

WITH: Human-Computer Collaboration for Data Annotation and Enrichment

Alexandros Chortaras, Anna Christaki, Nasos Drosopoulos, Eirini Kaldeli, Maria Ralli,
Anastasia Sofou, Arne Stabenau, Giorgos Stamou, Vassilis Tzouvaras*
National Technical University of Athens
Athens, Greece

achort@cs.ntua.gr, achristaki@image.ntua.gr, ndroso@image.ntua.gr, ekaldeli@image.ntua.gr, mariaral@image.ntua.gr,
natasa@image.ntua.gr, stabenau@image.ntua.gr, gstam@cs.ntua.gr, tzouvaras@image.ntua.gr

ABSTRACT

The transformation that has been accomplished in Cultural Heritage (CH) during the last decades has resulted in the production of vast amounts of content from many different cultural institutions, such as museums, libraries and archives. A large part of this rich content has been aggregated in digital platforms that serve as cross-domain hubs, which however offer limited usability and accessibility of content due to insufficient data and metadata quality. In our effort to make CH more accessible and reusable, we introduce WITH, an aggregation platform that provides enhanced services and enables human-computer collaboration for data annotations and enrichment. WITH excels existing cultural content aggregation platforms by advancing digital cultural data through the combination of artificial intelligence automation and creative user engagement, thus facilitating its accessibility, visibility, and re-use. In particular, by using image and free text analysis methodologies for automatic metadata enrichment, in accordance to the human expertise for enrichment and validation through crowdsourcing approaches with gamification elements, WITH combines the intelligence of humans and computers to improve the quality of digital cultural content and its presentation, establishing new ways of collaboration between cultural organizations and their audiences.

CCS CONCEPTS

• **Information systems** → **Crowdsourcing**; • **Human-centered computing** → **Social tagging**; • **Applied computing** → **Annotation**; *Document metadata*;

KEYWORDS

cultural heritage, metadata, annotation, metadata enrichment, human-computer collaboration, crowdsourcing

ACM Reference Format:

Alexandros Chortaras, Anna Christaki, Nasos Drosopoulos, Eirini Kaldeli, Maria Ralli, Anastasia Sofou, Arne Stabenau, Giorgos Stamou, Vassilis

* Authors appear in alphabetical order.

This paper is published under the Creative Commons Attribution 4.0 International (CC BY 4.0) license. Authors reserve their rights to disseminate the work on their personal and corporate Web sites with the appropriate attribution.

WWW'18 Companion, April 23-27, 2018, Lyon, France

© 2018 IW3C2 (International World Wide Web Conference Committee), published under Creative Commons CC BY 4.0 License.

ACM ISBN 978-1-4503-5640-4/18/04.

<https://doi.org/10.1145/3184558.3191544>

Tzouvaras. 2018. WITH: Human-Computer Collaboration for Data Annotation and Enrichment. In *Proceedings of The 2018 Web Conference Companion (WWW'18 Companion)*. ACM, New York, NY, USA, 9 pages. <https://doi.org/10.1145/3184558.3191544>

1 INTRODUCTION

In recent years, the Cultural Heritage (CH) sector has seen an incredible transformation: accelerated digital evolution in the form of massive digitisation and annotation activities along with action towards multimodal cultural content generation from all possible sources has resulted in vast amounts of digital content being available through a variety of cultural institutions, such as museums, libraries, archives and galleries. Initiatives to aggregate this content in international level have resulted in digital platforms such as the Europeana¹ and the Digital Public Library of America². Such platforms operate as cross-domain hubs, making content accessible to users, readily available for search and study, or through creative applications and web services that reuse and repurpose it. While the main strength of such platforms lays in the vast number of the items they contain, they offer limited usability and accessibility due to insufficient data and metadata quality.

There are many factors that affect the quality of metadata and there have been many initiatives towards metadata quality assessment [1], recommendation [3] and standardization in CH [7]. The lack of structured and rich descriptive metadata, the complex, heterogeneous and multi-channel aggregation workflow, possible shortcomings that appear in the data providing process, surpassing manual quality control of automatic metadata generation in digital repositories are some of the main causes that result to poor metadata descriptors. This drawback highly affects the accessibility, visibility and dissemination range of the available digital content, also limiting the potential of added-value services and applications that re-use the available cultural resources in innovative ways, limiting also the user experience.

Metadata quality improvement usually faces the problem of scale, since improving the metadata quality of hundreds of thousands or even millions of records coming from different sources often requires a huge amount of time, effort and resources that aggregators and cultural heritage institutions unfortunately cannot afford. In this context, metadata enrichment services through automated metadata processing [2] and feature extraction along with crowdsourcing annotation services [4], [13] available in a centralized

¹<https://www.europeana.eu/portal/en>

²<https://dp.la/>

way through a dedicated platform can offer a remarkable opportunity for improving the metadata quality of digital content stored in platforms such as Europeana [18], [17], [9] while at the same time engaging users and raising awareness about cultural heritage assets.

In this paper we introduce WITH³, a cooperative cultural heritage content sourcing, publication and crowdsourcing platform that targets the CH domain as well as creative and cultural industries, aiming to enrich and improve metadata quality, facilitate the searchability, visibility, and promote the re-use of digital cultural content that is available through a variety of sources.

