



Polifonia: a digital harmoniser for musical heritage knowledge, H2020

D1.10: Polifonia Web portal – 2nd Version

Deliverable information	
WP	WP1 Pilots and Web portal
Document dissemination level	PU Public
Deliverable type	R Websites
Lead beneficiary	UNIBO
Contributors	OU, KCL, NUIG, KNAW, DP
Document status	Final
Document version	V1.0
Date	26.02.2024
Authors	Marilena Daquino - UNIBO, University of Bologna; Giulia Renda - UNIBO, University of Bologna; Marco Grasso - UNIBO, University of Bologna



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

INTENTIONALLY BLANK PAGE

Project information

Project Start Date: 1st of January 2021

Project Duration: 40 months

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

Project contacts

Project Coordinator

Valentina Presutti

ALMA MATER STUDIORUM - UNIVERSITÀ DI BOLOGNA

Department of Language, Literature and Modern Cultures (LILEC)

E-mail: valentina.presutti@unibo.it

Project Manager

Marta Clementi

ALMA MATER STUDIORUM - UNIVERSITÀ DI BOLOGNA

Research division

E-mail: marta.clementi3@unibo.it

POLIFONIA consortium

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

Project Summary

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

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

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

Executive summary

Deliverable 1.10 describes activities and results of Work Package 1, Task 3 (WP1.T3). These include the development of an aggregator of digital musical heritage collections (later called musoW), a number of services for indexing, querying and browsing music data, and interfaces for experts (later called MELODY) and lay users (i.e. the Web portal). This deliverable builds on state of the art, methodology, approach, and preliminary results described in D1.9. In this document we report on updates of previously developed technologies, and we present the final prototype of the Web portal.

In detail, we report on:

- **Crowdsourcing Linked Entities via web Form (CLEF)**, the web application for cataloguing music data (usage) and **Music data on the Web (musoW)**, the online catalogue of music data (status) - see Section 4.1
- **Make me a Linked Open Data Story (MELODY)** the platform for music data analysis and storytelling (evaluation and usage) - see Section 4.2
- **Web portal**, a harmonizer and access point to Polifonia datasets (tests, evaluation and case studies) - see Section 4.3 and 4.4

Finally, we discuss the impact, limitations, and future plans of results of WP1 Task 3.

Document History

Version	Release date	Summary of changes	Author(s) -Institution
V0.1	2/01/2024	First draft released	UNIBO
V0.2	25/01/2024	Revisions and integration added in all sections after internal review	UNIBO, OU, KNAW, DP
V0.3	30/01/2024	Version released to all partners for final integrations	All
V1.0	26/02/2024	Final version submitted to REA	UNIBO

Table of contents

Project information	3
Project contacts	3
POLIFONIA consortium.....	3
Project Summary	4
Executive summary	5
Document History	6
1. Introduction.....	8
2. State of the art.....	10
3. Research problems, objectives, and approach.....	11
3.1 Research problems and objectives.....	11
3.2 Approach and workplan.....	11
4. Software outputs, results and evaluation	13
4.1 musoW and CLEF.....	13
4.2 MELODY	15
4.3 The Web portal.....	26
4.4 Case studies: Explore music heritage through the Web portal	44
5. Discussion and Conclusion	51
5.1 Impact and sustainability of results	51
5.2 Limitations and future works	51
6. Compliance to the FAIR principles	53
6.1 Research Ecosystem Components	53
Bibliography	55

1. Introduction

The second and final deliverable of Work Package 1, Task 3 (WP1.T3) focuses on activities relevant to (1) the improvement and evaluation of CLEF, musoW, and MELODY applications, (2) the development and preliminary evaluation of a prototype of the web portal, and (3) dissemination activities and sustainability plans.

The remainder of the report is the following. In section 2 we recall a few works from the state of the art to demonstrate that the work we pursued is tackling important gaps. In section 3 we remind the reader our goals, and we describe the approach we adopted to finalise aforementioned research products and validate our work. Activities, results, and their evaluation are presented in section 4 and discussed in section 5, where we highlight impact-driven actions and limitations. Components (i.e. CLEF, musoW, MELODY, and the Web portal) are accredited and monitored via the Polifonia Research Ecosystem, as part of Polifonia Data Management, and are detailed in section 6.

In summary, the following outputs and activities are presented.

The crowdsourcing platform **Crowdsourcing Linked Entities via web Forms (CLEF)**, the catalogue of **music data on the Web (musoW)**, and the platform for data storytelling **Make mE a Linked Open Data storY (MELODY)** have been developed during the first two years of the project (mid 2021-mid 2022) to respectively facilitate access to music data sources on the web and to support scholars in producing article-like stories based on their data. Follow-up activities connected to such outputs include user tests, evaluation, bug fixing, development, and maintenance - respectively described in Sections 4.1 (CLEF and musoW) and 4.2 (MELODY).

The Polifonia Web portal is a web application designed to allow lay people accessing data sources used and produced by Polifonia partners (see section 4.3). During the last year we have refined and improved mockups described in deliverable 1.9 (Daquino et al. 2022) thanks to iterative user tests, co-design sessions with lay people and experts' feedback, and we have developed a prototype web application¹ leveraging an initial set of data sources.

Scholarly publications in venues relevant to Information Science and Digital Humanities (also targeted to multidisciplinary fields, such as digital musicology and digital music history) have been successfully submitted and presented in national and international conferences. National and international stakeholders have been involved in the definition of new requirements, which currently guide the development and refinement of applications like CLEF and MELODY.

¹ <https://polifonia.disi.unibo.it/portal>

Recently awarded project proposals (PRIN2022²) and commissioned research (GEL,³ funded by UNESCO; KNOT⁴, funded by the Ministry of Italian Culture (MIC); Epigraphs⁵, funded by the University of Bologna) ensure sustainability of results in the mid-term, thanks to newly recruited personnel that continue working on developed technologies and develop new results on top of them (Polifonia output CLEF).

Long-term collaborations within national and international consortia PNRR CHANGES ⁶, PHAROS⁷ (2022-2025) have been established to customize existing data storytelling solutions (Polifonia output MELODY), and to guarantee maintenance and further development of methods for data reconciliation (Polifonia output Web portal).

² PRIN2022 Is an Italian funding program. The project ATLAS (<https://dh-atlas.github.io/>) is a two-year project (October 2023-October 2025) awarded to the University of Bologna to develop a knowledge graph of Italian Digital Humanities projects. CLEF has been proposed as a software solution to crowdsource data and host the final catalogue.

³ Global Education and Learning (GEL) database (<https://projects.dharc.unibo.it/digestgel/>) is a UNESCO funded project to publish a bibliographic catalogue of scholarly works on the topic of education to global citizenship. It uses and extends CLEF to publish the catalogue.

⁴ KNOT is an ongoing three-year pilot project (2022-2025) from the University of Bologna and the Central Institute for the Digitization of Cultural Heritage (ICDP) - Digital Library of the Italian Ministry of Culture to showcase the digital cultural heritage of Italian universities- (<https://icdp-digital-library.github.io/KNOT/>). CLEF is used to prototype the catalogue or research projects.

⁵ [Epigrafi Latine a Bologna \(https://projects.dharc.unibo.it/el\)](https://projects.dharc.unibo.it/el) is a project funded by the University of Bologna to track latin epigraphs around the city of Bologna. The project actively uses and customises CLEF to collect data and showcase results.

⁶ CHANGES (<https://sites.google.com/uniroma1.it/changes/home>) is a consortium of universities, research centers and schools gathered to develop methods for knowledge and technological transfer in the Cultural Heritage and Humanities domains. The Consortium has requested the development of new features of MELODY to use it as a gateway for displaying cultural analytics on the heritage preserved in a number of cultural institutions.

⁷ PHAROS (<http://pharosartresearch.org/>) is an international consortium of 14 photo archives that aims at creating a web portal for art historians wherein to search among millions of photographs of artworks. To provide a better user experience, data across institutes must be aligned (e.g. artworks of same artists). The consortium is interested in reusing methods developed for the Polifonia web portal and has recently funded a mid-term collaboration with the University of Bologna to continue developing such methods.

2. State of the art

A complete account of the literature review relevant to musoW, CLEF, MELODY, and the Web portal is available in deliverable D1.9 (Daquino et al. 2022) as well as in scholarly publications respectively addressing CLEF (Daquino et al. 2023), MELODY (Renda, Daquino and Presutti 2023) and the methodology used to develop the Web portal (Renda, Grasso and Daquino 2023; Scharnhorst et al. 2021). The design behind the portal has been discussed also in other deliverables of Polifonia concerning the socio-technical roadmap and the collaborative methodology d1.3 (Daga et al. 2021) and the Data Management Plan (Scharnhorst et al. 2023).

Each of the research outputs here presented address important research questions in the field of information science, addressing aspects like cataloguing and dissemination of data via semantic web technologies, and [sen](#) have a significant impact on the music domain (Daquino et al. 2022). It is worth noting that in the last year no new projects/competitors have addressed similar problems to those outlined at the beginning of the project wrt WP1.T3 objectives. To this extent, we believe that the research in Polifonia is still at the edge and shows innovative potential in a number of research areas. We now briefly list for each of the four research outputs the state of the art.

musoW. Existing **catalogues of music heritage data** have not been updated nor leveraged in new projects. Instead, musoW has been refined and updated by contributors, and it is actively used in several activities relevant to Polifonia tasks, e.g. support literature review in WP1-related activities, licenses extraction (see D2.6), and web portal population.

CLEF. To better compete with existing LOD-native cataloguing systems, CLEF has been improved with new cutting-edge features for **knowledge extraction**, and it is currently leveraged in a few national and international projects (Daquino et al. 2023).

MELODY. Neither new software solutions for **LOD-based data storytelling** have been developed, nor have existing ones been improved to integrate features like those offered by MELODY (i.e. a web-based platform for creating author-controlled narratives of charts, search tools, and curated content), which is a unicum in the landscape of data storytelling platforms (Po et al. 2020). Instead, MELODY is constantly improved to (1) serve as a teaching tool, and (2) be easily integrated into other web applications (Renda, Daquino and Presutti 2023).

Web portal. To the best of our knowledge, no out-of-the-box solutions exist to facilitate the **alignment of Linked Open Data** sources and, at the same time, populate user-friendly interfaces (Renda, Grasso and Daquino 2023). We developed a software solution that performs (1) customizable data reconciliation operations, (2) data indexing and text search, and (3) population of descriptive web pages. Notably, the web portal allows data sources leveraging different ontologies to be integrated, although a preferential route has been designed for sources reusing the Polifonia Ontology Network (PON).

3. Research problems, objectives, and approach

3.1 Research problems and objectives

Among the four objectives outlined in D1.9, i.e. cataloguing, crowdsourcing, discovery, design and exploration, the second half of the project has mainly focused on the completion of the latter one, namely:

***Design and exploration.** On the one hand, digital cultural heritage collections are often dedicated to specialists, and fail in engaging with lay people. On the other hand, solutions to enable experts to analyze Linked Data and share data stories are missing.* (Daquino et al. 2022)

More attention has been devoted to (1) continuous testing and evaluation of CLEF and MELODY, to deliver solid, reusable, software solutions, and (2) the development of the Web portal, to deliver a prototype for cultural heritage data reconciliation and dissemination. A common ground in the design and development of such software solutions is ensuring their **reusability** in new projects, hence fostering its adoption and maintenance in the near future.

3.2 Approach and workplan

The preliminary analysis performed to prototype CLEF, MELODY, and the Web portal, has been presented in D1.9 (see D1.9, Section 3.2.4). It included activities like the following ones:

- **Qualitative and quantitative analysis** of competency questions raised by Polifonia experts and stakeholders
- **Competitive analysis** of existing solutions for LOD-native cataloguing, data storytelling, and web applications for cultural heritage dissemination
- **Focus groups with experts** to collect desiderata and feedback on prototypes (mockups)
- **User studies with lay users** to understand habits, tasks, and expectations towards web environments for music heritage discovery

During the second half of the project, we have **continuously involved our target audiences** in user testing sessions, co-design activities, and focus groups (see Fig. 3.1).

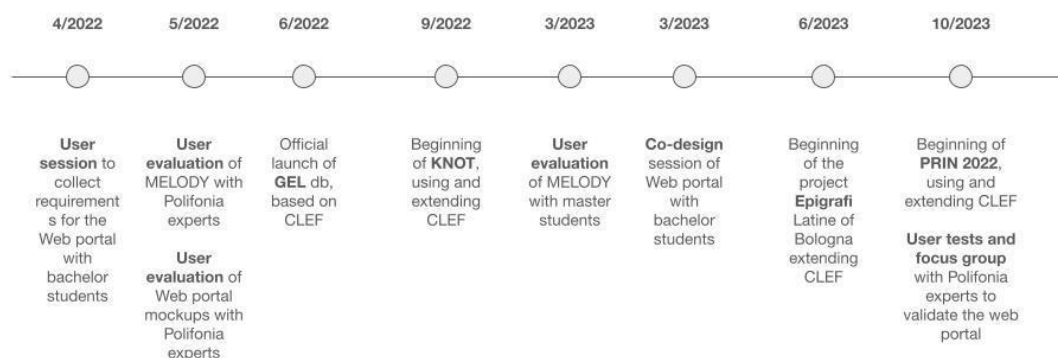


Fig. 3.1 Timeline of main activities relevant to evaluation and exploitation

CLEF. We are currently reusing and extending functionalities of CLEF in **four new projects**, namely: GEL, funded by UNESCO; KNOT, funded by the Italian Ministry of Culture; ATLAS, a project funded by the Italian funding program PRIN; Epigraphs in Bologna, a project of the Department of Classical Philology and Italian Studies of the University of Bologna. The adoption of the software in external projects contributed to its continuous evaluation and improvement.

MELODY. We first tested MELODY with **experts** of the Polifonia consortium in dedicated hands-on sessions, producing data stories that exemplify the potential of pilot datasets, and we have iteratively collected feedback to develop new UI templates. Secondly, we evaluated MELODY with **lay users**, i.e. master students with limited technological skills enrolled in an introductory course on Semantic Web technologies, to test its applicability as a teaching tool.

Web portal. We performed a user study with **lay users**, i.e. bachelor students in Communication Science, to co-design user journeys of the Web portal, and we performed a user testing session with **experts** of the Polifonia consortium to collect feedback on the prototype and to co-design interesting case studies.

4. Software outputs, results and evaluation

In this section we present the outputs of WP1.T3, namely musoW and CLEF, MELODY, and the Web portal. While preliminary analysis and design considerations are detailed in D1.9, here we briefly introduce their peculiarities, aspects of novelty, their current usage in Polifonia and in external projects, and their evaluation.

4.1 musoW and CLEF

Summary. *musoW* is the online catalogue of music heritage data available on the web. It provides structured information about sources, content, access points, and licenses. *musoW* records are constantly updated and reviewed by partners and external contributors. A constant number of users view the website every day, and it is used in several activities relevant to Polifonia. *musoW* is published online via *CLEF*, a crowdsourcing platform for Linked Open Data native cataloguing developed as part of Polifonia objectives. *CLEF* has been continuously extended and refined to comply with requirements of projects for crowdsourcing in the cultural heritage domain and it is currently used in four external projects. External funding has been received to guarantee its maintenance after the end of the project.

In the last year, **musoW** has been actively enhanced by external contributors and partners of the Polifonia consortium. Around 2K users have engaged with the platform and 10-20 users access the website every day (Fig. 4.1.1). While the number of indexed resources has not significantly grown up (~520 resources), the quality of cataloguing data has sensibly been improved. Existing records have been deduplicated, more information has been provided, and the scope of the collection has been refined.

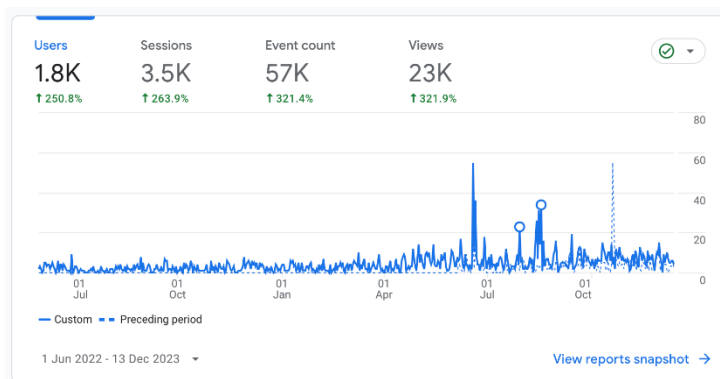


Fig. 4.1.1 User engagement in musoW from June 2022

To date, musoW is used in a few tasks in Polifonia, e.g. as a source of information for literature review in WP1 and WP2, and as a data source for refining Large Language Models in extracting licenses (Daga et al. 2023). To foster its reuse beyond Polifonia, in summer 2023 a dedicated social media campaign has been launched on Twitter/X, which regularly advertises a resource described in musoW via the Polifonia profile.

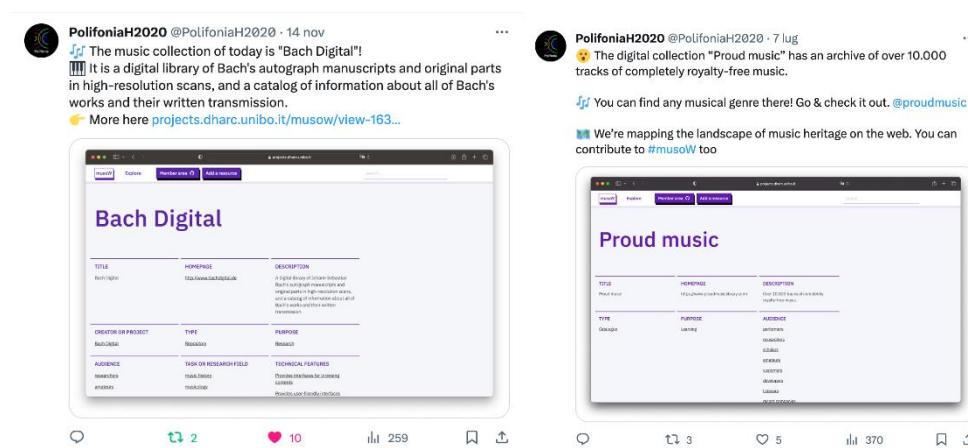


Fig 4.1.2 Examples of posts on Twitter/X advertising musoW resources

CLEF - the software underlying musoW - demonstrated to be an attractive solution for mid-size projects focused on cataloguing, crowdsourcing and dissemination of digital resources in the cultural heritage domain. The software has been presented at the *1st linked archives international workshop (LinkedArchives 2021)* co-located at TPD 2021 and has been selected among the best papers for publication on JOCCH (Daquino et al. 2023), i.e. an international venue appreciated in the communities of Information Science, Digital Libraries, and Digital Humanities.

