

Emerging Practices in the Cultural Heritage Domain - Engaging Users on a Large Scale

Johan Oomen
VU University Amsterdam
De Boelelaan 1105
The Netherlands
joomen@beeldengeluid.nl
+31-356773434

Lotte Belice Baltussen
Netherlands Institute for Sound and
Vision
Sumatralaan 45, Hilversum
The Netherlands
lbbaltussen@beeldengeluid.nl
+31-356773434

Sander Limonard
TNO ICT
Brassersplein 2, Delft
The Netherlands
sander.limonard@tno.nl
+31-152696900

Maarten Brinkerink
Netherlands Institute for Sound and
Vision
Sumatralaan 45, Hilversum
The Netherlands
mbrinkerink@beeldengeluid.nl
+31-356773434

Annelies van Ees
VU University Amsterdam
De Boelelaan 1105
The Netherlands
avanees@gmail.com
+31-205989898

Lora Aroyo
VU University Amsterdam
De Boelelaan 1105
The Netherlands
l.m.aroy@cs.vu.nl
+31-205989898

Just Vervaart
KRO
's-Gravelandseweg 80
Hilversum
The Netherlands
just.vervaart@kro.nl
+31-35-6713179

Kamil Afsar
Q42
Waldorpstraat 17F
Den Haag
The Netherlands
kamil@q42.nl
+31704452357

Riste Gligorov
VU University Amsterdam
De Boelelaan 1105
The Netherlands
rgligor@few.vu.nl
+31-205989898

Abstract

Cultural heritage institutions and their users are beginning to inhabit the same, shared information space. New, innovative services are launched, such as social tagging. Engaging in social tagging is beneficial for both parties, as it improves access to data and stimulates active engagement with the content. To explore the impact and success criteria of social tagging in the cultural heritage domain, a large-scale video labeling pilot was executed: Waisda?. It built on earlier work, and introduced three innovations: [i] Using gaming as method to annotate television heritage [ii] Actively seek collaboration with communities connected to the content [iii] use curated vocabularies as a means to integrate tags with professional annotations. Within a period of 7 months, 350,000 tags were added in Waisda?. An extensive evaluation was conducted, that provided input on the usability of the tags, the game design and so on. Based on this input, a roadmap for future developments towards a fully operational service was drafted.

Keywords

heritage, crowdsourcing, games with a purpose, tagging

1. Introduction

As the Web gets more “social” and as museums, libraries and archives are beginning to offer online access to digital representations of their collections, users and institutions are beginning to inhabit the same, shared information space. This is an exciting prospect, as we are now witnessing new paradigms for engaging users with our shared heritage. ‘Netizens’ (people actively involved in online communities) are using technological advances, offered by cultural heritage institutions, publishers and other commercial entities, as well as objects from a great variety of sources to shape this information space. The new paradigms imply, in many cases, the need for profound change in institutional practice. For instance, using the power of the Social Web to enrich the knowledge about our shared heritage. As a

result, republication and the reuse of heritage will be enhanced, and thus its value is increased.

This paper investigates the use of “social tagging” in the cultural heritage domain. Social tagging services are attracting a lot of attention recently as it can circumvent some of the disadvantages linked to traditional annotation practice, which revolved solely around professionals. We investigate the nature of tagging and how it is currently applied in the heritage domain, and for annotating videos. We present results from a large scale pilot with a video labeling game that was developed by one of the largest audiovisual archives in Europe. In this pilot project, different principles of both institutional and user involvement in the abovementioned ‘shared information space’ have been explored.

2. Social Tagging: a Way to Explore the Social Web

The heritage sector currently faces the challenge of reassessing its role in society. The search for this role transcends the quest for new sources for funding. As Willams writes. “The primary difference between the old and new system can perhaps best be captured by a single concept of profound ramifications: openness.” [3]. In 2005, social tagging - ad-hoc annotation by end-users – was introduced as one of the first ways to explore the possibilities, especially for heritage institutions, to make advantage of this web 2.0 functionality. Since the first experiments in this field by steve.museum, social tagging was embraced by institutions in the sector to explore how they could benefit from what ‘the crowd’ has to offer. While most of the early pilots were set up to seek ways to bridge the semantic gap, social tagging proved to offer a much richer array of potential benefits for heritage institutions. We were able to discern the following 4 motivations for heritage institutions to engage in social tagging:

1. Bridging the semantic gap between the terminology used by professionals and search terms of end users;

2. Enriching collections / cultural heritage with factual and contextualized information;
3. Increasing connectedness with the archive (community building)
4. Defining the annotation practice of the future;

The growing body of literature on this subject has validated as well as nuanced these claims in different ways. In the next sections, we sum up to what extent these motivations are valid in what contexts.

2.1 Bridging the semantic gap

One of the main reasons for archives to investigate the use of tags is that they might improve the findability of archival objects as they overcome disadvantages linked to traditional annotation practice. In traditional archival practice, professional cataloguers add keywords according to a set of specific criteria. One criterion in the audiovisual domain could for instance be that keywords should encapsulate the topical subject matter of a given programme. In many cases, these keywords are derived from controlled vocabularies such as thesauri. Professional cataloguers are familiar with the structure and contents of these resources. However, end-users are not aware with the structure and content of these resources, and thus have a disadvantage in searching. For instance, someone might query for ‘stapler’; but receives zero hits, as the more ‘broad’ concept ‘office supplies’ is used by the cataloguer.

