

# Analysing User Generated Content Related to Art History

Max Arends, Josef Froschauer, Doron Goldfarb and Dieter Merkl  
Institut für Softwaretechnik und Interaktive Systeme  
Technische Universität Wien  
Favoritenstrasse 9-11, 1040 Wien, Austria  
max@ec.tuwien.ac.at, josef@ec.tuwien.ac.at, doron@ec.tuwien.ac.at,  
dieter.merkl@ec.tuwien.ac.at

## ABSTRACT

For the canon of important works of art, lots of information is available on the Internet. Different Web platforms exist that show artworks with additional information and an Art historical description. While this information is suitable for experts in Art history, users without this expert knowledge may find this information hard to access because of a vocabulary mismatch between experts and laypersons. In order to allow users with different expertise to communicate and comprehend Art in their own vocabulary we have created the explorARTorium ([www.explorARTorium.info](http://www.explorARTorium.info)), which visualises artworks in context and allows users to browse along different narratives. The explorARTorium also offers the possibility to annotate artworks. We have therefore collected a Folksonomy about artworks. Applying Data Mining algorithms on this Folksonomy, we show that the reproduction of Art historical facts is possible. Moreover, we are able to analyse the extent of Art historical knowledge among people of the 21st century.

## Categories and Subject Descriptors

H.1.2 [Information Systems]: User/Machine Systems—*Human Factors, Human information processing*; J.5 [Arts and Humanities]:

## General Terms

Human Factors, Theory

## Keywords

Folksonomy, Tagging, Folksonomy Mining, Digital Libraries, Cultural Heritage Mining

## 1. INTRODUCTION

Images of artworks are available in lots of copies over the Internet and may therefore be observed by a huge amount of people. As early as 1998 [11] discussed the benefits for web-based Art education. One aspect was the tremendous storage capacity that allows information to be available to

the public. Different Web portals provide the public with open access to huge collections of Art with further background information. *Europeana*<sup>1</sup>, for example allows people to explore the digital resources of museums, and publishes information about more than 15 million items. *The Web Gallery of Art*<sup>2</sup> publishes more than 26.000 artworks and covers more than just the canon of the world most famous paintings. Recently *The Google Art Project*<sup>3</sup> was released, which gives visitors the possibility to explore museums in a "street-view" like fashion and offers high-resolution pictures of selected artworks. With the rise of the Semantic Web, it is even possible to link different kinds of collections together, as described in [4].

The second benefit of the Web for Art education mentioned in [11] is the easy access and search for the public. As the first prediction of [11] has become true and a lot of information about Art is available, easy access and search have not been made completely available. Even though all of the above mentioned Web portals provide loads of information, people without previous knowledge in Art history might find this information hard to access. The main reason might be that information is often created "by experts for experts". Like all other experts, Art historians use their own very specific vocabulary that allows them to communicate in a very precise fashion. People without knowledge about this expert vocabulary feel excluded from the discussion and have to go through a steep learning curve in order to comprehend the information provided. This obviously increases the entry barrier for laypersons to delve into Art history.

In order to bridge this communication barrier, we follow a crowd sourcing approach where people of different ages, gender and educational background are invited to annotate works of Art. Such crowdsourcing approaches have been used with collections of online museums, such as the *steve.museum*<sup>4</sup> [17], one of the first science projects that allowed users to annotate artworks. Another example is Flickr Commons<sup>5</sup>, where institutions like the Smithsonian Institute<sup>6</sup> or the Library of Congress<sup>7</sup> [16] publish pictures and encourage users to comment and annotate pictures. Oomen

<sup>1</sup><http://www.europeana.eu>

<sup>2</sup><http://www.wga.hu>

<sup>3</sup><http://googleartproject.com>

<sup>4</sup><http://www.steve.museum>

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

<sup>6</sup><http://www.si.edu>

<sup>7</sup><http://www.loc.gov>

and Aroyo [12] give a comprehensive overview regarding the opportunities and challenges of crowdsourcing in the cultural heritage domain.