WITH addresses (stakeholders) various types of users from cultural heritage domain and creative industries having different levels of expertise and offers enriched services based on published resources through the use of the platform's programmatic interfaces (APIs). The platform offers metadata management and organisation in collections and introduces the concept of spaces, defined to enable data visualization based on various intuitive levels of aggregation. Through linked data principles and automated metadata processing services that aim to enhance and enrich available web resources, WITH contributes to the improvement of metadata quality. Additionally, it offers the possibility to cultural heritage institutions and aggregators to launch ad-hoc crowdsourcing campaigns with gamification elements and measurable results, thus mobilizing and engaging users to execute useful tasks for the enrichment and validation of selected cultural heritage metadata. Through the offered crowdsourcing campaigns, targeted users are enabled to add annotations (e.g. semantic tagging, image tagging, geotagging etc), depending on the type of content and missing metadata, and validate existing annotations (e.g. by upvoting or downvoting) in user-friendly and engaging ways (e.g. through leaderboards or rewards).

The rest of the paper is structured as follows: Section 2 illustrates the overall design of the platform and its constituent parts. Section 3 provides a detailed description of the data management approach in terms of metadata aggregation workflow (3.1), organisation and management of content in terms of collections and spaces (3.2). Section 4 describes the metadata enrichment methodologies adopted in the platform, in terms of AI metadata extraction methodologies (4.1) and crowdsourcing annotation services (4.2). Finally, Section 5 concludes the work and addresses how the platform will evolve in the future. In order to help readers understand the underlying technologies, at the end of all sections, we provide a running example of WITH platform's use in the area of music collections, motivated by a real WITH usecase scenario.

2 WITH ECOSYSTEM OVERVIEW

The main motivation for WITH is to utilize CH repositories in unison and promote the digital cultural content by enhancing its accessibility and discoverability [6]. WITH can be viewed as a complete ecosystem that offers enriched services, aggregates content from multiple sources, resolves interoperability issues by performing automatic mappings in the backend, improves metadata quality through linked data principles and automated free text and image analysis services, and mobilizes and engages users to execute useful

tasks for the enrichment and validation of selected cultural heritage metadata. In its core, it is a platform especially designed for CH with focus on human-computer collaboration, offering a set of services such as content integration, management, retrieval and curation, automatic metadata enrichment, and crowdsourcing campaigns to accommodate human expertise and intelligence.

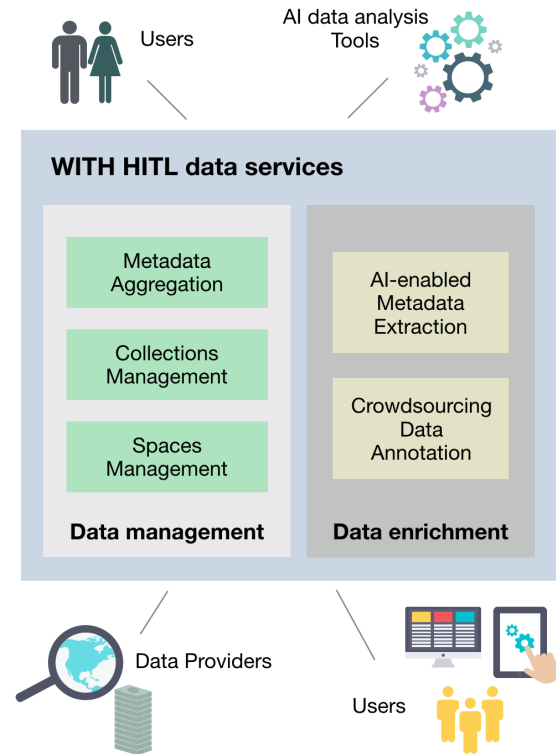


Figure 1: Platform Overview.

WITH addresses all kinds of users from cultural heritage domain and creative industries and serves as a multipurpose platform with varying level of functionality abstraction depending on user's level of expertise. Specifically:

- For *end users*, WITH serves as a starting point for the discovery of CH resources that reside in distributed repositories. They can combine different sources with their own material, allowing them to offer new interesting narratives for the appreciation and communication of culture. Records from user collections can be used to create virtual exhibitions that allow for a quick and easy way to publish a presentation or a narrative that can be shown to an audience, such as a group of students, without having to use any tools outside WITH, accessed and edited directly from a web browser. Groups of users can also collaborate in their shared exhibitions and add or modify content together.
- For *CH professionals*, WITH offers the tools to compose and contextualize eclectic collections, contributing to the promotion, improvement, and evolution of digital CH knowledge.
- For *content holders*, WITH offers an easy to use content & metadata repository and management system, that can

³<http://with.image.ntua.gr/>

ensure interoperability with standards, best practices and aggregator guidelines.

- For *CH organisations*, WITH offers Spaces to organise and promote their content and improve their content's metadata and engage with users through crowdsourcing campaigns.
- For *aggregators*, WITH integrates the metadata interoperability platform (MINT)⁴ [5], [11], an open source, web based platform that fully implements workflows for the ingestion, formal mapping, transformation and aggregation of metadata records.
- For *developers*, WITH serves as a back-end to build applications that reuse digital heritage, employing WITH API⁵ that publicly offers all the available data and services of the platform, fully accessible to any developers who wish to utilize it to their applications.

