CrowdHeritage: Improving the quality of Cultural Heritage through crowdsourcing methods

Maria Ralli*, Spyros Bekiaris*, Eirini Kaldeli*, Orfeas Menis - Mastromichalakis*, Natasa Sofou*, Vassilis Tzouvaras*
and Giorgos Stamou*
*Artificial Intelligence and Learning Systems Laboratory
School of Electrical and Computer Engineering
National Technical University of Athens, Greece

***Abstract*—The lack of granular and rich descriptive metadata highly affects the discoverability and usability of the digital content stored in museums, libraries and archives, aggregated and served through Europeana, thus often frustrating the user experience offered by these institutions' portals. In this context, metadata enrichment services through automated analysis and feature extraction along with crowdsourcing annotation services can offer a great opportunity for improving the metadata quality of digital cultural content in a scalable way, while at the same time engaging different user communities and raising awareness about cultural heritage assets. Such an effort is Crowdheritage, an open crowdsourcing platform that aims to employ machine and human intelligence in order to improve the digital cultural content metadata quality.**

***Index Terms*—crowdsourcing, automatic enrichment, user validation, annotations model**

I. Introduction

Cultural Heritage (CH) includes the sites, things, and practices a society regards as old, important, and worthy of conservation. It has been the the subject of increasing popular and scholarly attention worldwide, and its conceptual scope is expanding. In recent years, the Cultural Heritage 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. In addition, the evolution of web technologies, has contributed in making the Web the core platform for the circulation, distribution and consumption of a broad range of cultural content.

Initiatives aiming to aggregate digital cultural content in national and international level and make it easily available to cultural and creative sectors have appear during the last decades, amongst which the Europeana[1] and the Digital Public Library of America[2] stand out. They operate as cross-domain hubs, making content accessible to users, readily available for search and study and reuse through creative applications and web services. But although their main strength lays in the vast number of the items they contain, their main weakness is the lack of structured and rich descriptive metadata and/or the insufficient metadata quality. Such problem highly affects the accessibility, visibility and dissemination range of the available digital content, also limiting the usability and the potential of added-value services and applications that re-use these resources in innovative ways, limiting also the user experience.

Metadata quality improvement and enrichment are major challenges that receive increasing attention in the digital cultural heritage domain. They have been traditionally manual processes facing the problem of scale, since improving or even adding new metadata to hundreds of thousands or even millions of records coming from different sources requires significant investment in time, effort and resources, which cannot usually be afforded by aggregators and cultural heritage institutions. The bottleneck of scale in CH metadata enrichment can be currently surpassed owing to the evolution of Artificial Intelligence (AI) and the rise of crowdsourcing initiatives. The latest advancement in AI and Machine Learning (ML) technologies facilitate the metadata enrichment process by providing capabilities of ingesting and analyzing almost any amount and type of data. In addition, crowdsourcing initiatives and campaigns, viewed as efforts of harnessing the crowd or the potential of the crowd to solve complex problems at scales and rates that no one individual can, have proved to be a powerful tool to obtain input from the crowd and assist metadata enrichment. In this context, metadata enrichment services through automated analysis and feature extraction along with crowdsourcing annotation services available in a centralized way through a dedicated platform can offer a remarkable opportunity for improving the metadata quality of digital content stored in platforms such as Europeana while at the same time engaging users and raising awareness about cultural heritage assets.

In this paper we present CrowdHeritage, an open crowdsourcing platform that aims to employ machine and human intelligence in order to improve the metadata quality of digital cultural heritage collections available in the Europeana potal. Specifically, CrowdHeritage, utilizes advanced artificial intelligence automated process to extract metadata from CH

This work has been co-financed through (i) the Connecting Europe Facility (CEF) project *CrowdHeritage: CrowdSourcing Platform for Enriching Europeana Metadata* (No INEA/CEF/ICT/A2017/1564820), (ii) the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH – CREATE – INNOVATE (T1EDK-01728 – *ANTIKLEIA*) and (iii) the European Union's Horizon 2020 research and innovation project ”*CultureLabs*: recipes for social innovation” (No 770158).

[1] https://www.europeana.eu/portal/en

[2] https://dp.la/

content and exploits the power of the crowd by mobilising users to execute useful tasks for the enrichment and validation of selected cultural heritage metadata of Europeana. The users are engaged through crowdsourcing campaigns and 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 remaining of this paper describes the CrowdHeritage platform, Its functionalities and its campaigns so far.

