



HAL
open science

D1.8: Final ten-pilots validation report and lessons learned (V1.0)

Nicholas Canny, Jason Carvalho, Alba Catalina Morales Tirado, Enrico Daga, Jacopo de Berardinis, Raphaël Fournier-S 'Niehotta, Arianna Graciotti, Christophe Guillotel-Nothmann, Marco Gurrieri, Simon Holland, et al.

► **To cite this version:**

Nicholas Canny, Jason Carvalho, Alba Catalina Morales Tirado, Enrico Daga, Jacopo de Berardinis, et al.. D1.8: Final ten-pilots validation report and lessons learned (V1.0). Open University; King's College London; IReMus; CNAM; NUIG; KNAW; MiC; Università di Bologna. 2024. hal-04582760

HAL Id: hal-04582760

<https://hal.science/hal-04582760>

Submitted on 22 May 2024

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.



Polifonia: a digital harmoniser for musical heritage knowledge, H2020

D1.8: Final ten-pilots validation report and lessons learned (V1.0)

Deliverable information	
WP	WP1 Pilots and Web Portal
Deliverable dissemination level	PU Public
Deliverable type	O Other
Lead beneficiary	OU
Contributors	KNAW, CNRS, KCL, CNAM, UNIBO, NUIG
Document status	Final
Document version	V1.0
Date	April 29, 2024
Authors	Jason Carvalho, Alba Morales Tirado, Enrico Daga, Simon Holland, Nicholas Canny (OU), Jacopo de Berardinis (KCL), Christophe Guillotel-Nothmann, Marco Gurrieri (CNRS), Raphaël Fournier-S'niehotta (CNAM), James McDermott, Rory Sweeney (NUIG), Andrea Scharnhorst, Peter van Kranenburg (KNAW), Elena Musumeci (MiC), Eleonora Marzi, Arianna Graciotti (UNIBO)
Peer review	Each chapter was reviewed by a co-author. Overall document reviewed by the editors Enrico Daga, Jason Carvalho (OU), and Andrea Scharnhorst (KNAW).



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004746

PAGE INTENTIONALLY BLANK

Project Information

Project Start Date: 1st January 2021
Project Duration: 40 months
Project Website: <https://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	MINISTERIO 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 VOOR BEELD EN GELUID	Netherlands
9	KNAW	KONINKLIJKE NEDERLANDSE AKADEMIE VAN WETEN- SCHAPPEN	Netherlands
10	DP	DIGITAL PATHS	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

This deliverable is the final deliverable on Pilots in Workpackage 1 of Polifonia, next to the deliverable D1.10 about the Polifonia Web Portal. WP 1 (Pilots and Web portal) has been also called Polifonia's validation WP. It contributes to achieve the objectives O3 (Tailor) and O5 (Share & Engage)¹. As such it drives the whole development of Polifonia and in particular the monitoring and aligning of the 10 Pilots. This last deliverable builds on previous deliverables in this WP. Among them D1.1 *Roadmap and pilot requirements 1st version* [1], D1.2 *Roadmap and pilot requirements – 2nd version* [2], and D1.3 *Pilots development - collaborative methodology and tools* [3] defined the collaborative strategy of the project, the main knowledge sources and epistemological units (research questions), and how to best monitor and management the overall progress in an otherwise distributed knowledge exploration process. In this process each of the Pilots can be envisioned as a *scout* send out to explore other dimensions of the actual problem space. But, to achieve a scientific progress in such an interdisciplinary setting as Polifonia represents, the knowledge collected by those *scouts* needs to be brought together again and integrated. Polifonia did this in an iterative way, and dedicated a specific governance board - the Technical Board - to this task. The individual Pilots each on its own way also operated on fronts of research, methods, and data curation. For instance for the more methodologically operating Pilots (scope Studying) the work crucially includes cutting-edge experiments evaluating Large Language Models (LLM) in the context of various scenarios relevant to the stakeholders. The latter belongs to the unforeseen new challenges during the journey of the project. Another central challenge concerns how the consortium has been experimenting on how symbolic knowledge (Knowledge Graphs) can interplay with predictive, generative models. To both acknowledge the Pilot specific achievements and the contribution to the project-wide goals, this deliverable devotes one section to each Pilot, but uses the same kind of template (subsection headings) to report about the Pilots. In their report the focus lies on main advances and the validation of results. Moreover, each Pilot description details on which other (more elaborate Deliverables) this work is based. In a last section, this deliverable includes discussion of the reusability of Polifonia's research and outputs. We have gathered data on the reusability of different components of the Polifonia ecosystem via a self-reporting survey, and we report on this in Chapter 12. Further, we revisit a table of anticipated reusability of components, originally presented in the Polifonia Grant Agreement, and map it against our outcomes.

¹Description of Action Polifonia 101004746, document Ref. Ares(2020)5825178 - 23/10/2020 PART B, *internal document*, section 1.1.

Document History

Version	Release date	Summary of changes	Author(s) - Institution
v0.1	1/12/2023	Deliverable layout	Enrico Daga (OU)
v0.2	09/02/2024	Deliverable structured and subsections	Enrico Daga, Jason Carvalho (OU)
v0.7	11/10/2023	Authors develop first draft, Introduction and Executive Summary drafted	Jason Carvalho, Alba Morales Tirado, Enrico Daga, Simon Holland, Nicholas Canny (OU), Jacopo de Berardinis (KCL), Christophe Guillotel-Nothmann, Marco Gurrieri (CNRS), Raphaël Fournier-S'niehotta (CNAM), James McDermott, Rory Sweeney (NUIG), Andrea Scharnhorst, Peter van Kranenburg (KNAW), Elena Musumeci (MiC), Eleonora Marzi, Arianna Graciotti (UNIBO)
v0.8	10/04/2024	Version sent to reviewers for feedback	Enrico Daga (OU)
v0.85	22/04/2024	Integrate feedback from reviewers	Jason Carvalho, Alba Morales Tirado, Enrico Daga, Simon Holland, Nicholas Canny (OU), Jacopo de Berardinis (KCL), Christophe Guillotel-Nothmann, Marco Gurrieri (CNRS), Raphaël Fournier-S'niehotta (CNAM), James McDermott, Rory Sweeney (NUIG), Andrea Scharnhorst, Peter van Kranenburg (KNAW), Elena Musumeci (MiC), Eleonora Marzi, Arianna Graciotti (UNIBO)
v0.9	25/02/2024	Revised version.	Jason Carvalho, Enrico Daga (OU), Andrea Scharnhorst (KNAW)
v0.95	29/02/2024	Revised version.	Jason Carvalho, Enrico Daga (OU),
v1.0	29/02/2024	Submission to EU	Coordinator

Table of contents

1. Introduction	1
2. ACCESS	3
2.1. Overview	3
2.1.1. Haptic Bracelets	3
2.1.2. Harmony Space	5
2.2. Connection with research WPs	6
2.3. Progress since last deliverable	9
2.3.1. Haptic Bracelets	9
2.3.2. Harmony Space	15
2.4. Evaluation method/rationale	50
2.4.1. Haptic Bracelets	50
2.4.2. Harmony Space	50
2.5. Results and discussion	50
2.5.1. Haptic Bracelets	50
2.5.2. Harmony Space	51
2.5.3. Further work on both Harmony Space and the Haptic Bracelets beyond Polifonia	51
2.6. List of outputs (FAIR section)	52
2.7. Appendix 1	53
2.7.1. Anonymised PEER FEEDBACK from children after first visit	53
2.8. Appendix 2 Harmonising a Melody	55
3. ORGANS	56
3.1. Overview of the pilot	56
3.1.1. Connection with research WPs	57
3.1.2. Progress from last deliverable D1.4	57
3.2. Evaluation method/rationale	57
3.3. Results and discussion	58
3.3.1. Data extraction and SQL export	58
3.3.2. Ontology and Knowledge Graph	60
3.4. Summary of the outputs (FAIR section)	62
4. BELLS	63
4.1. Overview of the pilot	63
4.1.1. Connection with research WPs	63
4.1.2. Progress from last deliverable D.1.4	63
4.2. Evaluation method/rationale	64
4.3. Results and discussion	64
4.3.1. Quantitative	64
4.3.2. Qualitative	64
4.4. Summary of the outputs (FAIR section)	68
5. INTERLINK	69
5.1. Overview of the pilot	69
5.1.1. Connection with research WPs	69
5.1.2. Progress from last deliverables	70

5.2. Evaluation method/rationale	70
5.3. Results and discussion	71
5.4. Summary of the outputs (FAIR section)	72
6. FACETS	74
6.1. Overview of the pilot	74
6.1.1. Connection with research WPs	74
6.1.2. Progress from deliverable D1.5	75
6.2. Evaluation method/rationale	75
6.3. Results and discussion	76
6.3.1. Results	76
6.3.2. Discussion, lessons learned	77
6.4. Summary of the outputs (FAIR section)	77
7. TONALITIES	79
7.1. Overview of the pilot	79
7.1.1. Connection with research WPs	79
7.1.2. Progress from last deliverable D1.6	79
7.2. Evaluation method/rationale	80
7.3. Results and discussion	80
7.3.1. Quantitative validation of input data	80
7.3.2. Validation of input data quality by experts	82
7.3.3. Validation of the interface by users/experts	83
7.3.4. Validation of performance, requests/information passed between front and back end	83
7.3.5. Validation of results by ground truth	84
7.4. Future perspectives	86
7.5. Summary of the outputs (FAIR section)	87
8. TUNES	89
8.1. Overview of the pilot	89
8.1.1. Connection with research WPs	89
8.1.2. Progress from last deliverable D1.6	89
8.2. Evaluation method/rationale	90
8.3. Results and discussion	91
8.3.1. Similarity Relations	91
8.3.2. Ontology and Knowledge Graph	93
8.4. Summary of the outputs (FAIR section)	94
9. MUSICBO	96
9.1. Overview of the pilot	96
9.1.1. Connection with research WPs	96
9.1.2. Progress from last deliverable D.1.6	98
9.2. Evaluation method/rationale	99
9.2.1. Input validation	99
9.2.2. Validation of interface and performance by experts	99
9.2.3. Output/Goal Validation	100
9.3. Results and discussion	101
9.4. Summary of the outputs (FAIR section)	101
10. CHILD	103
10.1. Overview of the pilot	103
10.1.1. Background and connection with research WPs	103

10.1.2. Progress from deliverables D1.6, D4.3, and D4.4	104
10.2. Evaluation method/rationale	104
10.2.1. The Documentary Evidence Benchmark (DEB)	104
10.2.2. Prompt Engineering for Thematic Classification (Task 1)	105
10.2.3. Prompt Engineering for Metadata Curation (Task 2)	105
10.2.4. Datasets for the evaluation	107
10.2.5. Experiment 1: classification of thematic evidence	107
10.2.6. Experiment 2: curation of documentary evidence metadata	108
10.3. Results and discussion	109
10.3.1. Classification of Childhood Evidence (Experiment 1)	109
10.3.2. Cataloguing of documentary evidence text (Experiment 2)	110
10.3.3. Discussion	112
10.4. Summary of the outputs (FAIR section)	113
11. MEETUPS	114
11.1. Overview of the pilot	114
11.1.1. Connection with research WPs	114
11.1.2. Progress from last deliverable D4.4	115
11.2. Evaluation method/rationale	115
11.2.1. Knowledge Extraction pipeline - data evaluation.	115
11.2.2. Feedback questionnaire from domain experts.	116
11.3. Results and discussion	117
11.4. Summary of the outputs (FAIR section)	117
12. Reusability survey	119
12.1. Design and rationale	119
12.2. Reusability Survey	119
12.2.1. Questions 1-3: Names, IDs, links	119
12.2.2. Question 4: Project types	119
12.2.3. Question 5: Target audience	120
12.2.4. Question 6: Aspects of reusability	120
12.2.5. Questions 7-8: Outside reuse	121
12.2.6. Questions 9-12: Data formats and technologies	122
12.2.7. Question 13: TRLs	122
12.2.8. Question 14: Reuse strategies	123
12.2.9. Questions 15, 16 and 19: Reuse for specific project types	124
12.2.10. Questions 17-18: Extensibility and domain independence	125
12.3. Technology Readiness Level	125
12.4. Conclusion	126
13. Compliance to the FAIR principles	127
14. Conclusions	128
A. Polifonia Reusability Survey	134

1. Introduction

Polifonia's main goal is to enhance access to our musical cultural heritage. The opening statement on the <https://polifonia-project.eu> reads "From the soundscape of Italian historical bells, to the influence of French operas on traditional Dutch music, European cultural heritage hides a goldmine of unknown encounters, influences and practices that can transport us to experience the past, understand the music we love, and imagine the soundtrack of our future." To realise this huge aspiration during the limited life time of a project, Polifonia took a decentralised approach. Realising that the diversity characterising *Musical Heritage* is both a value and a curse, Polifonia organised its research around 10 Pilots, representing different types of actions when dealing with cultural heritage, in doing so responding to the scope and challenges of the original call this project responded to. Those types of actions are: preserving, studying, interacting and managing. As indicated in Figure 1.1 (Figure taken from [4], see also [5]).

Workpackage 1 for which this is the final deliverable documenting progress and final status of the pilots acts as main 'validation' work package. 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, and 2) to contribute to push the state of the art in relevant, though specific, musical heritage use cases, represented by the pilots.

The challenge in Polifonia is to contiguously balance between the necessary independence of the explorations done by the pilots and to wake about their interdependence in the light of the overall project goals. For this their work is aligned to a socio-technical roadmap [2][1], supervised by a Technical Board, and accompanied by various instruments of documentation (see here the three iterations of the Data Management Plan [6],[7], [8]).

This deliverable gives a final overview of the progress in the 10 pilots according to both their pilot specific goals and their function in the overall project. To create this overview each pilot reports according to a shared template: starting with an overview (including cross-connections with other WP's and where appropriate references to and updates on reports given in other deliverables). The central focal point in each Pilot case section is the description of the evaluation rational and method. Followed by a discussion and thoughts on FAIR.

Dedicated to FAIRness, Open Science and the highest degree reproducibility and re-usability possible, Polifonia through all project (and other meetings) discussed how its results can be best disseminate and having a substantial impact (academically and beyond; short mid and long-term). In the original Description of Action ¹ Polifonia defined success indicators for each objective. Among them are also indicators which are related to 're-usability' of results. For instance, indicators such as "# forks, downloads, and visits to GitHub: in the order of hundreds - Objective O5 Share and Engage" or "# interfaces per types of heritage objects: all types relevant to the pilots - Objective O4 Experience and Understanding" are related to reusability. At the 2023 Bologna meeting a whole afternoon was dedicated to discuss what re-usability means, for which groups and purposes. We also discussed what part of the knowledge production process can be documented in a way that it can be reproduced by others, and which parts contain specific or tacit knowledge which requires specific action to travel. Think here of training of students in a research and lab versus an interface which enables interaction with relatively light documentation. As a result, and governed by the Technical Board a typology of reusability was developed, and a reusability survey was set out and run via all Pilots. The results of this are described in the section Reusability 12. We would like to emphasis here, that again machine actionability was a concious engineering choice, making this re-usability survey to an instrument which could also be applied for and by other projects. This is (methodologically) comparable to the design of the Polifonia Ecosystem which represents a machine executable Research Data Management plan.

¹Description of Action Polifonia 101004746, document Ref. Ares(2020)5825178 - 23/10/2020 PART B, *internal document*, p.11



Figure 1.1: Overview about the Pilots

2. ACCESS

2.1. Overview

The ACCESS pilot explores technologically mediated ways to improve inclusion and engagement in musical activities via innovative haptic, gestural, and visual interaction. This is being achieved by music outreach and educational workshops, along with technologies that support inclusion of those with hearing and physical disabilities as first-class citizens.

To this end, the ACCESS Pilot is using two technology vehicles developed at the Open University:

- Haptic Bracelets, aimed at inclusive engagement with rhythmic music making,
- Harmony Space, aimed at inclusive engagement with harmony.

2.1.1. Haptic Bracelets

The haptic bracelets are self-contained lightweight wireless bracelets for wrists and ankles (Figure 2.1) that contain processors, motion sensors, fast Wi-Fi links, and rapidly acting powerful haptic motors (vibrotactiles). The bracelets have several modes of use, as described below. Depending on the specific task in hand, participants may wear four bracelets at once (wrists and ankles), or just one or two. Motion sensors make it possible for one (or several) participants to feel in real-time, in the appropriate limb, what each limb of another player is doing (Figure 2.2). It is also possible for several players to be coordinated using a recorded 'haptic score' (Figures 2.3, 2.4, and 2.5), or for haptic recordings to be played back slowed down for subsequent study. Finally, combinations of these modes have varied musical applications.

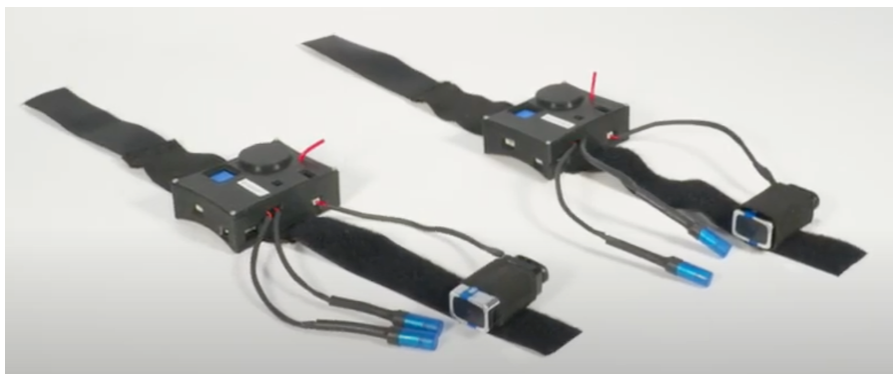


Figure 2.1: Two haptic bracelet units, worn on ankles and wrist as needed. Each bracelet can sense impacts created by the limb on which it is worn, which can be transmitted in real time to be felt on the corresponding limb of another wearer. The vibrotactile units are the small blue cylinders seen in the image. If two or more players are transmitting to each other at the same time, these units are typically fastened slightly away from the box-like sensing unit seen in the image, to avoid haptic feedback - analogous to microphone feedback.



Figure 2.2: Two drummers, each wearing four haptic bracelets, one for each wrist and ankle. The teacher in the foreground is transmitting a repetitive drum pattern to the learner in the background in real time (achieved with a latency under 10 ms.).

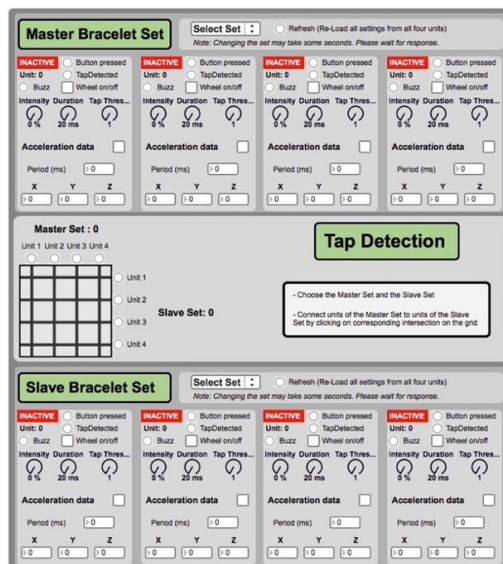


Figure 2.3: Any bracelet can be set to control any other bracelet using the graphic routing table in the middle of the control panel shown here.

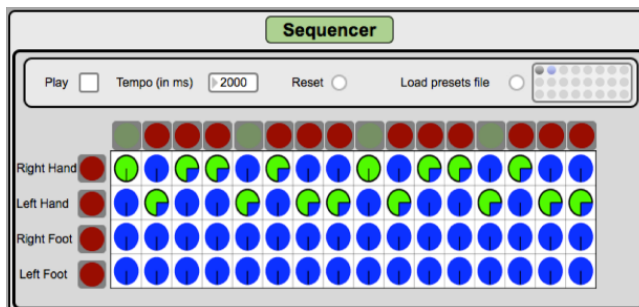


Figure 2.4: A simple, easy-to-program, four-track, sixteen-beat haptic sequencer. Different beats may be set with different relative intensities (as indicated by the green fractions of the blue circles).

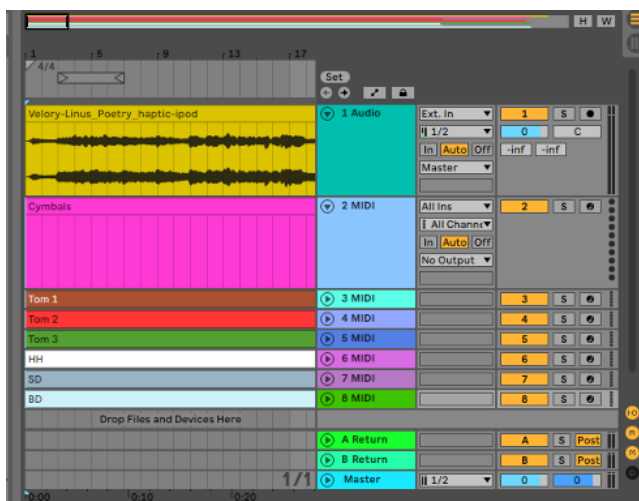


Figure 2.5: A stereo audio recording synchronised with a six-drum midi recording as seen in Ableton. The midi parts can drive the Haptic Bracelets mapped to different limbs, or different players.

Use scenarios

The haptic bracelets can support a range of user scenarios. For example, they have applications that include the following:

- support the development of practical rhythm skills and multi-limb coordination for one or more individuals using a drum kit
- support inclusion, engagement, and co-ordination for a group of players, each playing one drum in a samba group, (or one instrument in a gamelan)
- support the inclusion and engagement of d/Deaf participants as first-class participants in such activities

2.1.2. Harmony Space

The aim of Harmony Space within the ACCESS pilot is to explore technologically mediated ways to improve inclusion and engagement in harmonic aspects of musical activities. In order to understand the work with Harmony Space, it is useful to briefly consider some context and background.

Harmony concerns the organisation of multiple simultaneous voices playing independent but co-ordinated pitches. The structures and processes involved in this co-ordination can be complex. Indeed, harmony is widely considered

to be the most distinctive and complex aspect of the European musical heritage. Many other musics explore complex melodies (e.g. Carnatic and other Indian traditions) and many other traditions explore rhythm in depth (e.g. music associated with African and Indian cultures). However, no tradition appears to have explored harmony in more depth than the European heritage and closely associated traditions (such as jazz).

Obstacles to deeper engagement with harmony

For the musically untrained general public, listening to harmonies in appropriate musical contexts can produce rich emotional responses, and indeed many without formal training can spontaneously sing simple harmonies. However, generally, there are considerable barriers to developing deeper engagement with harmonic aspects of music - arguably more so than in the case of rhythm or melody. To engage deeply with any aspect of music (or indeed any other medium), it appears to be necessary to be able to play actively with the relevant elements, i.e. to gain experience of their active manipulation and modification - simple passive exposure to examples appears to be insufficient. Accordingly, for those who have not mastered a musical instrument that can produce multiple pitches simultaneously (such as a piano or guitar) it can be hard to gain meaningful experience of playing with, generating, modifying, and manipulating the stuff of harmony. This lack of experience can be a substantial barrier to deeper engagement. Harmony Space aims to overcome such barriers.

Obstacles for the musically untrained general public to deeper engagement with harmony go beyond the problem of lacking physical means to fluidly manipulate harmony. There is also the issue of language and terminology. This is not just a matter of labelling. Harmony involves complex concepts and relationships, and explicit discussion of tonal harmony involves articulating relationships between a wide range of abstract entities, relationships, and terminology. In conventional music theory these concepts are approached using somewhat confusing and inconsistent terminology, in a manner reminiscent of studying algebra or language grammars. These concepts span interlinked semi-hierarchical conceptual levels (for example, pitch, interval, scale, chord quality, voicing, mode, key, and modulation). Each of these levels is associated with its own technical vocabulary. Harmony Space aims to use a single consistent extended spatial metaphor to reduce the cognitive load needed to understand these abstractions if and when they are needed. At the same time, Harmony Space reveals distinctive visual patterns that can promote identification of, engagement with, and manipulation of, high level harmonic features associated with different genres of music.

2.2. Connection with research WPs

The ACCESS pilot has a two-way relationship with the Polifonia ChoCo JAMS Data set of some 20,000 chords.

How Harmony Space has benefited from the Polifonia ChoCo JAMS DataSet

Harmony Space has been interfaced with the Polifonia ChoCo JAMS Data set of some 20,000 chords, so that users of Harmony Space can look up, download, manipulate and benefit generally from engaging with these chord sequences.

However, this client relationship between Harmony Space and ChoCo raises some interesting and non-obvious issues. The representation that Harmony Space uses for harmony is distinctively different from the ChoCo ontology, so non-trivial heuristic translation steps had to be developed by ACCESS.

- ChoCo (and the underlying Harte Notation) assumes that the chromatic scale is scalar, whereas Harmony Space uses an ontology where this pitch set is inherently three dimensional (independently of exact tuning ratios). Indeed, it is the representation manipulation of these three dimensional relationships that give Harmony Space its power.
- Due to the inherently three-dimensional nature of pitch, chord sequences can be plotted spatially in Harmony Space, and the spatial shapes and paths give information about harmonic structure. This information is absent,

or at best implicit in ChoCo. Consequently spatial mapping heuristics were developed in Harmony Space to allow the spatial mapping of chord sequences from ChoCo, with the user able to compare the results of contrasting heuristics. Comparing different spatial translations can amplify harmonic insight (a simple example of how this can be afforded is discussed in the report on the second school visit - see "Group 3's novel chord sequence").

- ChoCo records keys and chords – but the key information is disconnected from the chord information. That is to say, the recorded key of a piece can be changed without making any difference to the following chords (which, for example, is not the case with common music notation, nor with Harmony Space - which responds fluidly to key information).
- Recordings in the Harmony Space library can go beyond ChoCo recordings in that typically embody a full harmonic analysis. In the Harmony Space internal representation, where not indicated otherwise, an assumption of diatonicity is used, which enables *diatonic compression*, whereby chord qualities are re-created from live key, mode, mode extension and root information, with departures from diatonicity noted as needed.
- Recordings in the Harmony Space library are spatial and generative, so that they can typically generate a family of related chord sequences by manipulating parameters such as chord size, inversions, chord maps, etc - unlike statically represented Harte sequences in ChoCo.

How Harmony Space feeds back into ChoCo

Harmony Space feeds back into the Polifonia ChoCo JAMS DataSet in two ways. Firstly Harmony Space opens up the possibility of democratizing or widening meaningful feedback to ChoCo in the way described below. Secondly Harmony Space expands the kinds of harmonic data that ChoCo can record.

Widening meaningful interaction with ChoCo

In the Polifonia ChoCo JAMS DataSet, sometimes sources contain wrong key information. Sometimes this is due to simple slips or omissions. At other times, modal pieces are wrongly interpreted as tonal, or modulations are missed. Clearly any well-trained musicians could identify such errors, and, if they have access to an editor that can output suitably packaged Harte notation, could send back corrections to ChoCo. However, Harmony Space opens up the possibility of democratizing, or lowering the barriers needed for such feedback, as follows.

- At present, using conventional means, there is a high bar to knowing enough about harmony to be able to correct key attribution errors. Typically this depends on explicit expert knowledge about the concepts and notations of scales, keys and modes, and the notation of chords and chord alterations. Even many competent amateur musicians lack this knowledge to a sufficient degree.
- By contrast, using Harmony Space, how to establish the key of a section can be communicated in principle to musical beginners, or those with limited musical knowledge as a visuospatial task. Roughly speaking, this could correspond, for example, to moving the white-shaded area in (Figure 2.6), so that as many shared circles as possible fall in the white rather than grey areas (Figure 2.7). See the discussion of children undertaking this task in the first school visit (section "Day 2 Activity three: spatially identifying the key of a piece").
- Of course, there would be downsides to flooding ChoCo with corrections from complete beginners, and checks and balances would be required. However, there are positives here for engagement, and for the lowering of barriers for relative beginners or amateurs. Finally, it would not be hard for Harmony Space to generate quantitative analytical data to check judgements arrived at visually using the containment metaphor.
- Harmony Space has the necessary export function to package corrections of this kind for ChoCo.

Extensions of the ChoCo ontology

Spatial examination of the popular music corpus in Harmony Space can reveal building blocks and extensions of harmony in different genres. Some of the simplest such extensions are variants of modes commonly found in popular

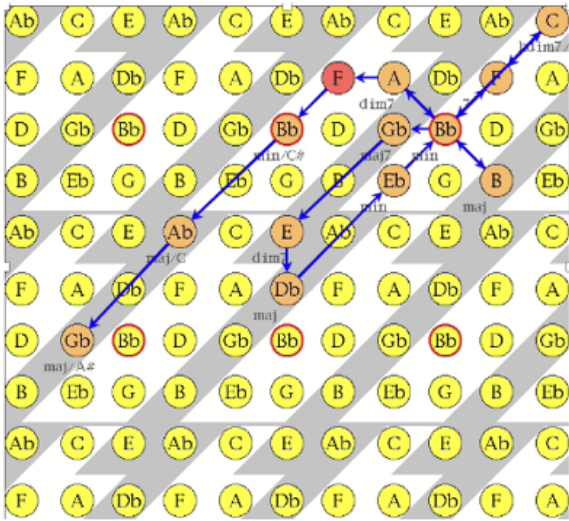


Figure 2.6: Visual key finding. This figure and the next demonstrate how to establish the key of a section of a harmonic sequence as a visuo-spatial task in Harmony Space. By sliding the white areas (representing the diatonic scale) to a position that maximises number chords they fully contain, candidates for the likely key and mode can be established. In addition, as the white area is slid, the system (slightly oversimplifying) marks as red the roots of chords contained in the white area, and as orange the roots of chords not fully contained in the white area.

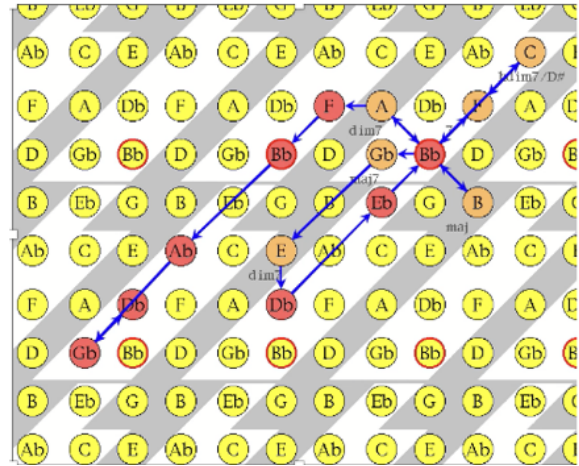


Figure 2.7: Following the previous figure, in this way, visual containment can act as an accurate physically manipulable embodied metaphor for diatonicity. The chord sequence from Schubert shown in this and the previous figure was marked in ChoCo as being in B major, but this visual exercise suggests that it is more likely to be in parallel B minor mode. These involve a different position for the white areas, and reveal the highly connected Bb to be at the location in the white area corresponding to the major, rather than the minor, modal centre.

music, such as the Picardian and Harmonic (i.e. dominant V) variants of modes other than the commonplace Aeolian. From a melodic viewpoint, it might be argued that only the Aeolian can be inflected as Picardian, but from a harmonic point of view, examples are widespread, especially in popular music. For example, Old Town Road, by Lil Nas X, based on a sample originally from Nine-Inch Nails, uses the Picardian Dorian to good musical effect (Figures 2.8 and 2.9). Feedback from Harmony Space allows extensions such as these to be evidenced for extensions to the JAMS and ChoCo namespaces.

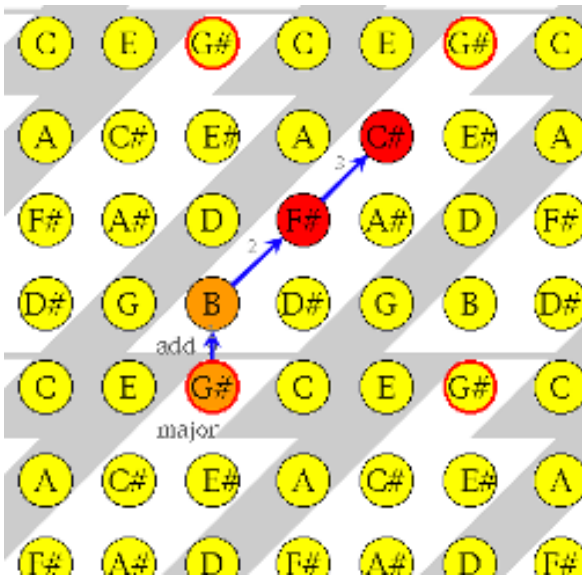


Figure 2.8: The chord sequence for Old Town Road, by Li'l Nas X: (G# Picardian Dorian) G# major G# major B add9 B major F# major F# major C# major C# major provides a good example of the use of the Picardian Dorian. The plot in this figure shows the root progression only, whereas the plot shown in the next figure shows all of the notes in each chord. Nowhere under strict diatonic constraints can a major triad be found a minor third above another major triad, so that the first two chords create a subconscious puzzle for the western educated ear, due to the stealthy use of the Picardian Dorian from the beginning of the chord sequence.



Figure 2.9: Following the narrative of the previous figure, for listeners who have grown up in a western tonal musical environment, this gives the chord sequence a twist or frisson on hitting the B major chord. In fact, the B major has an added 9, so that the scene-setting major third of the initial major triad is not just crushed by the root of the second chord but is squeezed a semitone above and below by the root and its ninth. Harmony Space lowers barriers for relative beginners to articulate and manipulate such relationships through physical and embodied means.

2.3. Progress since last deliverable

2.3.1. Haptic Bracelets

The previous validation (described in D1.7) had satisfied our Deaf educator Sean Chandler¹ that our approach with the Haptic Bracelets gives a clear path for deaf people to participate as first class citizens in musical activities alongside hearing people. At the same time we were keen to explore how well the haptic bracelets are capable of supporting a wider range of practical musical tasks, to understand current technical limitations, and to identify ways to improve the bracelets to support inclusion in a wider range of musical activities. This led on to the next set of validations, following on from the previous deliverables.

¹Sean is an accomplished and recognised leader of music workshops with deaf children across the UK. (Sean prefers the word 'deaf' to 'hearing impaired' – considering deaf people to have differences but not to be impaired in any way). Most music educators of the deaf are not themselves deaf – limiting their depth and quality of engagement with deaf people. By contrast, Sean is one of only a tiny handful of such leaders who is both a professional musician and profoundly deaf. Sean is unusual in that while being expert at lip reading and having perfect spoken English (having grown up with hearing parents) he is also a completely fluent BSL signer and is thus also totally integrated with (capital 'D') Deaf Culture. Thus, he describes himself as DeaF (with the 'F' standing for Fluid, bi-modal, bi-lingual, bicultural, and at home in both worlds).

Exploring the challenges of larger percussion ensembles

While the previous evaluation had established that the Haptic Bracelets could provide a path for deaf people to participate as first class citizens in musical activities alongside hearing people, that validation had been carried out in a relatively small-scale group setting, and we wondered whether additional considerations or refinements for the bracelets might be required to suit the context of larger and more complex percussion ensembles. We decided to begin exploring the relevant issues in a small pilot (see next section) with Ravin 'Raz' Jayasuriya, who has deep experience of large percussion ensembles.



Figure 2.10: Members of the Garden City Brazilian and Cuban Samba Ensemble working with the haptic bracelets to trial various synchronisation and learning tasks.

2.3.1.1. Large ensemble Brazilian and Cuban percussion

Pilot Workshop 14th June 2023

Participants

- Graeme Surtees (Head of Learning and Participation, The Stables)
- Simon Holland (Director Music Computing Lab, The Open University)
- Jason Carvalho (Project Officer and Developer, The Knowledge Media Institute, OU)
- Enrico Daga (PI Polifonia Project, The Open University)
- Paul Mulholland (Polifonia Project, The Open University)
- Nicholas Canny (Polifonia Project, The Open University)
- Ravin 'Raz' Jayasuriya Professional percussionist and Tutor specialising in Brazilian and Cuban percussion

Purpose of the pilot workshop

Brazilian and Cuban percussion performed by large ensembles (around 30 to 40 amateur players led by a master drummer) is a highly social activity, involving weeks of rehearsal leading to large-scale performances at public events, often outdoors. Based on feedback from the earlier validation, and in consultation with master drummer 'Raz' Jayasuriya we determined that there were aspects of communal synchronisation and signalling involved in this kind of drumming which had not been fully taken into account in our validations to date. Raz suggested exploring the

relevant issues in pilot workshop before trying out the technology with a large ensemble. Consequently we held a pilot workshop at the Music Computing Lab with Ravin 'Raz' Jayasuriya and members of the ACCESS pilot.

The previous workshops had shown that the haptic bracelets could be used to communicate rhythmic figures accurately by sense of touch alone; that the perceiver could perform such sensed rhythms accurately in synch with other players; and that they could feel how the rhythm they were performing meshed with other rhythms. However, this left unexplored various aspects of synchronisation and signalling relevant to Brazilian and Cuban ensemble drumming.

Activities

The workshop involved systematically working through tasks that the master drummer identified as going beyond activities in the previous validations.

We investigated the communication of smooth rapid tempo changes by the drum master to a drummer (without line of sight or clear hearing) whose task was to play a given rhythm. We tried this with the drum master beating the changing tempo on the side of his body or his leg, with each beat identified by impact detection via the haptic bracelets. We also explored three alternatives suggested by members of the workshop, the first using a microphone and a Max/Msp patch to detect a foot tap on a stomp box, the second using a button on an Ableton Push, and the third using a hand tap on a midi controller.

We investigated ways of haptically communicating *accented*, as opposed to unaccented figures.

We investigated ways of haptically communicating melodic rhythmic patterns (e.g. a figure that included both mid- and low- surdo drums to a single player).

We investigated the limits of temporal resolution for communicating figures haptically.

We investigated possible haptic substitutes for the whistle blasts conventionally used by a drum master to announce imminent changes in the rhythm, especially in out of doors contexts with spatial separation of tens of feet, no clear hearing, and no guaranteed line of sight.

Findings

- The impact detection system in the current version of the haptic bracelets sometimes misses the impact of a hand strike against the body. This is not a problem when using the bracelets to learn a repetitive rhythmic figure, as the pattern can be picked up over two or three repetitions. However, for transmitting now-or-never time critical commands from the drum master, additional sensing devices would be helpful.
- Melodic rhythm patterns (e.g. figures that involve two or more instruments) could be distinguished using two or more bracelets, but in a large ensemble context, the drum master needs some kind of routing or switching device for addressing different players.
- Using our current rotary vibrotactiles, temporal resolution blurred out at about quarter notes or eighth notes. Sixteenth notes could not be unambiguously distinguished.
- Differences in accent were hard to perceive clearly, since due to haptic masking (the obscuring of haptic signals when hitting things), the vibrotactiles needed to be run at maximum output all the time.
- The drum master suggested that a combination of haptic and visual signals could work well for d/Deaf players.

Next steps

The next step was to investigate these issues in more depth with a large percussion ensemble.

2.3.1.2. Validation: Workshop with the Garden City Brazilian and Cuban Samba Ensemble

Location: Jackman's Community Centre, Letchworth Garden City, UK

Date: 11th July 2023

Participants:

- Around 20 members of the Garden City Samba Ensemble
- Graeme Surtees (Head of Learning and Participation, The Stables)

- Simon Holland (Director Music Computing Lab, The Open University)
- Nicholas Canny (Polifonia Project, The Open University)
- Jason Carvalho (Project Officer and Developer, The Knowledge Media Institute, OU)
- Enrico Daga (PI Polifonia Project, The Open University)
- Paul Mulholland Polifonia Project, The Open University)
- Ravin 'Raz' Jayasuriya Professional percussionist and Tutor specialising in Brazilian and Cuban percussion



Figure 2.11: Garden City Samba members trying out the Haptic Bracelets

Purpose of workshop

Based on issues identified in the pilot, the purpose of the workshop was to investigate the demands of tempo changes; section and event marking; synchronisation; multiple patterns; accented beats; and melodic rhythmic figures in a Brazilian and Cuban ensemble context, and to seek candidate ways of overcoming limitations.



Figure 2.12: Garden City Samba members working with the Haptic Bracelets

Feedback from Players

Key items of feedback were as follows, paraphrased from interviews with the drum master and members of the ensemble after the session.

- **1 Tempo adjustment** How to increase tempo needs re-thinking. It worked fine when a signal was sent to the mid & low surdo players playing by themselves but not when there was a larger group playing on other instruments at the same time.

- **2 Learning patterns and temporal resolution** We sent a signal to three surdo players who each had two devices, one for tempo and the other for patterns. The signals sent were either a 1 or 2 bar patterns. This worked fine when the patterns consisted of crotchets &/or quavers but did not work when the patterns included semi-quavers.
- **3 Pattern against tempo** For a player using 2 devices, it was difficult when they simultaneously received a pulse on one device and a pattern on another device.
- **4 Tempo setting** Tempo settings worked better when the pulse was sent on its own to 1 device to establish the tempo for a few bars. This was then stopped, and the patterns were then sent to the 2nd device.
- **5 Strength of signal** Players with devices sometimes found that the signals weren't felt strongly enough when the whole group was playing, due to the natural vibrations produced by their own and the other instruments. Strength of signal needs to be stronger, and individually adjustable.
- **6 Placing of haptic signal** Some preferred to wear the device with the pulse on their weaker hand and the device with the patterns on their dominant hand. Wearing the device around the neck seemed to work well.
- **7 Temp setting and tempo adjustment** The stomp box [for triggering haptic tempo information] appeared to have some latency. The launchpad and bop-pad appear worth exploring.
- **8 Attention to individual difference and alternative strategies** Some players count as the primary method of understanding the patterns. But not everyone does. When someone is deficient in one skill, they make workarounds.

Comments included: "I can't count, so I use physical or visual cues to understand what I'm supposed to do in the following ways:

- On the surdo I'll spin my beaters in my hands and the rotation will denote when I'm supposed to hit the drum.
- On the surdo I'll exaggerate the movement of the beater either straight forward or up and around in a circle, and that will denote the time I'm supposed to hit the drum.
- On the shaker I watch the low surdo and make sure I punch forwards when they hit their surdo.

Deaf / blind people will undoubtedly also have their own workarounds to compensate for the sense they're deficient in - they may benefit from those more physical methods of timekeeping if the pulse isn't working for them."

Reflections on the workshop

While the potential for the inclusion of d/Deaf people in small scale musical activities using the haptic bracelets was clearly established in the D1.7 validation, the two workshops related to large ensembles helped to identify ways in which the bracelets might be improved to work better in larger scale contexts.

2.3.1.3. Public Workshop: Milton Keynes International Festival



Figure 2.13: Part of the publicity material for the Milton Keynes International Festival

Location: Milton Keynes International Festival 2023

Date: 22 July 2023 **Participants** Around 20 members of the public of all ages

Organisation: organised in collaboration with the Music Computing Lab, the Stables Theatre, and master drummer Raz Jayasuriya.



Figure 2.14: Participants in the Haptic Bracelets Workshop at the Milton Keynes International Festival

Purpose of the workshop

The purpose of the workshop was for outreach and engagement and to showcase to the general public the Haptic Bracelets technology and its applications for integration and inclusion. The workshop was open to the public, free of charge, and was heavily oversubscribed.

Part of the workshop was to demonstrate how people could learn rhythmic patterns, and how they interlocked with other patterns without seeing or hearing them, by sense of touch alone.



Figure 2.15: Master drummer Raz Jayasuriya co-ordinating workshop participants with invisible silent haptic signals

The workshop was generally felt to be highly enjoyable by participants, and after the event several participants stayed to talk about implications of the technology and the wider Polifonia project.

2.3.2. Harmony Space

Following on from the previous validation period, limitations that had previously restricted validation activities to the lab had been lifted by software development and upgraded hardware (see Results). With these new capabilities, it was clear that Harmony Space had the capacity to support collaborative music composition and performance by teams that included people with physical disabilities as first-class citizens.

For this reason, we co-organised and presented at a workshop on Music Interaction and Physical Disability to allow us to demonstrate Harmony Space to potential users, identify new collaborators, and to explore these possibilities further.

