



Project Acronym: Europeana Sounds
Grant Agreement no: 620591
Project Title: Europeana Sounds

D2.9 Evaluation report on implementation of semantic enrichment

Revision: Final

Date: 02/02/2017

Authors: Maarten Brinkerink (NISV), Giulio Andreini (Net7), Breandán Knowlton (Historypin), Hugo Manguinhas (EF), Vassilis Tzouvaras (NTUA), Sergiu Gordea (AIT), David Haskiya (EF)

Abstract: This report documents the implementation of semantic enrichment within the context of the Europeana Sounds project. It reports on the different technical components that have been (further) developed to support the objectives and application scenarios of the project. It then goes on to highlight the role of crowdsourcing within the semantic enrichment and the application scenarios and supporting activities that the project realised. Finally this report evaluates and reflects on the success of semantic enrichment in the context of Europeana Sound against the KPIs and provides a summary of conclusions.

Dissemination level	
Public	X
Confidential, only for the members of the Consortium and Commission Services	



Revision history

Version	Status	Name, organisation	Date	Changes
0.1	ToC	Maarten Brinkerink, NISV	16/12/2016	
0.2	1 st draft	Giulio Andreini (Net7), Breandán Knowlton (Historypin), Maarten Brinkerink (NISV)	05/01/2017	First technical partner contributions
0.3	2 nd draft	Hugo Manguinhas (EF), Vassilis Tzouvaras (NTUA), Maarten Brinkerink (NISV)	12/01/2017	Additional technical partner contributions
0.4	3 rd draft	Sergiu Gordea (AIT), Maarten Brinkerink	19/01/2017	Additional technical partner contributions
0.5	4 th draft	Maarten Brinkerink (NISV)	23/01/2017	Reworked individual input from respective technical partners into a narrative
0.6	5 th draft	Giulio Andreini (Net7), Breandán Knowlton (Historypin), Hugo Manguinhas (EF), Vassilis Tzouvaras (NTUA), Sergiu Gordea (AIT), David Haskiya (EF)	25/01/2017	Final input from technical partners
0.7	5 th draft	Sander Pieterse	26/01/2017	External review comments
1.0	Final	Richard Ranft, BL	31/01/2017	Layout, minor changes

Review and approval

Action	Name, organisation	Date
Reviewed by	Sander Pieterse (Naturalis Biodiversity Center)	26/01/2017
Approved by	Coordinator and PMB	31/01/2017

Distribution

No.	Date	Comment	Partner / WP
1	02/02/2017	Submitted to the European Commission	BL/WP7
2	02/02/2017	Posted on Europeana Pro website	BL/WP7
3	02/02/2017	Distributed to project consortium	BL/WP7

Application area

This document is a formal output for the European Commission, applicable to all members of the Europeana Sounds project and beneficiaries. This document reflects only the author's views and the European Union is not liable for any use that might be made of information contained therein.



Statement of originality

This document contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

Project summary

Europeana Sounds is Europeana's 'missing' fourth domain aggregator, joining APEX (Archives), EUscreen (television) and the Europeana film Gateway (film). It will increase the opportunities for access to and creative re-use of Europeana's audio and audio-related content and will build a sustainable best practice network of stakeholders in the content value chain to aggregate, enrich and share a critical mass of audio that meets the needs of public audiences, the creative industries (notably publishers) and researchers. The consortium of 24 partners will:

- Double the number of audio items accessible through Europeana to over 1 million and improve geographical and thematic coverage by aggregating items with widespread popular appeal such as contemporary and classical music, traditional and folk music, the natural world, oral memory and languages and dialects.
- Add meaningful contextual knowledge and medium-specific metadata to 2 million items in Europeana's audio and audio-related collections, developing techniques for cross-media and cross-collection linking.
- Develop and validate audience specific sound channels and a distributed crowd-sourcing infrastructure for end-users that will improve Europeana's search facility, navigation and user experience. These can then be used for other communities and other media.
- Engage music publishers and rights holders in efforts to make more material accessible online through Europeana by resolving domain constraints and lack of access to commercially unviable (i.e. out-of-commerce) content.

These outcomes will be achieved through a network of leading sound archives working with specialists in audiovisual technology, rights issues, and software development. The network will expand to include other data-providers and mainstream distribution platforms (Historypin, SoundCloud) to ensure the widest possible availability of their content.

For more information, visit <http://pro.europeana.eu/web/europeana-sounds> and <http://www.europeanasounds.eu>

Copyright notice

Copyright © Members of the Europeana Sounds Consortium, 2014-2017. This work is licensed under the Creative Commons CC-BY License: <https://creativecommons.org/licenses/by/4.0/>.



Contents

Executive summary: D2.9 Evaluation report on implementation of semantic enrichment 6

1	Introduction	7
2	Europeana Sounds semantic enrichment strategies	7
2.1	Enrichment during ingestion	7
2.2	Enrichment through crowdsourcing	8
2.3	Enrichment through alignment	8
3	Core systems for semantic enrichment	8
3.1	Semantic enrichment and the entity collection in the core infrastructure	9
3.2	Annotations API	10
3.3	Alignment applications	11
3.3.1	CultuurLink experiments	11
3.3.2	Phase 1: Vocabulary alignment of subject terms to MIMO	11
3.3.3	Phase 2: suggesting annotations in the crowdsourcing campaigns	12
3.3.4	Wikidata and Wikimedia Commons experiments	13
3.3.5	Linked Open Data vocabularies	15
4	Semantic enrichment through crowdsourcing	16
4.1	Types of enrichment	16
4.1.1	Semantic tagging	16
4.1.2	Geotagging	17
4.1.3	Object linking	17
4.1.4	Moderation (or validation through crowdsourcing)	17
4.2	Semantic tagging in Tunepal and Historypin	18
4.2.1	4.2.1 Traditional music pilot	18
4.2.2	Geo-Location crowdsourcing	20
4.3	Semantic tagging in Pundit	24
4.3.1	Features	24
4.3.2	Sound object annotation	24
4.3.3	MIMO vocabulary and thesaurus	26
4.3.4	Moderate previously created annotations	26
4.3.5	Multilingual support	27
4.3.6	Workflow improvements	28
4.3.7	Pundit integrations	29
4.3.8	Data model	30
4.3.9	Discussion	31
4.4	Semantic tagging in WITH	31
4.4.1	Crowdsourcing space	31
4.4.2	Tagging Process	33
4.4.3	Game mechanics	35



4.4.4	Discussion	36
4.5	Semantic tagging in Europeana Radio	36
4.6	Semantic tagging in the crowdsourcing campaigns	38
4.6.1	Semantic tagging in the first crowdsourcing campaign	39
4.6.2	Semantic tagging in the second crowdsourcing campaign	40
5	Semantic enrichment results	42
5.1	KPI-8	42
5.2	KPI-9	43
5.3	KPI-10	43
5.4	Lessons learned	44
6	Conclusions	45
	References	46
	Appendix A: Terminology	47



Executive summary: D2.9 Evaluation report on implementation of semantic enrichment

In order for Europeana Sounds to meet its objective to increase the quality of the metadata, WP2 deployed a threefold strategy to pursue semantic enrichment of the aggregated data: enrichment during ingestion, enrichment through crowdsourcing, and enrichment through alignment.

The enrichment strategies informed the (further) development of several core systems as components - namely the Entity API and Annotations API - in the Europeana core infrastructure. These components aimed to support the application scenario developed by the project (most prominently the crowdsourcing application), but also allowed for experimentation with various ways to align and link the data held by Europeana to external knowledge structures and resources (like MIMO and Wikidata). These experiments showed great potential for this semantic enrichment strategy for the future.

For semantic enrichment through crowdsourcing specifically, the project modelled and developed four types of annotations: semantic tagging, geotagging, object linking and moderation. This work contributed greatly to the state-of-the-art development of a W3C standard for web annotations (WADM, which all Europeana Sounds enrichments comply with).

The types of annotations developed by the project were deployed in various crowdsourcing applications: Traditional Music Pilot (object linking), Historypin (geotagging), Pundit (semantic tagging, moderation), WITH (semantic tagging, moderation), and Europeana Radio (semantic tagging).

The crowdsourcing applications have been promoted through two crowdsourcing campaigns. The results in terms of user engagement have varied (depending on maturity of the application, subject matter and expertise required for the crowdsourcing task). In some cases it has also suffered from late delivery. This calls for post-project continuation of the utilisation of these applications, in order for an apt valorisation of the development and impact of the enrichment through crowdsourcing. The project has already identified some opportunities to support this and NISV has committed to continuing some promotion towards end-users and liaising between data providers and the Technical Partners.

Looking at the results achieved with the various strategies, one can assess that enrichment during ingestion is already standard practice at EF. This strategies has mainly been improved through the project by connecting this to the Entity API, hence further normalising the semantic enrichment through a curated set of entities in the Entity Collection (which are all part of the data.europeana.eu namespace). Enrichment through crowdsourcing has performed less than was originally estimated (at



the beginning of the project), but still shows potential. Finally enrichment through alignment - although experimental of nature in this project - outperformed the expected results and shows great potential, especially since it can also be combined with crowdsourcing for a multiplier effect.

1 Introduction

This report documents the implementation of semantic enrichment within the context of the Europeana Sounds project (see Section 2: Europeana Sounds Semantic Enrichment Strategies). It reports on the different technical components that have been (further) developed to support the objectives and application scenarios of the project (see Section 3: Core Systems for Semantic Enrichment). It then goes on to highlight the role of crowdsourcing within the semantic enrichment and the application scenarios and supporting activities that the project realised (see Section 4: Semantic Enrichment through Crowdsourcing). Finally this report evaluates and reflects on the success of semantic enrichment in the context of Europeana Sound against the KPIs (see Section 5: Semantic Enrichment Results) and provides a summary of conclusions (see Section 6: Conclusions).

2 Europeana Sounds semantic enrichment strategies

The Description of Work (Ref 1) for Europeana Sounds explains the role of semantic enrichment for WP2 Enrichment & participation as following:

“Apply semantic web technologies to enable enrichment of the Europeana Sounds collections. This will **increase quality of the metadata** and **user satisfaction in terms of content discovery.**”
(Ref 1)

To meet this objective the project developed three strategic approaches related to semantic enrichment that are described below.

2.1 Enrichment during ingestion

The first strategy for applying semantic enrichment is implemented at the point of ingestion by Europeana Foundation (EF). This enrichment strategy directly targets the metadata records that are contributed by data providers and aims to automatically enrich the data as they become part of the Europeana aggregated database. The goals of this semantic enrichment during ingestion is to partly unify the data coming from various data providers using data sources¹ such as Geonames², DBpedia³, GEMET⁴ and Wikidata⁵. This makes search and retrieval of the individual metadata records in Europeana

¹ <http://www.europeana.eu/portal/en/rights/data-sources.html>

² <http://www.geonames.org>

³ <http://dbpedia.org>

⁴ <http://www.eionet.europa.eu/getmet>

⁵ <http://wikidata.org>



more uniform, and adds the benefit of being able to utilise the multilingual labels these various data sources have to offer.

The enrichment during ingestion currently focusses on the Sounds Genre Vocabulary (Ref 2) and entities related to concept, place, agent and timespan. The results of this semantic enrichment strategy are reflected in KPI 8 'Number of metadata records enriched through semantic enrichment' (Ref 1, see Section 5.1).

2.2 Enrichment through crowdsourcing

The second strategy for applying semantic enrichment is aimed at putting semantic web technologies into the hands of end users, as a (useful) form of engagement and participation. To facilitate this strategy WP2 has (further) developed several core services (see section 3) and crowdsourcing applications (see section 4) that allow end users to apply specific types of semantic enrichments (see Section 4.1) to particular collections in the form of crowdsourcing tasks. To promote these crowdsourcing applications and stimulate the semantic enrichment through crowdsourcing, the project organised several crowdsourcing campaigns (see section 4.6). The results of this semantic enrichment strategy contribute to KPI 9 'Number of annotations (e.g. tags) added by users' (Ref 1, see section 5.2).

2.3 Enrichment through alignment

The third strategy for applying semantic enrichment is aimed at experimenting with the combination of expert knowledge from data managers, curators and amateurs with smart technologies to enable semi-automatic methods of enrichment (see section 3.3). By (manually) aligning recurring terms extracted from the unstructured data present in the metadata records from specific data providers with Linked Open Data sources after they are published, this approach allows for semantic enrichments in knowledge areas or topics that are more specific than the general enrichment that is realised with the semantic enrichment during ingestion. The results of this approach can then be transformed into suggestions for annotations that can be validated through crowdsourcing at the metadata record level, creating a feedback loop for the semantic enrichment. The results of this semantic enrichment strategy contribute to KPI 10 'Number of new connections among records established by users and automatically' (Ref 1, see Section 5.3).

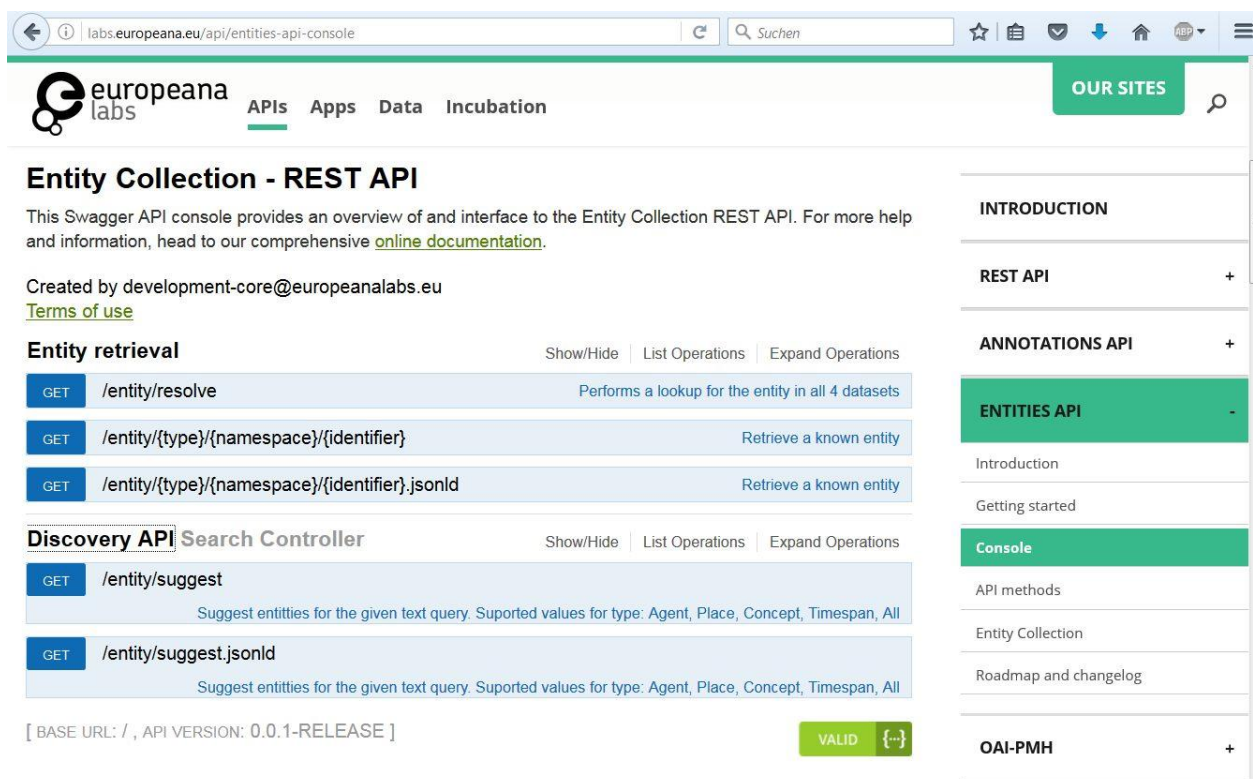
3 Core systems for semantic enrichment

To support the strategies mentioned above (see Section 2), and to enable the application scenarios described later in this document (see Section 4) we have (further) developed several components in the Europeana core infrastructure.

3.1 Semantic enrichment and the entity collection in the core infrastructure

The Europeana Entity API⁶ allows core applications and third parties to retrieve and search for named entities used in Europeana. These named entities (such as persons, topics and places) are part of the Europeana Entity Collection (EC), a collection of relevant entities in the context of Europeana harvested from and linked to controlled vocabularies that are available as Linked Open Data, such as Wikidata, Geonames and DBpedia. The Entity API allows one to facilitate search on or retrieve information from named entities (within or without the Europeana context). In 2016, the Entity Collection also took the role of the vocabulary used by Europeana to automatically semantically enrich all metadata records in Europeana including of course all records contributed by the Europeana Sounds project (see section 5.1).