The software is maintained by members of the project that are part of the Digital Humanities Advanced Research Centre (/DH.arc) of the University of Bologna, which also ensures long-term hosting and maintenance of musoW. The centre ensures maintenance and continuous development of resources thanks to commissioned research activities, and evaluated CLEF as a suitable solution in four new research projects, namely:

- Global Education and Learning (GEL)⁸ is a project funded by UNESCO, the European Union and the ANGEL network to publish and maintain a multilingual open bibliographic dataset of scholarly publications on the theme of global education. The online catalogue is served via a customized version of CLEF, which also includes advanced search mechanisms and exploration via data visualizations.
- KNOT⁹ (in progress) is a project funded by the Italian Ministry of Culture (MIC) to develop a prototype of the National Digital Library. It uses CLEF to store and enrich cataloguing records of scholarly projects and research outputs produced by Italian Universities (which is the scope of the cataloguing process). Project requirements include new features like semi-automatic knowledge extraction and extended data entry functionalities.

⁸ <https://projects.dharc.unibo.it/digestgel/>

⁹ <https://icdp-digital-library.github.io/KNOT/index.html>

- Epigraphs¹⁰ (in progress) is a project of the Department of Classical Philology and Italian Studies of the University of Bologna to geolocate epigraphs in Bologna.
- ATLAS¹¹ (PRIN2022, in progress). Scholars of Polifonia / DH.arc centre have been recently awarded a PRIN2022 grant by the Italian Ministry of Education for 250K euros to continue developing CLEF, extending its knowledge extraction functionalities, and integrate open-source solutions like SPARQL Anything¹² to facilitate the cataloguing process.

4.2 MELODY

Summary. *MELODY¹³ is a web interface designed for authoring data stories based on Linked Open Data. The application is designed for querying any SPARQL endpoint and sharing insights in the form of data visualizations, interactive charts, text searches, and curated content. It has been developed using a novel methodology that combines eXtreme Design and Design Thinking methodologies (Renda, Grasso and Daquino 2023). To foster its reusability, we (1) refactored the code base (React + Flask) to facilitate integration of new UI components (e.g. new charts), we (2) developed APIs to query MELODY platform and generate data stories as HTML snippets to be injected into external websites. Finally, we (3) evaluated the software with lay users, to foster its reuse as a teaching tool.*

Overview. MELODY (Renda and Daquino 2023) is a web platform that allows scholars to share cultural heritage data with users lacking technical expertise using data visualizations and storytelling techniques. In MELODY, a Graphical User Interface allows users to query any online SPARQL endpoint and plot results directly into charts, maps, and design interactive text searches. It was designed with a focus on requirements from the Polifonia ecosystem's datasets. However, it has been developed in a way that is sufficiently general, allowing for reuse with any (cultural heritage) data. For more details, please refer to Deliverable 1.9 (Daquino et al. 2022). MELODY source code¹⁴, documentation¹⁵, and public platform¹⁶ are available online.

¹⁰ <https://projects.dharc.unibo.it/el>

¹¹ <https://dh-atlas.github.io/>

¹² <https://polifonia-project.github.io/ecosystem/polifonia-project/external-components/components/sparql-anything/sparql-anything.html>

¹³ <https://projects.dharc.unibo.it/melody/>

<https://github.com/polifonia-project/dashboard>

<https://polifonia-project.github.io/dashboard/>

¹⁶ <https://projects.dharc.unibo.it/melody/>

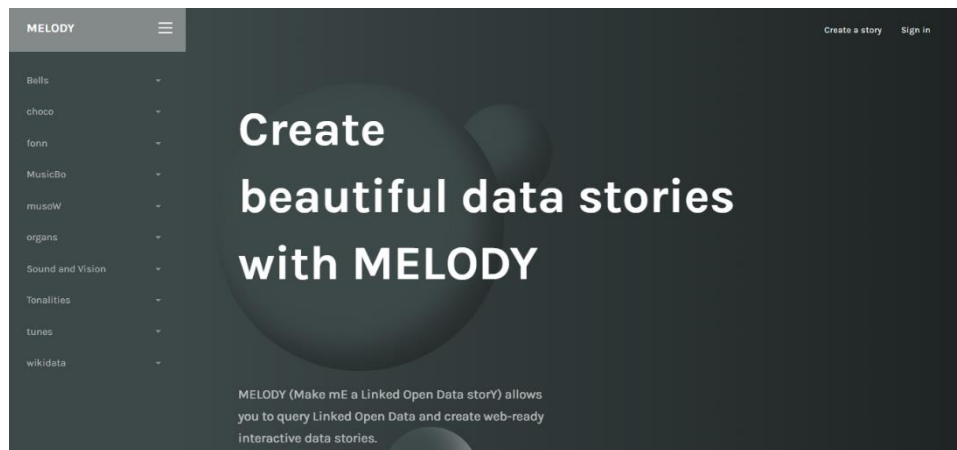


Fig 4.2.1 Melody public platform homepage.

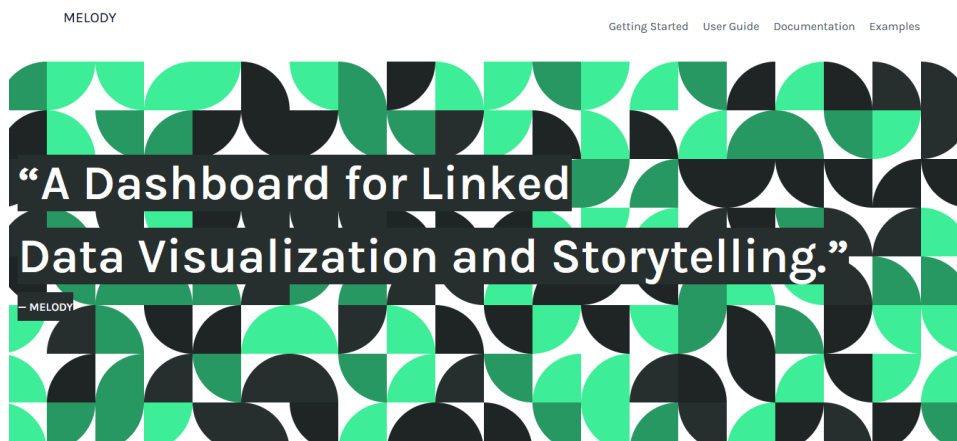


Fig 4.2.2 Melody documentation homepage.

MELODY allows three types of **user authentication**, namely:

1. Collaborator: authenticated users belonging to the Polifonia organization with creation, publishing, and editing rights on all stories created by collaborators
2. External: authenticated users outside the organization, can create, modify, and publish their data stories (publication on a separate catalogue)
3. Unauthenticated: can create stories but are limited to download, without editing or publishing functionalities.

Authentication configuration is adaptable for any organization interested in reusing MELODY. The prerequisites include having a GitHub account (for the organization) and a dedicated repository to store MELODY code. To set up the configuration, the `conf.py` file needs to be completed with the required information, as detailed in a dedicated section of the documentation website¹⁷.

¹⁷ See more detail in the documentation https://polifonia-project.github.io/dashboard/documentation/getting_started.html

To set up a new data story, a user must choose or create a section (shown in the left sidebar of the website), enter the main title, and specify the SPARQL endpoint. Afterwards users are guided through a **What-You-See-Is-What-You-Get** (WYSIWYG) interface, which allows them to create data stories by adding and combining various UI components in a single canvas. Components allow users to query the SPARQL endpoint specified and to preview results as charts or other interactive components (e.g. tables, counters) (Fig 4.2.3). Available components are the result of the preliminary analysis on data journeys identified during the requirements elicitation phase (i.e., dashboard-like interface, interactive text-search, geographic map), which are detailed in the previous deliverable. The **available components** are the following:

- Text: Enables the creation of text content (i.e. HTML) using a WYSIWYG rich text editor.
- Counter: Displays a card with a number and a label as shown in Fig. 4.2.3.
- Chart: include bar charts, pie charts, line charts, scatterplots, and doughnut charts.
- Map: Presents geolocated data on a map with optional filters and a sidebar for metadata.
- Table: Displays data in tabular form and can include text, images, audio files, and videos.
- Text Search: Provides a free-text search on a dataset, returning results in a table.
- Action: Executes a query on any result of a text search, displaying results in another table.

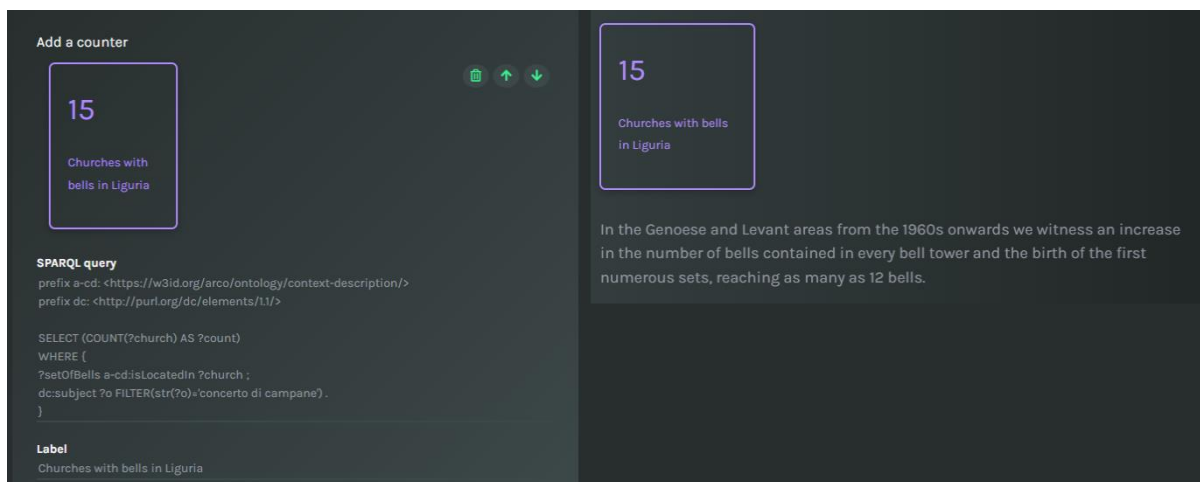


Fig 4.2.3 An example of WYSIWYG to create a Counter (left) and the final preview (right).

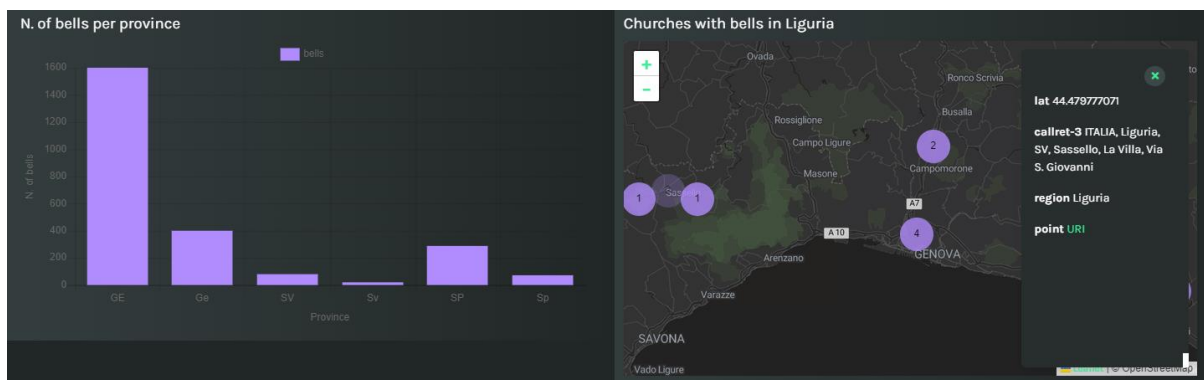


Fig 4.2.4 An example of a bar chart (left) and a map (right).

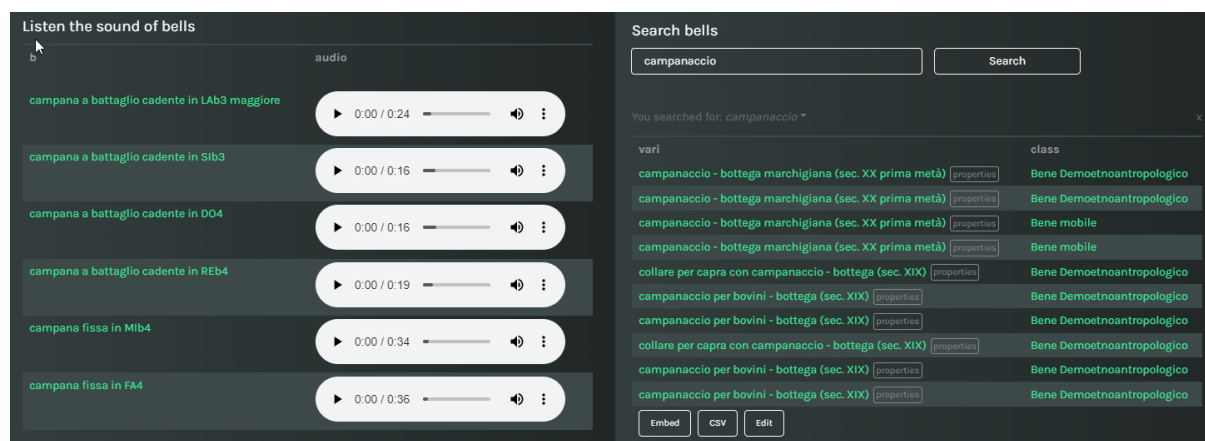


Fig 4.2.5 An example of a table (left) and a text search with actions buttons (right).

Users can further personalize their stories by choosing the color palette, include metadata (title, author, description), and can sort or delete components at any time.

The process of creating a data story can be illustrated as follows using the example of BELLS data story¹⁸. In the setup screen, BELLS dataset curators began by selecting a title and specifying the URL of the SPARQL endpoint from which to extract data. Once confirmed, they were directed to a form for filling in context information, including author names, a subtitle, and an overview of the content. As an initial step, the curators opted to showcase the number of bells present in Liguria. They chose the “Counter” component, which prompted them for a SPARQL query to extract the count, along with a descriptive label (Fig. 4.2.3). Subsequently, they added the “Text” component to provide textual information (Fig. 4.2.3) and introduced another counter to display churches with at least one bell in Genoa. To visually represent the geographical distribution of churches, they decided to incorporate the “Map” component, requiring a title and a corresponding SPARQL query (Fig. 4.2.4 right). The query was written following instructions in the supporting documentation available. The curators continued to fill the story by incorporating components such as “Chart” (Fig. 4.2.4 left), “Table” (Fig. 4.2.5 left), and “Textsearch” (Fig. 4.2.5 right), filling in the necessary fields. Since curators are members of the Polifonia project, their story is published on the same platform.https://projects.dharc.unibo.it/melody/modify/bells/overview_of_bells_in_liguria[https://project s.dharc.unibo.it/melody/modify/bells/overview_of_bells_in_liguria](https://projects.dharc.unibo.it/melody/modify/bells/overview_of_bells_in_liguria)

Indeed, once saved, the web-ready data story is accessible from the left sidebar of the website (if the user is a Collaborator of Polifonia) or on an external catalogue¹⁹ (if External User). The story can be shared through its URL, exported as a PDF file or as an HTML page. Likewise, each component of the story can be individually exported as an embeddable HTML snippet (all), static image (charts), SPARQL query, or csv file (tables). The data used to configure the web data story is also available for download as a JSON file.

¹⁸ https://projects.dharc.unibo.it/melody/bells/overview_of_bells_in_liguria

¹⁹ <https://melody-data.github.io/stories/>

MELODY reuse. In Polifonia, MELODY allows pilots experts to share their work via data stories, i.e. web articles including curated content and charts, which are published on the web platform. To date, **13 data stories** have been created showcasing peculiarities of Polifonia pilot datasets, and these have been advertised via Twitter/X. As aforementioned, MELODY can be used by any user, regardless of them being Polifonia collaborators, and it offers a separate venue²⁰ (Fig. 4.2.6) to publish data stories created by external users. To date the external catalogue counts more than **100 data stories** (see e.g. Fig. 4.2.7), mostly developed by students at the University of Bologna, who are currently using the platform to learn using Semantic Web technologies and music data.

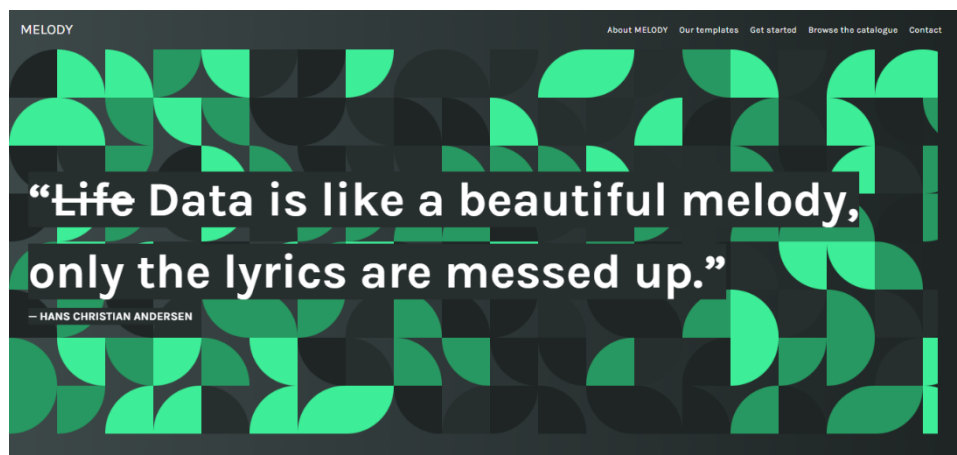


Fig 4.2.6 The catalogue of data stories created by external users.