We created the *explorARTorium*<sup>8</sup>, which allows users are able to intuitively browse through a collection of artworks and freely annotate, or tag, artworks. The thereby collected vocabulary later allows the users to search for artworks by using a vocabulary that is better suited for them and therefore returns the results they are actually looking for. By following the interlinked tags, users comprehend the annotations of other users and thereby get to know new keywords and terms that have not yet been familiar to them. Folksonomies also offer the possibility to get to know Art by following the links that are created from the same tags. In other fields, the combination of expert and lay vocabularies are pretty common. To give an example from academia, BibSonomy<sup>9</sup> allows the sharing of publication references and bookmarks as described in [7].

A particularly interesting aspect is the combination of an expert taxonomy with the collected Folksonomy in order to create a taxonomy of its own. As mentioned by Steward Butterfield, co-founder of the photo-sharing Website Flickr<sup>10</sup>, in [2] a Folksonomy is "like 90% of the value of a 'proper' taxonomy but 10 times simpler". By applying data mining to the provided annotations we are able to draw a conclusion about our users perception of Art. We analyse which (Art) historical figures are identified and what Art historical knowledge can be derived from a Folksonomy. Furthermore we give an insight into the Art historical understanding of people from the 21st Century.

Folksonomies and especially the application of data mining on Folksonomies is a relatively young application. A good overview on data mining on Folksonomies is given in [6], where data mining algorithms are applied to Folksonomies in order to detect spam, provide ranking and recommendation. As data source the academic bookmarking system Bibsonomy is used. The Application of association rules in Folksonomies in order to analyse and structure Folksonomies is discussed in [14]. Here, the social bookmarking system del.icio.us is used as a data source. As a description and visualization of meta-data from the Cultural Heritage area, which is enhanced by human feedback, a variety of applications exist. One example is the social online game ARTigo<sup>11</sup>, where two players provide tags for artworks in a playful environment. An example for using tag-clouds and to provide feedback and create new links between data entities is described in [15].

In Section 2 we describe our Web platform *explorARTorium* and how we combined it with the information provided by the Web Gallery of Art. In Section 3 we process the collected data and point out a statistically interesting elaboration of the collected tags. In Section 4 we describe which Art historical information can be derived from the gathered Folksonomy. Finally, Section 5 concludes the paper.

<sup>8</sup><http://www.explorARTorium.info>

<sup>9</sup><http://www.bibsonomy.org>

<sup>10</sup><http://www.flickr.com>

<sup>11</sup><http://www.artigo.org>

## 2. SYSTEM SETUP

We created the *explorARTorium* in order to combine a system that presents an Art historical taxonomy that was created by experts and offers the possibility for users to annotate artworks. We have focused on providing an intuitive web-based user interface that allows users with different expertise in Art history to navigate through the collection according to their interests.

Our first approach was the so-called *Tagging-tool*<sup>12</sup>. The main idea of the *Tagging-tool* was to allow users to make completely free associations without being influenced by Art historical information about a painting. Therefore only the artwork itself is presented without any further background information like the name of the artist, title of the artwork, and reference to the geographical region the artist originated from, etc. The user thereby annotates the artworks only based on her feelings, perception and background knowledge. Users are not only able to insert tags on their own; they are also able to see tags that were already inserted by other users. This was not done automatically - users had to click a button in order to see previously inserted tags. When doing so, the action is logged in the system in order to identify which tags were inserted without previously seeing the other tags. Our assumption was that by seeing the previously inserted keywords, users will be looking closer at details of a picture and provide more distinct annotations. After analysing the significance of the provided tags, however, we found the attributed tags are not more distinct.

In order to measure the significance of the tags we looked at how often tags were associated with artworks. The less frequent they are used, so our assumption, the more significant they are. A tag like *mann* (man), which is used 1,334 times is therefore less significant than the tag *prozession* (procession) which was used 25 times. The average amount of occurrences of tags which are immediately shown to the user is 3.40, while the average amount of occurrences for tags that are displayed after the user chooses to see all annotated tags, is 3.44. Therefore we cannot say that there is a difference in the significance of tags that are inserted when users see previously inserted tags. We kept this in mind for further updates of our systems.

Between October 2010 and January 2011 we were able to collect over 80,000 tags with the *Tagging-tool*. However, since the *Tagging-tool* returns no feedback to the users, the users got less interested in annotating pictures. We therefore started an interview-session with some of our most active users and talked with them about what features they enjoyed about the *Tagging-tool* and what they wanted to have improved. It was the consensus of the users that they would appreciate a system that provides more interaction and returns more information about the artworks. We also discussed ways to visualise information with Art historians which lead to the idea that it would be best to present each artwork in context with other artworks. With the feedback we received we have created a more interactive system – the *explorARTorium*.