This difference between the terms that professionals assigns from a controlled vocabulary and the search terms that the general public uses in search queries is called the semantic gap [4]. Jørgensen writes: “While natural language composes a searcher’s query, indexing languages typically employ highly precise and specific terms relevant to the community that uses the indexing language. This suggests that a closer look at the vocabulary generated in the tagging process might be useful to understanding and bridging the “semantic gap” between current indexing vocabularies and user’s natural language queries.” [5] So, one way of measuring the existence of this gap is by determining the information need of end users through the analysis of query logs of information retrieval systems, in which the terms entered by users are kept, (e.g. looking at orders). By comparing the differences between the terms from these query logs to and the keywords used by professional cataloguers the semantic gap can be determined.

Research has proven that tagging can indeed bridge the semantic gap and consequently increase the findability of archival objects. For instance, Veenstra et al. recently conducted experiments that focused on educational videos. In this experiment, participants was asked to add tags to 115 short video fragments. In the next stage, tests were performed to find out if these tags did indeed increase the findability of these videos. This was done by having various groups answer questions, the answers to which could be found in the videos. Each group was allowed to use only specific metadata that contained either professional metadata, tags, or both. The groups that were allowed to search by using the tags answered the questions most successfully, which indicates that the semantic gap can indeed be closed (further) through the use of tagging [6, 11].

In a typical tagging system, there is no explicit information about the meaning or semantics of each tag, and a user can apply new tags to an item as easily as applying older tags. As a result, terms

in folksonomies lack a hierarchical structure [13]. This ‘messy’ nature of tagging can also be seen as something positive according to some. David Weinberger’s book *Everything is Miscellaneous*’ premise is that in the third order of order, or the digital order, it is possible to create an infinite amount of tags for one object. Clay Shirky does not see a problem with this, and states that “you can extract a surprising amount of value from big messy data sets.” [9, 10] Also, although some keywords might mean more for one group of people than another, expanding the range of keywords means that multiple interpretations and viewpoints of one single object are possible. Because people usually know what they are interested in and what they are looking for, they will find their way in the ‘messy’ tag data set by specifically looking for keywords that they are interested in.

Open Linked Data can be used to improve the usefulness of the social tags. For instance, by integrating existing thesauri (“semantic tagging”) in the tagging environments, or by aligning user contributions

2.1.1 Semantic tagging

As demonstrated by Hildebrand, it is beneficial for museum professionals to use search aids that look for appropriate terms within multiple thesauri during an annotation task [36]. He performed a study with eight thesauri at the Print Room of the Rijksmuseum Amsterdam. Five of these thesauri came from sources besides the ones used by the Rijksmuseum in their current annotation practice. He writes that “The extra thesauri deployed provided mainly a quantitative addition, by providing more terms of a particular type. In some cases, the addition was more of a qualitative nature, for example where the more general WORDNET terms were better suited to describe photographs than the specific terms from ICONCLASS.” [36]

2.1.2 Tag Alignment

Linked Open Data can be used for semantic tagging (as used in the Print Room example) but also for tag alignment. Gligorov introduces a way of creating interoperability between end-user tags with professional annotations by linking (directly or indirectly) to the appropriate terms from controlled vocabularies. Gligorov uses Waisda? (see section 5) as case study [34]. This way, tagging works both ways: it closes the semantic gap, and can also help add more controlled terms to cultural heritage by aligning the tags with vocabularies.

2.2 Enriching with factual information

Opponents of using tags from the public claim that users need the authority and authenticity provided by trained professional [7]. In contrast, scholars that advocate the use of tags state that a great group of people can cumulatively be smarter than a selective group of experts. James Surowiecki underlines this statement in his book *The Wisdom of Crowds*. Individually, the ‘masses’ will not know more than the trained experts, because they simply do not have the experience or make irrational choices based on emotions. But as many case studies and experiments that Surowiecki describes have proven, “despite all these limitations, when our imperfect judgements are aggregated in the right way, our collective intelligence is often excellent.” [8] Thus, even though experts can provide a good array of keywords based on traditional taxonomies, allowing the general public to tag can result in even better keywords since cumulatively there is more knowledge out there than in an archive itself.

Through this cumulative knowledge, tagging provides the opportunity for more a more inclusive range of keywords. For instance, an archive might add the keywords to a clip from 1917 from The Netherlands that shows the construction of a bridge with ‘bridge’, ‘1917’ and ‘The Netherlands’. However, people might feel that ‘construction’, ‘black and white’ would be valuable keywords as well. It is also possible that cataloguers lack detailed factual knowledge such as the location and who the people in the clip are. Someone outside the archive might know additional factual information (like the architectural style of the bridge), and can add this as a tag. In short, tags can provide a wider and more inclusive array of keywords, and even add new factual information. The possible disadvantage of an excess of tags can be countered by the argument that this just provides more possibilities for people to retrieve archival materials.