The basic underlying system architecture can be viewed in Fig. 1 that mainly depicts the user involvement in the data services of the platform. WITH human-in-the-loop approach targets to the advancement of data quality and organisation, offering data services that the users can employ in order to better manage and enrich data. Specifically, with the aid of *Metadata Aggregation*, *Collection management* and *Spaces Management* services, users can upload, collect and organise content from several sources and also create interesting content views and presentations and store all this data in WITH database. Moreover, using the *AI-enabled Metadata Extraction* services that integrate automatic or semi-automatic annotation tools (like automatic text analysis, image annotation etc), users can select interesting ways to analyze the relevant content, developing and storing new content descriptions. Finally, by initiating and managing crowdsourcing campaigns, users can involve other users in recognition tasks that are difficult for machines to perform (like instrument recognition in music etc), thus further improving the content descriptions.



Figure 2: User involvement in WITH.

⁴<http://mint.image.ece.ntua.gr/redmine/>

⁵api.withculture.eu/assets/developers-lite.html

Based on the above system architecture, the user interactive involvement in WITH can be summarized in four interactive processes, as illustrated in Fig. 2. The *Content Search* and the *Content Management* processes that enable users to collect content and organise content, as well as the *Metadata Enrichment* and the *Crowdsourcing* processes that enable users to advance content descriptions, using AI content analysis tools or human annotations, respectively. In the next sections, we describe these main processes, providing technical details and usecase examples.

3 CONTENT SEARCH AND MANAGEMENT

3.1 Content aggregation

The content aggregation is mainly accomplished through searching in external repositories of CH Objects and digital libraries. WITH mashups APIs from different digital CH resources, such as Europeana, the Digital Public Library of America, Rijksmuseum, the British Library, National Library of Australia, YouTube, Historypin, Internet Archive etc, and provides a single powerful new service, which gives access to a huge set of heterogeneous items (images, videos, different metadata schemata etc). WITH supports a variety of different data models (e.g., EDM, LIDO etc) and formats (e.g. XML, JSON-LD), and resolves interoperability issues by integrating MINT, an open source metadata interoperability web based platform that fully implements workflows for the ingestion, formal mapping, transformation and aggregation of metadata records in the backend.

The User Interface (UI) offers federated and faceted search services, which enable the user to apply multiple search criteria in different combinations based on the metadata of search results, navigate through the results via different presentation views and retrieve more information about items of interest. Once a search term or phrase has been provided, WITH makes a parallel search to all selected repositories and creates facets that can be used to narrow down results. As most digital data in external repositories are accompanied with rich metadata, the user may choose from a variety of specialized filters such as rights, creator, media type and dates, resulting in the desirable objects among million others. The search results as well as the facets can be accessed through the API, so other developers can easily include this search feature in their applications, or even improve on it and adapt it to their needs.

WITH also provides a mechanism of mass data import from Europeana so as to facilitate the acquisition of multiple data and the creation of big thematic collections. The user can easily import and add to her profile a Europeana Collection by providing its ID or even acquire Europeana items corresponding to a specific search term and organised into a WITH collection. In addition, the user can upload and curate her own items by uploading the desired items (image, video, text) and adding the respective metadata.

All records imported in the platform are instantly transformed into the internal data model. The WITH data model follows the paradigm of major CH repositories, such as Europeana, including extensions in order to satisfy compatibility restrictions with various different data models. It is compliant with the Europeana Data Model (EDM) [10], [8] and supports two-way transformation capabilities, therefore all changes in the records in their internal data model can still be exported in different formats. The WITH API also

supports various serializations such as JSON, XML, RDF, etc. This allows for all resources in WITH to be shared with other external applications or platforms.

3.2 Collections and Spaces management

WITH users are able to create collections that become available under their profile. All records imported in the platform are instantly transformed in the internal data model, so they can be instantly included in collections created by users and stored in their user profiles, while their original record structure is kept intact as a backup and can still be accessed. This transformation facilitates the transition from items to web resources allowing for greater expressivity and more powerful collections. Along with collections, users can organise their content, which was either uploaded or collected from external CH repositories, into exhibitions that provide enhanced and more playful visualization features.

WITH platform also provides thematic content organisation in Spaces. Spaces is a concept that corresponds to specific, access-based views of stored data, enables the organisation of cultural content in different thematic categories and views, and allows different ways of interaction with the end user (e.g. users can re-use the provided content, comment and share etc).

This added feature enables interested CH organisations to design and host custom web spaces in order to promote their content and engage with users. Dedicated Spaces are associated with the content holder, for the visualization and web publication of digital collections, exhibitions and stories, while facilitating the public's access to content available for creative re-use. Spaces alleviate the need for specialized infrastructure by offering an easy alternative solution for the publication of Linked Open Data, an important task in the CH area [16].

The content access within a space (collections and exhibitions etc.) is limited to specified user groups. The scope of the search engine in each space can be customized as well, for example to exclude some sources or, to only search for video resources etc. Individual customization of the front end (descriptive texts, images, CSS) is also available within its space.

Example 3.1. As a case in point, let's introduce Simona a music enthusiast fascinated by Greek traditional music, who wants to explore and discover relevant musical artifacts and build, organise and contextualize her own library. First, Simona creates a user profile and initializes an empty collection, where all the assets will be saved. As she owns content from Greek traditional instruments, she uploads and curates it by providing name, description and rights metadata. To enrich her collection and add more assets, she uses WITH search service to look for Greek traditional instruments, like "Laouto" and "Lyre". Simona collects the assets from the Digital Public Library of America, National Library of Australia and Europeana, and further filters the Europeana results by setting the preferable rights and mime type illustrated in Fig. 3a. From a set of 11 music pieces returned from Europeana containing "Lyre Greek", she opens the ones she is interested in, as shown in Fig. 3b, listens to the music, inspects the accompanying metadata and collects the items she wants. Collected items are automatically transformed

into the WITH datamodel, an excerpt of which, is shown at Excerpt 1. Simona can then observe and manage her collection shown in Fig. 3c.