II. Platform Architecture and Functionalities

The development of the platform was conducted in collaboration with stakeholders who were involved in the definition of the user scenarios and evaluation tasks by actively participating in the the necessary discussions in order to concretize the functional requirements for the technological platform, identify the content for the campaigns and carefully shape every detail regarding the campaigns' execution. Cultural Heritage Institutions (CHIs) and associations such as the ModeMuseum Antwerpen[3], Network of European Museum Organisations[4] and the Philharmonie de Paris, strongly influenced the objectives and outcomes of CrowdHeritage, and broader target audiences (students, teachers) provided useful feedback for the proper functioning of the CrowdHeritage platform. The development of the platform was implemented in line with the agile principles, where three versions of functional requirements were developed, each one directing the next development sprint and taking into account the evaluation results of the previous iteration.

A. Platform overview

The platform consists of three basic components: (i) the content aggregation and collection management system (ii) the Crowdsourcing Web Spaces where end-users can navigate through selected collections of cultural records and add different annotation types and (iii) the administrative user interfaces facilitating the design, customization and validation process of the campaigns. The system architecture and the connections between the different components is illustrated in figure 4.

The platform is fully compatible with Europeana by following its standards [1] for modeling the CH records and Annotations and employing its services[5] to retrieve cultural content and publish the enrichments. In particular, the platform is connected to the Europeana Search and Record API and requests for specific cultural resources stored in Europeana in order to make them available for crowdsourcing. After the enrichment of the cultural records by the crowd, the annotations are validated by the campaign organizers and sent back to Europeana through the Europeana Annotation API.

The backend layer of the platform is built on top of existing technologies for the aggregation and reuse of Cultural Heritage content form Europeana. i.e. the WITH platform [2] [3], since the latter is already providing access to digital CH resources along with management services. The WITH database layer was extended in order to support the storage of Annotations and the backend of the application was also interlinked with the Europeana Annotation API to publish the produced Annotations to the Server.

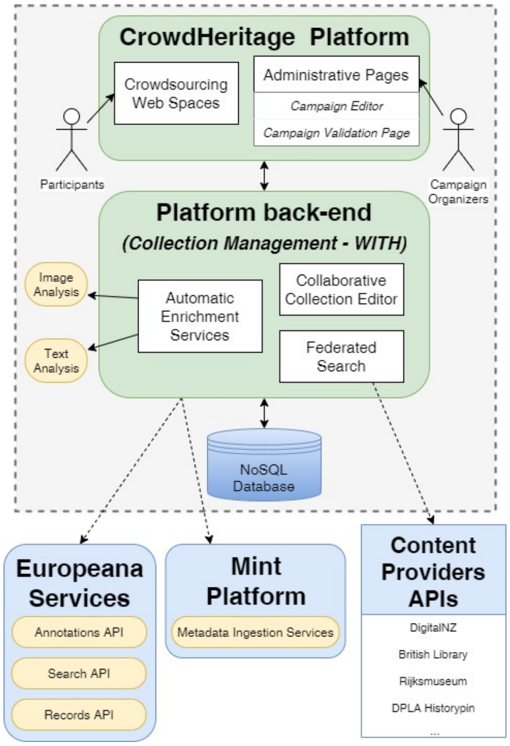
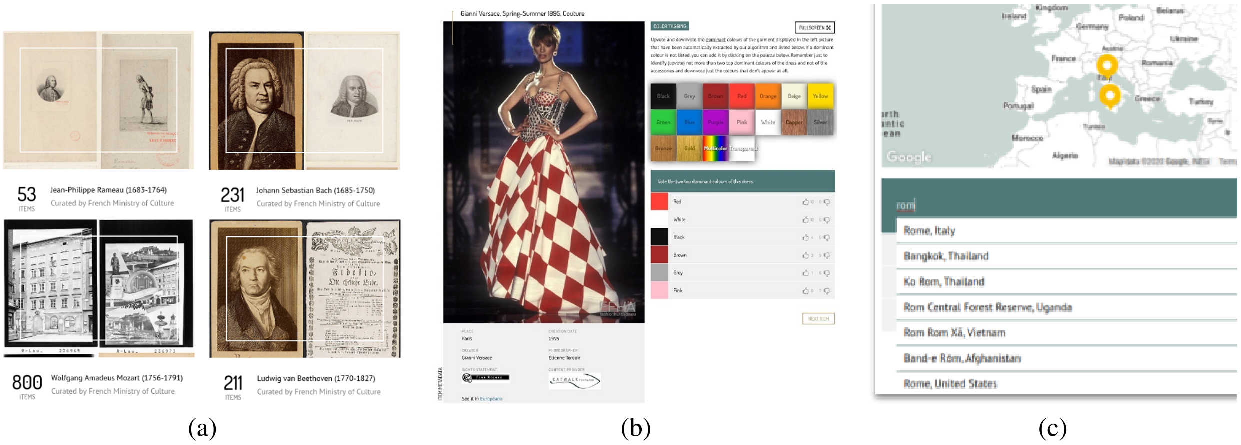