2.3 Defining the annotation practice of the future

User annotations can eventually play a role in the future annotation workflow. The influx of born-digital information is growing at an immense pace [4]. And at the same time, large-scale digitisation projects are resulting in more accessible collections, that however lack the supporting metadata that allows users to find them. Institutions often do not have the resources to manually add metadata to all this material. Especially in the case of digitised audiovisual heritage, for which adding metadata is a particularly time-consuming undertaking. Depending on level of the description, a cataloguer spends one hour for each hour of audiovisual material when describing it in some detail, although this ratio can be four to one in cases where high-level item descriptions suffice. Thus, it is a lot of work to provide ample metadata for audiovisual materials, even more so than in the case of static objects, since the temporal nature of the material adds a level of complexity [4].

So, archives are looking for alternative ways to create annotations, for instance by using technologies such as OCR and speech recognition. Tagging can also be a helpful source [12]. Cultural heritage organizations can add these contributions from end-users to their existing metadata records. This can in turn save time and money. However, a certain level of quality needs to be assured, as visitors to institutional websites expect the information provided to be factual. In the user interface level, showing provenance is, in any case, of key importance. Users can deal with some ‘fuzzyness’ as long as the interface states that tags are generated by non-experts [33]. The use tags in search and retrieval systems is a topic of growing interest in computer science. For instance, Wartena et al. introduce a methodology that uses second order co-occurrence and a related distance measure for tag similarities. This methodology allows them to analyse large quantities of tags and filter out the most relevant ones [14]. The reader is referred elsewhere for detailed information on this so-called “tag based recommendation” [26].

Integrating tagging as part of the workflow results in a win-win situation for both heritage institutions and taggers. Institutions can use the tags to increase the findability of their collections in an cost-effective way, and the taggers themselves at the same time have unprecedented access to their cultural heritage

2.4 Connectedness

YouTube reported in March 2010 that every minute, 24 hours of video is uploaded to the world’s most popular video sharing site

[31]. Also, February 2010 was the first time Facebook attracted more online visitors in the US than Google. These are just some of many clear indications that the ‘social’ Web is now firmly established [28]. As indicated in the Introduction, archives are aware of the power of the Social Web and create services that aim to connect to their online constituencies. Tagging is an excellent way for heritage organisations to tap into the enthusiasm of external users. The mutual benefit is often underlined in the interface. For instance the tagline of Flickr The Commons¹ is “Your opportunity to contribute to describing the world’s public photo collections.” GWAP² has a similar catchphrase: “When you play a game at Gwap, you aren’t just having fun. You’re helping the world become a better place”. As the next sections will show, the tagging initiatives in the cultural heritage domain are often hugely popular [22]. For instance, they present themselves in the form of games that are fun to play, or affirm the users that they’re part of a highly valued community of experts. But other strategies are also applied. Some crowdsourcing efforts (notably Amazon Mechanical Turk³) actually pay people to perform certain tasks.

3. Related work in the heritage domain

In the next sections we discuss two distinct areas, namely the use of tagging in the cultural heritage domain and the use of tagging of videos, as these are most relevant for our pilot.

One of the first organizations to embark on a large scale project with tagging is the Powerhouse Museum in Sydney [16]. In June 2006, the museum offered users of their online catalogue the possibility to add tags to objects. As soon as a tag is submitted by a user, it is incorporated in the online catalogue. Tags can also be corrected and removed by other users if they are deemed incorrect.⁴ Within six months, almost 4,000 tags were added to over 2,200 objects in the online catalogue. Over 500 of these tags “were deleted, edited for spelling, or removed by other users and the system administrator” [17].

In January 2008, the Library of Congress (LoC) published a set of about 3,000 pictures on Flickr, with the goal to reach out “to unknown as well as known audiences” and to collect information about these photos through the audiences’ comments and tags [15]. The launch was heavily advertised in the blogosphere and within a day after the launch of the project the pictures had been viewed over a million times. The LoC photo set on Flickr has been expanded gradually since, and still receives about 500,000 views monthly [15]. There are three ways in which users can leave information on Flickr images: either through tags, ‘stick’ notes on the images themselves (a sort of digital post-it) and by adding comments. During the first ten months of the project, the Flickr community placed more than 7,000 comments on over 2,800 pictures. People often commented on the aesthetic qualities of the pictures, but a lot of additional factual information was added as well. Within this timeframe, a total of 2,518 people left over 67,000 tags. Of these tags, 1,000 (21%) were unique. On average, 14 tags were added to each photo [15].

The LoC and Flickr teamed up to develop a communal page for other cultural heritage institutions with photography collections;

1 See http://www.flickr.com/photos/library_of_congress/.

2 See <http://www.gwap.com/>

3 <https://www.mturk.com/mturk/welcome>

4 In the online catalogue from the Powerhouse, tags that have been added to the catalogue directly and the ones that have been imported from Flickr are presented separately. See <http://www.Powerhousemuseum.com/collection/database/> for examples.

Flickr: The Commons⁵. This initiative was launched on the same date as the first photo collection on Flickr by the Library of Congress. Until now, over thirty organisations (including the Powerhouse Museum, Nationaal Archief and the Smithsonian Institution) have joined The Commons.