2.3.2.1. Harmony Space at the EPSRC CHIME Workshop

Location: St Cecilia's Hall Edinburgh

Date: 19th May 2023

Focus of Workshop: Music Interaction and Physical Disability

Harmony Space was demonstrated at the EPSRC CHIME Workshop on Music Interaction and Physical Disability, with a presentation focusing on how Harmony Space could be used to improve inclusion and integration of those with physical disabilities.

This workshop was a collaboration between the CHIME EPSRC Network on Music and Human Computer Interaction <https://www.chime.ac.uk/> and Drake Music Scotland <https://drakemusicscotland.org> which provides music making opportunities for people with disabilities, and which is the leading national arts organisation in this role.

The overall purpose of the workshop was to explore current research and practice in Music Interaction and Disability, and to provide a means for researchers and practitioners to make new connections. Some twenty contributors were present, including performers and composers with disabilities, established researchers, performers, composers, two SMEs and representatives from charities including the Amber Trust and Drake Music Scotland.

Notable elements of the workshop

- an extended presentation in person by Sound and Music New Voices 2022 Scottish composer & performer Chris Jacquin about his work and the challenges of making music with cerebral palsy, during which Chris introduced the video premier of his performance of Stefan Niculescu's "Echoes II" with violinist, Gordon Bragg.
- a demonstration by SME Digit Music of their CMPSR gaming-style music-making instrument, whose joy-stick-driven hardware is loosely based on the Allied Motion Shark 2 wheelchair controller.

Lesson from the workshop

- Digit Music's joy-stick-driven hardware, which is 100% compatible in its affordances with the device that many wheelchair users employ to navigate their wheelchairs, led us to explore adapting Digit Music's hardware for navigating chord sequences in Harmony Space in a way that would be well suited to people with physical disabilities.
- From presentations by researchers at the conference who had been working with the One-handed Musical Instruments Trust, it became evident that by interfacing Harmony Space to a hexaphonic midi guitar transducer (such as a Fishman Tripleplay interface, or Cycfi Research's Nu Multi 6) and a dance mat, Harmony Space could potentially become an expressive polyphonic instrument playable by one-handed musicians.
- From Chris Jaquin's demonstration of his tools for composition and performance, and extended discussions with him, we gained insights into constraints, affordances, and tools for severely physically disabled musicians, and considered how we might adapt the interfaces of Harmony Space to take these into account.
- After conversations with The Director of Drake Music, Pete Sparkes, we identified the potential for Harmony Space to contribute to Drake Music's Music Technology programme for school leavers with disabilities.

- Some of the most valuable insights came from follow-up discussions with Chris Jaquin and Pete Sparkes in the weeks following the workshops. One such insight stemmed from the observation that people with severe physical disabilities can be precluded from playing polyphonic instruments, blocking an important source experience in the precise live physical control of harmony. Both Chris and Pete felt there is serious potential for Harmony space to offer a way of gaining this experience. A second such insight, from Chris Jaquin, was his view that in practical terms, if we could make Harmony Space fully mouse-accessible, that would give it sufficient accessibility for him, and for a wide range of people with severe physical disabilities. This gave us a concrete target to aim for, that we were able to substantially meet in the quarter following this workshop.

This last insight suggested strong possibilities for contributions to Drake Music's Music Technology programme for school leavers with disabilities (See Results section).

Outcomes

- We began to investigate interfacing CMPSR with Harmony Space.
- We started systematically reviewing the existing Harmony Space User interface to improve accessibility by those with physical disabilities.
- We began prototyping a Fishman Tripleplay app for Harmony Space.
- We organised visits to work with Chris Jaquin and Drake Music in April 2024 to follow up the work outlined above. Unfortunately, due to illness, these visits had to be delayed, but see the Results section for the follow-on project that ACCESS has already begun in Partnership with Drake Music to pursue this work.

After the workshop, the next priority was to prepare for a three-day visit to the school I. C. Eleonora Pimentel Fonseca in Italy at the invitation of the deputy head mistress, to see how Harmony Space might best be integrated with Music education practices at the school. In the following sections we spell out some of the activities in detail, where we believe this helps to illuminate the strengths and limitations of this multi-dimensional formative evaluation, and where it can act as a resource for its continued evolution and development.

2.3.2.2. Workshop at IC FONSECA School, Italy

- School: I. C. Eleonora Pimentel Fonseca
- Location: Pontecagnano Faiano (near Salerno, southern Italy)
- Date: Three-day workshop with Children: 16th – 18th October 2023
- Age group of children: 11-12 years old
- Facilitators: Simon Holland and Nicholas Canny
- Translator: Carmina Somma
- Class music Teacher: Luigi Calabrese

Purpose of visit

- Investigate how Harmony Space could be integrated with music education practices at the school.
- Investigate new ways in which those with physical disabilities might be included in musical composition and performance.

Music at the school

Most children starting at the school have little or no formal knowledge of music composition, music performance, or music theory. A few children are learning to play instruments, such as piano, violin, guitar, and flute. The school uses the free and open-source software MuseScore in class with a touch-based wall-screen to enable the writing of melodies and accompaniments by children.

Specialised hardware and software

Specifically for the purpose of this visit we reimplemented Song Walker (a variant of Harmony Space, whose purpose and operation are outlined below). Song Walker is a cross-platform, whole-body, user interface for Harmony Space that works with one or more x-box controllers and a set of dance mats. This system with associated hardware was donated to the school at the end of the visit.

Context for the workshop

- The workshop involved 12 - 14 school children aged from 11 to 12 (one child spoke some English).
- Two facilitators, Simon Holland and Nicholas Canny (non-Italian speakers).
- A teacher was available some of the time, who was not musically trained, but could translate between English and Italian.
- The children's music teacher (who does not speak English) was present for observation (and occasional encouragement) all of the time.
- Two or three of the children had some skills playing a musical instrument (guitar and flute).
- We split the children split into four teams of three, each with a laptop running Harmony Space.
- Where possible, each team contained someone who played an instrument.
- Organising in teams in this manner worked well for playing Harmony Space collaboratively, with roles within each teams rotating.

Equipment

- Harmony Space running on a large wall screen controlled by a PC laptop.
- Three dance mats (attached to the same laptop).
- Two x-box controllers (attached to the same laptop).
- Various acoustic musical instruments (on day three).
- Small midi keyboard connectable to the laptop.
- The ability to display any team's laptop on the wall screen.

Principles of the sessions

- As few words as possible.
- Absorbing and fun, if possible.
- Involve everyone all the time - keep an eye open – don't let anyone get bored.
- Some achievement or useful tool from each activity.
- Simplest engaging activity that achieves productive end.
- Orientation: you (the children) can't do anything wrong - help us find our design mistakes.
- Each task is to be demonstrated to children as a purely physical task, with minimal theoretical interpretation, without technical vocabulary.

Despite best efforts, in the heat of the moment, these principles were far from always adhered to over the three days, but the workshop went better when they were.

A note on terminology

As with the rest of this document, technical vocabulary is used in discussions below for the musical expert reader's convenience, but was *not* used with the children. Instead embodied physical equivalents were used where possible.

Differences from previous activities with Harmony Space

We had previously run workshops with Harmony Space in different contexts both for expert musicians and non-musicians, although individuals had always required a lot of technical support.

In these previous sessions, expert musicians had generally understood the ideas quickly and many were able to make connections with deeper musical ideas embodied in the interface. In the case of children and beginners, small teams had typically been taught in a few minutes to play existing songs collaboratively, sounding the correct chords, forming the right inversions, and producing all of the right notes at the right time, albeit as a simple, patterned, collaborative physical, non-theoretical task. This was typically invigorating, fun and engaging, and provided food for thought for some participants about patterns underlying music.

However, in previous activities, there had rarely been time for beginners to explore the materials of harmony in detail, or from diverse points of view, nor to link their physical actions systematically to musical intuitions. With the school, because we had the luxury of a three-day workshop, we wanted to investigate the potential for establishing deeper links.

DAY ONE: WARM UP

The aim of the first activity was to give the children a basic familiarity with the controls and some idea of their effects, and to make sure that all of the laptop installations of Harmony Space were working and could be operated by the children.

ACTIVITY ONE: Exercise ONE

For each of the four teams to play by imitation, using a mouse, five different sequences of notes with the following Harmony Space settings, and to listen to the results produced (**Figure 2.16**). (If we had asked the two instrumentalists to bring their instruments on day one, they would have also been asked to mirror each of these sequences, comment on them, and perhaps identify them).

Harmony Space Settings

- Trace all on
- Piano trace off
- Chord size single note
- Inversion zero

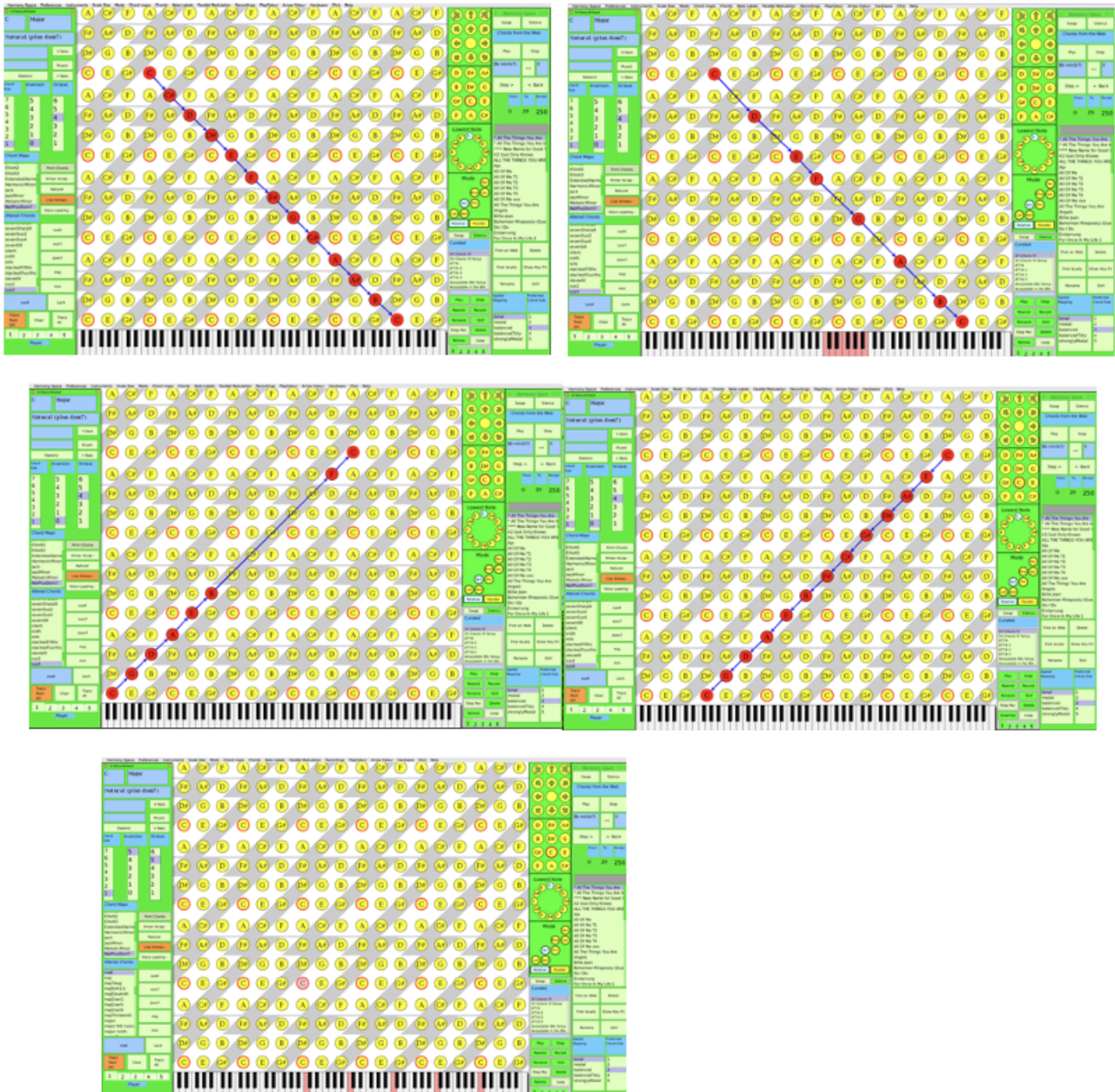


Figure 2.16: Five Harmony Space traces for exercise 1. In technical musical terms, these five figures show respectively: • An ascending chromatic scale • An ascending diatonic scale • A descending diatonic circle of fifths • A descending chromatic circle of fifths • A single note in different octaves. In embodied, physical terms they correspond respectively to the following instructions: • Play circles down the diagonal including circles in the grey area • Play circles down the diagonal leaving out circles in the grey area • Play circles down the reverse diagonal leaving out circles in the grey area • Play circles down the reverse diagonal including circles in the grey area • Use + and – to send any note up or down an octave while watching the shading.

In concrete physical terms, the first four of these paths are straight lines on one or other Harmony Space diagonal, with the grey areas either ignored or not. The fifth sequence is simply a single note played in different octaves using

the + and – keys. In technical musical terms (which we did not elaborate with the class) these correspond to:

- A chromatic scale
- A major scale
- A real cycle of fifths
- A diatonic cycle of fifths
- A single note in different octaves

However the key aim was simply to warm up and make initial engagement with the system.

ACTIVITY ONE: Exercise TWO

In physical, uninterpreted terms, this exercise involved playing a sequence of notes moving stepwise in a diagonal path while ignoring the grey area, but now with different Harmony Space settings (chord size two, instead of chord size one, and trying first with inversion zero then inversion one). The children were asked to compare and contrast how these variously sounded.



Figure 2.17

In technical terms, which we did not elaborate upon, these paths (and associated sounds) corresponded to: a C major scale in diatonic diads, in root inversion, then in first inversion.

ACTIVITY ONE: Exercise THREE

- Same as exercise 2 but in triads.
- Repeated with different inversions.
- Optionally have some teams do this in a different key.

Discussion

- Build a chord in steps from single note to diad to triad in different locations.
- Why are the shapes different in different locations?
- How can we predict the shapes?
- Water metaphor for understanding chord shapes.

In retrospect, for children familiar with basic music theory, this would have been an appropriate warm up exercise, but we had overestimated how much most children knew about music, so that it was not an ideal starting point. However, many useful lessons were learned from this, as discussed later, and as take into account in the later visit to the school, and not too much harm was done. The activities continued after a break for lunch.

ACTIVITY TWO

Playing a chord sequence for an existing song

The teams were shown on the classroom’s shared screen how, in physical terms, to play the chords for the Sam Cooke song ‘Wonderful World’ (also known as ‘Don’t know much about history’).

- Key of C: C major / A minor / F major / G major

Then each team worked together to make sure that each member of the team could play the chords in Harmony Space (In this exercise, Harmony Space is allowed to provide the appropriate diatonic chord qualities given simply the different root positions within the diatonic scale).

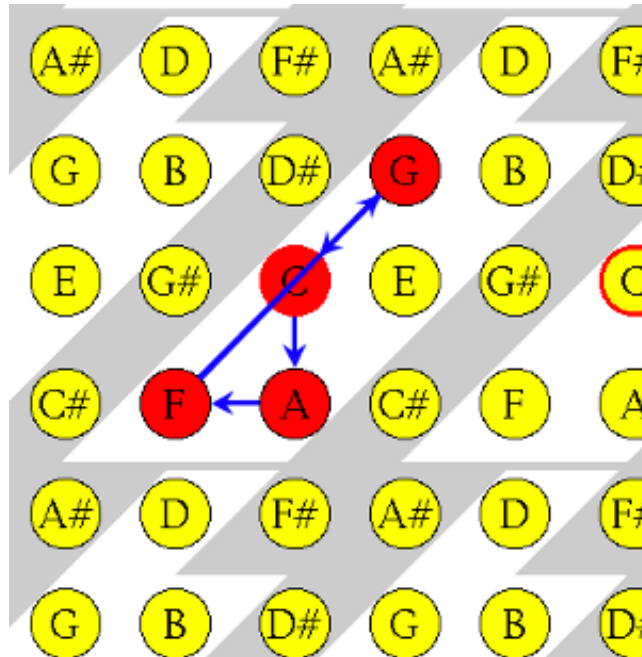


Figure 2.18: Roots of chords for wonderful world.

ACTIVITY THREE

Making up a new chord sequence

The four teams were each asked to make up a chord sequence, however they liked, in a game with the following rules.

- Make a short chord sequence with a beginning, a middle and an end.
- The chord should stay in the white area (avoid the circles in the gray area).
- Use any inversion you like.
- Each chord sequence must be different from the original & that of any other team.

During and after this process, time was spent sharing, discussing and reviewing the different chord sequences composed and played by each group. This activity was generally enjoyed, and on reflection, this would have been a better starting point for the day.

LESSONS learned from day one at the school

Lesson 1: Note notation

The children were not familiar with US/UK alphabetic note labelling A B C D, let alone labels like C# and Bb. Their teacher outlined the fixed do sol-fa system for the ascending and descending chromatic scale respectively, as follows:

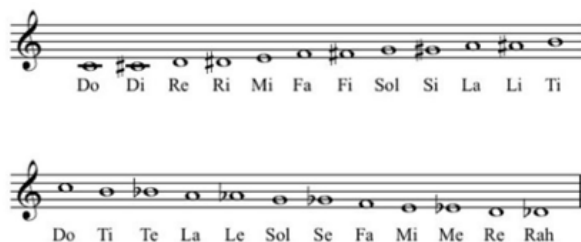


Figure 2.19: The fixed doh labelling used in Italy for pitch names (ascending and descending respectively).

We prided ourselves on being ready for this kind of eventuality, as Harmony Space already incorporates sol fa labelling. However, at the time, Harmony Space was limited to the UK solfa system with moveable doh (i.e. doh is always the tonic of the current key) – whereas the system the school favours is fixed doh (i.e. doh is C regardless of the current key). We have subsequently implemented the Italian notation in Harmony Space, but it was not available at the time.

Lesson 2: Localisation

Harmony Space has been developed up until now on the Mac. Before coming to Italy, we had taken pains to find midi drivers for our virtual machine on the Windows PCs that the school uses, and to ensure that Harmony Space would display correctly on the screen sizes used at the school (for technical reasons, the diverse screen tools comprising Harmony Space cannot scale in a uniform way – so that care is needed to ensure suitable visual layout across range of screen resolutions). This all worked as planned. However, for the purpose of rapidly playing or comparing chords while changing the number of notes or the inversion, keyboard shortcuts are very helpful. We had overlooked the fact that keyboard layout on an Italian PC is different from the UK version, and that spoiled the logical spatial progression of the keyboard shortcuts. We were able to fix the mapping on a single machine, but we had no mechanism to quickly propagate the change to all of the children’s laptops on the day. We have subsequently implemented comprehensive fixes to this class of problem (See Results section).

Lesson 3

We had worried that there would be too much noise leaking between teams, and that headphones would be needed, but in fact it was easy to turn volume levels down, so that this was not a problem in practice. In fact, the problem was the reverse - volume levels were too low, so Bluetooth speakers were found for the next day.

Lesson 4

Part of the original lesson plan had been for each team both to use Harmony Space to print out their chord sequence in conventional chord notation, and to record it for spatial playback using the step record function. Unfortunately this was not always remembered, but we took screenshots of each composition.

DAY TWO of Workshop at Fonseca

Plenary review of each team’s compositions

Each team was asked to perform their composition from the day before in front of the whole class on the big screen. All of the pieces were seen to be different, and all were found to have musically interesting features. Some of these features are noted below.

Chord sequence from Team 1

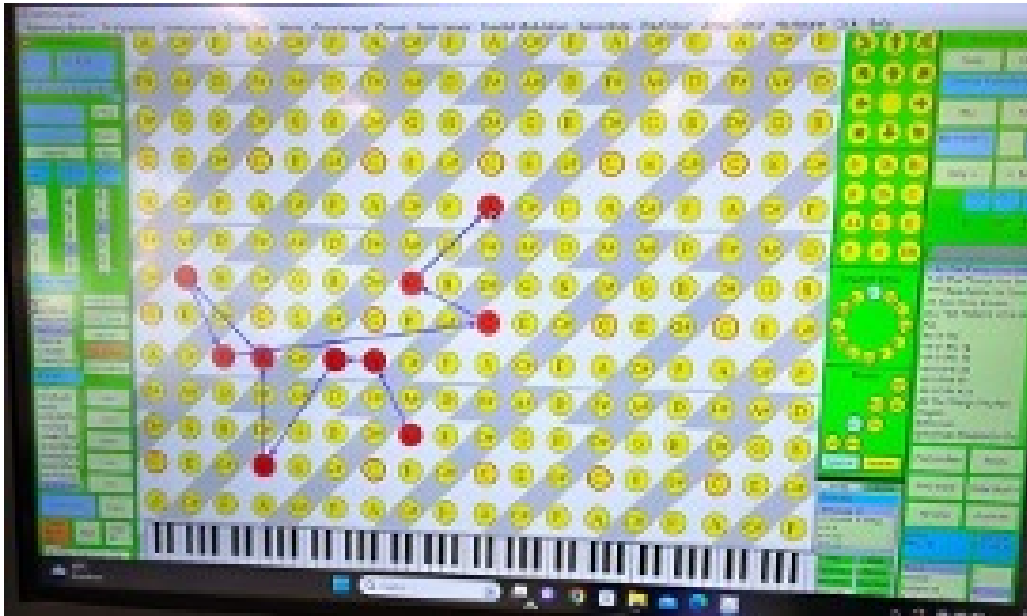


Figure 2.20: (C Major) G major A minor F major C major A minor G major F major C major G major A minor.

This chord sequence is unusual, full of interest, and rich for musical exploitation. Analytically, it can be redrawn to highlight the interesting mix of tonal and modal movement as follows:

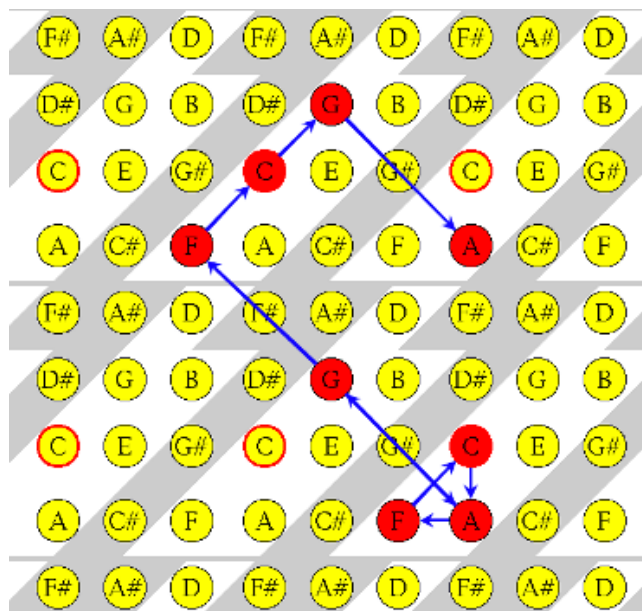


Figure 2.21: Team1's chord sequence, redrawn to reveal the balance of tonal, Aeolian and Mixolydian movement. See figure 2.24 for a remarkable point of comparison.

Chord sequence by Team 2

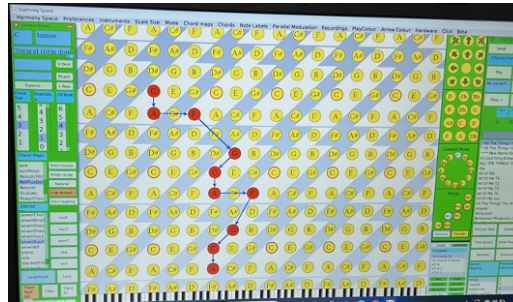


Figure 2.22: Teams 2’s striking chord sequence

Teams 2’s chord sequence was interesting in several ways. The chord sequence exactly follows that of Wonderful World. However, the team were not imitating it visually – the visual path they took is completely different. This created a useful opportunity to draw attention to the property of the representation that allows this, and which can have valuable interpretative properties (See ‘Group 3’s novel chord sequence’ from the second school visit for more details). As played by the team to the rest of the class, the relentless repetition of the same chords created a striking musical effect.

Chord sequence by Team 3

Team 3’s chord sequence was different again. Unfortunately a screen shot of the piece as performed on the day has not survived, but on the day, a facilitator demonstrated how to reorganise the spatial structure of any sequence (via ‘nearest-neighbour’ mapping)² while preserving exactly every note to reveal an underlying structure. A screen shot of this redrawing has survived.

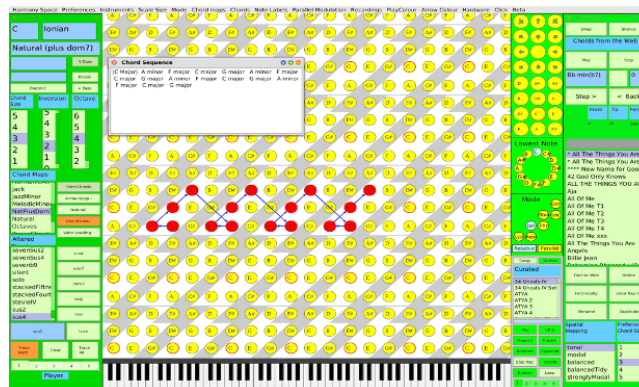


Figure 2.23: Team 3’s chord sequence (redrawn) A minor /F major/ C major/ G major (repeat x4)

This chord sequence is interesting. Although it uses the same chords as Wonderful World, their order is different, and this gives it a completely different musical effect. To highlight just one difference, it may be viewed as having a strong Mixolydian cadence

- (G Mixolydian: A minor/ F major /C major /G major)

quite different in effect from Wonderful World’s familiar four chord trick

- (C major: C major / A minor/ F major / G major).

²See the section ‘Group 3’s novel chord sequence’ from the second school visit for the basis of this mapping)

DAY TWO: Activity TWO

Relating each team’s compositions to existing songs

We did not discuss these features technically, but we did point out comparisons with existing songs that the children might know.

As already noted, the chord sequences created by the children all had considerable musical interest. For example, Team One’s chord sequence, after a seemingly conventional major minor beginning went into an Aeolian progression followed by a Mixolydian cadence. This is not to make any claim about the children’s’ conscious musical intentions or musical knowledge, but the sequence was their original creation, and it does relate interestingly to several celebrated songs.

We looked for related songs that the children might know. In terms of Mixolydian cadences, candidate celebrated songs include David Bowie’s Suffragette City and the Beatles’ Hey Jude, but searching for a more recent example that more of the children might know, we considered Good 4 U by Olivia Rodrigo.

In this song, the Aeolian progression has an air of being miserable or aimless (as noted by Bjornberg, 1985) whereas the Mixolydian cadence often embodies a switch to a feeling of empowerment and exultation (as exemplified by the chorus of Suffragette City and the final chorus of Hey Jude)

This was useful to illustrate how changes of movement in harmony can be use compositionally for channelling shifts of emotion.

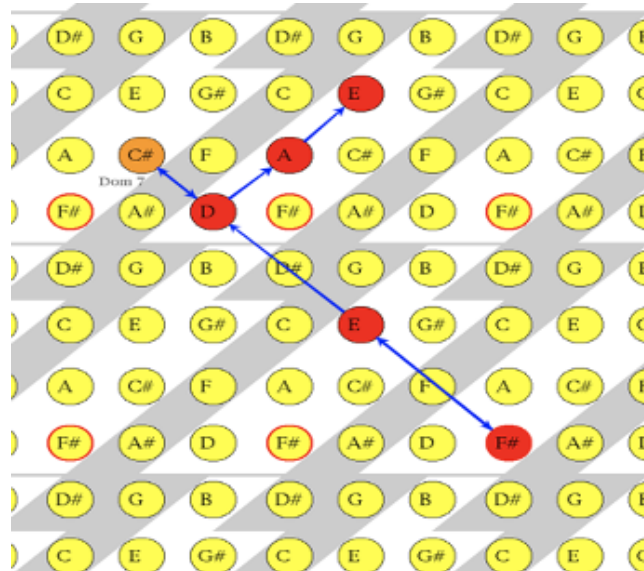


Figure 2.24: Chord Sequence of ‘Good 4 U’ by Olivia Rodrigo (a strong and striking chord sequence retroactively acknowledged as borrowed from The Paramore’s ‘Misery Business’).

This chord sequence begins with an Aeolian modal harmonic ostinato, with sharp emotional twists to a mixolydian cadence as follows:

- (F# Aeolian/Harmonic Minor)
- F# minor E major F# minor E major F# minor E major D major
- C# Dom 7 (*Flip from Aeolian to minor*)
- D major A major E major (*Mixolydian cadence*)

Although it was not appropriate to focus on technical detail, comparisons of features of songs that children have made with features of celebrated songs can be a valuable resource if done sensitively.



Figure 2.25: The children working in small groups on Harmony Space.

DAY TWO: Activity THREE
Spatially identifying the key of a piece

We played a video to the class of 'Fresh' by Kool and the Gang on YouTube, and gave them all a spatially and audibly replayable Harmony Space recording of the chord sequence. For the specific purpose of a key finding exercise, all of the chords were specified precisely and individually rather than being free to flex generatively and diatonically as the key changed, as is usual in native recordings in Harmony Space.³

- Kool and the Gang *Fresh*: B minor 7th / G Ma7 / F# minor/ (repeated)

We asked the teams to do two things on their laptop, and then show to the rest of the class on the wall screen, when they had a candidate solution.

- Task 1: Display the path of the *roots* of the chords only, and experimentally move the diatonic key window to find the best fit (i.e. the likely key).
- Task 2: Display the path of *all of* the notes in each chord to see if this allowed better informed visual disambiguation.

Two candidate solutions found by the children for task 1) were B Aeolian and F# Aeolian.

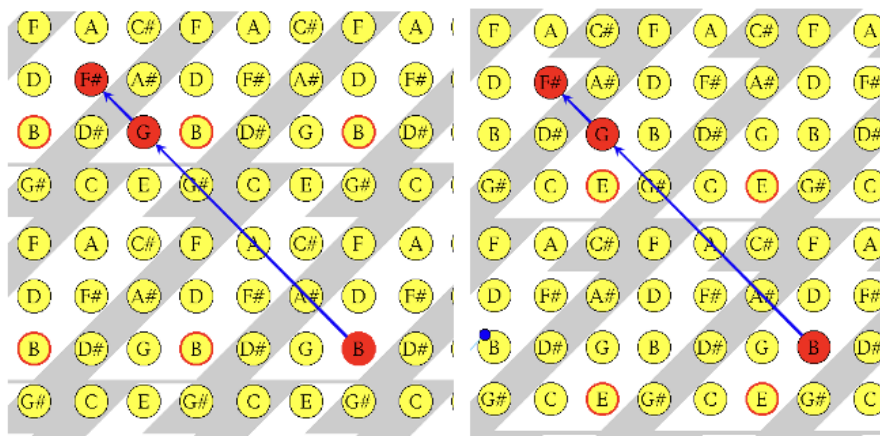


Figure 2.26: Two different solutions by visual fit for identifying the key of 'Fresh' by Kool and the Gang, using chord roots only: (a) B Aeolian (b) E Aeolian.

³A noted in bullet 3 of 2.2.1.

However, task 2 with the full chords admits only one solution, B Aeolian (NB not B minor).

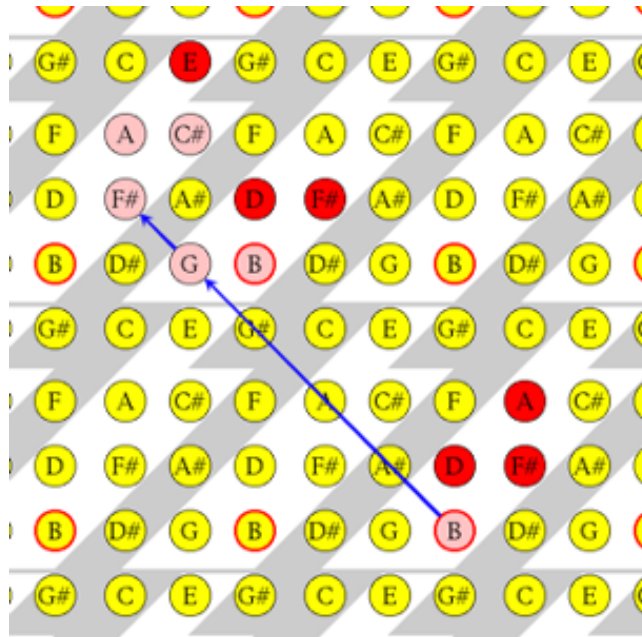


Figure 2.27: Trial visual fit of *full chords* (in sevenths) of ‘Fresh’ by Kool and the Gang, finding the only mode that fits - B Aeolian.

Only some of the children could solve this confidently in the very short time available to spend on it. We realised on trying this activity that we had needlessly complicated things, as we had not previously talked about modal centres, let alone the major and minor tonal centres. We should have kept things simpler - though with examples from current popular music, it can be hard to stick exclusively with major/minor examples, as modal movement and modal centres are so prevalent. Still, the embodied visual nature of this task, and the perceptual and cognitive strength of the containment metaphor suggested that future versions of this exercise have considerable promise for helping beginners to distinguish between different chords and different keys - a distinction that appears to escape many amateur musicians. However, this experience suggested that pedagogical reflection is needed to balance and order examples that introduce modes and keys.

Day THREE of Workshop at Fonseca
Collaboratively performing the children’s composition

For simplicity, we decided to focus with the whole class on just one of the new chord sequences composed by the four teams, picking the example from Team 3 that was short, interesting, and novel. We dubbed this the ‘Salerno Shuttle’.

DAY THREE: Activity ONE

Each team learned to play the Salerno shuttle on their laptops, with agreed slight harmonic elaborations. The elaborations were:

- that each chord was to be played twice, first in second inversion, then in first inversion,
- that the first of the final G major chords was to be played as an altered chord, namely a suspended fourth.

The Salerno Shuttle (composed by Team 3)

A minor (c) / A minor (b) / F major (c) / F major (b) / C major (c) / C major (b)
/ G sus4 / G (b) / REPEAT (lower case letters indicate inversions)

DAY THREE: Activity TWO

When each team could play the Salerno shuttle on the laptop, then they were asked to play it collaboratively as a team using Song Walker in front of class, with the wall display facing the class so that everyone could track the performance. To play the piece on Song Walker, three different roles were involved, which were rotated amongst the class. One member would navigate and fire (i.e. play) the chord root using the x-box controller; the second member would use a dance mat while shifting their weight from one foot to the other rhythmically to change the chord inversions; and the third member would select the desired altered chord at the desired time.

By default, chord qualities in Harmony Space are influenced by the visible constraint of positioning of the diatonic key window, and the thirds stacking ‘water metaphor’ – but can be overridden at any time with altered chords or altered chord maps.

Correct synchronisation threw up some interesting challenges. By design, the controllers for inversions and altered chords do not sound notes, rather they alter the quality of whatever chord is to be triggered next (in the present case by the x-box controller). Thus, these alterations must be triggered *before* the next chord is due.

To make this work smoothly, we first had the whole class clap and count together to establish a beat “uno, due, tre, quattro”; and then we had some of the class shift their weight from one foot to the other on the offbeat “uno, E due, E tre, E quattro”.



Figure 2.28: Counting and clapping to practice precise timing.

This then facilitated precisely synchronised performances, with the x-box controller navigator giving a count-in and leading the count. Generally dance mat players liked to look at the screen while working to see the piece unfolding visually. Teams rotated, and roles within teams were rotated.



Figure 2.29: Rehearsing precisely timed weight-shifting for dance-mat moves.

Day THREE: Activity THREE

On the third day two instrument players (a flautist and a guitarist) had brought their instruments, so we asked them to join in with the performances. Neither were sure what to play, and the guitarist had not learned to play chords yet, so we suggested that he play the root notes. With the flautist we assured her that nothing she played could be wrong. This proved to be true in practice. The overall effect of the ensemble was widely agreed to be affecting and impressive.



Figure 2.30: Using dance mat moves to control chord inversions and altered chords, these changes interlocking with the x-box controller navigating overall path (x-box navigator in foreground).



Figure 2.31: The navigator (on the right) uses an x-box controller to navigate harmonic paths, and also counts out loud to synchronise the mat crew, flute and guitar.

Lessons from first 3-Day workshop at school

As a direct result of the workshop and feedback from the children and teachers we learned many lessons about how to teach with Harmony Space and implemented many fixes and new features for Harmony Space.

Comments from the children can be found in Appendix 1.

Fixes and new features following from the visit.

- We created a new keyboard short cuts editor for end-users to localise Harmony Space to the different keyboard layouts of different countries and pre-created an Italian localisation.
- We changed the file format for saving work so that it works cross platform across Mac and PC.
- We created a new simpler way to record compositions.
- We created a new simpler way to edit single notes in compositions so that if you “made a mistake in your composition you [don’t] have to delete everything and start playing again”.
- We implemented a graphic editor for ChoCo chord sequences so that end users could make and return suggestions. This also simplifies the key fitting exercise on Day two.
- We added a lot of new how-to information to the Harmony Space website.
- We added fixed-doh note labelling as an option.
- We implemented a chord explorer for exploring the elements of altered chords in great details (as opposed to exploring how chords move).
- Following the suggestions of the music teacher, we have implemented facilities for spatial mapping of melodic Midi input.
- We implemented several helper programs for exploring the interaction of harmony and rhythm.

Revisiting the school.

After revising our teaching materials and software and refining our pedagogical approach in the light of feedback from the first visit, we revisited the same school again five months later.

Whereas previously we had been dealing with the oldest class in the primary school, on the second visit we were dealing with the first and second year of the secondary school.

2.3.2.3. SECOND visit to IC Fonseca School 25th to 27th March 2024

Introduction

The aim of the first day was to take a different collection of secondary students, with no previous exposure to Harmony Space, and to cover more material, more clearly, more economically and more systematically than in the first visit. An additional aim was to gather insight for future wider pedagogical extension. In summary, the plan for the first day was to:

- familiarise the children rapidly with the basics of controlling the interface to make music
- have the class collectively play a Stevie Wonder song as a laptop orchestra (some playing a bass line, some inverted triads, some fully voiced extended chords),
- learn to play a different and more harmonically malleable song (Wonderful world by Sam Cooke) and then, split into groups each to modify this chord sequence to create different new songs,
- to play at least one of the new songs using x-box controller, dance mats and wall screen, with accompaniment from some of the class who played acoustic instruments

The following account goes into low level details in the form of a lesson plan with the following sections, corresponding to the above goals (with an added preparation step).

- preparation
- class warm up
- laptop orchestra
- song writing by modification
- strongly embodied performance

Target users

This summarises a lesson plan designed for children in the second year of secondary school (aged 13) without any previous exposure to Harmony Space. This plan is aimed at a mixed class mostly of children with little or no musical training, interspersed with a few children who are taking instrument lessons, who are invited to bring their instruments.

Preparation: set up for the approximately two-and-a-half-hour session

Twelve children were arranged in four groups of three. Each group shared a laptop on which Harmony Space had been pre-installed. Ideally children should be able to share headphones when working as a group of three, but this was not practical - however each group should at least have a speaker when playing back results to the rest of the class (in practice two groups had a speaker).

A facilitator displayed Harmony Space running on a class machine on the wall screen where needed during each exercise (Generally, it is helpful if students' machines can be switched to the wall screen, but this is not essential and was not done during the workshop).

Localisation In the case of Italian school children using PCs, it is important that the keyboard shortcuts are localised to fit the keyboard layout of an Italian PC (as contrasted with, for example, with a UK PC or Mac – see **Figure 2.32**). This allows the changing or chord size and inversion and octave to be carried out with muscle-memory-friendly mnemonic keyboard shortcuts.

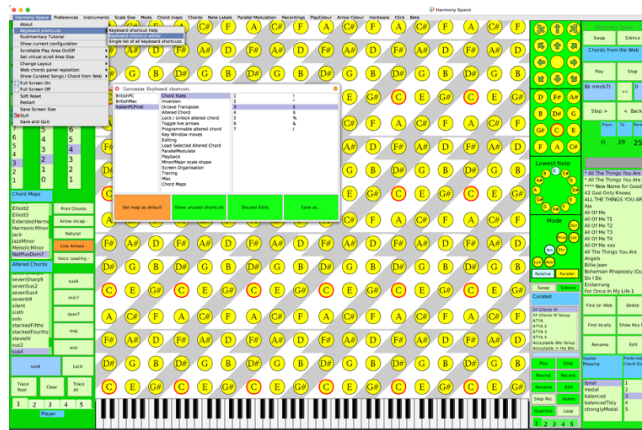


Figure 2.32: Localising the keyboard short cuts to fit an Italian PC keyboard.

Note labelling

For this initial exercise, the note labelling was initially set to Italian sol fa (**Figure 2.33** which, as previously explained, uses a fixed-doh, as opposed to moveable-doh system). For this specific exercise, the diatonic *notes only* in the key of C were labelled (with chromatic notes unlabelled) to focus initially on the diatonic (here major) scale, to avoid unneeded distraction.

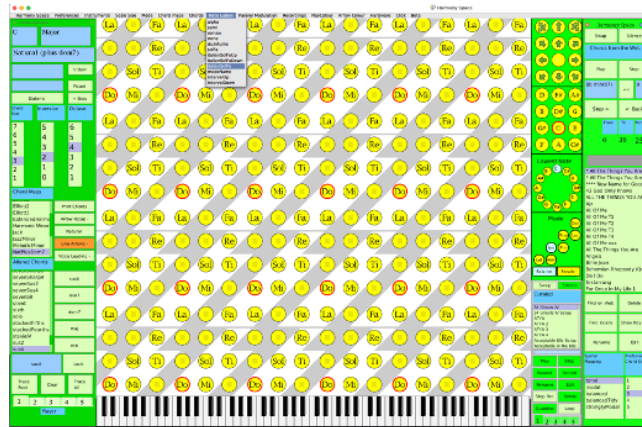


Figure 2.33: Switching note labelling to Italian sol fa, using a fixed vs. moveable doh.

Reflection on labelling

On reflection, in retrospect, a better decision might have been to start with no labelling at all. This is because if you ask students to play the diatonic (here major scale) 'do ray me fa so la te doh', some of them will translate this simply to finding and pressing the correspondingly labelled circles. By contrast, with no labels (**Figure 2.34**) the task can be reframed as a single visuospatial chunk: playing notes along a diagonal (**figure 2.34**) while omitting the those in the grey areas. This would foreground the notion of the diagonal as a salient entity, and the white and grey areas as important features of the interface, with the result 'do ray me fa so la te doh' being identified from listening. This framing also has the advantage that it is immediately independent of key. This approach is something to explore at a future date.

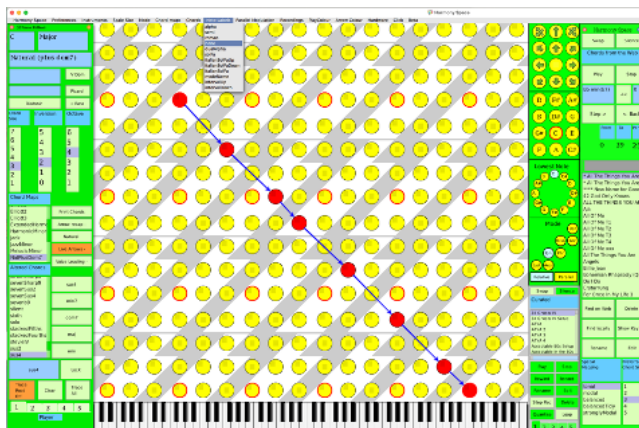


Figure 2.34: The major scale unlabelled, played as a diagonal line omitting the notes in the grey area.

Contesting Italian Sol fa

One of the few Italian students who spoke English contested the labelling:

- ‘do ray me fa so la te do’

objecting that it should be:

- ‘do ray me fa so la si doh’.

However, the music teacher had previously indicated to us that the ascending *chromatic* scale in Italian solfa (which we would show the children later) goes:

- ‘do di ray ri me fa fi so si la te doh’.

This suggests that ‘si’ should be used for B as it is already reserved for G#. Unfortunately we did not have time to further investigate this point.

Class warm-up

The purpose of the warm-up was for students to learn the operation of the basic controls of Harmony Space, and to tacitly absorb something of their musical effects - while also learning how to play a simple musical exercise in synchrony. The first step was to make sure each of the four groups had the correct initial settings (in part by referring to the wall screen projection for examples of the correct settings as needed).

Initial Harmony Space setting: Key of C, Chord size 1, inversion zero, octave 4, trace root quality lock off.