²⁰ <https://melody-data.github.io/stories/>

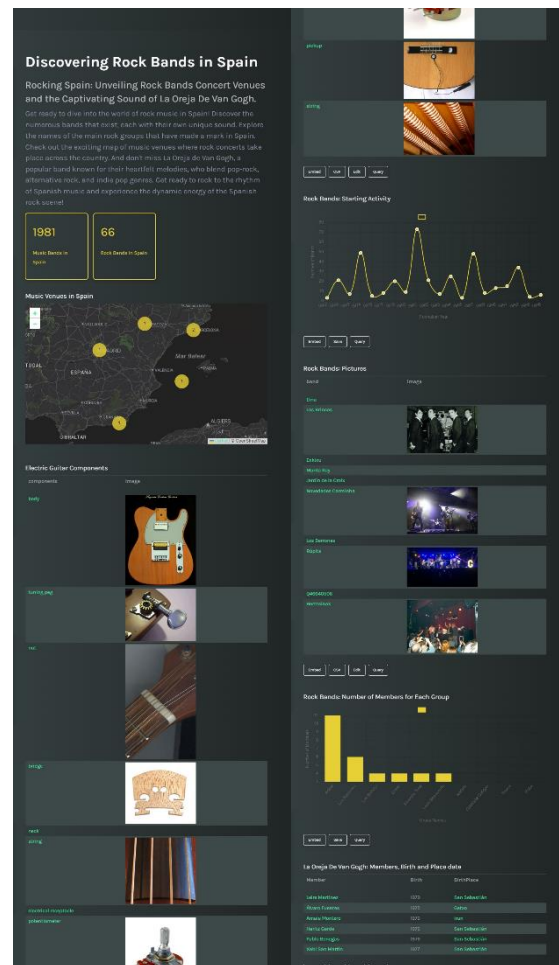


Fig 4.2.7 An example of data story²¹ produced by an external user and published on the external venue.

Technical improvements. MELODY is based on Flask²² backend framework. Recently its code base has been refactored and enhanced including **React**²³ JS library to design reusable UI components. Given React's focus on modular development, this design choice makes it easier to extend the code base, ensures versatility, and fosters new developers in including new UI components that suit their needs.

In the last year we continuously collected feedback from experts of the consortium, and we **refined the behavior of UI components** to support more requirements highlighted by pilot experts. For instance, we included the possibility to edit tables, which allows experts to collect feedback from colleagues and stakeholders and to exchange edited tables (Fig 4.2.8).

²¹ https://melody-data.github.io/stories/published_stories/story_1686331429.988095.html

²² <https://flask.palletsprojects.com/en/3.0.x/>

²³ <https://legacy.reactjs.org/>

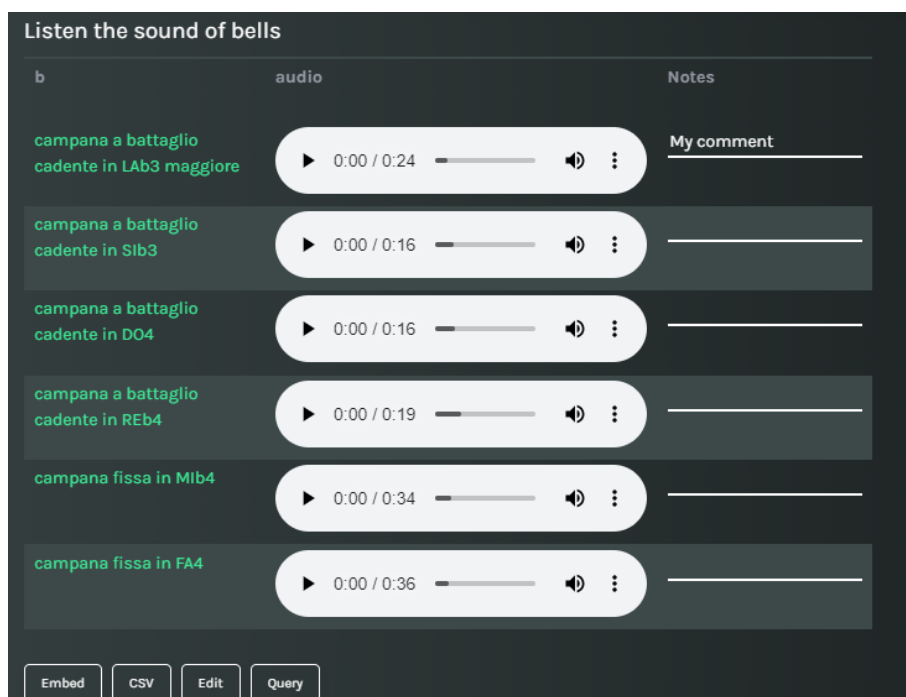


Fig 4.2.8 By clicking “edit” users can leave comments in the right column of a table

Moreover, the original idea of having separate templates to present data journeys, i.e. statistics, geographical exploration, and interactive text search, has been revised and all UI components have been merged into a **single template**. This solution allows editors of data stories to combine all kinds of components into the same canvas.

User testing. In spring 2023 we performed a user study with around 25 master students at the University of Bologna. The form²⁴ and results of the questionnaire²⁵ are available online (Grasso, Daquino, Renda 2023).

Participants are primarily students in their twenties, all coming from the Languages & Linguistics field, who have little or no familiarity with the SPARQL language (i.e. the query language used to interrogate Linked Open Datasets), ranging from novice to intermediate proficiency (Fig. 4.2.9). Similarly, the respondents' familiarity with data visualizations appears limited, with a significant cohort indicating a foundational understanding (Fig. 4.2.10).

²⁴ <https://forms.gle/PmtUt4nRR5sNHosw9>

²⁵ <https://doi.org/10.5281/zenodo.10442429>

From 1 to 5 (1 being not expert at all, 5 being extremely confident), how much experience do you have with SPARQL?

26 risposte

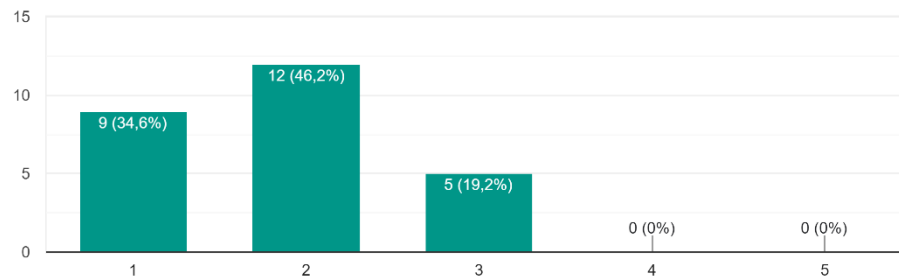


Fig 4.2.9 Users' level of experience with SPARQL is low.

From 1 to 5 (1 being not expert at all, 5 being extremely confident), how well you know data visualizations (reasons for use, usefulness, functioning)

26 risposte

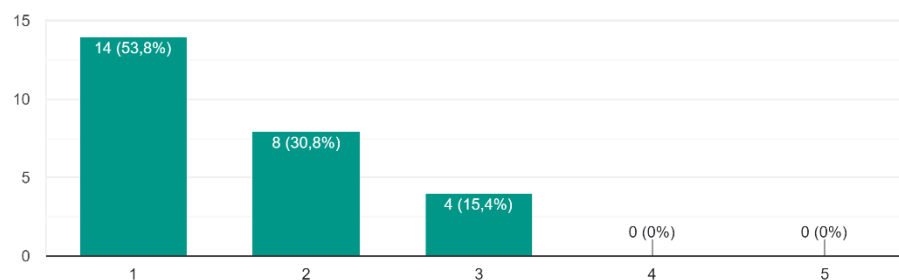


Fig 4.2.10 Users' data visualization literacy is low.

Participants were given 3 tasks to perform on MELODY online and they were requested to answer a few open questions. Whenever applicable, a Likert scale (1=low agreement to 5=high agreement) was proposed.

Task 1. Users were asked to visit the MELODY website and provide some feedback on the clarity of content organization (Fig. 4.2.11). The initial perceptions are diverse. Some users describe it as interesting, useful, and stylish, while others find it confusing and non-intuitive. The navigational aspects received a diverse responses, with most participants acknowledging ease in authenticating and identifying the structure of the platform. However, a subset of participants provided lower ratings, indicating potential areas for refinement in terms of clarity and coherence of the ecosystem.

I understood the differences between the three main pages (catalogue, dashboard, documentation)
26 risposte

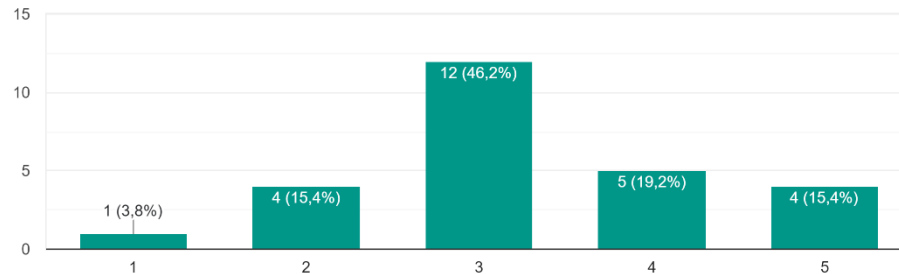
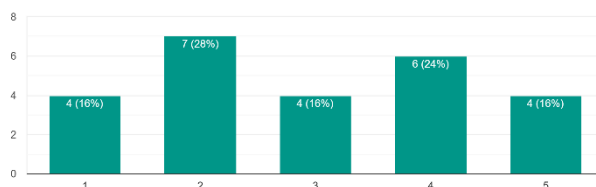


Fig 4.2.11 Users seem to moderately understand MELODY ecosystem on average (1 = Strongly disagree; 5 = Strongly agree).

Task 2. Users received instructions on how to use the MELODY dashboard from a moderator, who also requested participants to use the online user guide to perform their task (Fig. 4.2.12a). They were asked to select the Statistics template and to create a data story that included free text descriptions, charts, and data summaries. The level of use of the user guide in this task varied, showing that more than half participants required guidance beyond the explanation provided by the moderator of the user session. However, most users completed the task (4.2.12b).

I was able to perform the task without the user guide
25 risposte



How successful were you in accomplishing what you were asked to do?
26 responses

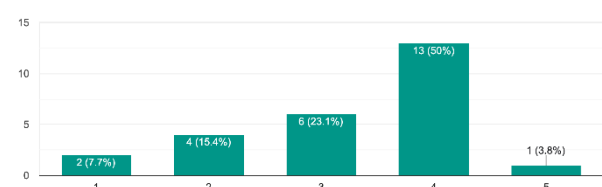


Fig 4.2.12 Users need guidance but can effectively achieve their task (1 = Strongly disagree/Very Low; 5 = Strongly agree/Very High).

Task 3. The last task required users to create iterative text searches, which we consider an advanced feature. Most participants required the user guide (Fig. 4.2.13a) and partially achieved their goal. However, none of the participants were 100% confident of their result. This result shows us other areas for improvement, e.g. in the way sophisticated features are presented and the way we support the user in correcting themselves when making mistakes.

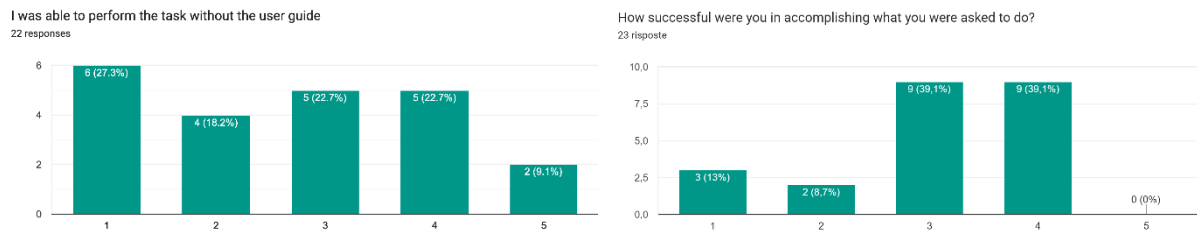


Fig 4.2.13 Users' level of success in performing task 3 (1 = Strongly disagree/Very Low; 5 = Strongly agree/Very High).

Components Evaluation. The questionnaire proposes a Likert scale evaluation of all UI/UX components (e.g. color picker, textbox, counter, line chart, etc.). Familiar components such as “delete” “reset” and “save” were considered the most useful, as well as counters and all components containing text. Despite the perceived complexity, components like actions and charts are deemed extremely useful (Fig. 4.2.14). Text components are deemed easy to use, while charts were considered slightly more complex to fill in. From an aesthetic point of view, the most moderately rated components were the organizational components (delete, reset, save, arrows) and the text search with its actions, highlighting the need for improvement.

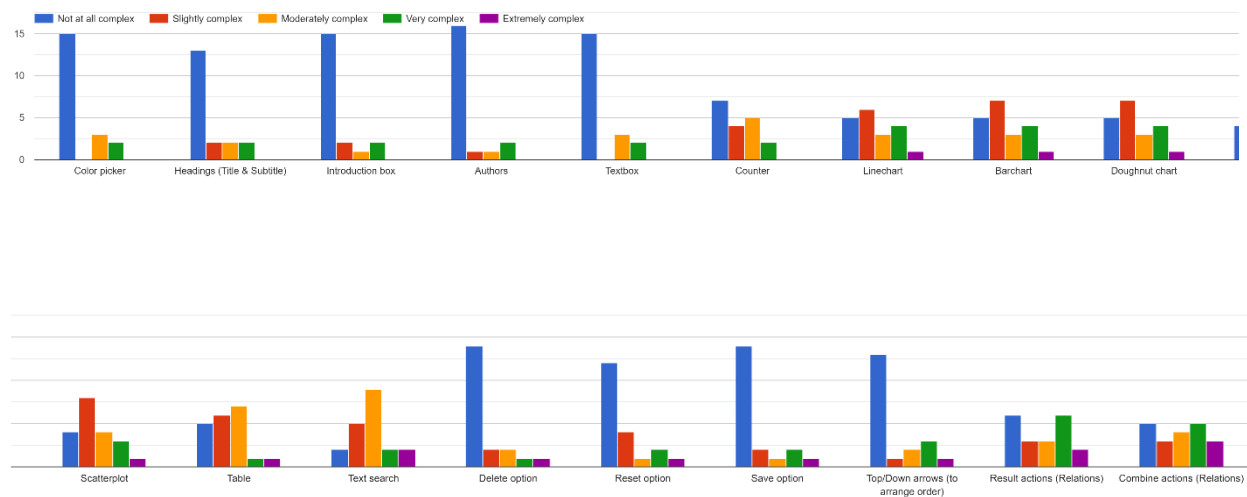


Fig 4.2.14 Components are perceived as useful and not particularly complex

Bug Reporting. Some issues that merit attention were reported. Users have consistently highlighted slow performance on their computers as a recurring concern, as well as confusion during the process of saving and publishing a story. This was to be attributed to limitations within the external catalogue, which is hosted on a GitHub repository that receives stories as commits. We have therefore changed the mechanism stories are published and now the system affords several competing stories being published at the same time. Furthermore, users have encountered difficulties in correctly displaying images, characterized by

images appearing excessively large, and have faced challenges in adding actions to text searches. We fixed this issue by harmonizing the size of thumbnails.

Final considerations show that participants have varying levels of likelihood to use the dashboard frequently. Notably, around 53.8% of respondents disagreed, possibly because they do not expect to work with LOD technologies soon again. 42.3% of the respondents believe the software can be of use in the future (Fig. 4.2.15). Based on the distribution of these scores, it appears that the majority of users grasped the utility and potential of MELODY. However, considering the limited experience many users have declared with data visualization and semantic web technologies, the overall experience and perception of the software might be somewhat 'naive.'. Nonetheless, participants acknowledged that MELODY made it possible for them to work with all such new technologies (SPARQL, Linked Open Data, web documents, charts), while without it, it would have been impossible to produce a similar result. This observation points to the need for developing mechanisms that guide users through the platform upon their initial access, ensuring a smoother and more informed user experience.

I think that I would like to use this dashboard frequently
26 risposte

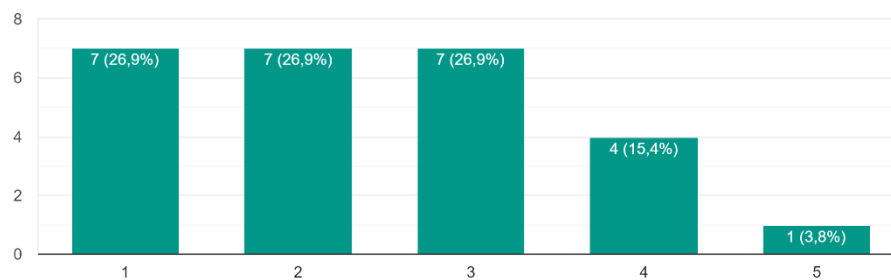
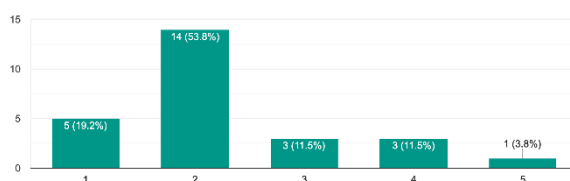


Fig 4.2.15 Half participants can imagine using MELODY again (1 = Strongly disagree; 5 = Strongly agree).

73% of respondents rated the complexity of the tool as low (i.e. rating 1 and 2 in Fig. 4.2.16a). A notable portion of respondents (36%) found the dashboard easy to use, and a sizable percentage (32%) provided a rating of 4, indicating a moderately positive perception (Fig. 4.2.16b).

I found the dashboard unnecessarily complex
26 responses



I thought the dashboard was easy to use
25 responses

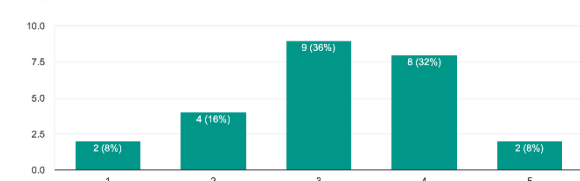


Fig 4.2.16 Overall perceived complexity is low, and ease of use is high (1 = Strongly disagree; 5 = Strongly agree).

Overall, users seem to have varying levels of confidence in using the dashboard independently, (Fig. 4.2.17a) but a total of 54.6% of respondents gave positive ratings (4 and 5) when describing the integration of functions (Fig. 4.2.17b). The absence of respondents who strongly disagreed (rating 1) suggests a generally positive sentiment towards the aspect of the dashboard, although new revisions of the look and feel could help to fill the gap in lack of confidence of beginners. Considering the limited expertise in SPARQL and data visualization, these results indicate that MELODY served, at least in part, as a tool to empower users who might not otherwise have the capability to engage with these technologies. Once again, this underscores the necessity to further reduce barriers, such as those related to SPARQL, that users faced during their interactions.