Another high-profile tagging initiative is *steve.museum*, which was launched in 2005.⁶ Museum professionals and other interested parties collaborate in this consortium to investigate the possibilities of using tagging as a way to make museum collections more accessible. It is one of the most thoroughly documented tagging projects to date. Steve has developed its own tagging software, based on the specific needs of the art museums that are involved in the project. Using the so-called *steve tagger*⁷ users can add tags to pictures of artworks that belong to various collections of the participating museums. A total of 1,621 people added almost 37,000 tags to objects from various museum collections through the *steve tagger* between March 2007 and July 2009. Trant writes: “of the 36,931 terms, 88% (32,609) were found to be useful in museum staff review.” [18] Almost 12,000 of these tags are unique. The level of participations from the taggers exceeded the expectations in all cases [15]. The usability of these tags, however, have not been evaluated thoroughly until now [18].

There are several examples of cultural heritage institutions that have developed picture tagging games,⁸ but their impact (in terms of visitors and the amount of tags that are collected) is smaller than the initiatives listed above.

4. Related work: audiovisual tagging

To date, there have been few tagging projects related to audiovisual material. We found no examples carried out within the heritage domain. The most well-known services are *PopVideo*⁹ and the *Yahoo! Video Tag Game*¹⁰. *PopVideo* is a project from Luis von Ahn, professor of Computer Science at Carnegie Mellon University [19]. The *Yahoo! Video Tag Game* was developed by Yahoo! Research [20]. Both of these projects were set up to make intra-video search possible by gathering user tags. People can add tags to the videos in the form of a game. When their opponents enter the same word within a certain time frame, both players receive points. The assumption is that when two separate players enter the same tag within a certain timeframe, it is a valid word to describe what can be seen or heard in that moment of the video. Because the moment a player enters a tag is registered, it becomes possible to link the tags to the specific moment they are added. Consequently, the tags have time-stamps that can be used for retrieval. This is a departure from the way general or item level descriptions.

The fact that both moving image tagging projects mentioned above are designed in the form of a game might be an indication that people need an extra motivation or incentive to add tags to moving images, as opposed to Flickr: The Commons, *steve* and

the Powerhouse Museum project. Here, people might have a stronger emotional connection to the objects; for instance because they love modern art. We might also consider the fact that it takes much more time to add tags to video (because of its temporal nature) than to still images and that therefore the form of a game stimulates people to keep engaged in the activity.

*Annotrax*¹¹ is a HTML5 based video-publishing platform, centred around annotation and opening up video content by making it possible for users to search for short fragments. One of the distinguishing features (besides the fact *Annotrax* is not based on Flash like *Video Tag Game* and *PopVideo*) is that it distinguishes between different information layers, for instance “speaker”, “subject” and “animals”. The code is available under the GPL open source license.

Late 2009, the BBC launched *Mooso*,¹² a game designed to tag music. *Mooso* is built in Ruby on Rails, and uses XMPP/Jabber Instant Messaging protocol to process real-time submissions from players and send them instant updates as they play [32]. It is the only example of a audio-only tagging game found we managed to identify.

5. Waisda?

This Section present results from a large scale pilot concerning the tagging of moving images. The pilot was initiated by the Netherlands Institute for Sound and Vision (one of Europe’s largest audiovisual archives), the VU University Amsterdam and KRO Broadcasting. The software development company Q42 built the application.

Tagging can have an even more profound impact on audiovisual archives than on libraries and museums. Annotating moving images is inherently more complex than tagging static images, since each shot can potentially have specific semantic meaning and the temporal nature of the material makes the annotation process particularly time-consuming. Professionals in the media domain are often looking for short clips with very specific content. Research has shown that they “often wish to order small fragments of broadcasts, but the audiovisual result pages show only short textual summaries of the entire audiovisual item that is returned for a search.” [4] They therefore need specific and time-related metadata to find these materials. However, this metadata is not often entered in the catalogue due to a lack of resources. Content-based retrieval (speech-to-text, visual concept detection) can offer a means to provide more fine-grained access. Several projects are developing technology in this area [21]. Using the tags of end-users (as in the case of *Waisda?*) is a second way to solve this issue.

5.1 Social tagging as an example to explore how value towards end users can be created

While exploring the benefits associated with social tagging, several authors also identified a fundamental bias at the basis of social tagging initiatives that might prove to be an obstacle for long term adoption by ‘the crowd’. Bakhshi and Throsby identify several broad categories of innovation that are common to cultural institutions [1]. One of the more fundamental ones is innovation in the creation of value towards end users. Key question here is “What kind of surplus value does the heritage sector offer that is distinct from other activities on the social web?” Despite efforts to

5 <http://www.flickr.com/commons>

6 See <http://steve.museum/>

7 See <http://tagger.steve.museum/>

8 Examples are *Mapit1418* (Nationaal Archief, <http://www.mapit1418.nl/>), *Tag!* (McCord Museum of Canadian History, <http://www.mccordmuseum.qc.ca/en/keys/games/24/>), *Tag! You’re it* (Brooklyn Museum, http://www.brooklynmuseum.org/opencollection/tag_game/start.php/).