Reflection on initial setup

Students were invited to carry out this setup manually (as opposed to by pushing a single button) to open the possibility of absorbing the effects of the controls.

Part of the aim here was to avoid verbal explanation and instead to privilege learning by doing. But, at the same time we wanted to lessen the scope for fruitless error. For this reason, to allow quick recovery in case of finger trouble, in future we will create in-built song-like recordings whose sole function is to prepare the settings for different tasks.⁴ But on this occasion, we preferred to let the students do this manually.

⁴Such recordings purely for the purpose of adjusting settings cannot be located in the ChoCo compatible library shown on the right-hand-side of figures 34, 35 and 36, as these contain only chords and key metadata. These more powerful “songs” (which can include sequences of setting) must instead be located in the reflective library shown in figure 4 whose recordings can reify any and all Harmony Space settings as well as chords, keys, and spatial paths. With large screen sizes (such as retina screens), both of these library panels are always visible, but on smaller screen where there is only room for one; they can be swapped using the swap button at the top of the RHS control panel.

Exercise 1: Play the diatonic scale

- Students were invited to play a diatonic scale using the trackpad (or a mouse) in the 2D Harmony Space area like so: Do re me fa so la te do. This was also displayed on the wall screen by the facilitator (**Figure 2.34**)
- Once all four groups could do this, a student was designated to count everybody in so that the whole class could play the scale in unison together.
- Depending on the skill of the class, the groups could have been invited to each play in unison in a different octave (to aid exploration of the octave dimension).

Reflection

As previously, it was pointed out to the class that the path is diagonal, avoiding the grey area.

NB Where students’ laptop screens are too small to permit performing the scale as a single straight line (which is the case in all of the student laptops) an alternative needs to be suggested (**Figure 2.35**). Arguably this doubling back path has the merit of drawing attention to the fact that all notes with the same label on the main screen with the same shading have the same pitch.

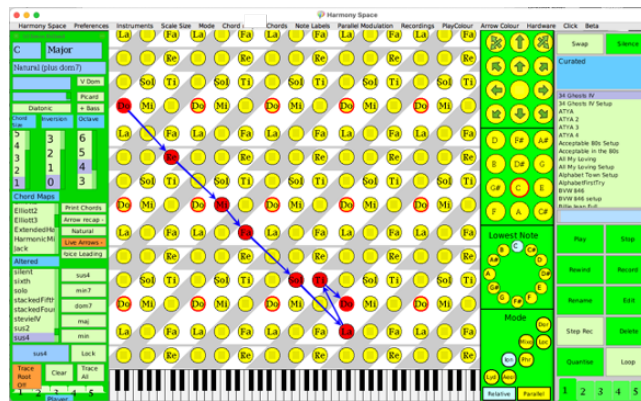


Figure 2.35: Fitting the diatonic diagonal on a smaller screen.

Additional instructions for exercise 1

- Use left click to play notes and alter settings
- If playing the scale again, and if the trace arrows get in the way, use the shortcut C to clear the screen.

Reflection on exercise 1

In the event, this went fairly quickly and smoothly, and we moved fairly promptly on to Exercise 2.

Exercise 2: Play the Chromatic scale

The students were asked to change the note label setting to the labelling Italian sol fa chromatic ascending (Figure 2.36)

- Students were invited to play a chromatic scale using the trackpad (or a mouse) in the 2D Harmony Space area like so: Do di re ri me fa fi so si la li te do. This path was also displayed on the wall screen by the facilitator.
- Once all four groups could do this, a student was again designated to count everybody in, and for the class to play the scale together.
- Again, if the class had picked this up very quickly, the groups could have been invited to each play in unison in a different octave.

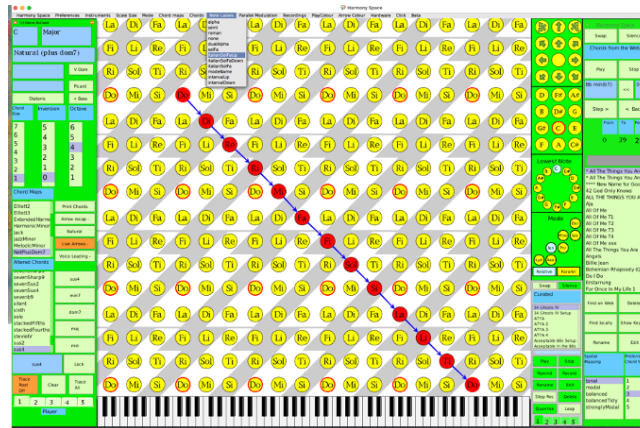


Figure 2.36: Playing a chromatic scale – the path is again diagonal, but now encompasses notes in both the white & grey areas – not just the white area.

Reflections on the two warm up exercises

It was pointed out that both paths were diagonal, with the second spanning both the white & grey areas – not just the white area.

The correspondence between what was played in Harmony Space and the lighting up of notes on the graphic keyboard at the bottom of the screen was also pointed out. A few children were already aware of the UK/US note naming, so this was also briefly mentioned and demonstrated (**Figure 2.37**). The aim of the warm up exercise, to a greater or lesser extent, was to pilot and approximate a lesson in which the children learn (from a sensory-motor point of view, rather than using technical musical terms) how to play diatonic and chromatic chord sequences, change chord size, change inversion, change octave, work with the root trace on or off and to learn when and why to use the clear screen command. This was achieved in part with some, but not all children. An additional aim was to gain a geometrical (as opposed to note-label-based) insight on where to find and play the diatonic and chromatic scales, and to relate this to the piano keyboard and to familiar note names. Subsidiary aims were to understand that all notes with the same labels and the same shading have the same pitches, and how shading corresponds to octave height.

The idea was not to articulate these concepts orally, but to have them to be inferable from what happened in simple tasks.

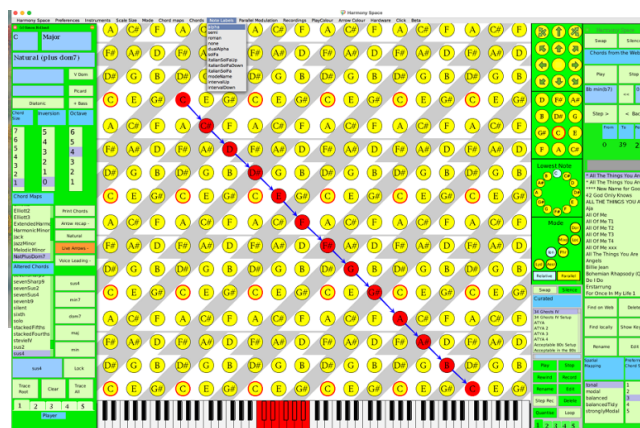


Figure 2.37: Showing English/US note naming, with the chromatic scale matched by an automatic trace on the piano keyboard.

Exercise 2: Massed jazz laptop orchestra

The idea for this exercise was to have the class collectively play a Stevie Wonder song (Isn't She Lovely) as a laptop orchestra (with some children playing a bass line, some inverted triads, some fully-voiced and properly voice-led extended chords). Part of the reason for choosing this song was to demonstrate tacitly that the opposite diagonal to the diatonic/chromatic diagonal is the circle of fifths axis (which the verse and middle eight -for want of a better term - of this song respectively traverse).

Crucially, each of the four groups were asked to make different Harmony Space settings.

All four groups were asked to play the same root progression (demonstrated as a root progression on the wall screen).

- La re so do
- A D G C

It was mentioned without any music theoretic comments that this was geometrically the opposite diagonal from the previous exercise.

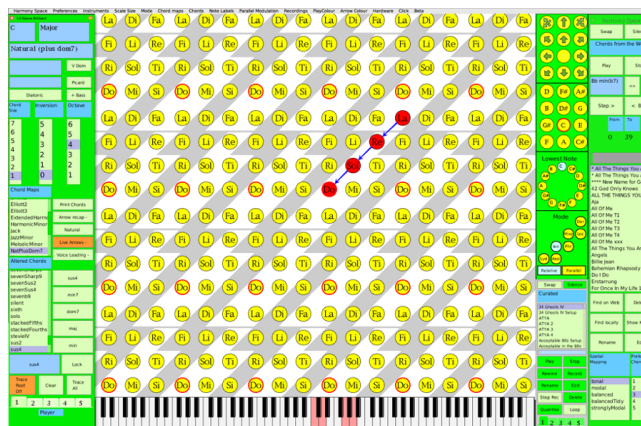


Figure 2.38: Root progression of the verse of “Isn’t She Lovely” by Stevie Wonder. The trace on the graphic piano keyboard also shows the four root notes played.

Group 1 Role

Settings: Chord size 1, octave 3, inversion zero, trace root, live arrows on.

Musical Function: This group served as the 'bass player' for the class for the chord sequence.

Group 2 Role

Settings: Octave 4, chord size diads, inversion zero, no trace live arrows on.

Musical function Providing harmonic support.

Pedagogic function: The pedagogic function here was to explore diads, how they sound and how they flex visually and aurally as they move (Figure 2.39); experiencing how the varied sound of the inversions is mirrored by the varied three-dimensional shading, allowing students to hear the effects of inversion and make aesthetic choices.

Discussion with the group. The group was asked to play diads with root progression:

- La re so do
- A D G C

with inversion zero and inversion one, and to go with whichever sounded better to them.

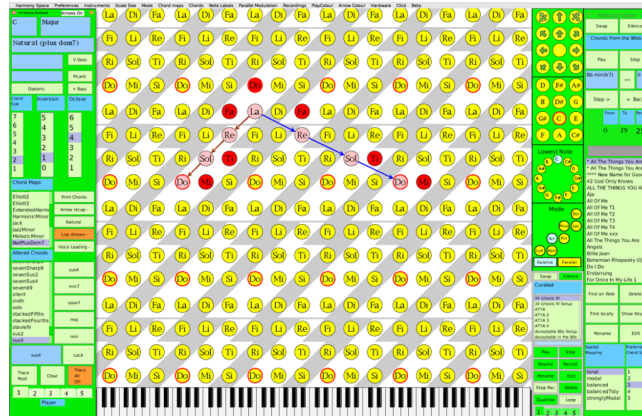


Figure 2.39: Two different ways of plotting the verse of “Isn’t She Lovely” by Stevie Wonder played in inverted diads (the choice of first inversion here is indicated by the lighter shading of root compared with the other note in each diad). Only one of these paths (red arrows) would likely be used in real life, the other (blue arrows) is used here on the page pedagogically simply to space out the diads so their elements can be seen more clearly. At the risk of labouring this, the canonical spatial path (here red arrows) is that shown in the left-hand trajectory; canonical paths are readily found as they always minimise nearest-neighbour distance. The right-hand trajectory (here blue arrows) was not used with the students but is used in the diagram simply as it gives more space to show the reader in a static medium how the shapes of the diads flex as they move. In fact with students, the trace was deliberately kept as trace root only, although the elements of each diad could be clearly seen while they played. This kept the resulting final trace simpler

Group 3 Role

This setup was similar to group 2, but with inverted triads instead of diads.

Setting: Octave 4, chord size triads, inversion zero, no trace, live arrows on.

This group was asked to try playing triads with root progression:

- La re so do
- A D G C

with inversions zero one, and two, and to go with whichever sounded better to them.

Group 4 Role

This was similar to group 3, but using a customised chord map which assigns a voiced, extended, possibly altered chord to each degree of the diatonic scale. When moving along the circle of fifths, this chord map is nicely voice led (Figure 2.40). Harmony Space does have a voice leading module, but it is currently buggy, so this particular chord map was hand-crafted.

Setting: Octave 4, chord size 5 (ninths) ‘Stevie’ chord map, inversion zero, no trace, live arrows on.

If there had been time, we would have asked this group to compare this customised chord map with simple thirds stacked sevenths (chord size 4 with inversion 0, 1, 2 or 3) and to go with whichever sounded better to them.

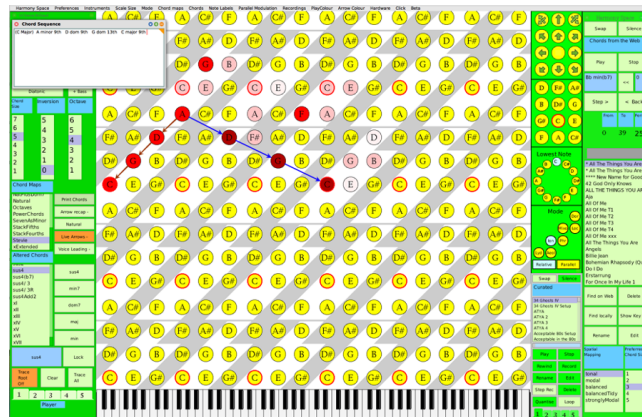


Figure 2.40: As with figure 2.39, two different ways of showing the verse of “Isn’t She Lovely” by Stevie Wonder, this time played using well-voiced and well-voice-led ninths. The canonical spatial path is that shown in the left-hand trajectory (here red arrows). The right-hand trajectory (here blue arrows) was not used with the students, but is used in the diagram simply as it gives more space to show the reader how the shapes of the ninths flex. Note, in an aside that might be mentioned to more advanced students, that the A (Ia) is simple thirds-stacked minor 9th whereas the D (re) is dominantised (as can be seen by its jutting out into the grey area) in accordance with the jazz idiom here used by Stevie Wonder, in which chains of II Vs are used to give additional tonic-directed momentum. Note also, as an aside for experts, that the G (soh) is a full-blown 13th (with the third and 5th conventionally omitted). The voicing used here which can be written as Fma7/G is readily identifiable from the trace (for those familiar with the shapes of sevenths chords in Harmony Space). (C Major) A minor 9th D dom 9th G dom 13th C major 9th.

Playing together

As previously, a member of the class was then deputed to give a count in class as a whole and for all four groups play their diverse voicings simultaneously together. This sounded good and was generally considered fun to do.

Learning the middle eight

The root progression of the middle eight, that follows after playing the verse twice, was then demonstrated by the facilitator on the class wall screen (Figure 2.41).

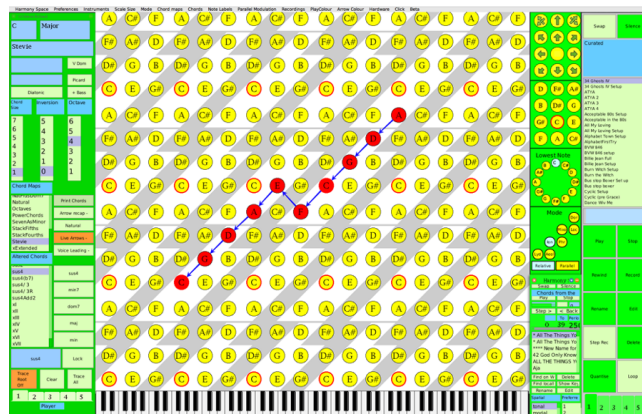


Figure 2.41: The middle eight involves continuing the root progression in descending fifths, past the tonic (C) down to F (IV), then swerving sideways chromatically down to E(III) which has then built up a head of steam to a longer III VI II V I extended cadence.

The class playing the song as a whole

In playing the song as a whole, the verse is repeated twice, then the middle eight once. In playing the middle eight, both the final G and the final C chords are repeated twice. Then the song as a whole can be repeated.

For any chord size, Harmony Space can, if desired, apply inbuilt diatonic chord formation constraints to form the needed chord qualities, simply from playing the root. It is interesting to contrast these purely diatonic qualities with the chord qualities from the Stevie Wonder chord map (Figure 2.42).

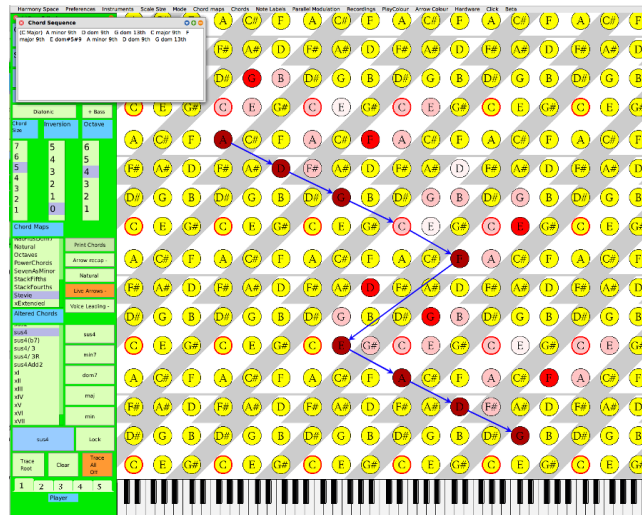


Figure 2.42: Again, this has been plotted using a somewhat perverse non-nearest-neighbour trajectory, to spread out the trace of successive chord elements for the benefit of readers in a static medium. The F (IV chord) is a simple major ninth, whereas the D(II) can be seen to be dominantised (where it juts out from the white into the grey). In music-theoretical terms, this is effected by a sharpened 9th, so that both the major and minor third are present in a close crunch.

Reflection on rhythm and synchronisation

The class was able to learn to play the entire piece together tightly quite quickly (at about an hour into the lesson). At least two interesting things occurred as regards rhythm. Firstly it was necessary to debug the counting process. The deputed student counted out loud ‘uno duo tre quattro’, clicking their fingers and flexing their arm. However, while the physical movements kept good time, the oral count was somehow rushed and untethered. Fortunately, simply pointing out this discrepancy was enough to remedy it.

The second interesting thing as regards rhythm concerned the idea that ‘It’s too late to move the mouse on the one’. The students had Harmony Space set to sound as different MIDI instruments, depending principally on how loud each laptop was (Figure 2.43). The quietest computers tended to be set by the facilitator to, for example, “overdriven guitar” to help them cut through, whereas the louder computers could be set to the softer e.g. “string synth 1” which is relatively pleasant sounding.

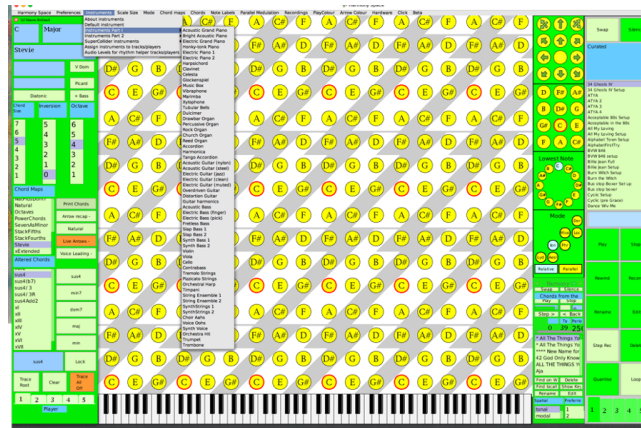


Figure 2.43: Some of the available General MIDI timbres.

However, the amplitude envelope on the string synth sample takes a moment to ramp up, so that it is too late to be moving the mouse towards the needed chord on the instant of the first beat in every bar. Instead, the preceding chord should be quite early so that one can anticipate being in position on the first of the bar before the needed mouse click. Thus did the discussion revisit Bootsy Collins’ advice on the importance hitting the one (though in a slightly different context).

Playing along with acoustic instruments

Some of the class could play acoustic instruments, so at this point the class split into the laptop orchestra and instrumentalists (violin flute & guitar) playing together.

This sounded excellent, but as facilitators we missed a trick, in that we expected the instrumentalists to be able to improvise a melody to the verse. This was utterly unreasonable, and in fact all soloists simply played the root notes (or in the case of one guitarist, the chords). Excitingly the guitarist could play Jingle bells in C, which might have fitted well as a bebop solo. Still, what the facilitators should have done was to have pre-suggested a few melodic four note phrases as solos. Nonetheless, the combination of laptop orchestra and acoustic instruments sounded very well and was absorbing for the entire class.

The above exercises took about an hour, and after a 15-minute break we went on to part 2.

Part 2 Invent a new chord sequence

Although in part 2 we went on to revisit something we had tried five months early with a different group (modify “Wonderful World”), here the class was moving much faster, the context was different, and we tried different rules.

‘Isn’t She Lovely’ has a robust, euphonious chord sequence, but it is hard to modify the harmonic trajectory (as opposed to simply decorating or modifying particular qualities) without destroying its harmonic logic. For part 2, we wanted students to play and modify a chord sequence as the basis for their own new songs. So for this exercise, we turned to a malleable four-chord trick in the major mode, namely “Wonderful World: by Sam Cooke, also known as “Don’t know much about history” (Figure 2.44).

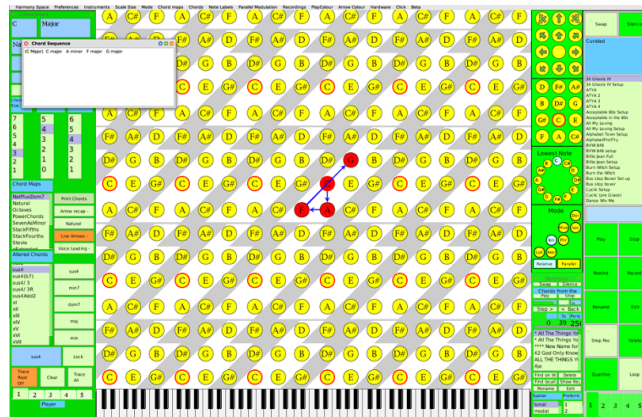


Figure 2.44: “Wonderful World: by Sam Cooke (also known as “Don’t know much about history”).

Part 2

First part of the exercise: play the original four-chord sequence

The first part of the exercise was simply for all four groups to play the chord sequence of this song, as demonstrated on the wall screen.

Setting: Octave 4, chord size triads, inversion 1 or 2, trace root, live arrows on.

Once this was done, a set of rules were given for each group to independently modify the chord sequence.

Initial preparation to help with documenting results: switch on basic live record (not step record).

The rules for modifying the chord sequence.

- Make up a new chord sequence.
- The La chord is not allowed.
- Use 4 chords in some order that sounds OK.
- The chords should be able to loop around.
- Every group should make up a different chord sequence.

Follow up tasks

- When finished save the recording.
- If the recording was not played in time, use the quantise function in the editor, and save a quantised copy.

Note that the disallowing of ‘la’ is an arbitrary rule to ensure the chord sequence is different from the original. Other constraints could do just as well.

The introduction of the editor and quantising was intended to address feedback from students from the first visit - though in the event this did not get well explored.

The four compositions by the four groups of students

Group 1's novel chord sequence

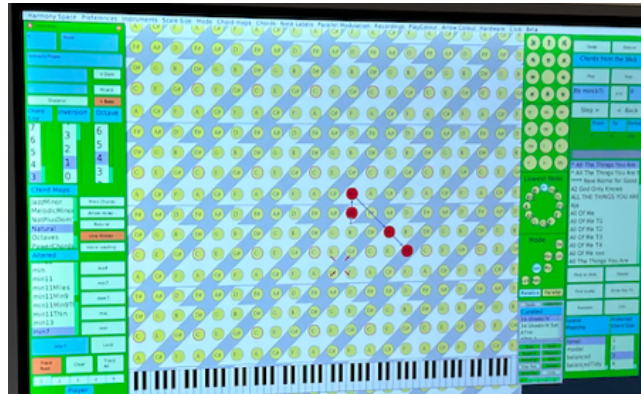


Figure 2.45: The chord sequence that Group 1 composed.

- D minor F major G major G sharp major

This chord sequence is full of strong possibilities.

Due to time pressure, we did not have time to discuss each of the new chord sequences in detail. If we had had time, and with a more advanced group, we might have discussed some of the following points. For the readers convenience we will frame these discussions primarily using standard musical terminology. However, all of these discussions could be translated into embodied physical terms, using Harmony Space.

After the robust D minor F major beginning (these triads visibly have two notes in common), some interesting modal then chromatic movement ensues. Possibly the centre is Dorian (using D minor as the home chord). Because the fourth chord is in the grey area, its quality is undefined. By default in Harmony Space, these undefined qualities are set as major (or dominant, in the case of chord size four), but it is legal and fair to try any altered chord quality in the grey area. Dim or Dim 7 might be useful candidates as non-primary shapes, as they refer back to the D and the F roots – or maybe that might be too four square and conformant – an augmented chord might be another possibility (Figure 2.46) – in any case, the ear should decide.

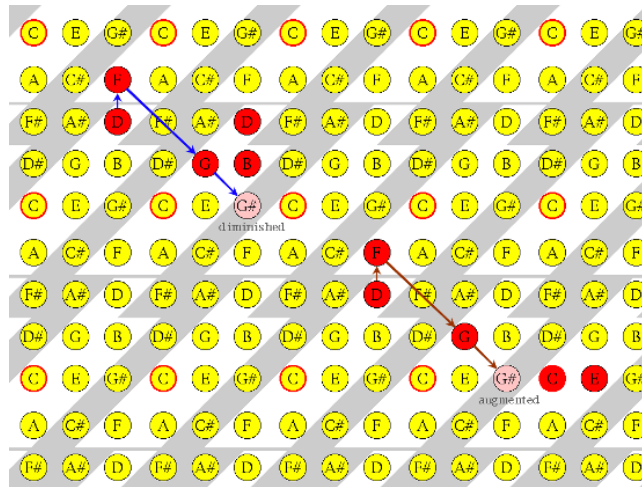


Figure 2.46: Two candidate ways of starting to develop Group One’s chord sequence, interpreting the undefined quality of chord 3 as diminished or augmented respectively.

Group 2’s novel chord sequence



Figure 2.47: The chord sequence that group 2 composed.

- C major F major Bb Major Eb major

This chord sequence is again full of interest, albeit in a completely different way. It is a resolute chord sequence in descending chromatic perfect fifths. An example point of comparison we mentioned to the children is that this chord sequence is like Jimi Hendrix’s “Hey Joe” but backwards. This suggests that a walking bass like Noel Redding’s walking bass in Hey Joe, which plots very easily and geometrically in Harmony Space, might be worth trying here. Or, perhaps more realistically, this chord sequence cries out to be extended something like this (Figure 2.48 - ignoring for now the last two chords in this figure):

- (C Major) C major F major A# Dom 7 D# Dom 7 G# Dom 7 C# Dom 7 F# Dom 7

But by this point this recomposed version would have had 7 events, so if we tucked in one more event with a cadential feel, we would have hit a round 8 events and the song could repeat nicely. So this suggests (Figure 2.48):

- (C Major) C major F major A# Dom 7 D# Dom 7 G# Dom 7 C# Dom 7 F# Dom 7 G major C major

This sounds good (Figure 2.48).

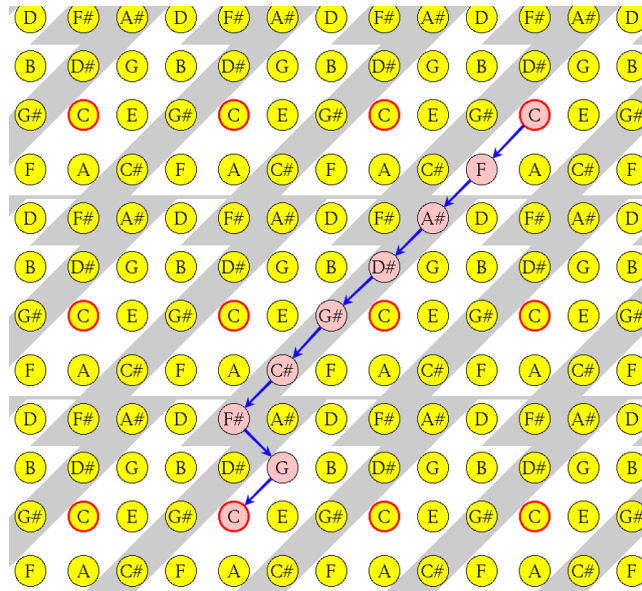


Figure 2.48: One way of extending Groups 2’s very interesting chord sequence.

Group 3’s novel chord sequence

This is yet another interesting chord sequence.

- (C Major) C major F major E minor D minor

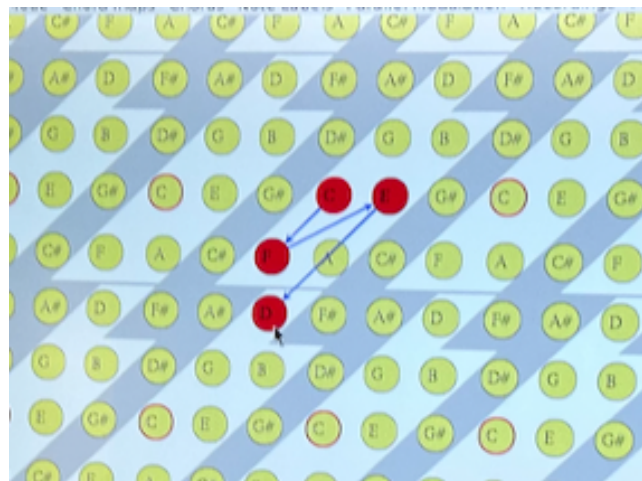


Figure 2.49: Group 3’s new chord sequence.

Modal insights on Group 3's composition

If we redraw Group 3's idea using the nearest-neighbour least-distance heuristic, we obtain (Figure 2.50):

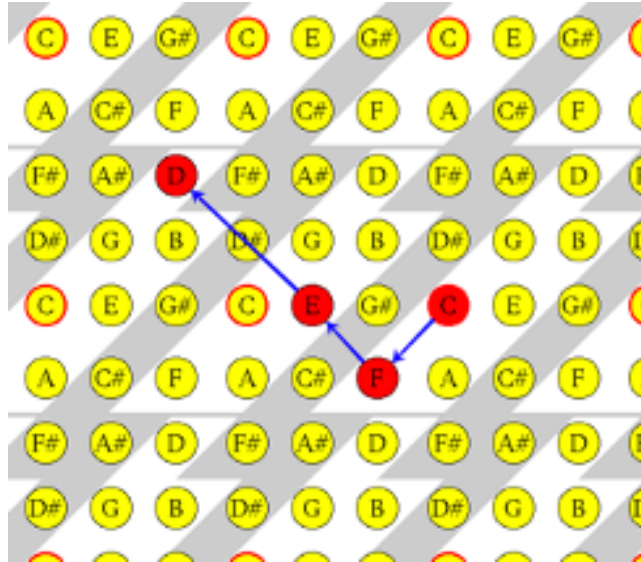


Figure 2.50: Group 3's new chord sequence redrawn using a nearest-neighbour least-distance heuristic.

And if we shift the hyper-metrical phase in the obvious way, and repeat the sequence we get:

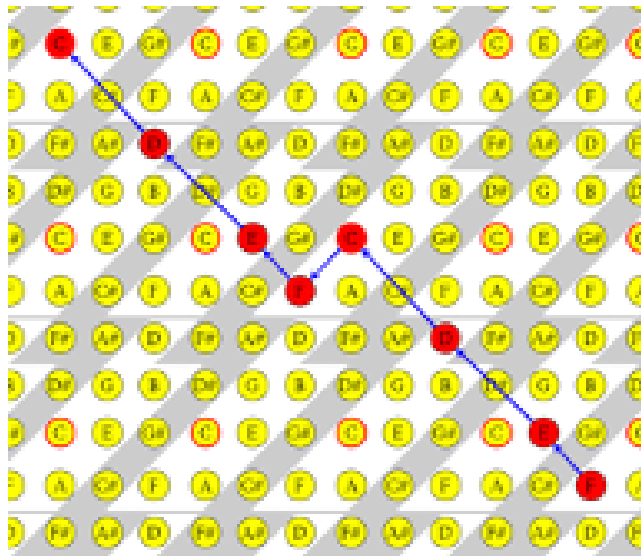


Figure 2.51: Group 3's new chord sequence phase shifted.

Which suggests something Ionian and modal (but not tonal). This certainly could work as the basis for a song. One way of developing it might be a modal modulation.

When considering how to modify this further to suggest new variants, one obvious exercise might be to see how well variants such as these work (Figures 2.52 and 2.53).

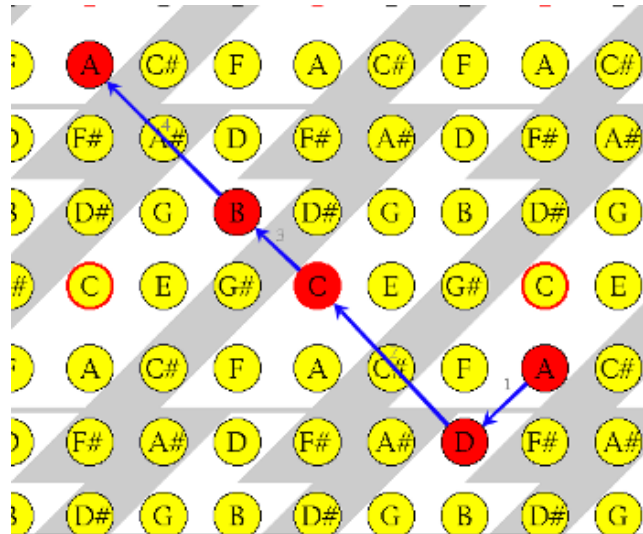


Figure 2.52: Trying Group 3’s idea with a different starting point – the ear should be the judge.

- (C major or A minor) A minor D minor C major B dim A minor

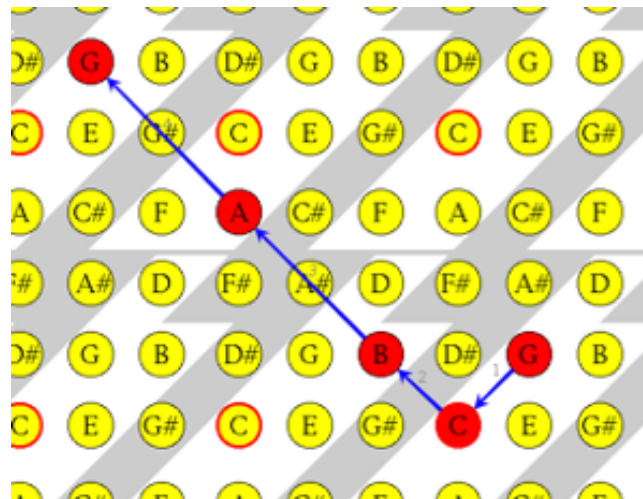


Figure 2.53: Trying Group 3’s idea with another different starting point.

- (C Major) G major C major B dim A minor G major

Do these modal variants work better or worse? If so, why? How might they be improved?

Tonal insights

Plotting Groups 3’s composition in different ways prompts a valuable brief aside. An interesting feature of Harmony Space’s spatial plotting heuristics is that while nine intervals (including the null move - i.e. not moving) all have unambiguous shortest distance moves (Figure 2.54).

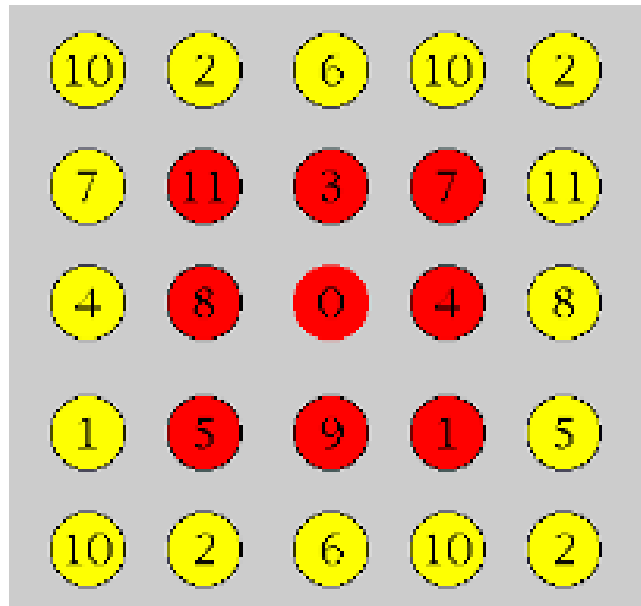


Figure 2.54: The 9 unambiguous nearest-neighbour moves from pitch zero (including not moving at all).

There are, however, two intervals that can be plotted in two different ways, the tone and the tritone, and in the case of the tone, these alternatives can be given interpretational significance (**Figure 2.55**). Loosely speaking, plotting moves of a tone on the modal (also chromatic, semitone) axis typically draws the eye to possible modal interpretations, whereas choosing to plot these moves on the tonal axis can suggest tonal interpretations (e.g. secondary dominant, etc).

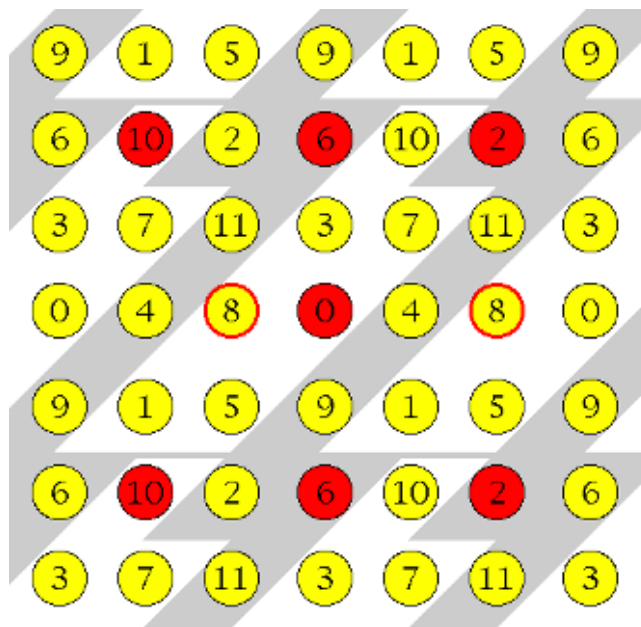


Figure 2.55: Two intervals with a tie for the nearest neighbour – whole tone up (and whole tone down) and the tritone. In the case of movement by a tone, this turns out to have suggestive interpretational value in Harmony Space.

Depending on the surrounding harmonic movement, one interpretation may make much more sense than the other, however, typically it is interesting to look at both interpretations, which can give richer insights.

Applying this idea to group 3's novel chord sequence, we can see that another way to look at it might be (Figure 2.56):

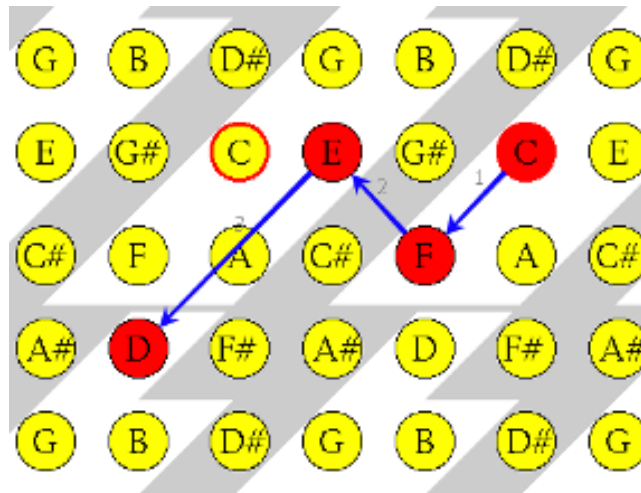


Figure 2.56: A suggestive alternative way of plotting Group 3's composition, that gives rise to tonally-related (as opposed to modally-related) ways of extending or modifying it.

Does this viewpoint suggest any different interesting modifications?

Well, by taking advantage of the implied dominant-powered momentum it might suggest something like this (Figure 2.57):

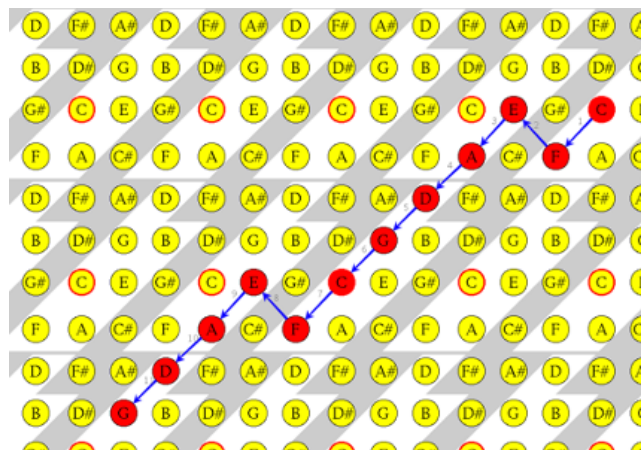


Figure 2.57: A different way of building on Group 3's composition, based on a tonal as opposed to modally-oriented spatial plotting heuristic.

Group 4's novel chord sequence

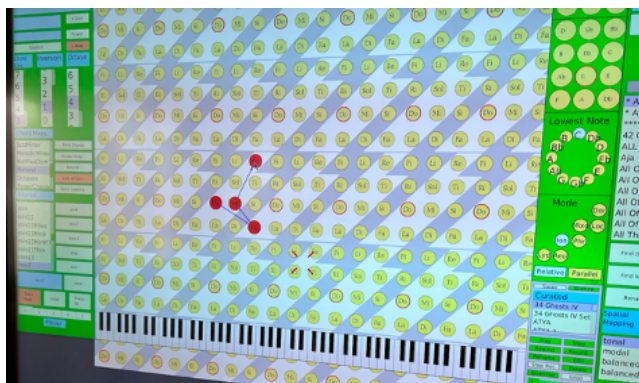


Figure 2.58: Group 4's chord sequence.

Finally, looking at Group 4's chord sequence (Figure 2.58) although the (non-nearest neighbour) path is different from group 3's path, when played using the next neighbour heuristic, this turns out to be the same chord sequence as group three.

- (C Major) C major F major E minor D minor

However, this is far from a problem, since, as we have already seen, the space of interesting variants of this idea to be explored is quite big.

Day Two of Second visit to Fonseca Harmonising melodies

At the request of Fonseca's music teacher, part two of day two was focused on an inverse problem – harmonising melodies previously composed by two students. We approached this as a one-on-one lesson with just the two students who had previously composed long modal minor melodies in MuseScore. We prepared by importing these melodies into Harmony Space from MuseScore. The approach we intended to take is described in Appendix 2. However, in retrospect we comprehensively conducted this session wrongly. Far too much talking and not enough simple tasks. Reflecting on this session, we realise that we should have pre-prepared very short melodies and played or imported them into Harmony Space (perhaps eight or so notes maximum). As a warm-up, we should have asked students to play and reflect on a small strategic selection of songs in Harmony Space to observe a variety of cadential mechanisms, and simple tonal and modal progressions. Then we could reasonably have proceeded as suggested in Appendix 2. This was a valuable lesson learned.

Other activities on second visit to Fonseca

Several other activities were undertaken at Fonseca, including work with three individual students with cognitive impairments. This was complicated by the fact that their helpers did not speak English, and we did not have translators for each student. One of the three students was able to learn to play "Isn't She Lovely" in about 15 minutes, and appeared pleased by this, but then understandably became tired. We noted that another of the students in this session was able to unintentionally demolish the entire screen layout with a sweep of the mouse in under a second – suggesting that we should block off certain meta-click functions. We carried out another day's worth of activities revisiting the students we had visited on the first trip. This was highly productive but did not reveal validation issues particularly different from those explored above.

Reflections on both visits to Fonseca

- Harmony Space allowed instrumentalists and those with little or no musical training to work together in an integrated way to perform music. This was evidenced, for example, by the laptop orchestra collaborating with the instrumentalists to play Stevie Wonder.
- This collaboration also allowed beginners to begin to articulate and manipulate fine grained technical harmonic properties such as inversion, chord size, and chord quality, generally reserved for instrumentalists.
- Harmony Space allowed musicians and beginners to work in an integrated way to both compose music together and perform together music that they had jointly created. This was evidenced partly by the composition exercises on the second school visit, but perhaps even more vividly by the dance mat and x-box controller collaborative performances in the first visit, of pieces that the children had composed.
- An initial proof of concept demonstration was carried out of identifying key by containment metaphor.
- Although we did not get the opportunity at the school to work with children with severe physical disabilities, the staff who witnessed song creation with dance mat and x-box controller performances were strongly of the view that this approach could allow children with physical disabilities to compose and take part via mouse and x-box controller in an integrated physical musical performance as first class citizens.

2.4. Evaluation method/rationale

2.4.1. Haptic Bracelets

Validation of the Haptic bracelets involves interviewing and gathering the views of educationists, inclusion specialists and end users, and gathering views on how the technology might be improved and better applied.

The valuation method rationale for the Haptic Bracelets focused on trialling them in a variety of use scenarios. This included work that we carried with profoundly d/Deaf expert music educator Sean Chandler, who established empirically that the bracelets have the potential to allow deaf people to participate as first class citizens in musical activities alongside hearing people. This was evidenced by the fact that, in small workshop context, from the vibrations alone, transmitted via the haptic bracelet, wirelessly communicated from the drum master tapping his own haptic bracelet, our profoundly deaf educator was able to learn intricate cross rhythms and track tempo changes.