Excerpt 1 WITH datamodel

```
"descriptiveData": {
  "label": "Greek from Festival of Song",
  "description": "This image has been taken from Festival
    of Song: a series of Evenings with the Poets",
  "keywords": [
    "Greek",
    "kylix",
    "lyre",
    "symposium"
  ],
  "isShownAt": "http://www.europeana.eu/api/ANnuDzRpW",
  "isShownBy": "http://farm8.staticflickr.com/7406.jpg",
  "rdfType": "http://www.europeana.eu/schemas/
    edm/ProvidedCHO",
  "country": "united kingdom",
  "dclanguage": "English",
  "dctype": "scanned image",
  "dcredits": "Public Domain",
  "dctermspatial": "New York, 1866",
  "dcformat": "jpg"
}
```

4 METADATA ENRICHMENT

Metadata enrichment can be defined as the process of improving the metadata accompanying an object by adding new statements about the object itself. In the cultural heritage domain, metadata enrichment attracts the interest of many organisations and researchers [2] since low metadata quality is a very common problem that imposes restrictions in the accessibility and discoverability of the actual content. As a process, it is accomplished by applying specifically designed methodologies such as co-referencing, alignment, contextualisation and annotation, and results in new metadata created at the end of the process. It is used to "standardize data" by linking it to authority resources, improve multilingual coverage in datasets [18], [17], contextualise resources and much more.

WITH ecosystem enables automatic enrichment implemented via a set of image and text analysis methodologies, as well as manual enrichment in the form of content annotation accomplished via crowdourcing initiatives as described in Sections 4.1 and 4.2 respectively.

4.1 Annotation Generation

The items that are collected and aggregated in WITH initially contain the original metadata that are provided directly by the sources. These are stored in the respective representation of each item using the WITH data model. To allow enrichment of these metadata with additional, either manually or automatically generated, metadata, WITH supports the use of annotations. Additional metadata are assumed to have the form of Linked Data resources or other IRIs, thus a WITH annotation essentially associates a WITH item, or a part of it, with a Linked Data resource or other IRI. Individual

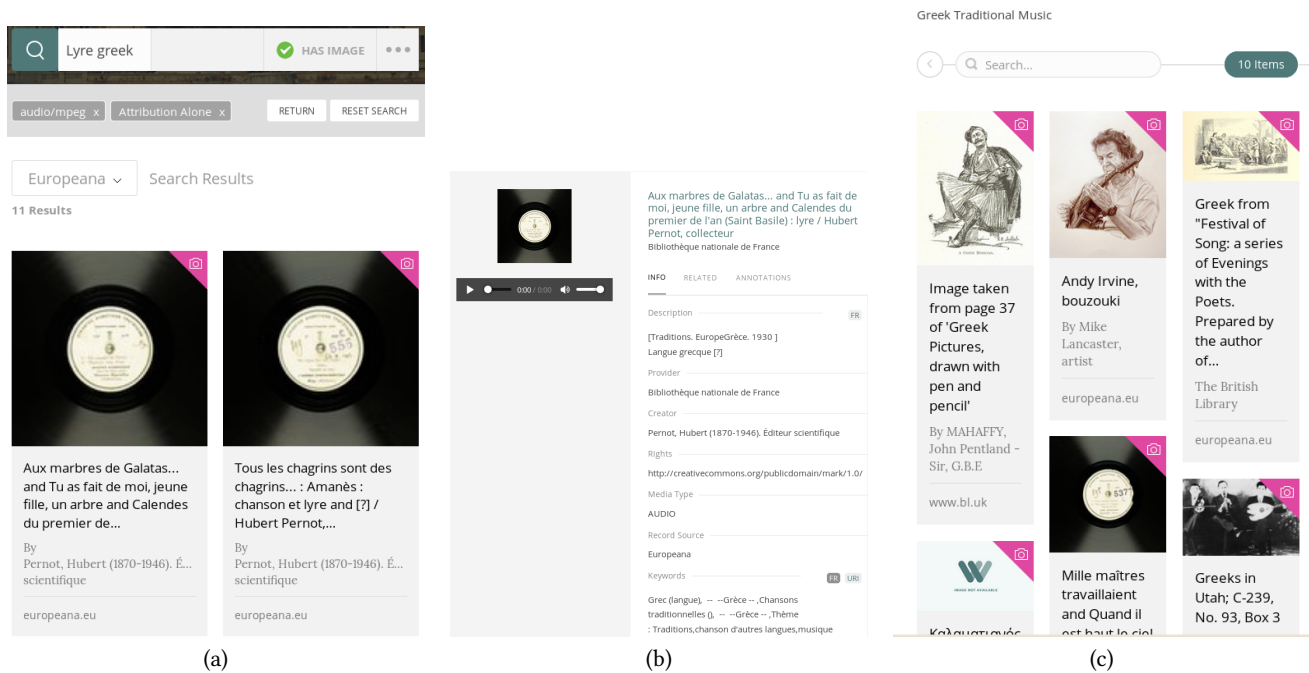


Figure 3: User data management: (a) search results, (b) selected item from search and (c) collection.

users can be involved in the validation process of annotations by upvoting or downvoting annotations.

