35 musical passages; Uniformise the articulation of 2 sections or internal parts.



b) Validation of tools/means/software

The results of this section of the Validation Frameworks concern the Tonalities interface.

Validation of interfaces by users/experts

A first iteration of the satisfaction questionnaire has been made in November 2022. This questionnaire has been structured into two parts: a first one taking into account the main features and functionalities already available in the interface; and a second, more general part, questioning about frustrations and needs of the involved experts, about the usability of the interface, and about its ergonomics. The group of experts has been chosen from a list of fifteen stakeholders formed by:

- 8 musicologists/music analysts: specialists on different periods, from the Renaissance to the 20th century (from Belgium, Canada, Italy, Netherlands, USA);
- 3 musicologists/computer scientists (from Switzerland, Great Britain, Netherlands);
- 2 digital musicologists (from Great Britain, USA);
- 1 music theorist (from Germany);
- 1 composer/philosopher (from the USA).

For this first iteration a small group of three experts has been built from the above-mentioned list. Others will be asked to give their feedback for the forthcoming iterations of the Questionnaire that will be adapted according to the improvements and developments of the interface. Furthermore, next versions of the Questionnaire will include the possibility to define and explore analytical scenarios.

As of today, on the basis of the Satisfaction Questionnaires collected⁴¹, the interface obtained a score of 2.1 while a score of 3.0 (“Moderately satisfied”) has been defined as the threshold. Two stakeholders declared that they would unlikely recommend this interface at this stage, while only one stakeholder would likely recommend it. All stakeholders suggested several improvements and made emerge some general problems, critics/suggestions on ergonomics, and corrections on the terminology adopted in the interface:

- the process of selection requires too many passages and manipulations;
- the process of annotation is time-consuming;
- a list of shortcuts for all the typologies of selection should be added;
- a system of pop-ups could be adopted to manage the classes and properties of the theoretical models;
- to implement an automatic system for generating some basic analysis (e.g. to track all the dissonances);
- ergonomics and intuitiveness/clarity need to be improved;
- the hotspot of the elements in the score must be enlarged in order to improve the selection options;
- some terminology needs to be corrected.

Improvements must, therefore, be made in the perspective of reducing clutter. A possible solution is to involve an expert in Web ergonomics. A general strategy also is to study the functionalities and the options available in other similar music webservices, to meet the end user’s habits and gestures. An extra effort must also be made to reduce the time-consuming of the music analytical

⁴¹ https://github.com/polifonia-project/tonalities_pilot.



process on the screen. This should be potentially be lesser than the same process conducted in the traditional way. Another important aspect is to let the end user suggest a personal and adapted analytical scenario, going beyond the possibilities currently offered by the Interface.

Validation of performance, requests/information passed between front and back end

The results of this specific validation method allowed to automatically verify some aspects concerning the production and registration of analytical annotations in the interface. In particular:

- a. to process all the analytical annotations initially entered in the interface in an automatic way through a SparQL query verifying that each of them has an author and bears a date;
- b. to make punctual verifications (randomly about 5) about the traceability of an analytical annotation (independently of its typology) in relation to the work of reference.

All the adopted mechanical validations between front and back end have given positive results showing that the Tonalities' interface is operative and its functionalities work correctly.

c) Output/goal validation:

The application of the two specific methods of this section is still in progress and gave only partial results.

Validation of results by ground truth

Two historical theories have been modelled so far: Zarlino 1558 and Praetorius 1619 (with about 2500 axioms, 500 classes and 100 object properties each). The Praetorius model has been presented⁴² and applied⁴³ in former works. Initial tests have been undertaken during the last weeks to infer the mode of individual works from these models using a reasoner (HermiT). In the coming months, it is planned to test the Zarlino model on a sample of 150 works and to validate the whole workflow against ground truth. We consider that a 90% recognition rate within the sample validates this feature.

Validation of objectives through experts

After the described internal and initial process of validation of Tonalities' objectives, the current issue is to make them validate and to spread their number and action through a questionnaire to be distributed to an external group of twelve potential end users (formed by musicologists, music analysts, and Conservatoire students) in the coming months⁴⁴. Other iterations of this

⁴² Christophe Guillotel-Nothmann. "Knowledge extraction and modelling in the project *Thesaurus Musicarum Germanicarum*", paper presented at the *Prague DH Workshops Session III: Editions*, Masaryk Institute and Archives of the Czech Academy of Sciences, 2020.

⁴³ Christophe Guillotel-Nothmann and Anne-Emmanuelle Ceulemans. "Das diatonisch-chromatische System zur Zeit des Michael Praetorius. Eine digitale Neuerschließung des *Syntagma musicum* (1619) in Verbindung mit dem Tanzzyklus *Terpsichore* (1612)," in *Musik im Umbruch. Michael Praetorius zum 400. Todestag*, Harrassowitz, 2021, pp.109-150; Anne-Emmanuelle Ceulemans and Christophe Guillotel-Nothmann. "Praetorius' *Polyhymnia caduceatrix* (1619): Diatonische Logik und Skalenlogik in mehrchörigen und konzertierenden Werken des Frühbarocks," *Zeitschrift der Gesellschaft für Musiktheorie, Gesellschaft für Musiktheorie, 2020, Alte Musik und Musiktheorie zwischen Historiografie und Pädagogik*, 17/2 (2020): 51–93.

⁴⁴ The list of interested participants has been collected by the Tonalities research team in the occasion of the *Journée d'analyse musicale* organised by the French Society for Music Analysis (SFAM) on the 10th December 2022 at the Berlioz Conservatoire in Paris.



questionnaire with a larger number of potential end users (up to 50) are expected for the M30 and M36.

Lessons learned and future perspectives

After having encountered some problems in the input validation protocol, we realised that the whole editing and verification process is extremely time-consuming. For this, we plan to integrate some further mechanical checks, like GioQoso⁴⁵.

Feedback from the stakeholders played a central role in identifying some important problems relating to the interface – missing documentation, ergonomics, terminology – that will be key for defining the next steps in the developing process.

The idea to associate usage scenarios to the Satisfaction questionnaire addressed to our team of stakeholders turned out to be successful, linking concrete step-by-step actions to the functionalities made available in the interface.

A major challenge is to take advantage of the collected feedback, evaluating its pertinence on the scientific level. Some feedback made us realise that we should involve not only experts in musicology but also experts in web conception and ergonomics. Another important lesson learned has been to communicate to our team of IT experts the problems merged through the stakeholders' feedback. In this perspective a dedicated working space has been created on GitHub⁴⁶, to list issues for each request, correction, or change.

For the moment, we notice that our team of stakeholders is very reactive and positively involved. However, we are aware that they almost all belong to the same musicological and analytical milieu, so that a larger and more heterogeneous community will be involved in the future.

Concerning the results, only the outcomes from the analytical annotations (e.g. this is a cadence, this is a syncopation...) have been evaluated for the moment. Beyond these annotations, there is a need to evaluate the ability to comment these annotations and, to produce argumentative chains and discussion networks. At this stage, this kind of result is not explicitly considered in the validation protocol. We are considering the possibility to go in this direction, eventually by means of a satisfaction questionnaire on the heuristic contribution of this dimension. At a second stage, once the progress towards the achievement of the objectives is more advanced, one of the issues will also be to check not only whether the objectives are important, but whether they have been implemented in a satisfactory way for the user.

Following the user feedback, a number of measures have already been taken⁴⁷ against the danger to fall into the problem of “paralysis of choice”. The problem is generated by the high number of properties and classes present in several theoretical models. We reduced the high number of

⁴⁵ <http://giogoso.irisa.fr>.

⁴⁶ <https://github.com/sherlock-iremus/sherlock-tonalities>.

⁴⁷ https://github.com/polifonia-project/tonalities_pilot/blob/main/validation/interface/Séance%20de%20travail%20Thomas%2C%20Marco%2C%20Félix%2C%20Antoine%2C%20Christophe.pdf; https://docs.google.com/document/d/1lp_zwYTrlUEzQDghAfDae5h1Bbb14Mk9SCj4wj2rqD8/edit?invite=COcn1KQG.



options by limiting the choice to only the classes of a theoretical model, and by making a property implicitly linked to each class.⁴⁸

Tunes

Presentation of the validation protocol

The Tunes pilot aims to both implement the core validation criteria as defined in the Polifonia validation framework⁴⁹ and a set of pilot-specific validation criteria. We will discuss both validation protocols in this session.

1. Core validation criteria

a) XD and UX Design

The user experience design is driven by two personas: Mark, a music historian, and Brendan, a traditional musician and ethnomusicologist. These personas are complementary in their profile. Where the Mark persona is interested in investigating origin and dissemination of tunes, the Brendan persona is interested in finding suitable tunes for his performance practice. For both personas, a set of competency questions has been established which will be a reference in designing and evaluating user interaction.

b) Ecosystem and Rulebook

By adhering to the general guidelines for the Polifonia Ecosystem, the pilots secure the coherence of the ecosystem. Components that have been, and will be released by the TUNES pilot, conform to these rules.

c) Web portal

Tunes contributed two personas, two stories, and two sets of competency questions to the Web Portal. Two 'data stories' have been released within the MELODY framework. The first version of the Tunes knowledge graph has been deployed on the Polifonia server, and is available for components of the Web Portal.

d) Survey

The validation methods, and protocols for the Tunes pilot are recorded in the survey. The following table 9 presents the details for Tunes.

Survey's validation section	Tunes pilot
In order for the pilot to be successful, what criteria would the new tools and methods need to satisfy?	The tool should facilitate the discovery of connections between the melodic contents of different data sets, where 'connection' means the same (or related) melodies or melodic motifs in different data sets.
What will be the validation setup?	*Output: Validate the output by using the existing manual annotation of tune origins as ground truth.

⁴⁸ When a clause is assigned in a cadence, the property ("hasCantizans") that links the object (e.g. cantizans) to the subject ("cadence") does not have to be selected by the user. It is directly inferred by the device.

⁴⁹ https://github.com/polifonia-project/socio-technical_roadmap/tree/main/validation_framework.



	*Interface: User study with domain experts on how the tool can support scholarship.
What are the pilot's validation methods?	*Validation of interfaces by users/experts *Validation of results by ground truth
What KPIs or metrics will be used for validating it?	*User satisfaction. *Retrieval evaluation measures (including precision, recall, mean average precision, accuracy, F1 measure) of identified melodies in the ground truth.
What values for KPIs/metrics shall be achieved for claiming success?	To be determined
Can you name at least three (additional) stakeholders from different institutional backgrounds and countries who could contribute to the definition/validation of the pilot?	Rebekah Ahrendt (Music Historian, Utrecht University), Ellen van der Grijn (Collection Specialist, Meertens Institute), Frans Wiering (Computational Musicologist, Utrecht University), Ewa Dahlig-Turek (Ethnomusicologist, The Institute of Art of the Polish Academy of Sciences), Dániel Biró (Composer, Grieg Academy, Bergen, Norway)

Table 9. Tunes' answers to the Survey's validation section

2. Pilot-specific methods

a) User validation

To validate the "Mark" persona and the Mark#1 FolkMusic story, we organized a session with one of the project's stakeholders, Dr. Rebekah Ahrendt (Utrecht University). The Mark persona was partly based on a interview we conducted with her at the beginning of the project.

b) Input validation

For each of the data sets that will be incorporated in the Tunes pilot, we need to perform a series of conversions to prepare the data for processing. This involves conversion into Humdrum ****kern** encoding⁵⁰, rendition of the scores as images (.png), and extraction of basic features from the melodies: the pitch and the metric weight for each note.

The extent to which the conversions are successful is an important validation criterion for the input data.

c) Classification Performance

The data from the Meertens Tune Collections includes tune family labels. If melodies have been recognized as variants of the same melody by the collection specialists, they were assigned the same tune family label. The tune family label groups variants of the same tune. We use these

⁵⁰ <https://www.humdrum.org/Humdrum/representations/kern.html>.



labels as ground truth in a retrieval experiment. Using one melody as query, we are interested whether our methods are able to retrieve the other members of the variant group (tune family) from the data set. As matching method, we employ an existing alignment algorithm⁵¹, which gives an alignment score for every pair of two melodies. We computed a full distance matrix for all melodies from Ceol Rince na hÉireann (CRE), Meertens Tune Collections (MTC), Session, and Essen corpora. From these distances, we computed the mean average precision, and the recognition rate using melodies from MTC as queries. The mean average precision indicates to what extent the relevant items are at the top of the ranked lists. The recognition rate is the percentage of queries with a relevant item at the top rank of the result list. This performance measure is relevant for the aim of identifying a melody. It is not required to have all relevant items at the top of the ranked lists for the purpose of identifying a melody. It can be sufficient to have a relevant item among the top 5 or top 10 in the list. But here we take the strict measure of having a relevant item in the top rank of the list.