9 See <http://www.gwap.com/>

10 See <http://videotaggame.sandbox.yahoo.com/>

11 <http://www.annotrax.nl>

12 <http://www.mooso.fm/>

Increases in the number of registered players are strongly related to targeted promotional activities. For instance, at some point, a prize was given to the people that scored the most points within a specific week. The use of the site peaked as a result. There is also a clear link between the most tagged content, and the efforts by the Dutch public broadcaster (and project partner) KRO. They promoted Waisda? through the immensely popular programme Web site of the “Farmer Wants a Wife” website. This underlines the importance of extensive external promotion of a project like this, aimed at specific target groups, in this case viewers of this reality show.

6. Methodology

The evaluation methodology combined qualitative and quantitative research techniques. Qualitative evaluation consisted of three separate activities. First of all, an online questionnaire¹⁴ (completed by 42 people) was sent out to people playing the game. It included questions on gameplay, motivational aspects and so on. Secondly, a focus group was organized. In a moderated discussion, five people elaborated on their experiences playing Waisda? and proposed suggestions for improvements. Three of the participating focus group members were already acquainted with the environment. Thirdly, usability tests have been conducted, with five subjects that never played the game. Primary aim of this work was to evaluate interface design.

Quantitative evaluation was carried on all tags added from May until November 2009. Within this period, 42,068 unique tags have been added. The total amount of tags added by players is 340,551, of which 40.3% (137,421 tags) consists of matching tags (tags added by two more players within a time frame of 10 seconds). Each tag entry is represented by an instance of a ternary relation that relates the player that entered the tag, the game the tag was entered in (which indirectly determines the video the tag was attached to), and the tag itself. Additionally, a tag entry is associated with the point in time —relative to the beginning of the video — when the tag was entered. It also includes a score computed taking into consideration agreement with other tag entries in the temporal neighborhood.

6.1 Quantitative Evaluation

The evaluation has shown – and supports earlier research on similar projects – that altruism is an important motivation for playing Waisda?. Therefore the updated ‘about’ section of the Web site emphasizes the benefits of player activity to the (public) accessibility of the content and further research on tagging. The current research on tagging shows that taggers that are explicitly invited to help an institution by tagging, are notably more active. In the case of Steve, for instance, members of the Metropolitan Museum of Art were asked to tag via the Steve Tagger, and four times more works were tagged than through the Steve Tagger in general [18]. To further promote Waisda? a strategy that targets these altruistic players should be developed. Besides that, players should be given a sense of the impact of their activity by experimenting with ways to demonstrate the usefulness of the tags for searching through the content.

Apart from altruism, the evaluation showed that the video content itself is also a motivational factor for players to play the game. The most popular channel on Waisda? contained episodes of the reality show, “Farmer Wants a Wife”; which has a weekly

viewing audience of millions. To attract a broad constituency of users, it is important to expand the diversity of the content available on Waisda?, and experiment with different types of content. [12] Research has shown a particular interest from users in popular talk shows reflecting on recent events, programmes aimed at children and historical footage. Also, it is important to keep the content fresh. For example, at the moment there are already 29 items that each contain over 2,000 tags.

Literature studies, user research and practical experience with Waisda? have shown that both intrinsic and extrinsic factors play a role in the motivation of players. The recent literature also supports the initial concept of the Waisda? project, that assumes that a game setting is a good way to motivate people to tag (audiovisual) archive material. This shows it is important to make sure that the game design also motivates players that are not particularly interested in tagging per se, or feel that in general tagging is too much of an effort. Besides that, it is crucial to provide a good game design and game play, so the altruistic players also enjoy Waisda?. Ideally, the intrinsic and extrinsic factors come together in the game and interface design.

Although Waisda? can be played in solitude (against so-called bots), user research has shown that the vast majority of players prefer playing against others. This shows the importance of a substantial and active community of players.

6.2 Quantitative Analyses and Usefulness of Tags

The majority of players (1,051, or 45.8%) added between one to ten tags. A smaller number of players (810, 35.3%) added between ten and a hundred tags, and less than half of that number (372, 16.2%) added between a hundred and a thousand tags. Only a few players added more than a thousand tags (63, 2.7%), but together, they were responsible for adding the largest number of all contributed tags. The longest session lasted about three hours, in which one player entered 3,329 tags. This indicates that a project like Waisda? shouldn’t only aim for a wide audience, but should also find a way to specifically target these ‘super taggers’ [22].

Analyses of the most recent database dump of tags shows that 5.8% of the tags match with the terms in the GTAA thesaurus the Netherlands Institute for Sound and Vision uses to classify their collection. Apart from this, 23.6% corresponds with Cornetto¹⁵, the Dutch version of the lexical database WordNet. These are conservative estimates, since plural version of words are not recognized, and neither are conjugations and words that are spelled correctly, but that are missing the appropriate diacritic.

Since only a small number of tags is present in both databases (1,135, or 2.7%), it can already be assumed that almost a third of the tags is an existing and correct word. A senior cataloguer (employee of the Netherlands Institute for Sound and Vision) has assessed the usefulness of the tags added to two episodes. The selected episodes were the best-tagged episode (from the popular Dutch reality show mentioned before, with 19,322 tags added to it) and an episode that was tagged with an averaged number of tags (Westerman’s New World, a documentary series about a former Dutch news correspondent situated in the U.S. returning to the Netherlands, with 738 tags).