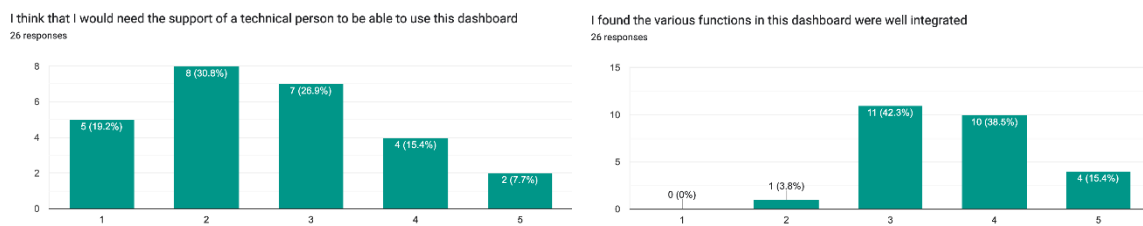


Fig 4.2.17 Independence in usage is still to be achieved (1 = Strongly disagree; 5 = Strongly agree).

4.3 The Web portal

Summary. *The Web portal²⁶ (Renda, Grasso and Daquino 2024) is the main entry point into the extensive musical heritage produced by Polifonia. While MELODY was solely focused on expert users, the Web Portal caters to a broader and more complex target audience, encompassing both expert and generic users. The audience includes enthusiasts, scholars, musicians, and educators, and facilitates exploration of a varied spectrum of musical Linked Data. This contrasts with MELODY, which was geared more towards the creation of data-driven stories by experts. Key functionalities of the portal include text search and on-demand insights. The web portal currently ingests a selection of pilot datasets and allows their exploration via bespoke user journeys designed to foster serendipitous discovery. We continuously redesigned and evaluated the portal UI/UX via user studies, focus groups, and user testing sessions.*

Goals and overview. The web portal application is designed to accommodate informative needs of generic users as well as providing user journeys of incremental difficulty to engage also with experts (see *User experience overview*). Moreover, the development of the web portal aims at accomplishing three technical objectives, namely:

1. facilitate the **alignment** of datasets ingested, hence facilitating the inclusion of datasets external to Polifonia too (see *Data ingestion, alignment, and processing*)
2. **scale** efficiently when new datasets are included, as well as being **customised** when new UI requirements are identified (see *Customisation and reusability*)

²⁶ <https://polifonia.disi.unibo.it/portal>

3. be **agnostic** with respect to the ontologies used by ingested sources, despite facilitating the ingestion of data that conform to the Polifonia Ontology Network (see *Ontology-based retrieval*)

User Experience overview. Fig. 4.3.1 visually maps connections among pages, UI components, and actionable features, providing a depiction of the user journey.

The **Home page** guides users from a broad understanding of the functionalities available on the web portal to more specific ones. The *Intro* component shows a quick take home message, and a *Carousel* shows key pages and resources connected to the web portal. As the user scrolls, the *Highlights* section offers a list of example data available on the platform. Like a music box playing tunes, this section shows featured content and navigation options, creating a welcoming introduction to the diverse access points to the datasets indexed in the web portal. Each highlight is representative of a category (e.g. genres, artists, places) which corresponds later on in the homepage to a text search section. *Search* sections allow users to explore data connected to a topic at hand. Initially, relations of the highlighted topic are shown as an example of possible searches and their results. However, users can change the input and look for other entities of interest. Relations between the topic and other entities are listed and users can filter by type of relation, category of connected entities, and the source of information. Lastly, the *Insights* card provides more information about the connected entity.

Throughout the user journey, the intentional arrangement of components on the Home page ensures a seamless transition from a general overview to more specific areas of interest and searches. The user journey is meticulously designed to not only cater to diverse needs but also foster serendipitous discoveries by ensuring that users can effortlessly access information, guiding them step by step through our platform's offerings and revealing information on demand.

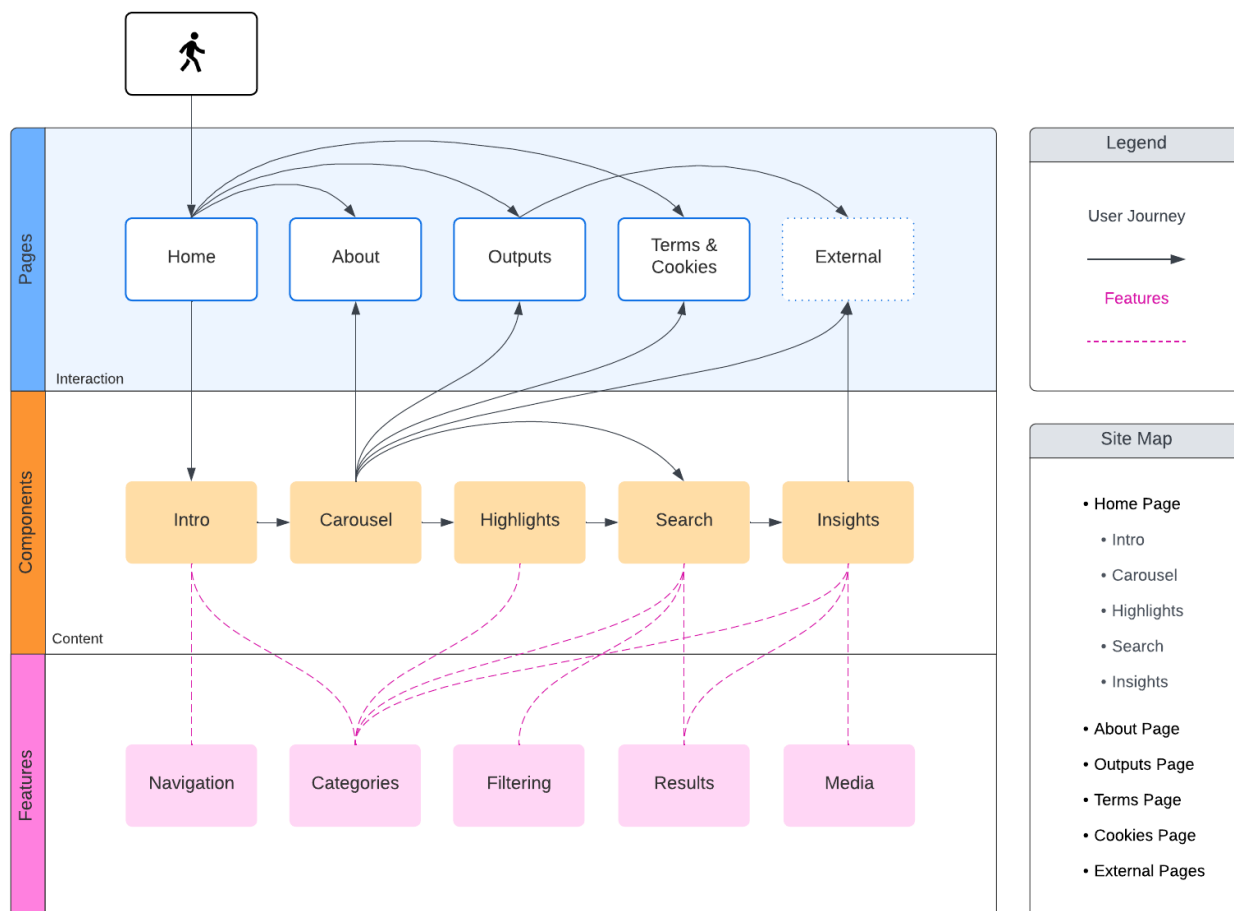


Fig 4.3.1 Pages, components, and features of the Web portal

The **About** page presents the mission, approach, and interplay between the overarching Polifonia project and its dedicated portal. It articulates the mission and goals of Polifonia, while providing insights into the datasets ingested in the portal. The **Output** page groups various outcomes and deliverables that have originated from the Polifonia project and that are integrated in the web portal.

User Interface Components. The main UI components of the Web portal are the following: Intro, Carousel, Highlights, Search, and Insights.

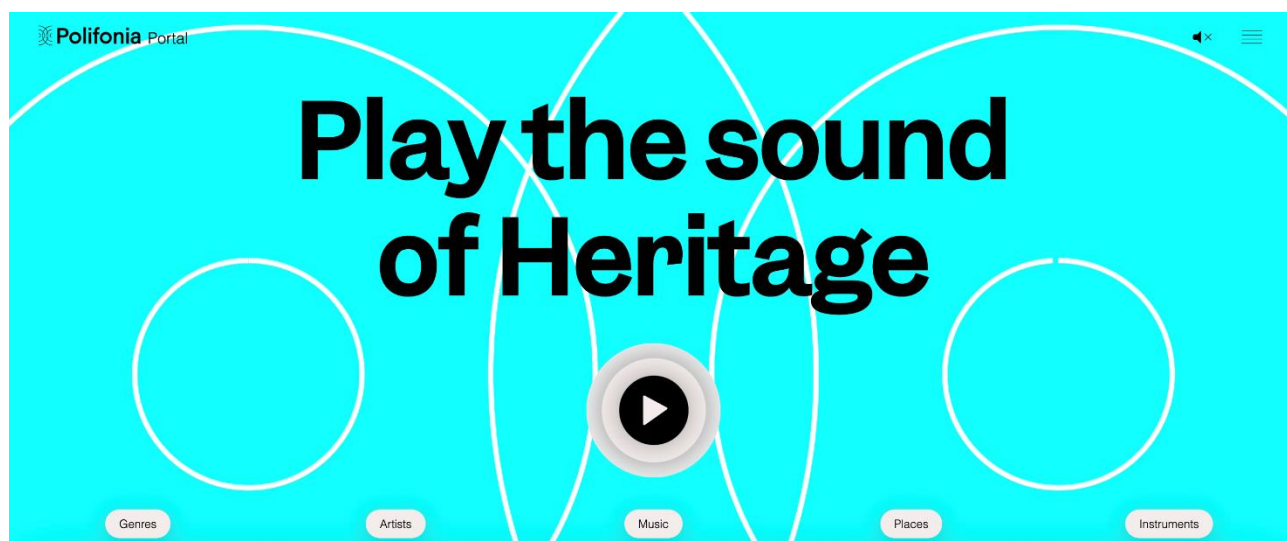


Fig 4.3.2 The homepage of the Polifonia web portal: introductory banner

Intro

The Intro section serves as the gateway to the web portal, offering users an initial orientation and a glimpse into the featured content categories. A background video with intersecting circles echoes the project logo. The Play button prompts users to start the exploration, which can continue with manual/automatic scrolling of the page. In the bottom border, users find the main **categories** that drive the exploration of the Web portal, namely: genres, artists, music, places, and instruments. Categories are lined up in five columns in which the homepage is divided. As they scroll, a parallax effect elevates these categories to form a fixed navigation bar.

The Intro section also contains a **navigation menu**, an audio on/off button, and two back-to-top buttons that facilitate smooth scrolling.

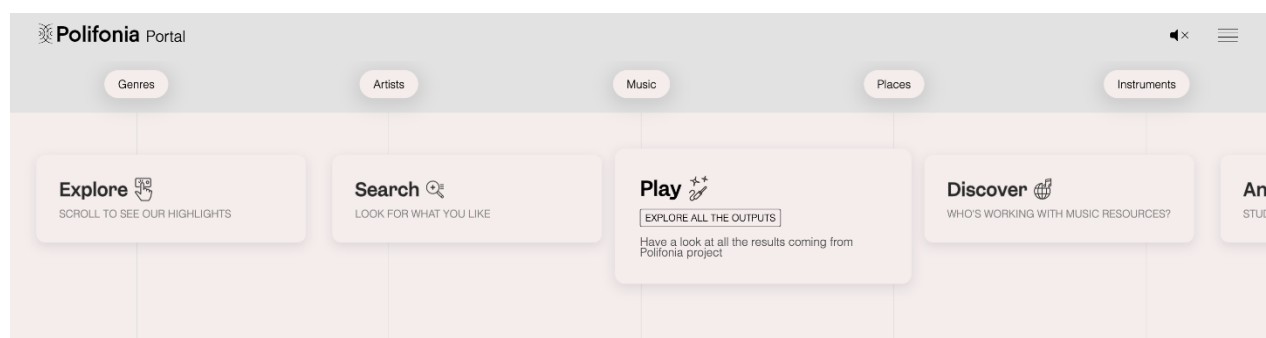


Fig 4.3.3 The homepage of the Polifonia web portal: carousel of resources

Carousel

Like the intro section, the carousel serves as an orientation and summary of the contents of the web portal. It groups **links to internal or external pages** in a sequence of boxes, facilitating the exploration of specific sections of the homepage or redirecting to external Polifonia outputs (e.g. musoW catalogue, the corpus, MELODY). Each box specifies actions users can perform on the website, such as "Explore" (scroll to

Highlights), "Search" (scroll to Search), "Play" (link to Outputs page), "Discover" (link to musoW), "Analyze" (link to Corpus), "Create" (link to MELODY), and "Learn" (link to the page About).

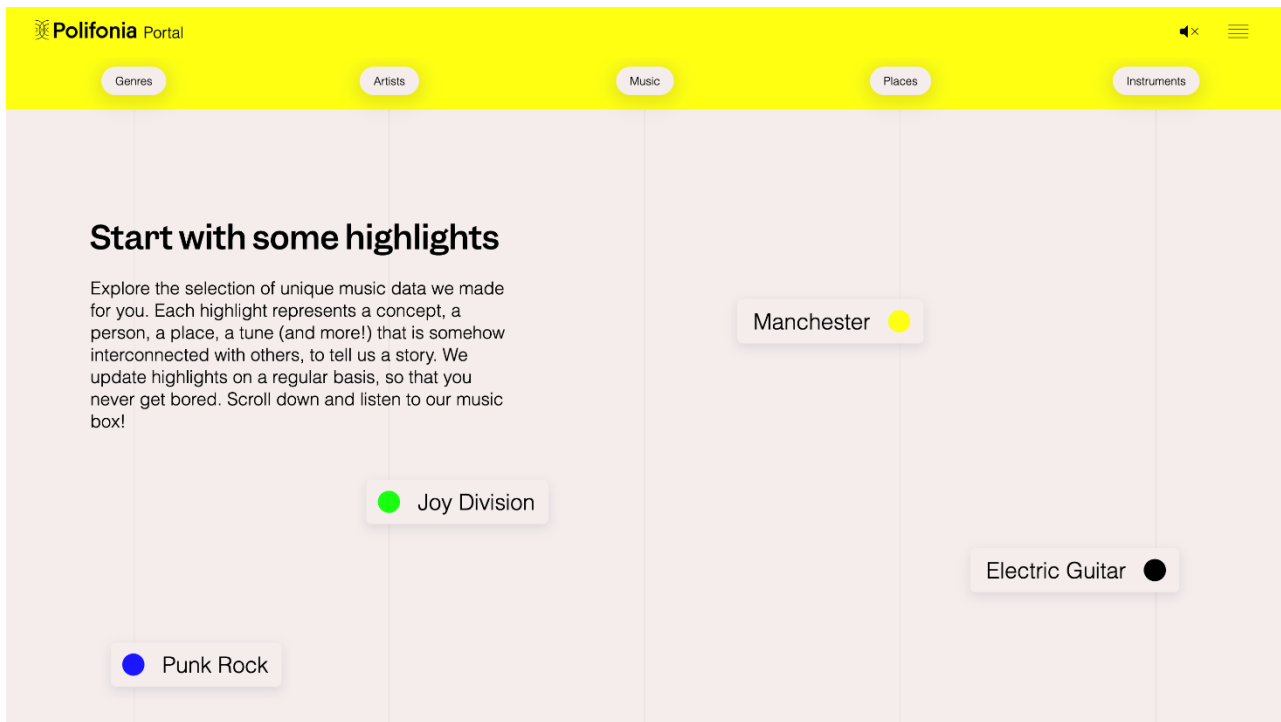


Fig 4.3.4 The homepage of the Polifonia web portal: highlights of featured topics

Highlights

The Highlights section is a **curated collection of featured entities** (i.e., topics) belonging to the different datasets that populate the web portal. The idea is to offer users a preliminary glimpse into the diverse and interconnected data available within the web portal. Highlights not only introduce users to the content but also suggest the types of relationships among entities that can be discovered via the Web portal (e.g. the place where a band was born, its genre, its instruments). Highlights are arranged in 5 columns, resembling a music box, each representing one of the five aforementioned categories, and all together they are related by an associative relation. By clicking on a highlight, users are redirected to the related search section. The highlights are manually curated and configurable: all users see the same entities when accessing the web portal, to guarantee an optimal first entry point to Polifonia datasets designed by experts. Each highlight is accompanied by an audio sound (that can be turned on/off from the menu) and it is associated with a colored dot. The webpage requires users to manually enable audio to respect individual preferences and environments, avoiding the intrusion of automatic playback. The **color** is chosen among those of the Polifonia logo, and it is recursively used in the following sections to identify that category (e.g. yellow – places).

