

THIS IS A DRAFT WAITING FOR FINAL PUBLICATION: NOT FOR CITATION  
CIRCULATION OR COPYING

## **Flickr The Commons: what 51 million community members actually do?**

Erik Hekman, MSc (PhD researcher Crossmedialab)  
Prof. Harry van Vliet (Head of Crossmedialab)

Utrecht University of Applied Science  
Faculty of Communication and Journalism  
Padualaan 99, 3512 CH Utrecht  
Erik.hekman@hu.nl; Harry.vanvliet@hu.nl

Keywords: Flickr The Commons, cultural heritage, Web 2.0, user-generated content, enrichment

### **Abstract**

Various institutions report that joining Flickr The Commons has had a positive impact in the sense of reach and user-generated content. However, we found no empirical research to backup this statement or even research that gives a general understanding of how the Flickr community interacts with the participating institutions. By extracting a rich dataset from The Commons we constructed a general understanding on The Commons with regard to institutional and community participation. Looking at our findings we can conclude that Flickr The Commons is indeed a good alternative to share hidden treasures momentarily stored away in institutional archives and websites. We found that in total 167 thousand community members actively did something with the photos since The Commons started in January 2008. However the total amount of user-generated content is quite minimal and institutions hardly interact with the community consisting of 51 million members.

## 1. Introduction

In a recent blogpost Miller (2012) gives an impression of the current state of The Commons after four years of usage. *“The Commons started with 1,500 photos from a single institution and has grown steadily to 56 institutions in 12 countries exhibiting over 200,000 photos. And over the last four years we have seen a ton of engagement from the Flickr community with over 130,000 comments and nearly 7 million favorites on these images”*.

However our literature research shows that empirical research on The Commons is scarce. Most of the studies are limited to surveys, best practices and a few insightful blog posts. Some institutions try to analyze and comprehend the user-generated content in a quantitative or qualitative manner but this is quite limited. One of the main reasons for this shortcoming is that gathering and analyzing the data is labor-intensive when it is not automated (Bray et al., 2011). While Flickr offers some basic statistics and overviews of recent community activities for analysis, it does not get much more granular than that (Vershbow, 2008). Institutions are able to track activities surrounding their photos but would like more functionality in regards to statistics, such as the ability to export tabular data and more flexibility with sorting, filtering and manipulating the information (Graf, 2009).

The focal point of this study is to construct a better understanding of how the Flickr community interacts with the photo collections of the institutions and how the institutions react to these activities of the community. We achieved this by extracting a rich dataset from The Commons. We believe that by extracting a large and rich dataset, we can give a more detailed understanding about the current state of The Commons. This study is of an explorative nature, examining the various types of data we could extract, this enabled us to construct an overall picture of The Commons and the photo collections of the 56 participating institutions.

## 2. Flickr

Flickr can be considered as a poster child of the Web 2.0 era. The predominately image-hosting site was founded in 2004 and sold for a reported \$35 million to Yahoo! in 2005. Flickr stored more than five billion images in 2010 and passed the seven billion mark in the first quarter of 2012. This number continues to grow steadily. Thousands of photos are uploaded every minute bringing it to a total of 4.5 million daily (Sheppard, 2010; Leung, 2012; Yahoo!, 2012). Flickr’s total traffic consist out of 20 million unique visitors from the United States and in total nearly 80 million worldwide. These visitors spend an average of 2.7 minutes per visit (Kopytoff, 2011; Yahoo!, 2012).

Yahoo! (2012) positions Flickr as being more than ‘just’ a repository of photos. The company describes Flickr as a far-reaching community consisting of 51 million registered members that are eager to share and interact with each other. It is a community ranging from ultra-active members running the show, to occasional participants adding comments and tags, and lurkers watching behind the scenes. While Flickr is well known and to date widely

used, traffic to the site seems to be declining. According to comScore, unique visitors to Flickr from the United States dropped 16 percent in 2011. One of the reasons could be the rise of Facebook serving as a popular alternative for image-sharing.

Flickr combines the possibility of sharing photos with social media functionalities. Registered members can create and maintain a profile, upload and organize photos, explore other photostreams (i.e. a way of displaying images chronologically), befriend each other and enrich photos by assigning tags or notes to self-assigned areas. They have the possibility to add comments, to share an experience, ask a question or simply give a compliment. Photos can be added to a member's favorite list or to other self-organized groups with a common interest or topic.

A basic free Flickr account allows members to upload 2 videos and 300MB worth of photos each calendar month. On top of that, photostream views are limited to the 200 most recent images. Photos are allowed to be shared to a maximum of 10 groups and only smaller (resized) images are accessible. A pro account (1 year subscription for \$24.95) offers unlimited photo uploads (max 50MB per photo), unlimited HD video uploads (90 seconds max, max 500MB per video), unlimited storage and unlimited bandwidth. High-resolution versions of original images are archived, photos can be shared in up to 60 group pools and view count and referrer statistics are shown. Signing up for a pro account also ensures ad-free browsing and sharing of photos (Flickr, 2012a).