- Looking at the best-tagged episode 45% of the

14 The online service Survey Monkey was used.

15 <http://staff.science.uva.nl/%7Emdr/Research/Projects/Cornetto>

tags were deemed useful with 27% having a low and 12% having a high accuracy.

- The averagely tagged episode contained 73% tags deemed useful with 26% having a low and 19% having a high accuracy.

The senior cataloguer noted that in general the useful tags describe the material in a different way than keywords that cataloguers add do. Firstly because the tags focus on describing what is seen and heard within a programme, while the professional metadata for audiovisual content focuses on the topical subjects that a programme refers to. Apart from that, the tags also describe instances from a programme, instead of a logical segmented part and or entire episode. The fact that the tags collected by Waisda? differ from professional metadata is no surprise, and possibly even an indication that the tags contribute to bridging the semantic. However, further research on the usefulness of the tags is needed, for example by conducting search experiments with different types of end users. To describe the episode as a whole, only two tags from the top 20 most added tags of the averagely tagged episode proved to be useful. For the best-tagged episode none of the top 20 tags were deemed useful to describe the complete episode.

We found that tags added to the documentary series episode were notably more often useful than tags added to the reality show. (see fig. 3 and fig. 4 below) They were more defining (descriptive?) and specific. The reality show contained more general tags and lacked specificity. These findings contradict the assumption that the more times a tag is added to an episode, the higher the usefulness of this tag is to the audiovisual archive.

Figure 3. Usefulness of the Tags: Reality TV series

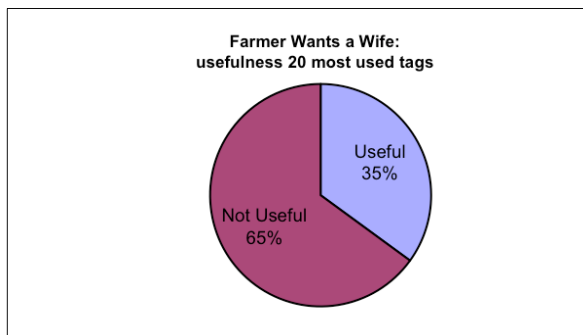
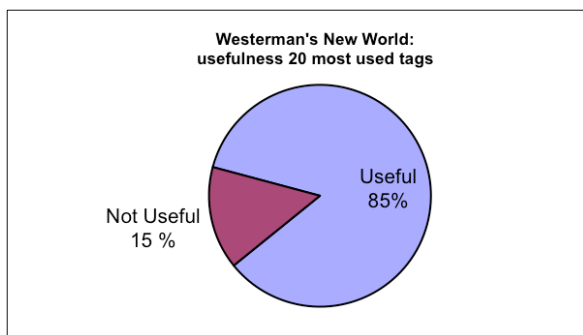


Figure 4. Usefulness of the Tags: Documentary



is also striking that in the case of the reality show more tags correspond with the GTAA thesaurus and the Cornetto lexical database, but still the tags added to the documentary series episode were deemed more useful by the professional senior cataloguer. This suggests that when a programme contains a

multitude of specific items or topics, this might result in more specific and useful tags. The way in which programme content itself influences tagging behaviour demands further research. Since the metadata for audiovisual collections mostly only describes collections on an item level, time-based metadata like tags can result in an important progress in servicing media professionals looking for specific fragments. It is therefore important to further develop this research to discover how and to what degree the tags can be used within the professional metadata in the catalogue of the Netherlands Institute for Sound and Vision.

For a more exhaustive account of the quantitative evaluation carried out, the reader is referred to the paper of Gligorov, also presented at WebSci 2010 [33].

6.3 Discussion

The transformative notion that heritage institutions and their users are now exhibiting a shared and online information space was our central notion in studying a variety of innovations currently taking place in the heritage domain. Tagging initiatives such as Waisda? proved to be a suitable platform to explore principles of both institutional and user involvement. We indicated there are at least four motivations for cultural heritage organizations to start tagging services. Some general observations from the Waisda? pilot are listed below.

For instance, the need to reach out to a critical mass of users. The success of Waisda? relied on the collaboration with existing and popular television broadcasters channels related to the content. However, the exceptional effort put into the game by a small number of users indicates that a project like Waisda? shouldn't only aim for a wide audience, but should also find a way to specifically target these 'super taggers'. Trying to leverage the power of existing social networks (i.e. Facebook, Hyves) and micro blogs (i.e. Twitter, Identi.ca) will be studied further.

It can be assumed that a third of the tags are existing and correct words. As mentioned in 6.2, this is a conservative estimate. Further research on the usefulness of the tags is needed, for example creating (directly or indirect) links to the appropriate terms from controlled vocabularies used by professionals. This work will be carried out within the scope of the PrestoPRIME IP.

Altruism proved to be a crucial motive to play the game, but altruistic players should be given a better sense of the impact of their activity by experimenting with ways to demonstrate the usefulness of the tags for searching through the content, for instance by showing a retrieval interface that is based on end-user tags. Apart from altruism, the video content itself and the gaming element have also proven to be important motivational factors. The qualitative research also showed areas for improvement regarding the game logic and game design. These are currently being implemented by the VU in an experimental prototype called TAG16. Some of the main improvements compared to the operational Waisda? environment are the introduction of a short term game goal, levels, and a so-called 'game recap'. The short term game goal is to earn more points than your co-players before the video ends. Videos in TAG are at most three minutes long instead of full-length programs and the levels that are used in TAG will 'unlock' additional video channels. The game recap shows the user how his tags can be used to search through the video he or she just tagged and this should improve the sense of

usefulness of playing the game. All these new features should motivate the player to play the game longer and more often [35].