The Entity Collection was developed to also support specific functionality used within the Europeana Collections (WP4). However it is also used as a repository of controlled vocabularies used within the Annotation API. The first public alpha release of the Entity API will be launched in early 2017 and similarly to the other Europeana APIs, this is empowered with a Swagger⁷ based console⁸, which allows developers to test its full functionality using standard web browsers (see Figure 1).



The screenshot shows the Swagger API console for the Entity Collection REST API. The browser address bar shows the URL: `labs.europeana.eu/api/entities-api-console`. The page header includes the Europeana Labs logo and navigation links for APIs, Apps, Data, and Incubation. A search bar is present in the top right corner.

The main content area is titled "Entity Collection - REST API" and includes a description: "This Swagger API console provides an overview of and interface to the Entity Collection REST API. For more help and information, head to our comprehensive [online documentation](#)." It also mentions the creator: "Created by development-core@europeanalabs.eu" and provides a link to "Terms of use".

Under the heading "Entity retrieval", there are three API endpoints listed:

- `GET /entity/resolve`: Performs a lookup for the entity in all 4 datasets
- `GET /entity/{type}/{namespace}/{identifier}`: Retrieve a known entity
- `GET /entity/{type}/{namespace}/{identifier}.jsonld`: Retrieve a known entity

Below this, the "Discovery API Search Controller" section lists two endpoints:

- `GET /entity/suggest`: Suggest entities for the given text query. Supported values for type: Agent, Place, Concept, Timespan, All
- `GET /entity/suggest.jsonld`: Suggest entities for the given text query. Supported values for type: Agent, Place, Concept, Timespan, All

At the bottom left, the base URL is shown as `[BASE URL: / , API VERSION: 0.0.1-RELEASE]`. A green "VALID" button with a refresh icon is located at the bottom right.

On the right side, there is a sidebar menu titled "OUR SITES" with a search icon. The menu items are: INTRODUCTION, REST API (+), ANNOTATIONS API (+), ENTITIES API (-), Introduction, Getting started, Console (highlighted), API methods, Entity Collection, Roadmap and changelog, and OAI-PMH (+).

Figure 1. Entity API Console

⁶ <http://labs.europeana.eu/api/entities>

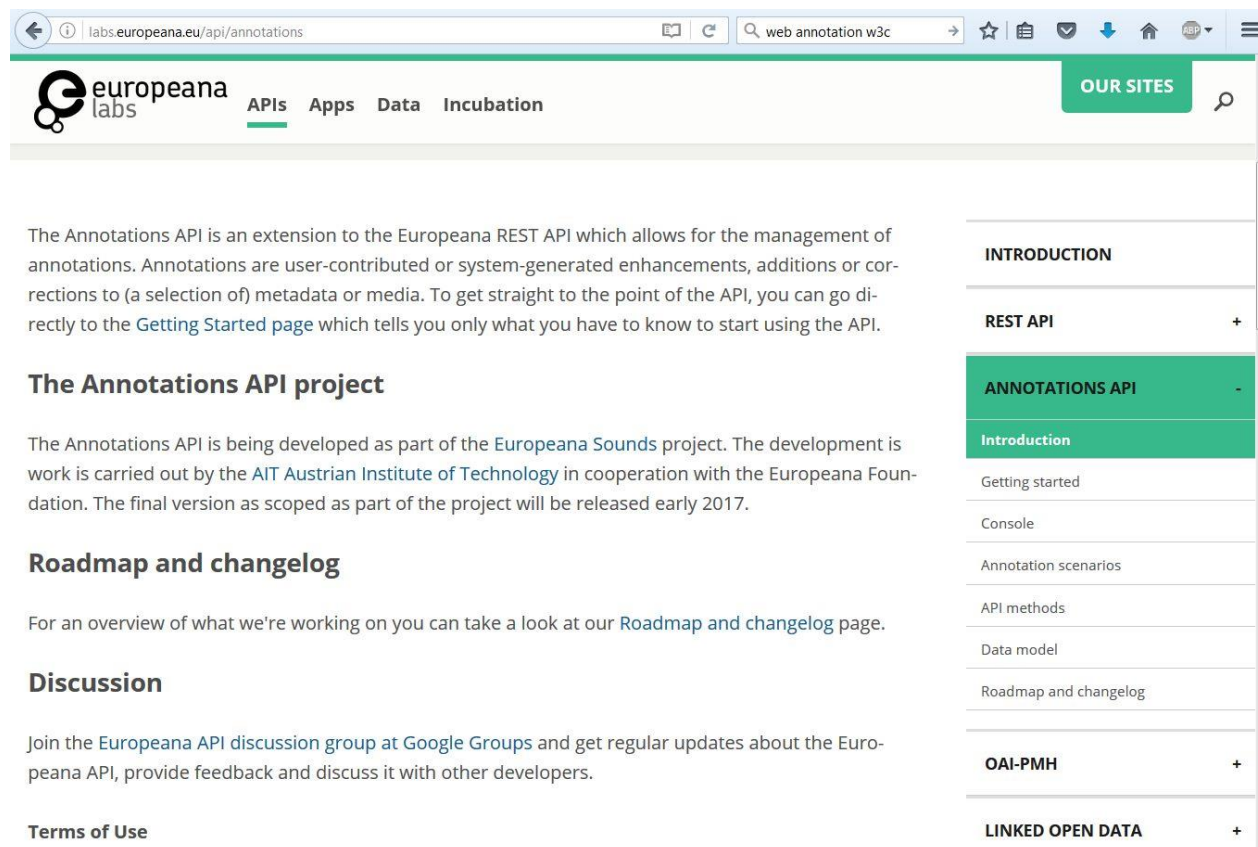
⁷ <http://swagger.io/>

⁸ <http://labs.europeana.eu/api/entities-api-console>

3.2 Annotations API

The Annotations API is an extension to the Europeana API which allows for the management of user-contributed or system-generated semantic enrichments. The Annotation API is used as the central repository storing the semantic enrichments generated through the crowdsourcing campaigns. The description of the first public release was introduced in D2.10 *Development of the crowdsourcing infrastructure* (Ref 3).

In 2016, there were several subsequent releases of this API, the latest one being deployed into the production environment in December 2016 (Version 0.2.6)⁹. The full documentation for the Annotations API is published in Europeana Labs¹⁰ (see Figure 2).



The screenshot shows a web browser displaying the Europeana Labs Annotations API documentation page. The browser address bar shows 'labs.europeana.eu/api/annotations'. The page header includes the Europeana Labs logo and navigation links for 'APIs', 'Apps', 'Data', and 'Incubation'. A search bar with 'OUR SITES' is also visible. The main content area features a paragraph describing the API as an extension to the Europeana REST API for managing annotations. Below this is a section titled 'The Annotations API project' followed by a paragraph about its development as part of the Europeana Sounds project. A 'Roadmap and changelog' section follows, with a link to a 'Roadmap and changelog page'. A 'Discussion' section invites users to join a Google Groups discussion. At the bottom, there is a 'Terms of Use' section. On the right side, a table of contents lists various sections: 'INTRODUCTION', 'REST API', 'ANNOTATIONS API' (highlighted in green), 'Introduction', 'Getting started', 'Console', 'Annotation scenarios', 'API methods', 'Data model', 'Roadmap and changelog', 'OAI-PMH', and 'LINKED OPEN DATA'.

Figure 2: Documentation of Annotation API in Europeana Labs

The main goal of the development work was to ensure full compatibility with the W3C specifications for the Web Annotation Data Model¹¹ (WADM) and Web Annotation Protocol¹² (WAP). EF and AIT were actively contributing¹³ to the development of these standards by providing feedback on the specifications, and by acting as an official implementer¹⁴ of the standard. This contributed to the

⁹ <http://labs.europeana.eu/api/annotations-roadmap-changelog>

¹⁰ <http://labs.europeana.eu/api/annotations>

¹¹ <https://www.w3.org/TR/annotation-model/>

¹² <https://www.w3.org/TR/annotation-protocol/>

¹³ <https://www.w3.org/TR/2017/PR-annotation-model-20170117/#acknowledgements>

¹⁴ <https://github.com/w3c/test-results/blob/gh-pages/annotation-model/README.md>



standardisation process, supporting the efforts to move the standard forward from candidate recommendation status into proposed recommendation¹⁵. The compliance results for the WADM computed against the 0.2.4 Version of the Annotation API are available online¹⁶, while the compliance tests for WAP are currently worked on, and they are expected to be published after the project end (January 2017).

3.3 Alignment applications

The third strand of work was more experimental of nature and consisted of further enhancing the semantic enrichment supported through the standards workflows (aggregation, ingestion and user engagement) in the Europeana Core Infrastructure with more tailored and domain-specific actions. To enable this, these experiments made use of external applications and/or data sources for the semantic enrichment, through semi-automatic workflows.

3.3.1 CultuurLink experiments

These experiments focused on exploring the feasibility of vocabulary alignment to semantically enrich the terms appearing in the source metadata - as contributed by the data providers - with semantic resources from controlled vocabularies. It focussed in particular on linking the elements of an internal (local) vocabulary - as used by a data provider - to semantically related elements from a richer vocabulary, that were then used to generate suggestions for enrichment of the related metadata records as annotations.

This experiment was divided into two phases: A first phase where the vocabulary alignments were identified, resulting in a first list of possible annotations for the metadata records. And a second phase in which these annotations would be used as suggestions for the crowdsourcing campaigns to validate.

3.3.2 Phase 1: Vocabulary alignment of subject terms to MIMO

A vocabulary alignment tool, named CultuurLink¹⁷, was selected for this experiment as it applies a semi-automatic approach to alignment, to make the task less labour-intensive, yet still taking benefit from the expertise of a data manager required to assert the right alignments. It is the successor of the Amalgame framework¹⁸ developed in the Europeana Connect project¹⁹, but presently being developed by Spinque²⁰.

We asked data providers that were contributing data to Europeana as part of WP1 to select a collection of metadata records pertaining to sound object that could contain terms related to musical instruments.

¹⁵ <https://www.w3.org/TR/2017/PR-annotation-model-20170117/#candidate-recommendation-exit-criteria>

¹⁶ <https://w3c.github.io/test-results/annotation-model/all.html>

¹⁷ <http://cultuurlink.beeldengeluid.nl/>

¹⁸ <http://semanticweb.cs.vu.nl/lod/tpdl2011/paper.pdf>

¹⁹

http://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/EuropeanaConnect/Deliverables/ECO_NNECT-D1.3.1-Semantic%20Workflow%20Automation%20Method%20Implementation.pdf

²⁰ <http://www.spinque.com/>



For this experiment we considered only musical instruments terms within subject fields (dc:subject), as instructed by data providers. A total of 6 datasets containing a total of 10,406 metadata records were obtained from the data providers and used in this experiment.

As a significant number of terms within subject fields of the Europeana Sounds metadata records are related to musical instruments we chose the Musical Instruments Museums Online Vocabulary and Thesaurus²¹ (MIMO) as target vocabulary for this experiment, following the recommendations made in the report²². MIMO is a multilingual controlled vocabulary of musical instruments built to ensure consistency of classification for the musical instruments²³. It is a result of an alignment of a vernacular classification with the professional “HornbostelSachs” classification²⁴. The vocabulary has been built with English as pivot language, and translations in seven other languages have been added.

We decided to focus on the vocabulary terms as they are used, i.e. present within the subject fields of the metadata delivered to Europeana. We chose to do this, as opposed to aligning the full vocabulary used by the data providers, since: these were not available for use outside the organisation and/or in a data structure that suits a vocabulary alignment tool (e.g. SKOS), and furthermore, we did not have the opportunity, nor the resources, to develop an export to SKOS for each vocabulary; and, we preferred to report on alignments for the subjects that were actually used in the source datasets that were contributed and not on all possible subjects.

We asked the data providers, namely BL, CREM, MMSH, and NISV, to design and apply alignment strategies in CultuurLink and then evaluate the alignments (i.e., validate the links and assign them a SKOS mapping property). Once all the participants had finished their task we collected the alignment results and summarised the findings.²⁵ This phase of the experiment resulted in 187 aligned terms, which generated 35,006 suggested annotations to the metadata records. We consider this a contribution to KPI-8 (Ref 1). This is an updated calculation in relation to previously reported results in Q10 of the project.

As well as the interest from WP2 to obtain annotations for the crowdsourcing campaign, this experiment also had the ambition to introduce data providers to vocabulary alignment tools. In this regard, there was a consensus from the data providers that the experiment was successful and they were able to understand and work with the vocabulary alignment tool with good level of success.

3.3.3 Phase 2: suggesting annotations in the crowdsourcing campaigns

In D2.10 *Development of the crowdsourcing infrastructure* (Ref 3) we have identified suggesting annotations in the crowdsourcing campaigns as an additional strategy:

“Alignment as a (semi-) automatic enrichment mechanism works at scale, because it does not apply enrichment at the metadata record level, but aggregates the unique entities found in a dataset and tries to align these terms to controlled vocabularies. If relevant candidates for

²¹ <http://www.mimo-international.com/vocabulary.html>

²² <http://pro.europeana.eu/taskforce/evaluation-and-enrichments>

²³ <http://www.mimo-db.eu/InstrumentsKeywords/>

²⁴ <http://www.mimo-international.com/documents/Hornbostel%2520Sachs.pdf>

²⁵ https://docs.google.com/document/d/1oAenFVa7lul77eP5L16rBHCYPJ_QLGmodeWBNaQAdpc/edit



alignments are found, collection experts and/or data managers can assess the results. These alignments are then applied to all instances of the original (unstructured) unique entity in the metadata records that belong to that dataset. Because alignment is applied at the dataset level, validity of a specific enrichment at the metadata record can vary. This is where validation through crowdsourcing is of great added value to further improve upon the (semi)-automatic enrichments.” (Ref 3)

Given that the project has exceeded performance on KPI-8 (Ref 1), this potential is indeed significant. In the second crowdsourcing campaign that was executed by Europeana Sounds, several crowdsourcing applications supported “moderation” as a type of semantic enrichment.

3.3.4 Wikidata and Wikimedia Commons experiments









Wikidata - one of the most recent Wikimedia projects - is an interesting resource for Europeana. It is “a free and open knowledge base that can be read and edited by both humans and machines. Wikidata acts as central storage for the structured data of its Wikimedia sister projects, including Wikipedia. The content of Wikidata is available under a free license, exported using standard formats, and can be interlinked to other open data sets on the linked data web.”²⁶

As part of the Europeana Sounds Wikidata experiment, the idea to display contextual information coming from Wikidata and Wikimedia Commons on the Europeana Portal and Collections was investigated. The experiment focused on establishing bi-directional links between Europeana and the two aforementioned Wikimedia projects. Establishing bi-directional links, especially between Europeana and Wikidata, opens up the possibility to - in the future - also enrich cultural heritage objects in Europeana with metadata from Wikidata. To be more concrete, this could for example be to establish bi-directional links between a bird sound recording in Europeana and the corresponding species entity in Wikidata. Once the bi-directional link is established, using the Europeana Annotations API, it can later be extended by Europeana to further annotate the bird sound recording using the information Wikidata provides, e.g. annotating the record with the common name for the bird in multiple languages in order to improve its discoverability in search.

The first step of this process was to create a link between Europeana and Wikidata/Wikimedia Commons by adding the Europeana URL in the Wikimedia Commons metadata of the content that was uploaded by partners in Europeana Sounds. The objects uploaded to Wikimedia Commons for the second British Library edit-a-thon (8th of October 2016) were chosen for this Wikidata experiment. NISV performed the batch upload of this material on the 22nd September, including all the links to Europeana (see Figure 3). During the edit-a-thon 19 Wikidata entries were enriched with items from this batch, for instance the Wikidata entry for the bird species Banded Parisoma²⁷.

²⁶ https://www.wikidata.org/wiki/Wikidata:Main_Page

²⁷ <https://www.wikidata.org/wiki/Q509128>

Title	African Grey Hornbill (<i>Tockus nasutus</i>)								
Description	African Grey Hornbill calls, recorded in Kongelai, Kenya								
Source	 <p>This file has been provided by the British Library from its digital collections. It is also made available on a British Library website.</p> <p>भारतीय Deutsch English Español Euskara Français Македонски 中文 +</p>								
Permission (Reusing this file)	<p>This file is licensed under the Creative Commons Attribution 4.0 International license.</p>  <p>You are free:</p> <ul style="list-style-type: none"> • to share – to copy, distribute and transmit the work • to remix – to adapt the work <p>Under the following conditions:</p> <ul style="list-style-type: none"> • attribution – You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work). 								
British Library Shelfmark	W1CDR0001419 BD20								
Copyright Holder	The British Library Board								
Recordist	<table border="1"> <tr> <td>A.R. Gregory</td> <td></td> </tr> <tr> <td>Alternative names</td> <td>Roy Gregory</td> </tr> <tr> <td>Description</td> <td>British <i>sound recordist</i></td> </tr> <tr> <td>Authority control</td> <td>: Q27182537</td> </tr> </table>	A.R. Gregory		Alternative names	Roy Gregory	Description	British <i>sound recordist</i>	Authority control	 : Q27182537
A.R. Gregory									
Alternative names	Roy Gregory								
Description	British <i>sound recordist</i>								
Authority control	 : Q27182537								
Recording Date	1 May 1974								
Europeana link	http://www.europeana.eu/portal/en/record/9200253/BibliographicResource_3000055618328.html								

File history

Click on a date/time to view the file as it appeared at that time.