The evaluation also included work we carried out with a large ensemble Brazilian and Cuban percussion group, and a public engagement workshop with members of the public as part of the Milton Keynes International Festival.

2.4.2. Harmony Space

The validation of a tool such as Harmony Space involves interviewing and gathering the views of educationists, inclusion specialists; working with end users; and gathering views on how the tool might be improved and better applied. In the early stages of the project, technical limitations restricted evaluation to the lab (D1.7). However, Harmony Space is now relatively robust and capable, allowing in-depth evaluations with the school I. C. Eleonora Pimentel Fonseca in Italy during two visits over six days.

2.5. Results and discussion

2.5.1. Haptic Bracelets

Enabling d/DeaF people to participate as first class citizens in musical activities alongside hearing people

Our d/Deaf educator Sean Chandler demonstrated empirically (D1.7) that that the bracelets have the potential to allow deaf people to participate as first class citizens in musical activities alongside hearing people. In particular, we were able to demonstrate that using haptic signals, transmitted via the haptic bracelet, a profoundly d/Deaf musicians

was able to learn intricate cross rhythms and track tempo changes. And the same time, from feedback working with large ensembles, we were able to identify various limitations.

- The impact detection system has limitations for precise signalling tasks.
- Temporal resolution currently blurs out at about quarter notes or eighth notes. Sixteenth notes cannot always be unambiguously distinguished.
- Haptic masking can sometimes occur when worn on the wrists.
- When directing larger ensembles or when working with multiple figures, a more convenient haptic signal routing system is needed.

The following issues have been addressed:

- to address more accurate signalling, the haptic bracelets were integrated with external control devices such as a stomp box and bop pad,
- to address better routing, the bracelets were integrated with an Ableton push routing system.

Other issues are being investigated in a follow-on project (see Further Work).

2.5.2. Harmony Space

Promoting inclusive engagement with harmony

- Engaging with extensive feedback has led to major enhancements in the capabilities and pedagogy of Harmony Space.
- We have demonstrated how Harmony Space can help those with no instrumental ability and negligible musical knowledge to quickly learn to manipulate complex harmonic materials and engage with music making.
- Harmony Space has been shown to enable those both with, and without, musical training to work as fully integrated teams on musical tasks.
- Discussion with Drake Music, and observations by staff at Fonseca strongly suggest that Harmony Space could allow children with profound physical disabilities to compose and take part in integrated physical musical performances as first class citizens.
- We have demonstrated proof of concept evidence that Harmony Space has the capacity to lower barriers for people to engage with material in the ChoCo database, engage with its harmonic structure and make evidenced corrections.

2.5.3. Further work on both Harmony Space and the Haptic Bracelets beyond Polifonia

The work of the ACCESS Pilot is being carried forward by Simon Holland, Nicholas Canny and Paul Mulholland in an Open University-funded Pilot project entitled “Music Making, Human Computer Interaction, Inclusion and Disability”, working with collaborators IC Fonseca, Drake Music Scotland and The Stables Theatre Milton Keynes. We have initial funding for three months and will seek to secure longer term funding. We have signed an agreement with IC Fonseca to collaborate over the next three years. The overall aim of the project is to further develop interactive resources to facilitate deep long-term engagement with music that is accessible and inclusive for all, especially those without musical training, irrespective of sensory or physical disability. We plan to create new partnerships with local schools and with music education researchers in the UK. Through our collaboration with Drake Music Scotland, we will investigate ways to use Harmony Space as part of their schools’ programme for people with disabilities. We are preparing research papers for The Journal of Music, Technology and Education, Musicae Scientiae, and Visual Languages and Human-Centric Computing. As part of the follow-on pilot initially funded by the Open University, we will conduct preliminary work with Professor Oliver Hodl (Technical University of Vienna) to specify an open source redesign of the haptic bracelets focusing on:

- higher temporal resolution,
- higher dynamic range, and
- lower latency,

with the longer-term aim of further improving integration with d/Deaf and deaf-blind people.

2.6. List of outputs (FAIR section)

The principal materials from the ACCESS validation are qualitative interview materials, photographs, images and questionnaires, and analyses thereof.

With the support of the Open University Library Research Support, we have implemented a number of actions to ensure that these outputs are in compliance with the FAIR Guiding Principles for scientific data management and stewardship.

Being integrated as part of the Polifonia Ecosystem, all data and metadata produced will be stored, and anonymized versions made available, where appropriate, in order to ensure its use beyond the project deadline.

Anonymised empirical data on the use of haptic and gestural technology will, wherever possible be published with appropriate metadata and documentation in an openly available repository such as Open Research Data Online (ORDO) - <https://ordo.open.ac.uk>, which guarantees the persistence of such resources.

The results of the ACCESS pilot and the results obtained from this validation will be part of future published work.

name	Harmony Space
component-id	harmony-space
type	Software
work-package	WP5
contributors	Simon Holland
link	https://www.harmonyspace.co.uk/
release-version	25-03-2024

2.7. Appendix 1

2.7.1. Anonymised PEER FEEDBACK from children after first visit

- We were divided into 4 groups and we tried to create chord sequences, then I was on a carpet and some of my companions brought the tempo and the chords with the joystick. My favourite part was the last day when we created a small orchestra with the flute, violin, guitar, and piano. It was really very beautiful because even with just a few instruments we created a small orchestra. It was a little annoying when we had to stop for a snack and return, but beyond that it was a wonderful project.
- During the project my classmates and I learned to use an application whose name I don't know. On the first day we started composing some songs thanks to the help of a professor who came from London, at first it was difficult but then together with my classmates I created a song, On the second day we started to modify the tonality of the notes using the slightly more "technical" part, we did it thanks to the help of the teacher,
- On the third day we brought the instruments, one of us used the controller, the other stood on the platforms to modify the notes and me and some other guys who play instruments played the notes. I liked this experience not only because we learned something new, but also because we did group work (and played an instrument) and I would like to do it again if I have the chance to do so.
- During the lessons, I used the controller for the sequence of the chosen musical notes and acted as an interpreter for the teachers. I really liked this project but I was disappointed that it lasted so short.
- In the project we built a melody and I participated both with the controller and playing the guitar. This project was fun and we learned to use a very useful English program to compose melodies and songs. I liked it a lot, but it only lasted 3 days, in 3 days there is obviously no possibility of learning the program well especially for boys or girls who don't know English well and therefore had difficulty understanding when it wasn't translated.
- In this project we learned to use a music program, called Harmony Space, and we understood how to write some compositions with it. It is a little difficult to use, but little by little you will be able to gain a little more confidence with it. In my opinion, however, they should change a little something, for example, there should be an easier way to be able to record, or there should be a button that only deletes a note, not the entire composition. All in all though, I think it's a nice program, and one that can easily be used to write compositions.
- This project was very captivating because we managed to discover many things about music with respect to how notes are used here, for example. The program is very beautiful and fun even if it is a little difficult to understand. In my opinion they should include a quickest key to record compositions, one to delete a single note (even if I don't know if there is one) because if for example you made a mistake in your composition you have to delete everything and start playing again, and finally make sure that the letters could become known so that if this program starts to be used somewhere other than in England it could be understood better. A very nice thing, however, was that there were mats like those for dancing which were used to change the inversions to the rhythm of the composition while clicking the notes with the joypad.
- The project with the English was very interesting, dealing with people from other countries was fascinating. My role within the work was to move the notes with the Xbox joystick and coordinate the work of the small orchestra. It was like being the director. The Harmony Space program is a little difficult to use, but with a little practice we will learn perfectly. The thing I didn't like about this work was not the complete explanation of all the buttons in the program, but only some of them, perhaps because they were too complicated. However, I really liked the organization of the project and Simon and Nicholas were very skilled and patient in their explanations.
- I participated in the music project with the English teachers. My task was, on the first and second day, to compose a very short piece using the chords and take inspiration from the corrections, on the third day to use the various keys of the chords. As for what needs to be improved, I won't add anything else, because I think this project is an excellent way to learn to use chords and their tonalities, especially for those who, like me, don't play an instrument and haven't studied music.
- I played the chords with the controller and I played the guitar. I liked it because they taught me how to use a new program.
- I brought the tempo to the mat and used the controller to play the notes. I liked the course because it was nice to see how work is done outside our school and I enjoyed working with other kids my age to achieve the goal

of completing this project together.

- I took the tempo to the mat and played with the digital piano, I really liked it because it was a lot of fun to use this program and it was nice to use it together with others.
- A few days ago we concluded the Polifonia project. We were divided into 4 groups each of which had 2 musicians and a person using the computer. On the screen we had notes and chords and we had to create a sequence of chords. There were 4 professors, plus two English ones, who created this project. This lasted 3 days, where we did more or less 3 hours of "theory". I played the guitar and accompanied the orchestra made up of a flute, another guitar, and a boy who set the tempo. The only flaw of this project was too much time wasted and too many hours of lessons were lost.
- The part I liked the most was when we did the orchestra. The part I didn't like was the waste of time and when we had to have a snack we would stop and then go back to class.
- My favourite part was the last day when we tried to create a small orchestra with: flute, guitar, violin and piano. Still, it was a great experience.

2.7.1.1. Feedback from the class music teacher

During the three days of participation in the Polifonia project we worked on the inclusion of students in the Harmony Space music program, obtaining good participation from all students.

It still needs to be verified that it works on Italian songs that the students are familiar with.

The various work plans proposed by the English teachers were carried out with great participation by all students and teachers.

Personally, I am interested in the interaction between the Muse Score and Harmony Space programs to facilitate harmonious accompaniment to the monody composed by the learners. Basically the work to be done should be:

Compose a melody and insert a harmonic sequence

Write a harmonic sequence and compose a melody on top.

2.8. Appendix 2 Harmonising a Melody

General heuristic advice

1/ What key or keys does it seem to fit? Might the key change? If so, where? (Try using visual best-fit methods – sometimes such a check can reveal unexpected alternatives).

2/ Does the melody have a home note? Might it change? If so, where? (This can reveal separate information from question 1 – there may be harmonisations – including possible modal modulations).

One way of approaching questions 1 and 2 is to consider: if you were allowed a single bass note to play along with the whole thing, which bass note would you choose? This can be tried with any bass instrument or the voice.

How to harmonise a tune -general advice

Try noodling around with a bass. Try to find a bass line that fits. If the melody is long or busy, try just one bass note per bar. Don't expect anything to work straight away – just keep trying till it starts getting better.

Things to try in Harmony Space

Basic ideas

Segmentation

If the melody is long, look for places it pauses, or seems to partially finish one idea before starting the next part.

Consider breaking up the melody into these parts.

Cadences

Would any of these partial endings make good cadences?

If previously you have looked at songs you like, or instructive songs in Harmony Space, you should be aware of various kinds of cadence e.g.

major V I, major II V I, major I IV V I, Dorian I VII I, Dorian III II I, Aeolian VI VII I

Mixolydian VII IV I etc

Do any of these fit?

One good way to harmonise a melody is to work back from any cadences you can find that seem to fit.

Non cadential segments?

If the melody breaks into pause points that do **not** seem cadential, does the flow of just the pause points suggest some way to craft a cadence for the whole thing?

Tracking in thirds or sixths

Would any of the melody track interestingly in diatonic thirds or sixths?

Try the one per bar bass line as triads?

Look at the plot of the melody in Harmony Space

(see “How to import a melody into Harmony Space)

Look at the plot of the melody in Harmony Space for fragments of diads & triads (some songs harmonise more interestingly in diads than triads)

Can you see a way to give the fragmentary bits direction, contrast or thematic unity?

(Consider how a variety of songs you like or know in Harmony Space give the harmony direction, contrast, or thematic unity.

3. ORGANS

3.1. Overview of the pilot

In this pilot, we have assembled a knowledge graph containing detailed information about the histories and characteristics of virtually all important historic organs in the Netherlands, and basic information about c. 25,000 organs across Western Europe, harvested from various national databases.

The main data set for the ORGANS pilot consists of the complete texts of the so-called Dutch Organ Encyclopaedia (Official title: *Het historische orgel in Nederland*), published by the National Institute for Organ Art (Nlvo)¹ between 1997 and 2010. These are the input data for the pilot. The printed edition consists of 15 volumes with 4,500+ pages. This encyclopedia is currently the primary source on Dutch organ history, describing activities of organ builders and characteristics of the instruments in a uniform manner. In the pilot project, this vast amount of knowledge has been made fully accessible in the form of a knowledge graph.

This enables detailed study of the history of organs and organ building in the Netherlands and abroad. Such a knowledge base will be highly valuable for music historians, but also for organ builders and organ advisors who are involved in restoration, maintenance, or reconstruction projects. The resulting data is a valuable resource of knowledge on Dutch organs, used by music historians, organ advisors, the Cultural Heritage Agency of the Netherlands, organ builders, and others.

The texts have been made available for the Polifonia project as a collection of Word documents, and have been provided by the Nlvo. These Word documents contain the version of the text that was submitted to the printer for preparing the printed volumes of the encyclopaedia. Redactional changes and corrections that were made during the printing process have been lost. Although the digital data as we now have in these documents include a wealth of information on virtually all Dutch pipe organs of historic importance, the presence of uncorrected errors is problematic. The Nlvo currently is building an online relational database that has ingested the data extracted from the Organ Encyclopaedia. This database is hosted at www.pipeorgan.nl. In a series of projects that will run over the next two years, the contents of the database will manually be corrected and updated. After finalizing this process, the updated data will be integrated into the Organs Knowledge Graph.

The main aims of the ORGANS pilot are

- to extract data from the input text;
- to export a SQL representation of the data suitable for importing into the Nlvo database;
- to design an organs ontology as part of the Polifonia Ontology Network;
- to construct a knowledge graph with the extracted data;
- linking the knowledge graph with mined data from the Polifonia Corpus (in collaboration with WP4);
- linking the knowledge graph with various other external datasets;
- designing interface components to explore and query the data.

The following concrete results of the ORGANS pilot are subject to validation in the current report:

1. second iteration of data extraction from the input data;
2. export of the extracted data in SQL for the Nlvo database;
3. the organs ontology;
4. the final version of the ORGANS knowledge graph;
5. the interface to the data as developed in collaboration with the Open University London.

¹<http://www.nationaalinstituutorgelkunst.nl>

3.1.1. Connection with research WPs

There has been close collaborations between the ORGANS pilot and the teams of workpackages 2 and 5.

Workpackage 2 The development of the Organs module of the Polifonia Ontology Network (PON)² was a result of intensive collaboration with the WP2 team. Through iterative refinement, the ontology evolved into its current state, which is able to capture the complex history of pipe organs, their technical details, and the events they've been part of. The interaction between the ORGANS and WP2 was crucial for this result. The pilot provided the domain knowledge, and WP2 the necessary skills of ontology design.

Workpackage 5 A user interface has been designed in close collaboration with the WP5 team at Open University, London. This user interface is crucial for the success of the pilot, as the targeted user groups may lack the skills to construct SPARQL queries for accessing the knowledge graph.

3.1.2. Progress from last deliverable D1.4

The following results have been achieved during the final phase of the project:

1. Data extraction and enrichment:
 - a) Extraction of dating and builder information for the deliveries.
 - b) Extraction of playing aids.
 - c) Extraction of literature references.
 - d) Disambiguation of builder names.
 - e) Disambiguation of stop names.
 - f) Disambiguation of division names.
 - g) Addition of images.
 - h) Addition of basic information (location, builder, dating, image) for 8,066 organs from the *Inventaire Nationale des Orgues*.³ These are mainly organs in France.
 - i) Addition of basic information for 15,412 organs from the *Organ Index*.⁴ These are mainly organs in Germany, Austria, and Switzerland.
 - j) Addition of basic information for 1,384 organs from the *Catalogo generale dei Beni Culturali* of the Italian Ministry of Culture.⁵ These are mainly organs in Italy.
2. SQL export for the relational database of the NivO. The full contents of the knowledge graph has been exported for ingestion in the relational database of the NivO. In consultation with the NivO, they adapted their database structure to fit better the data that results from the Organs pilot.
3. The competency questions based on the user stories have been refined.
4. The Organs Ontology has been refined and put into use.
5. The final iteration of the Organs Knowledge Graph is fully aligned with the Polifonia Ontology Network.
6. A first iteration of the Pathways user interface has been developed in collaboration with Open University London, and has been evaluated with expert users.

A detailed account of the progress concerning the data extraction and the affordances of the Knowledge Graph is included in the tables in the next section.

3.2. Evaluation method/rationale

The different results have their own validation procedure. The data extraction and the SQL export to the NivO database are validated in terms of data fields covered. A full table will be presented in the next section. The ontology and the knowledge graph will be validated by congruence with the competency questions as identified in the Polifonia

²<https://github.com/polifonia-project/organs-ontology>

³<http://inventaire-des-orgues.fr>

⁴<https://organindex.de>

⁵<https://catalogo.beniculturali.it>

stories. The Pathways interface has been validated in an informal session with three intended expert users. This has been reported in Deliverable 5.2.

3.3. Results and discussion

3.3.1. Data extraction and SQL export

The following table presents an overview of the data fields for each organ to be extracted from the text of the organ encyclopaedia. The status at the time of the previous report is added in the columns dated Feb. 2023. The column 'Extracted' indicates whether the data extraction has been done, and the column 'SQL export' indicates whether the data has been exported to the SQL structure for the NlVO database. The status is 'DONE' if the data has been extracted and can be considered of sufficient quality, 'PARTLY' if more work is needed to correct remaining errors or solve remaining issues, and 'TODO' if the extraction has not been done yet.

Data field	Extracted Feb. 2023	Export Feb. 2023	Extracted Apr. 2024	Export Apr. 2024
Base Information:				
Location – Place	DONE	TODO	DONE	DONE
Location – Building	DONE	TODO	DONE	DONE
Location – which organ in this building?	DONE	TODO	DONE	DONE
Builder	DONE	DONE	DONE	DONE
Disambiguated Builder Name	TODO	TODO	DONE	DONE
Year of First Delivery (building)	DONE	DONE	DONE	DONE
Original Location	DONE	N/A	DONE	DONE
Historic Locations	TODO	TODO	DONE	DONE
Monument Number	DONE	DONE	DONE	DONE
Organ Number	DONE	DONE	DONE	DONE
History of the Instrument:				
<i>Years of delivery (major change). For each delivery:</i>				
- Dating	TODO	TODO	DONE	DONE
- Builder	TODO	TODO	DONE	DONE
- Disambiguated Builder Name	TODO	TODO	DONE	DONE
<i>Maintenance and restoration projects. For each project:</i>				
- Dating	DONE	PARTLY	DONE	DONE
- Builder	DONE	DONE	DONE	DONE
- Disambiguated Builder Name	TODO	TODO	DONE	DONE
- Changes made to the organ	DONE	DONE	DONE	DONE
- Information extracted from the changes	TODO	TODO	TODO	PARTLY
<i>Dispositions. For each disposition:</i>				
- Source	DONE	DONE	DONE	DONE
- Dating	DONE	PARTLY	DONE	DONE
- Divisions. For each division:				
-- Division Name	DONE	DONE	DONE	DONE
-- Disambiguated Division Name	TODO	TODO	DONE	DONE
-- Remarks	DONE	N/A	DONE	DONE
-- Playing Aids	TODO	TODO	PARTLY	PARTLY
-- Stops. For each stop:				

--- Name	DONE	DONE	DONE	DONE
--- Partitioning	DONE	DONE	DONE	DONE
--- Specification	DONE	DONE	DONE	DONE
--- Remarks	DONE	DONE	DONE	DONE
Current State of the Instrument:				
<i>Current disposition:</i>				
- <i>Divisions. For each division:</i>				
-- Division Name	DONE	DONE	DONE	DONE
-- Remarks	DONE	N/A	DONE	N/A
-- Stops. For each stop:				
--- Name	DONE	DONE	DONE	DONE
--- Partitioning	DONE	DONE	DONE	DONE
--- Specification	DONE	DONE	DONE	DONE
--- Remarks	DONE	DONE	DONE	DONE
- Playing Aids	PARTLY	PARTLY	PARTLY	PARTLY
Console Location	DONE	DONE	DONE	DONE
Pitch	DONE	DONE	DONE	DONE
Keyboard Ranges	PARTLY	PARTLY	PARTLY	PARTLY
Temperature	DONE	DONE	DONE	DONE
Wind Pressure	DONE	DONE	DONE	DONE
Type of Wind System	DONE	DONE	DONE	DONE
<i>Composition of Compound Stops.</i>				
<i>For each stop:</i>				
- Identification of the Stop	DONE	TODO	DONE	DONE
- Keyboard ranges. For each range:				
-- Lowest key	DONE	TODO	DONE	DONE
-- Composition (list of foot heights)	DONE	TODO	DONE	DONE
Description of the Case (running text)	DONE	DONE	DONE	DONE
Information from Case Description	TODO	TODO	TODO	TODO
Peculiarities (running text)	DONE	DONE	DONE	DONE
Information from peculiarities	TODO	TODO	TODO	TODO
Literature				
<i>Literature References. For each reference:</i>				
- Author	TODO	TODO	DONE	DONE
- Year	TODO	TODO	DONE	DONE
- Title	TODO	TODO	DONE	DONE
- Collection Title	TODO	TODO	DONE	DONE
- Volume	TODO	TODO	DONE	DONE
- Pages	TODO	TODO	DONE	DONE
- Publisher	TODO	TODO	DONE	DONE

As can be seen in the Table, most of the desired data fields have successfully been extracted from the input text. Some of the fields are not available in the database structure of the NIV database (status 'N/A'). Some of the fields require further work (status 'PARTLY' or 'TODO'). The main remaining challenge is the following. Both the description of the case, and the peculiarities are two chunks of unstructured, running text, the content of which varies a lot among the organs. Sophisticated natural language processing techniques are required to extract information from these texts. This has not been accomplished within the project. It will be addressed in a follow-up project.

3.3.2. Ontology and Knowledge Graph

In close collaboration with the WP2 team, an ontology module has been developed specifically for organs. The work on this has been reported in Deliverable 2.3.

From the extracted data a knowledge graph has been constructed, which is fully aligned with the Organs ontology. The KG is accessible from the SPARQL endpoint at the Polifonia server.⁶ The following table shows for each of the competency questions (CQs) from the relevant user stories, which of those competency questions can be addressed with this version of the knowledge graph.

Competency Questions	Covered Feb. 2023	Covered Apr. 2024
Paul#1_OrganComparison		
CQ1: Which are all organs at location X?	Y	Y
CQ2: Which are all organs in city X?	Y	Y
CQ3: Which are all organs near to geographic coordinates x, y?	N	Y
CQ4: Which are all organs that have stop X?	Partly	Y
CQ5: Which are all organs with playing aid X?	N	Partly
CQ6: Which are all organs with more than X stops?	N	Y
CQ7: Which are all organs with more than X keyboards?	N	Y
CQ8: Which are all organs with manual range starting lower than X?	N	Partly
CQ9: Which are all organs with manual range ending higher than X?	N	Partly
CQ10: Which are all organs with pedal range starting lower than X?	N	Partly
CQ11: Which are all organs with pedal range ending higher than X?	N	Partly
CQ12: Which are all organs with wind system type X?	N	Y
CQ13: Which are all organs with temperament X?	N	Y
CQ14: Which are all organs with year of construction after X?	Y	Y
CQ15: Which are all organs with year of construction before X?	Y	Y
CQ16: Which are all organs built by organ builder X?	Y	Y
CQ17: Which are all organs which has been maintained or restored by builder X?	Y	Y
CQ18: Which are all organs that are linked to person X?	N	N
CQ19: Which are all organs with more than X bellows?	N	Partly
CQ20: Which are all organs with pitch heigher than X Hz?	N	Y
CQ21: Which are all organs with wind pressure higher than X mm?	N	Y
CQ22: Which are all organs with wind pressure lower than X mm?	N	Y
CQ23: Which are all organs with key action type X?	N	N
CQ24: Which are all organs with stop action type X?	N	N
CQ25: Which are all organs with pitch lower than X Hz?	N	Y
CQ26: Which are all organs with console location X?	N	Y
CQ27: Which are all organs based on search term X?	Y	Y
CQ28: Who was the builder of organ X?	Y	Y
CQ29: What is the current location of organ X?	Y	Y
CQ30: What were previous locations of organ X?	N	Y
CQ31: When have changes been made to organ X?	Y	Y
CQ32: What changes have been made to organ X?	Y	Y
CQ33: Why was change X made to organ Y?	N	N
CQ34: What is the current disposition of organ X?	Partly	Y
CQ35: What is the composition of stop X in organ Y?	N	Y

⁶<https://polifonia.disi.unibo.it/organs/sparql>

CQ36: Of which material is stop X made?	N	N
CQ37: Of which material is the case of organ X made?	N	N
CQ38: Which materials have been used to build organ X?	N	N
CQ39: What historic dispositions are known for organ X?	N	Y
CQ40: What is the source for historic disposition X of organ Y?	N	Y
CQ41: What is the current pitch of organ X?	Y	Y
CQ42: What is the current temperature of organ X?	N	Y
CQ43: What is the current console location of organ X?	N	Y
CQ44: What is the current wind system type of organ X?	N	Y
CQ45: What is the current wind pressure of organ X?	N	Y
CQ46: What are the characteristics of the case of organ X?	N	Partly
CQ47: What are decorative elements of the case of organ X?	N	Partly
CQ48: What are inscriptions on the case of organ X?	N	Partly
CQ49: What are the dispositions of all organs made by organ builder X?	Partly	Y
CQ50: What are the sources for fact X?	N	Partly
CQ51: Given organ X, what organs are similar according to similarity measure Y?	N	N
CQ52: What literature exists about organ X?	N	Y
CQ53: What literature exists about organ builder X?	N	N
CQ54: What sources are available regarding the sound of organ X?	N	N
CQ55: Does organ X stil exist?	Y	Y
CQ56: When was organ X deconstructed?	N	N
CQ57: Why was organ X deconstructed?	N	N
CQ58: Which parts of deconstructed organ X are still in existence?	N	N
CQ59: Where are parts of deconstructed organ X currently located?	N	N
CQ60: Which components of organ X have been reused from other (earlier) organs?	N	N
CQ61: What is the origin of component X of organ Y?	N	N
CQ62: Which are all organs that have components made by organ builder X?	N	N
Story: Paul#2_ResourceReliability		
CQ42: What are the sources for fact X?	N	Partly
Story: Amy#1_OrganTrends		
CQ1: What are geographically distinct features of organs from region X? (e.g., in what differ 17th century Southern German organs from 17th century Spanish organs?).	N	Indirectly
Story: Frank#1_OrganKnowledge		
CQ1: What does the organ look like?	N	Y
CQ2: Which organs are built in a similar style?	N	N
CQ3: What is the concert agenda for organ X?	N	N
CQ4: Where to find audio/video resources featuring organ X?	N	N
CQ5: Where to find audio/video resources featuring organist X?	N	N
CQ6: What is the address of the owner/maintainer of organ X?	N	N

The reason not all CQs are addressed in the current version of the knowledge graph is that the required information is not available in the source text of the Organ Encyclopaedia. Given the general aim of the pilot, these questions could be considered 'could haves'. At the start of the project, we were aware that these aims were ambitions. They are not critical for the success of the pilot.

3.4. Summary of the outputs (FAIR section)

name	Organs Knowledge Graph
component-id	organs-knowledge-graph
type	KnowledgeGraph
work-package	WP1
related-components	pipe-organ-pathways, Paul#1_OrganComparison, Paul#2_ResourceReliability, Amy#1_OrganTrends, Frank#1_OrganKnowledge
licence	CC-BY
contributors	Peter van Kranenburg, Eline Duijsens, Eoin Kearns
link	https://github.com/polifonia-project/organs-knowledge-graph
release-version	v1.0

4. BELLS

4.1. Overview of the pilot

In the collective perception, bell tower is certainly one of the architectural elements that best characterize the Italian landscape. To better understand their value could be considered how Bell Towers can play a cultural role in the Italian society, thanks to their recurring presence in the space landscapes. Bell structures are widespread both in urban and rural areas and contribute to the distinctive shape of a landscape, defining its soundscape and playing as markers of daily, festive and ritual times. Bell heritage is complex and fascinating and influences our perception of the places we live daily. Moreover, bell tower could be mostly considered a very large musical instrument capable of hosting a belfry, a set of bells and related bells, that can produce different notes. This musical instrument can have different characteristics and be manually played in different ways, defining very specific identity values for each territorial area of reference, depending on traditional culture. The sounds produced, as well as the formal characteristics of every part of this instrument, are related to the materials used, the construction process, the tradition of the players, the musical repertoires that characterize a geographical area, as well as values are also related to intangible aspects such as human know-how and the ways of transmitting this knowledge and practice. The historical Bell heritage can therefore be defined as a complex set of two kinds of elements that take place in a specific territorial area: material (bell tower, belfry, armor, bells) and immaterial (oral tradition, musical repertoire, language, sound practices, traditional techniques). This pilot intended to collect descriptions of set of bells and bells in their tangible and intangible dimensions and encode the information in a knowledge graph.

4.1.1. Connection with research WPs

The pilot contributed on the activities of the following Wps:

- WP1: BELLS contributed to Wp1 activities by filling out the survey ¹ and collecting research needs and providing data for the data stories represented in the Melody portal ².
- WP2: BELLS contributed to Wp2 with the formalization of BELLS Ontology ³ and BELLS Knowledge Graph ⁴
- WP4: although being a pilot based mainly on structured data, BELLS has collected a corpus of texts available in the Polifonia Corpus Web Application. ⁵
- WP6: BELLS has contributed to WP6 with the participation and organization of different events in cooperation with bell ringers associations. The participation of bell ringers in Polifonia was mentioned in the Unesco ICH submission file for "Traditional bell art" ⁶

4.1.2. Progress from last deliverable D.1.4

The first validation was mainly focused on data collected, in terms of quantity (according to scenarios and competency questions), and quality (consistency and compliance with specific domain standards). The results showed minor action to be taken. Further feedbacks were requested for the validation of tools/means/software and a general validation of outputs and results. The first validation round indicated the need to improve the usability of the inter-

¹<https://github.com/polifonia-project/survey>

²https://projects.dharc.unibo.it/melody/bells/overview_of_the_bells_in_liguria

³<https://github.com/polifonia-project/bells-ontology>

⁴<https://github.com/polifonia-project/bells-knowledge-graph>

⁵<https://polifonia.disi.unibo.it/corpus/>

⁶<http://www.iccd.beniculturali.it/getFile.php?id=8740>

face. The main correction actions were therefore directed towards better usability, in particular giving centrality of the geographical dimension through a map.

4.2. Evaluation method/rationale

Following last validation round, the Pilot's validation actions are:

- 1) Quantitative
- 2) Qualitative, in terms of
 - a) Data quality
 - b) Usability

In particular, following users' feedback on usability, the validation has been focused mainly on interface and on the general evaluation of the pilot in term of utility, usability and data reusability.

4.3. Results and discussion

The evaluation process led to the following results

4.3.1. Quantitative

A first validation method follows a quantitative approach, measuring the progression of data collected: the BELLS knowledge graph contains 249.050 triples, and includes 204 distinct bells, related to 30 distinct sets of bells. Therefore, the KG includes approx. 230 instances of musical instruments. Other notable entries in the knowledge graph are:

- 1280 roles played by some agents;
- 507 time indexed typed locations, linked to 68 distinct addresses;
- 295 photographic documentations;
- 267 events
- 203 audio documentations;

4.3.2. Qualitative

BELLS involved a group of 27 experts, asking their feedback through interviews and a specific questionnaire (Figure 4.1). Experts' profiles for the validation are:

- Personnel involved in the protection of cultural heritage
- Musicologist/Ethnomusicologists
- Bell-ringers
- Digital Humanists
- Campanologists
- Generic Users

Validation of data quality and reusability

- How would you overall evaluate the quality of the data collected for the "BELLS" pilot project, which can be consulted at this link: <https://catalogo.beniculturali.it/search/typeOfResources/MusicHeritage?region=Liguria> (Figure 4.2).

2. Qual è la tua attività principale?

[Altri dettagli](#)

● Operatore dei Beni Culturali	12
● Musicologo/Etnomusicologo	4
● Campanaro	3
● Digital Humanist/informatico	5
● Campanologo	3
● Utente generico	6
● Altro	2

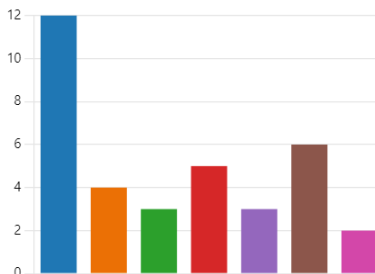


Figure 4.1: Participants

3. Come valuteresti complessivamente la qualità dei dati raccolti per il progetto pilota "BELLS", consultabili a questo link: <https://catalogo.beniculturali.it/search/typeOfResources/MusicHeritage?region=Liguria>

[Altri dettagli](#)

[Dati analitici](#)

4.96
Valutazione media

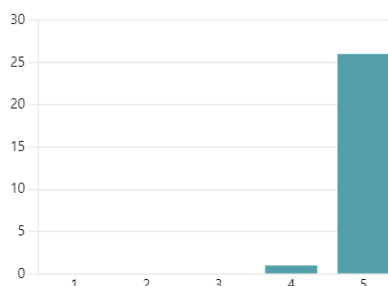


Figure 4.2: Data quality

Validation of tools/means/software validation of interface-Usability

- How would you evaluate the interface developed at this link: <https://bells.fly.dev/> (Figure 4.3)
- How would you overall evaluate the usefulness of the project with a view to future developments for the activities of knowledge, protection, safeguarding and dissemination of the cultural heritage and specifically of the bell tower heritage?(Figure 4.4)
- Do you believe that the data collected can be used by you for research, consultation and reuse purposes? (Figure 4.5)
- What is your level of satisfaction with each of the following items? (Figure 4.6)
 - Innovation
 - Navigation of results
 - Data quality
 - Data reusability
 - Future development potential

4. Come valuteresti l'interfaccia sviluppata a questo link: <https://bells.fly.dev/>

[Altri dettagli](#) [Dati analitici](#)

4.93
Valutazione media

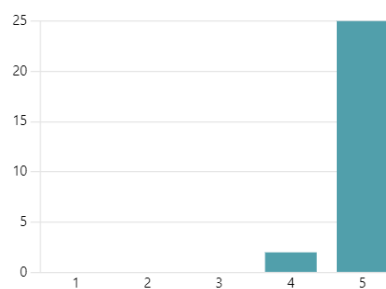


Figure 4.3: Interface

5. Come valuteresti complessivamente l'utilità del progetto nell'ottica di sviluppi futuri per le attività di conoscenza, tutela, salvaguardia e divulgazione del patrimonio culturale e nello specifico del patrimonio campanario?

[Altri dettagli](#) [Dati analitici](#)

4.93
Valutazione media

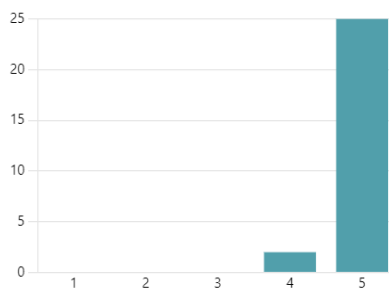


Figure 4.4: Usability (1)

6. Ritieni che i dati raccolti possano essere da te utilizzati a scopo di ricerca, consultazione e riuso?

[Altri dettagli](#) [Dati analitici](#)

4.93
Valutazione media

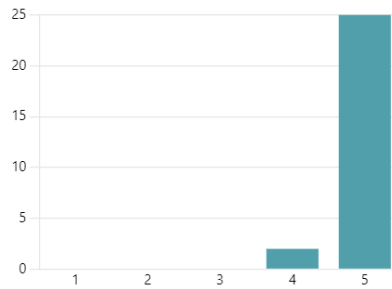


Figure 4.5: Usability (2)

7. Qual è il livello di soddisfazione per ciascuno degli elementi seguenti?

[Altri dettagli](#)

■ Per niente soddisfacente
 ■ Piuttosto soddisfacente
 ■ Molto soddisfacente
 ■ Completamente soddisfacente
 ■ Non applicabile

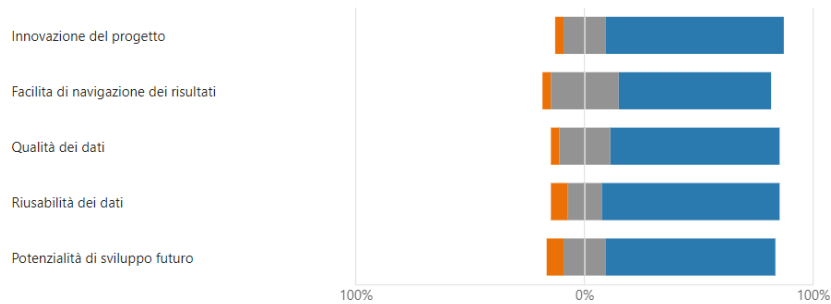


Figure 4.6: General feedback

4.4. Summary of the outputs (FAIR section)

List all components related to this Pilot

name	BELLS Knowledge Graph
component-id	TBD
type	Knowledge Graph
work-package	WP2
licence	CC-BY v4
contributors	Elena Musumeci
link	https://polifonia-project.github.io/ecosystem/polifonia-project/bells-knowledge-graph/README.html#bells-knowledge-graph
release-version	v1.0
doi	https://zenodo.org/doi/10.5281/zenodo.7633704

name	BELLS Ontology
component-id	https://w3id.org/polifonia/ontology/bells-ontology/
type	Ontology
work-package	WP2
licence	CC-BY v4
contributors	Valentina Carriero Elena Musumeci
link	https://github.com/polifonia-project/bells-ontology/blob/main/ontology/bells.owl
release-version	v1.0
doi	https://zenodo.org/doi.org/10.5281/zenodo.7919970

5. INTERLINK

5.1. Overview of the pilot

The INTERLINK pilot aims at connecting knowledge graphs of European musical cultural heritage, which are notoriously large and use a diversified vocabulary of relations. The pilot is not bound to a specific institutional dataset, collections or use cases; but rather summarises the common needs of all pilots of representing their data as music knowledge graphs, and develops technologies that will enable finding new meaningful connections among them. Initial efforts were devoted at creating ChoCo [9] – the largest knowledge graph of music harmony, and harmonising the stakeholders' requirements into the Polifonia Ontology Network – to facilitate entity linking and ontology alignment upfront. We then looked into linking at the musical content level, and contributed methods and KGs on harmonic patterns as part of the Harmory project [10], also showing how these enable novel applications in music discovery and computational creativity. Overall, these efforts provided the infrastructure and the requirements to address the last challenge of the pilot: i.e. linking knowledge graphs at scale to ensure the long-term reuse of our resources.

Learning numerical representations from knowledge graphs is currently the most prominent approach to address *link prediction*, in addition to knowledge discovery and reasoning tasks. This is often done by allocating and learning embedding tables for all (or a subset of) the entities. As this scales linearly with the number of entities, learning embedding models in real-world KGs with millions of entities is computationally intractable. Driven by these motivations, we proposed a novel embedding model that aims to address the scalability and inductive embedding problems faced by current methods. Specifically, our model aims to (i) address the scalability requirement by only storing relation representations and learning to compute entity representation by designing a model architecture that draws inspiration by sequence models; (ii) leverage path information to contextualise nodes and the connectivity patterns between them aiming to produce useful structure aware entity representations without having to conduct message passing across the entire graph as Graph Neural Networks (GNNs) do; (iii) address the inductive embedding of new nodes by only using relations and paths to calculate node embeddings, thus being able to produce embeddings for new nodes without additional training. The embedding models are trained for and evaluated on link prediction and relation prediction, and can hence be used to complete (triples in) knowledge graphs and finding links between their entities. Finally, to complement our vision, the pilot also developed a new computation method for modelling musical influence – which enables the detection of creative influence between artists using techniques based on graph theory and complex network science.

5.1.1. Connection with research WPs

The pilot contributed to and relies on the technologies developed in:

- WP2, focusing on ontology requirements for interconnecting musical resources, both at the metadata and at the content levels. This also includes the collaborations established with the Polifonia Ontology Network, and in particular with Deezer and the Digital Music Observatory, to evaluate the suitability of our ontologies outside the scope of the project, and with the central goal of interlinking musical resources.
- WP3, central to the development of state of the art methods for harmonic segmentation and similarity, both in the symbolic domain. This also include collaborations within the consortium to facilitate the reuse of our resources to address similar requirements.
- WP5, to promote the reuse of the ChoCo KG to support its integration with Harmony Space.

5.1.2. Progress from last deliverables

Results are still consistent with D1.5 *Intermediate validation reports for pilots: INTERLINK and FACETS* [11] and D2.5 (*Methods for interlinking knowledge graphs*) [12]. We performed additional experiments to better understand the interplay between relation and link prediction to complement the evaluation of our KG embedding models.

5.2. Evaluation method/rationale

Given the heterogeneous variety of the pilot's outputs, we have employed different evaluation strategies to assess the effectiveness and quality of our data, models, and results.

To evaluate ChoCo, as carefully outlined in [9], we carried out various experiments focusing on two components of the integration workflow: (i) the integration of annotation formats within the JAMS [13] annotation standard (the *JAMifier* component); and (ii) the chord conversion module that integrates all the notations within the Harte [14] system (the *Chonverter* module). As the goal of the *JAMifier* is to automatically generate a JAMS dataset given a music collection providing chord annotations and metadata in different formats, notations, and conventions, we carried out a series of tests to compare a sample of generated JAMS files with those that are expected from this process. This required the creation of a groundtruth dataset of JAMS files that were manually produced by two human annotators from a given template (the backbone of a JAMS file), and through manual inspection of the original collections. Instead, evaluating the output of the chord conversions requires musical expertise and familiarity with different chord notations. Therefore, we performed a 2-step evaluation with music experts to validate the alignment and the conversion rules. The first step focused on validating the context-free grammars used to parse chords in the original formats and aligning them to the corresponding chord families. Participants were presented with 3 different grammars, including 250 mapping rules to validate. Whenever a rule was deemed incorrect, participants were asked to provide the expected mapping. As for step 2, Once chords were converted, the final result of the conversion was validated. This step also allowed for the validation of other conversion types that were not validated in Step 1, such as Roman numerals and Polychords. In addition, even for annotations originally provided in Leadsheet, this step allows for the validation of added/removed notes and inversions. Overall, 4 participants with at least 5 years of musical training were recruited for this experiment. Participants were first introduced to the task, and asked to express their level of familiarity with the different chord notations, and the validation methodology.

To evaluate Harmony, as outlined in [10], we focused on the two state of the art methods driving its creation: the harmonic segmentation, and the harmonic similarity models. In particular, we evaluated the DTW harmonic similarity by comparing our implementation with other algorithms for the *cover song detection* task – a common benchmark for similarity algorithms in the symbolic music domain [15, 16]. In this comparison, performance is evaluated using two standard metrics: *First Tier* and *Second Tier*. The former measures the ratio of correctly retrieved songs within the top $(C_t - 1)$ matches to $(C_t - 1)$, where C_t is the size of the song class (e.g. the same composition, or performance) for track t . To validate our harmonic segmentation, we measure the overlap between the resulting structures with a collection of well-known chordal patterns. This exemplifies the hypothesis that a good segmentation would maximise the “reuse” of harmonic patterns – as building blocks that can be found in other pieces.

To evaluate our KG embedding methods, we measure link and relation prediction performances using the Mean Reciprocal Rank (MRR) and Hits@K as done in the literature [17]. Both these metrics are computed using the scores of the triples produced by the model and evaluated by the ranking induced from those scores. Hits@K measures the ratio of true triples that are ranked among the top K, whereas the MRR averages the reciprocal ranks of true triples and drops rapidly as the ranks grow. These measures are computed by sampling $N \times 2$ corruptions (negative triples) for each positive: N negatives for head corruptions, by replacing the head with other entities in $|\mathcal{R}|$; and N negatives for tail corruptions, which is analogous to the former case.

To evaluate the influence prediction models, we also used Mean Reciprocal Rank (MRR), along with the Mean Average Precision (MAP) and Discounted Cumulative Gain (DCG). All the listed metrics measure how high are ranked appropriate values, e.g. how high are ranked actual influential artists with respect to a reference artist. In this context, MRR can be interpreted as how far is the first influential artist in the ordered list. MAP is the average

of the number of relevant entries within the first k results, where k is the number of influences of each artist. DCG evaluates the results by penalising when relevant entries are not positioned at the top of the list.