<sup>12</sup><http://vsem.ec.tuwien.ac.at/taggingtool/>

## 2.1 explorARTorium

The aim of the *explorARTorium* is to offer an interactive environment that allows users without previous knowledge in Art history to navigate through Art history. At the same time users with existing Art historical knowledge shall also be challenged in using the tool. In order to suit both needs we make use of contextualisation. The selected artwork is put into context with (possibly) 25 other artworks – five of each category. Categories are the artist name, the title of the artwork, the school (Italian, Dutch, German, etc.) and the motive (like portrait, religious, mythological, etc). Additionally we randomly show five others so that the user is able to escape from the current context and browse to a completely different artwork. All the contextualised artworks are selectable – as soon as a users clicks on an artwork, the artwork is again put in context with other artworks.

The contextualisation of artworks shall make it understandable to the user, that each artwork is part of a greater environment. Images can thus be put into context with other images and provide a way for visitors to move intuitively through Art history. Further information such as the artist name, title, region, school etc are presented in a fold-out menu. That way, users are not distracted by textual information and again able to examine and possibly annotate the artworks completely freely. Also an artist name that is not familiar to the user might have a feeling of intimidation and therefore counteract with the basic approach. Additionally the already assigned tags are displayed and provide interaction. This way we make use of serendipity, as described in [1] where users are invited to browse along a narrative path between artworks that share the same set of tags and, thus, get the chance of discovering other artworks. By clicking on one tag, we show five randomly chosen pictures that are also associated with the same tag. By clicking on another tag, a user is able to combine different tags with a logic AND. Thereby it is possible to search for images without prior knowledge of names, titles or categories. Additionally we provide the users with the possibility to rate a picture on a scale from 0 to 5 to give an indication about how much they like the picture. This additional information will in later stages be used to build an artwork recommender on top of the *explorARTorium*.

Figure 1 shows a screenshot of the *explorARTorium*. On the left side the currently selected image is shown and the 5x5 matrix on the right provides the context for the selected artwork. Directly underneath the pictures the expert textual information is provided in the fold-out menus. Just below that the user-generated-context section is located. The already annotated tags are displayed with the possibility to provide more tags. The tags can be selected in order to create a selection of artworks from a combination of tags. At the time of taking this screenshot, the following tags were assigned for the artwork: *beautiful, frau, headcloth, kopftuch, portrait, ring* and *woman*. These tags were provided by two different users in German and English. The possibility to have multilingual tags associated with an artwork is an advantages of collecting user-generated content, as it allows users from different languages to explore and search the collection.

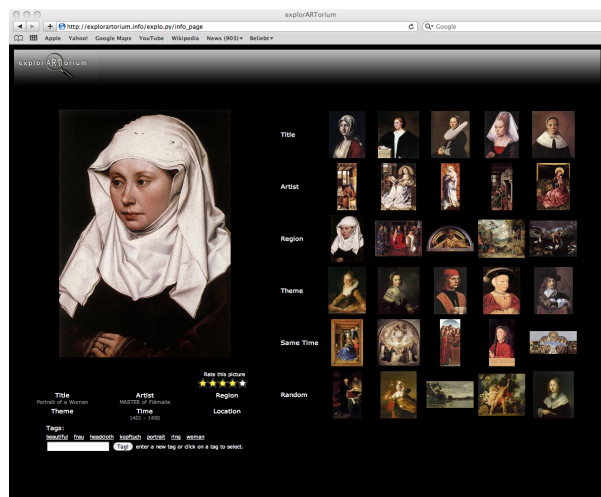


Figure 1: Screenshot of the *explorARTorium*.