Fig. 1. CrowdHeritage System overview

The CH content is aggregated through federated and faceted search services, allowing for the simultaneous search of multiple searchable CH repositories such as Europeana, DPLA and Rijksmuseum, giving access to a huge set of heterogeneous items (images, videos, different metadata schemata etc). The integration with MINT [4] [5] offers an alternative data import mechanism, providing workflows for the ingestion, formal mapping and transformation of metadata records that are not hosted in the popular CH repositories. The aggregated content is converted into a homogeneous data model based on EDM [1], organized into thematic collections by one or more collaborators and stored in a NoSQL database making them available for crowdsourcing via the Campaign Editor. The content can be further enriched with the use of automatic enrichment methods, described at section IV.

The CrowdHeritage platform was designed to fully support multilingualism by providing translated texts in English, Italian and French for the platform's text, campaigns' descriptions and instructions. The online thesauri and vocabularies integrated into the platform were also chosen to be multilingual so as the annotating process could support translated tags in the three above languages.

[3] https://www.momu.be/en/
[4] https://www.ne-mo.org/
[5] https://pro.europeana.eu/page/apis



Upvote and downvote the dominant colours of the garment displayed in the left picture that have been automatically extracted by our algorithm and listed below. If a dominant colour is not listed, you can add it by clicking on the palette below. Remember just to identify (upvote) not more than two top dominant colours of the dress and not of the accessories and downvote just the colours that don't appear at all.

Vote the two top dominant colours of this dress.

Fig. 2. CrowdHeritage annotation process: (a) collections, (b) color-tagging and (c) geo-tagging.

B. Use cases and functionalities for end users

The registered CrowdHeritage users are able to contribute in whichever running campaigns, via a simple and user-friendly interface, with a quick learning curve, without any professional knowledge for most of the campaigns. The content in the platform is organized under thematic collections of cultural records enabling end-users to navigate, choose a collection, browse through the records and their metadata and select the ones to enrich. The platform facilitates the semantic annotation of records with terms form controlled online vocabularies and thesauri, color-tagging and geotagging items by pinpointing their location.

CrowdHeritage also provides information and statistics about each campaign, like its percentage-based progress depending on the set annotation goal, the total count of the contributors and gamification elements such as leaderboards consisting of the most active users, making crowdsourcing a user-friendly and engaging experience. The user is encouraged to add more annotations and gain two points for new annotations and one point for an up-vote or down-vote pursuing the gold, silver or bronze badge which, depending on the campaign, can be accompanied with a prize.

In every campaign page, CrowdHeritage provides statistics for each user regarding their contribution on the campaign, e.g. total number of new annotations, down-voted or up-voted, the number of digital cultural objects they have annotated, and their ranking in the campaign leaderboard, based on the awarded points for their contribution and determining the badge they have earned. Furthemore, the user karma points are calculated i.e the percentage of inserted annotations which have not been down-voted by other users, defining a quick way of identifying malicious users who may insert quick and unrelated annotations in order to gain points.

Non logged-in users can browse the list of the available campaigns with basic information: banner, title, description, thumbnail, start date, end date, contributors, annotation target, current number of annotations, and percentage of completion. They can filter the list of campaigns according to their status: active, upcoming, and complete and sort the campaigns alphabetically or by ascending the start date. By opening a campaign, more detailed information and statistics appear as well as grid with the collections available for crowdsourcing containing their records, visualization (photo, video, sound), metadata and existing annotations. The campaign leaderboard is also visible illustrating 12 most active users of the campaign with their rank, avatar, name and points.