Sound and Vision will use both the new features of TAG and the vocabulary linking work from PrestoPRIME in the design of a new operational version of Waisda?, to be launched later this year. This will be the next important step towards the implementation of video labeling as a regular service.¹⁷

Finally, we can conclude from our investigations into innovative practices in the social heritage domain that a lot can be gained for both heritage institutions and the public alike. The costs and benefits of the initiatives described in this paper will become clearer over time and some (such as Waisda?) will be developed into fully operational services that will make online collections more open and increase their findability. This will allow both existing and new users to further find, explore, and connect with their heritage.

Acknowledgements

This work is part of the PrestoPRIME IP, a project co-funded by the EU within FP7 and Images for the Future.

References

- [1] Maron, Nancy L., Smith, Kirby K. and Matthew Loy. *Sustaining Digital Resources: An On-the-Ground View of Projects Today*, 2009.
- [2] Hasan Bakhshi and David Throsby, *Innovation in Arts and Cultural Organisations*, NESTA interim research report, December 2009. <http://www.nesta.org.uk/library/documents/Innovation-in-arts-and-cultural-interim.pdf>
- [3] Williams, Anthony. *Wikinomics and the Era of Openness: European Innovation at a Crossroads*. The Lisbon Council e-brief issue 05/2010. (2010)
- [4] Bouke Huurnink, Laura Hollink, Wietske van den Heuvel, Maarten de Rijke. *The Search Behavior of Media Professionals at an Audiovisual Archive: A Transaction Log Analysis*. *Journal of the American Society for Information Science and Technology*, 2010 (to appear)
- [5] Jørgensen, Corinne, "Image Access, the Semantic Gap, and Social Tagging As a Paradigm Shift" in *Proceedings 18th Workshop of the American Society for Information Science and Technology Special Interest Group in Classification Research*, ed. Joan Lussky. <http://dlist.sir.arizona.edu/2064/> (accessed on 13 March 2010).
- [6] Melenhorst et al., "Tag-based Information Retrieval for Educational Videos," *EBU Technical Review*, Q2 (2008), p. 12, http://tech.ebu.ch/docs/techreview/trev_2008-Q2_social-tagging.pdf (accessed on 13 March 2010).
- [7] Keen, Andrew. *The Cult of the Amateur: How Today's Internet Is Killing Our Culture and Assaulting Our Economy*. London: Nicholas Brealey Publishing, 2007
- [8] Surowiecki, James, *The Wisdom of Crowds: Why the Many are Smarter than the Few* (New York: Random House, 2005): p. xiv.
- [9] Weinberger, David, *Everything is Miscellaneous: The Power of the New Digital Disorder* (New York: Henry Holt and Company, 2008).
- [10] Shirky, Clay, "Ontology is Overrated: Categories, Links and Tags," *Clay Shirky's Writings About the Internet Economics & Culture, Media & Community* http://www.shirky.com/writings/ontology_overrated.html/ (accessed on 13 March 2010).
- [11] Houten van, Ynze Abe (2009) *Searching for videos : the structure of video interaction in the framework of information foraging theory*.
- [12] Baltussen, Lotte Belice, Maarten Brinkerink and Johan Oomen, *Evaluatieverslag Waisda? Taggen van Audiovisueel Archiefmateriaal in Spelvorm*, Hilversum: Nederlands Instituut voor Beeld en Geluid, 2010, p. 7, http://www.beeldenvoordetoeekomst.nl/assets/documents/rapport_waisda_def.pdf (accessed 19 January 2010).
- [13] Golder, Scott A. Huberman, Bernardo A. (2005). "The Structure of Collaborative Tagging Systems." *Information Dynamics Lab, HP Labs*. Visited November 24, 2005.
- [14] Christian Wartena, Rogier Brussee, Martin Wibbels: *Using Tag Co-occurrence for Recommendation*. ISDA 2009: 273-278.
- [15] Springer et al., *For the Common Good: The Library of Congress Flickr Pilot Project* (Washington: Library of Congress, 2008), p. iii, http://www.loc.gov/rr/print/flickr_report_final.pdf/ (accessed on 13 March 2010).
- [16] Chan, Sebastian, "Powerhouse Museum Joins the Commons on Flickr – The What, Why and How," *Fresh + New(er) blog*, posted on 8 April 2008, <http://www.powerhousemuseum.com/dmsblog/index.php/2008/04/08/powerhouse-museum-joins-the-commons-on-flickr-the-what-why-and-how/> (accessed on 13 March 2010).
- [17] Chan, Sebastian, "Tagging and Searching – Serendipity and Museum Collection Databases," in *Museums and the Web 2007: Proceedings*, Toronto: Archives & Museum Informatics, ed. Jennifer Trant and David Bearman. Published 1 March 2007. <http://www.archimuse.com/mw2007/papers/chan/chan.html/> (accessed 9 December 2009).
- [18] Trant, Jennifer, *Tagging, Folksonomy and Art Museums: Results of steve.museum's Research – 31 January 2009*, (steve.museum / Archives & Museum Informatics: 2009): p. 18-21, [http://conference.archimuse.com/files/trantSteveResearchReport2008.pdf/](http://conference.archimuse.com/files/trantSteveResearchReport2008.pdf) (accessed on 10 October 2009).
- [19] Luis von Ahn: *Games with a Purpose*. *IEEE Computer* 39(6): 92-94 (2006)
- [20] R. van Zwol, Lluís Garcia, G. Ramirez, B. Sigurbjornsson and M. Labad. *Video tag game*. In *proceedings of the 17th International World Wide Web Conference (WWW2008)*, Beijing, China, April 2008
- [21] Laura Hollink, Bouke Huurnink, Michiel van Liempt, Johan Oomen, Annemieke de Jong, Maarten de Rijke, Guus Schreiber, Arnold Smeulders. *A Multidisciplinary Approach to Unlocking Television Broadcast Archives*. *Interdisciplinary Science Reviews*, Vol. 34 No. 2, June, 2009, 257-271. Copyright: Institute of Materials, Minerals and Mining 2009.