## 2.2 Data Source

As a data-source we chose the Web Gallery of Art (WGA), an open available data-source on the Web. The WGA offers a collection of about 26,000 European works of Art from over 3,000 artists and a time span between 1000 AD until 1900 AD. Besides the pictures themselves, basic curatorial information is available for each artwork. As an example consider Figure 2. In this case we show the information that is available from the WGA for a painting by Frans Hals<sup>13</sup>. This meta-information for each artwork consists of biographical information about the artist, like the date and place of birth and death, title, date of creation, size of artwork, and location of the artwork, the school of the painting like Italian, Dutch, Flemish, etc. and information about its motive. In order to work properly with this information, we transformed the information from the Web Gallery of Art into the CIDOC CRM, "a formal ontology intended to facilitate the integration, mediation and interchange of heterogeneous cultural heritage information" [3].



```
HALS; Frans ; (b. 1580,  
Antwerpen, d. 1666, Haarlem) ;  
Portrait of a Man ; c. 1655 ;  
Oil on canvas ; 108 x 80 cm ;  
Liechtenstein Museum ; Vienna  
; http://www.wga.hu/html/h  
/hals/frans//08-  
1658/78noport.html ; painting  
; portrait ; Dutch ; 1601 -  
1650
```

Figure 2: "Portrait of a Man", Frans Hals, 1655 and the information as it is available on the Web Gallery of Art

## 3. PROCESSING THE COLLECTED DATA

Between October 2010 and March 2011 we have collected more than 90,000 tags for more than 10,000 images that have been tagged by more than 120 different users. However, we

<sup>13</sup><http://www.wga.hu/art/h/hals/frans/08-1658/73noport.jpg>

did not collect demographic data from our users. Since users are encouraged to annotate the images in any language, they have created a multilingual Folksonomy. The distribution is roughly two thirds German, one fourth English and the rest is annotated in other languages. In the current state of the evaluation, all languages are used equally and no translation is done from language to the other. Therefore some tags with identical meaning are not considered as identical. In further research we will aim to separate the tags in different languages.

The annotated tags are then standardised as discussed in [9]. All spaces and special characters are removed, and each tag is transformed into lower case letters and concatenated so that it consists of one single word. Furthermore German umlauts are transformed into combinations of vowels, such that for example "ä" is transcribed as "ae". We thereby improve the matching between similar tags.

In the current analysis, various word stems are not taken into consideration. Hence we do differentiate between singular and plural. In future work we plan to use stemming algorithms in order to possibly improve results.

### 3.1 Identifying the Overall Vocabulary

By asking people of different expertise, age, gender and educational background to annotate artworks, we have collected a Folksonomy. As of March 24th 2011, we have collected exactly 93,906 tags, of which are 14,931 distinct, for 10,823 pictures. That means that each picture is annotated by an average of 8.67 tags, median of 7 and standard deviation of 6.5. Note, however, that artworks might be annotated using one, two or more languages.

Table 1 shows the most frequently assigned tags. Since most of the collected tags are German, we will also give an English translation in brackets. The 5 tags together make up around 7% of the total amount of tags and were annotated by 47 users. We can therefore argue that these tags are part of a more common vocabulary which is familiar to a lot of users. Please note the average number of tags per artwork is astonishingly diverse among the most frequently used tags. So, for instance, an artwork tagged with *wolken* (clouds) or *himmel* (heaven, sky) has on average more than 15 additional tags. In contrast, an artwork tagged with *mann* (man) or *frau* (woman) is on average annotated with 8 additional tags. A possible explanation might be, that in the latter group people are the main subject of the artwork with fewer additional items to observe. Artworks that are associated with the tags *wolken* and *himmel* probably deal with landscapes and thus are richer on additional items to be observed by users. A deeper analysis of this observation is part of our future work.

Out of the 14.931 uniquely annotated tags, 135 tags are annotated more than 100 times. 1,324 tags are annotated at least ten times and 5,545 tags are annotated at least twice. Therefore there are 9,386 unique tags that have been annotated only once. One reason for this might be typic. Another reason might be that these tags are very specific. Further improvements on the analysis will focus on increasing the matching of these tags, so that they can be better included in future analysis.

In Figure 3 we show the distribution of the tags. On the x-axis the distinct tags are plotted (about 15,000), the y-axis shows their respective frequency of occurrence in our collection of artworks. Since there are so few frequently used tags, we use a logarithmic scale. The number of very frequently used tags drops very fast as shown in the diagram.

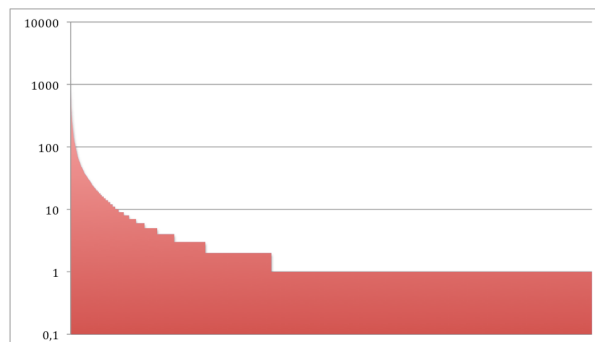


Figure 3: Tag distribution according to number of occurrences.