Next, to the quantitative evaluation, we perform a qualitative evaluation of the retrieval results. We provided the collection specialist of the Dutch Song Database with the ranked lists, and we asked here to examine those in light of tune family membership of the included tunes.

Application of the validation protocol and results obtained

a) User validation

During the meeting with the project's stakeholders, Dr. Rebekah Ahrendt (Utrecht University), we presented the Mark#1_FolkMusic story and corresponding competency questions. Dr. Ahrendt confirmed the completeness of the questions. Three recommendations were recorded based on the discussion:

- At any point, it should be clear to the user on what authority the provided information is based. In other words, we always have to include the source of information. Especially for academic users (musicological experts) information should not be presented as 'truth', but in such a way that the interpretation of it is left as much as possible to the user.
- It would be advisable to include more data sets, as there are many, and musicology currently lacks the centralization of databases.
- There are a number of musicologists interested in scribes and copyists. To automatically identify copyists based on their hand writing would be a fantastic future goal.

⁵¹ Peter Van Kranenburg. *A Computational Approach to Content-Based Retrieval of Folk Song Melodies*, Ph.D. thesis, Utrecht University, 2010.



b) Input validation

Five data sets have been imported into the project. The data are stored in the tunes-dataset repository in github⁵².

Ceol Rince na hÉireann (CRE) corpus (Irish) **c. 1,200 tunes**

The abc encoding of the corpus was taken from the folk_ngram_analysis repository, which is being maintained by the WP3 team⁵³. The abc source file was split into separate .abc files and these have been converted to **kern, midi, and MusicXML. Visual renditions of the scores have been generated (.png and .svg), and features have been extracted (stored in the json directory). All conversions are successful.

Meertens Tune Collections – MTC-FS-INST-2.0 (Dutch) **c. 18,000 tunes**

All tunes are available in **kern, and midi in the official release of the data set. Also, visual renditions of the scores are available as .png images. Features have been extracted and stored in .json format.

ESAC Folksong Databases (mostly German) **c. 8,300 tunes**

The official web page of the Essen collection⁵⁴ refers to the KernScores library for the .krn versions⁵⁵. That is where the data was obtained. It appeared, however, that there is an error in this data. The files deut1328 till deut2271 are corrupted in the KernScore set. We reconstructed those. The correct version is in the tunes-dataset github repository.

The tunes were converted to midi, and visual renditions have been generated (.png and .svg files). Features have been extracted and stored in the .json format.

The Session Corpus (Irish) **more than 40,000 tunes**

The abc encoding of the corpus was taken from the folk_ngram_analysis repository, which is being maintained by the WP3 team⁵⁶. The abc source file was split into separate .abc files and these have been converted to **kern, and MusicXML. Visual renditions of the scores have been generated (.png and .svg), and features have been extracted (stored in the json directory). During the process, many edits have been made to the original abc file. These were mainly to correct the inconsistencies which prevented parsing the file. Not all conversions were successful, but given the sheer size of the data set, we accepted the loss of the tunes that couldn't be converted or parsed.

Early American Secular Music and Its European Sources, 1589–1839: An Index **c. 66.000 Incipits, metadata**

The data was scraped from <https://www.cdss.org/elibrary/Easmes/Index.htm> and stored in the tunes-dataset repository. The data fields were extracted from the .html files, and stored in .json format in the directory output in the tunes-dataset repository.

⁵² <https://github.com/polifonia-project/tunes-dataset>.

⁵³ https://github.com/polifonia-project/folk_ngram_analysis.

⁵⁴ <http://www.esac-data.org>.

⁵⁵ <https://kern.humdrum.org/cgi-bin/browse?l=/essen>.

⁵⁶ https://github.com/polifonia-project/folk_ngram_analysis.



c) Classification Performance

From the distance matrix, as computed by the alignment algorithm, we construct for each tune a ranked list of all other tunes ordered by descending alignment score. The result lists are currently available at https://syrinx.knorrie.org/~pvk/polifonia/tunes-dataset/output/distmat/cre_mtc_essen_ses_json.ExactPitch40_IMA_0.6_0.2_min/.

Using the tune family labels of the MTC data set as ground truth, we computed the mean average precision and the recognition rate. The members of the same tune family as the query were taken as relevant items, and all melodies that have a tune family label have been used as queries.

We obtain the following results:

Mean Average Precision: 0.68

Recognition Rate: 0.88

Realizing that for each query we only have a very small number of relevant items, typically in the order of 10, and the size of the full set is in the order of 10,000, these results are impressive. This indicates that the alignment algorithm is able to retrieve tune family members quite reliably.

For the qualitative evaluation, we provided the collections specialist of the Meertens Institute with the retrieval results. Based on these, she was able to make a considerable number of changes and additions to the Dutch Song Database, in the order of tens. These are various types of changes and additions. Some examples:

1. Merging tune families.

Tune NLB127036_01 was assigned to the tune family “Nog staat daar een roos”. The top item in the result list⁵⁷ for this tune is another tune from the Dutch Song Database, NLB181070_01, which is assigned to tune family “Letzte Rose”. The tunes are clearly related. Therefore, the tune families “Nog staat daar een roos” and “Letzte Rose” were merged.

2. Removing tune family labels

In the result lists for all tunes from tune family “Ohne Lieb und Wein”, two tune family members (NLB147023_01, and NLB151281_01) are structurally at a very low rank. Upon inspection, it appeared that these tunes are not related to the tune family, so they were removed from the tune family.

3. Indicating relations

Based on the result lists, some tune families seem to be related, but not to the degree that they can be considered the same tune. These kinds of relationships have been indicated in the database. An intriguing example is the resemblance between tune families “In Holland staat een huis”⁵⁸ and “Al is ons prinsje nog zo klein 1”⁵⁹. The first in 4/4 meter, and the second in 6/8 meter. However, both groups of melodies have their own characteristics, such that they are not considered the ‘same’ tune, but related.

⁵⁷ https://syrinx.knorrie.org/~pvk/polifonia/tunes-dataset/output/distmat/cre_mtc_essen_ses_json.ExactPitch40_IMA_0.6_0.2_min/rankedlists/NLB127036_01.html.

⁵⁸ <https://www.liederenbank.nl/resultaatlijst.php?zoek=4321&actie=melodienorm&sorteer=jaar&lan=en>.

⁵⁹ <https://www.liederenbank.nl/resultaatlijst.php?zoek=288&actie=melodienorm&sorteer=jaar&lan=en>.



4. Additional information from other sources

In the result lists for tunes in the tune family “Cecilia”⁶⁰, a concordance for the melody was found in the collection of The Session entitled “Italian Rant”. Upon inspection and literature research, it appeared that this tune, which has numerous occurrences in Dutch sources, indeed has Italian roots. This was previously not known, and was added to the database.

Lessons learned and future perspectives

Although the retrieval performance is impressive, the quantitative evaluation of the retrieval results indicates that there is room for improvement. However, the qualitative evaluation showed that at least a part of the inaccuracies is caused by errors in the ground truth. We will do an effort to correct the tune family labels as much as possible based on the current results and rerun the experiment, which will result in improved values for the retrieval evaluation measures.

The qualitative evaluation of the results yielded many improvements and discoveries concerning the metadata of the tunes.

Currently, the results are presented as static web pages. An important next step will be to design and build interactive interfaces to explore the network of melodies and sources, based on different kinds of similarities.

MusicBo

Presentation of the validation protocol

Four specific methods in the validation framework pertain to MusicBo: 1) validation of input data quality by experts, 2) validation of interface by experts, 3) performance validation, and 4) validation of a sub-corpus/case study and transfer to a larger corpus.

1. Core validation criteria

a) XD and UX Design

The MusicBo corpus undergoes a process that automatically transforms the unstructured information in its documents into structured knowledge modelled in a Knowledge Graph (KG). Through this KG, many different types of situations have been modelled in the Polifonia Ontology Network⁶¹. For example, the situation of an entity (e.g. a musician) having a (nick)name that may be valid only in a specific time interval and within an organization; the situation of an artist creating an art entity (e.g. a musical composition) at a specific time and place and with a possible beneficiary; or the situation of an agent (e.g. a composer) writing something to someone else. The role modeling aspect within the music domain is complex, as is the linkage to other existing ontologies, and has been addressed particularly in deliverables related to modeling D2.1 and D2.2.⁶²

⁶⁰ <https://www.liederenbank.nl/resultaatlijst.php?zoek=1151&actie=melodienorm&sorteer=jaar&lan=en>.

⁶¹ <https://github.com/polifonia-project/ontology-network>.

⁶² Meroño-Peñuela, Albert *et al.* D2.1 *Ontology-based knowledge graphs for music objects. 1st version*, European Commission, Grant Agreement number: 101004746 — Polifonia — H2020-SC6-TRANSFORMATIONS-2018-2019-2020 / H2020-SC6- TRANSFORMATIONS-2020, Research Report, 2021; D2.2 *Ontology-based knowledge graphs for music objects. 2nd version* is due for M30 (30 June 2022).



b) Ecosystem and Rulebook

The conformity of the pilots to the Ecosystem and its Rulebook ensures that all components are correctly referenced and documented. MusicBo contributes with CQs and Stories as Carolina⁶³, the MusicBo module in Polifonia Corpus available on Github⁶⁴ and Zenodo⁶⁵, annotated for lemmas, tokens, and part of speech. Because the content of the corpus is copyrighted, only metadata such as the author, date, title, and source have been published.

c) Web portal

As a place for disseminating, enhancing, and interconnecting the pilots, the Web portal helps to harmonize the way Polifonia’s resources can be presented, explored, and searched by both amateur and expert users. MusicBo’s main contributions to the Web portal can be summarized as follows:

- a) participating in the data Stories to collect certain users’ needs.
- b) providing a first release of data via Melody⁶⁶ for communicating results and showing the progress of the pilot.

d) Survey

MusicBo contributes to the Survey⁶⁷ that helps organize data, bring to light pilot needs, and refine the validation structure.

Survey’s validation section	MusicBo pilot
In order for the pilot to be successful, what criteria would the new tools and methods need to satisfy?	New tools that are easy to use and accessible to the general public
What will be the validation setup?	Validation of some case study by submitting a questionnaire
What are the pilot’s validation methods?	<ul style="list-style-type: none"> *Validation of input data quality by experts *Validation of interfaces by users/experts * Performance validation *Validation of a sub-corpus/case study and transfer to a larger corpus
What KPIs or metrics will be used for validating it?	<ul style="list-style-type: none"> User satisfaction *Retrieval evaluation measures (BLEURT Score)

⁶³ <https://polifonia-project.github.io/ecosystem/pages/pilots/musicbo.html>.

⁶⁴ <https://github.com/polifonia-project/Polifonia-Corpus>.

⁶⁵ <https://zenodo.org/record/6672165#.Y7f32XbMJpk>.

⁶⁶ https://projects.dharc.unibo.it/melody/musicbo/music_in_bologna_knowledge_graph_overview.

⁶⁷ <https://github.com/polifonia-project/survey>.