	Date/Time	Thumbnail	Dimensions	User	Comment
current	09:24, 22 September 2016	 0:00	22 s (337 KB)	Beeld en Geluid Collecties (talk contribs)	GWToolset: Mediabestand wordt aangemaakt voor Beeld en Geluid Collecties.

Figure 3: Metadata of one of the audio objects on Wikimedia Commons used for the WikiData experiment (including a link to Europeana)

Europeana has queried Wikimedia Commons²⁸ and Wikidata databases to identify all²⁹ entries in them where there is a link pointing to a Europeana object and exported the results. The exports have been imported to Europeana, via the [Europeana Annotations API](#), annotating each concerned Europeana object with the links to Wikimedia Commons and/or Wikidata. This work resulted in a total of 3,511 Europeana records to Wikimedia Commons and around 75,000 annotations of Europeana records to Wikidata. Given the crowdsourced nature of the latter data source, we consider the links to Wikidata a contribution to KPI-9 (Ref 1), through an external crowdsourcing platform (Wikidata). The former is considered a contribution to KPI-8 (Ref 1), since most cultural heritage content on Wikimedia Commons is contributed by the data providers themselves. The link-annotations are now searchable and retrievable in the Europeana Annotations API, which in turn makes it possible to display them in the Europeana Portal and Collections.

After the end of the Europeana Sounds project, EF will continue to further develop this Wikimedia-Europeana data connection. As mentioned above, one strand of development will go beyond establishing the bi-directional links to additionally annotate Europeana objects with Wikidata metadata

²⁸ Using the Wikimedia Commons API proved unfeasible so instead we used the Wiki community developed tool [Quarry](#).

²⁹ So not only on such links created by and during Europeana Sounds but all such links ever added to Wikimedia Commons and Wikidata.



as annotations or enrichments (like for instance multilingual labels), another will be to automate the connection, ensuring the bi-directional links are regularly updated.

As part of this development Europeana has also updated its policy³⁰ to allow the display of Wikidata-sourced metadata in Europeana Collections, as long as its origin is clear in the display and distinguishable from metadata provided by the institutional Data provider.

3.3.5 Linked Open Data vocabularies

There are other collections in Europeana for which references to music instruments are not available within the (structured) subject classification fields (i.e. like the dc:subject) but they are available in free text fields such as dc:title and dc:description. There are two such collections from the Europeana Sounds project that were analysed in order to develop semantic enrichment processes that are able to address this situation.

One of these collections was provided by Austrian National Library (ONB) and contains manuscripts of music scores of classical music created by famous Austrian composers³¹. There is a special particularity for this kind of records: the fact that the music scores are created for a concrete set up of interpretation, including the specific music instruments that are meant to play the music. Therefore these records are rich in references to music instruments. However, this information is represented in free text fields only.

The librarians from ONB were creating a vocabulary of music instrument terms, by manually inspecting the record descriptions. This vocabulary was mapped to MIMO, by following the process described in section 3.4. During the mapping process it was also decided to include the broad matches generated with Cultuurlink for generating semantic enrichments (e.g. a record mentioning the “viola d’amore” instrument, was also enriched to indicate that the given music score is played with instruments from the “violins” family).

As a next step in this process, the Europeana API was used to find all records that relate to each mapped music instrument in order to generate the semantic enrichments using MIMO. There are 1,396 records (out of 1,691 available in the dataset) in this dataset for which 1,539 enrichments were generated.

Another collection submitted by Netherlands Institute for Sound and Vision (NISV) contains a subset - with European relevance - of live recordings of contemporary music from the Internet Archive³² (a networking partner in the project). This subset consists of songs played by European bands or songs played by international bands at performances in Europe. Rich information about the music bands and individual musicians performing these songs can be found on the MusicBrainz website³³. In many cases

³⁰ <http://www.europeana.eu/portal/en/rights/data-sources.html>

³¹

http://www.europeana.eu/portal/en/search?q=europeana_collectionName%3A2059216_Ag_EU_eSOUNDS_1001_ONB&per_page=24&view=grid

³²

http://www.europeana.eu/portal/en/search?q=europeana_collectionName%3A%222059210_Ag_EU_eSOUNDS_1014_NISV%22&view=grid

³³ <https://musicbrainz.org/>

the performers of the musical acts are listed together with the instruments they play, which are assigned using a controlled vocabulary maintained by the initiative (partly through crowdsourcing)³⁴.

Within this scenario we exploited the power of Linked Open Data and collected the list of unique music instruments played by the performing artists that are related to this concrete dataset. Also in this case we used the process described in Section 3.4 to align the vocabulary of MusicBrainz instrument terms to the MIMO.

The version of the dataset analysed within this experiment contained 10,562 records, from which we extracted 6,063 names of performing artists. However, only 1,860 of them are known to play music instruments (i.e. often the vocalists of music bands do not play any instruments). By summarising the list of unique instruments played by the given artist, we built a vocabulary with 93 entries, out of which 51 could be automatically mapped to MIMO using the Cultuurlink tool. The mapping of the other part of vocabulary required extra effort to search the appropriate MIMO term, and in many cases, these could be mapped only to a broader term describing the instrument family. This is because the MusicBrainz vocabulary includes some very specific variations of music instruments. Finally, a list of 8,805 unique enrichments was computed for this dataset.

Given the fact that MusicBrainz contains crowdsourced information, it is expected that the completeness of the data will increase with the passing of time, which will result in a higher number of enrichments to be generated if the same experiment is repeated in the future.

4 Semantic enrichment through crowdsourcing

This section highlights the role of crowdsourcing within the semantic enrichment. It documents the application scenarios and supporting activities that the project realised to enable semantic enrichment through crowdsourcing.

4.1 Types of enrichment

In previous deliverables (Ref 3 & 4) we have reported on the various types of enrichment that have been supported by the project, mostly from a technical and data modelling perspective. Here we highlight only the types of enrichment that are of a semantic nature.

4.1.1 Semantic tagging

In Europeana Sounds, and other similar projects, Semantic Tagging is used to enrich the objects with semantic resources from Linked Open Data repositories, or with semantic classification schemes (like genre concepts from the Europeana Genre Vocabulary created by WP1, Ref 2).

³⁴ <https://musicbrainz.org/instruments>



4.1.2 Geotagging

With Geotagging - a more specific type of semantic tagging - users can add geo-coordinates to Europeana metadata records building on the existing spatial metadata of the record, or suggest a more precise location for the locations that are already present in the Europeana metadata records. For example, most of the existing locations do not identify a specific location, but only indicate a town, a village or a state. This information will enable alternative explorations of Europeana metadata records, for example, through map representations or location- based search.

4.1.3 Object linking

There are two types of annotations for linking Europeana metadata records (i.e. Object Linking). The first - and the one currently supported by the WADM, but not strictly 'semantic' of nature - is used to relate two or more Europeana metadata records together, without specifying the nature of the relationship.

However, many types of relationships may exist between metadata records (e.g. same object, same work or original-copy). Some of these relationships are symmetric, in the sense that both resources have the same relationship to each other. Others are asymmetric, meaning that there are different relationships between two metadata records, depending on the direction (i.e. parent-child).

To be able to support an explicit type of relationship between two metadata records within an annotation - hence creating a semantic, or meaningful relation - a second type of annotation for the linking of Europeana objects was defined and implemented. These use the RDF graph specification³⁵ which makes it possible to express the specific type and direction of the relationship. The model makes no restriction on the type of relation for the link, however, it is encouraged that these correspond to the EDM properties that can be used to relate objects together in the metadata (i.e. all extensions of dc:relation available in the Europeana Data Model³⁶).

4.1.4 Moderation (or validation through crowdsourcing)

Some of the annotations are created by users, which are biased by the user level of domain knowledge or interest, and other annotations are created by software, which do not always inspire a high level of confidence. In such cases, the end users may be interested in providing their feedback on existing annotations.

The main goal of the moderation functionality is to assess the quality of user annotations, providing a mechanism for the users to indicate which are the most useful and precise annotations. In addition, the moderation functionality allows the leveraging of crowdsourcing as a validation mechanism for (semi-) automatically generated enrichments.

³⁵ <https://www.w3.org/TR/trig/>

³⁶

http://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_Documentation//EDM_Definition_v5.2.7_042016.pdf

4.2 Semantic tagging in Tunepal and Historypin

4.2.1 4.2.1 Traditional music pilot

The Traditional Music Pilot (as previously reported in D2.10, Ref 3 & D2.7, Ref 4) was a find-by-playing interface to search for matches within archives of traditional music, as contributed by three data providers in the project. The pilot was conducted by modifying the existing TunePal software³⁷, already in use as a mobile app (and based on a crowdsourced shared 'tunebook'), to create a new web interface for the application with an integrated connection to the Europeana API. The feature is currently in production and used by TunePal's online users, powering approximately 1,000 queries per day. The software was initially deployed in July 2015 and continues to exist online in the production version of TunePal.org as of the end of 2016 (see Figure 4). Tunepal receives around 73,000 unique visitors a year and roughly 40% of these (or 29,200 visitors) see and interact with the Europeana recordings, as around 40% of the approximately 24,000 tunes available in Tunepal (9,600 tunes) have been found to have relevant Europeana-provided recordings associated with them.

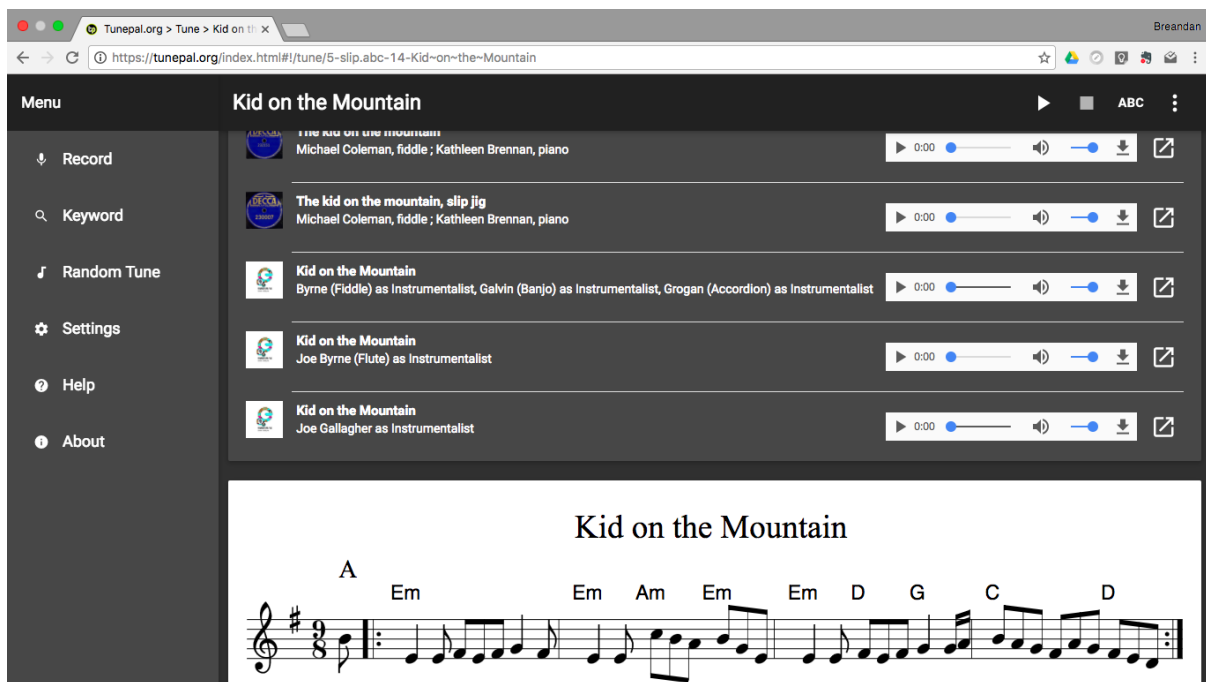


Figure 4: Screenshot of the Tunepal web application with relevant matches from Europeana displayed as suggested links

The Traditional Music Pilot targets amateur and learning traditional musicians. We initially targeted the 20,000 users of the TunePal application and the 50,000 regular users of TheSession.org traditional music web platform³⁸. Supervised user testing was carried out with approximately 50 users at Fleadh Cheoil na hÉireann³⁹, the largest traditional music festival in the world (see Figure 5).

³⁷ <http://tunepal.org/tunepal/index.php>

³⁸ <https://thesession.org/>

³⁹ <http://www.fleadhcheoil.ie/about/tradition>



Figure 5: Recruiting poster used to attract musicians to help evaluate the prototype

Our expectation was that users would query Europeana by directly playing fragments of tunes on an instrument, or by uploading recordings or transcriptions made earlier. For example, a musician might query a particular setting of an Irish reel by playing it on an instrument and then learn from the software the name of the tune and be able to play an authoritative recording of the tune sourced (via Europeana) from a music archive. The fact that a musician made this association (link) between a contemporary live recording or transcription and a particular archival track could then be stored within the Europeana Annotations API. The archive could then query Europeana for the links made to their own recordings, thus connecting the historic and contemporary practice for researchers.

When using Tunepal for the Europeana Sounds corpus, dance tunes were far easier to work with than songs, due to the design of the matching algorithm. This type of traditional music performed much better with the more consistent fast rhythms of dance music than the more melismatic character of traditional vocal songs.

Initial user testing of the find-by-playing interface created for this pilot, indicated a high perceived value and a novel and interesting experience for end users.



However, the data providers unfortunately withdrew access to the full length versions of archival recordings needed by the pilot, which delayed production implementation of the application in TunePal.org and eventually led to the cancellation of the planned integration with TheSession.org.

Data provider ITMA first withdrew from the pilot, and the recordings from data provider TAD turned out not to be suitable for the Tunepal algorithm. Data provider CCE then withdrew the ability to reference full-length audio tracks from the Europeana API, which led to the pilot being abandoned by Historypin.

Reaching beyond research audiences into the world of amateur cultural heritage practitioners is practical, but intellectual property and other access issues must be considered early and addressed comprehensively in the design of systems in order to be accepted by the practitioners.

Working with existing crowdsourcing or community-driven platforms brings the huge advantage of starting with existing audiences, and this pilot undoubtedly benefited from the existing sophistication of existing algorithms and software for Tunepal. However, working with existing platforms carries risks, particularly when the interests of the project and the platform diverge, when it can become difficult to maintain cooperation. In this case, frustrations with the (technical) access to tracks available via the Europeana API caused important platform providers to abandon the project, which affected delivery. Without the integrated management accountability structure that applies to a full project partner, there is really nothing other than mutual interest that can continue such collaboration with an external party.

There is an appetite for novel interfaces to audio archives and an opportunity to make such archives very relevant to contemporary practice, providing that the access issues can be dealt with. While the Europeana API remains deployed on the production version of TunePal.org, the associations being discovered by users between contemporary and archive performances of these tunes are not being tracked or recorded. As a result, we do not anticipate significant contributions to KPI-9 (Ref 1) by means of this channel.

4.2.2 Geo-Location crowdsourcing

With the understanding that users sometimes have more accurate understanding of key metadata elements than the producers or curators of archives, we intended to encourage users of the Historypin.org⁴⁰ crowdsourcing platform to contribute suggestions of more accurate places, time periods or keywords to objects sourced within the Sounds project. This pilot functionality was completed and tested in April 2016 (see Figure 6) and was used to support a crowdsourcing campaign to source geotags for 258 Dutch WWII radio programmes for two weeks in June.