### 3.2 Identifying the Vocabulary of Users

Basically the collection of all annotations a user has made for the artworks is called Personomy, while the collection of all these Personomies is called Folksonomy [8]. Among the vocabularies of users there is a high amount of variation. Besides the fact that users annotate in different languages the vocabularies of different users vary a lot. By looking at the vocabulary of each user it is possible to make findings about how people perceive Art and what aspects and details they focus on.

In order to visualize the Personomies we make use of tag-clouds. Tag-clouds are a simple and widely used visual interface model on the Web. Basically a tag-cloud is a list of tags that is visually weighted by font size [5]. The larger the used font is, the more frequently a tag has been used. The presented tag-clouds are alphabetically sorted to make it easier to compare and look for certain keywords.

In order to compare two vocabularies, we chose the tags of the two users that have provided most tags. We only look at tags from artworks, that were tagged by both of these users. Figure 4 shows tag-clouds that represent the vocabularies of the two users by visualizing the 150 most frequently used tags. User A seems to be more focused on landscape details of artworks - the tags *wolken* (clouds), *himmel* (heaven, sky) and *baeume* (trees) are used frequently. User B seems to be more focused on details of humans. *vollbart* (full beard), *hut* (hat), *nackt* (naked) and *kopftuch* (headscarf) all focus on the look of people.

The difference in perception of the users becomes even clearer when we look at their uniquely used tags. By showing the exclusively used tags of User A and User B for the identical pictures it is possible to more closely describe their respective observational preferences. Again, it is obvious to see that the focus varies a lot among both users, as seen in Figure 5. In the tag-cloud of User A we frequently see exclusively used tags like *kapitelle* (capital, cap), *schwarz-weiss* (back-and-white) and *duester* (dusky). In contrast User B







Figure 6: The tag-clouds visualise the most frequently used tags of a region. Top: Italian Art; Bottom: Dutch Art

## 4.2 What (Art) Historical Figures are People Able to Identify?

After giving a general overview on what Art historical information can be derived from a Folksonomy we will focus on the question whether the users of the *explorARTorium* are able to identify historical and iconographical figures.

Our data source includes paintings of historical figures. Even though we have not explicitly asked the users to identify the depicted people, users frequently annotate the name of the historical figures. To give a few examples: Napoleon has been identified in 9 out of 17 cases; Maximilian I has been recognized in 1 out of 3 cases, whereas the Empress Josephine of France has never been identified in any of the 10 artworks that portray her and that have already been tagged. A reason for this observation might be that various historical figures are present in the memories of 21st century people to a highly different extent. We will focus on that issue in future research.

Out of all Christian figures, people are mostly able to identify Jesus and Virgin Mary. Virgin Mary has been annotated as *maria* 787 times and as *madonna* 251 times. Jesus has been identified with the tag *jesus* in 685 and *jesukind* (young jesus) in 366 cases. However, we have not checked manually in how many cases Virgin and Mary are definitely portrayed in those artworks. Besides Jesus and Mary other Christian figures are identified as well. St. Joseph is tagged about 180 times, Judas is tagged 12 times and St. Anne is tagged only 5 times.

The next step to understand the perception of figures by our users is to look at the titles of artworks, in which the users were able to identify the figures. Taking the above examples of Judas, there are basically two scenes in which our users were able to identify a person as Judas. One is "The Betrayal of Christ" and the other is "The Last Supper". For the latter scene this is particularly interesting, since there are 11 other apostles present, but none of our users was able to identify

any of them. The 5 tags of St. Anne were all annotated in scenes that show her together with Virgin Mary and Jesus.

Thus it may seem that there are very few religious and historical figures that people are commonly familiar with. As Virgin Mary and Jesus are often recognised, other biblical figures are very rarely identified by users – and probably by users with a better knowledge about Art History.