¹⁷ Separately, Sound and Vision is experimenting with building retrieval interfaces that are based on the tags.

<http://www.omroep.nl/subsites/mashdag/pages/de-mash-ups>

- [22] Trant, Jennifer. Tagging, Folksonomy and Art Museums: Results of steve.museum's Research – 31 January 2009. *steve.museum / Archives & Museum Informatics: 2009*. <http://conference.archimuse.com/files/trantSteveResearchReport2008.pdf>. (accessed on 13 March 2010).
- [23] Malone, Thomas M. "Heuristics for Designing Enjoyable User Interfaces: Lessons from Computer Games," in *Proceedings of the Conference on Human Factors in Computing Systems* (New York: ACM Press, 1982), p. 65.
- [24] Frank, Anders, "Balancing Three Different Foci in the Design of Serious Games: Engagement, Training Objective and Context," in *Conference Proceedings DiGRA 2007*, <http://www.digra.org/dl/db/07312.29037.pdf> (Tokyo: DiGRA, 2007), p. 567.
- [25] Hauben, M. H., Ronda. (1996). *Netizens on the history and impact of usenet and the internet*. London: Peer To Peer.
- [26] Hung, C.C., Y.C. Huang, J.Y. Hsu, D.K.C. Wu. Tag-Based User Profiling for Social Media Recommendation. <https://www.aaai.org/Papers/Workshops/2008/WS-08-06/WS08-06-006.pdf> (accessed on 13 March 2010).
- [27] O'Neil, Nick. What Exactly are Social Games? Ion July 31st, 2008. <http://www.socialtimes.com/2008/07/social-games/> (accessed March 16, 2010)
- [28] Deuze, M., & Martin, C. (2008). Aphra Kerr, The Business and Culture of Digital Games: Gamework/Gameplay; Jon Dovey and Helen W. Kennedy, Game Cultures: Computer Games as New Media. *EUROPEAN JOURNAL OF CULTURAL STUDIES*. 11 (2), 197-202.
- [29] Arthur, Charles. Facebook passes Google as most-viewed site in US in past week. <http://www.guardian.co.uk/technology/blog/2010/mar/15/facebook-passes-google-share-us>. (accessed on 13 March 2010).
- [30] Cashmore, Pete. FarmVille Surpasses 80 Million Users. <http://mashable.com/2010/02/20/farmville-80-million-users/> (accessed on 13 March 2010).
- [31] Parr, Ben. YouTube Is Huge: 24 Hours of Video Now Uploaded Every Minute. <http://mashable.com/2010/03/17/youtube-24-hours/> (accessed March 20, 2010)
- [32] Ferne, Tristan. Mooso - a game with a purpose. <http://www.bbc.co.uk/blogs/researchanddevelopment/2009/12/mooso-developing-a-tagging-gam.shtml> (accessed 20 March 2010)
- [33] Carmichael, J. et.al. "Multimodal indexing of digital audio-visual Documents: a case study for cultural heritage data" In: *Proceedings of the Sixth International Conference on Content Based Multimedia Indexing (CBMI '08)* (London: Queen Mary University, 2008)
- [34] Gligorov, Riste. Towards Integration of End-User Tags with Professional Annotations
- [35] Ees, Annelies van. Motivating people to annotate videos with a video tagging game.
- [36] Niet, Marco de. Business Model Innovatie Cultureel Erfgoed. *Digitaal Erfgoed NNederland*, 2009.
- [37] Hildebrand, M., van Ossenbruggen, J. R., Hardman, L., and Jacobs, G. Supporting Subject Matter Annotation Using Heterogeneous Thesauri, A User Study In Web Data Reuse. In: *International Journal of Human-Computer Studies* (volume 67, number 10, pages 888 - 903), October 2009
- [38] Oppenheimer, Elisabeth. Introduction: Ubiquitous Human Computing. <http://futureoftheinternet.org/introduction-ubiquitous-human-computing>