To facilitate annotation creation, retrieval, management and interoperability, WITH includes a thesauri manager that is responsible for importing, through an offline process, the Linked Data vocabularies and datasets that make up the pool of potential annotation resources. The thesauri manager converts the imported vocabularies from their source format (e.g. SKOS thesauri, OWL ontologies, N-triples datasets) to a common model, stores them in the WITH thesauri database and indexes them to allow fast search and retrieval. Currently, several widely used Linked Data vocabularies, datasets and ontologies have been incorporated in WITH, including the Getty Art & Architecture Thesaurus (AAT)⁶, the GEMET Thesaurus⁷, the Musical Instruments Museums Online (MIMO) Thesaurus⁸, the Europeana Fashion Thesaurus⁹, the Europeana photoVocabulary, Wordnet¹⁰, and datasets such as DBpedia¹¹ and Geonames¹². In all the above vocabularies and datasets, each resource is always accompanied by one or more textual labels, possibly in several languages. These labels provide textual representations for the specific resource and are used for indexing the resources and facilitating lookup.

WITH's annotation model is based on W3C's Web Annotation Model [14], which is a structured model and format to enable annotations to be shared and reused across different hardware and

software platforms. In brief, a WITH annotation consists of an id, a list of annotators, a body, a target, and a list of scores. An annotation may be generated either automatically by a content analysis software, a web-service etc., or manually by a human annotator. Thus, the list of annotators contains all relevant information about the origins of each annotation. The core part of the annotation is its body which identifies the relevant Linked Data resource or IRI. The target of an annotation identifies the WITH item, or the part of a WITH item, which the particular annotation relates to the body resource. A part of a WITH item may be a particular metadata field value, or a part of it (e.g. a span of text or an image area). Finally, the list of scores holds information about the users that have upvoted or downvoted the particular annotation.

Based on the above annotation model, WITH provides a number of tools for the manual and automatic generation of annotations. In the manual annotation process, the user has to directly choose a resource from the underlying thesauri database and add it as annotation for a particular item. The user starts typing in a keyword and the user interface assists him through an autocomplete functionality, which consults the underlying thesauri index, and suggests to the user resources having textually matching labels. The user can restrict the scope of autocomplete suggestions by selecting only specific vocabularies.

Part of the manual annotation tools is the geotagging tool, which allows the generation of annotations with a geo URI in their body to represent geographical locations. To generate such an annotation the user clicks on the desired points of the provided map.

The automatic annotation tools are aimed to provide textual and visual metadata analysis. Textual metadata are the main descriptive metadata of the items, namely their title and description fields.

⁶<http://www.getty.edu/research/tools/vocabularies/aat/>

⁷<http://www.eionet.europa.eu/gemet/en/themes/>

⁸<http://www.mimo-db.eu/InstrumentsKeywords/>

⁹<http://thesaurus.europeana-fashion.eu/thesaurus/>

¹⁰<http://wordnet-rdf.princeton.edu/>

¹¹<http://wiki.dbpedia.org/>

¹²<http://www.geonames.org/ontology/documentation.html>

These can undergo natural language processing and machine learning based named entity recognition and disambiguation (NERD) with the purpose to identify in them occurrences of named entities (persons, locations and organisations) as instances of the supported Linked Data resources [15]. Currently NERD is performed using DBpedia Spotlight API¹³ and translating its results to the WITH annotation data model. Textual item data can also be used to automatically generate annotations through the dictionary lookup tool that is implemented in WITH. This tool, uses efficient dictionary generation techniques to compile a fast to search dictionary of resources from a selected subset of the available thesauri, and then scans the textual metadata to find occurrences of the dictionary terms. The lookup process is assisted by natural language processing techniques [12] so as to use lemmatisation and part of speech information to improve detection accuracy. Multilinguality in the above processes is handled automatically: metadata and resource label languages should match.

Apart from textual metadata, image related metadata can also be extracted using the CultIVML service, that implements an automatic image annotation tool for WITH. The tool uses state-of-the-art computer vision algorithms based on deep neural networks [19] to analyze item images. The automatic annotation process can produce localization of human faces, localization of a diverse set of common objects, as well as generic image classification. The CultIVML annotator is trained using the Image-net database¹⁴, so the annotations it generates are WordNet resources.

All annotations for each item are displayed along with the other item metadata. Users are allowed to upvote or downvote an annotation, enabling calculation of validity confidence statistics for each annotation.

Example 4.1 (continued). Unrolling our running example, Simona wants to annotate, contextualize and enrich her collection using WITH automatic annotation tools. She chooses the automatic collection annotation tool to enrich the records using the MIMO Thesaurus which contains an exhaustive list of all known musical instruments. The process results in an enriched music collection, with MIMO annotations derived from each record's metadata. For instance, Simona has uploaded an item with the following description: *"The three main Greek musical instruments brought to America: the Cretan lyra played by Stelios Mavrikis, the clarinet played by Louis Kosta, and the laute played by Andrew Bathemess, leading Greek musician."* The enriched item, now contains the relevant annotations, namely *"mimo : Lyre"*, *"mimo : Lute"* and *"mimo : Clarinet"*. These annotations are represented in WITH as shown in Excerpt 2.