## 4.3 What Does the Folksonomy Tell Us About Iconography?

Iconography is a branch of Art history that has to do with the "identification, description and interpretation of the content of images" [20]. Especially in Christian Art biblical scenes found in paintings are highly standardised so that they can be easily identified. As an example for the perception of iconographical figures we will use the theme of "Adoration of the Magi".

Adoration of the Magi is the title of paintings that show the subject of Nativity of Jesus at the moment the three kings, in earlier times called Magi, have found Jesus after following a star [18]. The three Magi, called Melchior, Caspar and Balthazar then present their gifts: gold, incense and myrrh [13]. In our collection, we have 155 paintings titled "Adoration of the Magi". Two paintings from different centuries and different regions, one by Fra Angelico and Fra Filippo Lippi<sup>14</sup> and one by Rembrandt<sup>15</sup> are shown in Figure 7. Despite the fact that these paintings were created in different regions at different times they still share some similarities. They both show Virgin Mary holding the newborn Jesus in her arms and at least one of the kings on his knees, presenting his gifts and showing his adoration to the newborn Jesus. St. Joseph is shown standing behind Virgin Mary. Also, a donkey and an ox are present.



Figure 7: Left: "Adoration of the Magi" by Fra Angelico and Fra Filippo Lippi, 15th Century ; Right:"Adoration of the Magi" by Rembrandt, 17th Century

When analyzing all 155 paintings titled "Adoration of the Magi" we see that people are able to identify Virgin Mary and Jesus in 33,5 percent (52 out of 155) of paintings. A woman holding a child can be easily identified as Virgin

<sup>14</sup><http://upload.wikimedia.org/wikipedia/commons/e/e0/Fra-Angelico-Adoration.jpg>

<sup>15</sup><http://upload.wikimedia.org/wikipedia/commons/3/37/Magi-rembrandt.jpg>

Mary and Jesus. However, people are not as familiar with the portrayal of St. Joseph. Interestingly enough St. Joseph is identified in 17 percent of the paintings showing this particular scene in the life of Christ. By making use of association rules, we see that St. Joseph is never identified without a co-occurrence of other figures. The tags *joseph*, *josef* and *hljosef* (St. Joseph) are associated to tags like *maria*, *jesus* (or similar) with a confidence of 1 for all portraits titled "Adoration of the Magi". The reason for this obvious correlation is that people tend to remember stories, make out certain figures and draw conclusions about who the other figures might be. In our examples people with a Christian background are probably familiar with the famous iconographical subject of a woman holding a child, are able to make out Virgin Mary and the newborn Jesus. Since they also recall the story of "The Adoration of the Magi", they are able to make the connection that St. Joseph has to be present during this scene and are therefore able to identify him.

#### 4.4 Is It Possible to Derive an Iconographic Taxonomy From a Folksonomy?

As Mathes points out in [10] "examining the quantitative aspects of Folksonomies is an area that could yield some interesting data on the makeup and use the terms used to describe items" and that "the use of a Folksonomy to supplement existing classification schemes [...] is a possible area for research". By applying a quantitative analysis on the tags associated to artworks titled as "Adoration of the Magi" we create a taxonomy of keywords for these artworks. By creating item sets we are able to create a taxonomy for the theme "Adoration of the Magi". Table 2 shows a few item sets. The most obvious combination is the set with Virgin Mary, Jesus and St. Joseph. The other sets include important iconographical information, like *geschenke* (gifts) in combination with Virgin Mary and Jesus. By forming these item set for every common title from our connection, it is possible to derive a user-generated taxonomy over the pictures.

#### 4.5 Reproducing Art Historical Knowledge With a Folksonomy

As the collected Folksonomy is created on the observation of the users who are not necessarily familiar with Art historical knowledge, it is interesting to verify whether it is possible to reproduce Art historical information. Since the tag *heilighenschein* (halo) is very common in our collection, we examine the decline of the halo as an iconographical symbol over time.

Halos were used in Christian Art to identify important figures like Jesus, the Virgin Mary, Old Testament prophets, angels and saints. By portraying important figures with, and the less important without haloes, observers of an artwork were visually directed to focus on the key figures. Haloes have been used in Christian Art since the 5th century. The more realistic paintings became and the more painters made use of perspective, the smaller and increasingly subtle the haloes became. In the early 15th century painters like Jan van Eyck quit using haloes to point out key figures at all, but the use of the halo was still pretty common. The usage of halos in Art declined further until it had become completely

unusual by the 18th century [19].