The end-users who are logged in the platform can also contribute to the campaigns, meaning that they can add new labels to the records. In color-tagging, this is accomplished by selecting the desired color from the palette of available colors. In tagging and geotagging the contributors can tag the records by typing the desired tag into the relevant text field, displaying a list of suggested terms derived from the relevant thesauri or vocabularies. The users can validate existing annotations by up-voting or down-voting them depending on whether or not they agree with them or even delete their own annotations. The annotation process for the end-users is illustrated in figure 2.

C. Administrative use cases and functionalities

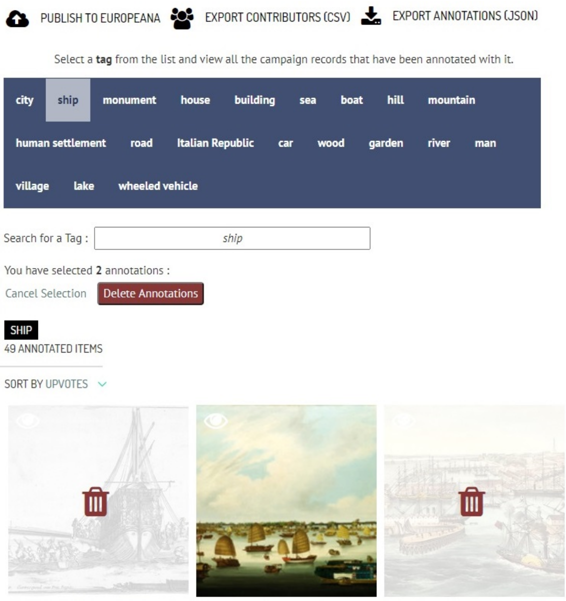
The platform also offers administrative functionalities for the campaign organizers such as a custom campaign editor and an annotation validation interface. Via the platform, a group can launch its own customized crowdsourcing campaign, in order to procure annotations (tagging, color tagging, geo-location tagging) from the public, for digital cultural collections the group has selected and finally moderate the campaign results by validating the produced annotations.

The campaign organizers can launch ad-hoc crowdsourcing campaigns through the campaign editor, which enables the creation, editing, deletion and preview of custom crowdsourcing campaigns. The campaign editor provides access to various campaign parameters related to its appearance, annotation process and content thus enabling the organizers to define the basic features of the campaign i.e. title, description and duration, choose banner, populate the campaign with collections either by directly importing a Europeana collection or by searching into Europeana and curating their own collections. Subsequently, they are able to design the desired annotation process, by setting a target for the campaign, selecting the vocabularies and thesauri for the derived annotations and choosing the type of annotations they need from tagging, geotagging or color-tagging. They can also compile the instructions for participants and describe the prizes for the top three contributors.

The campaign organizers can access the validation interface where they can moderate the crowd produced annotations. Even though the crowd validates the annotations, the moderation step by the organizers was deemed necessary. In some cases the annotations required expert knowledge and even some trivial cases can cause ambiguity e.g. the dominant colour of an outfit. Furthermore, the moderation step aims to remove some correct yet unhelpful information about the records which were tagged with obvious but general annotations (e.g. womenswear). Through the validation interface, the organizers can view the popular tags of campaigns, click on them, find out the records tagged with each term and untag the irrelevant records assuring useful and valid annotations. The process is depicted in figure 3.

III. Annotations metadata model

Platform users can add annotations to cultural heritage objects. An annotation is a note/comment associated with a resource that can be added on top of the resource without modifying the resource itself. The project's annotation model is based on W3C's Web Annotation Model [6], which has a structured model and format to enable annotations to be shared and reused across different hardware and software platforms. The CrowdHeritage annotations are all derived from a thesaurus or a vocabulary such as Wikidata[6]. By using vocabularies, we actually link the cultural records with unique URIs, which are multilingual and have specific semantics.