What values for KPIs/metrics shall be achieved for claiming success?	To be determined
Can you name at least three (additional) stakeholders from different institutional backgrounds and countries who could contribute to the definition/validation of the pilot?	To be determined

Table 10. MusicBo answers to the Survey’s validation section

2. Pilot-specific methods

Specific methods	MusicBo
a) Input validation	/1) Validation of input data quality by experts
b) Validation of tools/means/software	/1) Validation of interface by experts /2) Performance validation
c) Output/goal validation	/1) Validation from a sub-corpus/case study and transfer to a larger corpus

Table 11. MusicBo’s specific methods

a) Input validation

Validation of input data quality by experts

The design of the corpus was completed in view of the two criteria: 1) representativeness of the Pilot’s research needs – the role that music played in the life of the city of Bologna from a historical and social perspective; and 2) balance of textual genres and variety of languages. The input of two expert musicologists who indicated the textual sources to be collected following the above criteria was a crucial step in identifying the texts to be included in the corpus. The representativeness is thus ensured by the alignment of the topics covered in the corpus texts with the needs of the Pilot; the balance includes coverage of the following textual genres: essays, historical texts, biographies, autobiographies, correspondences, media, catalogs of performances. Regarding language coverage, it was not possible to guarantee an equal portion for each language (Italian, English, French, and Spanish); due to the topic-specific nature of the corpus, the sources are mostly in Italian. Following the copyright license, they cannot be disclosed in their entirety, but metadata that allows for the reproduction of the corpus is released. In each document in the corpus, the metadata contains the author, title, publication date, textual genre information, and source link.



b) Validation of tools/means software

Validation of interface by experts

The WP4 implemented a linguistic query interface of the Polifonia Corpus⁶⁸, to which MusicBo contributed by 1) providing data and 2) operating initial queries. Some roundtable discussions and brainstorming meetings were carried out with language experts to gather requirements. The resulting mock-up (figure 2) allows for three types of interrogation: 1) the keyword search that can be used to select the sentences in the corpus that contain the keyword specified in the query; 2) Concept Search that can be used to select the sentences in the corpus that contain an occurrence of a polysemic word in one of its specific acceptations; and 3) Entity Search that can be used to select the sentences in the corpus that contain an occurrence of a word recognized as a Named Entity.

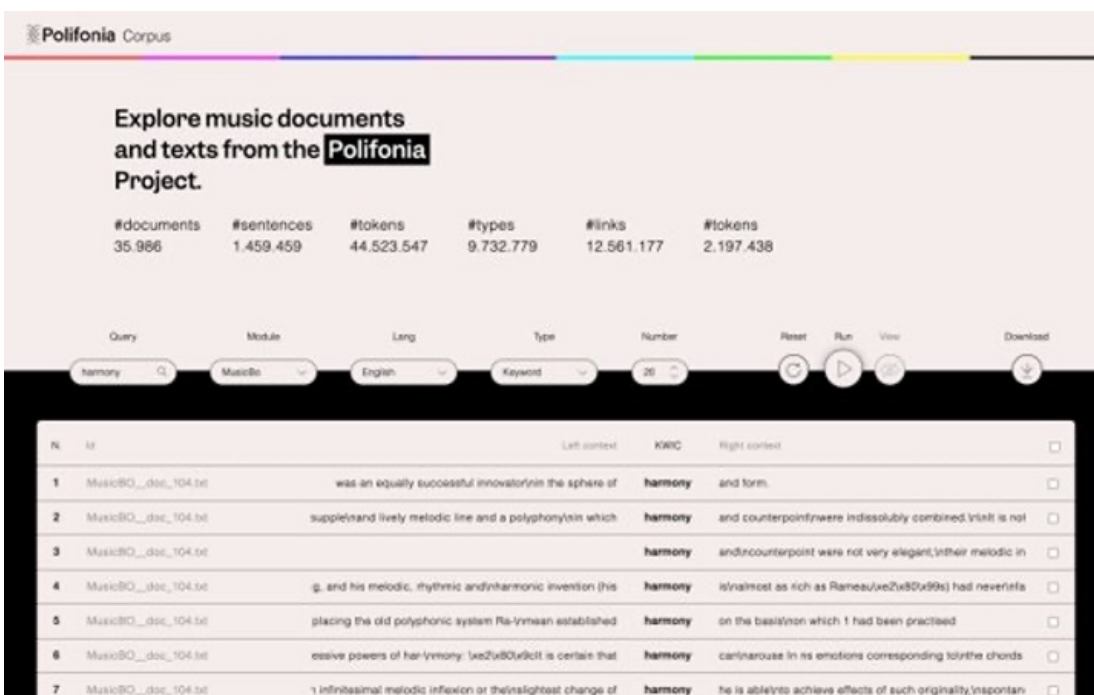


Figure 2. Interrogation Polifonia Corpus Interface

The language query interface will also undergo validation. In particular, a subdivided group will be composed following the variable of expertise (Senior and Junior experts). The interface will be evaluated according to the richness of the linguistic data and using the SUS criteria⁶⁹.

Performance validation

This paragraph describes the methods implemented to evaluate the quality of the KG generated from text. Our text-to-Knowledge Graph process leverages Abstract Meaning Representation (AMR) as an intermediate semantic representation. AMR graphs obtained through the semantic parsing of MusicBo textual corpus' sentences are then transformed into RDF/OWL knowledge graphs based

⁶⁸ <https://polifonia.disi.unibo.it/corpus/>.

⁶⁹ <https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html>.



on FRED's logic form⁷⁰. AMR2Fred tool⁷¹ realizes this transformation. AMR implements neo-Davidsonian semantics and relies heavily on PropBank framesets⁷². Those features make AMR particularly suited to capture “who is doing what to whom” in a sentence. Equally, FRED-like knowledge graphs are based on neo-Davidsonian semantics and are event-centric. Therefore, the knowledge graphs obtained through the AMR2Fred tool are appropriate to fulfill the MusicBo pilot's requirements. They inherently organize knowledge pivoting on framesets and frame arguments, making it easier to encode information about events and situations about the role of Bologna in European music history.

Given that the MusicBo corpus includes non-standard texts, as historical documents, the results of the state-of-the-art AMR neural parser may be inaccurate. Human validation is time-consuming, and there are no standard benchmarks for the semantic parsing of historic and OCRed text. For this reason, we decided to use a back-translation approach that converts the generated AMR graphs back to sentences (AMR2text)⁷³. With the back-translated sentences, we can compute Bilingual Evaluation Understudy with Representations from Transformers (BLEURT) score, a similarity score between the original sentence and the generated one. We hypothesize that generated sentences with a high BLEURT score correspond to high-quality graphs. As a first step, we decided to discard all the graphs of our AMR graphs bank corresponding to AMR2Text-generated sentences having a negative BLEURT score. In fact, according to our error analysis, negative BLEURT scores correspond to low-quality generated sentences, severely affected by issues such as OCR errors, and consequentially to low-quality graphs. We hypothesize that generated sentences with a high BLEURT score correspond to high-quality graphs. In future work, we would like to prove this hypothesis by performing human validation of the graphs via a custom WebApp⁷⁴ we have already developed. We started the human validation with a pilot tester.

⁷⁰ <http://wit.istc.cnr.it/stlab-tools/fred/>.

⁷¹ <https://github.com/infovillasimius/amr2Fred>.

⁷² <https://propbank.github.io/>.

⁷³ <http://wit.istc.cnr.it/stlab-tools/fred/>.

⁷⁴ <http://framester.istc.cnr.it/amr-eval>.

Sentence #21564 (BLEURT: 0.8210752010345459)

Walter was then scheduled to conduct the Berlin Philharmonic on March 20, but its management was warned by Joseph Goebbels that "unpleasant demonstrations" might occur at the concert, and the Propaganda Ministry clarified this by saying that there would be violence in the hall.

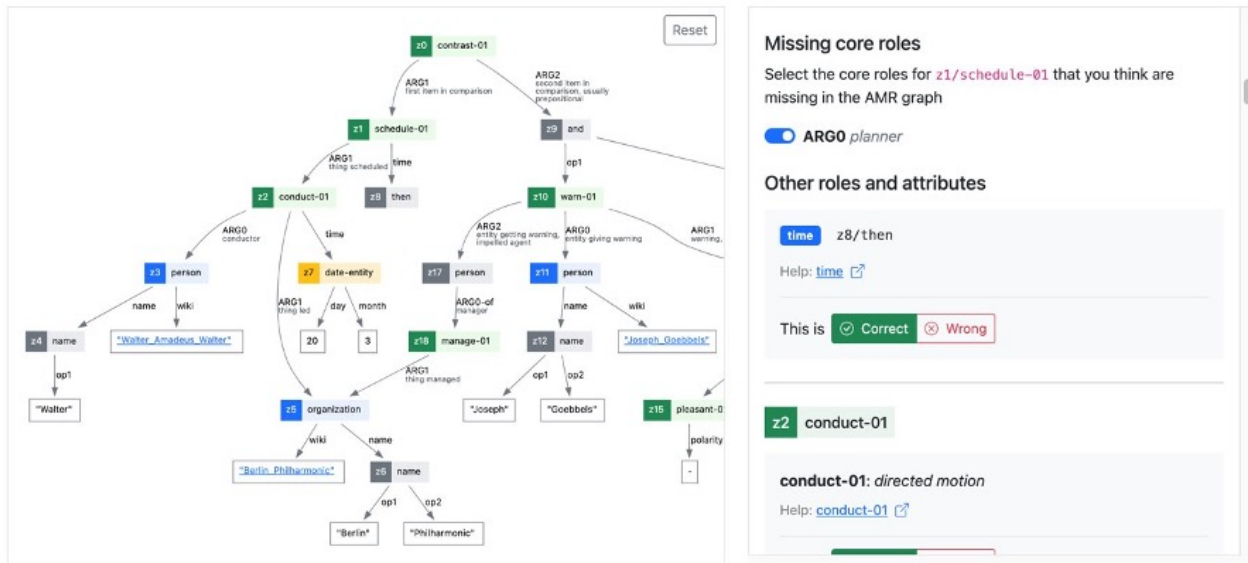


Figure 3. WebApp for human validation of the AMR graphs

Finally, we transformed the “best” graphs according to BLEURT score into OWL/RDF knowledge graphs. The graphs of our MusicBo’s AMR graphs bank corresponding to AMR2Text-generated sentences that had a positive BLEURT score and that were eventually transformed into FRED-like graphs were 5.965 out of 51.814 (11.5% of the total sent-AMR graph pair sample).

c) Output/goal validation

Validation from a sub-corpus/case study and transfer to a larger corpus

The multilingual nature of the corpus allows for a validation methodology that can be done on a representative sample and then extended to the whole corpus. The validation of the output of the pilots involves the submission of a questionnaire to assess the capability of the KGs generated from text to answer domain-specific questions.

This consists of the following steps:

1. Ask to a Validation Group to formulate some research questions related to the MusicBo Pilot. Research questions must have two characteristics: 1) be in relation to the data in possession (to avoid that data quality and not output performance is evaluated; 2) the proponent of the question must know the answer in order to provide the most appropriate answer in the questionnaire.
2. Transform the research questions into SPARQL queries;
3. Submit the SPARQL results to the Validation Group;
4. Submit a survey to the Evaluation Group that concern: accuracy and richness of results, correspondence between research questions and results, and level of expectation by the validators.

Considering that the Validation Group proposes the questions it evaluates, a division of the expertise of its constituent population is important. For this reason, a subdivided composition was envisioned between common user and expert, which in turn is divided into degree of experience,

junior and senior. This division allows for an accurate assessment that is reflected in better questioning. The validation Group is composed as follows:

	Musicians/music historians	Generic users
Senior Experts	30%	50%
Junior Experts	20%	

Table 12. Composition of the Validation Group

The survey consists of 4 questions (table 14) in line with the needs of the pilot. The capability measurement is reflected by the rating system formalized in the following table (table 13):



Rating				
0= Absolutely not	1= Basic information is missing, but some additional non-relevant information is present	2= Basic information is present, but a key part is missing	3= Key information is present, but some contextual/minor details are missing	4= Totally yes

Table 13. Rating system

Questions	
1)	The quality of the information extracted with respect to the question asked (Rating indicates if information is correct).
2)	The richness of information extracted with respect to the question asked (Rating indicates if information is complete)
3)	The novelty of information extracted with respect to the question asked (Rating indicates if new information is provided)
4)	Level of expectation by the validators

Table 14. Survey’s questions

Application of the validation protocol and results obtained

We are currently finalizing the composition of the Validation Group to which the survey will be submitted. In M27 we will begin the evaluation by submitting the Survey and applying it to the Pilot portion of the corpus, and we will be able to obtain initial results that will be useful in making changes for a second full iteration.