⁴⁰ <http://www.historypin.org>

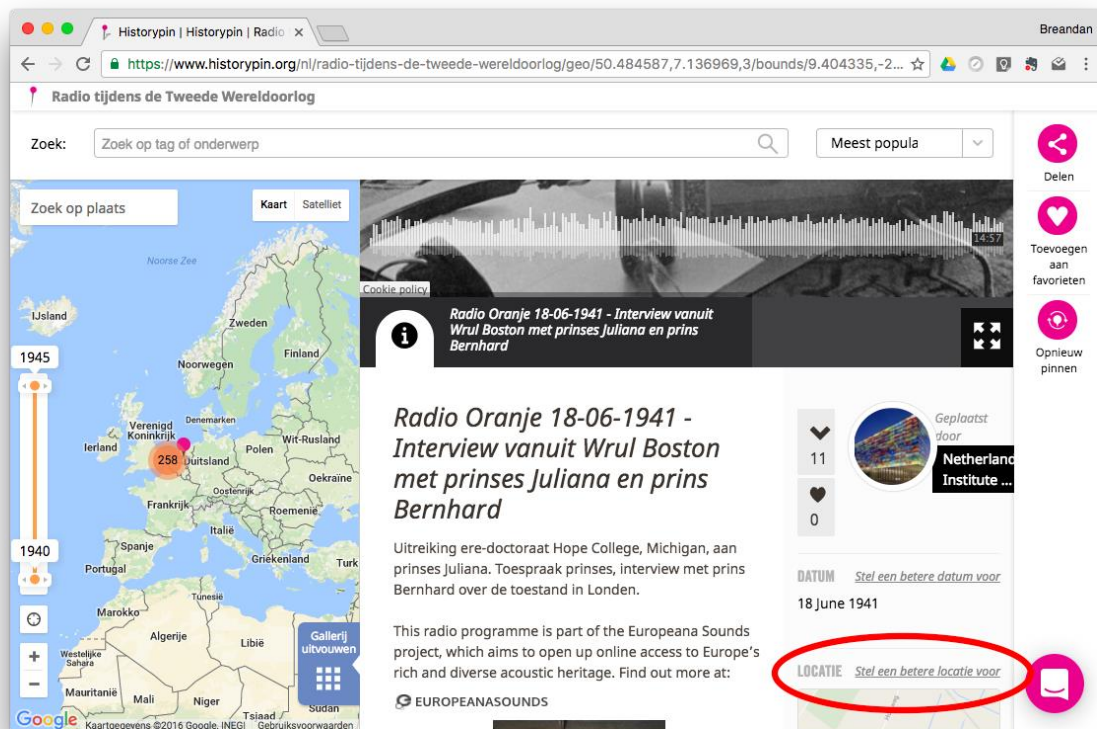


Figure 6: Screenshot of the Historypin platform with the functionality (in Dutch) to start suggestion an improvement of the location data through crowdsourcing

Users with knowledge of the Dutch language were recruited by NISV and Historypin over a brief period of time to contribute these more accurate metadata annotations to a WWII radio collection of NISV on the Historypin platform. However, there were no significant contributions of geotags from general members of the public in this time. The publicity of the campaign did lead to an increase in user traffic over baseline (>10x average) and to a significant user engagement metric of 05:19 average visit duration, which would be envied by many platforms (based on Historypin analytics). This additional amount of user attention unfortunately did not translate into a significant amount of crowdsourced contribution of metadata.

As the location of the radio broadcast was not always spoken near the start of the broadcast, a contributor would have to listen to a substantial amount of the audio recording in order to be able to contribute meaningfully, which may not have been a good match for a “casual” social media campaign, which often benefits from short, easily-understood tasks.

Historypin’s Dutch-language interface (see Figure 7) was appreciated by users, in user testing, as representing ‘their’ culture. This seems a useful learning point for future crowdsourcing campaigns.

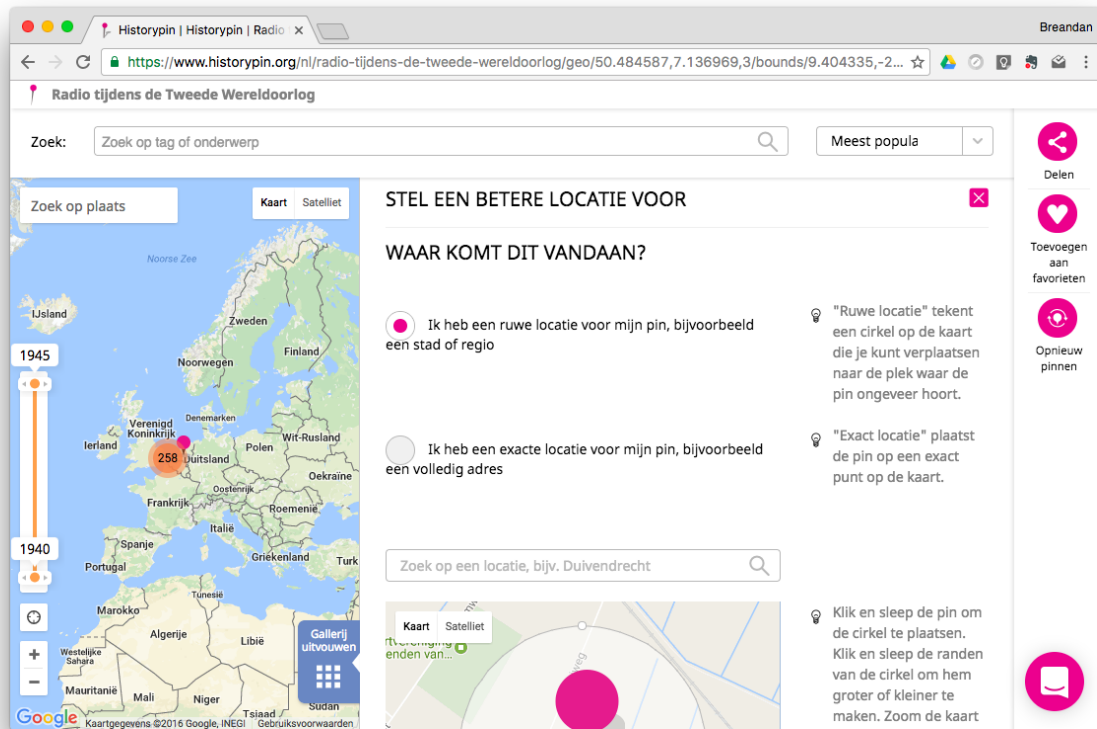


Figure 7: The geotagging procedure in Dutch

The actual geotagging functionality on Historypin (see Figure 8) performed well in user testing. Based on this there do not seem to have been significant issues with the usability of the interface. This suggests that low involvement was due to a combination of content- and campaign-related factors.

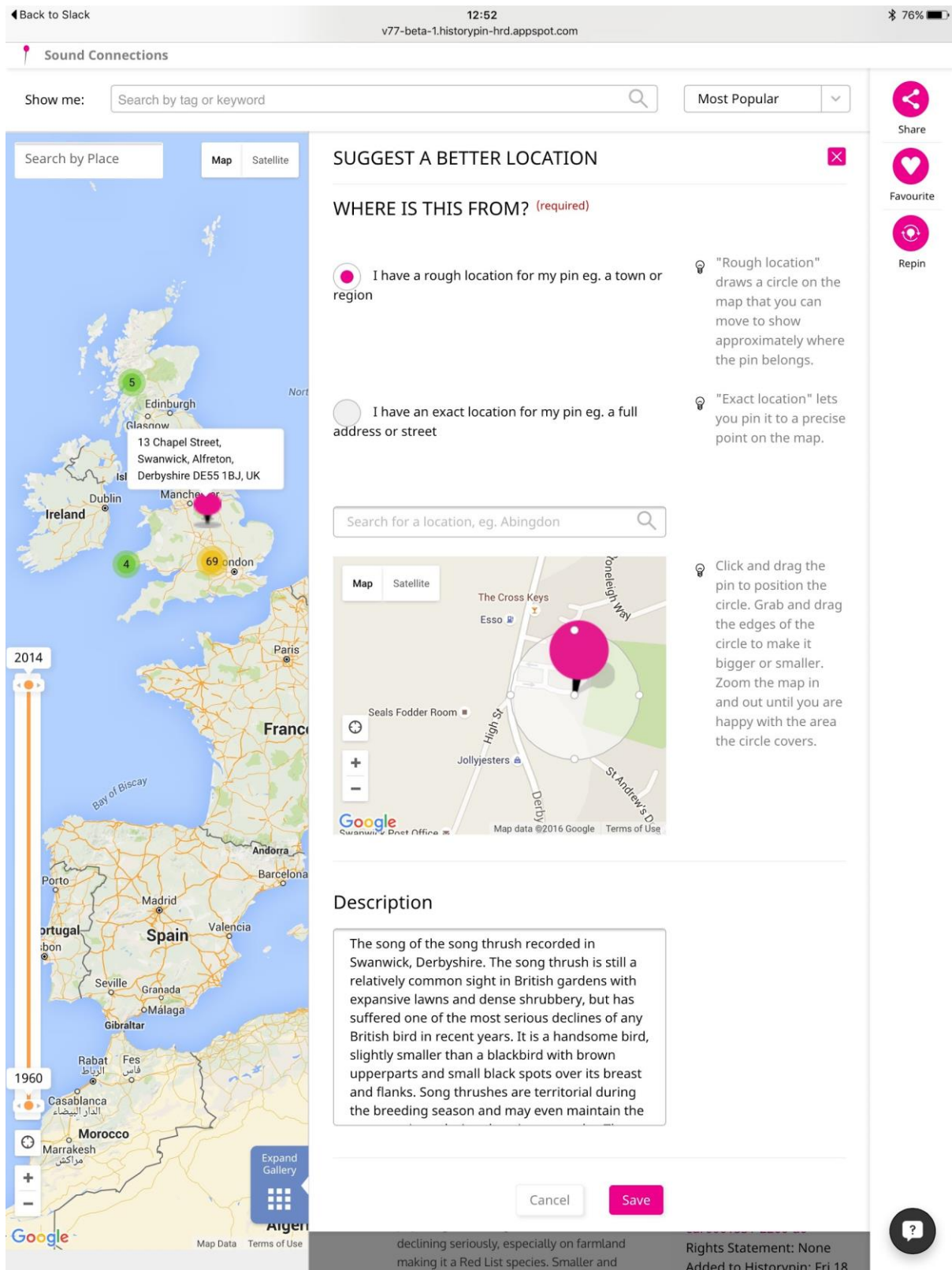


Figure 8: The geotagging functionality expanded

The limitation of both subject matter (old radio broadcasts related to World War II) and linguistic context (Dutch-speakers only) combined with a relatively short (but intense) social media campaign of



two weeks may have combined to keep the campaign from reaching a ‘tipping point’ of contributory behaviour.

The geotagging feature for the Historypin platform has been written in a very generalised way, and as such will remain a production feature on Historypin Collections into the future, including support for any future Europeana-related campaigns. Now that the first campaign has been completed, the tool remains available for any other Europeana Sounds partner to identify further collections of sounds from within the project that might be enriched by adding more precise locations. Sounds collections that have been added to Europeana as part of other projects might also be considered for enrichment.

Historypin does however - for the Europeana Sounds project - not anticipate substantial progress toward the meeting of KPI-9 (Ref 1), given the results mentioned above.

4.3 Semantic tagging in Pundit

Pundit is a web application that allows user to create semantic annotations on web resources. It was selected to integrate semantic tagging through crowdsourcing in the Europeana Sounds project. Pundit is developed since 2011 by Net7 Srl, an IT company based in Pisa.

One overall goal for the Europeana Sounds project WP2 was to put in place a web software environment to enable semantic enrichment of Europeana Sounds content:

“Apply semantic web technologies to enable enrichment of the Europeana Sounds collections. This will increase quality of the metadata and user satisfaction in terms of content discovery.”

(Ref 1)

The enrichment was to be done using domain specific controlled vocabularies, therefore providing the users a tool to create semantic triples that connect an entity of the Europeana Sounds dataset to a linked open data entity (LOD) (i.e. a term describing an instrument from a controlled vocabulary like MIMO). Another essential requirement was to allow users to review annotations from other users and the results of automated semantic enrichment.

4.3.1 Features

The Pundit software available at the beginning of Europeana Sounds was missing some essential features we designed and developed during the project.

4.3.2 Sound object annotation

Pundit is a web annotation tool and therefore its native goal is to allow users to annotate passages of text in web resources. In the context of the Europeana Sounds project it was necessary to implement a system to build annotations that target the Sound object (audio recordings, digitised scores, etc.) itself, identified by its unique identifier on Europeana.

The project couldn't use the native Pundit targeting system (that relies on XPointer⁴¹, a language for pointing to sections and elements of HTML documents) therefore we needed to implement a new feature to meet this requirement.

We extended the native Pundit targeting system to allow using CHO entities⁴² on Europeana (entity identified by a unique URI) as the target of the annotation. The target of the annotation is the resource itself and related annotations are displayed in any page where the resource is included: This was a welcome side effect of the implementation because the sound object can be displayed on Europeana and partly in the Europeana Music Collection, the Data provider's website or WITH platform and the same annotations will always be correctly retrieved for display and/or manipulation.

In order to achieve this effect, Pundit was modified to look for a snippet of code like this in the HTML⁴³:

```
<span about="http://your-resource-unique-identifier/123456" class="pnd-resource"></span>
```

When Pundit finds this snippet in the code it knows that an entity with the URI defined by the "about" attribute (i.e. "http://your-resource-unique-identifier/123456") is displayed in the page and it shows a button to activate the Pundit Annotator Pro.

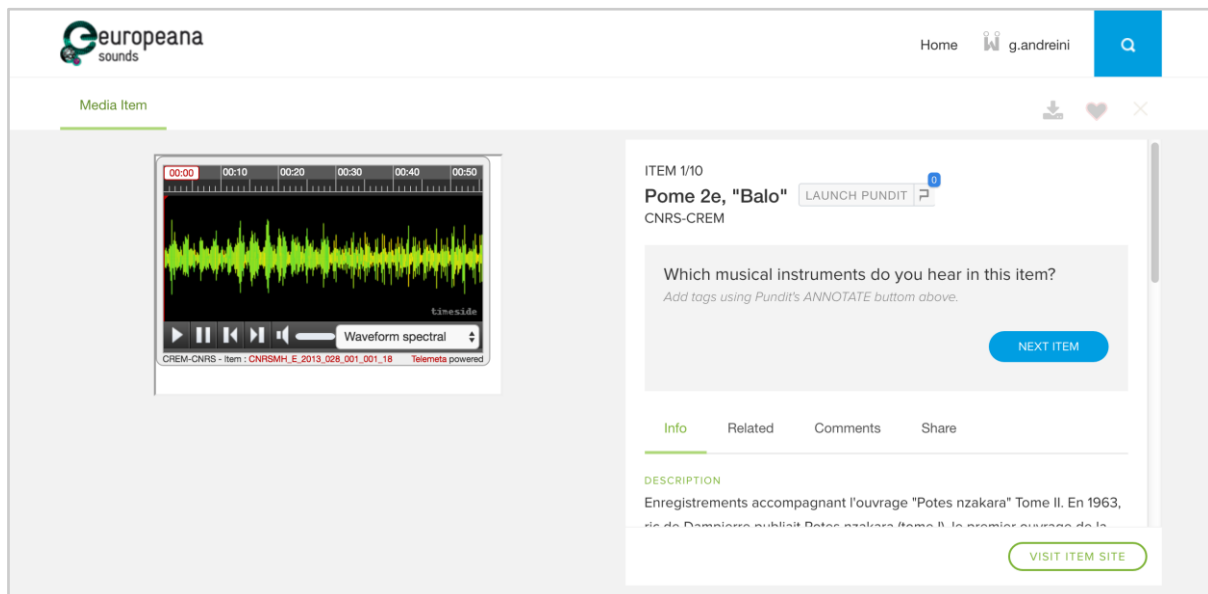


Figure 9: The “Launch Pundit” button that allows to launch the Annotator Pro. The button is automatically inserted in the page when Pundit finds the snippet of code

Once the “Launch Pundit” button is clicked the Annotator Pro is added to the page: When a user clicks on the “Annotate” button a drop down will be displayed allowing the user to create a semantic enrichment with the CHO entity as the subject (see Figure 9 & 10).