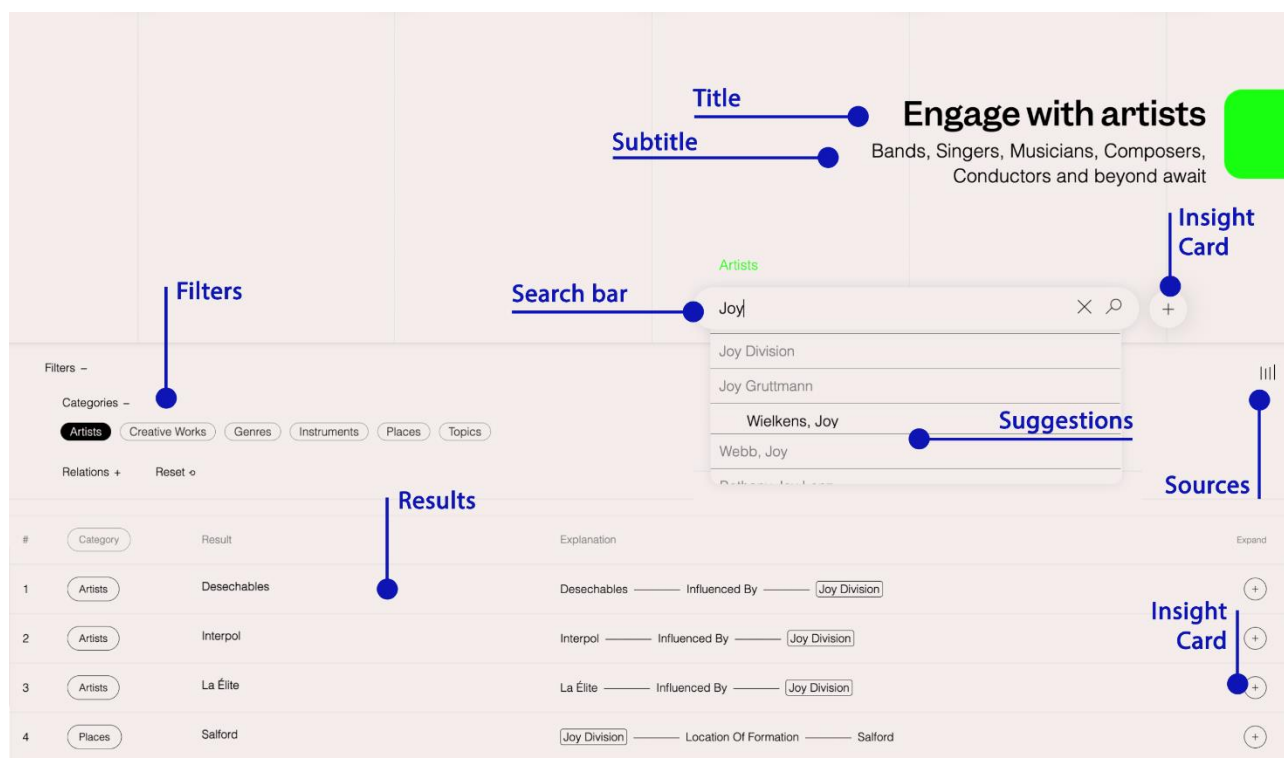


Fig 4.3.5 The homepage of the Polifonia web portal: a Search section

Search

While a generic search interface is missing in the Web portal, **5 Search sections** are currently available, each representing one of the 5 categories that first appeared in Intro, namely: genres, artists, music, places, and instruments. By providing separate search inputs, we ensure that users can delve deeply into specific areas of interest without the noise and clutter of unrelated results that a generic search might yield. In addition, separate search inputs allow for a more tailored and relevant searching experience in each category. This decision is influenced by a desire to avoid the "Google effect," as articulated in Whitelaw's theory of Generous Interfaces (Whitelaw 2015), presenting users with rich, exploratory, and visually engaging interfaces, rather than the minimal, task-based approaches to search collections as typified by Google. A search section is introduced by a title and a subtitle, and includes a textbox already filled with the highlight previously shown for the category at hand. The search bar is pre-filled with the highlight from the corresponding category, and results are displayed based on this initial highlight. Additionally, each search section employs the color of the respective category to ensure visual coherence throughout the features.

A search via the textbox performs a **lookup in an index of entities** falling under the category at hand, which returns autocomplete suggestions, as shown in Fig. 4.3.5. So doing users immediately know what is available on the web portal. Secondly, users are required to select a specific value from the list of suggestions. Such design choices, as well as the lack of a generic text search, derive from the discussion with experts in the consortium, which required (1) specialized interfaces rather than generic ones, and (2) clear, relevant, although limited, responses to their queries.

Each index includes entities of the same type belonging to different datasets and integrates **alignments across datasets**, therefore preventing duplicates as much as possible in the list of results. For instance, if a user types “Tchaikovsky ” to search for the famous Russian composer and this entity appears in more than one data source, only one suggestion will anyway appear in the search results.

Beside the search textbox and the list of suggestions, the search section includes a list of results, a box with filters on results, and a filter on data sources (i.e. the bar chart on the right side).

Results represent relations between the entity searched and other entities included in the ingested datasets. Results are displayed in a tabular format, where each (numbered) line represents a relation to a distinct entity. Columns provide information about the linked item, namely: its name, its category (which can be one of the main 5 ones or a new one, e.g. events, creative works), and its relation to the user's input. Results are displayed in a **semi-random order** to favor serendipitous discovery. To be precise, for each search category and each data source indexed under the category at hand, a bespoke rule for ranking and sorting results is defined by data providers (see **Customization and reusability**). So doing, relevance of results is defined by Polifonia partners according to peculiarities of their datasets, without enforcing one mechanism only.

Filters help to **refine the list of search results**. Filters encompass categories of linked entities (e.g. “return only <Place[s]> related to the artist at hand”) and relations (e.g. “return anything that <has subject> the artist at hand”). Additionally, a bar chart visually represents the **provenance of search results**, i.e. showing the distribution of relations across different data sources. While it is not possible to prune results based on the data source, by selecting a column from the bar chart it is possible to highlight results belonging to that source. This graphical element offers users a clear and concise overview of the diversity of content sources, supporting the assessment of search results relevance and credibility. This approach is designed on preferences collected in general user surveys, and allows for a broader, more inclusive range of content to be shown. To enhance user experience and manage the display of large volumes of data, the results are added incrementally. This is facilitated by a 'load more' feature, activated when users scroll through the list of results. This approach ensures a smoother, more responsive interface, preventing overwhelming the user with too much information at once. It also aids in faster page loading times, as only a portion of the results are loaded initially, with more being added as the user scrolls down.

Finally, an expand button is placed alongside each result, allowing users to open a full-page card (later called *Insights* card) that provides details on the specific result.

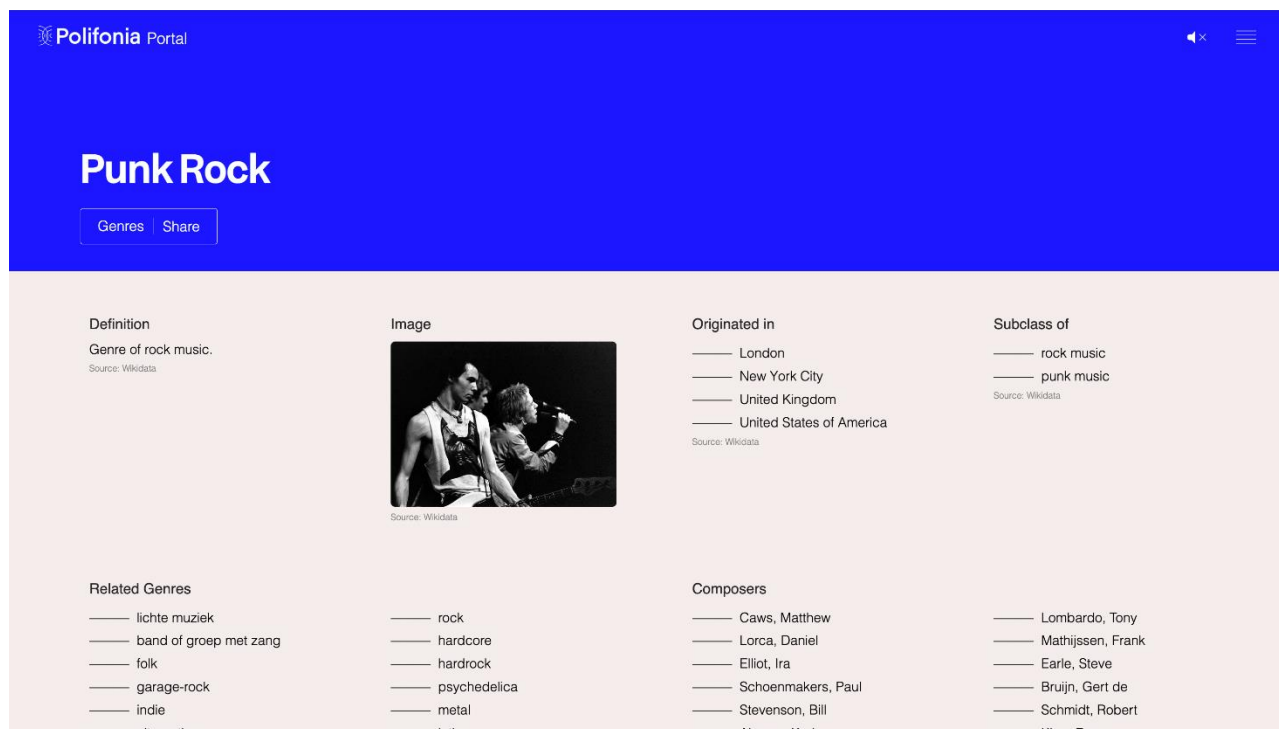


Fig 4.3.6 The homepage of the Polifonia web portal: an Insights full-page card

Insights Card

The Insights card allows users to **expand search results** and delve into details of linked entities. A card mainly includes texts, links to external resources, multimedia, and associations between the entity at hand (i.e. the entity mentioned in a row of search results) and other entities belonging to any of the ingested datasets.

An Insights card is generated for each entity indexed in the Polifonia web portal and includes some **recurring elements** defined by data providers and pilot experts. The heading of the card usually includes a title, category, and sharing options (i.e. prefilled posts to be shared on social media, and a link to be copied and pasted). The background color of the heading recalls that of the category the entity belongs to. At the moment, seven categories of cards have been set, namely: genres, artists, music, places, and instruments, creative works, and listening experiences.

The body of the card presents **customized lists of relations/links grouped into blocks**. Blocks can include long texts (e.g. “Biography”), lists of related entities (e.g. “artists that collaborated with”) that appear as links to other Insights cards, multimedia (images, videos, and audio files) that can be organized in carousels when multiple documents appear to be linked to the same entity, and links to external resources (e.g. “website of the artist”). Blocks can appear in three different formats – small, medium, large – based on the preferences of the card designer. The content of blocks can be customized according to the category of the entity at hand and according to the source to be queried to fill that block. For example, if the entity pertains to a person, we can query Wikidata to have a small-width portrait and a text description, and we can retrieve the list of digital archives preserving the artist’s materials from musoW. When describing a music genre, the card may accommodate lists of related songs, albums, or artists retrieved from CHILD, INTERLINK, TUNES, etc. It’s worth noting that long lists of related entities are paginated: only a selection

is shown and can be expanded on demand. This is facilitated by a 'load more' feature. Again, the order and the relevance of related entities is defined by data providers, who configure the queries underlying each block.

Ideally, a card offers a venue to **discover meaningful connections** and wherefrom to continue the user journey towards either cards of entities indexed in Polifonia Knowledge Graph or to external resources. Relations address some of the objectives originally outlined in user stories²⁷, and are indeed manually curated by data providers, who decide which information to show for a certain category. For instance, musoW developers define which relations to show between artists and digital archives including relevant materials, experts involved in the pilot CHILD define which relations to show when browsing listening experiences connected to places. Currently, we have designed cards representing the 5 categories, listening experiences, and creative works.

Technology stack. Fig. 4.3.1 illustrates the elements of the technology stack that compose the Web portal, highlighting their interactions through CRUD (Create, Read, Update, Delete) operations. The technology stack of the web portal includes a Flask backend and a React-native frontend application. Seamless communication between the frontend and backend is ensured by RESTful APIs. Sonic²⁸ search backend is used for fast **indexing** of data ingested in the web portal and ensures a responsive user experience. Blazegraph triplestore is used to store post-processed data. In particular, data ingested from external sources is analyzed to extract alignments between entities described in different data sources (e.g. people, places, genres) and the links across datasets are stored in a dedicated **linkset**, which is systematically updated whenever changes in the sources happen.

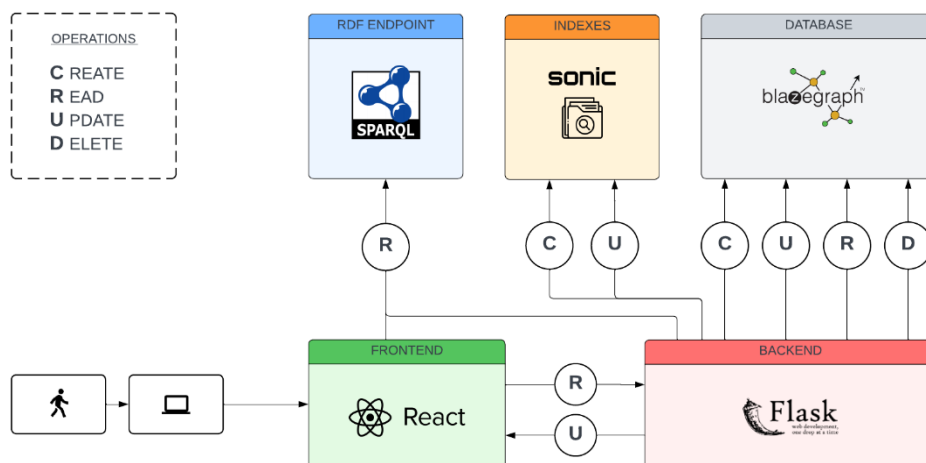


Fig. 4.3.7 Overview of the technology stack of the Web portal

Data ingestion, alignment, and processing. The prototype currently includes a selection of Polifonia pilot datasets that have been released at the time of this deliverable submission (e.g. musoW, NISV datasets, MEETUPS). Others will be included when a stable version is released.

²⁷ <https://github.com/polifonia-project/stories>

²⁸ <https://github.com/valeriansaliou/sonic>

Data is extracted from selected sources as prescribed in the categories configuration file (see *Customisation and reusability* for further detail on this file), which contains the SPARQL queries to be used on each dataset (Fig. 4.3.8, operation 1). In particular, for each dataset and for each category relevant to the dataset at hand (e.g. category: Creative works is available in dataset: musoW, while it is not available in dataset: MEETUPS), a query is performed to retrieve a list of URIs and labels (Fig. 4.3.8, operation 2) that will populate the indexes of entities used to populate suggestions in Search sections. URI/label lists are temporarily stored as files on the file system and become the input of the reconciliation process (Fig. 4.3.8, operation 3). Indeed, to present cross-datasets information in coherent ways and to prevent duplicate information to hamper the user experience, the Web portal acts as a post-processing hub that performs the reconciliation of data sources.

Entity reconciliation is a process for harmonizing and integrating data about entities that come from different sources. For instance, Pilot 1 stores information about an entity *Tchaikovsky 1*, and Pilot 2 stores other information about entity *Tchaikovsky 2* that is potentially the same as *Tchaikovsky 1*. In the web portal we allow users to search for *Tchaikovsky*, retrieve one occurrence only along with data from both sources when presenting insights about *Tchaikovsky*.

Entity reconciliation is a long-standing problem in the Semantic Web community for which several solutions have been proposed. In Polifonia we use a **rule-based approach** to extract explicit assertions about alignments (i.e. owl:sameAs statements) and we use transitive properties to external third-party authorities such as DBpedia and Wikidata to infer links across Polifonia datasets even when a direct alignment is not available in the data sources.

In detail, the reconciliation process unfolds in five stages:

1. **Named graph creation.** For each URI retrieved in the data extraction phase, we create a named graph (Fig. 4.3.8, operation 4).
2. **Internal search.** We query Polifonia datasets and look for asserted owl:sameAs | skos:exactMatch | schema:sameAs | ^owl:sameAs | ^skos:exactMatch | ^schema:sameAs statements having as subject or object the URIs previously extracted (Fig. 4.3.8, operation 5). We include retrieved links in named graphs (Fig. 4.3.8, operation 6).
3. **External search.** We query third-party authority datasets (i.e. Wikidata and DBpedia) looking for links to any of the aforementioned entities (Fig. 4.3.8, operation 7). We include found links in the aforementioned named graphs (Fig. 4.3.8, operation 8).
4. **Transitive links inclusion.** We expand relations using a transitive property function (A sameAs B; B sameAs C > A sameAs C) and perform a lookup to retrieve entities appearing in multiple graphs. If matches are found, we merge sameAs statements into a new graph, delete the old ones, and we mint a new Polifonia URI. The result is a pruned linkset that contains as many named graphs as the number of unique Polifonia entities (Fig. 4.3.8, operation 9).
5. **Linkset and indexes population.** Explicit and inferred connections are stored in a **linkset** served via a triplestore deployed along with the Web portal. Newly minted URIs and labels representing entities shared across datasets, and entities/labels that have not been reconciled among datasets, are ingested in Sonic search backend (Fig. 4.3.8, operation 10) and populate category-based indexes (Fig. 4.3.8, operation 11).

Indexes are queried by the web portal React application to disambiguate entities when a text search is performed, and the linkset is interrogated when showing integrated information in the cards.

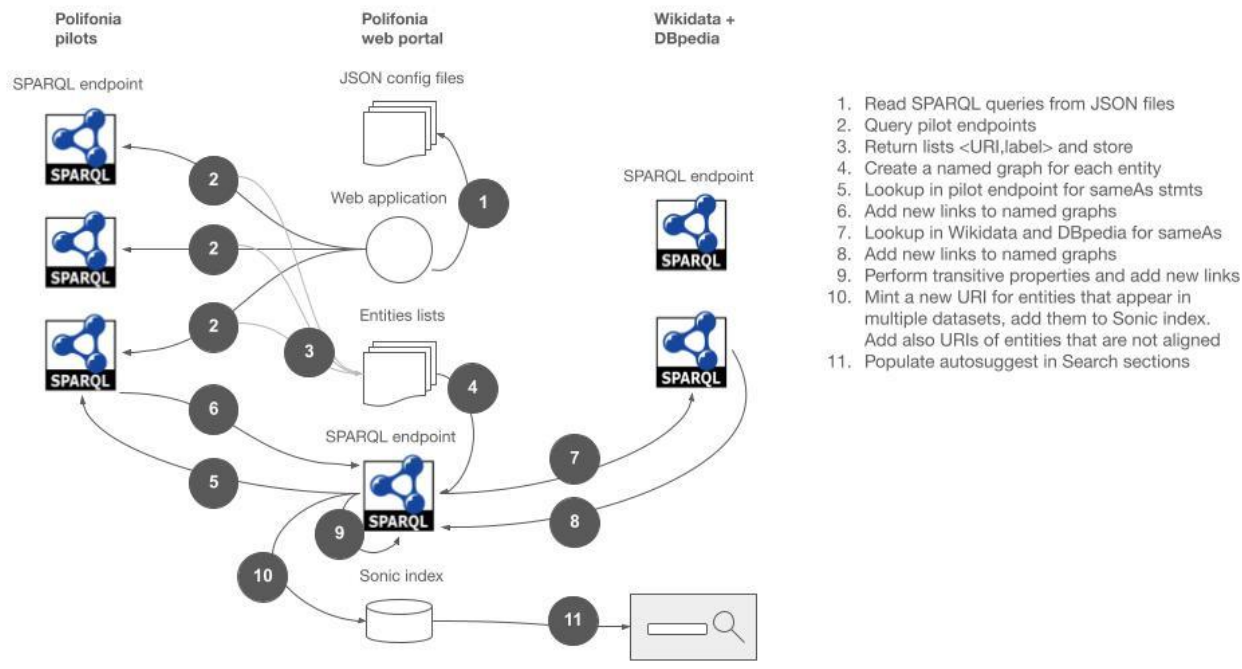


Fig. 4.3.8 Overview of the data extraction and reconciliation process

Customization and reusability. Each component of the Web portal is crafted to be reused regardless of a specific number or type of categories and can be easily extended to new datasets populating the final application. The design philosophy hinges on adaptability, ensuring that the website can evolve to address requirements derived from newly ingested datasets, new ontologies, and new UI/UX requirements. In fact, five configuration files respectively allow us to tailor access points to data sources and UI components. In detail:

- **Datasets:** a configuration file dedicated to datasets²⁹ allows us to incorporate new datasets, each described with a set of metadata. Metadata includes unique identifier, dataset name, description, IRI base, query method, a SPARQL endpoint or REST API.
- **Categories:** the configuration file of categories³⁰ affects several aspects of the portal, from the definition of the number of highlights that appear in the homepage to the number of search sections, their related indexes, and the templates for insights cards. Each category is here described with an identifier, a name, and a color, as well as a title, subtitle, and optional highlighted words that appear in the search section. Most importantly, it includes search patterns tailored on each dataset defined in the previous configuration file. In detail, for each dataset that is relevant to a certain category, the data provider must include (1) a query to retrieve the list of entities that will populate search indexes (e.g. the list of URIs and labels of entities falling under the category

²⁹ https://github.com/polifonia-project/portal/blob/main/backend/conf_datasets.json

³⁰ https://github.com/polifonia-project/portal/blob/main/backend/conf_categories.json

“Artists” respectively in musoW, MEETUPS, CHILD, etc.) and (2) query to set up filters on results (e.g. the list of relations to be shown in search results between an artist and places, other artists, works, etc.).