Lessons learned and future perspectives

In automatically generating KG from unstructured text, the main challenge we encountered was related to probing state-of-the-art semantic parsing models on documents containing novel characteristics compared to the models' original training sets, which are mainly made of contemporary texts from the web. The characteristics of the MusicBo corpus, such as being diachronic and plurilingual, pushed us to deal with poor-quality scanned documents that we needed to treat with OCR technologies, whose output is often noisy. We tackled this challenge by pre-processing the texts with rule-based post-OCR correction scripts. Also, state-of-the-art semantic parsing models perform better if given single sentences as inputs. Therefore, we



segmented MusicB corpus documents into sentences. As a consequence, we faced the problem of losing information at a sentence level, such as coreference information. We pre-processed the texts with rule-based coreference resolution tools to mitigate this issue. In future work, we will improve the coreference resolution and post-OCR correction algorithms developed for minimising the loss of information in the source text due to its breaking down into single sentences for semantic parsing. To evaluate the quality of the KG generated from text, we implemented a semi-automatic method based on the hypothesis that generated sentences with a high BLEURT score correspond to high-quality AMR graphs. In future steps, we plan to prove this hypothesis by expanding the human validation of the AMR graphs via the custom WebApp we have already developed, scaling the number of testers involved and AMR graphs analysed.

Child

This section of the report presents validation performed in the CHILD pilot. It consists of three main sections. In section 1, we discuss the CHILD pilot validation protocol. Section 2 offers the application of the validation protocol and results obtained, and section 3 provides the lessons learned. The report concludes by outlaying considerations for moving forward with further development.

Presentation of the validation protocol

This section provides the pilot's validation protocol to the overall validation framework defined in the roadmap. It details the core validation criteria and the selection of pilot-specific methods.

1. Core validation criteria

The compliance of the pilots to a set of standard core tools, principles and guidelines grants methodological and conceptual coherence to the whole while leaving enough space for individual pilot-specific validation protocols.

a) XD and UX Design

In performing the core validation for the CHILD pilot, the story and competency questions developed by Ortenz in WP1⁷⁵ were utilised. These story and competency questions are necessary for the CHILD pilot because they help assess the accuracy and effectiveness of the Listening Experience Database (LED)⁷⁶ content relevant to the pilot from the user's perspective. LED is a database that brings together a mass of data about people's experiences listening to music of all kinds, in any historical period and culture. The LED likely contains information about children's experiences of listening to music. This could include information about the types of music that children enjoy, how they listen to music (e.g., through a device or at a live concert), and the impact that music has on their overall experience and development. By bringing together this information in a single database, it may be possible to better understand the role that music plays in children's lives and how it contributes to their overall development and well-being. By regularly evaluating the content through competency questions and taking steps to address any issues that are identified, the team will be able to improve the user experience of the LED and ensure that it is a valuable resource for users.

⁷⁵ https://polifonia-project.github.io/ecosystem/polifonia-project/stories/Ortenz:%20Music%20Historian/Ortenz%231_MusicAndChildhood.html.

⁷⁶ <https://led.kmi.open.ac.uk>.



This aligns with the principles of XD and UX, which focus on designing products and services that are user-centred and create a positive experience for the user.

b) Ecosystem and Rulebook

CHILD's documentary evidence benchmark development in WP1 is a standard to evaluate the quality and reliability of the data sources used to create a knowledge graph relevant to the CHILD pilot. It is vital in LED because it helps ensure that the data is accurate, relevant, and up to date. Using the documentary evidence benchmark, we can ensure that the data sources used in LED for CHILD are reliable and trustworthy, which can help improve the overall quality and usefulness of LED. In addition, a documentary evidence benchmark helped to ensure that the knowledge graph for CHILD in LED is based on a solid foundation of evidence and is not simply a collection of unverified or biased information.

c) Web portal

The CHILD pilot is a crucial component in the dissemination and harmonization of other Polifonia resources through the use of MELODY. MELODY is utilized to create data stories that showcase trends or patterns discovered through the LED and to create interactive visualizations that aid users in comprehending and delving deeper into the data in the LED. By aligning MELODY with the LED for the CHILD pilot, we aim to create a more interactive and engaging experience for users and aid them in better understanding and utilizing the information within the LED. At present, the Melody's view of the CHILD pilot is still under construction. We plan to include this feature in the final report as soon as it is completed. Below is the link to the Melody presentation under construction.

https://melody-data.github.io/stories/published_stories/story_1667938609.943534.html



d) Survey

The contribution of CHILD to the survey is instrumental in organizing the data and ensuring that the focus remains on the pilot's needs, leading to a refinement of the validation design. Table 15 provides comprehensive information on the validation methods, protocols, and metrics utilized in the CHILD pilot, as recorded in the survey.

Survey's validation section	CHILD pilot
In order for the pilot to be successful, what criteria would the new tools and methods need to satisfy?	<ul style="list-style-type: none"> Automatic retrieval and characterization of documentary evidence quicker and more accurate.
What will be the validation setup?	<ul style="list-style-type: none"> Quantitative: ability of the system to produce accurate recommendations to curators User study on usability of thematic explorer User study on assisted formulation of research hypotheses
What are the pilot's validation methods?	<ul style="list-style-type: none"> Validation of input data quality by experts Validation of interfaces by users/experts Validation of objectives through experts
What KPIs or metrics will be used for validating it?	<ul style="list-style-type: none"> A] Component 1: Accuracy, F1, ... B] Growth of LED C] Number of scholarly activity supported / qualitative methods (Qualitative interviews with domain experts), Adoption of tools and methods in OU courses of Music Department
What values for KPIs/metrics shall be achieved for claiming success?	<ul style="list-style-type: none"> B] Growth of LED C] Number of scholarly activity supported / qualitative methods *Adoption of tools and methods in OU courses of Music Department
Can you name at least three (additional) stakeholders from different institutional backgrounds and countries who could contribute to the definition/validation of the pilot?	<ul style="list-style-type: none"> Europeana Rutgers DH Initiative RISM

Table 15. CHILD answers to the Survey's validation section

2. Pilot-specific methods

This section presents the selection of pilot-specific methods for CHILD pilot validation as required. It covers the three main methods: i) validation of input data quality by experts, ii) validation of interfaces by users, and iii) validation of objectives through experts.

i) Validation of input data quality by experts

This specific validation was done by matching the competency questions, CQ⁷⁷, against the Documentary Evidence Benchmark (DEB)⁷⁸. We used 20 randomly selected benchmarks to compare against the 12 competency questions related to the CHILD pilot. The CQ used in this validation was developed in WP1 by Ortenz, a Music historian with a background in art history and literature. It contains children's music experience, which aligns with the CHILD pilot's focus. The documentary evidence benchmark is the extraction of documentary evidence, taking the LED as a reference. The CHILD exemplar file in the Documentary Evidence Benchmark, a curated list of sources from LED that also includes childhood experiences, was used for this purpose.

By matching, we mean answering the competency questions with the content of each selected DEB. We focused on section 2 of the competency questions (CQ), which has 12 questions with a strong emphasis on music in Childhood. For example, in the CQ, we have a question: "*Where is the source/event produced?*" This question was answered by checking if there is a location (place) in the content of the DEB as indicated in figure 4, if there is a location mentioned such as Brighton (in the case of the figure 4 below), a score of "YES" is marked, and if no location/place is found, a score of "NO" is entered. We presented the results in table 16, with each column having a CQ as the header.

from **Journal, 1804**, page 16:



Shortly after the beginning of August I went on a visit to my sister at Brighton, taking Isabella and Edmond [her children] with me. There I remained nearly six weeks. During that time I was invited six or seven times to the Pavilion (the Prince of Wales). One of the nights there was a very pleasant ball; the other nights his band played. It is a prodigious fine band, as he has taken a great deal of pains with them, being passionately fond of music. He sung two nights himself, which he does very agreeably.

Figure 4. Extract of a DEB

To analyse, we assigned a score of 1 to YES and 0 to NO, then calculated the percentage of the sum on the total number of overall scores.

ii) Validation of interfaces by users

⁷⁷ https://polifonia-project.github.io/ecosystem/polifonia-project/stories/Ortenz:%20Music%20Historian/Ortenz%231_MusicAndChildhood.html.

⁷⁸ <https://polifonia-project.github.io/ecosystem/polifonia-project/documentary-evidence-benchmark/README.html>.



The validation of the interface was done using Nielsen ten usability heuristics for user interface design⁷⁹. These ten heuristics were matched against a user journey (curation process) relevant to CHILD on the LED by the users. The user validation was performed by four experts involved in LED's day-to-day curation activities. As data collection is essential in this validation, we collected the data using Nielsen's matrix (table 17), which was classified using rating according to Nielsen as explained below:

- 0 = I don't agree that this is a usability problem at all
- 1 = Cosmetic problem only: need not be fixed unless extra time is available on project
- 2 = Minor usability problem: fixing this should be given low priority
- 3 = Major usability problem: important to fix, so should be given high priority
- 4 = Usability catastrophe: imperative to fix this before product can be released

To analyse, points were assigned to each rating. That is 0 = 10 points, 1 = 9 points, 2 = 8 points, 3 = 7 points, and 4 = 6 points. The sum total of the points forms the overall usability point for the system.

iii) Validation of objectives through experts

The validation of objectives through experts was achieved by matching the LED against the competency questions⁸⁰ developed by Ortenz in WP1. We randomly selected ten experiences relevant to the CHILD pilot to match against the 12 competency questions. The validators were asked to check and score if the content of the experience answered the competency questions. For example, in the CQ, we have a question: when is the source produced? This question was answered by checking if there is a time (year) in the content of the selected experience. If there is a time mentioned in the content of the Childhood experience, a score of "YES" is marked, and if no time is found, a score of "NO" is entered. We presented the results in table 16, with each column having a CQ as the header.

To analyse, we assigned a score of 1 to YES, and 0 to NO, then calculated the percentage of the sum on the total number of overall scores.

Application of the validation protocol and results obtained

This section presents a fine-grained evaluation and analysis of the results obtained concerning the KPIs and success criteria defined for validation. It covers the results obtained from the three pilot-specific methods: validation of input data quality by experts, validation of interfaces by users and validation of objectives through experts.

Validation of input data quality by experts

The objective of this evaluation is to ensure the accuracy of the data being used. Table 16 provides the scores and validated datasets. After computation, the results show that the data quality is 70.8%, which is considered good but can still be improved.

⁷⁹ Jakob Nielsen. *Ten usability heuristics*, online PDF file published in *Semantic Scholar (S2CID)*, n. 59788005, 2005.

⁸⁰ https://polifonia-project.github.io/ecosystem/polifonia-project/stories/Ortenz:%20Music%20Historian/Ortenz%231_MusicAndChildhood.html.

Competency questions / Benchmark	01	02	03	04	05	06	07	08	09	10	11	12
memoriesofmusici00ganzuoft_djvu.txt	Y	N	N	N	Y	Y	Y	Y	Y	Y	Y	Y
musicstudyingerm00faya_djvu.txt	Y	N	N	N	Y	Y	Y	Y	Y	Y	Y	Y
musicandmanners005622mbp_djvu.txt	N	N	N	N	Y	Y	Y	Y	Y	Y	Y	Y
cu31924022435766_djvu.txt	N	N	N	N	Y	N	Y	Y	Y	Y	Y	Y
E14072_Memoirs_of_Baron_Bunsen_vol_I_djvu.txt	Y	N	N	N	Y	Y	Y	Y	Y	Y	Y	Y
E14073_Memoirs_of_Baron_Bunsen_vol_II_djvu.txt	Y	N	N	N	Y	Y	Y	Y	Y	Y	Y	Y
autobiographyand002432mbp_djvu.txt	Y	N	N	N	Y	N	Y	Y	Y	Y	Y	Y
musicalmemories017451mbp_djvu.txt	Y	N	N	N	Y	N	Y	Y	Y	Y	Y	Y
musicandfriends00gardgoog_djvu.txt	Y	N	N	N	Y	Y	Y	Y	Y	Y	Y	Y
thirtyyearsofmus00klei_djvu.txt	Y	N	N	N	Y	Y	Y	Y	Y	Y	Y	Y
reminiscencesofm02kellrich_djvu.txt	Y	N	N	N	Y	Y	Y	Y	Y	Y	Y	Y
lifeandlettersf02haregoog_djvu.txt	Y	N	N	N	Y	Y	Y	Y	Y	Y	Y	Y
lifelettersofsir00halluoft_djvu.txt	Y	N	N	N	Y	Y	Y	Y	Y	Y	Y	Y



LouisSpohrsAutobiography_djvu.txt	Y	N	N	N	Y	Y	Y	Y	Y	Y	Y	Y
lettersannasewa09consgoog_djvu.txt	Y	N	N	N	Y	N	Y	Y	Y	Y	Y	Y
recollectionsofo00ryanuoft_djvu.txt	N	N	N	N	Y	Y	Y	Y	Y	Y	Y	Y
selectedcorrespo002644mbp_djvu.txt	Y	N	N	N	Y	Y	Y	Y	Y	Y	Y	Y
musicallettersf00unkngoog_djvu.txt	N	N	N	N	Y	N	Y	Y	Y	Y	Y	Y
mymusicallife00haweuoft_djvu.txt	N	N	N	N	Y	Y	Y	Y	Y	Y	Y	Y
Praeterita_djvu.txt	Y	N	N	N	Y	Y	Y	Y	Y	Y	Y	Y

Table 16. Input data quality validation by experts. Yes(Y) = 1, No(N) = 0

Result Summary

	Responses	Percentage (%)
Yes	170	70.8
No	70	29.2

Validation of interfaces by users

This evaluation aims to ensure the users' Usability and ease of use of the system. Table 17 gives Nielsen usability scores obtained from the users. The system recorded a usability score of **86%**, which is considered good but can still be improved.