The evaluation of the latter two contributions (embedding and influence models) is detailed in D2.5 [12].

5.3. Results and discussion

Overall, the main findings and contributions of the pilot are summarised below.

- **ChoCo: the Chord Corpus** (*linking music datasets*) is the largest dataset for musical harmony knowledge graphs [9]. It comes with a data workflow to curate, transform, and integrate more than 20,000 human-made, high-quality harmonic annotations from 18 highly heterogeneous chord datasets, following the JAMS data structure as an annotation model. The resulting annotations are rich in provenance data (e.g. metadata of the annotated work, authors of annotations, identifiers, etc.) and refer to both symbolic music notation and audio recordings, while encompassing different notation systems. After semantically enriching, extending, and standardising these annotations under the JAMS definition, we used Polifonia ontologies [18] (and in particular Music Meta [19] and the Music Annotation pattern [20]) to release the ChoCo Knowledge Graph – providing fine-grained semantic descriptions of chords, opportunities for chord interoperability, and 4K+ links to external datasets. ChoCo achieves interoperability of harmonic datasets at three levels: metadata, annotation format, and chord notation. The interoperability at metadata and annotation format levels is implemented by integrating metadata from different sources, at the parsing level, and by leveraging the JAMS annotation standard to store harmonic annotations, consistently. Chord notation interoperability is achieved by converting chords to three reference notational systems – bridging them via the Harte notation. The outcome of this approach enables the use of these integrated collections as if they belonged to the same dataset and underpins the automatic generation of Music Knowledge Graphs. In addition to the conversions, ChoCo provides the original annotations in each JAMS file, along with rich provenance descriptions that keep track of the original sources.
- **Harmory: the Harmonic Memory** (*linking musical patterns*) is a Knowledge Graph of interconnected harmonic patterns, and a family of music models, aimed to support creative applications in a transparent and musically plausible way. We leveraged the Tonal Pitch Space - a cognitive model of Western tonal harmony to project chord progressions into a musically meaningful space. Then, we use novelty-based methods for structural analysis to segment chord sequences into meaningful harmonic structures. The latter are then compared with each other, across all progressions and via harmonic similarity, to reveal common/recurring harmonic patterns. A KG is created to semantically establish relationships between patterns, based on: (i) temporal links, connecting two patterns if they are observed consecutively in the same progression; and (ii) similarity links among highly-similar patterns. By traversing the KG, and moving across patterns via temporal and similarity links, new progressions can be created in a combinational settings; but also, unexpected and surprising relationships can be found among pieces and composers of different genre, style, and historical period. This is also enabled by the scale and diversity of Harmory, which is built from ChoCo, the largest existing collection of harmonic annotations. Currently, Harmory contains 26K harmonic segments from 1800 harmonic annotations (10% of ChoCo, corresponding to all the audio partitions). Out of all segments: 13667 (16%) correspond to the same pattern families, 66175 (53%) are pattern-friendly (they share non-trivial similarities with other segments), whereas 8176 (32%) are inherently unique (they are found in other songs).
- **Parameter-efficient KG embedding methods** (*linking arbitrary knowledge graphs at scale*). Our knowledge graph embedding model demonstrates state-of-the-art performance on relation prediction and competitive results on link prediction tasks across four benchmark datasets (FB15k-237, WN18RR, CoDEX-Large, YAGO 3-10) that are commonly used to evaluate embedding models. Notably, it achieves this using less than one million parameters and consumer-grade hardware. Ablation studies indicate that computing entity embeddings using multiple paths outperforms baseline models. We observe that model performance is optimal on KGs with large vocabulary sizes and diverse relational contexts. This aligns with the model's reliance on unique paths to reconstruct entity identities. Our model's MRR scores correlate positively with certain KG properties, providing insights into its behaviour and how the use of relations influences its effectiveness. The model's parameter

efficiency and inductive nature are particularly well-suited to the dynamic nature of Polifonia’s KGs. Our results demonstrate that (i) relations provide sufficient information to learn effective KG representations; while (ii) leveraging path information results in competitive performance with increased scalability and parameter efficiency. Our key contributions are threefold. Firstly, the model effectively addresses the challenge of dynamically evolving knowledge graphs through its scalability and parameter efficiency. Secondly, the path-based approach to entity embeddings offers advantages over traditional entity embedding tables. Finally, the model’s capabilities are rigorously validated through extensive empirical testing on standard benchmarks.

- **Influence prediction models** (*mining latent connections*) Our influence prediction method prioritises interpretability by treating relationships individually within a graph-theoretic framework. This approach outperforms standard baselines, and the results resonate with socio-cognitive and psychological theories regarding the influence of different relationship types. While our initial attempt to integrate machine learning yielded less accurate predictions, its potential remains. An expanded dataset could improve the performance of machine learning models. Additionally, attention-based architectures may provide benefits if more training data becomes available. Clustering techniques could be used to detect artist communities, revealing artistic influence networks, and integrating method with the similarity-based techniques developed in this pilot may further enhance accuracy and interpretability.

Overall, INTERLINK has developed a rich infrastructure of datasets, resources, and methods contributing to the long-term goal of establishing connections between music knowledge under different lenses.

5.4. Summary of the outputs (FAIR section)

name	ChoCo Project
component-id	ChoCo
type	Project
work-package	WP2
related-components	choco-dataset, choco-kg, choco-software
licence	Various (see each related component)
contributors	Jacopo de Berardinis, Andrea Poltronieri, Albert Meroño Peñuela, Valentina Pre-sutti
link	https://polifonia-project.github.io/ecosystem/pages/pilots/interlink.html
release-version	v1.0.0
doi	https://doi.org/10.5281/zenodo.7706751

name	ChoCo
component-id	choco-dataset
type	Dataset
work-package	WP2
related-components	Choco, choco-kg, choco-software
licence	<i>Attribution 4.0 International</i> (CC BY 4.0) with the exception of data derived from the following subsets: Chordify Annotator Subjectivity Dataset, Mozart Piano Sonata, and Jazz Audio-Aligned Harmony data. The latter are released under the <i>Attribution-NonCommercial-ShareAlike 4.0 International</i> (CC BY-NC-SA 4.0).
contributors	Jacopo de Berardinis, Andrea Poltronieri, Albert Meroño Peñuela, Valentina Pre-sutti
link	https://polifonia-project.github.io/ecosystem/pages/pilots/interlink.html
release-version	v1.0.0
doi	https://doi.org/10.5281/zenodo.7706751

name	ChoCo Knowledge Graph
component-id	choco-kg
type	Dataset
work-package	WP2
related-components	Choco, choco-dataset, choco-software
licence	Follows the same licensing schema of choco-dataset.
contributors	Jacopo de Berardinis, Andrea Poltronieri, Albert Meroño Peñuela, Valentina Pre-sutti
link	https://polifonia-project.github.io/ecosystem/pages/pilots/interlink.html
release-version	v1.0.0
doi	https://doi.org/10.5281/zenodo.7706751

name	ChoCo integration workflow
component-id	choco-software
type	Dataset
work-package	WP2
related-components	Choco, choco-dataset, choco-software
licence	<i>Attribution 4.0 International (CC BY 4.0)</i>
contributors	Jacopo de Berardinis, Andrea Poltronieri, Albert Meroño Peñuela, Valentina Pre-sutti
link	https://polifonia-project.github.io/ecosystem/pages/pilots/interlink.html
release-version	v1.0.0
doi	https://doi.org/10.5281/zenodo.7706751

name	Harmory Knowledge Graph
component-id	Harmory
type	Knowledge Graph
work-package	WP2
related-components	Choco, Harmory Software
licence	Follows the same licensing schema of choco-dataset.
contributors	Jacopo de Berardinis, Andrea Poltronieri, Albert Meroño Peñuela, Valentina Pre-sutti
link	https://polifonia-project.github.io/ecosystem/pages/pilots/interlink.html
release-version	v1.0.0
doi	https://doi.org/10.5281/zenodo.8021211

name	Harmory Software
component-id	harmory-software
type	Software
work-package	WP2
related-components	Harmory, Choco
licence	<i>Attribution 4.0 International (CC BY 4.0)</i>
contributors	Jacopo de Berardinis, Andrea Poltronieri, Albert Meroño Peñuela, Valentina Pre-sutti
link	https://polifonia-project.github.io/ecosystem/pages/pilots/interlink.html
release-version	v1.0.0
doi	https://doi.org/10.5281/zenodo.8021211

6. FACETS

6.1. Overview of the pilot

Music documents have been gathered in collections for ages. With the advent of dedicated formats for representing music documents in a digitised fashion, the need for specific tools to manage collections has risen. Musicians, students, educators, and enthusiasts want seamless exploration of music resources, to find precisely what they are looking for, whether it is to learn new compositions, study music theory, or simply enjoy listening to their favourite tunes. On the other end, cultural heritage institutions have some collections of music documents that they want to expose to end-users, integrating an exploration tool inside their existent applications.

The music content and the metadata of documents should be represented in an adequate manner, so that an exploration engine may properly store and index music documents. Once music information is structured, a *search-engine* may propose some pattern-based and metadata-based search modes. The search results are usually presented in a standard way, with highlighting and counting the pattern occurrences. Then, those results may be enhanced with clustering and hard-coded or dynamic filtering (faceting). Since the content is music, handling the proper visualisation and listening of the content could be another feature. Finally, augmenting the indexed data with external knowledge based through ontologies or with LLM would enable the end-users to have a better experience, not limited to the collection data. Cultural heritage managers are also intermediate users, and they should get some dedicated features. The interested reader is referred to Deliverable D3.2 (section 5) [21] for an in-depth discussion of the motivation and related works in the literature on music search-engines and collection managers.

The FACETS pilot aims to develop such a music collection manager tool, relying on a core search engine, with extended features. FACETS is disseminated through a demo website ¹, and a significant effort has been done to propose a standalone Docker image (see outputs in section 6.4 below), so that the tool may be embedded very easily in an existing cultural heritage platform.

All development work for FACETS is open source and publicly available on GitHub². Collaboration of team members was centralized through GitHub issues³, where ideas were contextualised, organised, linked, timestamped and discussed, which were then transformed into features, planned and marked as completed when necessary.

6.1.1. Connection with research WPs

The pilot contributed to and relies on the technologies developed in:

- WP1: collaboration with other pilot developers was crucial for FACETS, to get preliminary users and feedback. Especially, the team at NUI (Danny Diamond, James McDermott, Abdul Shahid Khattak and Pushkar Jajoria) became users and collaborators. We had fruitful exchanges with the team behind Tonalties (Christophe Guillotel, Marco Gurrieri).
- WP3: FACETS contributed to WP3 activities by enhancing the state of the art methods for representing symbolic music (core of a search engine, back-end) and exchanging around user interfaces for music-oriented digital tools (e.g., Pattern UI from NUI). The frequent meetings in the WP helped develop interesting exchanges and collaborations.

¹See <http://neuma-dev.huma-num.fr>.

²See <https://github.com/polifonia-project/facets-search-engine>.

³See <https://github.com/polifonia-project/facets-search-engine/issues>.

6.1.2. Progress from deliverable D1.5

In Deliverable D1.5 [11], we reported evaluation of FACETS performance, which was done by benchmarking it against other music search techniques. The results of this performance analysis showed that the FACETS search engine core was highly efficient and could handle complex pattern searches in large collections of score with unprecedented speed and memory usage.

At that time, a group of experts from Polifonia Stakeholders network had already been contacted to evaluate other aspects of the pilot (usability, features, etc.), but their feedback had not been yet collected (when the Deliverable was released). Those results have now been gathered, analysed and used (they are presented and discussed below). Working with those experts helped shape the development of FACETS during the second part of the project. A heightened focus on to metadata alignment and user needs were central to the pilot. Notably, the group collected and analysed a list of user searches from the music library of a prestigious institution (Philharmonie de Paris).

A second validation campaign was launched later in 2023 to test FACETS' distribution and reuse. Offering a Docker image and source code is good, but it is always best to check that people can reuse it. Taking advantage of the available Docker image and the open source code, a key developer of FACETS tested the installation and operation of the tool on other people computers, which lead to improvements in documentation efforts and minor changes. A survey was distributed to a new group of people, and feedback collected.

6.2. Evaluation method/rationale

A comprehensive evaluation methodology was employed to ensure a thorough estimation of FACETS' performance and quality. The evaluation process comprised two main parts: speed and memory testing, and quality assessment through expert evaluation and surveys. This second part evaluated the interface of the tool and its distribution modes, while the first part estimated its efficiency.

The technical evaluation methodology was published in [22]. FACETS' core, relying on Elasticsearch⁴, was compared with regular-expression-based pattern search in a large corpus of music scores (14k+), from various sources and belonging to different music genres. Several patterns of different lengths and with different distribution occurrences were sampled, and tests were run to gauge the scalability and speed of the system.

The quality evaluation of FACETS was first run with a focus group of Polifonia stakeholders, *i.e.*, members of institutions close to Polifonia's interests, such as Cité de la Musique/Philharmonie de Paris⁵ (large French public institution), Deezer⁶ (Paris-based streaming platform), Muziekweb⁷ (Netherlands-based musical library) or Podiumkunst⁸ (a network of performing arts). The focus group featured a presentation of the rationale behind FACETS, followed by a demo and then a discussion on all aspects on the tool, to collect qualitative data. Stakeholders were asked a series of precise, pre-defined questions:

- Do you think this tool could be useful for you or your institution?
- Who would be the potential users? How the tool fits with your work practices?
- In what ways could FACETS improve one of your (or your users') process?
- What is your most favourite feature of the interface?
- What is your less favourite feature of the interface?
- What tasks/scenarios would you like to carry out with this tool? Is it possible yet? If not, what's missing?
- Do you have any suggestion(s) for improving the interface?

The second qualitative evaluation round took place a few months later, as a workshop during a Polifonia meeting. Participants were recruited on a voluntary basis. They were allowed to use the FACETS website for 5 minutes before being asked some standardised questions about the interface

⁴<https://www.elastic.co/>

⁵<https://philharmoniedeparis.fr>

⁶<https://www.deezer.com/>

⁷<https://www.muziekweb.nl/>

⁸<https://www.podiumkunst.net/>

- How easy was it to use the tool?
- Were the search options clear and intuitive?
- Did the tool accurately identify the pieces or patterns that you look for?
- Did you find the ranking convincing?
- Did you find the faceted search helpful in refining your search?
- How was the response time during searches?
- How satisfied are you with the overall design?

Sixteen participants have taken part in the qualitative evaluation. 9 participants are music amateurs who have played an instrument or studied music theories, 5 participants are professional musicians, only 2 have no previous music training. Regarding previous experience with content-based music search engine, 10 participants have only used text-based music search engines, while 6 of them have previous experience with content-based ones. For each question in the survey, they were asked to rate their satisfaction or the performance of the tool on a scale of 1 to 5, with 1 indicating the lowest level of satisfaction or poor performance, and 5 a very satisfied experience or excellent performance. Their answers were gathered and analysed, and the results are presented below (Section 6.3.1).

This workshop also helped some users to install FACETS on their computers. Although there was no survey dedicated to this specific part, some feedback was collected and some changes were made to take into account some platform specifics.

6.3. Results and discussion

6.3.1. Results

The plots showing FACETS' performance were included in the D1.5 report [11] (Section 3) and not reproduced here.

The main feedback from the first focus group can be summarised as follow:

- standardization of metadata, alignment. FACETS need a taxonomy (for instruments, genres, etc.). The Music21 library⁹ may help, as well as some Doremus vocabularies¹⁰ ;
- The way users interact with search is evolving, influenced by technologies such as ChatGPT and other chatbot-based interfaces. Users ask questions in natural language and expect search results that go beyond a simple collection to include information from across the web. They also expect to receive some relevant answers to any given query;
- some extended features could be developed, such as fuzzy search (to allow for pattern mistakes), or more harmonic features (chords).

Combining the aforementioned feedback with that from the Polifonia ecosystem (members and reviewers), it guided the direction of FACETS towards becoming a more versatile tool with semantic capabilities. This adaptation aimed at streamlining the integration of knowledge into FACETS, expanding its data and facilitating connections with other tools. Several initiatives were taken to achieve these goals. Firstly, a link was established between FACETS data and knowledge bases such as Wikidata to align composer and period metadata. In addition, in collaboration with the OU, the process of exporting FACETS data in JSON-LD format was initiated.

The feedback from the second survey are displayed in Table 6.1.

While the overall results are encouraging or great (accuracy, response time, and faceted search), they are not excellent, and they call for improvements of the tool. Especially, FACETS is not perceived as easy nor intuitive, with scores below 4.0.

⁹See <https://web.mit.edu/music21/>

¹⁰Doremus, Doing Reusable Data. See <https://data.doremus.org/vocabularies/>

Question	Median score	Average score
How easy was it to use the tool?	4	3.66
Were the search options clear and intuitive?	3	3.40
Did the tool accurately identify the pieces or patterns that you look for?	5	4.56
Did you find the ranking convincing?	4	4.0
Did you find the faceted search helpful in refining your search?	4	4.31
How was the response time during searches?	5	4.31
How satisfied are you with the overall design?	4	3.87

Table 6.1: Survey results.

6.3.2. Discussion, lessons learned

The technical performance of FACETS' core is at the state-of-the-art, and proved capable of handling long and frequent pattern-based searches in large collections: relying on Elasticsearch was a sound technical choice (Deliverable D1.5). Similarly, using GitHub proved beneficial for organizing tasks, enhancing visibility within the open-source community, and facilitating reusability (*e.g.*, by *forking* the project).

However, while the FACETS pilot has led to a technically good tool, there is room for improvement in the user-friendliness and overall design. Involving users in an earlier stage of the design of the tool would have probably helped a lot.

Although the specification of the semantic functionalities for FACETS has been conceptualised, the implementation in the final pilot has only started, primarily due to a lack of manpower. This situation underlines a key lesson from the FACETS development process: while PhD students can benefit from tool development, they often lack specialised programming skills, as may tenured academics. Similarly, hosting a demo website on an external platform comes with a lot of technical management tasks: making sure that environments between development and production are compatible, handling software updates, monitoring downtime, etc. While those tasks may be brief, they are numerous, infrequent, and sometimes complex for a non specialist. It's therefore advisable to allocate resources to hiring dedicated programmers who can address such issues more efficiently and potentially with higher quality.

Lastly, it's important to acknowledge the significant evolution in the realm of "search" or "knowledge retrieval" since November 2022 due to the advent of powerful chat-based tools. Although these generative tools are not (yet) good with exact symbolic data such as music scores (and cannot replace FACETS for content-based centered searches), they have changed users' expectations. Therefore, FACETS may need to adapt, either by refining its interface or by establishing connections with these tools.

6.4. Summary of the outputs (FAIR section)

List all components related to this Pilot

name	FACETS-project
component-id	facets-search-engine
type	Project
work-package	WP3
related-components	facets-docker, facets-website
licence	CC-BY 4.0
contributors	Tianghe Zhu, Raphaël Fournier-S'niehotta, Philippe Rigaux
link	https://github.com/polifonia-project/facets-search-engine/
release-version	1.0
doi	https://doi.org/10.5281/zenodo.7671449

name	FACETS Docker image
component-id	facets-docker
type	Software
work-package	WP3
related-components	facets-docker, facets-search-engine
licence	CC-BY 4.0
contributors	Tiange Zhu, Raphaël Fournier-S'niehotta, Philippe Rigaux
link	https://github.com/polifonia-project/facets-search-engine/
release-version	1.0
doi	https://doi.org/10.5281/zenodo.7671449

name	FACETS-project
component-id	facets-website
type	Project
work-package	WP3
related-components	facets-docker, facets-search-engine
licence	CC-BY 4.0
contributors	Tiange Zhu, Raphaël Fournier-S'niehotta, Philippe Rigaux
link	http://neuma-dev.huma-num.fr/
release-version	1.0
doi	https://doi.org/10.5281/zenodo.7671449

7. TONALITIES

7.1. Overview of the pilot

TONALITIES leverages web technologies to grasp how distinct theoretical viewpoints bring to light different, sometimes conflicting musical properties; confront different interpretations; and, ultimately, provide documented and authored analyses of musical pieces. To this end, *TONALITIES* a) makes use of theoretical models, which b) can be associated manually and/or by means of artificial intelligence with arbitrary selections on the score and c) lead to critical analyses through collaborative approaches.

The focal point of *TONALITIES* is an online collaborative annotation interface for music analysis that addresses the following challenges

- Select different models, corresponding to different theoretical and analytical viewpoints;
- Select any item on the score (roots, groups of notes, etc.) at any level of granularity;
- Associate concepts derived from the models with these analytical elements
- create nested annotations (for instance the tonic function within a cadence), edit a selection or add/remove elements (including other selections);
- Comment on the analytical annotations;
- Compare the annotations made on the same score either by different users or on the basis of different models.

7.1.1. Connection with research WPs

TONALITIES is the result of a strong interaction with WP2. The collaboration between modelling experts and domain experts has led to the modelling of musical theories, and in particular the design of the Zarlino ontology (see below section 7.3.1.1 Theoretical models). As the aim of *TONALITIES* is to combine human annotations and machine annotations, interaction with WP2 has also involved the experimentation and implementation of inference tools. *TONALITIES* was the trigger for the design of the SANDRA reasoner, currently implemented for the modal classification of Renaissance polyphonic works (see section 7.3.5 Validation of results by ground truth below). Finally, the collaboration focused on the alignment between the PON and the *TONALITIES* CIDOC-CRM knowledge graph. The import and export routines implemented allow a complete bidirectional transformation between *TONALITIES* and the PON (ontology modules involved: music representation, music projection, music-meta, music analysis).

WP5 played a key role in the design and validation of *TONALITIES*' interface, based on UX design. The collaboration focused on the visual organisation of the interface, its evaluation by a community of experts and the collection of suggestions for its incremental improvement in two evaluation campaigns.

Less sustained connections also existed with WP3. *TONALITIES* interacted with this work package for the automated detection of cadences and root notes. To date, c. 100,000 root annotations are available in *TONALITIES* knowledge graph (Bach and Terpsichore collections).

7.1.2. Progress from last deliverable D1.6

The following progress has been made since deliverable D1.6[23] (February 2023):

- Number of MEI files made available: 411 => **1.168**;
- Number of theoretical models: 3 => **5**;
- Number of generated triplets with music-analytical information: 34.169 => **2.459.535**;
- Iterations of the satisfaction survey: 1 => **2**;

- Number of participants to the survey: 3 => **27**;
- Evaluation score of the annotation interface : 2,1 => **1,58** (on a scale from 1 = "very satisfied" to 5 = "unsatisfied").

7.2. Evaluation method/rationale

The following validation methods have been applied to *TONALITIES*:

Validation method	KPI
<i>a. Input validation</i>	
1) Quantitative validation of input data	- 5 theoretical models + 4 annotation models - 1000 MEI scores
2) Validation of input data quality by experts	100% compliance with TONALITIES' Score Quality Validation Protocol
<i>b. Validation of tools/means/software</i>	
1) Validation of interface by users/experts	At least score 3 ("moderately satisfied") in satisfaction questionnaire / User scenarios
2) Validation of performance, requests/information passed between front and back end	- SPARQL Query for analytical annotations. No quantitative KPI is associated with this design decision - Shortcuts between each annotation and the score. No quantitative KPI is associated with this design decision
<i>c. Output/goal validation</i>	
1) Validation of results by ground truth	Modal Classification Reasoner, precision: 0.75, recall: 0.75

Table 7.1: Pilot-specific validation protocol

7.3. Results and discussion

7.3.1. Quantitative validation of input data

As shown in **Table 7.2**, the goal of 1000 scores in MEI format has been reached.

These music collections were imported into *TONALITIES* for the specific needs of the project, to test the interface and to support the case study on modal-tonal classification carried out during the development of the pilot. In particular, the Zarlino corpus contains all the compositions quoted by Gioseffo Zarlino (1558), book 4 on the modes. Zarlino's own modal attributions of these works were used as ground truth to test the automated mode inference carried out in *TONALITIES* (see section *Theoretical models* below). Users are free to import any new MEI scores by simply dragging and dropping them, and to annotate them. **The tools created are therefore fully transposable to any MEI score, of any period and of any kind available on the Internet.**

7.3.1.1. Theoretical models

In line with our objectives, five theoretical models have been produced in *TONALITIES*. These models shed light on the modal-tonal classification of monodic and polyphonic works composed between the end of the Middle Ages and the Classical period. Depending on the adequacy between the models and the works, and given the analytical problems to be tackled, their application through the collaborative annotation interface allows different aspects of the structural organisation of the pieces to emerge:

1. Cochlaeus_1511: This ontology models Johannes Cochlaeus' (1511) elementary musical theory, including the definition of concepts relating to the diatonic system (*scala, vox, clavis, littera*, etc.) and musical modes (*tonus*,

Corpus	N° of MEI files	Polifonia Github
Bach/chorals	371	Bach_Chorals
Bach/fugues	1	Bach_Fugues
Dufay	73	Dufay
Josquin/chansons	110	Josquin_Chansons
Mangani/Sabaino	3	Anonyme Fontanelli Lechner
Praetorius/Terpsychore	531	Praetorius_Terpsychore
Zarlino	79	De La Rue De Mantua De Rore Gombert Hellinck Isaac Josquin_Sacred_Music Morales Verdelot Willaert Zarlino
TOTAL	1.168	

Table 7.2: List of *TONALITIES*' corpora and MEI files

finalis, confinalis, ambitus, etc.). The model lets coexist 8-note scale conceptions with 6-note scale conceptions, the latter being derived from solmization practice. This dual perspective is key for charting the conceptual background that underpins musical works of the 15th and 16th centuries.

2. Zarlino_1558: In *Le istituzioni harmoniche* published in 1558, Zarlino introduces a 12-mode theory that supercedes the long-standing octoechos tradition. The careful consideration of the octave species of each mode, the inclusion of the triad as a unifying principle, and the distinction between a natural diatonic system and a diatonic system with one alteration, thoroughly renew the conception of modal coherence.

3. Praetorius_1612_1619: This ontology models the modal theory described by Michael Praetorius (1612 and 1619). This theory stands out from earlier descriptions by a substantial extension of the modal transpositions, which have risen from 15 in Zarlino to 27. This extension, together with the stronger focus on the outer voices (*cantus-bassus*) rather than on the *cantus-tenor* pair for modal identification, provides an important perspective on the tonal organization of polyphonic works composed in the second half of the 17th century.

4. d'Alembert_1752: Jean Le Rond d'Alembert provides a condensed version of the theories presented by Jean-Philippe Rameau in his *Traité de l'harmonie* of 1721 and subsequent works. The modelling focuses on the concepts of scale, fundamental bass, harmonic functions and mode. It reflects a tonal conception based on the modern distinction between major and minor modes, transposed onto the 12 equal tempered degrees of the chromatic scale. The model provides a way to classify works of the common practice based on a historically situated conceptual framework.

5. Powers_1981: Harold Powers' Theory of Tonal Types (1981) proposes a set of ethical concepts – i.e. concepts external to the cultural environment of the music – to classify works according to categories close, but not identical, to "emic" modal categories (resulting from the production environment). The ontology formalises the concepts linked to this "external" point of view and thus projects one of the theoretical perspective adopted in the 20th century to understand modal polyphony.

In addition to these five full-fledged ontologies, with their detailed concept definitions, four lighter annotation models have been provided for the following analytical areas:

- **Cadences (Renaissance)**: annotation of cadences in Renaissance polyphonic works;
- **Fugue**: identifying the constituent parts of a baroque fugue;
- **Chords** - common practice: annotation of root notes and chord inversions;
- **Interpretation**: analysis of performance (articulation, dynamics, intonation, phrasing, tempo, etc.).

7.3.2. Validation of input data quality by experts

The origin of the scores imported in *TONALITIES* is heterogeneous:

- External online score databases (as [ChoralWiki](#) or [Petrucci online](#));
- Score collections created at the [IReMus](#) and published in the [NEUMA](#) score library;
- Scores created in the specific context of *TONALITIES*

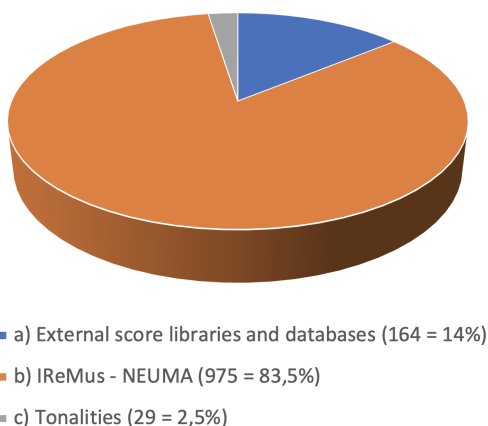


Figure 7.1: Origin of *TONALITIES*' MEI files

All MEI files have been enriched (metadata) and verified (music data) according to the following parameters (see [Table 7.3](#)):

Metadata	Music data
<ul style="list-style-type: none"> - composer - title - encoder/editor - licence - genre 	<ul style="list-style-type: none"> - pitches - rhythmic values

Table 7.3: Enriched and/or verified metadata and data

The metadata in Tonalities' native MEI files has been encoded according to a protocol published on [GitHub](#).

As regards the external musical collections, the metadata is heterogeneous and depends on the encoding practices adopted by the individual projects. We manage this heterogeneity through a dedicated [script](#) for injecting metadata into a homogeneous model (cidoc), on the basis of which we contribute to Polifonia's knowledge graph in [music-meta](#)[19, 18].

All the MEI files are available in *TONALITIES*' GitHub repository (see the third column of [Table 7.2](#)). Each item is covered by a licence. When the constraints of the initial licence allowed it, the CC BY-NC-SA 4.0 licence was systematically adopted in line with Polifonia's requirements.

7.3.3. Validation of the interface by users/experts

The interface has been at the centre of a radical revision till the current V2.1. To date (April 2024), the interface has been adopted by around fifty users from Belgium (2), Canada (2), China (2), France (35), Germany (1), Italy (1), the Netherlands (1) and the USA (1). The current number of analytical annotations is of **222.269**, for a total number of **1.405** analytical projects.

Two evaluation campaigns, based on user scenarios[24][25], have guided its development (see **Table 7.4**).

Satisfaction survey's iterations	Date	Number of participants	User scenarios	Overall evaluation score (1 ("very satisfied") to 5 ("unsatisfied"))
Campaign 1	February 2023	3	- Microstructure: verticalities, roots, selection, nested selection, annotation	2,1
Campaign 2	February 2024	27	- Microstructure: cadences - Macrostructure: fugue - Free user scenario	1,58

Table 7.4: Evaluation campaigns 1 and 2

In the second evaluation campaign, two scenarios were used to evaluate the interface's ability to annotate ergonomically the microstructure (cadences) and macrostructure (fugue) of a score. In addition, a free user scenario has been introduced, which enables the user to carry out a customised analytical task (for instance annotating music-text relations in late 16th century madrigals and motets).

The requests and critical suggestions obtained were subjected to a detailed analysis and then converted into [GitHub issues](#). Of the **139 opened issues**, **107** have been solved and closed at this stage. Beyond fixing obvious bugs, the improvements have focused on the following points:

- Ergonomics
 - Enhanced note selection, with selection contours, for all notes;
 - Multi-element selection across different systems;
 - Shortcut for selecting the whole score;
 - Enhanced search options within score lists (filters);
 - Help menu (documentation and list of keyboard shortcuts).
- Presentation
 - Nested annotations presented according to their hierarchies;
 - Annotation comments displayed.
- Functionalities
 - Deletion of annotations and of whole analytical projects;
 - Application of several analytical concepts to the same selection;
 - Import and export of analytical annotations.

The overall appreciation score of the interface has strongly improved between the first and second campaigns. The current score is of **1,58 on a scale of 5**, the score strongly validates the design, development and user adoption of the interface. As a consequence, a patent request has been submitted on 21 March 2024.

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

We needed to ensure that each analytical annotation made through *TONALITIES* Web application has an author, a date, and belongs to an "analytical project". This has been confirmed with a SPARQL query checking the conformity

of the graph for all the 222.269 annotations provided, as shown hereafter:

```

PREFIX dcterms: <http://purl.org/dc/terms/>
PREFIX crm: <http://www.cidoc-crm.org/cidoc-crm/>
SELECT ?project_name ?annotation ?date ?creator ?orcid WHERE {
GRAPH
<http://data-iremus.huma-num.fr/graph/sherlock> {
    ?project crm:P2_has_type <http://data-iremus.huma-num.fr/id/21816195-6708-4bbd-
        a758-ee354bb84900> .
    ?project crm:P1_is_identified_by ?project_name .
    ?project crm:P9_consists_of ?annotation .
    ?annotation dcterms:created ?date .
    ?annotation dcterms:creator ?creator .
    OPTIONAL {
        GRAPH <http://data-iremus.huma-num.fr/graph/users> {
            ?creator crm:P1_is_identified_by ?orcid_appellation .
            ?orcid_appellation crm:P2_has_type <http://data-iremus.huma-num.fr/id/73ea8d74-
                3526-4f6a-8830-dd369795650d>. #ORCID NAME IDENTIFIER
            ?orcid_appellation crm:P190_has_symbolic_content ?orcid
        }
    }
}
}
}

```

Endpoint: <https://data-iremus.huma-num.fr/sparql>

Yasgui: <https://data-iremus.huma-num.fr/sherlock/yasgui>

7.3.5. Validation of results by ground truth

An important objective of *TONALITIES* is to exploit low-level annotations to incrementally infer higher-level musical knowledge. To this end, *TONALITIES* relies on a) the concept definitions provided in the theoretical models and b) the annotations, available in the knowledge graph. This process is demonstrated through the modal classification of polyphonic works from the Renaissance, according to Zarlino's modal theory (see 7.3.1.1).

Zarlino describes twelve authentic and plagal modes with their transpositions, leading to a total of 27 possible classifications for a musical work. These modes are illustrated in the theoretical source by around 80 compositions from the 15th and 16th centuries, including works by Josquin Desprez, Nicolas Gombert, Pierre de la Rue, Cipriano de Rore, Adrian Willaert and Zarlino himself (see **Table 7.2**). Our strategy is to compare the modal attributions made by Zarlino with those inferred from *TONALITIES*. Because of the ambiguous nature of music and of the possible distance between theory and practice, a major challenge is to infer the best hypothesis (or hypotheses) when all premises are not met. Another challenge is to make explicit the criteria on the basis of which a modal attribution has been made and the criteria that are not respected in the work under consideration.

7.3.5.1. Reasoning on the ontology

One possible way of implementing the modal classification task in OWL2 is through the use of SWRL [26]. While it is possible to implement all the required rules by this means, we are faced with three main issues. Firstly, the formalisation process is difficult and requires a tight collaboration between a logic expert and a domain expert. Secondly, the inherent limitations of first order logic (FOL), namely the undecidability, complicates the reasoning process: it has been proved that it is not possible to always obtain an answer from FOL systems. Finally, when following good modeling practices, it is natural to define disjoint concepts. While disjointness is not an issue when relying on SWRL, it prevents the inference of modal attribution that are only *partially* possible. This is however an

important requirement, as it is not always possible to faithfully represent music theories, such as in the case of Zarlino, due to their inherent ambiguity. For these reasons, we rely on a reasoning method, distinct from SWRL, called SANDRA [27]. Sandra is a neuro-symbolic reasoner that is able to infer all the possible perspectives from which an observation can be interpreted. In the context of the Zarlino ontology, it is possible to define rules that are directly used by SANDRA to infer the modal attribution of a score. Rules are formalised as premises and conclusions. The set of premises are identified during the annotation phase, while the conclusions are inferred during the reasoning process. Sandra provides a probability of a modal attribution, which can be interpreted as the degree of confidence of that attribution, alongside an explanation of the performed inference. The explanation identifies which premises are fulfilled by the annotation and which ones are missing, given the rule.

7.3.5.2. Human-machine interaction

The knowledge production process implemented in *TONALITIES* is based on a virtuous circle involving human annotation, critical analysis and machine annotation. This cycle is represented in **Figure 7.2** below.

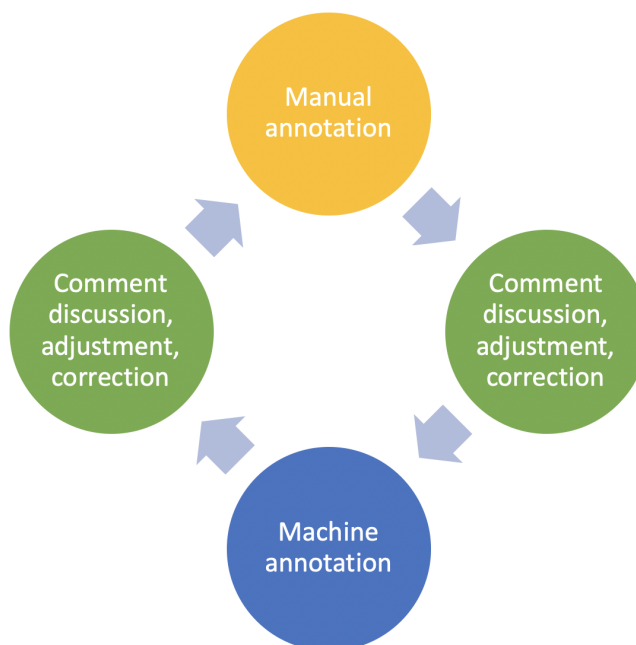


Figure 7.2: Human-machine interaction in Tonalities

In the use case under consideration, the musical mode of a work is inferred from the manual annotation of cadences. This information is then enhanced by automatic extraction of other parameters such as vocal ranges, initial notes and final notes. A set of rules extracted from the Zarlino ontology then serves as premises for identifying modal categories with the reasoner according to the low-level observations available. The modal inferences are fed back into the interface and the expert analyst may then validate, correct and/or comment on these annotations. The analyst may also choose to enrich or correct the initial data to explore how these changes affect the inference. At the current stage, automated inference and subsequent injection of machine annotations cannot be directly launched from the annotation interface by the user. An implementation of these functionalities is planned after the end of Polifonia in a follow-up project. We also plan to extend the inference of musical concepts beyond the use case on modal classification carried out in Polifonia (see below 7.4). Through its iterative and dialectical approach, the pilot goes beyond knowledge production by means of artificial intelligence and bypasses a strictly deductive process that quickly reaches its limits when it comes to music semantics [28]. An important contribution of *TONALITIES* thus

lies in its ability to combine manual and machine annotation in a coherent environment, suitable for the incremental production of contextualised knowledge about music.

7.3.5.3. Results

The inference currently achieves a precision of 0.56 and a recall of 0.73. These apparently weak results do not, however, invalidate the overall approach. They reflect strong latent tensions between theory and practice at the time of Zarlino. These tensions are due in particular to two phenomena. 1. The differentiation between authentic and plagal variants of the same mode (e.g. Dorian and Hypodorian modes 1 and 2), as used in theory, is partly disconnected from musical practice [29]. If this distinction is suspended in favour of six mixed authentic-plagal categories with their transpositions, the precision rises to 0.73 and the recall to 0.77. 2. The reliability of the attributions fluctuates significantly due to the different modal classes (see **Table 7.5**).

Mode	Precision	Recall
1	0.78	0.78
2	0.41	0.54
3	0.18	0.29
4	0.00	0.00
5	0.50	0.50
6		
7	1.00	1.00
8	0.00	0.00
9	0.92	0.92
10	1.00	1.00
11	0.88	1.00
12	0.69	0.90

Table 7.5: Modal inference: Precision and recall according to modal categories

While the Aeolian (modes 9 and 10) and Ionian (modes 11 and 12) modes achieve high scores, the Phrygian mode (modes 3 and 4) is identified with very high reliability. Rather than calling into question the reasoner's inference, the overall results in fact conceal considerable heterogeneity in the fit between individual modal categories and musical practice. In this context, it is worth remembering that the best-recognised modes - the Ionian and Aeolian modes - evolved at the turn of the 16th and 17th centuries into the major and minor modes of harmonic tonality.

7.4. Future perspectives

TONALITIES' objectives, as defined within Polifonia, can be considered as being achieved or even outperformed (number of scores and models, evaluation score of the interface). In particular, the collaborative annotation interface, which is now stable and tested by the scientific community, has been the subject of a patent application (see 7.3.3). Given its success, the needs of the laboratory and the demands of the user community, the development of the *TONALITIES'* interface will continue over the coming months and years funding applications have already been submitted for this purpose. The following features are planned or in the process of being implemented:

- Interface for managing scores and analytical projects (in the process of being implemented, scheduled for May 2024);
- Enable users to launch, directly from the interface, automated inference of concepts by machine learning and/or rigorous formalisation (see 7.3.5.2, scheduled for June 2024);
- Autonomous design and import of annotation models by users (Summer 2024);

- Adaptive and selective display of annotations and labels on the score: filters, search options, hierarchies, free comments, etc. (starting in Autumn 2024);
- Autonomous import of annotation sets (starting in Autumn 2024);
- Score synchronisation with audio sources (starting in Autumn 2024).

7.5. Summary of the outputs (FAIR section)

List all components related to this Pilot

name	Tonalities
component-id	tonalities-app
type	Application
work-package	WP1
licence	Apache-2.0
contributors	- Thomas Bottini (IReMus) - Christophe Guillotel-Nothmann (IReMus) - Marco Gurrieri (IReMus) - Antoine LeBrun (IReMus) - Félix Pouillet-Pagès (IReMus)
link	https://github.com/polifonia-project/external-components/tree/main/components
release-version	V2.1
doi	https://doi.org/10.5281/zenodo.7659992

name	Tonalities' Corpora
component-id	tonalities-scores
type	Corpus
work-package	WP1
licence	- CHORAL PUBLIC DOMAIN LIBRARY LICENSE - CC BY-NC 4.0 - CC BY-NC 2.5 - CC-BY-SA 4.0
contributors	- Margaret Greentree - Christophe Guillotel-Nothmann (IReMus) - Marco Gurrieri (IReMus) - <i>et al.</i>
link	https://github.com/polifonia-project/tonalities_pilot/tree/main/scores
release-version	V5.1
doi	https://doi.org/10.5281/zenodo.10940069

name	Tonalities annotations
component-id	tonalities-annotations
type	Dataset
work-package	WP1 and WP2
licence	- CC BY-NC-SA 4.0
contributors	- Adam Filaber (McGill University - IReMus) - Christophe Guillotel-Nothmann (IReMus) - Marco Gurrieri (IReMus) - <i>et al.</i>
link	https://github.com/polifonia-project/tonalities_pilot/tree/main/annotations
doi	https://doi.org/10.5281/zenodo.10940035

name	Tonalities Annotation Models
component-id	tonalities-annotationModels
type	Ontology
work-package	WP2
licence	- CC BY-NC-SA 4.0
contributors	- Adam Filaber (McGill University - IReMus) - Christophe Guillotel-Nothmann (IReMus) - Marco Gurrieri (IReMus)
link	https://github.com/polifonia-project/music-analysis-ontology/tree/main/annotationModels
release-version	2.1
doi	https://doi.org/10.5281/zenodo.10940196

name	Tonalities Theoretical Models
component-id	tonalities-theoreticalModels
type	Ontology
work-package	WP2
licence	- CC BY-NC-SA 4.0
contributors	- Christophe Guillotel-Nothmann (IReMus) [d'Alembert, Cochlaeus, Powers, Praetorius, Zarlino] - Marco Gurrieri (IReMus) [Zarlino] - Nicolas Lazzari (Università degli Studi di Bologna) [Zarlino] - <i>et al.</i>
link	https://github.com/polifonia-project/music-analysis-ontology/tree/main/theoreticalModels
release-version	2.1
doi	https://doi.org/10.5281/zenodo.10940238

name	Tonalities Meihead Parser
component-id	tonalities-meiheadParser
type	Script
work-package	WP2
related-components	tonalities-scores
licence	- CC BY-NC-SA 4.0
contributors	- Thomas Bottini (IReMus)
link	https://github.com/polifonia-project/tonalities_pilot/tree/main/scripts/meihead-parser
release-version	1.0
doi	https://doi.org/10.5281/zenodo.10974021

name	Tonalities Alignment (Import/Export)
component-id	tonalities-scriptAlignment-import-export
type	Script
work-package	WP2
related-components	tonalities-scores - tonalities-theoreticalModels - tonalities-annotationModels
licence	- CC BY-NC-SA 4.0
contributors	- Christophe Guillotel-Nothmann (IReMus)
link	https://github.com/polifonia-project/tonalities_pilot/tree/main/scripts/alignment-import-export
release-version	1.0
doi	https://doi.org/10.5281/zenodo.10974193

8. TUNES

8.1. Overview of the pilot

The TUNES pilot focusses on influences between oral and semi-oral music traditions over centuries. 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 has interlinked the entire melody collection of the Meertens Institute with a large number of other European collections. The linked melodic data sets are highly valuable for musicologists and music historians interested in cultural evolution of musical style and in oral transmission and variation.