- **Highlights:** the configuration file³¹ of highlights allows designers to curate the selection of featured entities to be shown at the beginning of the user journey. Each entity is described with a unique identifier, a name, a URI, and the ID of a category. Moreover, a highlight can be accompanied by a section title, a description, and some emphasized terms.
- **Carousel:** the Carousel configuration file³² includes the list of boxes shown under the introductory section of the website (e.g., "Search," "Explore," or "Discover"). For each box designers can specify a brief title, a description, an SVG or PNG logo icon, the action to be performed on click, like scrolling within the page, linking internally or externally, and the destination of the action (a URL or element id).
- **Insights cards:** In the Insights configuration file³³ data providers have extensive control over the dynamic process of cards population. The richness of Insights cards lies in the ability to populate them with diverse content blocks. Users can incorporate four types of content blocks: text, multimedia, relations, and web links. For each block, users can specify sizes (small, medium, or large), titles, and descriptions. A SPARQL Query can be specified for each data source to be queried to retrieve connections between the entity described in the card and other entities available in the Polifonia Knowledge Graph. Media blocks introduce a class key facilitating media type specification, with options for video, image, or audio. Similarly, relation blocks allow users to specify the category of linked entities, defining a cohesive thematic association. Lastly, link blocks can be associated with labels, providing an additional layer of context to the linked content.

Such a design philosophy ensures that the Web portal can be effortlessly reused by other projects with a focus on cultural heritage data dissemination, hence facilitating its adoption and maintenance in the future.

Ontology-based retrieval. Another peculiarity of the Web portal is that data retrieval is entirely based on the ontology specifications of the ingested data sources. The application is designed to be **agnostic with respect to the ontologies** used in source datasets, that is, any dataset can be ingested regardless to the ontology used, as long as the ontology can be effectively used to retrieve (1) lists of entities and labels that loosely fall under one of the categories of the website, and (2) relations between entities.

Nonetheless, when sources consistently adopt the same ontology classes and predicates, the configuration of the application becomes easier, and results are shown using a coherent vocabulary that does not require human intervention in polishing and harmonizing terminology. To this extent, the **usage of the Polifonia Ontology Network** (PON) is of great advantage in the configuration of the web portal and its interfaces. Even though the portal currently includes only a selection of pilot datasets, the ingestion of other datasets

³¹ https://github.com/polifonia-project/portal/blob/main/backend/conf_feed.json

³² https://github.com/polifonia-project/portal/blob/main/backend/conf_carousel.json

³³ https://github.com/polifonia-project/portal/blob/main/backend/conf_cards.json

complying with PON in a second time is likely to be seamless and can be performed by pilot leaders without much effort.

For instance, the following query allows us to retrieve artists across all Polifonia datasets.

```
PREFIX core: <http://w3id.org/polifonia/ontology/core/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
SELECT DISTINCT ?entity ?entityLabel
WHERE {
  ?class rdfs:seeAlso core:Person .
  ?entity a ?class ; rdfs:label ?entityLabel .
}
```

Listing 4.3.1 Example query to retrieve artists across Polifonia datasets

Web Portal User testing and evaluation. The continuous evaluation of the Web portal encompasses tests and feedback collection on (1) co-design aspects (gathered during user studies with lay users), and (2) UI/UX-related aspects like effectiveness of the interfaces and user satisfaction (collected from experts during dedicated user testing sessions). On the one hand, lay participants were asked to imagine themselves in an everyday scenario (e.g. “look for a music-related venue to visit”) and envision a web portal that allows them to perform the task, hence describe features of such platform, desiderata, and possible user journeys. On the other hand, expert users were asked to interact with the prototype, perform some tasks, and describe their experiences, providing feedback and overall impression. Such inquiries aim at identifying potential areas for improvement while enforcing user-centric design principles.

In total we performed four sessions divided into two phases. The initial phase of user tests was performed prior to development and consisted of one session with lay users (March 2022) and one with experts (April 2022). The insights gained from these tests played a pivotal role in shaping the design of the functionalities of the web portal, as detailed in the previous deliverable D1.9. The second series of user tests was carried out again with general users (March 2023) to better frame co-design aspects, and with experts (October 2023) to test the final prototype.

Co-design User Testing (March 2023)

We involved 186 graduating students with varying levels of expertise in the field of music. Participants actively participated in scenario-based question-and-answer sessions tailored to model the portal User eXperience. They were presented with four scenarios representing real-world situations, and they were asked to choose the one they felt closer, namely:

- Scenario 1: You are at home and studying. You would like to explore new music but would like to find something newer than the usual playlists. You start browsing the web... (selected by 74.7% participants)
- Scenario 2: You're travelling, and you want to learn more about the music scene of your location. You decide to look for some information about it... (19.4% participants)
- Scenario 3: You like playing music. You have a melody/chord progression in mind and would like to... (4.3% participants)

- Scenario 4: You are interested in analyzing a linguistic theme/phenomenon based on a large corpus of texts (1.6% participants).

Participants were not informed we were developing a software solution dedicated to such tasks. Rather, they were asked to imagine the web portal without any suggestion or image that could bias their expectations. In fact, participants' answers gave us a clear overview of their habits, preferences, and expectations when searching for music-related information. The questionnaire also included prompts for participants to rate specific features that were under test in the prototype, providing us with both qualitative and quantitative insights to reframe some UI/UX choices. The questionnaire³⁴ and results³⁵ are available online. Notice that scenarios 3 and 4 were not selected by enough participants, therefore they have not been evaluated.

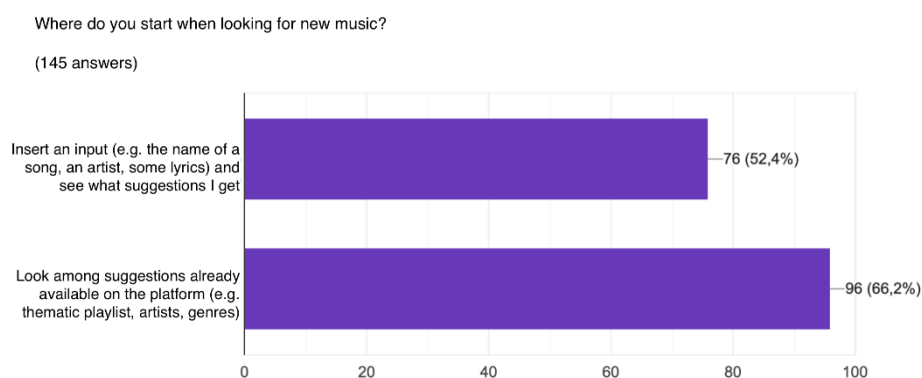


Fig 4.3.9 Users need platforms to suggest them new music (66,2%)

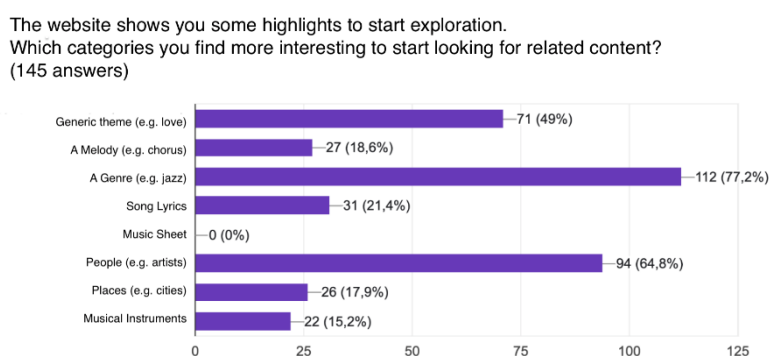


Fig 4.3.10 Users appreciate suggestions based on genres, artists, and topics.

Starting from a category some related content is presented to you (e.g. lyrics, artists and places related to a theme). Would you like to be able to filter contents?

(145 answers)

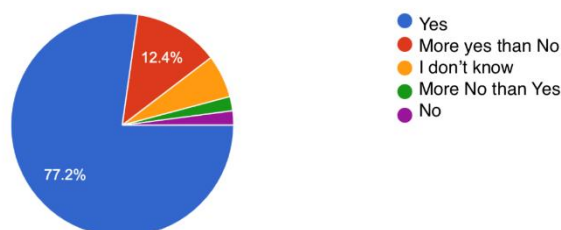


Fig 4.3.11 Users appreciate the possibility to filter their results.

If yes, how would you like to filter results?

(129 answers)

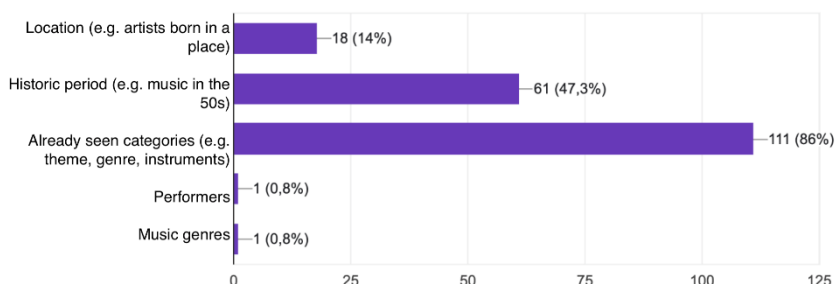


Fig 4.3.12 Users would like to have filters based on known categories.

Results of the user tests indicate that participants would appreciate a web portal where music data is integrated (86,9% of participants claimed to be interested), that they would enjoy serendipitous discoveries (57 participants shared positive personal experiences with previous serendipitous discovery journeys,) as well as guided journeys through diverse information (61,4% claimed to be interested in searches starting from specific inputs). 66.2% of participants expect platforms to provide suggestions as a starting point for their journey (Fig. 4.3.9) and to retrieve personalized user journeys. Journeys should start from or being somehow based on music genres, artists, and general topics (music genre is also the most recurring word in open-ended answers with 154 occurrences) (Fig. 4.3.10). Such a result guided us in defining the five categories that became the main entry points to Polifonia knowledge graph, i.e. the categories, the highlights, and the search sections. Interestingly, the novel contribution of the survey was to shed light on a rather simple feature that is not available in existing platforms for music data discovery, that is, filter options. 77.2% of participants indeed appreciate the possibility of filtering search results (Fig. 4.3.11) using known categories to prune results of their searches (Fig.4.3.12).

Expert Usability Review (October 2023)

We gathered 18 experts from the consortium to perform user tests on the prototype. Participants were chosen for their proficiency in user experience design and their extensive knowledge of the music domain. We asked them to perform time-constrained tasks, guided by moderators. During exercises, participants were prompted to explore the web portal, look for some information, perform a search or retrieve some link to resources. In detail, tasks were the following:

- Task 1: Browse the homepage and find the search section dedicated to music genres
- Task 2: Retrieve the country that originated K-Pop music and identify three notable artists.
- Task 3: Identify three artists influenced by the American singer Madonna and provide the links to their Insights page.
- Task 4: Freely explore the web portal.

After the task was completed, participants were asked to fill in a questionnaire where to share (1) the result of the task, (2) the process they followed, and (3) their satisfaction level. The questionnaire³⁶ and the results³⁷ are available online.

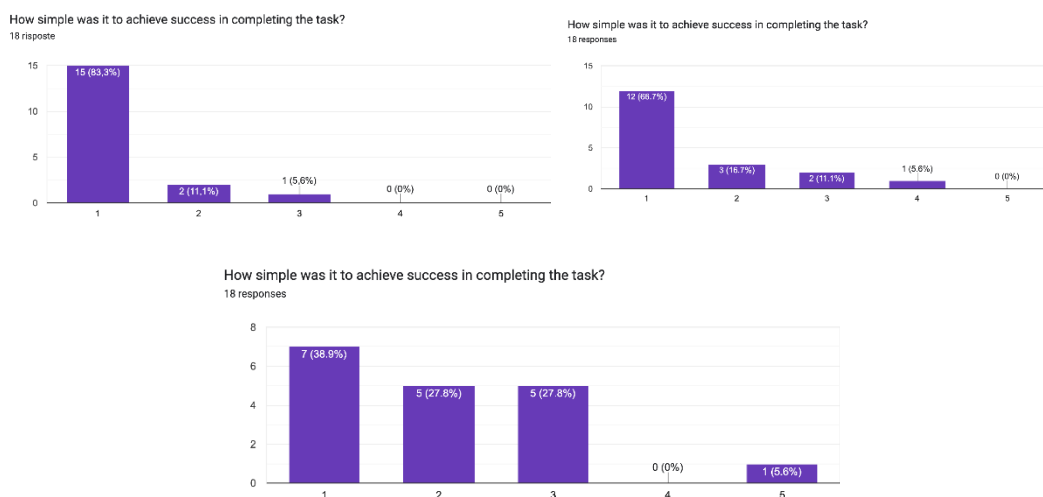


Fig 4.3.13 Despite tasks (1-3) increasing complexity, the success rate is always high (1 = Very easy; 5 = Very difficult)

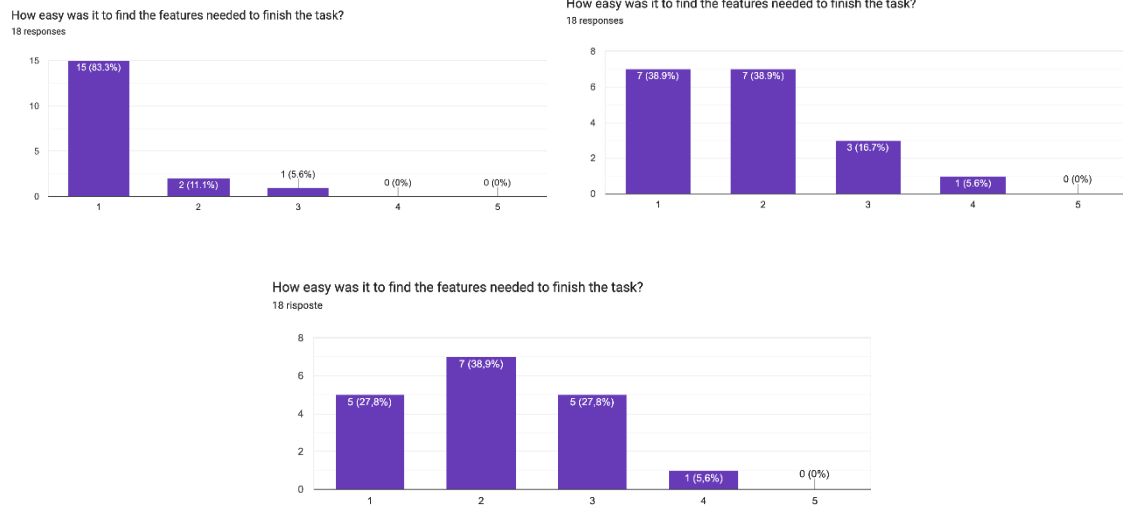


Fig 4.3.14 Features to accomplish tasks are easy to find regardless of their complexity (1 = Very easy; 5 = Very difficult)

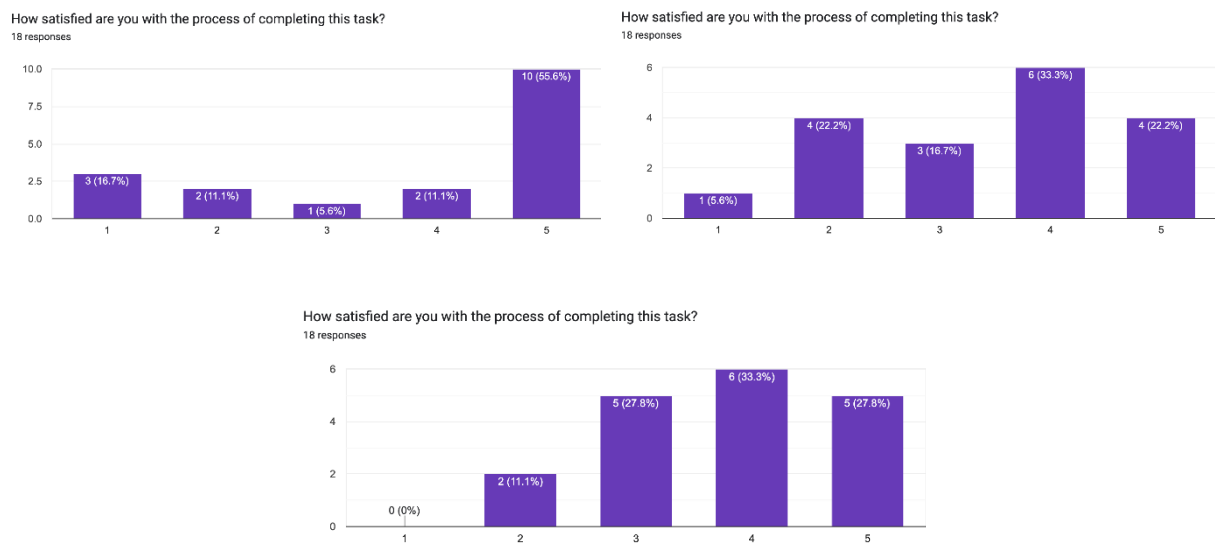


Fig 4.3.15 Satisfaction is generally high (1 = Very low; 5 = Very high)

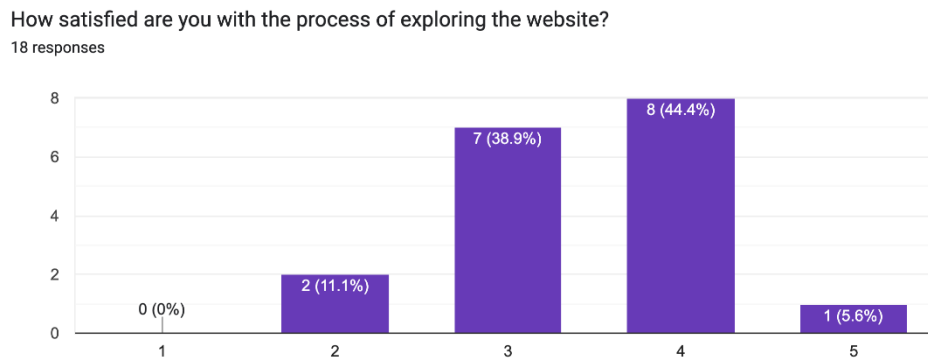


Fig 4.3.16 Exploring freely the portal is generally appreciated (1 = Very low; 5 = Very high), despite some improvements were requested

Results of the user tests indicate overall satisfaction (Fig. 4.3.15, Fig. 4.3.16) and simplicity (Fig 4.3.613) in completing the tasks proposed, while showing mixed responses in determining how easy it was to find the features needed to accomplish the task (4.3.14).