⁴¹ <https://en.wikipedia.org/wiki/XPointer>

⁴² See EDM documentation available on: <http://pro.europeana.eu/page/edm-documentation>

⁴³ The procedure for integrating Pundit in this way is documented here: <http://net7.github.io/pundit2/resource.html>

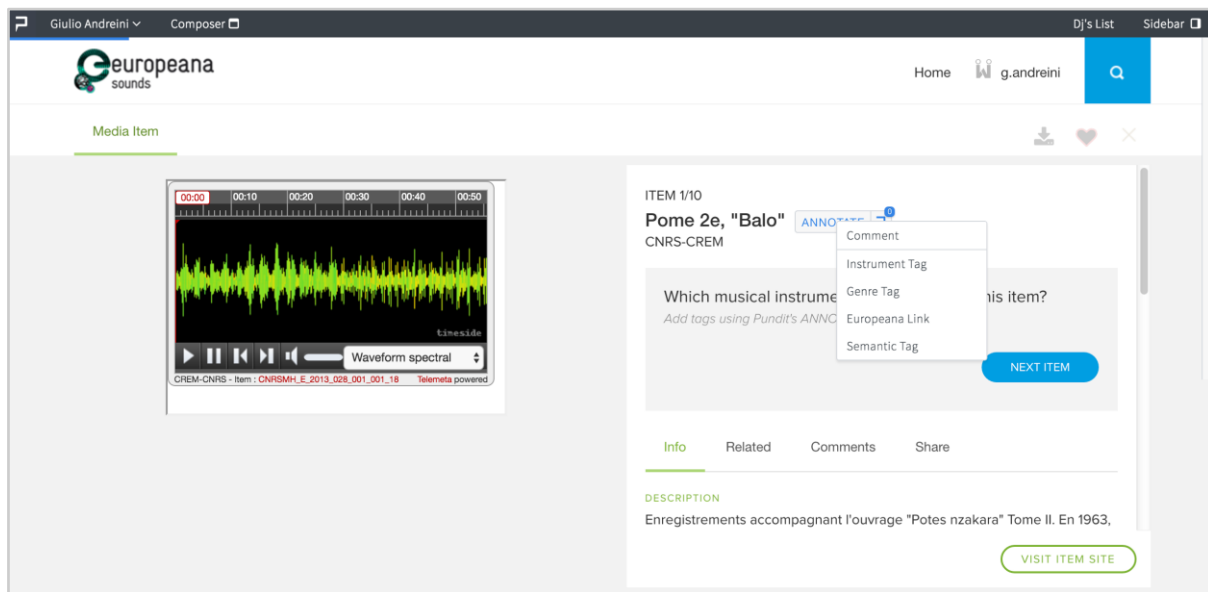


Figure 10: Pundit Annotator Pro and the drop down that allows the annotation

4.3.3 MIMO vocabulary and thesaurus

Pundit originally allowed users to create annotations using entities of DBpedia but did not provide any connection to other controlled vocabularies for the sounds domain.

We implemented the connection to MIMO to allow users to select from a wide range of controlled terms for musical instruments and state that a certain instrument is played in the recording as a semantic enrichment.

4.3.4 Moderate previously created annotations

One of the additional objectives of the project was to allow users to revise, correct and enhance previously created annotations (both results of crowdsourcing and automatic semantic enrichment, see Figure 11). Pundit was missing this feature and did not allow the creation of annotations targeting existing annotations. To respond to this project requirement we designed and implemented a feature that allows users to 'moderate' previously created annotations. The types of reactions are:

- Reply: users can reply to an existing annotation with a text comment.
- Like/Dislike: users state they like or not dislike an existing annotation using a pattern widely used in social networks.
- Endorse: the user endorse the annotation if the existing annotation is correct.
- Report: user report the annotation if the existing annotation is not correct.

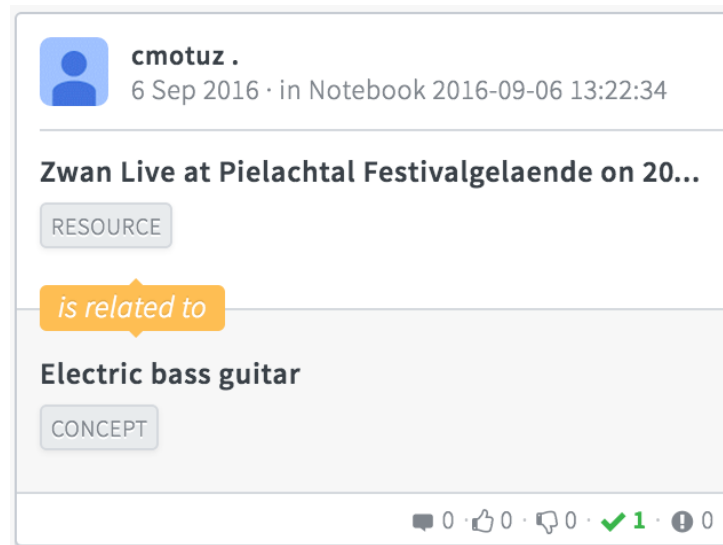


Figure 11: Pundit annotation with “respond” buttons: reply, like, dislike, endorse, report. In this example the annotation was endorsed by a user.

The data model of this new type of annotations was discussed with the other technical partners of the project and designed to be compliant with the Web Annotation Data Model recommendation by the W3C Web Annotation Working Group.⁴⁴

4.3.5 Multilingual support

Data providers and users who are involved in the crowdsourcing campaigns live in different countries and speak different languages. The default language used by Pundit is English and user testing sessions revealed a critical issue concerning the language of the object of the semantic enrichment: vocabulary entities (such as MIMO instruments) rely on expertise and domain knowledge and many users do not know the exact English term for each entity. This makes very hard for them to select the right object and they are forced to translate each term on Google Translate or similar translation services. This problem applies both to the creation of an annotation and the moderation of previously created annotations on the right sidebar.

We decided to overcome this Pundit limitation and implement a new feature to show the label for the MIMO entities, using the language selected by the user for the browser: in this way Pundit will always show the entities label in a language comprehensible to the user. This feature allowed Pundit and the project to benefit from the multilingual nature of MIMO. From a technical standpoint Pundit retrieves the label for each entity in all the languages available and it shows the user the correct label based on the browser language settings (see Figure 12).

⁴⁴ <http://labs.europeana.eu/api/annotations-scenarios>

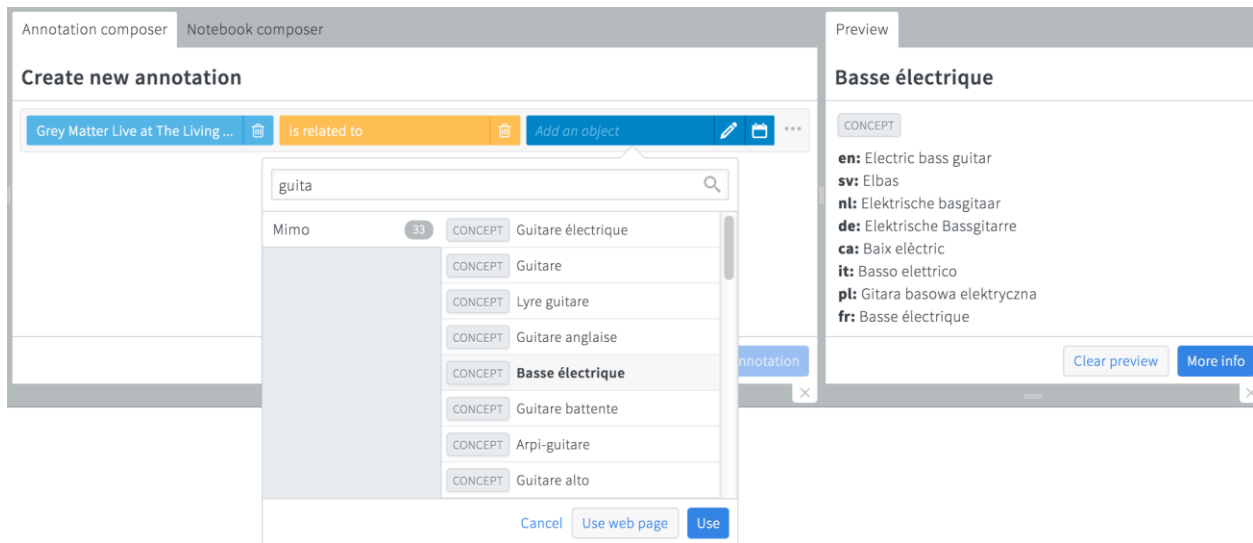


Figure 12: Multilingual support in Pundit: the name of the MIMO instrument is displayed in the user’s selected language (i.e. French) and in the Preview panel it is possible to view all the available translations.

4.3.6 Workflow improvements

The default workflow in Pundit when a user wants to create an annotation is to click on the “Annotate” button and then click on “Use as Subject”; then the user is asked to select a Predicate and an Object. This is quite straightforward on a technical level, but can sound too technical for users who don’t have a background in semantic web technology. Hence, we decided to simplify this workflow and improve the overall user experience: we changed the drop down content that opens up when a user clicks on the “Annotate” button, showing more understandable instructions describing the action the user is about to perform (i.e. “Instrument Tag”, see Figure 13). Moreover when the user selects one of these actions the semantic triple is precompiled with the Subject and the Predicate, and the user has to choose only the Object. In this way we guarantee a faster and more user-friendly workflow.

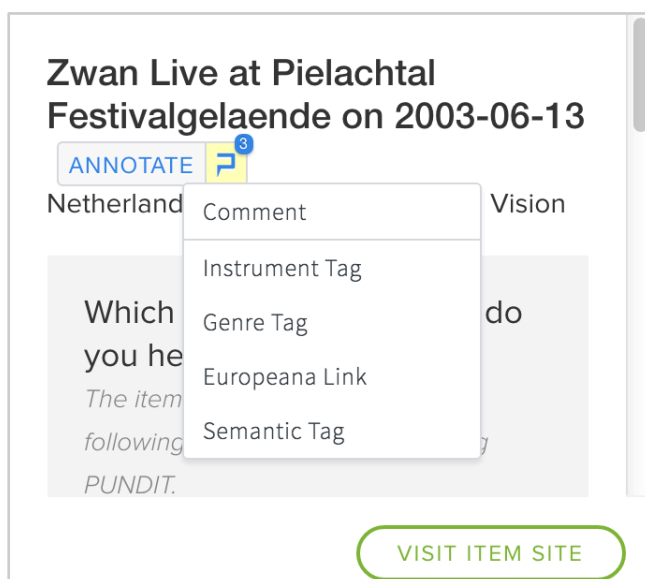


Figure 13: The drop down with the custom instructions for Europeana Sounds: when the users selects “Instrument Tag”, “Genre Tag” or “Europeana Link” the Predicate of the annotation is automatically set.

4.3.7 Pundit integrations

Users can annotate sound resources with Pundit in three different ways:

1. WITH: Pundit is embedded in the WITH platform and when users navigate to a resource page a “Launch Pundit” button shows up allowing to open Pundit and annotate the sound resource (see Figure 14). This is the main application of Pundit that was used for the crowdsourcing campaigns.

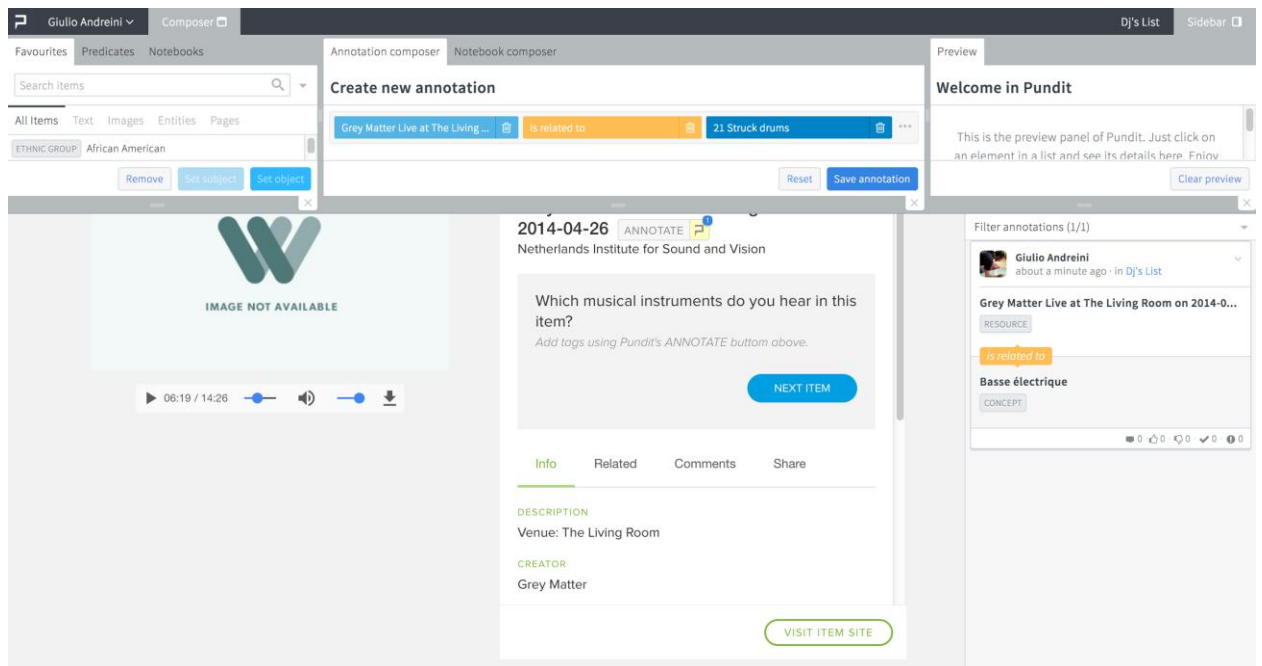


Figure 14: Pundit in the WITH platform.

2. Europeana Sounds Chrome Extension (see Figure 15): we created a Google Chrome Extension that can be used to annotate resources directly on the Europeana Portal or Europeana Music Collection. Each time a user navigates to a resource page on Europeana the extension will add the “Launch Pundit” button to the page to open Pundit and annotate the resource. The Extension is available on the Google Web Store⁴⁵.

⁴⁵ <https://chrome.google.com/webstore/detail/pundit-annotator-pro-for/gnadajfkeolhfngkhkffdobmbfekghico?hl=en&authuser=1>

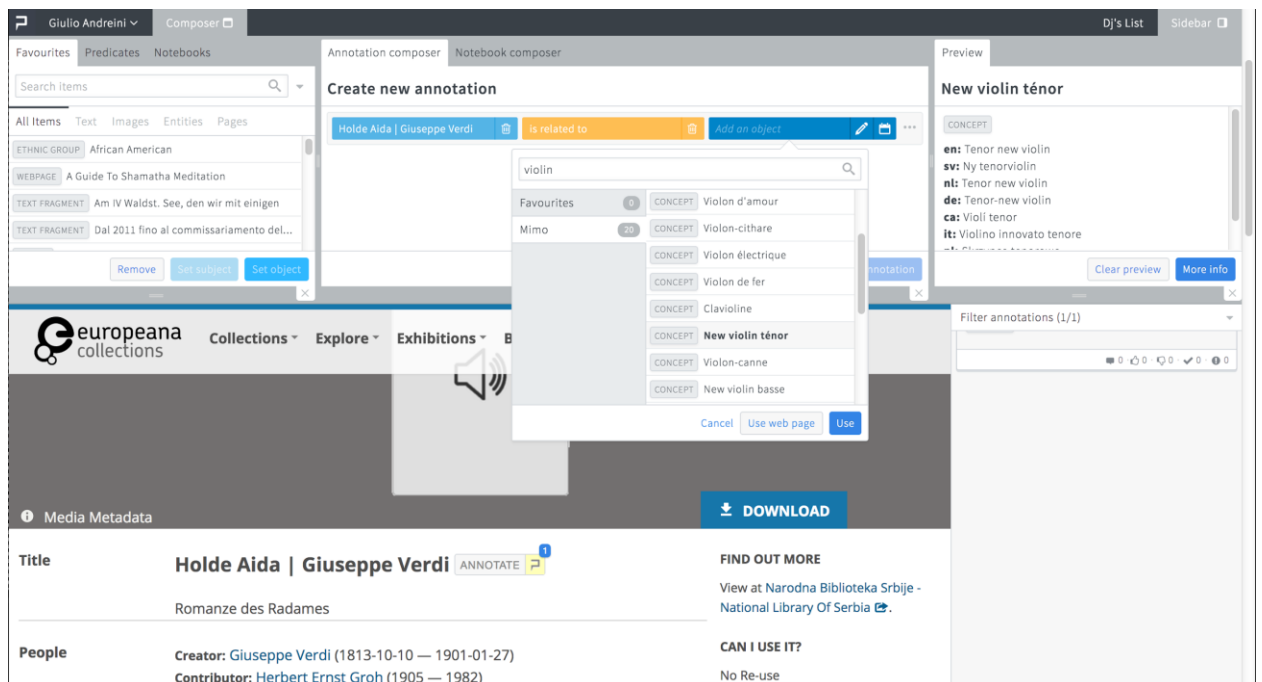


Figure 15: The Pundit for Europeana Chrome Extension launched on the Europeana website: it allows the annotation of any type of resource (not only sounds).

3. CNRS/Telemeta integration: this is the third type of Pundit integration and it's currently still under development. Pundit will be embedded on the Telemeta platform⁴⁶ to allow users to annotate entities directly on the CNRS archive: when a user will navigate to a resource page the "Launch Pundit" button will show up. The Pundit-Telemeta integration will be completed in March 2017.