Usability Heuristics	Severity					Issues	Recommendation
	0	1	2	3	4		



<p>Visibility of the status</p> <p>The system should always keep users informed about what is going on</p>		++++				<p>No system feedback</p>	<p>The system should provide feedback or support the curator to validate or recommend missing entities</p>
<p>Match between system and the real world</p> <p>The system should speak the user's language</p>	++++					<p>The source field has always been very sticky; once a title has been typed in, it holds it, no option to delete</p>	<p>The system should provide users, especially curators to a deletion option</p>
<p>User control and freedom</p> <p>Supports undo and redo actions</p>		+	+++				
<p>Consistency and standards</p> <p>Users should not have to wonder whether different words, situations, or actions mean the same thing</p>		+	+++			<p>It presents choices which the user cannot be expected to know how to choose between</p>	<p>The system should be able to introduce correct information into the entries</p>
<p>Error prevention</p> <p>Prevents a problem from occurring in the first place</p>	+	+++					

<p>Recognition rather than recall</p> <p>Making objects, actions, and options visible</p>				+++ +			Listening Experiences need to be enhanced with an open-ended set of themes that users can use to explore the data (e.g. CHILDHOOD)
<p>Flexibility and efficiency of use</p> <p>Speed up the interaction for the expert users with accelerator</p>				+++ +			Users need support for filling the forms
<p>Aesthetic and minimalist design</p> <p>Dialogues should not contain information that is irrelevant or rarely needed</p>			+++ +		Previous and next buttons appear even when no search result is found.		The previous and next buttons should not appear when no results are found
<p>Help users recognise, diagnose, and recover from errors</p> <p>Error messages should be expressed in plain language</p>		++++					
<p>Help and documentation</p> <p>Use prompt and contextual help related the task, allow easy search</p>	++++						

Table 17. Nielsen validation framework for interfaces



Validation of objectives through experts

Table 18 presents the scores of the objective validations through experts. After computation, the results show that the objective through experts is 69.2%, which is considered above average but can still be improved.

LED / Benchmark	01	02	03	04	05	06	07	06	09	10	11	12
https://led.kmi.open.ac.uk/entity/lexp/1433763797230	Y	N	N	N	Y	Y	Y	Y	Y	Y	Y	N
https://led.kmi.open.ac.uk/entity/lexp/1409482741175	Y	N	N	N	Y	Y	Y	Y	Y	Y	Y	Y
https://led.kmi.open.ac.uk/entity/lexp/1396292425643	Y	N	N	N	Y	Y	Y	Y	Y	Y	Y	N
https://led.kmi.open.ac.uk/entity/lexp/1365518167	Y	N	N	N	Y	Y	Y	Y	Y	Y	Y	Y
https://led.kmi.open.ac.uk/entity/lexp/1425902082920	Y	N	N	N	Y	Y	Y	Y	Y	Y	Y	N
https://led.kmi.open.ac.uk/entity/lexp/1437644129613	Y	N	N	N	Y	Y	Y	Y	Y	Y	Y	Y
https://led.kmi.open.ac.uk/entity/lexp/1435839378719	Y	N	N	N	Y	Y	Y	Y	Y	Y	Y	N
https://led.kmi.open.ac.uk/entity/lexp/1436610229204	Y	N	N	N	Y	Y	Y	Y	Y	Y	Y	N
https://led.kmi.open.ac.uk/entity/lexp/1558351681445	Y	N	N	N	Y	Y	Y	Y	Y	Y	Y	N

https://led.kmi.open.ac.uk/entity/lexp/1426844423738	Y	N	N	N	Y	Y	Y	Y	Y	Y	Y	N
---	---	---	---	---	---	---	---	---	---	---	---	---

Table 18. Objectives validation through experts. Yes(Y) = 1, No(N) = 0, N= 120

Result Summary

	Responses	Percentage (%)
Yes	83	69.2
No	37	30.8

Lessons learned and future perspectives

The CHILD pilot was the focal point of the validation presented in this section of the document. The guidelines provided by the validation protocol were instrumental in achieving its intended goals. The results of the validation have provided valuable insights and have led to the redesign of the LED system with specific emphasis on the CHILD pilot, to make it more suitable for curators in the future. It was evident from the validation that the LED system requires further enhancements with respect to the CHILD component. We aim to develop a tool that will assist curators in validating or recommending missing entities in the LED system in view of the CHILD pilot's overall objectives. We also hope to support users in filling out the forms. In general, the LED needs to be enhanced with an open-ended set of themes that users can use to explore the data (e.g. CHILDHOOD). Finally, in terms of the progress that still needs to be made in preparation for the final validation, we plan to finish the construction of the Melody presentation, and make the Polifonia ecosystem link to the CHILD exemplar file easily accessible. In addition, we also aim to apply the validation protocol to the tool that will be developed in order to validate its performance. We will make the results available in the upcoming deliverables.

FAIR Protocol

The following sections summarize the way each individual pilot follows the FAIR protocol defined at the project level⁸¹.

Tonalities

- Documentation (Persona: Sethus, Story#1: Conflicting Theoretical Interpretations, Story#2: Create Relevant Corpus, Story#3: Conflicting Analytical Annotations, Survey)
- Reusable software (User Interface: Collaborative Annotation Interface for Music Analysis)
- Data (Modal-Tonal Ontology, Satisfaction questionnaire)

⁸¹ https://liveunibo.sharepoint.com/:w:/r/sites/polifonia/_layouts/15/Doc.aspx?sourcedoc=%7B10BC16F1-A847-429A-99BE-5E157DF21FC2%7D&file=D7.2%20Updated%20Data%20Management%20Plan%20Polifonia%20Version%202022_V1.docx&action=default&mobileredirect=true.



Documentation – Persona: Sethus

- Does the component adhere to the Polifonia 10 Rules for Open Science? Yes: R1, R3, R8.
- Link to the Polifonia Github organisation repo: <https://github.com/polifonia-project/stories/tree/main/Sethus:%20Music%20Theorist>.
- Is the component part of the Ecosystem? Yes: <https://polifonia-project.github.io/ecosystem/polifonia-project/stories/Sethus:%20Music%20Theorist/readme.html>.
 - o The component is compliant with the Rulebook of the Ecosystem.
- Is the component already part of Zenodo? No, it is an internal documentation.

Documentation – Story#1: Conflicting Theoretical Interpretations

- Does the component adhere to the Polifonia 10 Rules for Open Science? Yes: R1, R3, R8.
- Link to the Polifonia Github organisation repo: https://github.com/polifonia-project/stories/blob/main/Sethus:%20Music%20Theorist/Sethus%231_ConflictingTheoreticallInterpretations.md.
- Is the component part of the Ecosystem? Yes: https://polifonia-project.github.io/ecosystem/polifonia-project/stories/Sethus:%20Music%20Theorist/Sethus%231_ConflictingTheoreticalInterpretations.html.
 - o The component is compliant with the Rulebook of the Ecosystem.
- Is the component already part of Zenodo? No, it is an internal documentation.

Documentation – Story#2: Create Relevant Corpus

- Does the component adhere to the Polifonia 10 Rules for Open Science? Yes: R1, R3, R8.
- Link to the Polifonia Github organisation repo: https://github.com/polifonia-project/stories/blob/main/Sethus:%20Music%20Theorist/Sethus%232_CreateRelevantCorpus.md.
- Is the component part of the Ecosystem? Yes: https://polifonia-project.github.io/ecosystem/polifonia-project/stories/Sethus:%20Music%20Theorist/Sethus%232_CreateRelevantCorpus.html.
 - o The component is compliant with the Rulebook of the Ecosystem.
- Is the component already part of Zenodo? No, it is an internal documentation.

Documentation – Story#3: Conflicting Analytical Annotations

- Does the component adhere to the Polifonia 10 Rules for Open Science? Yes: R1, R3, R8.
- Link to the Polifonia Github organisation repo: https://github.com/polifonia-project/stories/blob/main/Sethus:%20Music%20Theorist/Sethus%233_ConflictingAnalyticalAnnotations.md.
- Is the component part of the Ecosystem? Not yet.



- Is the component already part of Zenodo? No, it is an internal documentation.

Documentation – Survey

- Does the component adhere to the Polifonia 10 Rules for Open Science? Yes: R1, R3, R8.
- Link to the Polifonia Github organisation repo <https://github.com/polifonia-project/survey>.
- Is the component part of the Ecosystem? Not yet.
- Is the component already part of Zenodo? No, it is an internal documentation.

Reusable software – User Interface: Collaborative Annotation Interface for Music Analysis

- Does the component adhere to the Polifonia 10 Rules for Open Science? No
- No link to the Polifonia Github organisation repo. It is an analytical tool, not reusable by other pilots of the Polifonia project.
- Is the component part of the Ecosystem? Not yet.
- Is the component already part of Zenodo? No, It is an Alpha version of a webservice, not yet ready to be published.
- The component does not oblige the [Polifonia Software Deliverable Guidelines](#)
 - o It is an Alpha version under development.

Data – Modal-Tonal ontology

- Does the component adhere to the Polifonia 10 Rules for Open Science? Yes: R1, R3, R8.
- Link to the Polifonia Github organisation repo: <https://github.com/polifonia-project/modal-tonal-ontology>.
- Is the component part of the Ecosystem? Not applicable.
- Is the component already part of Zenodo? No, it is a component for the Tonalities' user interface.

Data – Satisfaction questionnaire

- Does the component adhere to the Polifonia 10 Rules for Open Science? Yes: R1, R3, R8.
- Link to the Polifonia Github organisation repo: [https://github.com/polifonia-project/tonalities_pilot/blob/main/Tonalities%20Interface%20Satisfaction%20Questionnaire%20\(last%20version\).pdf](https://github.com/polifonia-project/tonalities_pilot/blob/main/Tonalities%20Interface%20Satisfaction%20Questionnaire%20(last%20version).pdf).
- Is the component part of the Ecosystem? Not yet.
- Is the component already part of Zenodo? – No, it is for internal use.

Tunes

- Documentation: Persona Mark, Story Mark#1_FolkMusic
- Documentation: Persona Brendan, Story Brendan#1_FindTraditionalMusic



Documentation – Persona: Mark

- Does the component adhere to the Polifonia 10 Rules for Open Science? Yes: R1.
- Link to the Polifonia Github organisation repo: <https://github.com/polifonia-project/stories/tree/main/Mark:%20Computational%20Musicologist>
- Is the component part of the Ecosystem? Yes: <https://polifonia-project.github.io/ecosystem/polifonia-project/stories/Mark:%20Computational%20Musicologist/readme.html>
 - o The component is compliant with the Rulebook of the Ecosystem.
- Is the component already part of Zenodo? No, it is an internal documentation.

Documentation – Story Mark#1_FolkMusic.md

- Does the component adhere to the Polifonia 10 Rules for Open Science? Yes: R1.
- Link to the Polifonia Github organisation repo: https://github.com/polifonia-project/stories/blob/main/Mark:%20Computational%20Musicologist/Mark%231_FolkMusic.md
- Is the component part of the Ecosystem? Yes. https://polifonia-project.github.io/ecosystem/polifonia-project/stories/Mark:%20Computational%20Musicologist/Mark%231_FolkMusic.html
 - o The component is compliant with the Rulebook of the Ecosystem.
- Is the component already part of Zenodo? No, it is an internal documentation.

Documentation – Persona: Brendan

- Does the component adhere to the Polifonia 10 Rules for Open Science? Yes: R1.
- Link to the Polifonia Github organisation repo: <https://github.com/polifonia-project/stories/tree/main/Brendan:%20Traditional%20Musician>
- Is the component part of the Ecosystem? Not yet.
 - o The component is compliant with the Rulebook of the Ecosystem.
- Is the component already part of Zenodo? No, it is an internal documentation.