4.2 Crowdsourcing Data Annotation

Automatic enrichment techniques, no matter how well designed they are, do not always result in a high success rate, nor do they cover the growing need for as high quality data as possible. This is due to a variety of factors, such as lack of quality training datasets, low accuracy for many automatic annotation algorithms and extreme cases that are too difficult for the algorithm to handle. This problem is magnified when it comes to cultural data as they usually are old, multilingual, sometimes can even be obsolete and in

Excerpt 2 Annotation for MIMO term "Lute"

```
{
  "annotators": [
    {
      "generator": "Dictionary Annotator",
      "created": "2017-01-06T12:59:12.387Z"
    }
  ],
  "motivation": "Tagging",
  "body": {
    "uri": "http://www.mimo-db.eu/InstrumentsKeywords/3394",
    "label": {
      "en": [
        "Lute"
      ]
    },
    "uriVocabulary": "MIMO"
  },
  "target": {
    "recordId": "58bd48594c74794c8fb0018e"
  }
}
```

general pose difficulties in handling. In this context, crowdsourcing is the medium for human-computer collaboration and can offer a remarkable opportunity for improving the metadata quality of digital content stored in the platform, while at the same time engaging users and raising awareness about cultural heritage assets.

WITH implements a crowdsourcing infrastructure for annotations, completed with validation and user voting functionalities. The crowdsourcing component essentially complements any automatic enrichment technique by enabling the crowd to validate the automatic results and decide on difficult data, thus improving the machine's accuracy. WITH aims to mobilize and engage users to execute useful tasks for the enrichment and validation of selected cultural heritage metadata. The platform enables the deployment of crowdsourcing web spaces that present end-users with specific challenges for the enrichment of selected cultural items' metadata. Three main scenarios can be supported:

- enrich metadata by adding new annotations, namely add values to missing metadata fields or add additional tags by selecting terms from controlled vocabularies, depending on the annotation field;
- validate and suggest changes to existing metadata entries, especially the ones identified via automatic methods (e.g. image processing, free text analysis);
- up/downvote annotations that have been added manually by other users.

WITH enables the initialization and management of crowdsourcing campaigns by CH institutions or individuals with cultural content which is either already annotated with automatic content analysis systems or is difficult to analyze by automatic techniques. To initiate a crowdsourcing campaign the stakeholders should:

- (1) import their cultural content into WITH, through the methods provided by the platform;
- (2) organise their data into collections;

¹³<http://www.dbpedia-spotlight.org/>

¹⁴<http://www.image-net.org>

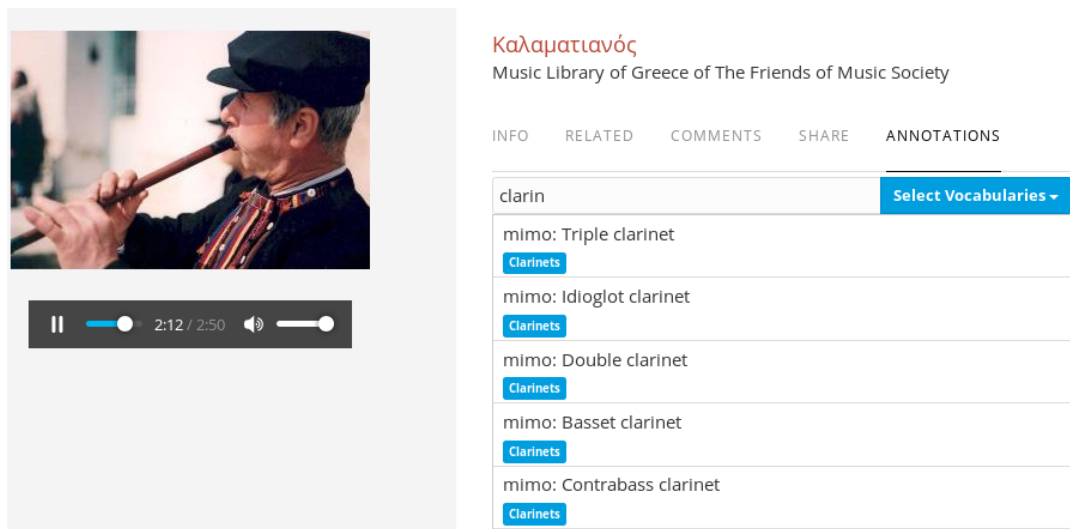


Figure 4: The process of adding new instrument tags from the MIMO thesaurus.

- (3) make a thematic Space according to their content;
- (4) use the WITH automatic annotation tools to enrich their data where possible;
- (5) contact the WITH administrators and specify the desired crowdsourcing features about their campaign such as duration, target annotation number, desired annotation type (semantic tagging, image tagging, geotagging, etc.), vocabularies and thesauri.

During the crowdsourcing campaign period, targeted users are able to add annotations, and validate existing annotations by up-voting or downvoting annotations in a user-friendly and engaging way. In order to inspire users to invest time and effort and make the campaign more immersive, playful and compelling, some key game mechanics were incorporated. Badge rewards (e.g. bronze, silver and golden badge), depending on the number of contributed or validated tags and a leaderboard, keeping track of the most active "taggers", namely the users with the most contributed tags and validations, were used to make the campaign more appealing. Useful metrics and statistical results about the campaigns are also presented in a dashboard (e.g. number and types of annotations added, number of users etc) to help the campaign organisers track the progresses and monitor the outcome.

