Introduction

Accessing images from online digital art collections remains challenging because traditional text-based search techniques are inadequate. We explored these issues through subject categorization of user-generated tags assigned to art images and a lab-based study using a Tobii eye tracker and a simple tagging interface.

This project is part of the T3: Text, Tags, and Trust project, which explores the application of social tagging and computational linguistic techniques to enrich the metadata of images and improve access to digital collections of art images. For more information about the T3 project, please visit: http://umiacs.umd.edu/research/t3

Subject Categorization Using the Panofsky-Shatford Matrix

Overview: For this pilot project, a collection of 495 tags that were assigned to 7 images from the Indianapolis Museum of Art were categorized by at least two individuals using a matrix that describes the subject matter of an image on two axes. One axis describes specificity, or an individual's depth of knowledge about the content of an image (Generic, Specific, and Abstract). The second describes the type of subject matter expressed (Who, What, When, and Where) (Panofsky, 1962; Shatford, 1986). This core matrix was modified to include a Visual Elements category as well as an Unknown category to capture information not related to subject matter.

Findings: In those cases where there was intercoder agreement, the majority of tags (164 or 51%) were in the Generic Who category. Looking at each axis separately, 204 or 63% of the tags were categorized in the Pre-iconographic or Generic category on the specificity axis and 176 or 54% were categorized in the Who category on the type axis.

Implications: Taggers favor tags that represent the generic meaning of subjects represented in an image (for example, individuals tagged an image with “war” rather than “Civil War”). In addition, tags covering persons or things occur more frequently than those expressing events, locations, or time periods. An expanded categorization project is underway to see if machine learning techniques can be used to automatically apply these subject categorizations.

Lab-based Study Using an Eye Tracker to See Where and What People Tag

Overview: As part of a larger study on understanding tagging behaviors in relation to viewing images, subjects were shown six images, one at a time, for 1 minute and 45 seconds each. Subjects were asked to describe the images they saw by typing tags into a white box below the image. They were permitted, although not required, to use that entire time for tagging or finish
early and wait for the screen to advance. A total of 51 subjects participated in the lab study ranging in age from 20-52 years old. Thirty-two participants were female and 19 were male. Subjects came from a variety of backgrounds, including social sciences (n=19), math/engineering/computer sciences (n=16), humanities (n=9), and physical science (n=1).

Findings: Specific tags showed up earlier in the rank order of tags than other tag types. This was seen most clearly in the Virgin and Child image (ca. 1520, attributed to Simon Bening) with 21 of the 49 first tags entered consisting of “Madonna”, “Virgin”, or “Jesus”. Abstract images like Sunny Beach Life (1974, Karel Appel) received far more visual elements tags (43.3% of tags for this image) than any of the representational images. We also found that users who had seen the image before and were then asked to tag it upon second viewing provided more tags on average (mean=10.6 tags) than those who were asked to tag an image upon first review (mean=8.1 tags), significant at the p<.01 level. Users who had already seen the image were also more likely to produce visual elements tags (10.2%) than users who were looking at the image for the first time (5.7%), significant at the p<.05 level.

Implications: First, when taggers can identify a specific person or thing, they will record that tag before more generic concepts like “woman” or “baby”. Knowing this may allow for new algorithms that get us one step closer to tag disambiguation. Second, a landscape image with definable trees may better represent the color green than an abstract image that receives this tag frequently. Tag-based search engines may bias the results of a user’s search for “green” to such images that are not as well represented by their tags. Finally, the result that users tagging behaviors differed with images they had seen before provide implications for tag collectors, like Google Image Labeler, to consider new strategies for engaging users with visual content and then asking them to tag those images.

The Road Ahead…

The digitization of art has enabled libraries and museums to reach global audiences. In this context, the language barrier is more present than ever due to the multicultural nature of items in the collections and users’ linguistic diversity. One of our project's future goals is to explore multilingual tagging for improving access to the collections across boundaries.

Following the path set by the steve.museum project, we are inviting Internet users in Spain to tag a set of images that were previously tagged in English by users in the U.S. The tags will be categorized to explore the semantic similarities and differences in both languages, and the best ways to exploit the bilingual tags for improved access. The same protocol can be applied to a number of languages in incremental steps.

Selected bibliography:

This project was supported in part by Grant No. LG-30-08-0117-08 from the Institute of Museum and Library Services (IMLS). Any views, findings, conclusions or recommendations expressed in this poster do not necessarily represent those of IMLS.