Documentation – Story Brendan#1_FindTraditionalMusic

- Does the component adhere to the Polifonia 10 Rules for Open Science? Yes: R1.
- Link to the Polifonia Github organisation repo: https://github.com/polifonia-project/stories/blob/main/Brendan:%20Traditional%20Musician/Brendan%231_FindTraditionalMusic.md
- Is the component part of the Ecosystem? Not yet
 - o The component is compliant with the Rulebook of the Ecosystem.



- Is the component already part of Zenodo? No, it is an internal documentation.

MusicBo

- Documentation (Persona: Carolina, Story#1: SourcesCrossAnalysis, Persona: Valeriana Story#1: Discourse Analysis, Story#2: Terminology, Survey)
- Data (MusicBo Corpus, SPARQL Endpoint)
- Melody DataStory

Documentation

Documentation – Persona: Carolina

- Does the component adhere to the Polifonia 10 Rules for Open Science? Yes: R1, R3, R8.
- Link to the Polifonia Github organisation repo: <https://github.com/polifonia-project/stories/tree/main/Carolina:%20Music%20Historian>
- Is the component part of the Ecosystem? Yes: <https://polifonia-project.github.io/ecosystem/polifonia-project/stories/Carolina:%20Music%20Historian/readme.html>
 - o The component is compliant with the Rulebook of the Ecosystem.
- Is the component already part of Zenodo? No, it is an internal documentation.

Documentation – Story#1: SourcesCrossAnalysis

- Does the component adhere to the Polifonia 10 Rules for Open Science? Yes: R1, R3, R8.
- Link to the Polifonia Github organisation repo: https://github.com/polifonia-project/stories/blob/main/Carolina:%20Music%20Historian/Carolina%231_SourcesCrossAnalysis.md
- Is the component part of the Ecosystem? Yes: https://polifonia-project.github.io/ecosystem/polifonia-project/stories/Carolina:%20Music%20Historian/Carolina%231_SourcesCrossAnalysis.html
 - o The component is compliant with the Rulebook of the Ecosystem.
- Is the component already part of Zenodo? No, it is an internal documentation.

Documentation – Persona: Valeriana

- Link to the Polifonia Github organisation repo: <https://github.com/polifonia-project/stories/tree/main/Valeriana%20:%20Linguist>
- Is the component part of the Ecosystem? Not Yet
- Is the component already part of Zenodo? No, it is an internal documentation.

Documentation - Story#1 Discourse Analysis



- Does the component adhere to the Polifonia 10 Rules for Open Science? Yes: R1, R3, R8.
- Link to the Polifonia Github organisation repo: https://github.com/polifonia-project/stories/blob/main/Valeriana%20:%20Linguist/Valeriana%231_DiscourseAnalysis.md
- Is the component part of the Ecosystem? Not Yet

Documentation - Story#2 Terminology

- Does the component adhere to the Polifonia 10 Rules for Open Science? Yes: R1, R3, R8.
- Link to the Polifonia Github organisation repo: https://github.com/polifonia-project/stories/blob/main/Valeriana%20:%20Linguist/Valeriana%232_Terminology.md
- Is the component part of the Ecosystem? Not Yet

Documentation – Survey

- Does the component adhere to the Polifonia 10 Rules for Open Science? Yes: R1, R3, R8.
- Link to the Polifonia Github organisation repo <https://github.com/polifonia-project/survey>.
- Is the component part of the Ecosystem? Not yet.
- Is the component already part of Zenodo? No, it is an internal documentation.

Data

Data - Pilot Corpus

- Does the component adhere to the Polifonia 10 Rules for Open Science? Yes: R2, R6,
- Is the component already part of Zenodo? Yes
- Provide DOI of MusicBo Corpus: <https://doi.org/10.5281/zenodo.6672165>
- Link to the Polifonia Corpus on Github : <https://github.com/polifonia-project/Polifonia-Corpus>

Data – SPARQL Endpoint

- Link to the MusicBo Knowledge Graph GitHub repository: <https://github.com/polifonia-project/musicbo-knowledge-graph>
- MusicBo Knowledge Graph is available in the MusicBo SPARQL Endpoint: <https://polifonia.disi.unibo.it/musicbo/sparql>

Melody DataStory

- Link to the MusicBo DataStory on Melody: https://projects.dharc.unibo.it/melody/musicbo/music_in_bologna_knowledge_graph_overview

Child

- Documentation (Persona: Ortenz, Story: MusicAndChild, Survey)



- Executables (CLI-tool) - SPARQL Anything
- Reusables – Library - SPARQL Anything
- Data (Child pilot Corpus, SPARQL Endpoint)
- CHILD DataStory

Documentation – Persona: Ortez

- Does the component adhere to the Polifonia 10 Rules for Open Science? Yes: R1, R3, R8.
- Link to the Polifonia Github organisation repo: https://github.com/polifonia-project/stories/blob/main/Ortez:%20Music%20Historian/Ortez%231_MusicAndChildhood.md
- Is the component part of the Ecosystem? Yes: https://polifonia-project.github.io/ecosystem/polifonia-project/stories/Ortez:%20Music%20Historian/Ortez%231_MusicAndChildhood.html
 - o The component is compliant with the Rulebook of the Ecosystem.
- Is the component already part of Zenodo? No, it is an internal documentation.

Documentation – Story: MusicAndChildhood

- Does the component adhere to the Polifonia 10 Rules for Open Science? Yes: R1, R3, R8.
- Link to the Polifonia Github organisation repo:

https://github.com/polifonia-project/stories/blob/main/Ortez:%20Music%20Historian/Ortez%231_MusicAndChildhood.md

- Is the component part of the Ecosystem? Yes:

https://polifonia-project.github.io/ecosystem/polifonia-project/stories/Ortez:%20Music%20Historian/Ortez%231_MusicAndChildhood.html

Documentation – Survey

- Does the component adhere to the Polifonia 10 Rules for Open Science? Yes: R1, R3, R8.
- Link to the Polifonia Github organisation repo <https://github.com/polifonia-project/survey>.
- Is the component part of the Ecosystem? Not yet.
- Is the component already part of Zenodo? No, it is an internal documentation.
 - o The component is compliant with the Rulebook of the Ecosystem.
- Is the component already part of Zenodo? No, it is an internal documentation.

Documentation – Executables (CLI tool)

- Link to the Polifonia Github organisation repo: <https://polifonia-project.github.io/ecosystem/polifonia-project/external-components/components/sparql-anything.html>
- Is the component part of the Ecosystem? Yes



<https://polifonia-project.github.io/ecosystem/polifonia-project/external-components/components/sparql-anything.html>

Documentation – Library - SPARQL Anything

- Does the component adhere to the Polifonia 10 Rules for Open Science? Yes: R1, R3, R8.
- Link to the Polifonia Github organisation repo: <https://github.com/polifonia-project/external-components/blob/main/components/sparql-anything.md>
- Is the component part of the Ecosystem? Yes

<https://polifonia-project.github.io/ecosystem/polifonia-project/external-components/components/sparql-anything.html>

Data

Data - Pilot Corpus

- Does the component adhere to the Polifonia 10 Rules for Open Science? Yes: R2, R6,
- Is the component already part of Zenodo? Yes
- Link to the Polifonia Corpus on Github : <https://github.com/polifonia-project/documentary-evidence-benchmark>
- Link to the CHILD GitHub repository:

<https://polifonia-project.github.io/ecosystem/pages/pilots/child.html>

Data – SPARQL Endpoint

- CHILD Knowledge Graph is available in the LED SPARQL Endpoint:

<https://data.open.ac.uk/index.php?path=sparql>

Melody DataStory

- Link to the CHILD DataStory on Melody:

https://melody-data.github.io/stories/published_stories/story_1667938609.943534.html



Conclusion

This deliverable has presented and partially applied the validation protocols for the pilots Tonalities, Tunes, MusicBo, and Child. These protocols are all part of Polifonia's general Validation framework, combining core validation criteria and pilot specific methods. The framework provides homogeneity to the whole validation process within the project. Namely, ensures that the pilots adopt the same methodological guidelines (XD and UX design), rules for referencing the components (Ecosystem), dissemination channel (Web portal), and tool for monitoring and centralizing information about the pilots (Survey). Within this framework, each pilot has defined individual strategies for validating their personal a) input data; b) tools, means, and software components; and c) outputs and results.

a) Tonalities, MusicBo, and Child apply qualitative approaches to assess the standard of their input data (musical scores, historical texts) and identify what changes have been made to suit the pilot's requirements: Tonalities put in place protocols for data curation and enrichment, MusicBo defined its corpus relying on experts, and Child used experts to check the quality of randomly selected data samples. Concerning Tunes, the validation focuses on the conversion of external data sources in various formats into Humdrum ****kern encoding**.

b) The evaluation of the tools/media/software components focuses on the validation of interfaces by amateur and/or expert users. This is to ensure the pilots' adherence to the XD-UX-design methodology. Tonalities and Child have put in place usability surveys and/or interviews to validate their interfaces, while MusicBo has gathered requirements based on roundtable discussions with language experts. In addition, the pilots validate their performances by tests on data samples (Tonalities), automatic data completeness checks (Tonalities), and back-translation methods (MusicBo).

c) As regards the validation of outputs and goals, Tonalities and Tunes make use of ground truth to validate the results obtained (i.e. the modal-tonal classifications of musical works and the grouping of tunes according to melodic patterns). The strategy adopted by MusicBo is to validate a larger set of results from narrow case studies. Additionally, Tonalities and Child have their objectives validated through experts.

In all subgroups, validation by users is a central criterion. This approach is in line with the project's overall commitment to obtaining early feedback from musical heritage communities. It aims at avoiding technology-driven development and to foster user-centered pilot developments. By doing so, the project intends to create a continuous dialogue between the users' feedback and the subsequent adaptation/refocusing of the pilots.

At this stage, the four pilots have identified and defined all components of their validation protocols. These components may however be further adapted and updated according to the validation results obtained in the months to come. The pilots are at different stages in the actual application of these protocols. MusicBo will apply its validation protocol starting in M27 (March 2023). Tonalities has completed a first iteration of some of its validation components (quantitative validation of input data, validation of input data quality by experts, validation of interfaces by experts, validation of performances, and requests/information passed between front and back end), with the validation of results and objectives to be addressed in the coming months. Tunes plans a new iteration of ground truth validation once the datasets from a previous iteration have



been refined. Child has carried out a first iteration in the validation of input data, interface and results. Further iterations are planned based on the lessons learned from the initial validation data.

Beyond the validation work that remains for the individual pilots, more global efforts situated at the level of WP1 and/or WP5 will be pursued in the coming months:

- The optimization of the Ecosystem model to provide differentiated referencing of all components (including validation components such as surveys, questionnaires, user scenarios, etc.);
- The refinement of validation methods not only to test the final product but also to inform the actual and still ongoing pilot design process;
- The targeted and more systematic involvement of user communities in line with the pilot's issues, namely communities relating to the study of musical heritage based on textual (MusicBo and Child) and musical (Tonalties, Tunes) sources.

Given Polifonia's methodological and technical foundation and the work already completed with respect to the identification and application of the validation protocols, it is reasonable to expect a successful final validation carried out at the end of the project.



References

Blomqvist, Eva *et al.* “Experimenting with eXtreme Design”, in *Proceedings of EKAW 2010*, ed. by P. Cimiano, H.S. Pinto, Springer, 2010, pp. 120-134.

Blomqvist, Eva *et al.* “Engineering Ontologies with Patterns - The eXtreme Design Methodology”, in *Ontology Engineering with Ontology Design Patterns*, vol. 25. *Studies on the Semantic Web*, ed. by P. Hitzler, A. Gangemi, K. Janowicz, A. Krisnadhi, V. Presutti, IOS Press, 2016, pp. 23-50.

Ceulemans, Anne-Emmanuelle and Guillotel-Nothmann, Christophe. “Praetorius’ *Polyhymnia caduceatrix* (1619): Diatonische Logik und Skalenlogik in mehrhörigen und konzertierenden Werken des Frühbarocks,” *Zeitschrift der Gesellschaft für Musiktheorie, Gesellschaft für Musiktheorie*, 2020, *Alte Musik und Musiktheorie zwischen Historiografie und Pädagogik*, 17/2 (2020): 51–93.