PUBLISH TO EUROPEANA EXPORT CONTRIBUTORS (CSV) EXPORT ANNOTATIONS (JSON)

Select a **tag** from the list and view all the campaign records that have been annotated with it.

city ship monument house building sea boat hill mountain human settlement road Italian Republic car wood garden river man village lake wheeled vehicle

Search for a Tag : ship

You have selected **2** annotations :

Cancel Selection Delete Annotations

SHIP
49 ANNOTATED ITEMS

SORT BY UPVOTES

Fig. 3. Validation page

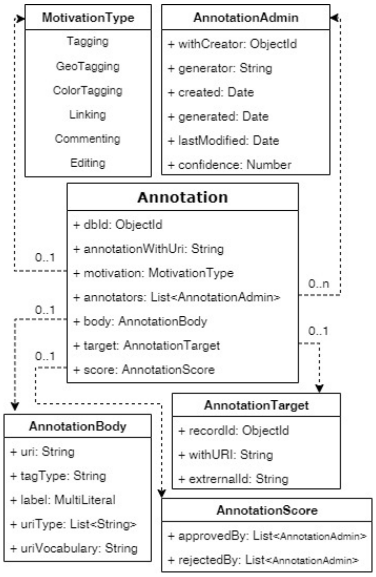


Fig. 4. Annotation Class Diagram

In brief, the platform's annotation consists of an id, a motivation, 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 record which the particular annotation relates to the body resource. Finally, the list of scores holds information about the users that have up-voted or down-voted the particular annotation. The annotation model used in the platform is depicted in a class diagram, at Figure 4.

IV. Interlinking with AI techniques that produce automatic annotations

Artificial Intelligence tools have been integrated into CrowdHeritage in order to automatically annotate big amounts of records that would require an important amount of time if done manually. Therefore, some campaigns can have a significant head start because most of the required annotations are already provided to the users and they only have to evaluate them, saving a remarkable amount of time and resources that are needed to create the annotations.

A. Automatic Color Extractor (ACE)

ACE [7] is a smart system that recognizes the main color of an image. It can detect tones of all the color spectrum and it is very useful in fashion where the color of a piece is a primary categorization factor. ACE was used to extract the dominant color of fashion items from catwalk photographs, and then the users evaluated the information that was extracted from the smart tool.

[6] https://www.wikidata.org

B. Graphical Entity Extraction Kit (GEEK)

GEEK [8] is a named entity recognition and disambiguation tool that extracts named entities in text and links them to a knowledge base using a graph-based method, taking into account measures of entity commonness, relatedness, and contextual similarity. It works in two steps: (i) It extracts text spans that refer to named entities, such as persons, locations, and organizations. (ii) It jointly disambiguates these named entities, by generating sets of candidate entities from external knowledge bases, and then iteratively eliminating the least likely ones, until we are left with the most likely mapping of textual mentions to their corresponding knowledge base entities. GEEK was used to extract important locations, organizations, artists, providers and many more, from the metadata of the cultural heritage records. The fields of metadata (e.g. description, creator, title, etc.) were evaluated and named entity recognition and disambiguation was performed for the fields that were considered adequate for each case.

C. CurAItor

CurAItor is a state-of-the-art deep learning system that was developed in the Artificial Intelligence and Learning Systems Laboratory of NTUA in 2020. The main part of the system consists of a deep ensemble network that uses Convolutional Neural Networks to process images of paintings and recognize their art style. The network was trained on thousands of images of artworks from two datasets: the Paintings Dataset for Recognizing the Art Movement (Pandora 18K) collection [9] and a dataset collated from publicly available fine art collections from WikiArt[7] [10]. The system is able to recognize 24 distinct art styles and can take any painting as input and return the main art style of the artwork. There have also been modifications to provide an extra functionality that includes fuzzy logic and allows the user to get a distribution of probabilities among all available classes instead of only the one that matches best. CurAItor was used in the context of CrowdHeritage to recognize automatically the art style of various paintings and was assisted by the crowd in cases that the system was not certain about the result. The fuzzy logic functionality provided the users with the top scoring options and they could evaluate the results.