Participants were not only expected to accomplish all the tasks but were also encouraged to proactively report any bug or issues encountered during the exercise. Participants reported 38 messages, which included comments, issues, and suggestions. This proactive feedback mechanism served as an invaluable additional layer, facilitating the identification of potential technical challenges or system errors. The **implications in UI/UX design** were several and led us to modify original prototypes presented in D1.9. To address user preferences, a focus was placed on simplifying the design. Keywords were strategically highlighted across the entire portal for a more intuitive experience. Filters in search sections have been divided into subcategories accompanied by a reset option. Clickable and non-clickable elements are now easy to distinguish. Pagination options ("Load more") were implemented to show content on demand, preventing cognitive overload. Cross-browser problems were also resolved to ensure a consistent experience. Expert users' feedback also guided us in refining lexicon used in titles, subtitles, and paragraphs, and in refining the names of the five main categories (i.e. genres, artists, music, places, and instruments) as the main entry points to Polifonia data.

To supplement quantitative insights, we also performed a focus group to capture qualitative perspectives while freely navigating the web portal and exploring the data available (Task 4, Fig. 4.3.16). Participants engaged in group discussions, offering collective opinions and uncovering shared desires, such as the need for more intuitive navigation or cross-browser compatibility. This qualitative approach unveiled aspects of user perception that may not have been evident in task-based performances, like how users value unexpected discoveries or the ease of finding links and sharing content across other media. It helped us to frame UI/UX aspects peculiar to generous interfaces, i.e. those designed to foster serendipitous discovery without assuming the user has a specific goal in mind. For instance, this feedback directly influenced our decision to add clear textual references near search sections, guiding users on what to type or what kind of queries to perform. Another example is the inclusion of 'load more' buttons on insight cards to give

users the flexibility to explore additional content at their own pace, without overwhelming them with information.

4.4 Case studies: Explore music heritage through the Web portal

Together with experts of the Polifonia consortium, we explored a series of case studies in which our web portal could be used to accomplish practical, real-world tasks, as well as serendipitous journeys. Case studies are derived from suggestions given during user testing sessions and are revisions of [Polifonia user stories](#), collected and refined since the beginning of the project. Scenarios highlight the potential impact that the portal can have on users, offering insights beyond traditional usability evaluations and encapsulating real user experiences.

Use Case 1: Serendipitous discovery

This use case describes a music enthusiast using the web portal for the first time. She wants to discover which content related to her favorite punk band, Rancid, is available. The case study is inspired by Polifonia user stories (Laurent, Keith) and a real user interaction happened at the “Researchers’ night” in Bologna (September 2023), i.e. an event where researchers present their work to passers-by in interactive sessions.

Objective: Engage with materials related to an artist, discover something unexpected.

Datasets: Wikidata, Musow

Stories: Laurent³⁸, Keith³⁹

³⁸ https://github.com/polifonia-project/stories/blob/main/Laurent_Music_Journalist/Laurent%231_MusicArchives.md

³⁹ https://github.com/polifonia-project/stories/blob/main/Keith_Music_Producer/Keith%231_MusicConnections.md

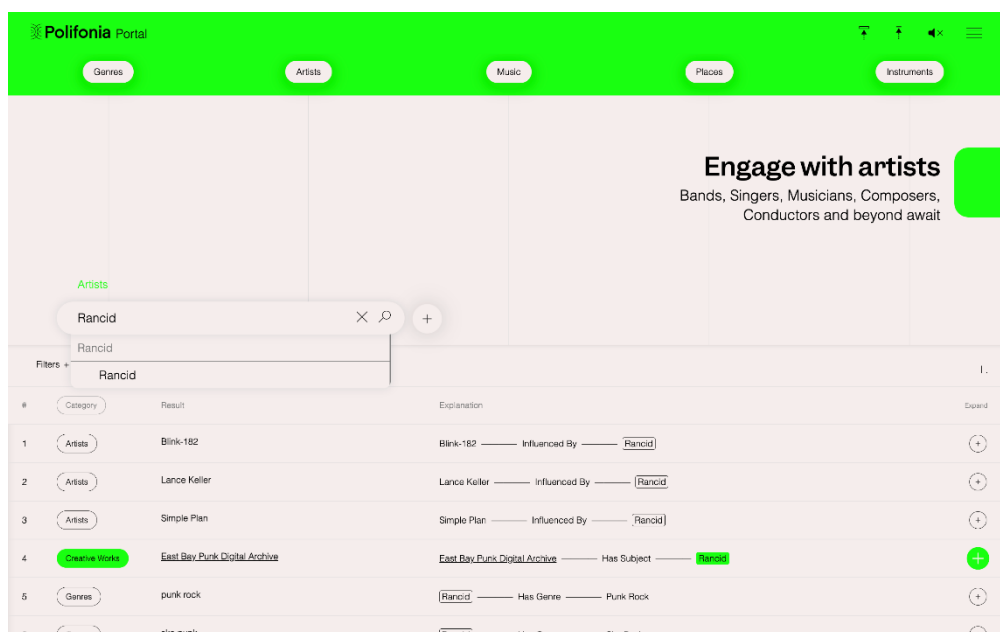


Fig. 4.4.1 Search section

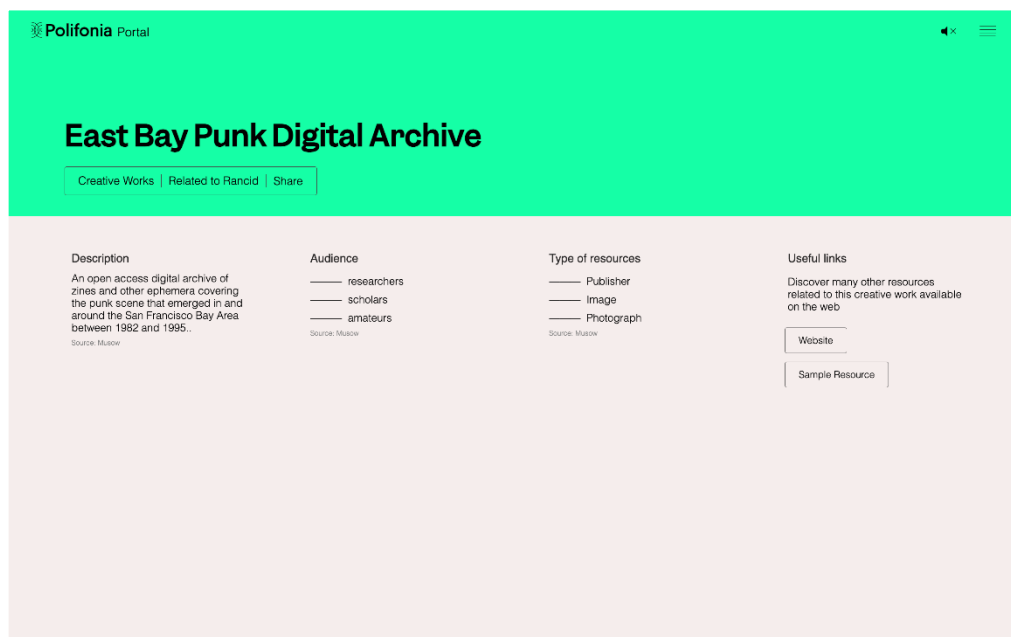


Fig. 4.4.2 Insight page

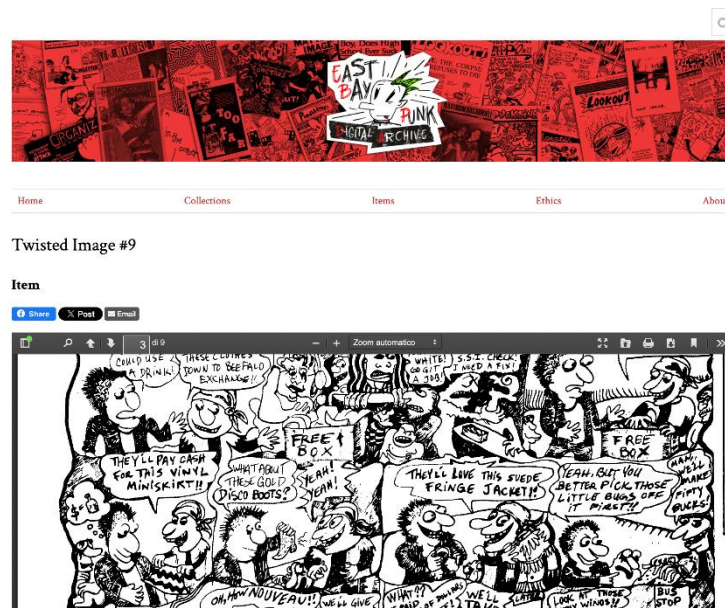


Fig. 4.4.3 External page reached at the end of use journey

Example: The user lands on the home page⁴⁰ and scrolls to the section dedicated to artists. S/he enters "Rancid" in the search bar. The search returns a list of results connected to the band, including related music genres and digital archives that include materials relevant to the band (Fig. 4.41). To fine-tune the search results, the user applies some filters, narrowing down results to creative works related to Rancid. Among the results, the user is interested in "East Bay Punk Digital Archive,"⁴¹. S/he clicks on the expand button, which shows a page dedicated to the online archive (Fig. 4.4.2), including metadata about the archive and a link that leads to the website of the archive itself and another to an exemplar resource available on the website of the archive⁴². The user navigates the website and discovers the digitization of comics relevant to Rancid.

Use Case 2: Interconnected archives

This use case describes a user who wants to retrieve all the archives, datasets, and online resources relevant to a certain artist. The case study is inspired by Polifonia user stories (William) and a real user interaction happened during the Expert Usability Review (October 2023).

Objective: Find external resources and archives related to an artist.

Datasets: MOZ (NISV), Wikidata, musoW, MEETUPS (OU)

⁴⁰ <https://polifonia.disi.unibo.it/portal/>

⁴¹

<https://polifonia.disi.unibo.it/portal/card?title=East%20Bay%20Punk%20Digital%20Archive&cat=creative%20works&input=Rancid&hasinput=false&uri=https://w3id.org/musow/1666725248-9482715>

⁴² <https://eastbaypunkda.com/s/east-bay-punk-digital-archive/item/6112>

Stories: William⁴³

Biography

Pyotr Ilyich Tchaikovsky (Тѣа ѣѡѡѡ/ chy-KOF-skee; 7 May 1840 – 6 November 1893) was a Russian composer of the Romantic period. He was the first Russian composer whose music would make a lasting impression internationally. He wrote some of the most popular concert and theatrical music in the current classical repertoire, including the ballets Swan Lake and The Nutcracker, the 1812 Overture, his First Piano Concerto, Violin Concerto, the Romeo and Juliet Overture-Fantasy, several symphonies, and the opera Eugene Onegin. Although musically precocious, Tchaikovsky was educated for a career as a civil servant as there was little opportunity for a musical career in Russia at the time and no system of public music education. When an opportunity for such an education arose, he entered the nascent Saint Petersburg Conservatory, from which he graduated in 1865. The formal Western-oriented teaching that he received there set him apart from composers of the contemporary nationalist movement embodied by the Russian composers of The Five with whom his professional relationship was mixed. Tchaikovsky's training set him on a path to reconcile what he had learned with the native musical practices to which he had been exposed from childhood. From that reconciliation, he forged a personal but unmistakably Russian style. The principles that governed melody, harmony and other fundamentals of Russian music ran completely counter to those that governed Western European music, which seemed to defeat the potential for using Russian music in large-scale Western composition or for forming a composite style, and it caused personal antipathies that dented Tchaikovsky's self-confidence. Russian culture exhibited a split personality, with its native and adopted elements having drifted apart increasingly since the time of Peter the Great. That resulted in uncertainty among the intelligentsia about the country's national identity, an ambiguity mirrored in Tchaikovsky's career. Despite his many popular successes, Tchaikovsky's life was punctuated by personal crises and depression. Contributory factors included his early separation from his mother for boarding school followed by his mother's early death, the death of his close friend and colleague Nikolai Rubinstein, his failed marriage with Antonina Miliukova, and the collapse of his 13-year association with the wealthy patroness Ildizhdza von Meck. His homosexuality, which he kept private, has traditionally also been considered a major factor though some musicologists now downplay its importance. Tchaikovsky's sudden death at the age of 53 is generally ascribed to cholera, but there is an ongoing debate as to whether cholera was indeed the cause, and also whether the death was accidental or intentional. While his music has remained popular among audiences, critical opinions were initially mixed. Some Russians did not feel it was sufficiently representative of native musical values and expressed suspicion that Europeans accepted the music for its Western elements. In an apparent reinforcement of the latter claim, some Europeans lauded Tchaikovsky for offering music more substantive than base exoticism and said he transcended stereotypes of Russian classical music. Others dismissed Tchaikovsky's music as "lacking in elevated thought" and derided its formal workings as deficient because they did not stringently follow Western principles.

Source: Metapops

20

Image



Source: Wikimedia

Related Digital Collections

- Generative Theory of Tonal Music

Source: Metapops

Related places

- Moscow
- Saint Petersburg

Source: Wikimedia

Meeting with

- Yehudi Menuhin
- Cole Porter
- Edward Elgar
- Landon Ronald
- Jawaharlal Nehru
- Rodgers and Hart
- Pierre Monteux
- Max Bruch
- Émile Sauret

Source: Metapops

Useful links

Discover many other resources related to this person, all available on the web

Discogs Page

Concerts

- Programma Mengelberg, Willem
- Programma Richter, S.
- Programma Mengelberg, Willem
- Programma Radio Filh Orkest - Kieft, Ronald - Hoerr, Peter - Demarquette, Henri - Gastinel, Anne
- Programma Radio Filh Orkest - Wigglesworth, Mark - Boreyko, Andrey - Mihara, Akihito
- Programma Leningrad Philharmonisch Orkest - Jansons, Mariss - Grubert, Ilya
- Programma Brabant Orkest - Bruins, Theo - Vries, Han de - Pieterse, George - Teney, Joep - Slagter, Jacob
- Programma Radio Filh Orkest - Kieft, Ronald - Versen, Quirine - Horsch, Gregor - Mukoyama
- Programma Leningrad Philharmonisch Orkest - Jansons, Mariss - Grubert, Ilya

- Programma Leningrad Philharmonisch Orkest - Jansons, Mariss - Alexeev, Dmitri
- Programma Radio Filh Orkest - Wigglesworth, Mark - Boreyko, Andrey - Mihara, Akihito
- Programma Radio Kamerorkest - Groot Omroepkoor - Montgomery, Kenneth - Schellenger, Dagmar - Claessen, Huub
- Programma Nationaal Jeugdorkest - Benzi, Roberto - Belkin, Boris
- Programma Browning, John - Radio Filh Orkest - Vonk, Hans
- Programma Gutman, Natalia - Lobanow, Wasly
- Programma Bashmet, Yuri - Radio Filh Orkest - Gergiev, Valery
- Het ochtendconcert
- Programma Radio Sy Orkest - Stulen, Jan - Muruzabel, Arturo

- Programma Marinierskapel - Buitenhuis, Gert
- Programma Radio Kamerorkest - Groot Omroepkoor - Montgomery, Kenneth - Schellenger, Dagmar - Claessen, Huub
- Programma Moskou Kwartet - Musica Antiqua Koeln
- Programma Leningrad Philharmonisch Orkest - Jansons, Mariss - Alexeev, Dmitri
- Programma Radio Sy Orkest - Fournet, Jean - Belkin, Boris
- Programma Nieuw Sinfonietta Amsterdam - Markiz, Lev - Verhey, Emmy - Meinders, Fredric
- Programma Orde der Vrijheid, Toos
- Programma Moskou Kwartet
- Programma Reizend muziekgezelschap

- Programma Radio Sy Orkest - Barsjai, Rudolf - Waal, Rian de
- Programma Rotterdams Ph Orkest - Vonk, Hans - Kavakos, Leonidas
- Programma Radio Filh Orkest - Sinasky, Wassily - Rodrigues, Eliane
- Programma Residentie Orkest - Gnn, Leonid - Hugh, Tim
- Programma Concertgebouw Orkest - Gergiev, Valery - Keulen, Isabelle van - Carr, Colin
- Programma Brabant Orkest - Joo, Arpad - Keulen, Isabelle van
- Programma Bruggem, Ronald - Radio Sy Orkest - Sinasky, V
- Programma Letland Nationaal Ph Orkest - Magi, Paul - Louisada, Jean Marc
- Programma Gelders Orkest - Dufallo, R - Paul, Pamela mia

Load more

Fig. 4.4.4 Insight page content of the artist

⁴³ https://github.com/polifonia-project/stories/blob/main/William_Curator_Europeana/William%231EuropeanFolkMusic.md