8.1.1. Connection with research WPs

The TUNES pilot has collaborated with workpackages 2 and 3, WP2 for ontology design and interlinking of data sets, and WP3 for modelling musical contents.

Workpackage 2 The development of the Tunes and MusicMeta modules of the Polifonia Ontology Network (PON)¹ was a result of intensive collaboration with the WP2 team. Through iterative refinement, the ontology evolved into its current state, which is able to capture the complex nature of metadata and the relations between musical compositions and tunes that were collected from oral culture. The interaction between TUNES and WP2 was crucial for this result. The pilot provided the domain knowledge, and WP2 the necessary skills of ontology design.

Workpackage 3 The modelling of musical contents and metadata in TUNES has been aligned with the activities in WP3. The MusicMeta module of the PON has been adopted by WP3. WP3 also made use of the TUNES Knowledge Graph for representing metadata on tunes. The motif-based analysis method as been developed in WP3 has been applied to one of the core datasets of the TUNES pilot, MTC-ANN.

8.1.2. Progress from last deliverable D1.6

The following results have been achieved since the previous validation report:

1. Metadata for 2.3 million incipits of the Répertoire International des Sources Musicales (RISM) have been added to the TUNES knowledge graph. This spans a substantial part of all known sources of composed music in European music history.
2. Metadata for the MTC-ANN-2.0.1 collection has been added to the knowledge graph. This is a small, richly annotated, dataset (360 songs) that has been used in various studies, and has been used in WP3.
3. Metadata for the Kolberg collection has been added. This is a dataset containing the digitization of c. 20k Polish folk songs as collected by Oskar Kolberg and his successors.
4. Metadata for sources of the tunes have been added. These can be manuscripts, printed editions, or audio recordings. The RISM, MTC, CRE, and Kolberg collections contain information on sources.
5. Similarity relations for melodies across the various collections have been computed and added to the knowledge graph.
6. Images with music notation (scores) have been generated and added for all melodies.
7. The Tunes and MusicMeta modules of the Polifonia Ontology Network have been released.
8. A new version of the knowledge graph has been published that is fully aligned with the Polifonia Ontology Network.

¹<https://github.com/polifonia-project/ontology-network>

9. A set of queries has been made available which can be used to add data to the knowledge graph.

The following data sets have been integrated into the TUNES Knowledge Graph:

Répertoire International des Sources Musicales (RISM) c. 2.3 million Incipits for c. 1.8 million works in c. 1.2 million sources. The RISM is an index for written musical sources in Europe, mainly manuscripts and printed editions of ‘classical’ music. For a large part of the musical works, an incipit is encoded, which consists of the musical contents of the first few bars of the piece or movement. Also data on persons and institutions has been extracted and stored in the knowledge graph. These include publishers, scribes, composers, lyricists, etc.²

Meertens Tune Collections – MTC-FS-INST-2.0 (MTC) (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.³

Meertens Tune Collections – MTC-ANN-2.0.1 (Dutch) 360 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.⁴

ESAC Folksong Databases – (ESSEN) (mostly German) c. 8,300 tunes. The official web page of the Essen collection refers to the KernScores library for the .krn versions. 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.⁵

The Session Corpus (SES) (Irish) more than 40,000 tunes (mostly Irish, Scottish and English). The abc encoding of the corpus was taken from the folk_ngram_analysis repository, which is being maintained by the WP3 team. 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.⁶

Ceol Rince na hÉireann corpus (CRE) (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 team. 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.

The Kolberg Collection (KOL) c. 20.000 melodies. A digital collection of melodies (instrumental songs and vocal melodies) from Oskar Kolberg’s Works, covering over 20,000 records. Kolberg (1814–1890) and his successors collected Polish folk melodies. The metadata have been extracted and included in the TUNES Knowledge Graph.⁷

Early American Secular Music and Its European Sources, 1589–1839: An Index containing c. 66.000 Incipits and corresponding 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. The music contents of the incipits have not been parsed.⁸

8.2. Evaluation method/rationale

In the evaluation, we focus on the affordance of the similarity relations that are included in the Knowledge Graph. These connect highly similar melodies within and between collections. Especially the similarity relations of melodies

²<https://rism.info>

³<https://www.liederenbank.nl/mtc>

⁴<https://www.liederenbank.nl/mtc>

⁵<http://www.esac-data.org>, <https://kern.humdrum.org/cgi-bin/browse?l=/essen>

⁶<https://thesession.org>, <https://github.com/adactio/TheSession-data/tree/4fdaaf16f921669ce3c7a0a60248176c4e774582>

⁷<https://kolberg.ispan.pl>

⁸<https://www.cdss.org/elibrary/Easmes/index.html>

from different collections are relevant, as these indicate a geographic distribution of the tune.

To evaluate the similarity relations, we provided the collection specialist of the Meertens Institute with lists of highly similar cross-collection pairs of melodies. We also provided her for each melody a ranked list containing the 100 most similar melodies across all collections, ordered according to similarity value. The similarity values are computed by an existing melodic similarity algorithm that makes use of sequence alignment. The details of this algorithm are included in a conference paper that we presented at the Digital Libraries for Musicology Conference in 2023.⁹ The performance of the algorithm also has been covered in Deliverable 1.6,

We asked the collection specialist to examine these lists to see whether they provide previously unknown information that can be incorporated in the Dutch Song Database.

The Tunes and MusicMeta ontologies and the Tunes Knowledge Graph are evaluated by congruence with the competency questions as identified in the Polifonia stories.

8.3. Results and discussion

8.3.1. Similarity Relations

Table 8.1 shows the number of highly similar pairs for all pairs of collections.

Pair	Count
RISM-MTC	36,901
RISM-ESSEN	17,227
RISM - SES	16,558
RISM - CRE	472
RISM - KOL	53,832
MTC - MTC	16,424
MTC - ESSEN	2,108
MTC - SES	285
MTC - CRE	0
MTC - KOL	301
ESSEN - ESSEN	2,753
ESSEN - SES	37
ESSEN - CRE	0
SES - SES	32,594
SES - CRE	2,057
SES - KOL	48
CRE - CRE	21
CRE - KOL	2
KOL - KOL	16,178

Table 8.1: Counts of highly similar pairs between collections.

It is not surprising to observe a fair overlap between MTC and ESSEN, since the German and Dutch folk music is highly related. It is also not surprising to see a large overlap between SES and CRE, since both focus on melodies from Irish traditions. Interestingly, there are also similar pairs between MTC and SES, and between ESSEN and SES. Both MTC and ESSEN have no melodies in common with CRE.

The collection specialist was able to incorporate dozens of corrections and additions into the Dutch Song Database based on manual inspection of these lists. Examples are provided in the DLfM paper.

⁹REFERENCE

The similarities between MTC and RISM were computed and added to the knowledge graph after submitting the DLfM paper. The collection specialist went over all highly similar pairs. Since the RISM only contains incipits, which are much shorter than full melodies, the results are more noisy: there are more false positives. Especially shorter incipits only containing a few notes tend to match a wide range of unrelated melodies. On the contrary, most of the matches of Dutch melodies with longer incipits comprise true positives. The presence of composer information in the RISM data is important. This could reveal the composership of tunes that appear anonymously in Dutch sources. As the reported count of MTC-RISM pairs in Table 8.1 indicates, there are many correspondences. It appeared that many of the those had been found earlier by using the RISM search engine. Now we have exhaustively linked the data sets, we have a full overview. The new findings are summarized in the following overview.

Song ID	Finding
	Tune Family Identification
138212	Daar waren twee koningskinderen 4
76077	Hier auf diesem Rasensitze
128234	A-section corresponds with: Oui je t'aime l'amour même
181022	Häsleins Klage
182338	Clé du Caveau
190533	Qual vive salamandra
192105	Clé du caveau
193449	Liebster Immanuel
175816	Qu'ils sont doux bouteille mamie
176389	Jadis un célèbre empereur
176386	Gavotte de Vestris
	Composer Information Added
113455	Christian Gotlob Neefe
125423	Allegretto by Charles-Auguste de Bériot
125757	Carl Gottlieb Reissiger
125849	Russische polka by Adam Rupp
126297	Frühlingsgruss by B. Klein
140990	Lachcanon by Luigi Cherubini
167134	Carl Ditters von Dittersdorf
167458	Canon by Michael Praetorius (?)
175932	Ignace Pleyel
176181	Song by Friedrich Heinrich Himmel
177037	Presto by Ignace Pleyel
177038	Rondo by Ignace Pleyel
177042	Rondo by Ignace Pleyel
177084	Duet by Ignace Pleyel
177143	Andante by Ignace Pleyel
177146	Rondo by Ignace Pleyel
180515	Air from an opera by Jean-Jacques Rousseau
181168	Song by Friedrich Burchard Beneken
198208	Menuet by Friedrich Schwindel
198209	Trio by Friedrich Schwindel
199745	Rondo by Ignace Pleyel
	Additional information found
177663	Georg Philipp Telemann used this melody
181078	German origin. Arrangement by Friedrich Silcher
183218	English title added

8.3.2. Ontology and Knowledge Graph

In close collaboration with the WP2 team, an ontology module has been developed specifically for Tunes. The work on this has been reported in a conference paper for the International Society for Music Information Retrieval Conference, 2023.¹⁰

From the extracted data a knowledge graph has been constructed, which is fully aligned with the MusicMeta ontology. The KG is accessible from the SPARQL endpoint at the Polifonia server.¹¹ The following table shows for each of the competency questions (CQs) from the relevant user stories,¹² which of those competency questions can be addressed with this version of the knowledge graph.

Competency Questions	Covered Feb. 2023	Covered Apr. 2024
Story: Mark#1_FolkMusic		
CQ1: Is a composer known for composition X?	N	Y
CQ2: What is the name of the composer specified in the source of this composition?	N	Y
CQ3: What is the similarity between compositions X and Y given similarity measure Z?	N	Y
CQ4: Which tunes are similar to tune X given similarity measure Y?	Y	Y
CQ5: Who (which source) attributed composition X to composer Y?	N	Y
CQ6: Which are all known concordances (same composition/tune in another source)?	N	Y
CQ7: Which concordances of composition X have a composer name associated?	N	Y
CQ8: What is the geographic origin of source X?	Y	Y
CQ9: Who were the owners of (manuscript) source X?	N	Y
CQ10: Who was/were the scribe(s) of (manuscript) source X?	N	Y
CQ11: What was the repertoire of scribe X (i.e. all compositions written down by X)?	N	Y
CQ12: Who was the publisher of (printed) source X?	Y	Y
CQ13: What is publication year of printed source X?	Y	Y
CQ14: Which are all compositions that are in source X?	Y	Y
CQ15: What is current location of source X?	N	Y
CQ16: Where to find a digital scan of source X (url)?	N	Y
CQ17: What is the title of composition X in source Y?	Y	Y
CQ18: What is the tune indication of composition X in source Y?	N	N
CQ19: On what page (or folio) is composition X in source Y?	Y	Y
CQ20: What is the serial number of composition X in source Y?	Y	Y
CQ21: What printed source shares content with manuscript X?	N	Y
CQ22: What is the language of the lyrics of tune X?	N	N
CQ23: Has composition X been identified as variant in a tune family?	Y	Y
CQ24: Which tune family does composition X belong to?	Y	Y
CQ25: Who assigned composition X to tune family Y?	N	N
CQ26: With what level of confidence is composition X a variant in tune family Y?	N	Y

¹⁰Carriero, V., de Berardinis, J., Meroño-Peñuela, A., Poltronieri, A., & Presutti, V. (2023, November). The Music Meta Ontology: A Flexible Semantic Model for the Interoperability of Music Metadata. In *Proceedings of the 24th International Society for Music Information Retrieval Conference*, Milano, Italy.

¹¹<https://polifonia.disi.unibo.it/organs/tunes>

¹²<https://github.com/polifonia-project/stories>

CQ27: What are all compositions in tune family X?	Y	Y
CQ28: What are the similarities / differences of all compositions in tune family X according to measure Y?	N	N
CQ29: To what tune families is tune family X related, given similarity measure Y?	N	Y
CQ30: What are alternative titles for composition X?	Y	Y
CQ31: What are the differences / similarities between two corpora of compositions concerning features Y1..Yn?	N	Y
CQ32: What are longitudinal differences / similarities within a corpus concerning features Y1..Yn?	Y	Y
CQ33: What are the differences / similarities between two corpora of compositions concerning occurrences of patterns?	(WP3)	(WP3)
CQ34: What are longitudinal differences / similarities within a corpus concerning occurrences of patterns?	(WP3)	(WP3)
CQ35: What patterns do the compositions in corpus X share?	(WP3)	(WP3)
CQ36: What patterns are overrepresented in corpus X compared to corpus Y?	(WP3)	(WP3)
Story: Brendan#1_FindTraditionalMusic		
CQ1: What tunes have similar geographic origin as tune X?	N	Y
CQ2: What tunes are similar to tune X, given similarity measure Y?	Y	Y
CQ3: Given a set of tunes, from which collections are these tunes?	Y	Y
CQ4: Given a set of tunes, what tunes are from collection X?	Y	Y
CQ5: What are the metadata for collection X?	N	Y

As can be observed in the table, virtually all competency questions have been addressed. The questions concerning occurrences of melodic patterns are covered in separate work that has been done in WP3. The few questions that are not covered are not of crucial importance for studying

8.4. Summary of the outputs (FAIR section)

name	Tunes Knowledge Graph
component-id	tunes-knowledge-graph
type	KnowledgeGraph
work-package	WP1
related-components	Mark#1_FolkMusic, Brendan#1_FindTraditionalMusic
licence	CC-BY
contributors	Peter van Kranenburg
link	https://github.com/polifonia-project/tunes-knowledge-graph
release-version	v1.0

name	Pitchcontext
component-id	pitchcontext
type	SoftwareLibrary
work-package	WP1
related-components	
licence	MIT
contributors	Peter van Kranenburg
link	https://github.com/polifonia-project/pitchcontext
release-version	0.1.9
doi	https://doi.org/10.5281/zenodo.8020644

9. MUSICBO

9.1. Overview of the pilot

Music has held a pivotal position in the history of the city of Bologna. From the renaissance over the centuries, the city garnered recognition from European intellectuals who lauded Bologna as a musical hub. In 2006, UNESCO designated Bologna as a "City of Music," underscoring its significance in the global musical landscape. Despite this acclaim, Bologna's musical heritage remains underappreciated relative to its potential partly because of the sparsity and non homogeneity of information. The MusicBo pilot is designed with the aim of filling this gap by extracting knowledge from a digital multilingual corpus representative of different discursive genres (manuals, biographies, epistles, etc.) thereby ensuring representativeness of information. This corpus forms the basis for a knowledge graph, facilitating the interlinking and publication of extracted knowledge as linked open data, thereby enhancing outreach to a wider audience.

9.1.1. Connection with research WPs

The pilot contributed on the activities of the following WPs:

- WP1. MusicBo contributed to WP1 activities i) by collecting requirements according to the Pilot's research needs, providing a complete panorama of the sources and to list the available digital corpora or datasets about musical cultural heritage in Bologna ii) to identify the software processes and artefacts that will make the data meaningful for users iii) to focus on the Socio-pedagogical aspect of the pilots, identifying the target groups, and their applications and functions of the technology developed. MusicBo pilot provides data for the stories represented in MELODY portal ¹, the cornerstone of the dissemination and valorisation of the knowledge produced within the project.
- WP2. MusicBo contributed to WP2 activities by providing an unlabeled dataset on which the influence prediction method is tested and the results are validated qualitatively. The data extracted from MusicBo is automatically enriched with novel information providing insights that are impossible to identify without converting the documents into a Knowledge Graph representation.
- WP4. MusicBo contributed to WP4 activities by providing a textual corpus to the wider Polifonia Textual Corpus ², a large-scale, multilingual and multigenre diachronic textual corpus described in Deliverable 4.1 [30] of WP4. The MusicBo corpus contains 137 texts in 4 languages (Italian, English, French, and Spanish) published between 1700 to the current era ³. It encompasses documents covering diverse textual genres ranging from historical to critical essays, media, correspondences and commercial documents. Expert musicologists and linguists collected the corpus with the intention of investigating the role that music played in the city of Bologna through a historical perspective, uncovering performances and encounters between key figures.

MusicBo corpus has been annotated through a Natural Language Processing (NLP) pipeline encompassing tokenization, part-of-speech (PoS) tagging, word-sense disambiguation, named entity recognition (NERC) and entity linking (EL) extensively described in [31]. The annotated corpus can be explored through a web application ⁴ specifically designed for meeting the requirements of linguists and terminology scholars.

¹https://projects.dharc.unibo.it/melody/musicbo/music_in_bologna_knowledge_graph_overview.

²<https://github.com/polifonia-project/Polifonia-Corpus>

³Due to copyright reasons, the documents of MusicBo corpus cannot be entirely disclosed. We release metadata sufficient for the reproduction of the corpus <https://doi.org/10.5281/zenodo.6672165>.

⁴<https://polifonia.disi.unibo.it/corpus/>

MusicBo leverages Text2AMR2FRED⁵ [32], the text-to-Knowledge Graph pipeline developed by WP4 and extensively described in Deliverable 4.6 [33], which is an enhanced version of FRED [34], to extract structured knowledge from its collection of textual documents. MusicBo's corpus, describing the role of Musical Heritage (MH) in the city of Bologna, is converted into an OWL-compliant RDF Knowledge Graph (KG). The KG is publicly accessible through a SPARQL endpoint⁶ enabling the creation of visual data stories⁷ using MELODY.

Table 9.1: Statistics describing the KG resulting from the application of Text2AMR2FRED to MusicBo corpus.

Language	#(sent, AMR graph) pairs	#(filtered sent, AMR graph) pairs	# triples
EN	51.814	5.798	412.911
ITA	10.563	1.759	118.162
Overall	62.377	7.557	531.073

The text-to-KG pipeline only includes MusicBo corpus' documents in English and in Italian (respectively 47 and 51). The initial formats of these documents encompassed *.pdf*, *images*, and *.docx*. We extract plain text from them through customised Optical Character Recognition (OCR) technologies⁸. Subsequently, we carry out co-reference resolution⁹, rule-based minimal post-OCR corrections¹⁰, and sentence splitting on the extracted plain texts. Following this pre-processing stage, we submit the processed sentences to neural models (SPRING for English [35] and USeA for Italian [36]) for text2AMR parsing.

AMR graphs, anchored to PropBank *Frames*¹¹, function as an event-centric representation of the MusicBo corpus' sentences, suited for extracting 'who-did-what-to-whom' information from a text. Through the application of the AMR2FRED tool¹² [37], accessible via the Machine Reading suite¹³, we transform AMR graphs into full-fledged RDF/OWL KGs aligned with FRED's theoretical framework. The outcome is a series of *named graphs*, enabling the tracking of each triple to its originating sentence in the corpus. We enrich the resulting KGs through Framester [38], which allows the alignment with external Knowledge Bases (KBs) such as DBPedia, Wikidata and VerbAtlas¹⁴.

For instance, consider the following triples¹⁵:

```

fred:Barbaja a amr:Person ;
    owl:sameAs dbpedia:Domenico_Barbaia ,
    wd:Q908235 .

fred:offer_1 a pbdata:offer-01 ;
    pblr:benefactive_or_entity_offered_to fred:Rossini ;
    pblr:commodity fred:engagement_1 ;
    pblr:entity_offering fred:Barbaja ;
    fschema:subsumedUnder va:0229f ,
    fnframe:Offering .

fred:Rossini a amr:Person ;
    
```

⁵<https://arco.istc.cnr.it/txt-amr-fred/>

⁶<https://polifonia.disi.unibo.it/musicbopilot/query>

⁷https://projects.dharc.unibo.it/melody/musicbo/music_in_bologna_knowledge_graph_overview

⁸<https://github.com/polifonia-project/textual-corpus-population>

⁹For English language documents, we implemented a co-reference resolution pipeline based on Spacy's *neuralcoref* (<https://spacy.io/universe/project/neuralcoref>). We are currently evaluating tools for Italian.

¹⁰<https://github.com/polifonia-project/rulebased-postocr-corrector>

¹¹PropBank Frames are the core lexicon of the PropBank paradigm and consist of predicate-argument structures named "rolesets".

¹²<https://github.com/polifonia-project/amr2Fred>

¹³<https://github.com/polifonia-project/machine-reading>

¹⁴<https://www.dbpedia.org/>, <https://www.wikidata.org/>, <https://verbatlas.org/>

¹⁵Extrapolated from the KG originating from the sentence "In the year 1814, Barbaja went to Bologna and offered Rossini a better engagement than before.", taken from the MusicBo corpus document *The Life of Rossini (Edwards, 1869)*, available at: <https://freeditorial.com/en/books/filter-author/henry-sutherland-edwards>

```
owl:sameAs dbpedia:Gioachino_Rossini ,
wd:Q9726 .
```

The reported triples encode the event of an engagement offer delivered from Domenico Barbaja, an opera manager, to the composer Gioachino Rossini¹⁶. Such knowledge is what scholars who supported the corpus collection aimed to disclose automatically and at scale from the original documents. Scholars can independently leverage such knowledge encoded in the KG and create data stories through MELODY, such as the one created by University of Bologna students¹⁷, focusing on the representation of Russian composers and classical music in the MusicBo corpus.

9.1.2. Progress from last deliverable D.1.6

Regarding the Corpus query tool from a linguistic point of view compared to the previous deliverable, the possibility of performing more detailed searches within the context has been added (c.f. Figure 9.1). As a result of brainstorming meetings with expert musicologists, it became necessary to add a filter in order to enable more in-depth queries. For example, searching for the word “symphony” yields 143 sentences in which this appears, but only in 8 of these does the word “orchestra” also appear. The ability to filter context makes the tool more powerful with regard to specific searches.

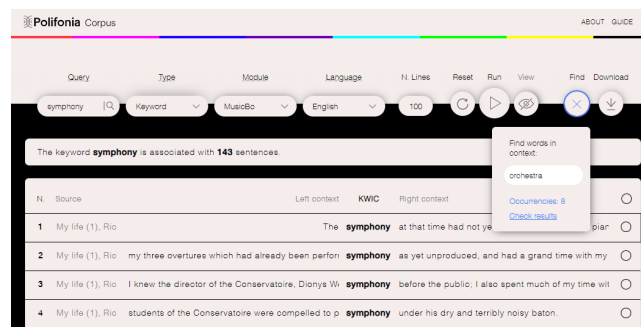


Figure 9.1: Find in context

Regarding the MusicBo KG and given the amount of noise extracted from the documents (c.f. Evaluation section), a pipeline to reduce the amount of noise and automatically enrich the data has been designed. In particular, Frames that describe relationship between two entities, which might be extracted from documents that mention two persons or a person and a place, are extracted by querying the KG. We computed a binary projection of each frame - i.e. a binary relation is created between the two entities - and a synthetic name is generated. The synthetic name is further processed to obtain a label that is more human-interpretable. This is done by leveraging an open-source Large Language Model (LLM), Mistral 7B [40], through few-shot prompting. By providing examples of synthetic binary projection names and the corresponding human-readable label written by an expert, the LLM produces a label that better reflects the meaning of the relation, allowing easier interpretation by the user. This process extends the techniques developed in WP2 for the ontology documentation method, described in [41]. The resulting relations are used as input for the influence prediction method developed within WP2 and described in [12]. We manually analysed the resulting labels and found that many labels generated by the LLM differ syntactically but matches semantically. We exploited this result by clustering the labels together. To do so, we compute sentence embeddings of each label using Sentence BERT [42] and use them as input for KMeans. We constraint the method to find 20 clusters. For each cluster, we generate a sentence that unifies the ones in the cluster by prompting Mistral 7B. The generated sentences are then refined manually.

¹⁶The named entities are automatically linked to their entry in Wikipedia by BLINK [39], the entity linker used by SPRING, and aligned to Wikidata and DBpedia in the AMR2RDF step of our pipeline

¹⁷https://melody-data.github.io/stories/published_stories/story_1687714706.423208.html

Description	Average score
Relations between a person and another object	3.3
Relations between two persons	3
Relations between a person and a place	2.4
Only highly relevant relations	1.9

Table 9.2: Expert evaluation on the generated labels

To validate the refined labels, we conducted a user evaluation with an expert. Each relation is validated in a 1-5 likert scale where 1 is assigned to a relation that does not make sense from a logical perspective and 5 to a relation that the expert would not have known otherwise. Table 9.2 reports the results.

9.2. Evaluation method/rationale

The Pilot's validation actions are of three types and relate to three different aspects:

- *Input validation* - validation of input data quality by experts
- *Validation of tools/means/software* - validation of interface and performance validation
- *Output/goal validation* - validation from a sub-corpus/case study and transfer to a larger corpus

9.2.1. Input validation

This step refers to the design and collection of the corpus. Two criteria were followed:

- 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;
- balance of textual genres and variety of languages. Textual sources following the above criteria were indicated by two expert musicologists

Representativeness was achieved by aligning the corpus content with the Pilot's research requirements, while balance was maintained by including a diverse array of textual genres such as essays, historical texts, biographies, autobiographies, correspondences, media, and performance catalogs. Although language diversity was desirable, the nature of the corpus's focus meant that the majority of sources are in Italian, with limited representation from English, French, and Spanish. Copyright restrictions prevent the full disclosure of the texts, but comprehensive meta-data facilitating corpus reproduction is made available, including authorship, title, publication date, textual genre, and source links for each document.

9.2.2. Validation of interface and performance by experts

Validation was performed on the linguistic data query tool¹⁸ developed within WP4 to which MusicBo contributed by providing data and operating queries. Some roundtable discussions and brainstorming meetings were carried out with language experts to gather requirements. For the validation of the interface, a questionnaire was conducted on a sample of students in order to test usefulness and ease of use of the interface through the Likert scale. The sample was asked to query the tool and evaluate individual features by investigating a linguistic phenomenon through the three available features: keyword search, concept search, and entity search. Figure ?? reports the result of the evaluation.

¹⁸<https://polifonia.disi.unibo.it/corpus/>

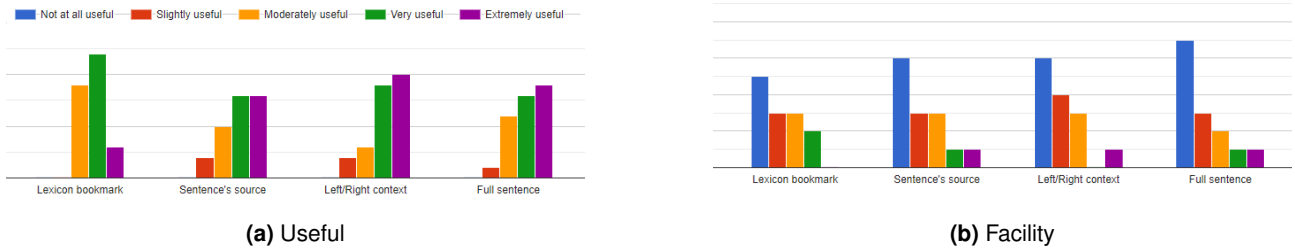


Figure 9.2: Linguistic interface evaluation

9.2.3. Output/Goal Validation

The tool allows for three types of interrogation which correspond to knowledge extraction based on different techniques and which have been the subject of individual evaluations.

1. Keyword/lemma 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;
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, as shown in Figure 9.5.

Through a questionnaire administered to a validation team composed of students and experts; these three research options were evaluated with respect to reusability on other applications (Figure 9.3); innovation compared to the state of the art (Figure 9.4); richness of information extracted (Figure 9.5).

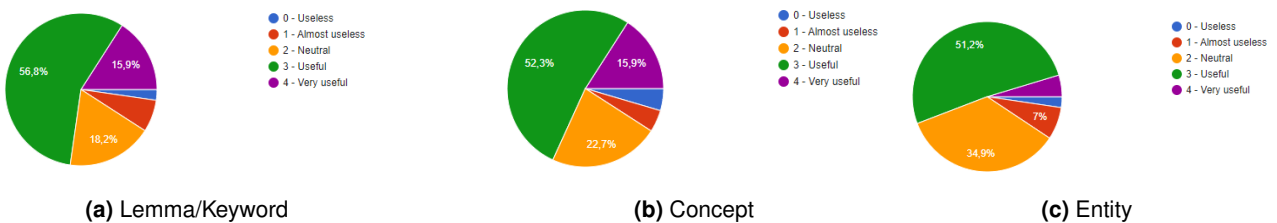


Figure 9.3: Usefulness for other applications

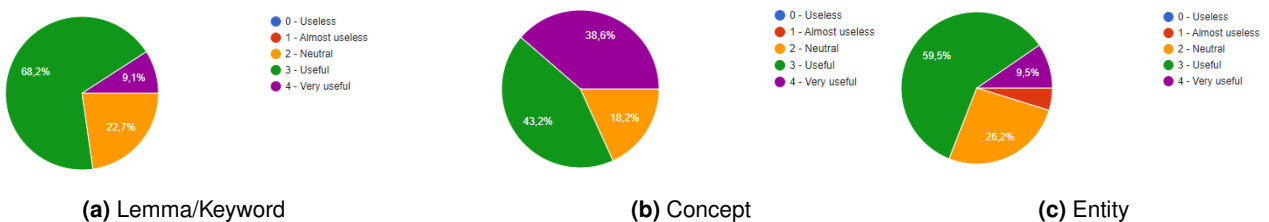


Figure 9.4: Innovation compared to the state of the art

Regarding the KG extraction from the documents, processing non-standard texts may lead to potential inaccuracies of text2AMR parsers, as such data is scarce in their training sets. Manual validation is time-consuming, and no standard benchmarks exist for semantic parsing of historic and OCR'd text. To address these challenges, we followed a back-translation [43] methodology. We converted the AMR graphs back to sentences using SPRING for

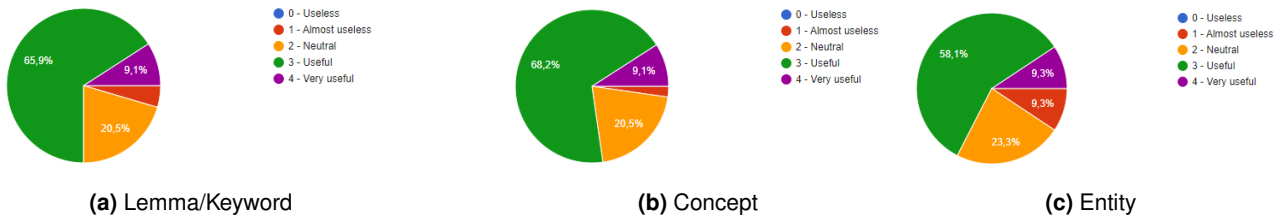


Figure 9.5: Richness of the results

English and m-AMR2Text for Italian¹⁹, followed by similarity score computations using BLEURT²⁰ for English and cosine similarity for Italian. We posit that high-quality graphs are associated with generated sentences exhibiting high BLEURT or cosine similarity scores. All AMR graphs paired with AMR2Text-generated sentences with a negative BLEURT score or a cosine similarity below 0.90 were discarded. Table 9.1 describes the statistics regarding the KG resulting from the application of Text2AMR2FRED to the MusicBo corpus, including insights regarding the automatic filtering. Raw data to recreate the KG are stored in a dedicated repository²¹.

9.3. Results and discussion

Regarding the linguistic query tool and considering the results of the evaluations, we find that the reusability rate is quite high, especially at the level of single word or lemma search. The tool is considered as a innovative when compared to the state of the art, and a strong interest in the function of terminological disambiguation by concept undoubtedly stands out from the data. Finally, users are highly satisfied by the richness of the results. From what it concerns the KG extraction pipeline, we find that the resulting KG contains a non-trivial amount of noise which might hinder its usability when queries are not equipped with sanity checks and filters. Our approach, involving finding binary relations, labeling through an LLM and enriching through graph-based techniques, increases the usability of the KG, particularly in the querying phase.

9.4. Summary of the outputs (FAIR section)

name	MusicBo Knowledge Graph
component-id	musicbo-knowledge-graph
type	Knowledge Graph
work-package	WP4
related-components	Polifonia-Corpus, Polifonia-Lexicon, machine-reading, amr2fred
licence	CC-BY_v4
contributors	Arianna Graciotti, Eleonora Marzi, Nicolas Lazzari, Rocco Tripodi
link	https://github.com/polifonia-project/musicbo-knowledge-graph
release-version	v0.1
doi	https://zenodo.org/doi/10.5281/zenodo.7633704

¹⁹<https://github.com/UKPLab/m-AMR2Text>

²⁰<https://github.com/google-research/bleurt>

²¹<https://github.com/polifonia-project/musicbo-knowledge-graph/tree/main>

name	MusicoBo Corpus Interrogation Tool
type	Corpus Interrogation Tool
work-package	WP1
related-components	Polifonia-Corpus, Polifonia-Lexicon, machine-reading, amr2fred
licence	CC-BY_v4
contributors	Arianna Graciotti, Eleonora Marzi, Nicolas Lazzari, Rocco Tripodi, Marco Grasso
link	https://polifonia.disi.unibo.it/corpus/

10. CHILD

10.1. Overview of the pilot

Focus of the pilot is on supporting music scholars, from the formulation of a hypothesis to the discovery and collection of resources relevant to the enquiry. The pilot develops tools for supporting the retrieval, collection, curation, and exploration of documentary evidence. Such a system is targeted to inform studies in historical musicology on the historical experience of music in childhood in Britain, using life writing – letters, diaries, memoirs, travel writing, currently undertaken by the Music Department of The Open University. Through this pilot, Polifonia supports research and education in music social history.

10.1.1. Background and connection with research WPs

In this section, we provide an overview of the research conducted in WP4 related to this Pilot, specifically, to complement what was already reported in deliverables D1.6 [23], D4.3 [44], and D4.4 [45].

Documentary evidence cataloguing from textual corpora is crucial to empirical humanities study. As the digital age unfurls, the wealth of documentary evidence available to humanity continues to proliferate exponentially, constituting a multifaceted tapestry of information encompassing everything from historical archives to contemporary multimedia repositories. Humanities research is increasingly relying on digital methods to aid the collection and organisation of material derived from textual digital and non-digital sources [46].

While a growing number of research efforts aiming to curate documentary evidence text such as Listening Experience Database (LED) [47] and UK Reading Experience Database [48], exist, the cataloguing of evidence is still a process that is performed manually, requiring a significant investment in terms of time and effort.

In the case of the Listening Experience Database (LED), once a piece of evidence is submitted through the online form, the submitted evidence goes through processes that include human annotation and validation. As the curation is mostly manual and done by human annotators, it is likely that background knowledge plays an important role, as the human annotator may know things that are not explicitly within the text, which comes from the metadata of the book. First, users need to read and interpret the content and decide whether it is useful to be included in the database. For example, that a given paragraph mentions a childhood experience of listening to music. Next, they need to identify in the text information that is useful to complete the required metadata. For example, that a certain event involved specific named individuals at a given time and location. Finally, they need to complement the information included in the text with background knowledge. This spans from the author of the source to the country of the town mentioned in the text, where the described event occurred. Therefore, at the forefront of documenting evidence, lies the serious challenge of suitable knowledge extraction techniques to support the curation activity [46].

LLMs like the GPT-4 [49] model being used in this paper have been the topic of much recent attention from the academic community in the last years. [50] provides a comprehensive survey of LLMs, including their evolution over time, the architectures and training methods used to build them, domains of application and challenges associated with the deploying of LLMs in real-world scenarios. [51] also provides an extensive survey on the evaluation methodologies and benchmarks on LLMs knowledge and capability and safety. [52] focuses instead on the challenge of evaluating these models and presents a comprehensive review of evaluation methods for LLMs focusing on three key dimensions: what to evaluate, where to evaluate, and how to evaluate. Parallel to the evaluation of these models has also been the question of their explainability. [53] introduces a taxonomy of explainability techniques and provides a structured overview of methods for explaining transformer-based language models like GPT-4. Debate has also taken place on the topic of how and how much these models can be said to understand language. [54] provides a summary of the pro and against arguments in the AI research community on whether large pre-trained language

models can be said to understand language and the physical and social situations that language encodes in any human-like sense. [55] surveys 194 relevant papers on arXiv to investigate the various applications of ChatGPT systems to research work, which is predominantly centred on direct natural language processing applications but also extends out to research in education, history, mathematics, medicine, and physics. However, LLMs limitation include the generation of contextually plausible yet potentially inaccurate responses [56].

In our work, we focus on the application of *in-context* learning prompt engineering techniques to leverage LLMs for automating two crucial tasks in knowledge curation of humanities databases: (a) the retrieval of relevant database records according to a research theme; and (b) the curation of database records of documentary evidence in textual form. Crucially, we explore the role of background knowledge in helping the classification of database records (a) and intervene to complement the information of textual sources so that an AI can recommend metadata annotations (b).

10.1.2. Progress from deliverables D1.6, D4.3, and D4.4

Recent years have witnessed a remarkable demonstration of Large Language Models (LLM), with important applications in natural language understanding [57], document generation [58], question answering [59], and text summarisation [60] in many different domains [61].

Yet, one intriguing and pressing question emerges: *to what extent can LLMs help to support the curation and classification of textual documentary evidence in the humanities?* The answer to this question holds profound implications for knowledge curation.

Our contributions are:

- an investigation into the integration of LLMs into the curation process, specifically
 - to classify textual evidence according to a theme of interest,
 - to extract valuable metadata from the text, and
 - to complement it with background knowledge.
- a benchmark of documentary evidence selected from the Listening Experience Database (LED) – previously introduced in D4.3, which has been extended and applied in the experiments reported here
- a suite of prompts for the LLM to automate the classification, extraction, and curation tasks, potentially reusable beyond our case study.
- experiments demonstrating value, opportunities, and limitations of this approach

10.2. Evaluation method/rationale

The primary objective of this research is to investigate the integration of LLMs into the curation process, with an analysis of their capacity to perform thematic classification of textual evidence (considering the case of music and childhood), and knowledge extraction, encompassing entity recognition, extraction of time, place, and mentions. This section presents the materials and methods employed in our research.

10.2.1. The Documentary Evidence Benchmark (DEB)

The Documentary Evidence Benchmark (DEB) was originally introduced in Deliverable D4.3 "*Software from Knowledge Extraction from Text*"¹. Motivation for this benchmark is outlined in this publications [63, 46].

The benchmark is based on the living case study of Listening Experience Database (LED). DEB includes a list of listening experiences curated by domain experts. Crucially, it provides a set of records that were classified as relevant to the research on music and childhood by a music historian (also coauthor of this article). All data was derived from the submitted documentary evidence in the Listening Experience Database.

¹The benchmark is published for reuse on Zenodo [62].

These are the data that can be used for benchmarking knowledge extraction processes:

- **child.csv** includes the list of listening experiences that were marked by domain experts to be relevant to childhood
- **experiences.csv** (columns: file, exp, excerpt, text, time, place, listening_to, environment, listener, listener_label, type, instrument, genre)
- **sources.csv** (columns: source, file, title, author, author_name, time)

In this work, we mainly focus on the child and experiences datasets and address the problem of supporting thematic classification and metadata curation using *in-context* learning LLM techniques.

10.2.2. Prompt Engineering for Thematic Classification (Task 1)

To approach the classification of music and childhood, we took a sample of around 1000 data items from the LED datasets that had already been reviewed manually by a domain expert; a music historian at The Open University. A number of these had been marked by the expert as describing the listening experiences of children, describing childhood or falling under the theme of childhood in some way. Our plan was to develop a suitable prompt to ask an LLM if it considered each of the listening experiences to fall under the thematic banner of childhood. To perform this experiment, we used OpenAI's API access to ChatGPT-3.5 model, making one call per entry from an iterative Python script².

In the case of Task 1, listening evidences were passed to the LLM, and the LLM was asked to not only provide a true/false answer if it considered the evidence to be related to music and childhood, but also to explain its reasoning. It was also asked to provide its answers and reasons in a JSON structure adhering to a pre-defined schema such that processing and rendering of the results could be automated. After an exploratory analysis on a selection of LED entries, the following LLM prompt was agreed upon: *"Does the following passage cover the theme of childhood/youth or describe or mention childhood/youth or children and young people in any way?"* An example response given by the LLM is given in the following listing:

```

1 {
2   "Childhood": "True",
3   "Reason": "The passage describes a little boy's excitement
4             and delight while watching a regiment band rehearsal.
5             The mention of the little boy, his interaction with
6             his father, and his enthusiasm for the music
7             indicates a focus on childhood/youth."
8 }
```

We study the effectiveness of this approach in Experiment 1, where we assess accuracy towards the annotated entries in DEB. In addition, we perform an error analysis with the aid of the domain expert.

10.2.3. Prompt Engineering for Metadata Curation (Task 2)

First, we performed an exploratory analysis to determine the Large Language Model (LLM) suitable for the cataloguing experiment. Specifically, we conducted a preliminary evaluation by designing a few-shot prompts and executed them on the web versions of some of the most widely used LLMs such as ChatGPT 3.5, ChatGPT 4.0 (GPT-4), and Google Bard to generate annotations. We then visually assessed the performance of these models and observed that the output of GPT-4 was most accurate, leading us to employ GPT-4 for this study. We then developed an annotation pipeline using OpenAI's API to GPT-4 model. We study the effectiveness of this approach in Experiment 2.

The experiment for Task 2 selects items from LED (included in DEB) with the related metadata and runs the designed operation with the LLM. The resulting recommended metadata are then stored alongside the human-curated ones

²We report here experiments that were performed before the introduction of GPT-4.

(see Figure 10.1). The documentary evidence text from the database record is passed to GPT-4 through the OpenAI API call for analysis and the generation of annotations in the same format as annotations performed by experts.

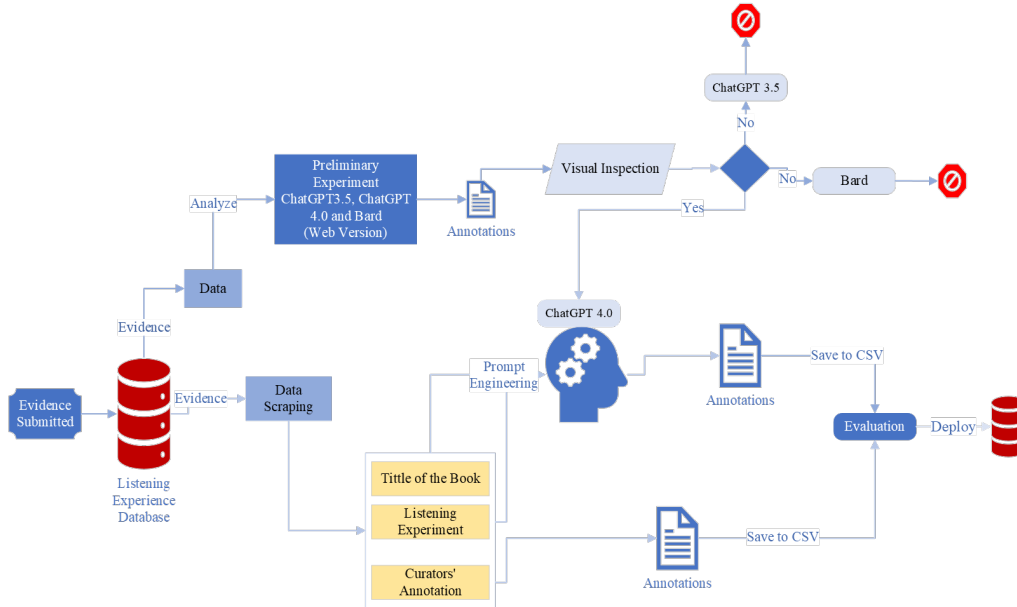


Figure 10.1: Overview of the knowledge extraction pipeline.