Example 4.2 (continued). Returning to our running example, Simona wants to recognize all the known instruments included in her songs but the automatic annotation mechanism recognizes only instruments extracted from the record's metadata. Consequently, a crowdsourcing campaign is needed, so that experts can annotate her music recordings and validate the existing automatic annotations.

The crowdsourcing campaign described above has already been conducted with success, using WITH in the framework of the Europeana Sounds project¹⁵, where semantic enrichment of content metadata was performed. The campaign was carefully designed, the goal of the campaign was determined to be musical instrument

semantic tagging of music recordings available in the Europeana Sounds project, the objects to be annotated were identified along with their annotation properties (concepts from structured vocabularies), the MIMO Thesaurus was selected, and niche communities were targeted (musicologists, music experts, music enthusiasts) through various reach out events. Parallel events were organised in different European countries and the tasks were running online.

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 recordings¹⁶. 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 into WITH from the Europeana database –using the Europeana API.

The semantic tagging Crowdsourcing Campaign lasted for a period of a month, in which the users were encouraged to visit the space and start adding their tags, as in Fig. 4, or validate the existing ones, as in Fig. 5, working together on the same goal. Parallel events were organised in different European countries and the tasks were running online. For the entire crowdsourcing period users were motivated through the homepage of the Sounds Crowdsourcing Space, where the goal of the campaign in terms of number of annotations to be achieved was highlighted. The overview of the progress, the number of contributed tags and the percentage of goal completeness were also provided. The user participation was more than satisfactory, with more than 5000 musical instruments identified by dozens of users.

Coming back to Simona again, in the end of the campaign, every item in her collection, even the items she uploaded, was enriched with several relative instrument tags coming from dozens of music experts and individuals across Europe. Her music collection ended

¹⁵<http://www.europeanasounds.eu/>

¹⁶<http://demo.withculture.eu/#/custom/sounds/>

up with records tagged with three or more instruments each, enabling her to better organise and search her collection and reuse her music archives more efficiently.

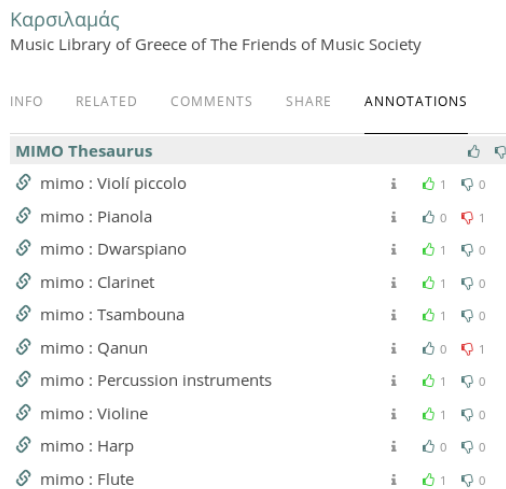


Figure 5: The tagging area with the existing instrument tags.

5 CONCLUSIONS AND FUTURE WORK

In this paper we presented a human-computer collaboration platform based on linked data and machine principles with unique features, such as automatic metadata enrichment and crowdsourcing services, specifically aimed for the CH domain.

We provided a detailed description of the workflow and the user engagement in the platform and illustrated the basic principles via a running example, and we also reported on a successful crowdsourcing campaign performed under Europeana Sound framework, which highlighted the need of a dedicated crowdsourcing space with various levels of freedom for campaign organisers.

WITH is an evolving ecosystem, so it changes with time as new repositories are aggregated, new spaces are created and new features and services are constantly designed and aimed to be deployed soon. Some of the features under development are:

- An automated musical instrument tag extraction service, based on specifically trained deep neural nets. This service will be used to provide automated tags which will be fed to users for verification and validation and then will be included in the training set so as to close the loop of human-computer collaboration.
- New visual feature extraction methodologies that will be used to enrich metadata of image items.
- Exploit the user annotations for training/improving machine learning algorithms.
- Fully automated crowdsourcing campaign creation.
- WITHcrowd: A dedicated crowdsourcing platform, satellite to WITH, which combines the best of human and machine intelligence to annotate, enrich, curate and validate cultural content, by providing the opportunity for obtaining information and metadata, through outsourcing simple tasks to an undefined public, in order to enrich and promote the digital cultural heritage from various sources. WITHcrowd

essentially will complement any automatic feature extraction technique by enabling the crowd to validate the automatic results and decide on difficult data, thus improving machine accuracy, and will provide various levels of freedom to campaign organisers.

ACKNOWLEDGMENTS

We acknowledge support of this work by the project *Human-in-the-Loop* (671048), which is funded by the General Secretariat of Research and Technology (GSRT) of Greek Ministry of Education, Research and Religious Affairs and the project *APOLLONIS* (MIS 5002738), which is implemented under the Action: Reinforcement of the Research and Innovation Infrastructure, funded by the Operational Programme: Competitiveness, Entrepreneurship and Innovation (NSRF 2014-2020) and co-financed by Greece and the European Union (European Regional Development Fund). WITH has also been partially funded by Europeana Space¹⁷, EUScreen¹⁸, Europeana Photography¹⁹, Europeana Food and Drink²⁰, and Europeana Sounds²¹.