Guillotel-Nothmann, Christophe. “Knowledge extraction and modelling in the project *Thesaurus Musicarum Germanicarum*”, paper presented at the *Prague DH Workshops Session III: Editions*, Masaryk Institute and Archives of the Czech Academy of Sciences, 2020.

Guillotel-Nothmann, Christophe and Ceulemans, Anne-Emmanuelle. “Das diatonisch-chromatische System zur Zeit des Michael Praetorius. Eine digitale Neuerschließung des *Syntagma musicum* (1619) in Verbindung mit dem Tanzzyklus *Terpsichore* (1612),” in *Musik im Umbruch. Michael Praetorius zum 400. Todestag*, Harrassowitz, 2021, pp.109-150.

Guillotel-Nothmann, Christophe *et al.* *D1.2 Roadmap and pilot requirements 2nd version*, European Commission, Grant Agreement number: 101004746 — Polifonia — H2020-SC6-TRANSFORMATIONS-2018-2019-2020 / H2020-SC6- TRANSFORMATIONS-2020, Research Report, 2022.

<https://iswc2022.semanticweb.org/index.php/workshops/>.

Meroño-Peñuela, Albert *et al.* *D2.1 Ontology-based knowledge graphs for music objects*, European Commission, Grant Agreement number: 101004746 — Polifonia — H2020-SC6-TRANSFORMATIONS-2018-2019-2020 / H2020-SC6- TRANSFORMATIONS-2020, Research Report, 2021.

Molinié, Georges, & Viala, Alain. *Approches de la réception : sémiostylistique et sociopolitique de Le Clézio*, Presses universitaires de France, 1993, pp. 43-46.

Nielsen, Jakob. *Ten usability heuristics*, online PDF file published in *Semantic Scholar (S2CID)*, n. 59788005, 2005.

Powers, Harold. “Tonal Types and Modal Categories in Renaissance Polyphony”, *Journal of the American Musicological Society*, 34/3 (1981): 428-470.

Scharnhorst, Andrea *et al.* *D7.2 Data Management Plan (Second Version)*, European Commission, Grant Agreement number: 101004746 — Polifonia — H2020-SC6-TRANSFORMATIONS-2018-2019-2020 / H2020-SC6- TRANSFORMATIONS-2020, Research Report, 2022.

Van Kranenburg, Peter. *A Computational Approach to Content-Based Retrieval of Folk Song Melodies*, Ph.D. thesis, Utrecht University, 2010.



Wiering, Franz. *The Language of the Modes: Studies in the History of Polyphonic Modality, Criticism and Analysis of Early Music*, Routledge, 2001.

Zarlino, Gioseffo. *Le istituzioni harmoniche*, Venice, the author, 1558.

Appendices

Term	Definition/description
<u>ArCo</u>	Architettura della Conoscenza. A project for developing the Italian cultural heritage knowledge graph.
AMR	Abstract Meaning Representation.
AMR2Fred	It translates from AMR to RDF according to the syntax of FRED.
AMR2text	AMR captures “who is doing what to whom” in a sentence.
BLEURT	Evaluation metric for Natural Language Generation.
CIDOC CRM	CIDOC Conceptual Reference Model (CRM) is a theoretical and practical tool for information integration in the field of cultural heritage.
CQ	Competency Question: a question that an ontology would need to answer.
CRE	Ceol Rince na hÉireann.
<u>CRMinf</u>	CRM _{inf} : the Argumentation Model. An Extension of CIDOC-CRM to support argumentation.
D	Deliverable.
DEB	Documentary Evidence Benchmark.
Discord	Instant messaging and digital distribution platform designed for creating communities.
DMP	Data Management Plan.
Domain specific part/aspect	Any scientific, epistemological and methodological considerations concerning one or more pilots.
ESAC	Folksong Databases.
eXtreme Design (XD)	Pattern-based ontology engineering methodology, which collects requirements in the form of Competency Questions.

FRED	FRED is a machine reading for the semantic web.
GioQoso	Gestion de la Qualité des partitions musicales ouvertes.
<u>Github</u>	Git is an open-source software for managing versions of documents. The documents typically managed in Git repositories are computer source code files, but it is in fact perfectly suitable for text data in general, including humanities data. The development of Git is closely related to that of GitHub a service for hosting Git repositories.
Goal	A short textual description of the goal(s) that the Persona needs to be addressed in the Story.
H	Human(s).
HCI	Human Computer Interaction.
Hermit	Reasoner for ontologies written using the OWL.
<u>IReMus</u>	Institut de Recherche en Musicologie, UMR 8223. Joint research unit composed of the CNRS, the BnF, Sorbonne University and the French Ministry of Higher Education and Research.
**kern	Core pitch/duration representation for common practice music notation.
KG	Knowledge graph.
Knowledge unit/type	New knowledge produced by the pilot through the analysis (by algorithms, machine learning and/or human annotation) of the input data. This unit may correspond for example to a cadence in a score, to the identification of a previously anonymous author of a source, to the type of a bell, etc. (see diagram 2. Survey – Knowledge units investigated by the pilots).
LED	Listening Experience Database.
M[1-40]	Month counting from the start of <i>Polifonia</i> on 1 January 2021: M6 (= June 2021), M18 (= June 2022), etc.
<u>MEI</u>	Music Encoding Initiative.
MH	Musical heritage.



MTC-FS-INST-2.0	Meertens Tune Collections.
NLP	Natural language processing.
Observation unit	A portion of input data on which the pilot relies to provide new units of knowledge through algorithms, machine learning and/or human annotation. These observation units may correspond for example to a sentence in a text, a score fragment (for example a cadence) or the dedicatee engraved on a bell.
OCR	Optical Character Recognition.
ODP	Ontology design patterns which correspond to small ontologies that work as reusable solutions to recurrent modelling problems.
OEI	Office of ethics and integrity.
OU	The Open University.
OWL	Web Ontology Language (OWL) is a Semantic Web language designed to represent rich and complex knowledge about things, groups of things, and relations between things.
Persona	A research-based description of a typical user. This description contains attributes such as name, age, occupation, and relevant characteristics such as knowledge and skills.
PON	Polifonia Ontology Network.
POPD	Protection of personal data.
PropBank	The Proposition Bank.
RDF	Standard model for data interchange on the Web.
Scenario	A description of how in a Story the Persona's task/need/problem is solved before, during and after interaction with the resource/software/service being developed.
SFAM	Société Française d'Analyse Musicale.

SMS tools	The three tools at the core of this deliverable: Story, <i>Maninpasta</i> and Survey.
Socio-pedagogical aspect	An item linked to a holistic and relationship-centred way of working in care and educational settings with the possible final users of <i>Polifonia</i> .
Source	Within a dataset, it is meant to define the origin of the input data, in opposition to the output data, that are the results collected by the pilot.
Story	A Story is a template for collecting requirements.
TB	The Technical Board, within the WP1, is in charge of the technical development, making sure that synergies are identified in a timely manner and issues around compatibility and technical interoperability are addressed harmonically during the development phase. The TB defines the common methodological approach and identifies the supporting collaborative tools to be used.
TSQVP	Tonalities' Score Quality Validation Protocol.
TSQVPEM	Tonalities' Score Quality Validation Protocol for Early Music.
UI	User interface.
<u>UNIBO</u>	Alma Mater Studiorum – Università di Bologna.
URI	A Uniform Resource Identifier is a unique sequence of characters that identifies a logical or physical resource used by web technologies.
UX	User Experience Design.
WG	Working group.
WP	Work package.

Appendix 1: Glossary

**Appendix 2: List of the pilots**

Polifonia's Pilot	Action
<u>CHILD</u>	This pilot will build a knowledge graph of the historical experience of music in childhood, using life writing (letters, diaries, memoirs, travel writing) and other historical texts as sources for adult reflections on music heard in childhood, third-party observations on children's engagement with music, and children's own first-hand accounts. The resulting knowledge graphs will inform an interface enabling the exploration and analysis through the dimensions of themes, time, and space.
<u>MUSICBO</u>	Music has always played a central role in the city of Bologna. Nevertheless, its musical heritage is only partly known and enjoyed compared to its full potential. This pilot will create multilingual (English, French, Spanish and German) digital corpora, containing the testimonies of scholars, journalists, travellers, writers and students from medieval to modern times through published documents showing diverse discourse styles such as stories, letters, reports, news, reportage, etc. These corpora will be the base to build a knowledge graph available to researchers, cultural institutes and public administrations for reuse.
<u>TUNES</u>	The digital music collection of the Meertens Instituut (Amsterdam) includes thousands of melodies from Dutch popular culture, spanning a period of more than five centuries. To trace possible international origins of Dutch early popular music culture, this pilot will interlink the entire melody collection of the Meertens Institute with a large number of other European collections. The linked melodic data sets will be highly valuable for musicologists and music historians interested in cultural evolution of musical style and in oral transmission and variation.
<u>TONALITIES</u>	The modal-tonal organisation of European music is decisive for its inner coherence, dramatic plot and, ultimately, artistic meaning. This pilot develops tools for the modal-tonal identification, exploration, and classification of monophonic and polyphonic notated music from the Renaissance to the twentieth century.

Appendix 3: WP1 participants, efforts, objectives, tasks and deliverables

WP number	1			Lead beneficiary	IREMUS					
Title	Web Portal and pilots									
Participant	1	2	3	4	5	6	7	8	9	10
	BO	OU	KCL	NUIG	ICCD	ICBSA	IREMUS	CNAM	NISV	KNAW
Person months	8	10	10	6	6	5	12	11	4	11
Start month	1			End month			40			

Objectives

WP1 is *Polifonia's* validation WP, it contributes to achieve O3 (Tailor), and O5 (Share & Engage). As such it drives the whole development. It also delivers a registry of all resources and materials retrieved, used and produced in the project, in the form of a Web portal. Its goal is twofold: 1) to demonstrate that the methods and tools developed in the technology provider work packages (WP2-5) are effective in facilitating management of large musical heritage collections and supporting enhanced understanding, preservation of, and interaction with, musical heritage, 2) to contribute to push the state of the art in relevant, though specific, musical heritage use cases. It includes seven tasks: T1 delivers a socio-technical roadmap to all partners ensuring supervision for the co-creation process; T2 provides a Technical Board supervising the development; T3 delivers a Web portal / registry and associated services; T4-7 develop ten pilots. Each task represents a theme relevant to the scope and challenges of the call: Preserving MH (T3), Managing MH (T4), Studying and understanding MH (T5), Interacting with MH (T6). Notice that ten “external” early adopters have already expressed their interest in participating in seven pilots, as explicitly indicated in the pilot descriptions (see also attached support letters).

Tasks

Task 1: Socio-Technical Roadmap (*Leader: IREMUS - Participants: ALL*)

This task provides a common framework to all pilots (T4-7) for: 1) coordinating collection of resources for use in the pilots, 2) gathering requirements from internal and external adopters/stakeholders, 3) setting objectives and challenges and mapping them from specific domains to the technology provider work packages (WP2-5), 4) monitoring validation within the pilot tasks. The goal is to maximise synergy and methodological soundness.

Task 2: Technical Coordination (*Leader: OU - Participants: ALL*)

This task will provide the pilots and the technology provider WPs with a reference Technical Board (TB) which, according to the roadmap defined in T1, will supervise the technical development making sure that synergies are identified in a timely manner and issues around compatibility and technical interoperability are addressed harmonically during the development phase. The TB will define the common methodological approach and identify the supporting collaborative tools to be used. It will



interact with WP6 to make sure that all the developed components are released according to the FAIR protocols defined in T7 of WP6. This task will also collect and report validation results from the pilots.

Task 3: Polifonia Web portal: an aggregator of digital musical heritage collections (*Leader: UNIBO - Participants: OU, NUIG, KCL*)

By extending musoW, a catalogue of Musical Data on the Web [<http://musow.kmi.open.ac.uk/>] (Daquino et al. 2018), this task will develop the *Polifonia* Web portal. It will include: 1) a reference registry of MH resources, including but not limited to all collections used and produced in the project, 2) methods for continuous indexing of MH resources, 3) methods for automatic and semiautomatic generation of metadata according to the *Polifonia* ontologies, 4) methods for searching, querying, browsing MH resources. The task will also provide input to WP6 about recommendations for the sustainability of the Web portal, beyond the lifespan of the project.