The process of transmitting documentary evidence to GPT-4 for analysis involves the utilisation of an API call as stated earlier, which after the analysis generates annotations. This process was achieved by engineering prompts within the Python script. In order to enhance GPT-4’s understanding of our context, we trained GPT-4 using a few-shot learning techniques, a paradigm that describes the process where a model is trained to perform a task with only a very small number of examples per class [64]. We provided the key entities or annotation keys that its values should be extracted as well as the description for the expected return, including the structure, thereby guiding the model’s output. An example of the training provided to GTP-4 is giving below. That is, at the prompt, we specify that we want the output to be in JSON structure with the seven annotation keys, namely: Listeners, Listening to, Performed by, Date/Time, Medium, Listening Environment, and Location. We also provided an explanation of each keys. The full code showing all the prompts can be accessed in the GitHub at ³. However, for the purpose of illustration, below is some example of the keys and explanations provided to train the model:

```

1 {
2     "Listeners": "Please enter the full name of the listener or listeners
3                 without any additional text.",
4     "Listening to": "Please specify only the title or name of what is
5                    being listening to, without any extra details.",
6     "Performed by": "Mention who performed or delivered the content
7                    being listened to."
8 }

```

Furthermore, the prompt underwent iterative process analysis, annotating the evidences as they are being scraped from the LED portal. An example of annotation from by GPT-4 is presented below:

```

1 {
2     "Listener": "John Ruskin",

```

³<https://github.com/polifonia-project/llm4led>

```
3 "Listening to": "German compositions, Italian modulations",
4 "Performed by": "Henry Watson, three sisters",
5 "Date/Time": "1820-00-00T00:00:00Z",
6 "Medium": "live",
7 "Listening Environment": "Indoors, In the company of others",
8 "Location": "London, United Kingdom"
9 }
```

Tuning parameters were largely left at default values and no specific configurations were used.

The dataset extraction process occurred in two stages: a) during preliminary experiments, where 50 pieces of evidence were randomly selected manually to identify the optimal model for the study, and b) during the actual experiment following the development of the annotation pipeline, 507 pieces of evidence were automatically extracted and processed using GPT-4 as the base engine for generating annotations. The extracted data from the LED include the URL of the evidence in the LED, the submitted evidence/text, and annotations curated by humans.

We conduct two different types of evaluation for Task 2. First, we apply string similarity measures to quantitatively assess the accuracy of the LLM in comparison to what provided by human annotators. Second, we perform a user survey, breaking down the analysis into different dimensions related to quality, completeness, and injection of beneficial or detrimental background knowledge.

We perform two experiments: i) the classification of evidence related to listening experiences in the context of music and childhood (Experiment 1) and ii) the cataloguing of documentary evidence text (Experiment 2).

10.2.4. Datasets for the evaluation

The evaluation for Task 1 was performed using 878 data items from the LED datasets that had already been reviewed manually by a domain expert and that are now part of the Documentary Evidence Benchmark (DEB).

For Task 2, we conducted an evaluation by randomly selecting 100 metadata annotations from a pool of 507 pieces of evidence from LED. These items were automatically extracted from the LED portal and annotated by GPT-4, and are now part of DEB. The selection process aimed to ensure a representative subset.

10.2.5. Experiment 1: classification of thematic evidence

For Experiment 1, a domain expert was asked to review and re-access the LLM classification for child related evidence against those classified by humans.

Experiment 1's initial results were somewhat unclear. Not only did the LLM miss a significant number of the entries flagged by the domain expert, it also flagged a large number of its own that were not originally marked by our expert. Some probing into what appeared to be false positive results revealed that many of these entries had in fact been correctly identified by the LLM, highlighting the difficulties faced by humans in exhaustively categorising large volumes of textual data. The LLM had also missed a number of entries too, so a slightly different approach to the LLM prompt was taken. As part of the wider CHILD project, a scenario had already been devised that describes a typical use case for the pilot.

The Scenario:

"Ortzen wants to characterize children's experience of music as witnessed in bibliographic and artistic sources. She is looking for primary sources (e.g. Personal journals, literary texts) wherein to find evidence of listening experiences. She needs to collect and analyze large corpora of texts and images recording or depicting children's experience with music. Documents include official sources (e.g. newspaper articles, reviews of concerts, paintings) and sources produced by "ordinary people". She prefers the latter as they provide more reliable feedback, and she looks at the context of production of such

sources (where, when, who created the source, the goal, which related events exist), contents (recurring motifs and themes), and elicited emotional responses. She collects sources belonging to different historical periods so as to characterize the development of identified phenomena.”

For our second attempt at thematic classification using an LLM, we instead presented the model with the above scenario and asked the LLM if each passage of text helps to address and answer the requirements of the pilot scenario. This experiment yielded more positive results than the previous prompt, with some overlap but also with many more passages now identified as falling under the banner of childhood. Since the original set of domain expert-identified entries showed to be incomplete, we employed the same music historian again to perform a qualitative evaluation the results of the LLM classification. They were presented with around 100 entries, specifically chosen where the LLM had flagged entries that were previously not in the original expert-selected subset. The purpose of this was to investigate the reasons for the discrepancies, rather than measure precise quantitative accuracy. Our expert was asked to review each entry and state whether they agreed or disagreed with the LLM’s inclusion of that entry in the CHILD subset. Where they disagreed, they were asked to explain their reasoning.

10.2.6. Experiment 2: curation of documentary evidence metadata

To evaluate the ability of an LLM to provide correct metadata annotations, we performed a quantitative analysis and a study involving human annotators. The first method employed a string similarity measure, specifically, the computation of the Levenshtein distance between the existing human annotation from LED and the GPT-4 generated annotation.

Furthermore, we employed a user evaluation process to assess the quality of annotations generated by GPT-4. Specifically, users were asked to compare the annotations produced by GPT-4 with the ones manually produced by domain experts in the LED database system.

We asked users to assess the response of the LLM according to the following criteria:

- LLM has new valid info (A)
- LLM has new invalid info (B)
- LLM is missing info from curator present in the text (C)
- LLM is missing external knowledge (D)
- LLM is introducing external knowledge (E)
- LLM is accurate (F)

The annotation metadata of each sample included: namely Listeners (1), Listening to (2), Performed by (3), Date/-Time (4), Medium (5), Listening Environment (6), and Location (7). In what follows, we refer to the dimensions of analysis as Ilmcode (alphabetical), while we refer to the metadata fields as infocode (numerical).

The process involved 100 randomly selected samples, clustered into five groups: Yellow, Blue, Green, Red, and Ash. Each group consisted of 20 items. For each group, three expert annotators were assigned to evaluate the annotations independently. A total of 15 expert annotators were recruited to participate in the evaluation process. To streamline the evaluation process, an automated Google form was used to generate the Evaluation Form. This form was designed to capture the submitted evidence, the human annotations, GPT-4 annotations and rating panel for each sample within the designated groups. The user study was conducted in presence, and a training was provided to ensure that the annotators possessed relevant expertise and were well-versed in assessing the quality of annotations. Thereafter, each annotator independently reviewed and assessed the annotations for each item within their assigned group.

To illustrate the process, an evaluator takes a field value in GPT-4 Annotations. For example, “Listener” in GPT-4 Annotation, then compare the field value against the extracted human Annotation. If they are exactly the same, at the rating panel, the evaluator ticks LLM is accurate as indicated below. However, if it is different, the evaluator checks the provided submitted evidence and then decide based on the options at the rating panel.

Example 1: Listener

GPT-4 Annotations

Listener: John Ruskin

Curators' Annotations

Listener: John Ruskin

In this case,

- Column 1: LLM has new valid information should be no, on the rating panel, untick = "No".
- Column 2: LLM has new invalid information should be no, on the rating panel, untick = "No".
- Column 3: LLM is missing information from text, should be no, on the rating panel, untick = "No".
- Column 4: LLM is missing external knowledge, should be no, on the rating panel, untick = "No".
- Column 5: LLM is accurate, on the rating panel, tick = Yes.

The expert evaluation results were processed using a scoring system where 'yes' was assigned a score of 1, and 'no' was assigned a score of 0.

The different expert groups, identified as Ash, Blue, Green, Red, and Yellow, which participated in the evaluation, were encoded as groupcode. For each llmcode, we matched it against all the corresponding responses for infocode according to the groupcode. For example, LLM has new valid info is encoded as 'A', and the 'Listeners' is encoded as 1. This gives rise to Cluster A1, which means all responses on how well the model performed in annotating the infocode, 'Listeners'. That is, all responses of LLM has new valid info on Listeners. A total of 42 Clusters were develop, and this form the basis of this analysis. We then performed a responses analysis on each annotation field with descriptive statistics such as percentages to have a holistic view, as well as understand the field in which GPT-4 has a great performance. To facilitate statistical analysis, we conducted inter-rater reliability testing for a comprehensive examination of the collected data.

10.3. Results and discussion

10.3.1. Classification of Childhood Evidence (Experiment 1)

Table 10.1 presents the classification of childhood music experiences derived from an analysis of a dataset comprising 878 instances. The cumulative results of the initial prompt and the second scenario-based prompt for Experiment 2 yielded a total of 85 entries classified as childhood-related by the LLM. In the table, the results from the initial prompt are indicated as (1), while the second classification with the updated scenario-based prompt is denoted as (2).

Table 10.1: Classification of childhood music experience.

Total listening experiences analysed	878
Flagged by domain expert	26
Flagged by ChatGPT-3.5 (1)	45
Flagged by ChatGPT-3.5 (2)	71
Flagged by ChatGPT-3.5 (total)	85

A total of 103 results were evaluated, with 78 instances demonstrating agreement between the LLM and the domain expert, as illustrated in Table 10.2. This yielded an agreement rate of 75.7%, indicating the percentage of cases where both the LLM and domain expert shared the same assessment. The observed moderate-high level of agreement suggests a promising alignment between the automated assessment provided by the LLM and the expertise of a human evaluator within the domain. These findings underscore the potential utility of LLMs in supporting domain experts in thematic classification tasks.

The initial results showed the challenges of LLMs in thematic text classification, with notable discrepancies between the LLM and domain expert as presented in Table 10.1. The LLM's propensity to flag both overlooked and spurious entries highlights the difficulties inherent in exhaustive human categorisation and the machine's pattern-recognition

Table 10.2: Expert evaluation

Total results evaluated	103
Instances where LLM and domain expert agree	78
Agreement rate	75.7%

capabilities. Subsequent adjustments to the LLM’s prompt, aligned with the CHILD project use case, enhanced classification accuracy, indicating the effectiveness of scenario-based prompting in improving LLM output. The more focused second attempt revealed a pattern in some of the LLM’s misclassifications, a tendency to overgeneralise childhood indicators. This appears to be a consistent overreach, interpreting any youthful reference or familial mention as a childhood experience. The human expert’s review illuminated this over-optimistic bias, offering clear examples where the LLM conflated youthfulness with childhood. Some comments from the evaluation:

“This doesn’t sound like a childhood experience to me. The protagonists are merely described as young, and the word ‘childish’ is used (suggesting they’re actually young adults).”

“Very borderline - the term ‘girls’ doesn’t really indicate children here I think.”

“No - this is simply a singing voice with a youthful quality.”

“Could be a childhood experience but I don’t think there’s much in the text to support this.”

The comments above, and more from the full evaluation, circle mainly around cases where the LLM has been over-optimistic in its analysis, often involving phrases referring to when the subject was young or when they were with their parents. The LLM has interpreted these as instances of childhood but there is often a lack of evidence to say this is definitely the case.

10.3.2. Cataloguing of documentary evidence text (Experiment 2)

The boxplot in Figure 10.2 presents the results of the Levenshtein distances for GPT-4 annotations across all the annotation fields. This provide insights into the levels of similarity and dissimilarity across different annotation fields. In the "Listener" field, the mean distance of 4.62 suggests moderate differences, while the median of 0.0 indicates that half of the annotations are identical, with a maximum distance of 83 indicating some significant disparities. For "Listening to," the higher mean of 17.98 and median of 15.0 point to more substantial differences, with a maximum distance of 125 indicating notable dissimilarities. Similarly, "Performed by" exhibits variations with a mean of 14.66 and median of 13.0. In contrast, "Medium" shows minimal differences with a mean of 0.87, and the majority (75%) having a distance of 0, indicating high similarity. "Listening Environment" presents consistent differences with a mean of 24.51 and median of 26.0, indicating a high level of dissimilarity. Finally, "Location" displays relatively low distances with a mean of 10.47 and median of 9.5, suggesting moderate differences between annotations. The dissimilarities observed in 'Location' resulted from the inclusion of 'Country' by GPT-4 in its annotations which is not present in the human annotations.

10.3.2.1. User study

The results in Table 10.3 presents the responses analysis in clusters of each llmcode against each infocde. The table provides the counts and percentages of "Yes" and "No" responses for each item within these groups. The results of the analysis indicate that all raters (100%) provided a "No" rating for the item concerning the LLM ability to generate new valid information on Date/time (A4). This suggests that GPT-4 model, captures the date/time information present in the submitted evidence, without generating additional valid data. The analysis also revealed that the LLM received a 100% "No" rating for missing external knowledge on Location (D7). This result implies that the LLM accurately captures the location information and it not missing external knowledge. Furthermore, the results showed that the LLM did not introduce external knowledge on Listener (E1), as indicated by the 100% "No". This finding suggests that the LLM incorporates only the names of listeners mentioned in the input text, without resorting to external knowledge

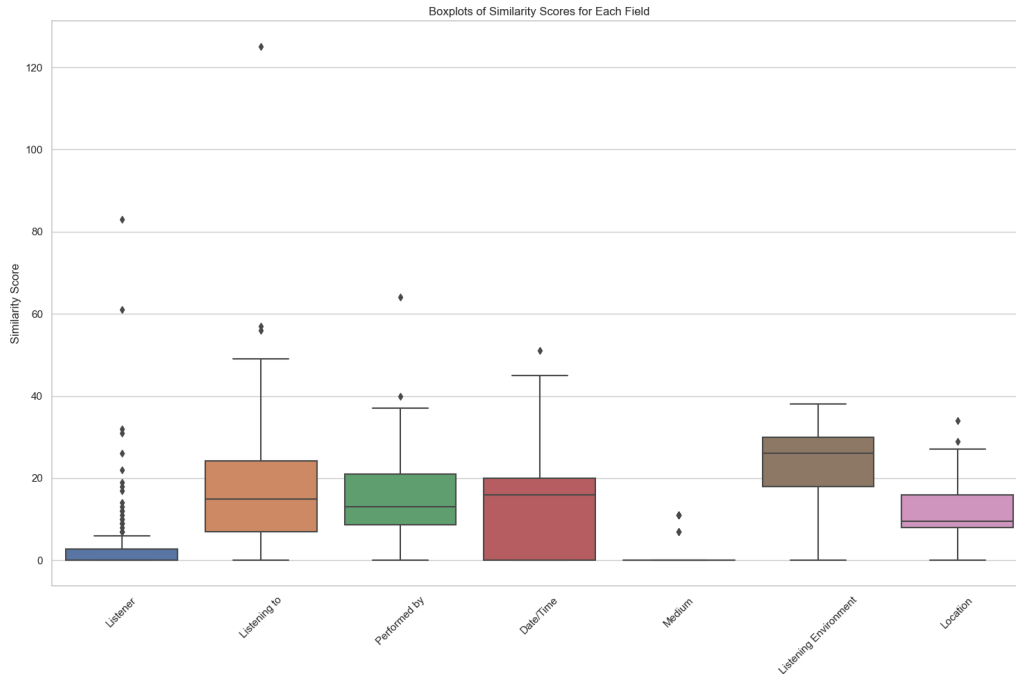


Figure 10.2: A Boxplot showing Levenshtein distances of the annotations

to supplement this information. The LLM also did not introduce external knowledge on Date/time (E4), with a 100% "No" rating, implies that the model does not require additional external knowledge to enrich its understanding of date and time-related details, relying solely on the information provided in the input text. Comparing across Ilmcodes, Clusters F1, F2, F3, F4, F5, and F6 which indicate the interaction of Ilmcode F (LLM is accurate) and all the infocode generally have a higher proportion of 'Yes' responses compared to 'No' responses, indicating more agreement on the performance of GPT-4 in annotating these annotation fields. Across all clusters, there is variation in the percentage of 'Yes' and 'No' responses, indicating differences in agreement levels across different items. For the Ilmcode B, which assesses "LLM has new invalid info," the majority of responses, ranging from 90.33% to 99.33%, indicate "No." This suggests that raters predominantly observed the absence of new invalid information in the LLM's output. Regarding Ilmcode D, which evaluates "LLM is missing external knowledge", in order to providing accurate annotations of infocode, the majority of responses, ranging from 90.67% to 100%, indicate "No." This indicates a consensus among raters that the LLM generally incorporates external knowledge sufficiently to annotate infocode accurately. The same trajectory was observed in "LLM is introducing external knowledge", except for the infocode 7 (Location), indicating "No" in most cases. However, for infocode 7 (Location), there was a marginal "Yes" response. This divergence is likely attributed to the inclusion of a country of in the LLM's annotations, representing a form of external knowledge introduction.

The shading of the rows in Table 10.3 gives an overview of the level of human rater agreement with the LLM annotations. Light shading indicates an agreement level above 80%. Dark shading indicates an agreement level below 80%. High levels of agreement are found for all LLM codes except F. This can be expected, as F would essentially accumulate any differences in rater response across the previous codes. This demonstrates the advantage of breaking down the human evaluator responses across LLM codes and information codes rather than a single overall judgement of the GPT-4 annotation.

Table 10.3: Summary of Yes/No Responses and Percentages for Each Cluster

Field	Items	Yes	No	Yes%	No%	Items	Yes	No	Yes%	No%
	A: LLM has new valid information					D: LLM is missing external knowledge				
1 = Listeners	A1	51	249	17.00	83.00	D1	1	299	0.33	99.67
2 = Listening to	A2	141	159	47.00	53.00	D2	10	290	3.33	96.67
3 = Performed by	A3	99	201	33.00	67.00	D3	24	276	8.00	92.00
4 = Date/Time	A4	0	300	0.00	100.00	D4	4	296	1.33	98.67
5 = Medium	A5	14	286	4.67	95.33	D5	3	297	1.00	99.00
6 = Listening Environment	A6	39	261	13.00	87.00	D6	28	272	9.33	90.67
7 = Location	A7	43	257	14.33	85.67	D7	0	300	0.00	100.00
	B: LLM has new invalid info					E: LLM is introducing external knowledge				
1 = Listeners	B1	15	285	5.00	95.00	E1	0	300	0.00	100.00
2 = Listening to	B2	29	271	9.67	90.33	E2	6	294	2.00	98.00
3 = Performed by	B3	29	271	9.67	90.33	E3	4	296	1.33	98.67
4 = Date/Time	B4	18	282	6.00	94.00	E4	0	300	0.00	100.00
5 = Medium	B5	7	293	2.33	97.67	E5	1	299	0.33	99.67
6 = Listening Environment	B6	12	288	4.00	96.00	E6	7	293	2.33	97.67
7 = Location	B7	2	298	0.67	99.33	E7	162	138	54.00	46.00
	C: LLM is missing info present in the text					F: LLM is accurate				
1 = Listeners	C1	5	295	1.67	98.33	F1	229	71	76.33	23.67
2 = Listening to	C2	18	282	6.00	94.00	F2	105	195	35.00	65.00
3 = Performed by	C3	32	268	10.67	89.33	F3	120	180	40.00	60.00
4 = Date/Time	C4	97	203	32.33	67.67	F4	188	112	62.67	37.33
5 = Medium	C5	7	293	2.33	97.67	F5	268	32	89.33	10.67
6 = Listening Environment	C6	97	203	32.33	67.67	F6	138	162	46.00	54.00
7 = Location	C7	17	283	5.67	94.33	F7	93	206	31.00	68.67

10.3.3. Discussion

The aim of this study was investigate into the integration of LLMs into the curation process, with an analysis of their capacity to perform entity recognition, extraction of different types of entity for LED annotation and thematic classification of evidence related to music at childhood (CHILD). The study has revealed that LLMs can provide support in the data curation process for documentary text and the classification of evidence in LED, while highlighting its limitations and discerning the potential impact of training the model to guide its output.

These findings not only reveal the promising capability of LLM in future applications but also emphasize the practical value of prompt engineering techniques for supporting humanities researcher in managing documentary evidence in digital archives. Specific instances of extracting text from submitted evidence and aligning them to the correct annotation fields and thematic classification of text based relevance in relation to music at childhood, highlight a commendable level of efficacy of LLMs in enhancing data curation within the humanities, particularly focusing on the cataloguing of documentary evidence text.

While this research highlights the successes achieved, it also identifies areas within the utilization of LLMs for supporting curation and thematic classification that require improvement. Some observed Levenshtein distances remain notably wide, particularly in fields related to listener and listening to, reaching up to 85 and 125, respectively. These significant dissimilarities indicate potential areas for enhancement. Scholars and researchers should carefully consider these insights to address and realign elements that currently fall short of achieving perfect annotation and classification of documentary evidence.

The central question of the extent to which LLMs can aid in the curation and classification of documentary evidence databases in the humanities reveals a strategic pathway for the comprehensive use of LLMs in performing annota-

tions and classifications. This is evidenced by the varying accuracy results of LLMs, which significantly differ across all other llmcodes. Additionally, the LLM's ability to introduce external knowledge to enhance the Location field is notable. This shift signifies adaptability and responsiveness to the evolving landscape of LLM use in applications, emphasizing the need for utilizing LLMs to support the curation of documentary evidence text in LED and thematic text classification such as music at childhood.

10.4. Summary of the outputs (FAIR section)

List all components related to this Pilot

name	Documentary evidence benchmark
component-id	documentary-evidence-benchmark
type	Dataset
work-package	WP4
licence	Apache-2.0
contributors	Enrico Daga, Alba Morales, Jason Carvalho, Chukwudi Uwasomba
link	https://polifonia-project.github.io/ecosystem/polifonia-project/documentary-evidence-benchmark/README.html
release-version	v1.0
doi	https://doi.org/10.5281/zenodo.6866982

name	LLM4LED
component-id	llm4led
type	Software
work-package	WP4
licence	Apache-2.0
contributors	Chukwudi Uwasomba, Enrico Daga
link	https://github.com/polifonia-project/llm4led
release-version	v1.0
doi	https://doi.org/10.5281/zenodo.6866982

name	Classification and curation of Listening Experiences with LLMs (Demo)
component-id	child-search-expansions
type	WebApplication
work-package	WP4
licence	Apache-2.0
contributors	Jason Carvalho, Alba Morales, Chukwudi Uwasomba, Enrico Daga
link	https://polifonia-project.github.io/ecosystem/polifonia-project/llm4led/README.html
release-version	v1.0
doi	https://doi.org/10.5281/zenodo.6866982

11. MEETUPS

11.1. Overview of the pilot

The MEETUPS pilot [65] focuses on supporting music historians and teachers by providing a web tool that enables the exploration and visualisation of encounters between people in the musical world in Europe from c.1800 to c.1945. It relies on information extracted from public domain books such as biographies, memoirs and travel writing, and open-access databases.

Such encounters were particularly significant before the period when musical ideas and influences were widely disseminated through broadcasting and recording technologies, and the collection of data about them will make it possible to trace key points of cultural and musical exchange and dissemination and meetings that were catalysts for musical change. These encounters will be explored in a timeline and map interface and may reveal unexpected connections and relationships that cast new light on aspects of European music history. The tool will provide persistent, citable identifiers in order to support referencing in scholarship outputs.

As an overview, the MEETUPS Pilot is focused on:

- collecting the knowledge requirements in terms of data expected and interface to be displayed for exploration in the Web tool,
- selecting and collecting data sources according to availability and objective,
- designing the knowledge extraction pipeline in charge of gathering all the data and,
- finally building the knowledge components that are the base of the Web tool, specifically the MEETUPS Knowledge Graph [66] and MEETUPS Ontology [67].

11.1.1. Connection with research WPs

The pilot contributed to the activities of the following WPs:

- WP4: Meetups Pilot¹ is strongly connected to WP4 and focused on developing methods for entity recognition and linking into a Knowledge Graph construction pipeline. The focus was on developing tools for the automatic extraction of people, places, and time expression entities, as well as the automatic classification of text according to events of interest for network analysis in the music domain: Developed tools for identifying named entities and link them automatically to DBpedia; Developed tools for identifying time expressions; Developed the workflow to identify and classify the text according to the list of meetup types (in collaboration with domain experts, these classes are the following: BusinessCareer, PersonalLife, Coincidence, Education, PublicCelebration and MusicMaking); and developed and algorithm to identify the historical meetups.
- WP1 and WP5: Meetups Pilot contributed to WP1 - task 1.7 by collecting data about personalities in the European musical world and extracting information about encounters between them from publicly available web sources. Furthermore, the pilot has developed and made available the Meetups Web Interface² leveraging the use of musical heritage knowledge graphs and providing accessible and improved visualisation of meetups through a timeline and map interface. This web platform showcases the historical meetups extracted from biographies and structured in the Musical Meetups Knowledge Graph [68].
- WP2: Regarding WP2 and Task 2.2, the MEETUPS Ontology [69] contributes to designing and implementing the Meetups Ontology that models the constituents of a historical meetup, including entities such as people, places, temporal relations and types of events. The ontology aims to support researchers in performing a thorough analysis of social networks and discovering potential new links, therefore, addressing challenges in

¹Repository https://github.com/polifonia-project/meetups_pilot

²Meetups web interface <https://polifonia.kmi.open.ac.uk/meetups/>

terms of interpretation and analysis complexity. The Meetups Ontology re-uses attributes of the Polifonia core Ontology and is integrated into the Polifonia Ontology Network.

11.1.2. Progress from last deliverable D4.4

In what follows, we list the latest updates on Meetups Pilot since the last deliverable:

- Minor updates to the Web tool, which enables the exploration and visualisation of encounters between people in the musical world via a web interface. Updates include bug fixes and query development to accommodate the latest versions of the Meetups knowledge graph.
- Knowledge extraction pipeline completed, feeding the Musical Meetups Knowledge Graph with data on encounters between personalities in European music history, a total of 33K biographies were processed.
- Evaluation of the data quality: building a gold standard to evaluate the accuracy and quality of the data produced by the Knowledge Extraction pipeline.

11.2. Evaluation method/rationale

In this section, we describe the evaluation performed on the data extracted from textual sources and present the survey results for assessing the importance and usability of the Musical Meetups Knowledge Graph (MMKG) to domain experts.

11.2.1. Knowledge Extraction pipeline - data evaluation.

As described previously, the Meetups Pilot work focused on collecting and extracting data to support the exploration of encounters of European musicians throughout history.

The knowledge requirements of scholars looking to uncover such encounters are identified and stated in the form of Persona stories and formalised as Competency Questions (CQs)³. These requirements were compiled in the Polifonia stories named Ortez, David, and Sophia.

From this analysis, we identified the elements of a *historical meetup*. Four components can be listed:

- Participant(s): people involved (who?)
- Location: the place where it took place (where?)
- Type of event: the reason for the event (what?)
- Time expression: the date or moment the meetup happened (when?)

These elements were identified and extracted from public domain sources, specifically Wikipedia. We approach this problem from the perspective of knowledge extraction tasks such as Named Entity Recognition and Linking (NER) in order to obtain entities such as people, places and temporal expressions. We include in the pipeline a Machine Learning approach to automatically classify the text according to the main theme, or purpose of interest for scholars. Finally, we process text using LLM tools to improve the classification of themes and the normalisation of temporal expressions. Finally, the pipeline includes the development of an algorithm to identify the historical meetups.

To evaluate the quality of the data extracted, we propose to build a gold standard for each element identified as an important part of historical meetups. In order to build the gold standard, we select randomly the sentences that will be annotated. We selected 300 sentences for each element: people, place, temporal expression, theme and historical element. We asked four people, to annotate the sentences with the corresponding elements. For instance, two annotators will read and annotate the entity and its resource (DBpedia or Wikidata). They will do a first round where they perform the task alone; in a second round, they will go through all the annotations together and solve any discrepancies or doubts. Table 11.1 summarises the gold standard statistics.

³Persona stories and CQs: <https://polifonia-project.github.io/ecosystem/pages/pilots/meetups.html>

	Person	Place	Temp. expression	Theme	Historical meetups
Entity identified	278	184	169 (normalised dates)	300	88
Entity linked	257	184	-	-	-

Table 11.1: Results Gold Standard

We performed an evaluation comparing the output of different experiments performed on each element identified. In this deliverable, we will concentrate on the final results, while we prepare a more thorough analysis to be published as a paper. Table 11.2 summarises the results found.

	People		Place		Temporal expressions		Theme		Historical meetups				
	Total	Accuracy	Total	Accuracy	Total found	Accuracy	Total	Precision	Correctly identified	Accuracy	F1		
Entity mention	214	0.77	162	0.88	206	0.986	@1	164	0.55	HM	51	0.58	0.52
Entity linking	173	0.67	161	0.88	-	-	@ 2	223	0.74	-	-	-	-

Table 11.2: Evaluation results on Meetups entities extraction process

11.2.2. Feedback questionnaire from domain experts.

In a survey with 12 domain experts from the Music Department of our University, we evaluated the value to users of MMKG⁴. Participants were either/or researchers (75%), musicians (33%), educators (33%), and historians (25%), with a significant 91.7% reporting daily engagement with music-related content. All the respondents agreed on the value of documenting musical history encounters, considering them either important (50%) or very important (50%). All respondents reported not being aware of any tool/database to store and organise historical music encounters. We asked about the value of specific elements of the ontology. All respondents rated the importance of documenting the people involved in musical encounters as "very important"(Fig 11.1). The place and time of the encounter were rated "very important" and "important" by 83.33% and 16.67%, respectively. Responses were varied for the purpose of encounters: 58.34% rated it as "Very Important," 33.33% as "Important," and 8.33% as "Moderately important."

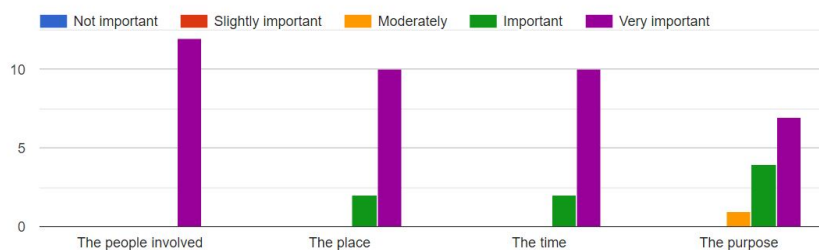


Figure 11.1: Dimensions in musical encounters.

Business meetings and music-making have been highly valued throughout music history, which received over 83% of the rating as "very important" and "important, respectively". The survey also revealed that 75% of the respondents respectively rated personal life and public celebration as "very important" and "important" purposes for musical encounters. While education and coincidence were rated the same at 66.67%.

⁴Complete results can be found at this address: https://docs.google.com/document/d/1PGknESmJm4f_-QwSdebTuiHIZ3bMaBrctcs4-9AKLPcc/edit?usp=sharing

In addition, all participants agree on the utility of exploring subjects' geographical, temporal, and thematic proximity, even outside direct evidence of documented encounters. Finally, the usefulness of knowledge graphs in teaching music history is largely (75%) acknowledged by experts.

11.3. Results and discussion

Regarding the software components, part of the knowledge extraction pipeline, we find that following a hybrid approach helped to include different technologies that improved the final quality of the data extracted. The feedback collected from domain experts indicates researchers strongly recognise the significance of resources such as a knowledge graph for the preservation of music's cultural heritage in a structured way.

11.4. Summary of the outputs (FAIR section)

name	Musical MEETUPS Knowledge Graph
component-id	meetups-knowledge-graph
type	KnowledgeGraph
work-package	WP4
related-components	meetups-ontology, meetups-corpus-collection
licence	CC-BY_v4
contributors	Alba Morales Tirado, Enrico Daga, Jason Carvalho
link	https://polifonia-project.github.io/ecosystem/polifonia-project/meetups-knowledge-graph/README.html
release-version	V1.0
doi	https://doi.org/10.5281/zenodo.7924618

name	MEETUPS Ontology
component-id	meetups-ontology
type	Ontology
work-package	WP4
related-components	meetups-knowledge-graph, meetups-corpus-collection
licence	Apache-2.0
contributors	Alba Morales Tirado, Enrico Daga, Jason Carvalho
link	https://polifonia-project.github.io/ecosystem/polifonia-project/meetups-ontology/README.html
release-version	V1.0
doi	https://doi.org/10.5281/zenodo.7928155

name	MEETUPS Pipeline
component-id	meetups-pilot
type	KnowledgeGraph
work-package	WP4
related-components	meetups-ontology, meetups-corpus-collection
licence	CC-BY_v4
contributors	Alba Morales Tirado, Enrico Daga, Jason Carvalho
link	https://polifonia-project.github.io/ecosystem/pages/pilots/meetups.html
release-version	V1.0
doi	https://doi.org/10.5281/zenodo.7928155

name	MEETUPS Web Application
component-id	meetups-application
type	WebApplication
work-package	WP5
related-components	meetups-knowledge-graph, meetups-ui-design
licence	Apache-2.0
contributors	Alba Morales Tirado, Enrico Daga, Jason Carvalho
link	https://polifonia-project.github.io/ecosystem/pages/pilots/meetups.html
release-version	V1.0.0
doi	https://doi.org/10.5281/zenodo.10017423

12. Reusability survey

Impact and research reusability are central to the Polifonia way of doing research. In the Polifonia Grant Agreement, Table 1.3.2 showed expected Technology Readiness Levels (TRLs) for various outcomes of the project. Reusability was emphasised by Reviewers and Stakeholders at several meetings, e.g. Bologna, October 2023. Following this, a reusability survey was designed and circulated among project members. The goal was to collect self-assessed information on various aspects of reusability of research outputs.

In this Chapter, we briefly present the rationale for the survey and its design, and then describe the results. We then include a mapping of TRLs of our outputs against the values expected in the original Grant Agreement Table 1.3.2.

12.1. Design and rationale

During the Bologna meeting (October 2023), a working group brainstormed on how to report on reusability in a rigorous way. A self-assessment online survey, organised by Polifonia Ecosystem Component, was chosen as the appropriate design. Organising by Component is appropriate because Polifonia's approach to open science and data management mean that all of the project's relevant outputs are collected as Components in the Ecosystem. A self-assessment survey is appropriate because individual Components have owners who can take responsibility for reporting on their components, and they have specific information on reuse not easily available to outside assessors. The survey content was drafted in Microsoft Forms and re-drafted at meetings of the Polifonia Technical Board. It was circulated to project members January - March 2024. 24 responses were received and were analysed by the Technical Board in March 2024.

12.2. Reusability Survey

The following [link](#) can be used to access a summary of the results. (In a few cases, below, the survey software has cut off a small amount of text in its visualisations. We have generally added extra text to clarify. In any cases of ambiguity, the link can be used.)

In the following sub-sections we briefly discuss the main results. Complete results are included as Appendix A.

12.2.1. Questions 1-3: Names, IDs, links

These initial questions requested Component names, Ecosystem Component IDs, and GitHub or other links. For all 24 responses, a GitHub or w3id link was provided. In these questions and all that follow, we used the term "Project" to refer to the respondent's sub-project within the Polifonia Ecosystem. However these correspond directly to Components, so we will use these terms interchangeably.

12.2.2. Question 4: Project types

For this question about project types, the results in Fig. 12.1 show a majority of Components are Software or Knowledge Graph or Similar.

4. Which of these best describes your project type?

[More Details](#)

● Software	10
● Data	1
● Knowledge graph, ontology, or ...	9
● User interface	3
● Other	1

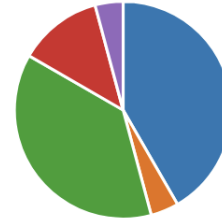


Figure 12.1: Project types. Most are of type Software, or type Knowledge Graph, Ontology, or Similar.

12.2.3. Question 5: Target audience

The target audience for a project speaks to its reusability. We found a wide range of target audiences, as shown in Fig. 12.2.

5. Who is the target audience for your project? Eg developers, or only Polifonia developers; musicologists; historians; etc.

[More Details](#)

● Developers (in the project team)	13
● Developers (any)	12
● Domain Experts	9
● Computational Musicologists	10
● Musicologists	6
● Historians	4
● Researchers	13
● Other	5

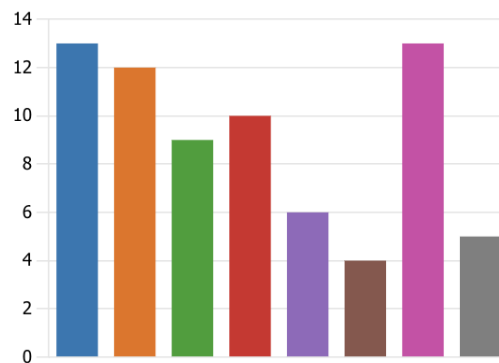


Figure 12.2: Target audience.

12.2.4. Question 6: Aspects of reusability

We asked: How reusable is your project in the following five aspects (where higher is better, i.e. more reusable):

1. When considering your breadth of project goals: if the project goals are highly specific to you (e.g. software not reusable because it is one-off curatorial work), and would not be of interest to others, this makes it less reusable. If the project goals address the goals of a broad class of users/tasks, this makes it more reusable.
2. When considering its reusability with respect to technical knowledge: if the user would need substantial, specific technical knowledge to reuse it, this makes it less reusable. Example: expertise in coding in C. If they would need to visit your lab or call you to get started, this makes it less reusable. If the project provides good "getting started" documentation, this makes it more reusable.

3. When considering its reusability with respect to domain knowledge: if the user would need substantial, specific domain knowledge to reuse it, this makes it less reusable. Example: expertise in musicology of 16th century Italian folk music.
4. When considering its reusability with respect to computational resources: if the user would need substantial, specific computational resources to reuse it, this makes it less reusable. Examples: requirement for a server with large RAM, a GPU, or web hosting.
5. When considering its customisability: if the project can be easily extended, customised, or re-branded, this makes it more reusable.

The results were overall positive for reusability in all aspects. However, several projects were not reusable or only slightly reusable.

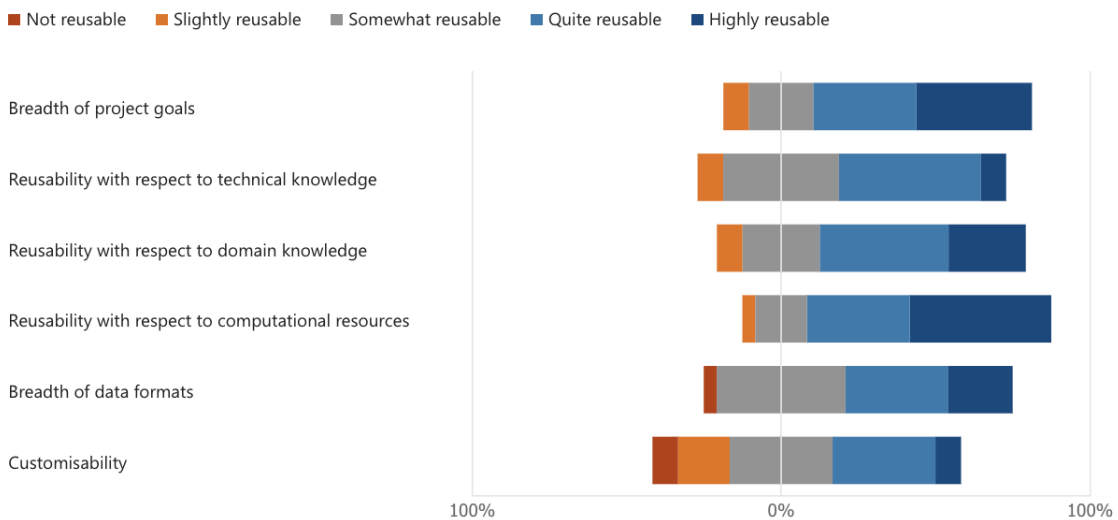


Figure 12.3: Aspects of reusability.

12.2.5. Questions 7-8: Outside reuse

7. Has your project already been reused by someone other than its core developers (ie, you or your group)? Even if that reuse is not "complete", please include it.

[More Details](#)

[Insights](#)

- No 12
- Yes, reused within Polifonia 6
- Yes, reused outside Polifonia 6



Figure 12.4: Outside reuse.

We asked respondents about reuse of their components by those outside the component or outside the Polifonia project. Results (Fig. 12.4) showed there was substantial outside reuse. We also asked for optional evidence, and for this we received substantial evidence:

- <https://www.eecs.qmul.ac.uk/~simond/pub/2023/ShaninDixon-WASPAA2023.pdf>
- <https://www.culturabologna.it/events/data-wanderings>
- <https://data.beeldengeluid.nl/showcases/moz-dataset-blog>
- GEL, a UNESCO funded project: <https://projects.dharc.unibo.it/digestgel/>; KNOT, a Ministry of Italian Culture funded project (in progress).
- Methods developed have already arisen interest in external projects, like PHAROS, which has recently kindly funded the research group of the University of Bologna to continue developing methods for data reconciliation, hence ensuring continuity in maintenance of the code base of the Polifonia web portal.
- MELODY can be used by any user, regardless of them being Polifonia collaborators, and it offers a separate venue to publish data stories created by external users (<https://melody-data.github.io/stories/>). To date the external catalogue counts more than 100 data stories, mostly developed by students at the University of Bologna, who are currently using the platform to learn using Semantic Web technologies and music data.
- 6 adopters for 5 different research projects: 1/ Mangani [Università degli Studi di Firenze]: The Interpretation of Textual and Music Meanings in the Late 16th-Century Madrigals and Motets 2/ Ceulemans [Université catholique de Louvain]: Dufay’s Musical Output 3/ Porot [Université de Reims] and Laurenti [Sorbonne université]: The Rethorical Interpretation of Musical Cadences in Lully’s Operatic Works 4/ Lalitte [IReMus]: Stravinsky’s Symphonie de Psaumes 5/ Freedman [Haverford College, USA]: Pedagogical Project on Renaissance Music Annotation
- The library is published open-source, we have evidence of adoption by several researchers and industry practitioners in the area of Knowledge Graph technologies.

12.2.6. Questions 9-12: Data formats and technologies

These questions asked respondents about the specific data formats and technologies used by the components, and tolerance for missing or imperfect data. As shown in Fig. 12.5, common technologies include SPARQL, JSON, RDF, and Python.



Figure 12.5: Data formats.

The large majority of components use standard data formats (one or multiple), with only a few using custom formats. Again, for the large majority of components, missing or imperfect data can be handled without a major problem. Often, if data is missing, it is simply missing, and the component operates without it.

12.2.7. Question 13: TRLs

We asked: “What is the approximate Technology Readiness Level (TRL) of your project? Higher TRL means closer to being in real-world use. See https://en.wikipedia.org/wiki/Technology_readiness_level. We will use the EU definition:

- TRL 1 – basic principles observed
- TRL 2 – technology concept formulated
- TRL 3 – experimental proof of concept
- TRL 4 – technology validated in lab
- TRL 5 – technology validated in relevant environment (industrially relevant environment in the case of key enabling technologies)
- TRL 6 – technology demonstrated in relevant environment (industrially relevant environment in the case of key enabling technologies)
- TRL 7 – system prototype demonstration in operational environment
- TRL 8 – system complete and qualified
- TRL 9 – actual system proven in operational environment (competitive manufacturing in the case of key enabling technologies; or in space)

In response, we found that most components were at a high TRL. The average was 6.8. The distribution is shown in Fig. 12.6. We will revisit TRLs below.

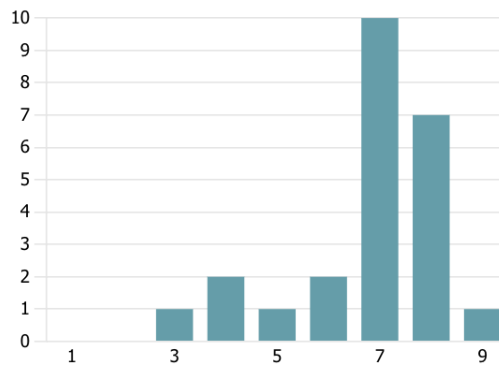


Figure 12.6: TRLs.

12.2.8. Question 14: Reuse strategies

This question is about strategies for ensuring exploitation and reuse. The strategies we suggested were:

- "Publicise via academic venues"
- "Publicise via social media"
- "Educate others in the use of the project"
- "Use the project in education"
- "Live services, eg a webapp maintained by you"
- "Third-party companies offer software as a service"
- "Deposit or index the data in well-known repositories"
- "Via GitHub - no commitment to maintenance or support"
- "Via GitHub - we have a commitment to maintain and support the project"
- "Via GitHub - a community maintains the software and provides support"
- "Other"

As shown in Fig. 12.7, almost all strategies were in place for different projects. Many project selected multiple strategies here.