By freely annotating the artworks, the users were indeed able to observe the decline of the halo in Art over time. Of course, users were not asked to search for haloes so it is not possible to draw a conclusion on the distribution of the tag *heilighenschein* (halo) alone.

By using association rules it is possible to identify the confidence between the tag *jesus* and the tag *heilighenschein* for each century – see Figure 8. For 13th century there is a confidence of 1, meaning that all identified Jesus' are portrayed with a halo. By the 18th century the confidence of co-occurrence of *heilighenschein* has reached 0, even though the tag *jesus* is annotated 10 times and the tag *heilighenschein* 14 times.

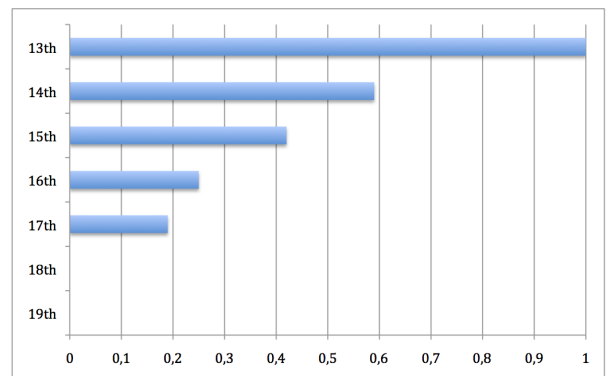


Figure 8: Confidence of co-occurrence between the tag *jesus* and the tag *heilighenschein* (halo) over the centuries.

## 5. CONCLUSION

The vocabularies of Art historical experts and laypersons often vary so much that it is difficult for them to communicate with each other. By applying the concept of a Folksonomy to artworks we aim to close this communication gap. In order to do so we take the expert knowledge from an Art database, The Web Gallery of Art, and combine it with the information of a Folksonomy. In our first approach we used an environment that only allowed users to annotate pictures while no expert information was presented. In order to provide further interaction for the users, we created the *explorARTorium*, a Web platform that allows users to observe an artwork in context with others. Users are able to understand the historical, geographical and iconographical circumstances of an artwork and have the opportunity to annotate paintings.

The thereby created Folksonomy is then analysed in order to obtain information about how people perceive Art and which (Art) historical and iconographical figures they are able to identify. We show how different (Art) historical figures are perceived, which ones are identified and how they are connected. Furthermore we answer whether Art historical knowledge can be derived from a Folksonomy. The findings in this paper show that it is indeed possible to reproduce several Art historical facts with an crowdsourcing approach and that an iconographical taxonomy can be de-

**Table 2: Samples of item sets that were created for the theme "Adoration of the Magi".**

Set 1	Set 2	Set 3	Set 4	Set 5
<i>maria</i>	<i>maria</i>	<i>maria</i>	<i>maria</i>	<i>esel</i> (donkey)
<i>jesus</i>	<i>jesus</i>	<i>jesus</i>	<i>jesus</i>	<i>ochs</i> (ox)
<i>josef</i>	<i>esel</i> (donkey)	<i>geschenke</i>	<i>melchior</i>	<i>stall</i> (stable)
			<i>weihrauch</i> (incense)	<i>gaben</i> (gifts)

rived from a Folksonomy.

Further work will include language identification and stemming to improve the analysis of the tags. An aspect of future analysis will focus on identifying the Art historical knowledge of people and to examine if they are enlarging their own vocabulary over time. This way it would be possible to show that users of the *explorARTorium* are actually picking up Art historical knowledge. The *explorARTorium* will be extended to provide additional interaction features to keep users interested in using the *explorARTorium*.

## 6. ACKNOWLEDGMENTS

This work was funded by the FWF (Fonds zur Förderung der wissenschaftlichen Forschung / Austrian Science Fund), Project No. L602, "The Virtual 3D Social Experience Museum".