Task 4: Preserving musical heritage through knowledge graphs (*Leader: KNAW - Participants: UNIBO, OU, KCL, NUIG, ICCD, ICBSA, Nationaal Instituut voor de Orgelkunst (NiVO) (external), Soprintendenza Archeologica Bologna (external), Soprintendenza Archeologica Genova (external)*)
This task includes two pilots focusing on building knowledge graphs for musical heritage that currently is hardly or not accessible, e.g. orally transmitted practices, embedded in texts, etc.

Task 5: Managing musical heritage collections through knowledge graphs (*Leader: KCL - Participants: ALL (except DP) + External adopters: BNF, CLARIN, DARIAH, CLARIAH, Europeana*)
This task includes two pilots focusing on building knowledge graphs for supporting curators, owners, etc. in managing their large collections

Task 6: Studying musical heritage through (interlinked) knowledge graphs (*Leader: IREMUS - Participants: UNIBO, OU, KCL, NUIG, KNAW, Biblioteca Universitaria di Bologna (external)*)
This task includes pilots for demonstrating how *Polifonia* technologies can be leveraged to demonstrate, discover, document, support etc. research hypotheses/theories.

Task 7: Interacting with musical heritage knowledge graphs (*Leader: OU - Participants: UNIBO, ICBSA, The Stables Theatre, Milton Keynes, UK (external)*)

This task includes pilots demonstrating novel interaction solutions. The goal is to show how musical heritage knowledge graphs can be leveraged for supporting accessible and improved user experience.

Deliverables

- D1.1: Roadmap and pilot requirements - 1st version (M6)
- D1.2: Roadmap and pilot requirements - 2nd version (M18)
- D1.3: Pilots development: collaborative methodology and tools (M6)
- D1.4: Intermediate validation reports for pilots: ORGANS and BELLS (M24)
- D1.5: Intermediate validation reports for pilots: INTERLINK and FACETS (M24)
- D1.6: Intermediate validation reports for pilots: TONALITIES, TUNES, MUSICBO and CHILD (M24)
- D1.7: Intermediate validation reports for pilots: MEETUPS and ACCESS (M24)



- D1.8: Final ten-pilots validation report and lessons learned (M40)
- D1.9: *Polifonia* Web portal - 1st version (M18)
- D1.10: *Polifonia* Web portal - 2nd version (M36)

Appendix 4: Work packages in *Polifonia* project

Acronym	Title	Tasks
WP1	Web portal and pilots	<ol style="list-style-type: none"> 1. Socio-Technical Roadmap 2. Technical coordination 3. <i>Polifonia</i> web portal: an aggregator of digital musical heritage collections 4. Preserving musical heritage collections through knowledge graphs 5. Managing musical heritage collections through knowledge graphs 6. Studying musical heritage through (interlinked) knowledge graphs 7. Interacting with musical heritage knowledge graphs
WP2	Musical heritage knowledge graphs	<ol style="list-style-type: none"> 1. Ontology-based knowledge graphs for music objects 2. Ontology-based knowledge graphs for music objects context 3. Interlinking knowledge graphs 4. Licences, ownerships, and conditions of use
WP3	Mining musical patterns	<ol style="list-style-type: none"> 1. Pattern extraction 2. Pattern recognition and definition in monodic and polyphonic music 3. Network of musical patterns 4. Music classification from musical patterns



WP4	Extracting musical heritage knowledge from text	<ol style="list-style-type: none"> 1. Building and evaluating multilingual text corpora on musical heritage. Themes, reception, and... 2. Automatic extraction of time, space, events, people and musical artifacts from text 3. Automatic extraction of socio-cultural and historical context of musical heritage 4. Evaluation of automatic knowledge extraction methods
WP5	Human music interaction and engagement	<ol style="list-style-type: none"> 1. Evaluation of interaction components 2. Multidimensional searching, browsing and exploration of musical assets and knowledge 3. Haptic and gestural interaction with music 4. Pattern exploration, composition and visualisation 5. Publishing, accessing and reusing musical scholarly objects
WP6	Dissemination and exploitation	<ol style="list-style-type: none"> 1. Plan for exploitation and dissemination of results (PEDR) 2. Web presence and <i>Polifonia</i> image 3. Dissemination of project results 4. Innovation drive: stakeholder network 5. Enhanced accessibility and inclusion strategies 6. Musical heritage promotion: the <i>Polifonia</i> open source artistic digital installation 7. Data Management Plan
WP7	Project coordination and management	<ol style="list-style-type: none"> 1. Overall coordination of project activities, monitoring, quality control, reporting 2. Overall legal and contractual management 3. Consortium agreement, management of the knowledge generated by the project and IPRs 4. Ethical compliance 5. Financial and administrative management and reporting



		6. Project governance and partnership communication
WP8	Ethics requirements	<ul style="list-style-type: none">1. OEI – Requirement No. 12. H - Requirement No. 23. H - Requirement No. 64. H - Requirement No. 75. H - Requirement No. 86. POPD - Requirement No. 97. POPD - Requirement No. 10



Appendix 5: Story (Persona–QC–Scenario) Template

○ **Persona**

- It is a research-based description of a typical user.
- It contains attributes such as name, age, occupation (if the Persona has more than one role, indicate which one is their primary role and which one(s) the secondary role(s)), and relevant characteristics of the person such as their knowledge and skills and their interests.

○ **Goal**

- It is a short textual description of the goal(s) that the Persona needs to be addressed in the Story.
- maximum number of characters: 1200.
- The goal(s) is(are) also represented by a short (maximum 5) list of keywords.

○ **Scenario**

- It is a Story describing how the Persona's task/need/problem is solved before, during and after interaction with the resource/software/service being developed.
- maximum number of characters: 1200.

○ **Competency questions (CQs)**

- Question(s) the Persona needs the resource/software/service to answer for satisfying their task/need/problem.

○ **Resources (optional)**

- List of resources (with references/links) where it is expected or known that the Persona can find what she's looking for.



Appendix 6: Survey template

1. DOMAIN SPECIFIC PART.

1.1. Identification & characterisation of the sources, and available digital corpora or datasets.

1.1.1. Availability

1.1.1.1. Are the sources digitized?

1.1.1.2. If so, are the sources born digital?

1.1.1.3. Can you provide links to the sources in a (partial) list?

1.1.1.4. Where is the data stored?

1.1.1.5. How scattered is musical knowledge data over different sources?

1.1.2. Source characterisation. What is the nature of the sources?

1.1.2.1. Score

1.1.2.2. Sound source

1.1.2.3. Image collection (iconographic items, diagrammatic items, etc.)

1.1.2.4. Texts (poetry, librettos, writings about music, correspondence, theoretical sources, documents about gestuality, technical documents, etc.)

1.1.2.5. Audiovisual sources

1.1.2.6. Material objects (bells, instruments, theatre equipment, sources listed above explored from the perspective of their materiality)

1.1.3. Legal issues

1.1.3.1. Are there any copyright or licensing issues?

1.1.3.2. Who has the rights to the used datasets?

1.2. Form, type, extent of research outcomes

1.2.1. What are your scientific assumptions and/or initial hypotheses?

1.2.2. What kind of scientific knowledge do you intend to produce (for example factual knowledge, conceptual knowledge, procedural knowledge, or metacognitive knowledge)? The data that embody the new knowledge generated should be characterized:

1.2.2.1. According to their nature (melodic patterns, intertextual units, networks of composers and compositions, ...)

1.2.2.2. According to their "documentary & technical" relationships with the sources:

1.2.2.2.1. Is the new knowledge anchored to a specific fragment in a source? (e.g. "The note identified by the xml:id "m-69" in a MEI file is analysed as a passing note")



- 1.2.2.2.2. Is the new knowledge about a whole document (e.g. "The composer of this piece is H.I.F. von Biber")
- 1.2.2.3. According to their "status" with regard to the sources and to socio-scientific practices:
 - 1.2.2.3.1. Is the new knowledge purely descriptive? (e.g. a diplomatic transcription)
 - 1.2.2.3.2. Can the new knowledge be the subject of dissensus?
 - 1.2.2.3.3. Does the new knowledge involve an interpretation? And should the interpretative context be made explicit through linked data? (e.g. "Mrs X hypothesizes that this specific note N1 is an escape note, basing her opinion on these three annotations A1, A2, A3 and on this theoretical statement S1.")
 - 1.2.2.3.4. Could the knowledge be inferred by algorithms, or can it only proceed from human interpretation?
 - 1.2.2.3.5. Should it be possible to represent hypotheses and/or statements with varying degrees of certainty?
 - 1.2.2.3.6. Could the knowledge proceed from collaborative scientific practices?
- 1.2.2.4. According to their relationships with formalized "meta-knowledge":
 - 1.2.2.4.1. Does the new knowledge rely on a controlled vocabulary? (e.g. "A note could be a 'Consonant note', a 'Passing note', a 'Neighbor note', an 'Anticipation', a 'Suspension', an 'Escape note';)")
 - 1.2.2.4.2. Does the new knowledge involve the creation of new controlled vocabularies to reflect specific scientific analytical concepts?
 - 1.2.2.4.3. Could the links to the socio-cultural context be made explicit?
- 1.2.3. How do the methods and results contribute to the state of the art?
 - 1.2.3.1. To which discipline/subdiscipline/specialty do you aim to contribute knowledge?
 - 1.2.3.2. Can you provide examples of research questions that the pilot will allow to explore?
- 1.3. Methods used to derive the results from the sources and to articulate them with other sources.
 - 1.3.1. Identification/creation of analytical categories
 - 1.3.2. Identification/creation of analytical and conceptual tools



1.4. Please provide a specific contact person for questions related to the pilot's scientific dimension (pilot coordinator)?

1.5. What else has to be reported that is not covered by this part? Do you have any remarks?

2. **TECHNICAL PART.** The technical part specifies the actual implementation of the conceptual model:

2.1. Characterisation of existing datasets (more specific technical questions will be addressed in the data management plan).

2.1.1. Metadata

2.1.1.1. What formats are used?

2.1.1.2. What standards are used?

2.1.2. Datasets

2.1.2.1. What formats are used?

2.1.2.2. What standards are used?

2.2. Knowledge organisation systems and interoperability

2.2.1. Please list the controlled vocabularies, taxonomies, thesaurus, ontologies, etc. that you will use in your pilot

2.2.2. Please indicate if you plan to expand existing knowledge organisation systems or if you plan to create a specific one

2.2.3. Do you plan to support interoperability towards third parties?

2.2.4. Do you have API's on your resources?

2.3. Are there planned overlaps (already defined or envisioned) between the pilot's resources and other resources used within *Polifonia*?

2.4. What algorithms will be developed for knowledge extraction or alignment?

2.5. How should the pilot's resources be presented, linked, or stored on the website dedicated to the pilot?

2.6. What type of knowledge you expect can contribute to the *Polifonia* Web portal?

2.7. What type of interaction do you expect should be supported by the Web portal?

2.8. In which ways the Web portal shall link to the (separate) pilot demonstrator (if any)?

2.9. What kind of 'augmentation/enrichment' do you expect to get out of being displayed via the portal?

2.10. What do you expect in terms of technologies from the technical providers?

2.11. Can you provide a specific contact person for questions related to the pilot's technical aspects?

2.12. What else has to be reported that is not covered by this part? Do you have any remarks?



3. **SOCIO-PEDAGOGICAL PART.** The socio-pedagogical part aims to identify the target groups and their conditions

3.1. What is the target group of your pilot?

3.1.1. Amateur Internet users,

3.1.2. Musicians

3.1.3. Students

3.1.4. Academics (specify, if possible, their field of expertise)

3.1.5. Curators

3.1.6. Others (please specify)

3.2. In order for the pilot to be successful, what criteria would the new tools and methods need to satisfy (e.g. certain tasks carried out more quickly or accurately, the amount or type of data that can be used in the task, number of people who are able to use the tools, the attractiveness of the interface, new tools being easy to learn, etc.)?

3.3. Do you have any thoughts on how the success of the new tools and methods could be measured (e.g. time and accuracy with which tasks are carried out, usability surveys, measures of interface aesthetics, observational studies, interviews, etc.)?

3.4. What facilities, if any, does the pilot make available for people with disabilities?

3.5. Can you provide a specific contact person for questions related to the pilot's socio-pedagogical implications?

3.6. What else has to be reported that is not covered by this part? Do you have any remarks?

4. Do you have any general comments and remarks?