REFERENCES

- [1] E. Bellini and P. Nesi. 2013. Metadata Quality Assessment Tool for Open Access Cultural Heritage Institutional Repositories. In *Information Technologies for Performing Arts, Media Access, and Entertainment*, P. Nesi and R. Santucci (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 90–103.
- [2] G. Berardi, A. Esuli, S. Gordea, D. Marcheggiani, and F. Sebastiani. 2012. Metadata Enrichment Services for the Europeana Digital Library. In *Theory and Practice of Digital Libraries*. Springer Berlin Heidelberg, Berlin, Heidelberg, 508–511.
- [3] M-C. Dangerfield and L. Kalshoven. 2015. *Report and Recommendations from the Task Force on Metadata Quality*. Europeana Technical Report. Europeana Foundation, Hague, Netherlands. https://pro.europeana.eu/files/Europeana_Professional/Europeana_Network/metadata-quality-report.pdf
- [4] C. Dijkshoorn, V. de Boer, L. Aroyo, and G. Schreiber. 2017. Accurator: Nichesourcing for Cultural Heritage. *CoRR* abs/1709.09249 (2017).
- [5] N. Drosopoulos, V. Tzouvaras, N. Simou, A. Christaki, A. Stabenau, K. Pardalis, F. Xenikoudakis, and S. Kollias. 2012. A Metadata Interoperability Platform. Museums and the Web, April 2012, San Diego, USA. <http://www.image.ece.ntua.gr/publications.php>
- [6] M. Giazitzoglou, V. Tzouvaras, A. Chortaras, E.M Alfonso, and N. Drosopoulos. 2017. The WITH platform: Where culture meets creativity. In *Proceedings of the Posters and Demos Track of the 13th International Conference on Semantic Systems - SEMANTICS2017 (CEUR Workshop Proceedings)*, Javier D. Fernández and Sebastian Hellmann (Eds.). Aachen. <http://ceur-ws.org/Vol-2044/>
- [7] C.A. Harper. 2016. Metadata Analytics, Visualization, and Optimization: Experiments in statistical analysis of the Digital Public Library of America (DPLA). *code4lib* 33 (2016).
- [8] A. Isaac and V. Charles. 2017. Europeana Data Model Definition. (2017). pro.europeana.eu/resources/standardization-tools/edm-documentation
- [9] A. Isaac, H. Manguinhas, J. Stiller, and V. Charles. 2015. *Report on Enrichment and Evaluation*. Europeana Technical Report. Europeana Foundation, Hague, Netherlands. <http://pro.europeana.eu/taskforce/evaluation-and-enrichments>
- [10] R. Isaac, A. and Clayphan. 2010. Europeana Data Primer. (2010). <http://pro.europeana.eu/documents/900548/770bdb58-c60e-4beb-a687-874639312ba>
- [11] I. Kolli, V. Tzouvaras, and G. Drosopoulos, N. and Stamou. 2012. A systemic approach for effective semantic access to cultural content. *Semantic Web* 3, 1 (2012), 65–80.
- [12] C.D. Manning, M. Surdeanu, J. Bauer, J.R. Finkel, S. Bethard, and D. McClosky. 2014. The Stanford CoreNLP Natural Language Processing Toolkit. In *ACL (System Demonstrations)*. The Association for Computer Linguistics, 55–60.
- [13] J. Oomen and L. Aroyo. 2011. Crowdsourcing in the cultural heritage domain: opportunities and challenges. In *Proceedings of the 5th International Conference on Communities and Technologies*. 138–149.

¹⁷<http://www.europeana-space.eu/>

¹⁸<http://www.euscreen.eu/>

¹⁹<http://www.europeana-photography.eu/>

²⁰<http://foodanddrinkeurope.eu/>

²¹<http://www.europeanasounds.eu/>

- [14] R. Sanderson, P. Ciccarese, and B. Young. 2017. Web Annotation Data Model. (2017). <https://www.w3.org/TR/annotation-model/>
- [15] N. Simou, A. Chortaras, G. Stamou, and S. Kollias. 2017. Enriching and Publishing Cultural Heritage as Linked Open Data. In *Mixed Reality and Gamification for Cultural Heritage*. 201–223.
- [16] N. Simou, J.P. Evain, N. Drosopoulos, and V. Tzouvaras. 2013. Linked European television heritage. *Semantic Web* 4, 3 (2013), 323–329.
- [17] J. Stiller, A. Isaac, and V. Petras. 2014. *EuropeanaTech Task Force on a Multilingual and Semantic Enrichment Strategy: Final Report*. Technical Report. <http://pro.europeana.eu/documents/468623/8b75b054-712e-432b-a0f7-761898e6f60e>
- [18] J. Stiller, V. Petras, M. Gäde, and A. Isaac. 2014. Automatic Enrichments with Controlled Vocabularies in Europeana: Challenges and Consequences. In *Digital Heritage. Progress in Cultural Heritage: Documentation, Preservation, and Protection*, M. Ioannides, N. Magnenat-Thalmann, E. Fink, R. Žarnić, A.-Y. Yen, and E. Quak (Eds.). Springer International Publishing, Cham, 238–247.
- [19] C. Varytimidis, G. Tsatiris, K. Rapantzikos, and S. Kollias. 2016. A systemic approach to automatic metadata extraction from multimedia content. In *SSCI*. IEEE, 1–7.