V. Campaigns and evaluation

A. Campaigns setup and implementation

During the CrowdHeritage project ten trans-European crowdsourcing campaigns were organized, concerning five different themes: fashion, music, European cities landscapes, sports and fifties. The objective of the campaigns was twofold: (i) to demonstrate how the platform can be used to improve the quality of cultural items from different collections/datasets suffering from poor metadata, thus facilitating the searchability, visibility and re-use of Europeana's cultural material; and (ii) to demonstrate the user potential of the proposed tools through the engagement of different target groups (e.g., pupils, music experts, fashion researchers, broader public, etc). Each campaign lasted three months and a final price, offered by the partners, was awarded to whom had enriched the highest number of records in the campaigns' framework.

1) Fashion campaigns: In the fashion thematic area, two campaigns were conducted by EFHA[8], focusing on data quality improvement of fashion-related content. The first campaign targeted a broad audience of fashion lovers and fashionistas who were invited to perform a quite elementary task of validating the dominant colours of fashion garments in catwalk photos from the Europeana Fashion datasets, previously identified through automatic machine learning analysis. The second campaign involved mainly fashion scholars and students with the goal to improve the metadata quality of content, by providing annotations that require expert knowledge, namely adding or validating object type information related to fashion garments and linking them with Fashion Thesaurus[9] terms.

2) Music campaigns: In the music thematic area, two more crowdsourcing campaigns were organised with the assistance of the French Ministry of Culture (FMC), in close collaboration with the Philharmonie de Paris and the DOREMUS (DOing REusable MUSical data) Consortium. The target audience was music professionals (musicians, music teachers and scholars etc) as well as the broader public interested in music (amateurs and music lovers). The first campaign focused on musical instruments and the description and recognition of early musical instruments on medieval depictions from manuscripts, in order to enrich the Europeana datasets of musical instruments exploiting human knowledge using the MIMO vocabulary[10].The second music campaign was about famous composers and focused on their representation on cultural objects and images. This campaign addressed the general public and music lovers to find the right and/or the best representation of a composer and complete appropriately the corresponding metadata on Europeana.

3) European Cities and Sports campaigns: Two campaigns were conducted in collaboration with Michael Culture Association (MCA)[11] targeting the annotation of a variety of different themed European Collections, including Art, Maps and Geography, WWI, Photography as well as sport-relevant content (e.g. sports event/game, competition, hobby and sport equipment) with Wikidata terms. The target audience were both cultural sector and teachers and pupils from elementary to middle schools including the "Arcangeli" School of Arts in Bologna and two pilot schools from France. The objective was to identify the potential interest of Educational community for the platform.

4) Fifties campaigns: For the fifties campaigns, participants were requested to get involved in four thematic campaigns about the 50s by adding annotations about outfits, characteristic architecture and stylish interiors, vehicles, traffic infras-

[7] "WikiArt, Visual Art Encyclopedia", www.wikiart.org

[8] https://fashionheritage.eu/

[9] http://thesaurus.europeanafashion.eu/

[10] http://www.mimo-db.eu/InstrumentsKeywords/

[11] http://www.michael-culture.eu/

tructure, suitcases and bags as well as photographic qualities (light, contrast, shadows, perspective). The campaigns were promoted both digitally and physically by the organization of events, focus groups and assignments for the Kuleuven university students.

B. Campaigns results

For each campaign a set of goals had been specified regarding the extent of participation, the user engagement and the quantity and quality of achieved annotations, which were assessed after the campaigns' completion. Overall, the targets of every campaign were met, since the crowdsourcing endeavour aggregated more than 30,000 annotations and 80,000 validations, produced by over 300 contributors. All the campaigns' results after the organizers' moderation can be found at table I.

C. Evaluation results

The evaluation of the platform has been performed all along the development process helping with the definition of functional requirements. The evaluation criteria were the clarity of the website, its global performance and its usability as a campaign organiser and as a contributor. To evaluate the crowdsourcing activity, a questionnaire had been set up. The Likert scale was used to elaborate the questions and appreciate the user's level of satisfaction. The questions concern the clarity of the CrowdHeritage platform and campaigns, the plarform's general performances, the things users liked the most and what was missing from the platform.

The online questionnaire and the feedback from the campaign organisers and stackeholders show that most of the contributors strongly agree on the clarity, the user-friendliness and the performance of the website. More than 85% of the contributors found the tagging process easy and intuitive, confirming the multipurpose nature of the the platform, i.e. engaging with professional communities and the general public, engaging with schools and pupils for education on art and heritage, cleaning/improving the metadata available on Europeana. The platform's gamification set up with the leaderboard and the ranking methodology (karma points) has been very effective to engage users and encourage the contributors. Participants were keen to share suggestions for improving the platform, mainly on the social level. Sharing on social media, further gamification, and interaction between users with features such as forum or a comment area are some examples of improvements claimed by the contributors. On the technical level, contributors asked for more languages, and the possibility to add more tags, having a larger choice of vocabularies and free texts as well.