4.3.8 Data model

The whole Pundit data model was redesigned to be more compliant with the W3C Web Annotation Working Group Guidelines⁴⁷: this required a deep refactoring of the whole structure. The triple store used by Pundit was also transferred to the new graph database Graph DB™ Standard Edition⁴⁸.

The W3C guidelines did not fully cover all the use cases of the Europeana Sounds project and we needed to extend the model. These are some application scenarios that required some design and improvement of the data model:

- Annotate a CHO with a semantic tag: An end-user wishes to annotate a CHO using a tag corresponding to a resource from a controlled vocabulary.
- Annotate a CHO with a language-specific tag: An end-user wishes to tag a CHO using a text literal, expressed in a specific language.

⁴⁶ <https://github.com/Parisson/Telemeta>

⁴⁷ <https://www.w3.org/TR/2014/WD-annotation-model-20141211/>

⁴⁸ <http://ontotext.com/products/graphdb/>



- Link a CHO to another using a qualified link: A user wishes to relate a CHO to another explicitly stating the nature of relation between them.
- Link a CHO to an external resource: A user wishes to relate a CHO to another resource which is not hold by Europeana, stating the nature of relation between them.

4.3.9 Discussion

The most relevant critical point which emerged during the project concerns the usability of the Pundit user interface for providers involved in the project. Pundit Annotator Pro is by origin a tool targeted at users with a background in semantic web technology or users who are taught how to use it. Building semantic triples in such an explicit manner is a relevant feature for expert users but it overcomplicates the experience of other users who are targeted in the enrichment through crowdsourcing.

A further development post-project that could improve the experience would be to simplify the semantic annotation creation workflow further and make it more straightforward and similar to the way users create comments and highlights in Pundit Annotator. This could be achieved not showing the triple composer but letting the user select a predicate from a form select and then the object in an autocomplete field or even skipping the predicate selection and let the user just choose the Object. This improvement would require a user experience design work including ideation workshops, wire framing and user testing.

Pundit will continue to be maintained by Net7 post-project as part of the Muruca⁴⁹ framework the company offers as a B2B service to digital libraries and research groups. The Pundit Chrome extension for Europeana will continue to be available to users and data providers for potential future crowdsourcing campaigns. To ensure a stable connection between Pundit and the Europeana Core Infrastructure, Net7 and EF will communicate important updates to their respective services to each other.

4.4 Semantic tagging in WITH

4.4.1 Crowdsourcing space

For the benefit of the Europeana Sounds project, a dedicated space was created using the WITH platform for the purposes of semantic tagging of music recordings. The tags are concerning musical instruments recognised inside the songs with values derived from MIMO. The Sounds crowdsourcing space contains musical records in the form of digital sound recordings along with their metadata, organised in thematic collections like “[Music from BNF](#)” or “[Viennese Songs with Hermann Leopoldi](#)”. These collections were imported from the Europeana database - using the Europeana API - into WITH for crowdsourcing purposes. The semantic tagging is achieved via a crowdsourcing campaign, namely a period in which the users are encouraged to visit the space and start adding their tags or validate the existing, working together to achieve the goal of the campaign. To motivate users and provide an overview of the progress with the current campaign, the goal is displayed in the homepage of the Sounds crowdsourcing space, along with the number of tags already contributed by the crowd and the

⁴⁹ <http://www.muruca.org>



percentage of completeness. The whole homepage of the Sounds crowdsourcing space is shown in Figure 16.

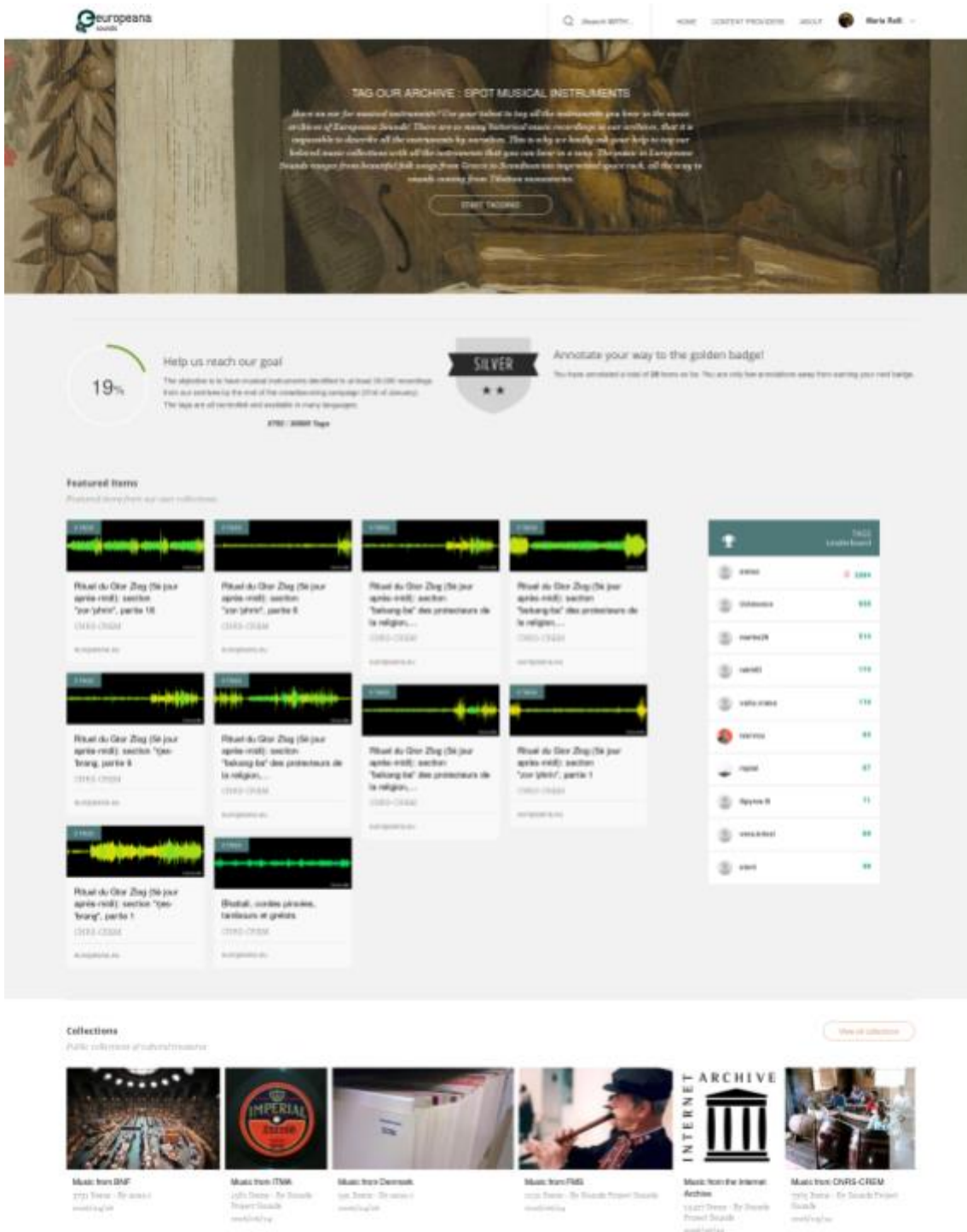


Figure 16: The homepage of the instrument crowdsourcing space

4.4.2 Tagging Process

The tagging process was designed to be simple, engaging and immersive for the users who want to contribute. There are several improvements from the previously reported (Ref 3, 4 & 5) crowdsourcing application developed by NTUA. The major one is the removal of Pundit as an intermediate annotating tool, which minimised the amount of clicks required by the user to add tags and solved the usability issue with the two logins needed, both on Pundit and WITH. In addition, the record player was improved, supporting the different streaming solutions used by the data providers. More specifically, apart from the typical .mp3 audio files played with HTML5, many records were displayed via the use of an iframe, such as SoundCloud streams and the Österreichische Mediathek iframe as shown in Figure 17.



KATALOGZETTEL	INFORMATION
Titel	Die schöne Adrienne hat eine Hochantenne (Tschintarata-radio)
Spieldauer	3:06
UrheberInnen	Leopoldi, Hermann [Komponist/in], Waldau, Theodor [Text]
Mitwirkende	Kuttner, Max [Gesang]
Datum	1925
Schlagworte	Musik ; U-Musik; Unterhaltung; Kabarett; Radio; Internet; Publierte und vervielfältigte Aufnahme; Österreich
Archivformat	Schallplatte, Schellack - 25 cm, 78/min
Typ	audio

Labels can. Copyright: Österreichische Mediathek

Figure 17: Example of the Österreichische Mediathek iframe as displayed on the crowdsourcing area

To start tagging, the user needs to login into WITH either via Google, Facebook or their private email address. They can either start tagging random music recordings from the collections by pressing the “START TAGGING” button, or choose a specific record from a collection they are interested in and add their tags to that record. The tagging page is shown in Figure 18. This page plays the sound of the music recording and displays its metadata under the “INFO” tab, giving further information about the record and thus aiding the tagging process. The user can either add new tags to the record or validate the existing tags. The tagging process is very simple, as a user can add new tags by typing the name of the instrument she/he recognises inside the input box. While typing, a list of suggested instruments will be displayed - through auto-completion - containing the matching MIMO thesaurus term and the category of the instrument (e.g. piano - pianofortes), from which the user will choose the appropriate one and add it to the object, as is shown in Figure 19. The already existing instrument tags are also displayed on

the right side of the page, including link to the MIMO description of the instrument, as shown in Figure 20. For these pre-existing tags, the user is given the ability to approve or disapprove the semantic tags from other users by pressing the dedicated buttons on the right, or even delete the tag if it is his/her own contribution.



Figure 18: The tagging page. On the left: the song player. On the right: the existing tags and the input box for the new tags

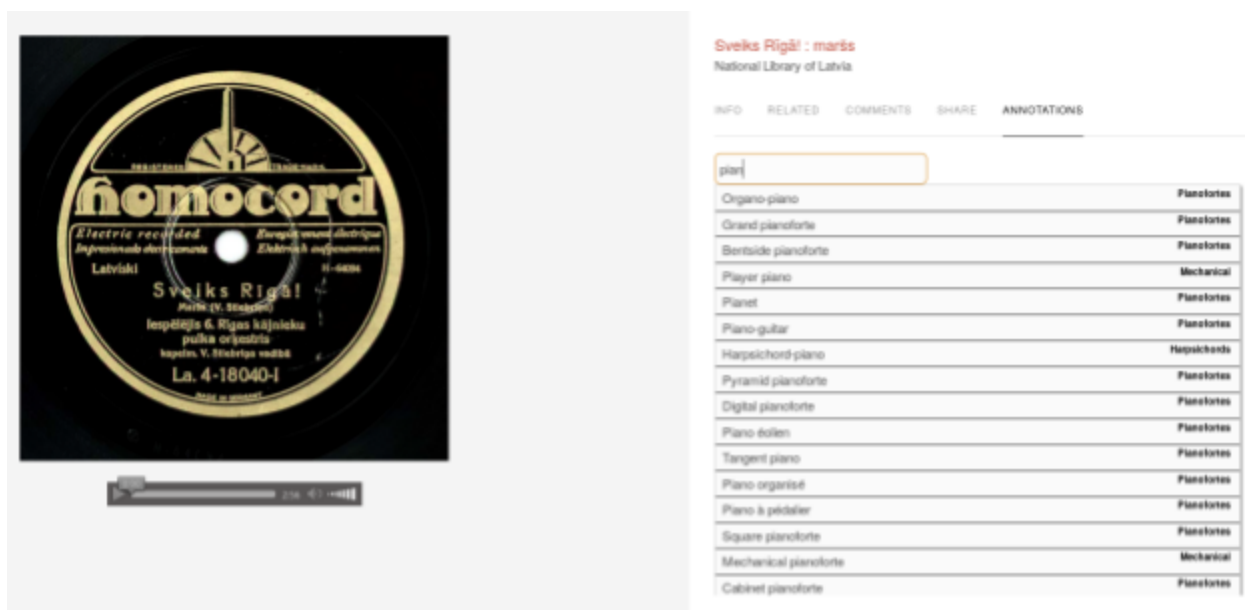


Figure 19: The process of adding new instrument tags from the MIMO thesaurus

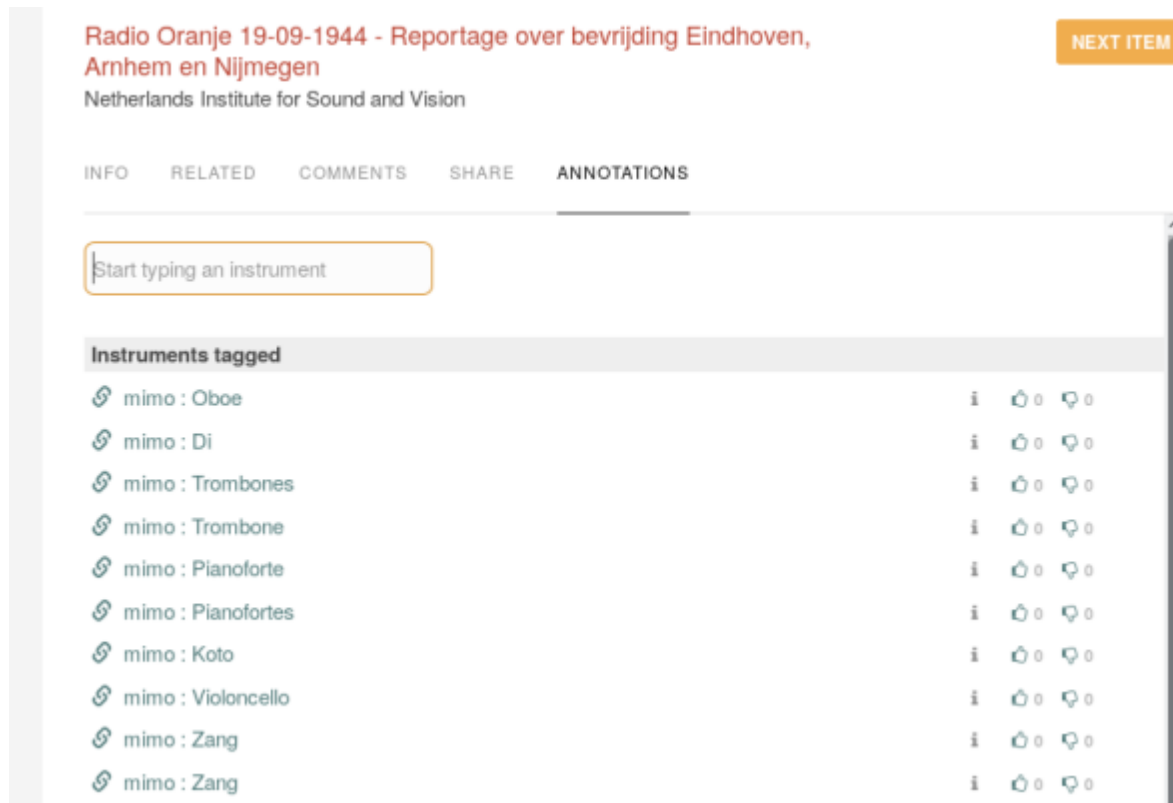


Figure 20: The tagging area containing the existing instrument tags and the input for the new ones

4.4.3 Game mechanics

In order to inspire the users to contribute their time and efforts towards our crowdsourcing endeavour, we made it more immersive and compelling through the integration of some key game mechanics. The contributors are rewarded with a badge - e.g. bronze, silver and gold badge - depending on the number of tags created or validated by them. In addition, there is a leader board showing the most active “taggers” namely the users with the most contributed tags and validations. These gaming features are shown in Figure 21 and Figure 22.