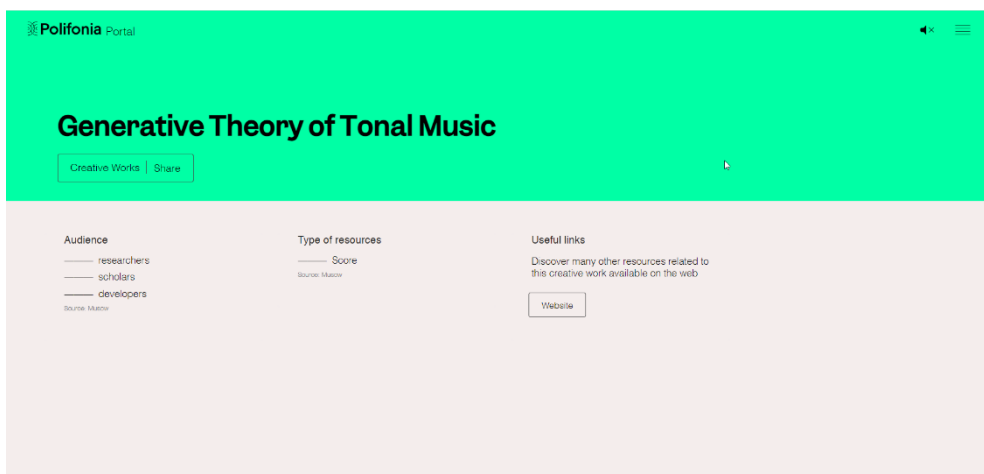


Fig. 4.4.5 Insight page of the related digital archive

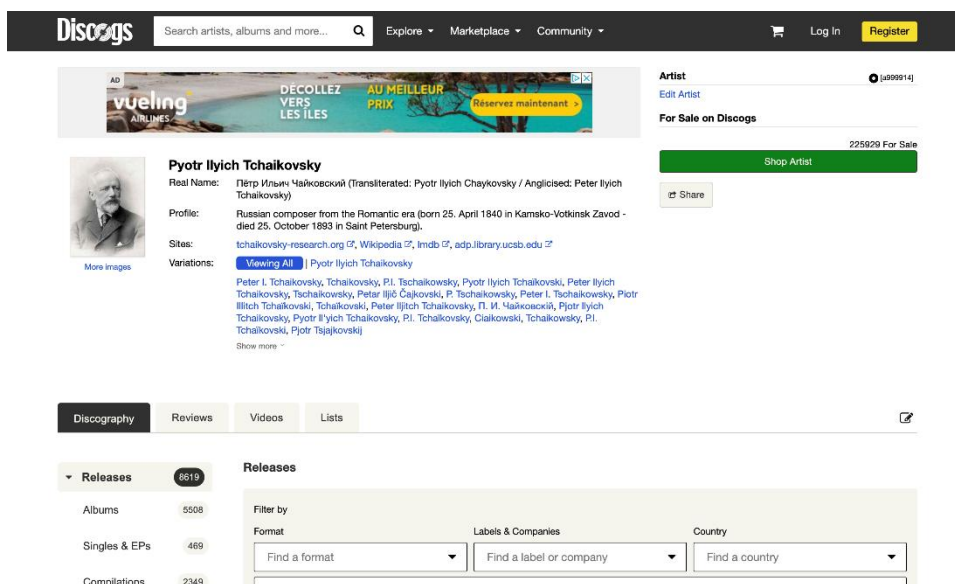


Fig. 4.4.6 Discogs page of Tchaikovsky linked from the web portal

Example: The user lands on the home page⁴⁴ and scrolls to the section dedicated to artists. S/he enters "Tchaikovsky" in the search bar. The search returns a list of results connected to the composer, including related music genres and digital archives that include materials relevant to the artist. To find more information about the artist, S/he clicks on the expand button, which shows an Insights page dedicated to the composer (Fig. 4.4.4), including pictures, textual descriptions, and related resources coming from four different datasets. The user scrolls the page⁴⁵ and discovers that the artist is related to a digital collection,

⁴⁴ <https://polifonia.disi.unibo.it/portal/>

⁴⁵

https://polifonia.disi.unibo.it/portal/card?title=Tsjajkovski,%20Pjotr%20Ijitsj&cat=artists&input=%20no%20input%20&hasinput=true&uri=http://w3id.org/polifonia/linkset/d_01_d_03_d_04_d_05_cat_02_e645e8a6-5957-4482-b92b-4c7f6ec3a853

which is also described in a dedicated page on the Polifonia web portal⁴⁶ (Fig.4.4.5), and he is linked to a Discogs page⁴⁷ (Fig. 4.4.6).

Use Case 3: Music tourism

This use case describes a user that wants to discover places connected by the influence that music genres had on them. The case study is inspired by participants' suggestions collected during the Co-design User Test held in March 2023 (see Scenario 2) and the Expert Usability Review held in October 2023.

Objective: Find places to visit related to a place and a genre.

Datasets: DBpedia, MEETUPS

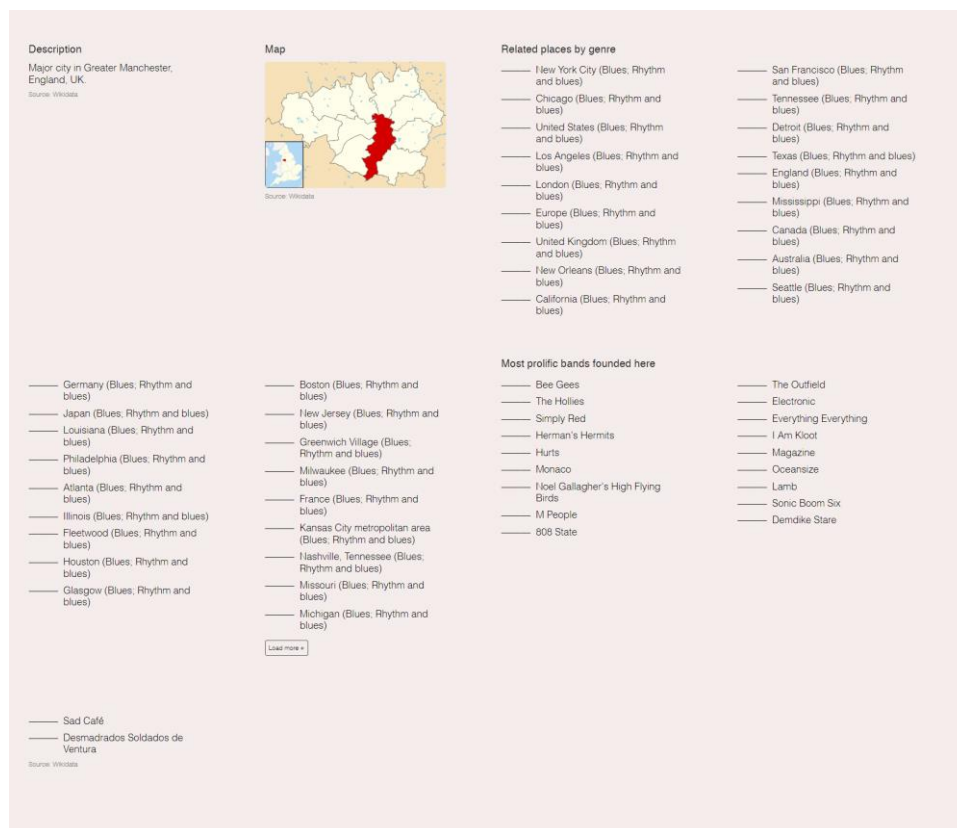


Fig. 4.4.7 Insight page of Manchester

Example: The user lands on the home page and scrolls to the section dedicated to Places. S/he enters "Manchester" in the search bar. S/he clicks on the expand button, which shows a page dedicated to the

⁴⁶

<https://polifonia.disi.unibo.it/portal/card?title=Generative%20Theory%20of%20Tonal%20Music&cat=creative%20works&uri=https://w3id.org/musow/1645534899-5232813>

⁴⁷ <https://www.discogs.com/artist/999914>

city (Fig. 4.4.7)⁴⁸. This includes metadata about the city and a list of related similar places influenced by the same music genres that influenced Manchester. This information is extracted from MEETUPS and DBPedia. Notably, each place appears alongside the music genres that are relevant to both the place at hand and Manchester (e.g. *New York City (Blues; Rhythm and blues)*). In detail, the relation is built on top of artists significative to the music history of Manchester (as recorded in MEETUPS), the music genres associated with them (as extracted from DBPedia), other artists relevant to the same music genre (from DBpedia), and finally the places where the related artists have performed or are somehow connected to (again from MEETUPS). Places are sorted by the number of events (e.g. meetups, concerts, performances) relevant to artists belonging to the music genre at hand (e.g. Blues) that happened in the related place (e.g. New York). This method allows us to create a network of places interconnected via the impact that a particular music genre has had on them, therefore showcasing “twin” cities or countries sharing the same music genres.

48

https://polifonia.disi.unibo.it/portal/card?title=Manchester&cat=places&uri=http://w3id.org/polifonia/linkset/d_05_d_02_d_04_cat_04_c91de2e6-24c6-4dfd-857e-e5da30d4e94e

5. Discussion and Conclusion

Expected results of WP1.T3 proposed in the Grant Agreement have all been accomplished. **A catalogue of music data sources (musoW), methods for indexing and population (CLEF), as well as tools for data exploration, browsing, search, and discovery (MELODY, Web portal) have been successfully delivered and evaluated.** The status of the web portal is still prototypical, since pilot datasets are to be released and will be integrated in later stages. Even though no sensible changes are expected in the UI/UX design, the web portal may be subject to further revisions to accommodate partners' requests.

5.1 Impact and sustainability of results

Overall, the software solutions developed in WP1.T3 have reached a maturity level sufficient to plan their reuse and estimate their potential impact in current and future projects.

musoW is an invaluable resource for mapping the landscape of music data on the web, which has a wider coverage than existing competitors and has applications in several scenarios. It leverages crowdsourcing techniques to ensure its continuous population and persistence in the future.

CLEF demonstrated to be an appealing solution to support mid-size projects in the cultural heritage domain, filling a gap in Linked-Open-Data-native platforms for crowdsourcing. Its reuse in 4 external projects shows promising results and external funding ensures its maintenance in the mid-term.

MELODY has been actively used as both a teaching tool and an exploratory platform for experts, therefore expanding its applicability in future scenarios. Research topics and further UI/UX developments relevant to MELODY are the subject of two PhD projects (candidates supervised by the UNIBO research group) and will ensure continuity in development and improvement.

Web portal. While the impact of the Polifonia web portal itself cannot be currently evaluated, since not all pilot datasets are available at the time of the deliverable submission, we can estimate the impact of the software solution underlying it. The focus on reconciliation and data alignment operations accommodates issues that are rather common in the LOD community. Methods developed have already arisen interest in external projects, like PHAROS⁴⁹, 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.

5.2 Limitations and future works

Current limitations are the subject of future work, which will continue after the end of the project.

musoW. The development of methods for **semi-automatic discovery and population** of musoW has been disrupted by the change of policy of Twitter/X APIs on which we based the initial development – X is now neglecting any form of free-of-charge usage for research purposes. In the future we will explore other venues for continuous discovery and extraction of music data sources.

⁴⁹ <http://pharosartresearch.org/>

MELODY. The main limitation in the wide adoption of the software concerns the technical skills required by users. Indeed, a medium-level **knowledge of SPARQL** and Semantic Web technologies is required to master the queries, as well as an understanding of how charts work. Likewise, data visualization literacy and sufficient knowledge of how a chart is built are necessary. In the future we aim to support lay users by designing interfaces for visual query and chart guessing. We also plan to integrate a guided tutorial for first-time users visiting MELODY. This tutorial will walk them through the process of creating a data story in a few simple steps. Existing tools like `intro.js`⁵⁰ can be utilized in the front-end of the website to facilitate this task.

Web portal. The prototype attempts to solve common issues in LOD datasets, i.e. entity alignment across sources that do not provide direct *sameAs* links. We are aware that there are still several limitations in our approach and that duplicates are still present in our indexes, which depend on limitations in the original data that we are not able to cope with on the web portal, and would require intervention in the original data sources. Moreover, (1) we cannot estimate whether we effectively addressed all potential data quality issues, which will be subject of future work in collaboration with new stakeholders (PHAROS), and (2) we cannot evaluate the efficacy of our work on pilot datasets that are not ready for publication, which will likely require revisions in the future. Nonetheless, as aforementioned, the data ingestion process designed for the portal is flexible enough to allow pilot leaders to integrate their data in the future without too much supervision, which should ensure smooth collaboration and improvement of the web portal even after the end of the project.

Among future activities we also plan **monitoring activities** on the Web portal, to estimate the engagement and evaluate user journeys. Lastly, we have **planned publications** in academic venues, namely: (1) a journal article on the methodology used to develop our software solutions (to be submitted to [Umanistica Digitale](#), an Italian A-ranked journal in Digital Humanities), (2) a conference paper to present the web portal (to be submitted to AIUCD, the Italian Conference of digital Humanities), and (3) an extended full article on the usage and evaluation of MELODY (to be submitted on [Digital Scholarship in the Humanities](#), an international A-Ranked journal).

⁵⁰ <https://github.com/usablica/intro.js/>

6. Compliance to the FAIR principles

This section encompasses information which helps to the knowledge management in the project and to make Polifonia FAIR (operating according to principles of Findable, Accessible, Interoperable, and Reusable)

6.1 Research Ecosystem Components

The four research outputs presented in section 4 correspond to four Research Ecosystem Components, which all comply to the Polifonia 10 Rules for Open Science. Main deviance from rules concern R3, i.e. the publication of results on the Polifonia server: musoW and MELODY are currently deployed on servers of the DH.arc centre of the University of Bologna to ensure their long-term maintenance and guarantee access to developers even beyond the end of the Polifonia project.

musoW (Dataset).

- Github: https://github.com/polifonia-project/registry_app
- Ecosystem: YES
- Zenodo: <https://zenodo.org/doi/10.5281/zenodo.5603222>
- PON alignment: No. It uses Schema.org ontology to foster findability of indexed music data sources on Google.
- License: CC0

CLEF (Reusable software)

- Github: <https://github.com/polifonia-project/clef/>
- Ecosystem: YES
- Zenodo: <https://zenodo.org/doi/10.5281/zenodo.6423932>
- License: ISC

MELODY (Reusable software)

- Github: <https://github.com/polifonia-project/dashboard>
- Ecosystem: YES
- Zenodo: <https://zenodo.org/doi/10.5281/zenodo.6637345>
- License: ISC

Web portal (Reusable software)

- Github: <https://github.com/polifonia-project/portal>
- Ecosystem: YES
- Zenodo: <https://zenodo.org/doi/10.5281/zenodo.10454048>
- License: ISC

Bibliography

1. **CHANGES**. Cultural Heritage Active Innovation for Next-Gen Sustainable Society.
<https://sites.google.com/uniroma1.it/changes/home>
2. **Daga et al. 2021**. Daga, Enrico, Albert Meroño Peñuela, Marilena Daquino, Fiorela Ciroku, Elena Musumeci, Marco Gurrieri, Andrea Scharnhorst, Femmy Admiraal, and Raphaël Fournier-S'niehotta. D1. 3: Pilots development–collaborative methodology and tools (V1. 0). Technical Report, 2021. doi: 10.5281/zenodo. 7712909, 2021.
3. **Daga et al. 2023**. Daga, E., Carvalho, J., Gurrieri, M., & Scharnhorst, A. (2023). D2.6: Ontology of licencing, ownership and conditions of use (V1.0). Zenodo.
4. **Daquino et al. 2022**. Daquino, Marilena, Albert Meroño Peñuela, Christophe Guillotel-Nothmann, e Paul Mulholland. «D1.9: Polifonia Web Portal - 1st Version (V1.0)». Zenodo, 2022. <https://doi.org/10.5281/zenodo.7116711>.
5. **Daquino et al. 2023**. Daquino, Marilena, Mari Wigham, Enrico Daga, Lucia Giagnolini, and Francesca Tomasi. "CLEF. A linked open data native system for crowdsourcing." *ACM Journal on Computing and Cultural Heritage*, 16: 3. DOI: <https://doi.org/10.1145/3594721>
6. **GEL**. <https://projects.dharc.unibo.it/digestgel/>
7. **Grasso, Daquino, Renda 2023**. Grasso, Marco, Marilena Daquino, e Giulia Renda. «MELODY User Survey». Zenodo, 29 dicembre 2023. <https://doi.org/10.5281/zenodo.10442429>.
8. **KNOT**. <https://icdp-digital-library.github.io/KNOT/>
9. **Po et al. 2020**. Po, N. Bikakis, F. Desimoni, and G. Papastefanatos, "Linked data visualization: techniques, tools, and bigdata, "Synthesis Lectures on Semantic Web: Theory and Technology, vol. 10, no. 1, pp. 1–157, 2020.
10. **PRIN2022**. The ATLAS of Italian Digital Humanities projects. <https://dh-atlas.github.io/>
11. **Renda, Daquino 2023**. Renda, Giulia, Daquino, Marilena. 2023. «Polifonia-project/dashboard: V1.0.0». Zenodo, 24 ottobre 2023. <https://doi.org/10.5281/zenodo.10037673>.
12. **Renda, Daquino and Presutti 2023**. Giulia Renda, Marilena Daquino, and Valentina Presutti. 2023. Melody: A Platform for Linked Open Data Visualisation and Curated Storytelling. In *Proceedings of the 34th ACM Conference on Hypertext and Social Media (HT '23)*. Association for Computing Machinery, New York, NY, USA, pp. 1–8. DOI: <https://doi.org/10.1145/3603163.3609035>
13. **Renda, Grasso and Daquino 2023**. Renda, Giulia, Marco Grasso, and Marilena Daquino. 2023. "From ontology design to user-centred interfaces for music heritage." In *Proceedings of AIUCD 2023* [in press]. Preprint available at: <https://arxiv.org/abs/2306.12973v1>
14. **Renda, Grasso and Daquino 2024**. Renda, Giulia, Marco Grasso, and Marilena Daquino. 2023. «Polifonia-project/portal: V1.0.0». Zenodo, 3 gennaio 2024. <https://doi.org/10.5281/zenodo.10454048>.
15. **Scharnhorst et al. 2023**. Scharnhorst, Andrea, Rene Van Horik, Enrico Daga, Marilena Daquino, Elena Musumeci, Peter Van Kranenburg, Christophe Guillotel-Nothmann, et al. «D7.2 Data Management Plan (second Version)». Zenodo, 2023. <https://doi.org/10.5281/zenodo.7660299>.

16. **Scharnhorst et al. 2021.** Scharnhorst, A., Admiraal, F., van Kranenburg, P., Guillotel-Nothmann, C., & Mulholland, P. (2021, September 7). The Polifonia portal: a confluence of user stories, research pilots, data management and knowledge graph technology. DARIAH Annual event 2021, Interfaces, virtual. Zenodo. <https://doi.org/10.5281/zenodo.5483727>
17. Whitelaw M. 2015. “Generous interfaces for digital cultural collections,” Digital Humanities Quarterly 9(1).