14. What are your exploitation and reuse strategies? (Choose zero or more.)

[More Details](#)

- Publicise via academic venues 21
- Publicise via social media 12
- Educate others in the use of the ... 17
- Use the project in education 10
- Live services, eg a webapp main... 10
- Third-party companies offer soft... 0
- Deposit or index the data in well... 8
- Via GitHub - no commitment to ... 7
- Via GitHub - we have a commit... 11
- Via GitHub - a community maint... 5
- Other 2

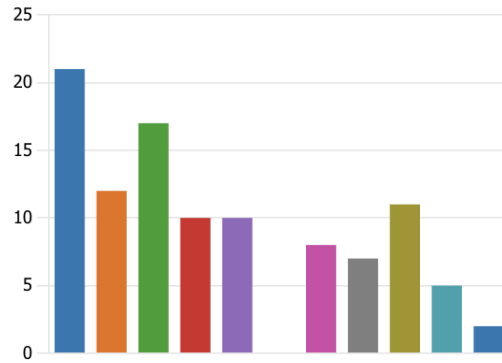


Figure 12.7: Reuse strategies.

12.2.9. Questions 15, 16 and 19: Reuse for specific project types

These questions asked about specific types of reuse for project types Data/Schema, Software, and User Interface, respectively. Results are shown in Figs. 12.8–12.10. The most natural types of reuse, such as reusing data in a similar task or domain, or reusing software by changing input data, or reusing a user interface by customising it, were the majority.

15. If your project is of type Data/Schema, please choose which type(s) of reuse are possible:

[More Details](#)

- Reuse in a similar domain 7
- Reuse in a similar task 7
- Reuse to produce new data conf... 4
- Other 0

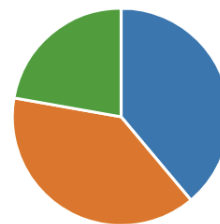


Figure 12.8: Reuse for Data/Schema projects.

16. If your project is of type Software, please choose which type(s) of reuse are possible:

[More Details](#)

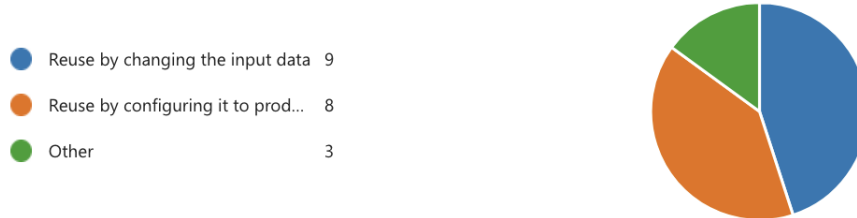


Figure 12.9: Reuse for Software projects.

19. If your project is of type User Interface, please choose which type(s) of reuse are possible:

[More Details](#)

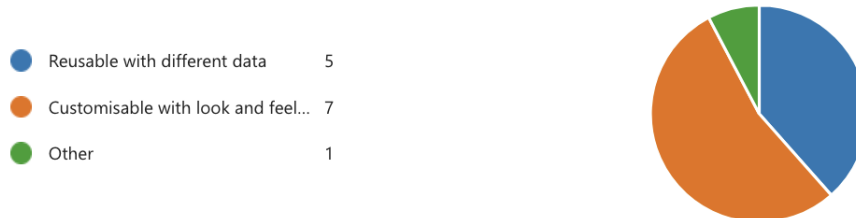


Figure 12.10: Reuse for User Interface projects.

12.2.10. Questions 17-18: Extensibility and domain independence

Question 17 asked about domain independence. 9 of the 24 responses claimed their project was domain-independence, i.e. not specific to the music domain.

Question 18 asked about extensibility. 3 components were not at all extensible, 15 were extensible, while 5 were specifically designed to be extended.

12.3. Technology Readiness Level

Table ?? is an extension of Table 1.3.2 in the original Grant Agreement. The first 4 columns are from the original table. The outputs here are the output concepts planned at the outset, which map, in the round, to the concrete components surveyed in Table 12.6. For each output here we show its type, as follows. MOD: a model specification; LIB: a reusable software component; API: an operating software exposed as Web API or Service; USE: a tool for our target consumers (citizens, artists, scholars, researchers). We also show the planned TRL. In the final column, which is new, we show the *Achieved* level. In every case, we have met or exceeded the expected TRL.

Table 12.1: Technology Readiness Levels: comparison between expected and achieved results. The expected results are shown in the Planned column in the format X/Y, where X is the *At least* TRL and Y the *Planned* TRL, as set out in the Grant Agreement

	Technology	Result	Type	Planned	Achieved
WP2	Music notation to RDF conversion pipeline	R2	API/LIB	5/6	6
WP2	Ontologies for the Musical Heritage domain	R1	MOD	5/7	7
WP2,4	Link discovery software	R2, R4	API/LIB	5/6	5
WP2	Transformation and inconsistency reasoning engine	R2	API/LIB	5/6	5
WP2	Music Knowledge Graph construction integrated software	R2	LIB	5/6	6
WP4	Software for knowledge extraction from text	R4	API/LIB	4/5	4
WP3	Software tool for pattern extraction	R3	LIB	4/5	6
WP3	Software tool for mining patterns in music repositories	R3	LIB	4/5	6
WP3	Software tool for extracting relations between musical patterns	R3	LIB	4/5	6
WP3	Software for music object classification	R3	LIB	4/5	5
WP1	Software to discover, index and query MH collections	R5	LIB	4/5	5
WP4	Plurilingual MH lexical corpus	R6	USE	4/6	4
WP5	Interfaces for interacting with MH assets and knowledge	R7	USE	4/6	6
WP5	Alternative modality interaction prototypes	R7, R8	USE	5/6	6
WP5	Visual Interaction software for MH experts	R7	USE	5/6	6
WP1, WP5	Web portal and registry of MH resources on the Web	R9, R10	USE	5/6	5
WP1, WP5	Meta-Search engine for MH resources	R9, R10	API/USE	5/6	5
WP1, WP5	Musical Patterns Explorer	R10	USE	5/6	6
WP6	Digital artistic installation	R13	LIB, USE	6/7	8

12.4. Conclusion

Overall, the results of the Polifonia Reusability Survey show that a large proportion of project outputs are reusable in natural ways, with projected TRL levels achieved and with substantial evidence of reuse outside the project already.

13. Compliance to the FAIR principles

For this deliverable we decided to list the information we usually have in the 'Compliance to the FAIR principles' section of the deliverable template as a specific subsection in each Pilot section. This way making the link to the Pilot's stronger, and providing a better overview than displaying a long table in one section. However, we would like to stress, that due to the finalisation of the Research Data Management Plan ([8], and the implementation and curation of the Polifonia Ecosystem, all of the Pilot's output can be also found via the Polifonia Ecosystem website. Moreover, the components (as central elements of the Polifonia Ecosystem) are also the units used for the reusability survey. The FAIR principles are part of the design of the Polifonia Ecosystem and of the annotation of the Polifonia Ecosystem components. This way FAIR principles are not a 'point to be checked' after the project has delivered output, but is already part of the knowledge production itself.

All components adhere to the Polifonia 10 Rules for Open Science [7].

14. Conclusions

In this (software) deliverable, we described the Research Ecosystem Framework (REECO), a suite of tools supporting the management of research objects in context. This includes to manage data in their context, but also goes beyond data. The context is interdisciplinary and innovative research collaborations executed in project funding. One could say that the REECO is a machine executable Data Management Plan. However, the aspiration of this approach goes beyond Data Management. It concerns in essence making explicit what are the fundamental knowledge units and actions needed in these kind of research collaborations.

Collaborations funded by projects come with their own dynamics. Their planning is already structured (usually in Workpackages and Tasks) and their accountability (scientifically and financially) is also regulated by a set of Milestones, Deliverables, and PersonMonths allocated to produce the promised output.

The REECO approach bridges between processes of knowledge production which are specific to each set of research questions, and the more administrative processes in the management of staff and other resources during research. It builds on formalisations which have been already proposed for the former (namely to describe concrete ingredients as research objects and to also define their inter-relatedness). It also embraces the formal structures existing for the latter, for project management. It does so by:

- finding the right agglomeration of research objects and relations - named components.
- providing a set of tools to create, annotate, evaluate and archive components
- creating an overview about the components and enable to monitor them

More concretely, the REECO encompasses the categories Data, Tools (software and applications) and Reports. The latter are representing specific documentations of output but also requirements which function as input for the research process. This way, the REECO extends Research Data Management, to managing Data in their Context.

In this DL we describe the generic approach of the REECO together with its first concrete implementation for the Polifonia project, in form of the Polifonia Ecosystem. This way, the REECO is made available for reuse and extension to future projects. For the Polifonia *use case or implementation* we also included details from a selection of Polifonia applications which support other ways in which musical data can be referenced and cited.

In Section 1 (see Table ??), we listed the REECO objectives and here we summarise how we addressed them (see Table 14.1).

The essence of the REECO is to find - collectively - a middle ground to identify those research components which form the cornerstones of a specific knowledge production process. The power of a formalisation of such cornerstones - named *components* in the REECO approach - is that main paths of transferring, combining and creating knowledge become visible, tangible and traceable. Finally, although we designed a rich metadata schema, metadata curation is considered in an open-ended way, where projects can extend the schema to incorporate new aspects and domain-specific concerns.

Table 14.1: Objections versus Solutions

Support Research Data Management	Data Management Plan Version 3 [8]
Present an organised view of project outputs to foster reuse beyond the project	http://github.com/reeco-framework/reeco-website-action A custom Github action to automate the generation of the ecosystem knowledge graph and website
Support an aware reuse of outputs by including copyright and licencing information	Specific field in the Annotation Schema http://github.com/reeco-framework/reeco-annotation-schema (see also [70])
Lower the barrier for users (developers, data scientists, researchers) to provide extensive metadata, building on good practices in (open-source) data/software development and research open data publishing	Python Library http://github.com/reeco-framework/reeco-python A software library in python to help processing and validating annotations. This is reused by various tools.
Facilitate the generation of metadata in different forms, to plug-in existing open science distributed infrastructures	Web Validator http://github.com/reeco-framework/reeco-web-repo-validator A utility to validate the annotations on a repository (running instance at http://reeco.kmi.open.ac.uk).

Bibliography

- [1] C. Guillotel-Nothmann, T. Bottini, V. Carriero, J. Carvalho, P. Cathé, F. Ciroku, E. Daga, M. Daquino, A. Davy-Rigaux, M. Gurrieri, P. Van Kemenade, E. Marzi, V. Presutti, and A. Scharnhorst, “D1.1 roadmap and pilot requirements 1st version,” Tech. Rep., Sep. 2022. [Online]. Available: <https://doi.org/10.5281/zenodo.7124253>
- [2] C. Guillotel-Nothmann, J. De Berardinis, T. Bottini, P. Cathé, E. Daga, M. Daquino, A. Davy-Rigaux, M. Gurrieri, P. Van Kranenburg, E. Marzi, J. McDermott, A. Meroño Peñuela, P. Mulholland, A. Scharnhorst, and R. Tripodi, “D1.2 roadmap and pilot requirements 2nd version,” Tech. Rep., Sep. 2022. [Online]. Available: <https://doi.org/10.5281/zenodo.7116561>
- [3] E. Daga, A. Meroño Peñuela, M. Daquino, F. Ciroku, E. Musumeci, M. Gurrieri, A. Scharnhorst, F. Admiraal, and R. Fournier-S’niehotta, “D1.3: Pilots development – collaborative methodology and tools (V1.0),” Tech. Rep., 2021. [Online]. Available: <https://doi.org/10.5281/zenodo.7712909>
- [4] P. Van Kranenburg, “Polifonia: A digital harmoniser for musical heritage knowledge,” Jan. 2024. [Online]. Available: <https://doi.org/10.5281/zenodo.10563013>
- [5] A. Scharnhorst, P. van Kranenburg, E. Daga, and M. Bontje, “D7.5 dmp webinar on long term preservation:,” Tech. Rep., Feb. 2024. [Online]. Available: <https://doi.org/10.5281/zenodo.10599999>
- [6] F. Admiraal, A. Scharnhorst, E. Daga, M. Daquino, E. Musumeci, P. Kranenburg, C. Guillotel-Nothmann, M. Gurrieri, V. Presutti, M. Clementi, A. Meroño Peñuela, M. Turci, E. Marzi, A. Puglisi, and R. Fournier-S’niehotta, “D7.1 first data management plan,” Aug. 2022. [Online]. Available: <https://doi.org/10.5281/zenodo.6963585>
- [7] A. Scharnhorst, R. Van Horik, E. Daga, M. Daquino, E. Musumeci, P. Van Kranenburg, C. Guillotel-Nothmann, M. Gurrieri, V. Presutti, M. Clementi, A. Meroño Peñuela, M. Turci, E. Marzi, A. Puglisi, and R. Fournier-S’niehotta, “D7.2 data management plan (second version),” Feb. 2023. [Online]. Available: <https://doi.org/10.5281/zenodo.7660299>
- [8] A. Scharnhorst, E. Daga, R. Fournier-S’niehotta, M. Gurrieri, J. de Berardinis, J. Mc Dermott, R. Tripodi, M. Daquino, G. Renda, A. Poltronieri, E. Musumeci, F. Magnani, P. van Kranenburg, C. Guillotel-Nothmann, T. Bottini, F. Ciroku, M. Clementi, A. Meroño Peñuela, M. Turci, E. Marzi, A. Puglisi, C. Uwasomba, J. Carvalho, A. Morales-Tirado, M. Bontje, R. de Groot, and V. Presutti, “D7.3 polifonia data management plan (final version),” Dec. 2023. [Online]. Available: <https://doi.org/10.5281/zenodo.7660299>
- [9] J. de Berardinis, A. Meroño-Peñuela, A. Poltronieri, and V. Presutti, “Choco: a chord corpus and a data transformation workflow for musical harmony knowledge graphs,” *Scientific Data*, vol. 10, no. 1, p. 641, 2023.
- [10] —, “The harmonic memory: a knowledge graph of harmonic patterns as a trustworthy framework for computational creativity,” in *Proceedings of the ACM Web Conference 2023*, 2023, pp. 3873–3882.
- [11] J. de Berardinis, A. M. Peñuela, A. Poltronieri, V. Presutti, R. Fournier-S’niehotta, P. Rigaux, and T. Zhu, “D1.5 intermediate validation reports for pilots: Interlink and facets (v1.0),” European Commission, The Polifonia consortium, Tech. Rep., 2022.
- [12] KCL, “D2.5: Methods for interlinking knowledge graphs (v1.0),” February 2024.
- [13] E. J. Humphrey, J. Salamon, O. Nieto, J. Forsyth, R. M. Bittner, and J. P. Bello, “JAMS: A JSON Annotated Music Specification for Reproducible MIR Research.” in *International Society for Music Information Retrieval (ISMIR)*, 2014, pp. 591–596.
- [14] C. Harte, M. B. Sandler, S. A. Abdallah, and E. Gómez, “Symbolic Representation of Musical Chords: A Proposed Syntax for Text Annotations,” in *International Society for Music Information Retrieval (ISMIR)*, vol. 5, 2005, pp. 66–71.

- [15] W. B. de Haas, R. C. Veltkamp, and F. Wiering, "Tonal Pitch Step Distance: a Similarity Measure for Chord Progressions," in *ISMIR 2008, 9th International Conference on Music Information Retrieval, Drexel University, Philadelphia, PA, USA, September 14-18, 2008*, J. P. Bello, E. Chew, and D. Turnbull, Eds., 2008, pp. 51–56. [Online]. Available: http://ismir2008.ismir.net/papers/ISMIR2008_252.pdf
- [16] W. B. de Haas, F. Wiering, and R. C. Veltkamp, "A geometrical distance measure for determining the similarity of musical harmony," *International Journal of Multimedia Information Retrieval*, vol. 2, no. 3, pp. 189–202, 09 2013. [Online]. Available: <https://doi.org/10.1007/s13735-013-0036-6>
- [17] A. Bordes, N. Usunier, A. García-Durán, J. Weston, and O. Yakhnenko, "Translating embeddings for modeling multi-relational data," in *Advances in Neural Information Processing Systems 26: 27th Annual Conference on Neural Information Processing Systems 2013. Proceedings of a meeting held December 5-8, 2013, Lake Tahoe, Nevada, United States*, C. J. C. Burges, L. Bottou, Z. Ghahramani, and K. Q. Weinberger, Eds., 2013, pp. 2787–2795. [Online]. Available: <https://proceedings.neurips.cc/paper/2013/hash/1cecc7a77928ca8133fa24680a88d2f9-Abstract.html>
- [18] J. de Berardinis, V. A. Carriero, N. Jain, N. Lazzari, A. Meroño-Peñuela, A. Poltronieri, and V. Presutti, "The polifonia ontology network: Building a semantic backbone for musical heritage," in *International Semantic Web Conference*. Springer, 2023, pp. 302–322.
- [19] J. de Berardinis, V. Carriero, A. Meroño-Peñuela, A. Poltronieri, and V. Presutti, "The music meta ontology: a flexible semantic model for the interoperability of music metadata," in *Proc. of the 24th Int. Society for Music Information Retrieval Conference*, 2023.
- [20] J. de Berardinis, A. Meroño-Peñuela, A. Poltronieri, and V. Presutti, "The music annotation pattern," *arXiv preprint arXiv:2304.00988*, 2023.
- [21] J. McDermott, D. Diamond, A. S. Khatkhat, P. van Krannenburg, J. de Berardinis, A. Poltronieri, A. M. Peñuela, R. Fournier-S'niehotta, and T. Zhu, "D3.2: Analysis of music repositories to identify musical pattern - 1st version (v1.0)," Sep. 2021. [Online]. Available: <https://doi.org/10.5281/zenodo.7116711>
- [22] T. Zhu, R. Fournier-S'niehotta, P. Rigaux, and N. Travers, "A framework for content-based search in large music collections," *Big Data and Cognitive Computing*, vol. 6, no. 1, 2022. [Online]. Available: <https://www.mdpi.com/2504-2289/6/1/23>
- [23] A. Filaber, A. Graciotti, C. Guillotel-Nothmann, M. Gurrieri, P. van Krannenburg, E. Marzi, and C. F. Uwasomba, "D1.6 intermediate validation reports for pilots tonalities, tunes, musicbo and child. 1st version," 2023. [Online]. Available: <https://shs.hal.science/halshs-04079571v1>
- [24] C. Guillotel-Nothmann, "Nested annotations in tonalities: Annotation of a cadence, with its individual melodic lines and cadence steps," 2024. [Online]. Available: <https://shs.hal.science/hal-04435142v1>
- [25] M. Gurrieri, "Annotation of large selections in tonalities," 2024. [Online]. Available: <https://shs.hal.science/halshs-04522940>
- [26] e. a. Horrocks, Ian, "Swrl: A semantic web rule language combining owl and ruleml," *W3C Member submission*, pp. 1–31, 2004.
- [27] e. a. Lazzari, Nicolas, "Sandra - a neuro-symbolic reasoner based on descriptions and situations," *arXiv preprint arXiv:2402.00591*, 2024.
- [28] C. Guillotel-Nothmann and T. Bottini, "L'Inférence de connaissances dans le projet PolyMIR : une approche hypothético-déductive pour l'analyse de patterns contrapuntiques complexes," in *Épistémologie de la musicologie numérique*, Paris, France, Apr. 2018. [Online]. Available: <https://hal.science/hal-02263255>
- [29] A.-E. Ceulemans and C. Guillotel-Nothmann, "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*, ser. Wolfenbütteler Forschungen 172, S. Limbeck, R. Schmitt, and S. Wirth, Eds. Harrassowitz, 2022, pp. 109–150. [Online]. Available: <https://hal.science/hal-03087363>
- [30] UNIBO, "D4.1: Plurilingual corpora containing source texts in English, French, Spanish and German (v1.0)," December 2021.

- [31] —, “D4.2: Interrogation and annotation of plurilingual corpora for discourse analysis (v1.0),” June 2022.
- [32] A. Gangemi, A. Graciotti, A. Meloni, A. Nuzzolese, V. Presutti, D. Reforgiato Recupero, A. Russo, and R. Tripodi, “Text2AMR2FRED, a tool for transforming text into RDF/OWL Knowledge Graphs via Abstract Meaning Representation,” in *22nd ISWC*. Athens, Greece: CEUR Workshop Proc., November 2023.
- [33] UNIBO, “D4.5: Software for knowledge extraction from text – context – 2nd version (v1.0),” August 2022.
- [34] A. Gangemi, V. Presutti, D. R. Recupero, A. G. Nuzzolese, F. Draicchio, and M. Mongiovi, “Semantic Web Machine Reading with FRED,” *Semantic Web*, vol. 8, no. 6, pp. 873–893, 2017. [Online]. Available: <https://doi.org/10.3233/SW-160240>
- [35] M. Bevilacqua, R. Blloshmi, and R. Navigli, “One SPRING to Rule Them Both: Symmetric AMR Semantic Parsing and Generation without a Complex Pipeline,” *Proc. of the AAAI Conf. on Artificial Intelligence*, vol. 35, no. 14, pp. 12 564–12 573, May 2021. [Online]. Available: <https://ojs.aaai.org/index.php/AAAI/article/view/17489>
- [36] R. Orlando, S. Conia, S. Faralli, and R. Navigli, “Universal Semantic Annotator: the First Unified API for WSD, SRL and Semantic Parsing,” in *Proc. of LREC 2022*. Marseille, France: European Language Resources Association, June 2022, pp. 2634–2641. [Online]. Available: <https://aclanthology.org/2022.lrec-1.282>
- [37] A. Meloni, D. Reforgiato Recupero, and A. Gangemi, “AMR2FRED, A Tool for Translating Abstract Meaning Representation to Motif-Based Linguistic Knowledge Graphs,” in *The Semantic Web: ESWC 2017 Satellite Events*. Portorož, Slovenia: Springer International Publishing, 2017, pp. 43–47.
- [38] A. Gangemi, M. Alam, L. Asprino, V. Presutti, and D. R. Recupero, “Framester: A Wide Coverage Linguistic Linked Data Hub,” in *EKAW 2016*. Bologna, Italy: Springer International Publishing, 2016, pp. 239–254.
- [39] L. Wu, F. Petroni, M. Josifoski, S. Riedel, and L. Zettlemoyer, “Scalable Zero-shot Entity Linking with Dense Entity Retrieval,” in *Proc. of the 2020 Conf. on Empirical Methods in Natural Language Processing (EMNLP)*, B. Webber, T. Cohn, Y. He, and Y. Liu, Eds. Online: ACL, November 2020, pp. 6397–6407. [Online]. Available: <https://aclanthology.org/2020.emnlp-main.519>
- [40] A. Q. Jiang, A. Sablayrolles, A. Mensch, C. Bamford, D. S. Chaplot, D. d. I. Casas, F. Bressand, G. Lengyel, G. Lample, L. Saulnier *et al.*, “Mistral 7b,” *arXiv preprint arXiv:2310.06825*, 2023.
- [41] KCL, “D2.8: Ontology testing and evaluation report; knowledge graph documentation and access interfaces,” January 2024.
- [42] N. Reimers and I. Gurevych, “Sentence-bert: Sentence embeddings using siamese bert-networks,” in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 11 2019. [Online]. Available: <https://arxiv.org/abs/1908.10084>
- [43] R. Sennrich, B. Haddow, and A. Birch, “Improving Neural Machine Translation Models with Monolingual Data,” in *Proc. of the 54th Annual Meeting of the ACL (Volume 1: Long Papers)*. Berlin, Germany: ACL, August 2016, pp. 86–96. [Online]. Available: <https://aclanthology.org/P16-1009>
- [44] A. Morales Tirado, A. Graciotti, and E. Daga, “D4.3: Software for knowledge extraction from text – entities – 1st version,” Sep 2022.
- [45] A. Morales Tirado, J. Carvalho, A. Graciotti, and E. Daga, “D4.4: Software for knowledge extraction from text – entities – 2nd version,” Oct 2023.
- [46] E. Daga and E. Motta, “Challenging knowledge extraction to support the curation of documentary evidence in the humanities,” in *Third International Workshop on Capturing Scientific Knowledge (Sciknow). Collocated with the tenth International Conference on Knowledge Capture (K-CAP)*, D. Garijo, M. Markovic, P. Groth, I. Santana, and K. Belhajjame, Eds., Los Angeles, CA, USA, November 2019. [Online]. Available: <http://oro.open.ac.uk/67961/>
- [47] A. Adamou, M. d’Aquin, H. Barlow, and S. Brown, “Led: curated and crowdsourced linked data on music listening experiences,” 2014.
- [48] M. Bradley, “The reading experience database,” *Journal of Victorian Culture*, vol. 15, no. 1, pp. 151–153, 2010.
- [49] J. Achiam, S. Adler, S. Agarwal, L. Ahmad, I. Akkaya, F. L. Aleman, D. Almeida, J. Altenschmidt, S. Altman, S. Anadkat *et al.*, “Gpt-4 technical report,” *arXiv preprint arXiv:2303.08774*, 2023.

- [50] M. U. Hadi, R. Qureshi, A. Shah, M. Irfan, A. Zafar, M. B. Shaikh, N. Akhtar, J. Wu, S. Mirjalili *et al.*, “A survey on large language models: Applications, challenges, limitations, and practical usage,” *Authorea Preprints*, 2023.
- [51] Z. Guo, R. Jin, C. Liu, Y. Huang, D. Shi, Supryadi, L. Yu, Y. Liu, J. Li, B. Xiong, and D. Xiong, “Evaluating large language models: A comprehensive survey,” 2023.
- [52] Y. Chang, X. Wang, J. Wang, Y. Wu, L. Yang, K. Zhu, H. Chen, X. Yi, C. Wang, Y. Wang *et al.*, “A survey on evaluation of large language models,” *ACM Transactions on Intelligent Systems and Technology*, 2023.
- [53] H. Zhao, H. Chen, F. Yang, N. Liu, H. Deng, H. Cai, S. Wang, D. Yin, and M. Du, “Explainability for large language models: A survey,” *ACM Transactions on Intelligent Systems and Technology*, 2023.
- [54] M. Mitchell and D. C. Krakauer, “The debate over understanding in ai’s large language models,” *Proceedings of the National Academy of Sciences*, vol. 120, no. 13, p. e2215907120, 2023.
- [55] Y. Liu, T. Han, S. Ma, J. Zhang, Y. Yang, J. Tian, H. He, A. Li, M. He, Z. Liu *et al.*, “Summary of chatgpt-related research and perspective towards the future of large language models,” *Meta-Radiology*, p. 100017, 2023.
- [56] Y. Zhu, H. Yuan, S. Wang, J. Liu, W. Liu, C. Deng, Z. Dou, and J.-R. Wen, “Large language models for information retrieval: A survey,” *arXiv preprint arXiv:2308.07107*, 2023.
- [57] P. Kumar and S. Kathuria, “Large language models (llms) for natural language processing (nlp) of oil and gas drilling data,” in *SPE Annual Technical Conference and Exhibition?* SPE, 2023, p. D021S012R004.
- [58] R. Tang, Y.-N. Chuang, and X. Hu, “The science of detecting llm-generated texts,” *arXiv preprint arXiv:2303.07205*, 2023.
- [59] D. Galat and M.-A. Rizoiu, “Enhancing biomedical text summarization and question-answering: On the utility of domain-specific pre-training,” *arXiv preprint arXiv:2307.04412*, 2023.
- [60] D. Van Veen, C. Van Uden, L. Blankemeier, J.-B. Delbrouck, A. Aali, C. Bluethgen, A. Pareek, M. Polacin, W. Collins, N. Ahuja *et al.*, “Clinical text summarization: adapting large language models can outperform human experts,” *arXiv preprint arXiv:2309.07430*, 2023.
- [61] J. Wei, Y. Tay, R. Bommasani, C. Raffel, B. Zoph, S. Borgeaud, D. Yogatama, M. Bosma, D. Zhou, D. Metzler *et al.*, “Emergent abilities of large language models,” *arXiv preprint arXiv:2206.07682*, 2022.
- [62] E. Daga, “polifonia-project/documentary-evidence-benchmark:,” Jul. 2022. [Online]. Available: <https://doi.org/10.5281/zenodo.6866982>
- [63] E. Daga and E. Motta, “Capturing themed evidence, a hybrid approach,” in *Proceedings of the 10th International Conference on Knowledge Capture*, 2019, pp. 93–100.
- [64] Y. Wang, Q. Yao, J. T. Kwok, and L. M. Ni, “Generalizing from a few examples: A survey on few-shot learning,” *ACM computing surveys (csur)*, vol. 53, no. 3, pp. 1–34, 2020.
- [65] A. M. Tirado and E. Daga, “polifonia-project/meetups_pilot: v0.1,” Jul. 2022. [Online]. Available: <https://doi.org/10.5281/zenodo.6866860>
- [66] A. Morales Tirado and E. Daga, “polifonia-project/meetups-knowledge-graph: v0.2,” May 2023. [Online]. Available: <https://doi.org/10.5281/zenodo.7924618>
- [67] —, “polifonia-project/meetups-ontology: v0.1,” May 2023. [Online]. Available: <https://doi.org/10.5281/zenodo.7928155>
- [68] A. Morales Tirado, J. Carvalho, M. Ratta, C. Uwasomba, P. Mulholland, H. Barlow, T. Herbert, and E. Daga, “Musical Meetups Knowledge Graph (MMKG): a collection of evidence for historical social network analysis,” in *European Semantic Web Conference*. Springer, 2024.
- [69] A. Morales Tirado, J. Carvalho, P. Mulholland, and E. Daga, “Musical meetups: a knowledge graph approach for historical social network analysis,” in *Proceedings of the ESWC 2023 Workshops and Tutorials co-located with 20th European Semantic Web Conference (ESWC 2023)*, May 2023.
- [70] E. Daga, J. Carvalho, M. Gurrieri, and A. Scharnhorst, “D2.6: Ontology of licencing, ownership and conditions of use (v1.0),” Tech. Rep., Feb. 2023.

A. Polifonia Reusability Survey

Polifonia Reusability Survey

24 Responses

13:27 Average time to complete

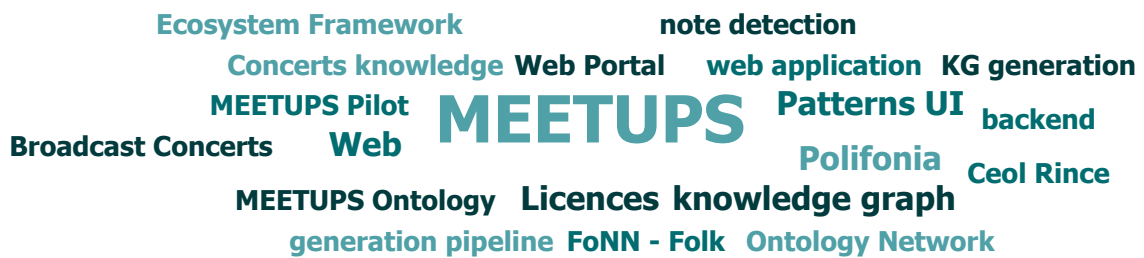
Active Status

1. What is the name of your Polifonia project or component? (If you have more than one, even if they are related, you may wish to fill in the survey once for each.)

24
Responses

Latest Responses
"PySPARQL-Anything"
"Licences KG generation pipeline"
"Licences Knowledge Graph"

4 respondents (17%) answered **MEETUPS** for this question.



2. If there is a GitHub repository of the component/project, please add it here

24
Responses

Latest Responses
"http://w3id.org/polifonia/ecosystem/sparql-anything-python"
"http://w3id.org/polifonia/ecosystem/licences-pipeline"
"http://w3id.org/polifonia/ecosystem/licences"

3. Please include the Ecosystem id of either the component or the container (if more components)

24
Responses

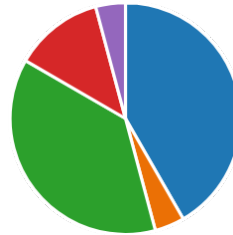
Latest Responses
"sparql-anything-python"
"licences-pipeline"
"licences"

4 respondents (17%) answered **knowledge-graph** for this question.



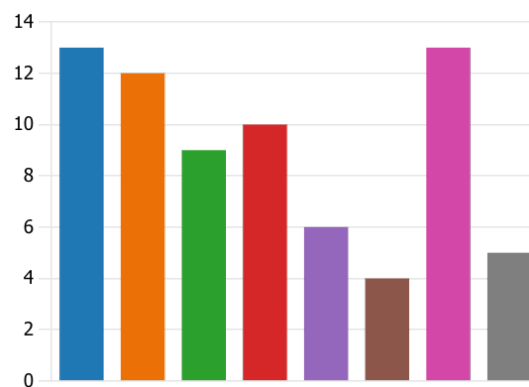
4. Which of these best describes your project type?

● Software	10
● Data	1
● Knowledge graph, ontology, or ...	9
● User interface	3
● Other	1



5. Who is the target audience for your project? Eg developers, or only Polifonia developers; musicologists; historians; etc.

● Developers (in the project team)	13
● Developers (any)	12
● Domain Experts	9
● Computational Musicologists	10
● Musicologists	6
● Historians	4
● Researchers	13
● Other	5



6. How reusable is your project in the following five aspects?

When considering your **breadth of project goals**: if the project goals are highly specific to you (e.g. software not reusable because it is one-off curatorial work), and would not be of interest to others, this makes it less reusable. If the project goals address the goals of a broad class of users/tasks, this makes it more reusable.

When considering its **reusability with respect to technical knowledge**: if the user would need substantial, specific technical knowledge to reuse it, this makes it less reusable. Example: expertise in coding in C. If they would need to visit your lab or call you to get started, this makes it less reusable. If the project provides good "getting started" documentation, this makes it more reusable.

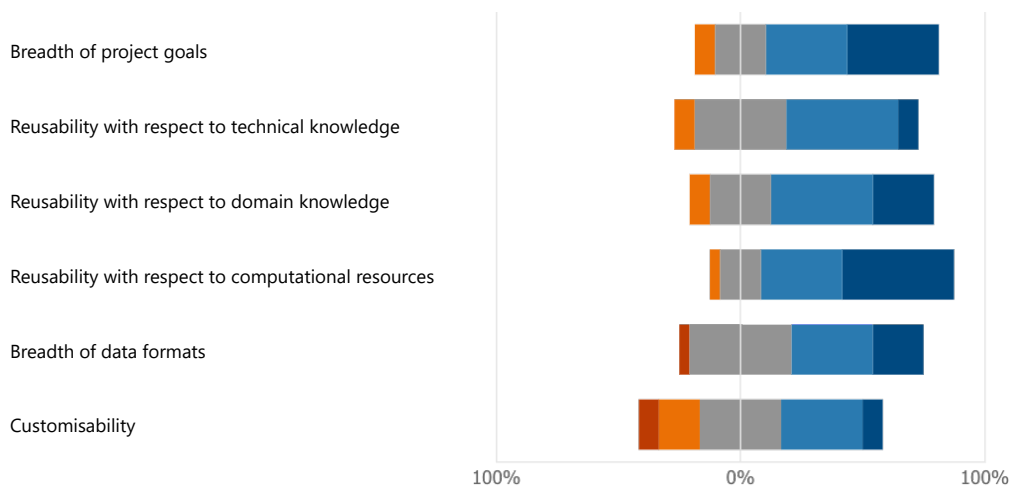
When considering its **reusability with respect to domain knowledge**: if the user would need substantial, specific domain knowledge to reuse it, this makes it less reusable. Example: expertise in musicology of 16th century Italian folk music.

When considering its **reusability with respect to computational resources**: if the user would need substantial, specific computational resources to reuse it, this makes it less reusable. Examples: requirement for a server with large RAM, a GPU, or web hosting.

When considering its **customisability**: if the project can be easily extended, customised, or re-branded, this makes it more reusable.

For the following statements, **higher is better (more reusable)**.

■ Not reusable ■ Slightly reusable ■ Somewhat reusable ■ Quite reusable ■ Highly reusable



7. Has your project already been reused by someone other than its core developers (ie, you or your group)? Even if that reuse is not "complete", please include it.

● No	12
● Yes, reused within Polifonia	6
● Yes, reused outside Polifonia	6



8. If the output is reused outside Polifonia, please provide some details, possibly with links to evidence.

22
Responses

Latest Responses

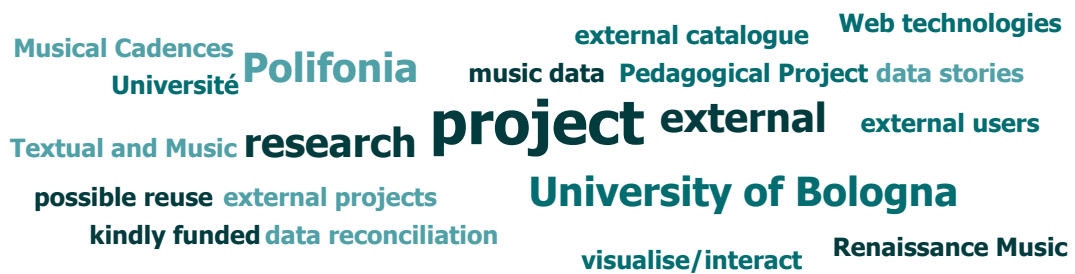
"The library is published open-source, we have evidence of adoption by se...

"N/A"

"N/A"

[Update](#)

3 respondents (13%) answered **project** for this question.



9. Please state the technical knowledge requirements as a list of comma-separated standards, languages (e.g.: "XML, RDF, Python, Java, SPARQL")

24
Responses

Latest Responses

"Python, SPARQL, RDF"

"Python"

"RDF, SPARQL"

15 respondents (63%) answered **SPARQL** for this question.



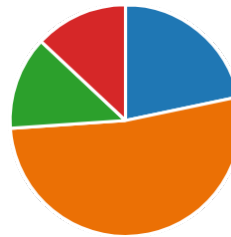
10. Specify what type and amount of formats the project deals with?

● Custom, ad-hoc input formats	3
● Well-known / standard input for...	13
● Single format	6
● Multiple formats	6
● Other	2



11. How much the project/component is able to tolerate or provide support for dealing with imperfect, incomplete, or low-quality data?

● Data completeness/quality affec...	5
● Requires a format but can handl...	12
● Deals with fully unstructured dat...	3
● The question doesn't apply to th...	3



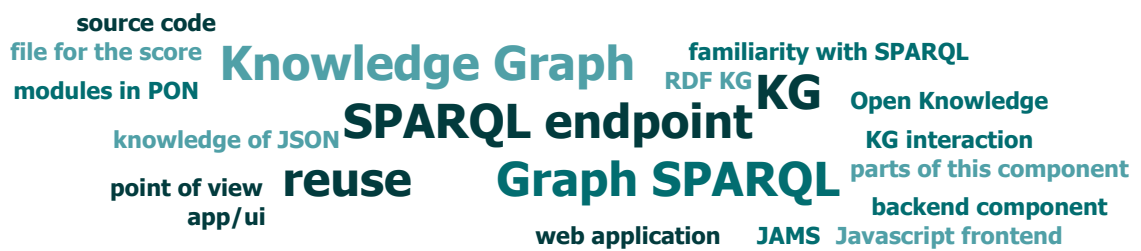
12. If you would like to give any comments to expand on your answers above, please enter them here. Eg, if you have constraints on data formats or computational requirements which are not captured by previous questions.

7

Responses

Latest Responses

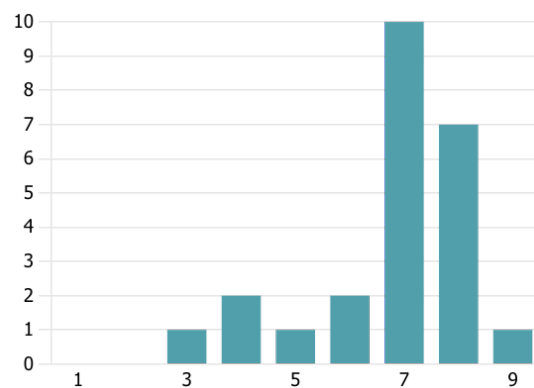
2 respondents (29%) answered **SPARQL endpoint** for this question.



13. What is the approximate **Technology Readiness Level (TRL)** of your project? Higher TRL means closer to being in real-world use. See https://en.wikipedia.org/wiki/Technology_readiness_level. We will use the EU definition:

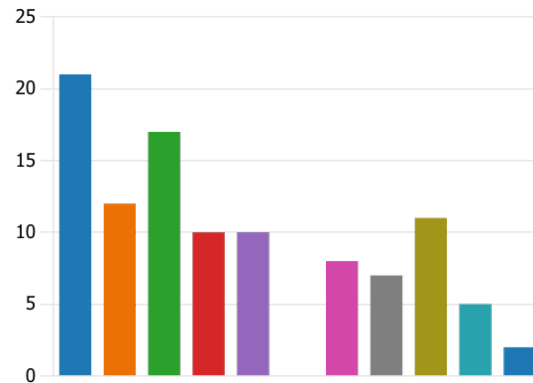
- TRL 1 – basic principles observed
- TRL 2 – technology concept formulated
- TRL 3 – experimental proof of concept
- TRL 4 – technology validated in lab
- TRL 5 – technology validated in relevant environment (industrially relevant environment in the case of key enabling technologies)
- TRL 6 – technology demonstrated in relevant environment (industrially relevant environment in the case of key enabling technologies)
- TRL 7 – system prototype demonstration in operational environment
- TRL 8 – system complete and qualified
- TRL 9 – actual system proven in operational environment (competitive manufacturing in the case of key enabling technologies; or in space)

6.79
Average Rating



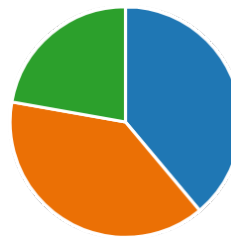
14. What are your exploitation and reuse strategies? (Choose zero or more.)

● Publicise via academic venues	21
● Publicise via social media	12
● Educate others in the use of the ...	17
● Use the project in education	10
● Live services, eg a webapp main...	10
● Third-party companies offer soft...	0
● Deposit or index the data in well...	8
● Via GitHub - no commitment to ...	7
● Via GitHub - we have a commit...	11
● Via GitHub - a community maint...	5
● Other	2



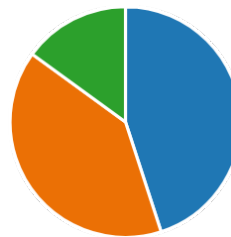
15. If your project is of type Data/Schema, please choose which type(s) of reuse are possible:

● Reuse in a similar domain	7
● Reuse in a similar task	7
● Reuse to produce new data conf...	4
● Other	0



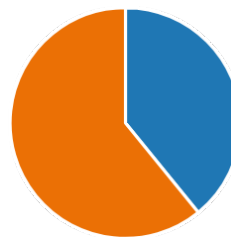
16. If your project is of type Software, please choose which type(s) of reuse are possible:

● Reuse by changing the input data	9
● Reuse by configuring it to prod...	8
● Other	3



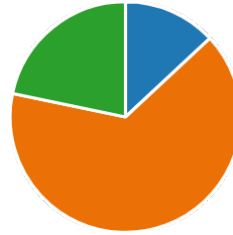
17. Is your project/component domain independent?

● Yes	9
● No (e.g. it is specific to Music)	14



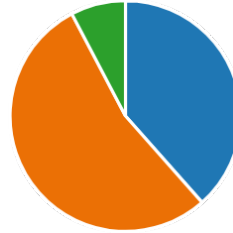
18. How much the data or software can be extended?

● Not at all	3
● It can be	15
● It is supposed to	5



19. If your project is of type User Interface, please choose which type(s) of reuse are possible:

● Reusable with different data	5
● Customisable with look and feel...	7
● Other	1



20. Please give your name so that we can contact you in case we need to clarify any responses.

24
Responses

Latest Responses

"*Enrico Daga*"

"*Enrico Daga*"

"*Enrico Daga*"

5 respondents (21%) answered **Enrico Daga** for this question.

Marilena Daquino **Marco Grasso** **Giulia Renda**
Jason Carvalho **Enrico Daga** **Jacopo**
Gurrieri **Paul Mulholland**
Mari Wigham **James McDermott**