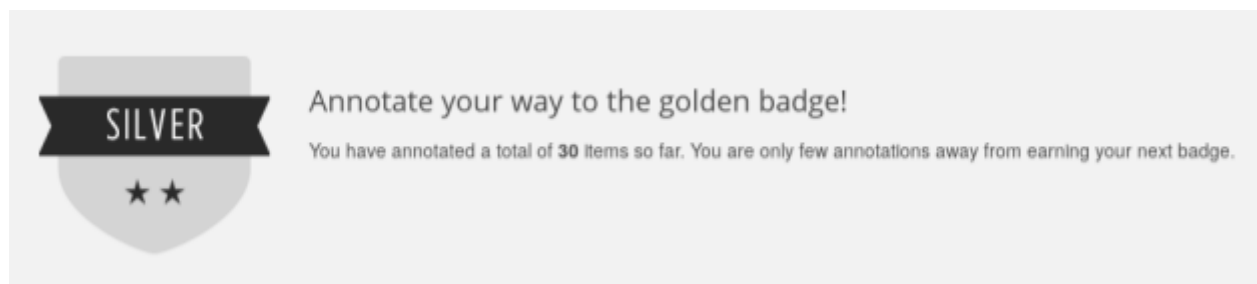


Figure 21: Badge reward

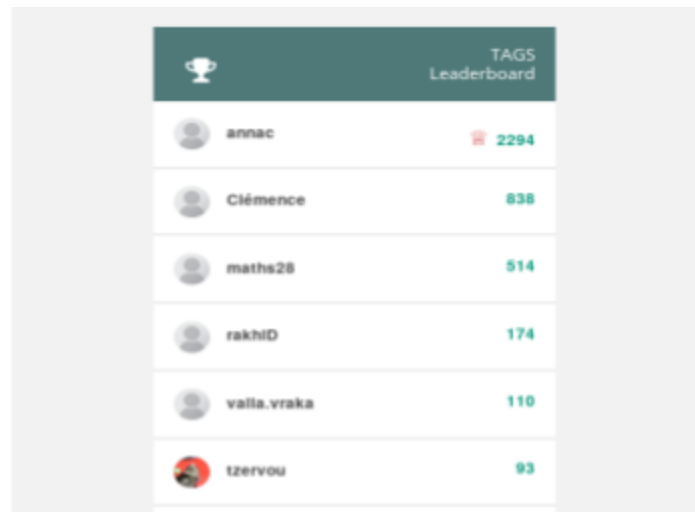


Figure 22: Leader board with the most active ‘taggers’

4.4.4 Discussion

After Europeana Sounds the WITH platform will be sustained by NTUA through national projects and B2B services offered in the museum domain. Crowdsourcing remains as a prominent topic on the roadmap for the platform. The space for Europeana Sounds on WITH will remain available to users and data providers post-project and future crowdsourcing campaigns are encouraged.

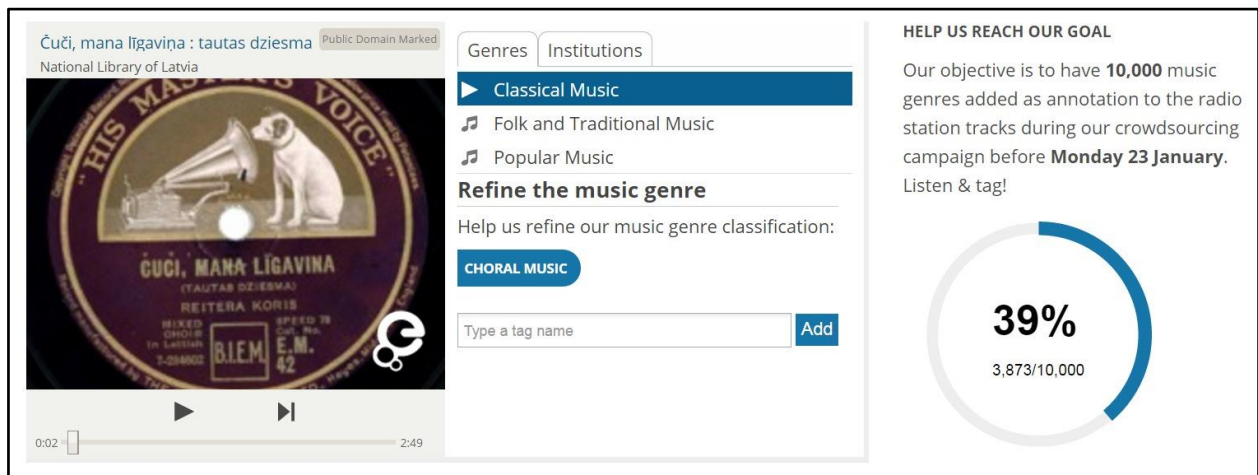
4.5 Semantic tagging in Europeana Radio

Europeana Radio⁵⁰ (see Figure 23) was developed as an attempt to make the discovery and experience of the music recordings in the Music Collections more playful and ‘frictionless’. It consists of two parts, first one is a radio station⁵¹ which retrieves tracks from the Europeana API for the radio stations, and the second is a radio player⁵² which is the user interface for a web browser to play tracks from the radio stations. The radio player can be embedded into any website to allow users to listen to streamable audio from the Europeana Music Collections. Europeana Radio was first released on October 20, 2016 with about 3,000 available music recordings.

⁵⁰ <http://www.europeana.eu/portal/en/radio.html>

⁵¹ <https://github.com/europeana/radio-station>

⁵² <https://github.com/europeana/radio-player>



The screenshot shows the Europeana Radio interface. On the left, a music player displays the title "Čučī, mana līgaviņa : tautas dziesma" from the National Library of Latvia. The album art features a gramophone and a dog. The player shows a progress bar from 0:02 to 2:49. On the right, there are tabs for "Genres" and "Institutions". Under "Genres", "Classical Music" is selected, with other options like "Folk and Traditional Music" and "Popular Music". A "Refine the music genre" section includes a text input field with "CHORAL MUSIC" and an "Add" button. To the right of the player, a "HELP US REACH OUR GOAL" section states the objective is to have 10,000 music genres added, with a progress indicator showing 39% (3,873/10,000) completed.

Figure 23: Europeana Radio

In an update on January 10, 2017 the number of music recordings available was increased to almost 200,000 and, powered by the Annotations API, users were enabled to add genre classifications to the music recordings. A follow-up release on January 25, 2017 increased the number of possible music genres classifications users could add to the songs playing in Europeana Radio and also added another 10,000 music recordings to the stations for users to tag, along with a percentage indicator for the crowdsourcing campaign. This allowed for real-time monitoring of the progress of the on-going campaign, as well as giving the contributors as sense of progress based on their activity (see Figure 23).

Europeana Radio acts as a direct client of the Annotations API and the genre classifications the users tag the music recordings with are saved as Wikidata URIs. Once saved in Europeana Radio, they appear on the corresponding item page on the Europeana Music Collection (see Figure 24).

The utilisation of Europeana Radio as a crowdsourcing application was a deviation from the original development plan as reported in D2.7 (Ref 4) by EF. The initial plan was to add genre tagging functionality to the Europeana Music Collection by October 2016. However, this implementation relied on the MyEuropeana⁵³ functionality to be ready for production. Since this external dependency could not be secured, EF developed the genre tagging functionality in Europeana Radio as an alternative. The later delivery of this functionality in January 2017 unfortunately has had a negative impact on KPI-9 (Ref 1).

⁵³ <http://labs.europeana.eu/api/myeuropeana>

The screenshot shows the Europeana Music Collection interface. At the top, there is a navigation bar with 'OUR SITES' and 'LANGUAGE' options. Below the navigation bar, there is a search bar and a 'Return to Home / Item' link. The main content area features a large image of a gramophone record label for 'CUCI, MANA LĪGAVIŅA (TAUTAS DZIESMA) REITERA KORIS'. Below the image, there is a 'Media Metadata' section with a play button icon. To the right of the metadata, there is a 'DOWNLOAD' button and a 'FEEDBACK' button. The metadata section is organized into several rows:

- Title:** Čuči, mana līgaviņa : tautas dziesma
- People:** Contributor: Reitera koris
- Classifications:** Subject: Folk songs, Latvian, Music; Has Type: Sound; Medium: audio disc
- Properties:** Format: Shellac record
- Time:** Date: [19--]
- User Annotations:** Relations: <http://www.wikidata.org/entity/Q1076513>

On the right side of the metadata section, there are several utility links: 'FIND OUT MORE' (View at National Library Of Latvia), 'CAN I USE IT?' (Free Re-use, Public Domain Marked), and 'SHARE' (with social media icons for Facebook, Twitter, Pinterest, Google+, and Tumblr).

Figure 24: User annotations displayed on the Europeana Music Collection (see for example: http://www.europeana.eu/portal/en/record/2059201/data_sounds_85.html)

Europeana Radio will continue to be available to end-users after the project. This includes the possibility to crowdsource genre classifications for streamable audio that is available in the Europeana Music Collection. The service will not see major updates until the end of the Europeana DSI-2 project, since this on-going project already populates the entire roadmap for the Product Development team at EF.

4.6 Semantic tagging in the crowdsourcing campaigns

The Europeana Sounds crowdsourcing campaigns - as reported in D2.7 (Ref 4) - functioned as a vehicle to coordinate:

“dedicated crowdsourcing tasks, organised centrally (meaning involving as many data providers with the same information needs and disseminating the crowdsourcing activities to the public through the main Europeana Sounds communication channels as a unified project)”.(Ref 4)

The aim of these crowdsourcing campaigns was to focus and enhance the impact of the crowdsourcing applications, as (further) developed by the project. The crowdsourcing campaigns were organised in collaboration with the Data providers (who curated collections suitable for specific crowdsourcing task and promoted the campaigns) and in collaboration with WP6 (who helped develop marketing and communication activities for the campaigns). Coherence in the communication through social media and the various channels of the project was achieved through the use of the hashtag #crowdsourcingmonth.

4.6.1 Semantic tagging in the first crowdsourcing campaign

The first crowdsourcing campaign - held in June 2016 and previously reported on in D2.7 (Ref 4) - encompassed two type semantic tagging. It started with promoting using the Historypin geotagging functionality (see Section 4.2.2) for collecting better location data for a collection of Dutch WWII radio broadcasts from NISV54 (see Figure 25).

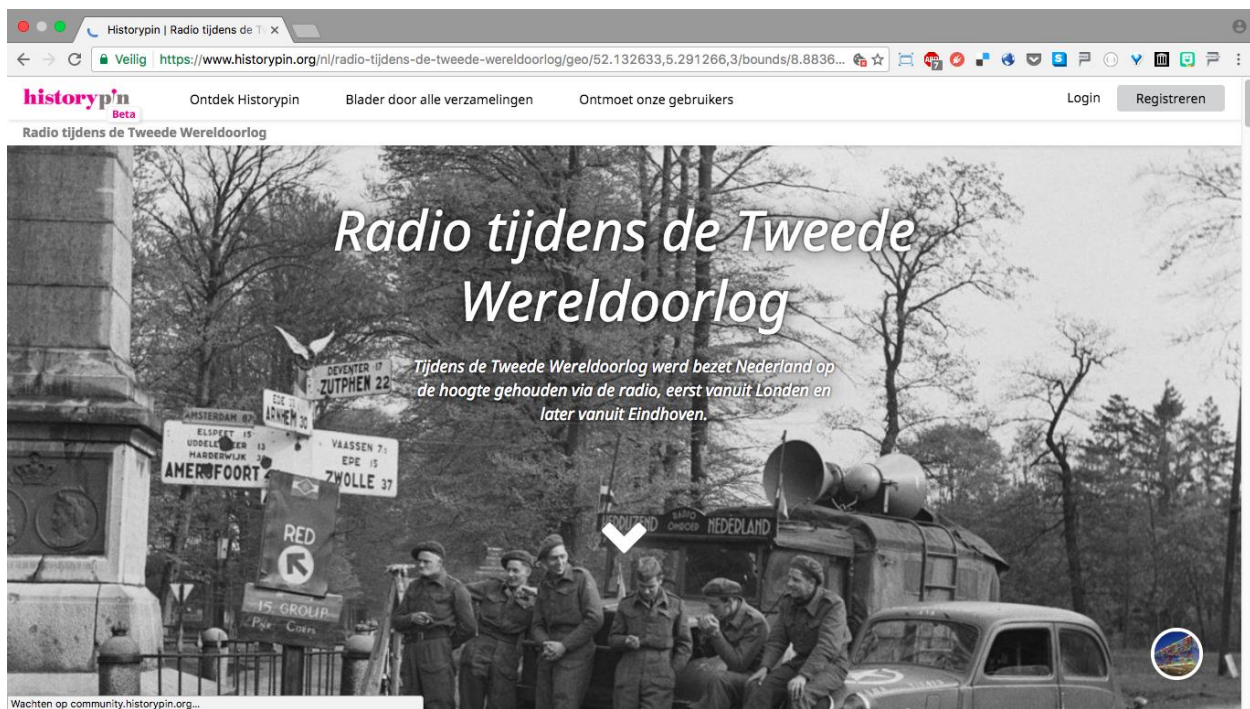


Figure 25: Landing page for the Geotagging Campaign on Historypin
(<https://www.historypin.org/nl/radio-tijdens-de-tweede-wereldoorlog>)

The first crowdsourcing campaign then went on to promote the use of the integrated Pundit Widget for musical instrument tagging (using MIMO terms) for curated collections from data providers on the WITH platform⁵⁵.

⁵⁴ <https://www.europeanasounds.eu/news/explore-the-world-of-sounds-and-help-out-audio-archives>

⁵⁵ <http://www.europeanasounds.eu/news/identify-musical-instruments-in-our-recordings>

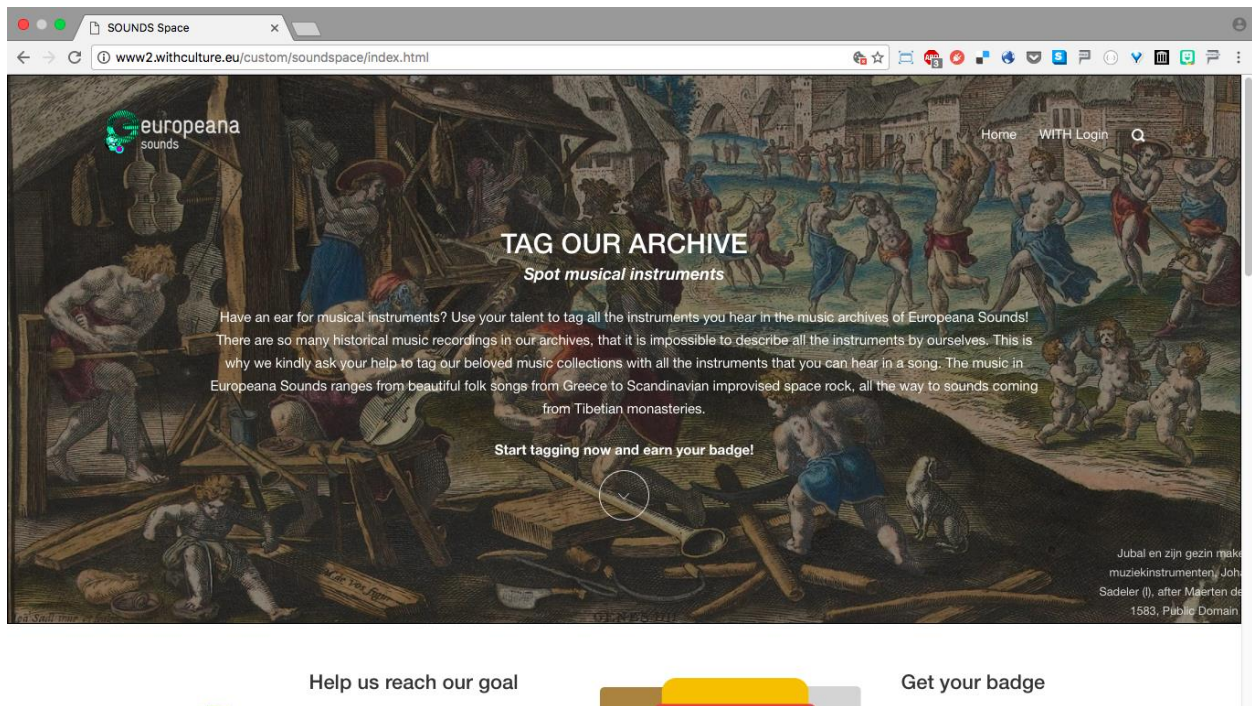


Figure 26: Landing page for the musical instruments tagging on WITH, through the Pundit Widget (<http://www2.withculture.eu/custom/soundspace/index.html>)

4.6.2 Semantic tagging in the second crowdsourcing campaign

The second crowdsourcing campaign started in December 2016⁵⁶ and ran until the end of the project. It started with another #crowdsourcingmonth, but promotion was extended from January onwards through social media campaigning on Thursdays under the #TagDayThursday hashtag (see Figure 27).



Figure 27: Promotional banner for the #TagDayThursday campaign in January 2017

⁵⁶ <http://www.europeanasounds.eu/news/december-crowdsourcing-month-identify-musical-instruments-in-our-recordings>

The second crowdsourcing campaign included various forms of semantic tagging. The Historypin geotagging functionality (see Section 4.2.2) was promoted for collecting better location data for various oral history collections from MMSH^{57 58}.

A further developed version of the WITH platform with native (semantic) tagging functionality was deployed to - again - collect musical instrument tags (using MIMO terms) for curated collections from data providers on the WITH platform (see Figure 28).