## 7. REFERENCES

- [1] N. Auray. Folksonomy: The new way to serendipity. *Communications & Strategies*, 65:67–89, 2007.
- [2] S. Butterfield. Sylloge. <http://replay.waybackmachine.org/20080703165628/http://www.sylloge.com/personal/2004/08/Folksonomy-social-classification-great.html>, 2004. [Online; accessed 21-April-2011].
- [3] CIDOC. The CIDOC CRM, or ISO 21127:2006. [http://cidoc.mediahost.org/standard\\_crm%28en%29%28E1%29.xml](http://cidoc.mediahost.org/standard_crm%28en%29%28E1%29.xml), 2004. [Online; accessed 25-March-2011].
- [4] B. Haslhofer, E. M. Roochi, M. Gay, and R. Simon. Augmenting europeana content with linked data resources. In *International Conference on Semantic Systems (I-SEMANTICS)*, ACM International Conference Proceeding Series. ACM, 2010.
- [5] Y. Hassan-Montero and V. Herrero-Solana. Improving Tag-Clouds as Visual Information Retrieval Interfaces. In *Proc. International Conference on Multidisciplinary Information Sciences and Technologies (InSciT) 2006*, Merida, Spain, Oct. 2006.
- [6] A. Hotho. Data mining on folksonomies. In *Intelligent Information Access*, volume 301 of *Studies in Computational Intelligence*, pages 57–82. Springer, Berlin / Heidelberg, 2010.
- [7] A. Hotho, R. Jäschke, C. Schmitz, and G. Stumme. Bibsonomy: A social bookmark and publication sharing system. In *Proceedings of the Conceptual Structures Tool Interoperability Workshop at the 14th International Conference on Conceptual Structures*, Aalborg, Denmark, July 2006. Aalborg University Press.
- [8] A. Hotho, R. Jäschke, C. Schmitz, and G. Stumme. Trend detection in folksonomies. In *Proc. First International Conference on Semantics And Digital Media Technology (SAMT)*, volume 4306 of *LNCS*, pages 56–70, Heidelberg, Dec. 2006. Springer.
- [9] M.-H. Hsu and H.-H. Chen. Tag normalization and prediction for effective social media retrieval. In *Proceedings of the International Conference on Web Intelligence*, pages 770–774. IEEE, 2008.
- [10] A. Mathes. Folksonomies - cooperative classification and communication through shared metadata. <http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html>, Dec. 2004. [Online; accessed 27-April-2011].
- [11] S. Milekic, C. Moreno, and A. Kazee. Museums and WWW-based art education. In *Proceedings of Museums and the Web Conference 1998*. Archives Museum Informatics, April 1998.
- [12] J. Oomen and L. Aroyo. Crowdsourcing in the cultural heritage domain: Opportunities and challenges. In *5th International Conference on Communities Technologies*, March 2011.
- [13] B. Riese. *Seemanns Lexikon der Ikonografie*. E.A. Seemann Verlag, Leipzig, 2007.
- [14] C. Schmitz, A. Hotho, R. Jäschke, and G. Stumme. Mining association rules in folksonomies. In *Proc. of the 10th International Federation of Classification Societies (IFCS) Conference*, pages 261–270. Springer, 2006.
- [15] R. Simon, C. Sadilek, J. Korb, M. Baldauf, and B. Haslhofer. Tag clouds and old maps: Annotations as linked spatiotemporal data in the cultural heritage domain. In *6th International Conference on Geographic Information Science (GIScience 2010)*, Zurich, Switzerland, Sep. 2010.
- [16] M. Springer, B. Dulabahn, P. Michel, B. Natanson, D. Reser, D. Woodward, and H. Zinkham. For the Common Good: The Library of Congress Flickr Pilot Project: Report Summary. Technical report, Library of Congress, October 2008.
- [17] J. Trant. Tagging, folksonomy and art museums: Results of steve.museum's research. In *Proceedings of Museums and the Web Conference 2009*, 2009.
- [18] Wikipedia. Adoration of the magi. [http://en.wikipedia.org/wiki/Adoration\\_of\\_the\\_Magi](http://en.wikipedia.org/wiki/Adoration_of_the_Magi), 2011. [Online; Version of 14 March 2011 23:30; accessed 25-March-2011].
- [19] Wikipedia. Halo (religious iconography). [http://en.wikipedia.org/wiki/Halo\\_%28religious\\_iconography%29](http://en.wikipedia.org/wiki/Halo_%28religious_iconography%29), 2011. [Online; Version of 14 March 2011 23:30; accessed 25-March-2011].
- [20] Wikipedia. Iconography. <http://en.wikipedia.org/wiki/Iconography>, 2011. [Online; Version of 24 February 2011; accessed 25-March-2011].