TABLE I
CAMPAIGN STATISTICS

Campaign theme	**Metric type**				
	Users	***Records***	***Tags***	***Upvotes***	***Downvotes***
Fashion	87	4113	11396[a]	23158	4113
Music	53	1265	560	5533	1573
Sports & Cities	67	1531	7470	23633	1034
Fifties	126	3719	11401	15433	149
All	333	10628	30827	67757	19927

[a] 1834 manual & 6549 automatic color annotations.

VI. CONCLUSIONS AND FUTURE WORK

CrowdHeritage can be used in an effective way for purposes such as education, discovery of cultural heritage or professional expertise. The platform has succeeded in leading on the same platform crowdsourcing campaigns as well as nichesourcing campaigns, e.g. it brought together experts in a field and the general public. Multilingualism of the platform is crucial to reach larger communities of users and foster common cooperation. CrowdHeritage is a major tool for curating and promoting the content available on Europeana. Furthermore, the reuse of content from Europeana made by CrowdHeritage reminds the content providers the importance of providing high quality images and good quality metadata for properly selecting the datasets for collection definition.

Based on our experience gained through the completed campaigns and the feedback gathered from participating users, the platform will be enhanced and further developed. A mobile version of the platform is scheduled to be released in the near future. Additionally, the functionalities of the platform will be further enriched by integrating more AI tools that are currently developed by the AILS group of NTUA, acting as a framework of training and optimizing machine learning algorithms through the active learning [11] methodology.

REFERENCES

[1] A. Isaac and V. Charles. Europeana data model definition, 2017.

[2] Alexandros Chortaras, Anna Christaki, Nasos Drosopoulos, Eirini Kaldeli, Maria Ralli, Anastasia Sofou, Arne Stabenau, Giorgos Stamou, and Vassilis Tzouvaras. With: Human-computer collaboration for data annotation and enrichment. pages 1117–1125, 04 2018.

[3] M. Giazitzoglou, V. Tzouvaras, A. Chortaras, E.M Alfonso, and N. Drosopoulos. The with platform: Where culture meets creativity. In Javier D. FernÃ¡ndez and Sebastian Hellmann, editors, *Proceedings of the Posters and Demos Track of the 13th International Conference on Semantic Systems - SEMANTiCS2017*, number 2044 in CEUR Workshop Proceedings, Aachen, 2017.

[4] N. Drosopoulos, V. Tzouvaras, N. Simou, A. Christaki, A. Stabenau, K. Pardalis, F. Xenikoudakis, and S. Kollias. A metadata interoperability platform. Museums and the Web, April 2012, San Diego, USA, 2012.

[5] I. Kollia, V. Tzouvaras, and G. Drosopoulos, N.and Stamou. A systemic approach for effective semantic access to cultural content. *Semantic Web*, 3(1):65–80, 2012.

[6] P. Ciccarese R. Sanderson and B. Young. Web annotation data model, 2017.

[7] C. Varytimidis, G. Tsatiris, K. Rapantzikos, and S. Kollias. A systemic approach to automatic metadata extraction from multimedia content. In *SSCI*, pages 1–7. IEEE, 2016.

[8] Alexios Mandalios, Konstantinos Tzamaloukas, Alexandros Chortaras, and Giorgos B. Stamou. Geek: Incremental graph-based entity disambiguation. In *LDOW@WWW*, 2018.

[9] C. Florea, C. Toca, and F. Gieseke. Artistic movement recognition by boosted fusion of color structure and topographic description. In *2017 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 569–577, March 2017.

[10] Sergey Karayev, Matthew Trentacoste, Helen Han, Aseem Agarwala, Trevor Darrell, Aaron Hertzmann, and Holger Winnemoeller. Recognizing image style, 2013.

[11] Burr Settles. Active learning literature survey. Computer Sciences Technical Report 1648, University of Wisconsin–Madison, 2009.