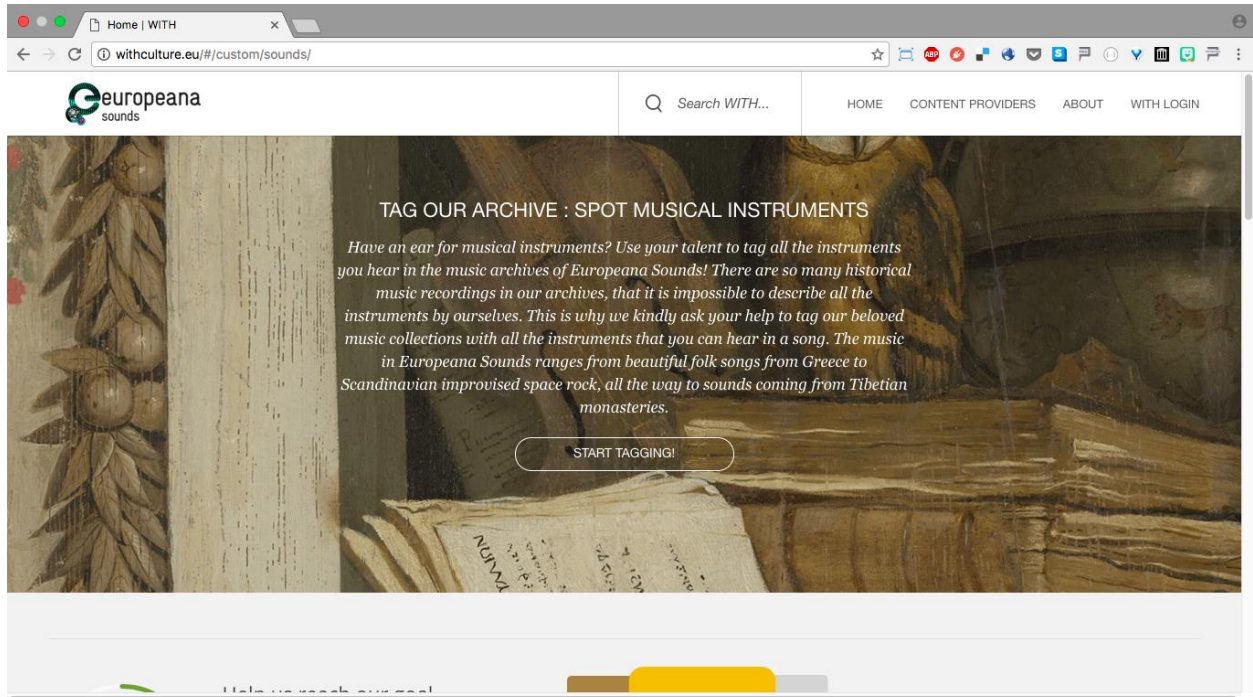


Figure 28: Landing page for the musical instruments tagging on WITH
(<http://withculture.eu/#/custom/sounds/>)

Finally in January 2017 the crowdsourcing campaign promoted the launch of Europeana Radio⁵⁹. This used a further developed version of the Europeana Radio application, which included functionality to crowdsource genre classifications for the streamable audio from the Europeana Music Collection. This crowdsourcing application offered three radio stations that users could pick from (Classical Music, Folk and Traditional Music and Popular Music). Whilst listening to the music they could enrich the genre classifications of the recordings with structured genre descriptions coming from the Wikidata repository (see Figure 29).

⁵⁷ <https://www.historypin.org/fr/oral-memory-of-armenian-people-living-in-mars>

⁵⁸ <https://www.historypin.org/en/recits-de-vie-de-francais-au-maghreb-1920-1962>

⁵⁹ <http://www.europeanasounds.eu/news/europeana-radio-discover-listen-and-tag-the-music-from-europeana-music-collections>

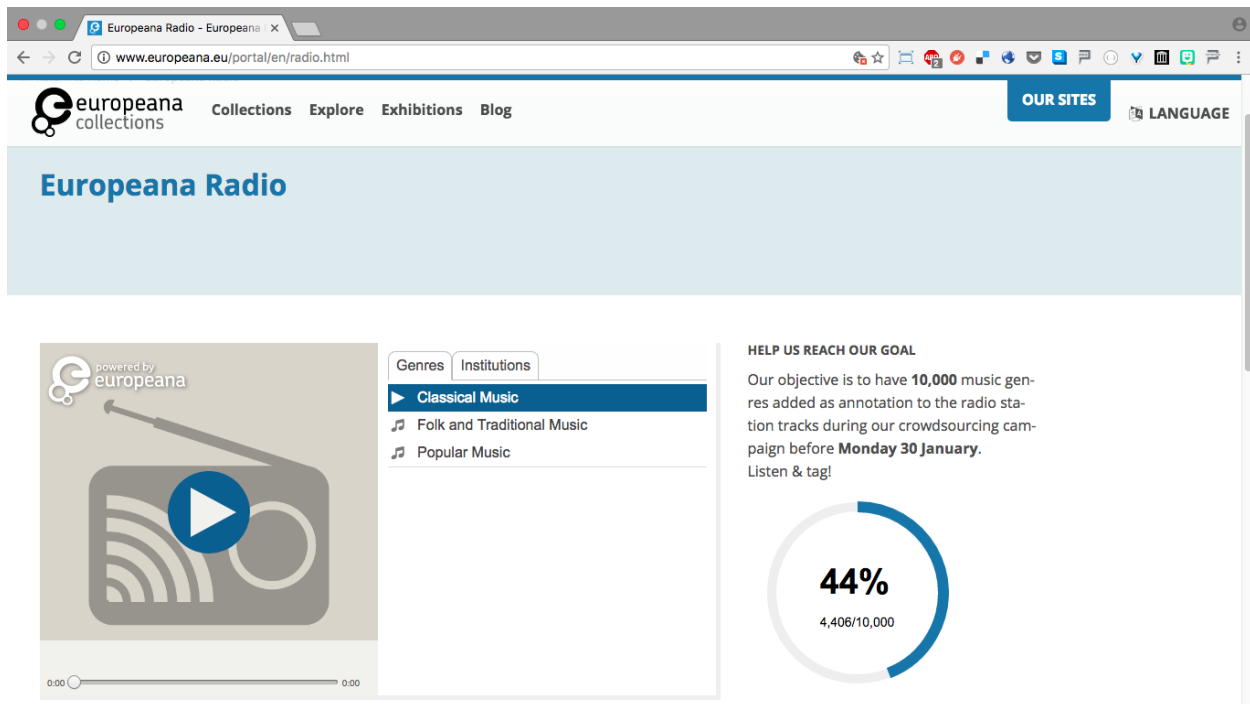


Figure 29: Landing page for the genre classification through Europeana Radio
(<http://www.europeana.eu/portal/en/radio.html>)

5 Semantic enrichment results

This section provided an overview of the KPIs related to WP2, focussing on the specific contribution to semantic enrichment of the metadata for records corresponding to sounds-related material in Europeana. These achievements were measured on the 27th of January, 2017. Final project results will be reported in the third and final periodic report.

5.1 KPI-8

This KPI relates to the automatic semantic enrichment during ingestion (see Section 2.1), as applied by EF. It corresponds to the amount of sound metadata records that pass through the workflow. Due to the timing of the writing of this document (before the end of the project) currently only the figure from quarter 11 is available, and the final number will be reported in the final periodic report. The figure for quarter 11 was 1,345,051. Since this number directly corresponds to the total amount of audio aggregated by Europeana as a whole (before and alongside of Europeana Sounds its own aggregation activities included), it reflects daily practice of the ingestion workflow. Given that fact, the accomplishment was not fully under the auspices of the Europeana Sounds project. For the amount of material that was ingested as a direct result of the aggregation by Europeana Sounds (WP1), the project ensured the semantic enrichment during ingestion included a genre classification that corresponded to the Europeana Sounds Genre Vocabulary (Ref 2).



5.2 KPI-9

This KPI relates to the (semantic) enrichment through crowdsourcing (see Section 2.2), as resulting from the various enrichment strategies (see D2.7, Ref 4) developed by the Europeana Sounds project.

In the crowdsourcing campaigns organised by the project (see Section 4.6), four crowdsourcing applications facilitated the collection of (semantic) enrichments in the form of crowdsourced annotations. At the time of writing this resulted in a total of 13,162 crowdsourced annotations being collected respectively: WITH (5,871), Europeana Radio (4,419), Pundit (2,846) and Historypin (26). Due to the various hurdles encountered during the development of the crowdsourcing applications - reliance on a yet unstable web standard (WADM) and changing priorities for the Europeana Collection development - the crowdsourcing campaigns started relatively late in the project (mainly Q10, 11 and 12). Although early estimations of the enrichment through crowdsourcing were not fulfilled (and have resulted in an adjusted KPI), we are now seeing results flowing in through the crowdsourcing applications. As WP2 lead, NISV is committed to further promotion of the crowdsourcing applications towards end users and data providers post project, in order to continue to ensure the value of the infrastructure and applications this project has delivered. Additionally NISV will assess possibilities to deploy the Europeana Sounds outcomes for its responsibilities related to user engagement in Europeana DSI-2.

As a result of the GLAM-Wiki activities promoted through task T2.3: *Wikimedia partnership* (Ref 1) another 75,000 crowdsourced enrichments have been collected for the benefit of the project, through data gathering of the relevant links created by volunteers between items on Europeana and items in the external crowdsourcing platform Wikidata. These links to Europeana in Wikidata are partly the by-product of the Edit-a-Thons (as reported in D2.8, Ref 6) as organised by the Europeana Sounds project, but not exclusive to Europeana. Europeana is recognised by Wikidata as a reliable source to link to and this has become a general community practice among the Wikidata volunteers focussing on cultural heritage (so also outside of the GLAM-Wiki activities 'pushed' by Europeana-related project).

This brings the total of (semantic) enrichments through crowdsourcing that Europeana Sounds has managed to collect to 88,162, which is more than double the expected outcome for KPI-9 (Ref 1). However, it is worth being mindful that the largest portion of the crowdsourcing resulted from a crowdsourcing platform external to the project (Wikidata). This can partly be explained by the relative late deployment of the crowdsourcing applications that the project (further) developed. Another important factor to consider is the established nature of the volunteer community that contributes to the various Wikimedia projects, which greatly surpasses the crowdsourcing capacity that cultural heritage organisations have been able to establish at this point in time.

5.3 KPI-10

This KPI relates to the semantic enrichment through alignment. The Europeana Sounds project has executed various successful experiments with the combination of expert knowledge from data managers, curators and amateurs with smart technologies to enable semi-automatic methods of enrichment. An early experiment with composers in the ONB dataset by AIT resulted in roughly 55,000 alignments. The CultuurLink experiments by EF with datasets from BL, CREM, MMSH and NISV resulted



in another 35,006 alignments between musical instrument terms and the controlled vocabulary that MIMO offers. The Wikimedia Commons experiment by EF resulted in 3,511 alignments between digital heritage objects uploaded to Wikimedia Commons and their corresponding metadata records on Europeana. And finally an experiment with linking Europeana metadata fields to various Linked Open Data vocabularies by AIT resulted in 1,539 alignments for an ONB dataset and 8,805 alignments for a NISV dataset.

This brings KPI-10 to a total of 103,861 alignments, thus greatly exceeding - more than tenfold - the expected outcome. This speaks to the potential of this type of semantic enrichment. It also shows that the feasibility of creating suggestions for annotations to be 'moderated' through is very present, based on the Europeana Data Model (of course depending on the data quality). This offers possibilities to enlarge the impact of future crowdsourcing campaigns in various Europeana domains.

5.4 Lessons learned

The work related to semantic enrichment that is documented in this deliverable has led to the following lessons learned:

- Crowdsourcing for sounds-related object and topics has been most impactful through established platforms, in terms of contributions.
- The niche that Europeana Sounds caters to is excellent in terms of having access to relevant domain knowledge and expertise, but is limited in volume.
- Constant communication about the crowdsourcing tasks is crucial for making a campaign successful. This requires a joined effort by the platforms themselves, data providers and Europeana.
- Technical hurdles to contribution need to be as low as possible for both contributors and data providers. Contributors will bounce when they encounter bugs or (for them) tedious steps in the crowdsourcing task. Data providers often lack technical expertise in their institutions, making the customisation that is often required for inclusion a limiting factor.
- Making the technical development dependent on an unstable web standard (WADM) makes sense from a data management perspective, but can have unwanted impact (unstable requirements, dependencies and delays) on the roadmap for completion of the end users services.

The experiments with enrichment through alignment demonstrate the potential for the possibilities to (semi-) automatically generate a large volume of (semantic) enrichment. The strategy for combining these with crowdsourcing through moderation (or validation) seems like a very interesting and potentially equally impactful area for user engagement.



6 Conclusions

In order for Europeana Sounds to pursue its objective to increase the quality of the metadata, WP2 deployed a threefold strategy to pursue semantic enrichment of the aggregated data: enrichment during ingestion, enrichment through crowdsourcing, and enrichment through alignment.

The enrichments strategies informed the (further) development of several core systems as components - namely the Entity API and Annotations API - in the Europeana Core Infrastructure. These components aimed to support the application scenario developed by the project (most prominently the crowdsourcing application), but also allowed for experimentation with various ways to align and link the data held by Europeana to external knowledge structures and resources (like MIMO and Wikidata). These experiments showed great potential for this semantic enrichment strategy for the future.

For semantic enrichment through crowdsourcing specifically, the project modelled and developed four types of annotations: semantic tagging, geotagging, object linking and moderation. This work contributed greatly to the state-of-the-art development of a W3C standard for web annotations (WADM, which all Europeana Sounds enrichments comply to).

The types of annotations developed by the project were deployed in various crowdsourcing applications: Traditional Music Pilot (object linking), Historypin (geotagging), Pundit (semantic tagging, moderation), WITH (semantic tagging, moderation), and Europeana Radio (semantic tagging).

The crowdsourcing applications have been promoted through two crowdsourcing campaigns. The results in terms of user engagement have varied (depending on maturity of the application, subject matter and expertise required for the crowdsourcing task). In some cases it has also suffered from late delivery. This calls for post-project continuation of the utilisation of these applications, in order for an apt valorisation of the development and impact of the enrichment through crowdsourcing. The project has already identified some opportunities to support this and NISV has committed to continuing some promotion towards end-users and liaising between data providers and the Technical Partners.

Looking at the results achieved with the various strategies, one can assess that enrichment during ingestion is already standard practice at EF. This strategies has mainly been improved through the project by connecting this to the Entity API, hence further normalising the semantic enrichment through a curated set of entities in the Entity Collection (which are all part of the data.europeana.eu namespace). Enrichment through crowdsourcing has performed less than was originally estimated (at the beginning of the project), but still shows potential. Finally enrichment through alignment - although experimental of nature in this project - outperformed the expected results and shows great potential, especially since it can also be combined with crowdsourcing for a multiplier effect.



References

Ref 1	Europeana Sounds Annex I - "Description of Work" http://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/Europeana_Sounds/Other%20documents%20related%20to%20the%20project/Dow%20Europeana%20Sounds%20620591%202015-10-19.pdf
Ref 2	D1.3 Ontologies for Sound http://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/Europeana_Sounds/Deliverables/EuropeanaSounds-D1.3-Ontologies-for-sound-v1.2.pdf
Ref 3	D2.10 Development of the Crowdsourcing Infrastructure http://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/Europeana_Sounds/Deliverables/EuropeanaSounds-D2.10-Development-of-crowdsourcing-infrastructure-v1-0.pdf
Ref 4	D2.7 Crowdsourcing Evaluation and Impact Assessment http://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/Europeana_Sounds/Deliverables/europeanasons-d2.7-crowdsourcing-evaluation-impact-v1.0.pdf
Ref 5	D2.4 Crowdsourcing infrastructure V1 Assessment and Recommendations http://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/Europeana_Sounds/Deliverables/EuropeanaSounds-D2.4-Crowdsourcing-infrastructure-V1-assessment-and-recommendations-v1.0.pdf
Ref 6	D2.8 GLAM-Wiki Collaboration Progress Report 2 http://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/Europeana_Sounds/Deliverables/europeanasons-d2.8-glam-wiki-collaboration-progress-report-2-v1.0.pdf



Appendix A: Terminology

A project glossary is provided at: <http://pro.europeana.eu/web/guest/glossary>.

Additional terms are defined below:

Term	Definition
AB	Advisory Board
APEX	Archives Portal Europe network of excellence
EC-GA	Grant Agreement (including Annex I, the Description of Work) signed with the European Commission
GA	General Assembly
PC	Project Coordinator
PI	Performance Indicator
PMB	Project Management Board
TEL	The European Library
UAP	User Advisory Panel
WP	Work Package