### **3. Flickr The Commons**

Recognizing the possibilities of Flickr, the Library of Congress contacted the image-hosting site mid-2007 to discuss the possibility to publish 14 million photographs from their collection. Though it was the Library of Congress who initiated this contact, it was Flickr designer George Oates who came up with the idea to create a new community space within Flickr entitled *The Commons* (Springer et al., 2008; Fysh, 2009). Flickr was, before The Commons, already a place for cultural heritage institutions to share their photographic collections and/or digitized physical object collections and "behind the scenes" photos. However these initiatives remained largely hidden and operate in isolation (Oates, 2008). The Commons could improve this by giving the world access to publicly-held photographic collections. On top of this it would give institutions the opportunity to gather context and information about the collections in the form of user-generated content added by the community members. In a pilot, the Library of Congress published on the 16<sup>th</sup> of January 2008, 1,500 photos from two collections of historical photographs entitled "News from the 1910s" and "1930s-40s in Color".

This pilot received immense exposure and praise on the Web. That same year 16 other institutions joined The Commons. In a period of four years, a total of 56 institutions joined Flickr The Commons ranging from large institutions such as the Smithsonian Institution to relatively smaller institutions such as the Museum of Reykjavík. The majority of the institutions

originate from countries with an English speaking population such as the United States, Australia and the United Kingdom. Noticeable is the lack of participation from institutions established in the continents South-America, Asia and Africa.

Appendix A shows the complete (alphabetically ordered) list of the 56 participating institutions, their location and the date they joined The Commons.

### 3.1 How to join The Commons

In order to join The Commons, institutions must meet the following three requirements (Oates, 2008):

1. *No known copyright restrictions*; All content must be marked "no known copyright restrictions". This means that there should be no third party restrictions on the use of the images, no royalties to be paid, no permissions for this use needed from photographers, 'models' or subjects in the images (Kurin, 2008)
2. *Type of photos*; The Commons will only be accepting content from photographic collections, and not photographs of physical objects. Images need to show people, not just objects, buildings or landscapes. Images should be chosen for historical and cultural significance, composition and aesthetic quality, and clarity. Images should not be overly repetitive i.e. not numerous shots of the "same" activity (Ibid, 2008)
3. *Terms of Service*; The institution must indemnify Yahoo! Inc. against potential copyright disputes, signing additional Terms of Service that also establish its presence on Flickr as an organization, not an individual.

### 3.2 Reasons to join The Commons

For the participating institutions the most important and primary reason to join The Commons is to share the hidden treasures momentarily stored away in their archives or on their own website. In short "be where the audience is" is a main reason to increase awareness and exposure by sharing photographs from collections with people who would enjoy these images but might not visit the institutions own website (Springer et al., 2008; Dalton, 2010). This sentiment is shared among the various institutions and communicated strongly in various press releases and reports (e.g. Kurin, 2008; Neighbors, 2008; State Library of New South Wales, 2008; Springer et al., 2008; Kalfatovic et al., 2009; Moortgat, 2009).

A second reason for participation in The Commons is the opportunity to gain additional knowledge about the shared collection. For example, the Australian War Memorial hopes that community members identify people on the photos from their Korean War series. Community members might notice a friend or member of their family, or have a story to share about these photos (Smedley, 2008). The Library of Virginia shares this aim by posting 200 photographs from the Rice Photograph Collection, which in total consists of nearly 16,000 original photographic negatives that document Richmond during the 1950s. With questions such as "Is there an interior view of your father's downtown office?" or "Does the collection contain an

image of a high school dance you attended more than 50 years ago?”. The Library of Virginia hopes to collect valuable information to enhance and enrich the quality of their bibliographic records (Neighbors, 2008).

A third reason for joining The Commons is to gain experience in interacting with such a large Web community. Now that the collections are “out there” with “no known copyright restrictions”, community members have the option to interact with the photos by adding tags, notes and comments. There is no telling what the community will do with these collections. Formerly in complete control by the archivist, the power automatically shifts the instant the photo is placed online (Edmunson-Morton, 2009). This shift is for some institutions a scary thought. However, the general consensus is that there is a need to experiment in order to gain experience in participating in emergent Web communities (Springer et al., 2008). Besides, the response of professionals is not necessarily negative; there is general recognition of the validity of both, users and professionals, in the enrichment of collections (Van Vliet & Hekman, 2012).

And while alternatives exist such as Picasa and Wikimedia Commons, Flickr The Commons is regarded as a safe and regulated space (Kalfatovic et al., 2009). Flickr has an online community already in place and its well-developed user interface encourages certain social norms and acceptable behaviours. Members are unable to delete each other’s tags, only the owner of the photograph can remove irrelevant, redundant, or inaccurate tags. This “less democratic” or “dysfunctionally democratic” nature gives the institutions a sense of control in comparison to e.g. Wikimedia Commons (Chan, 2010; Stvilia & Jørgensen, 2010).

#### **4. Related work**

Surprisingly enough, little empirical research has been done focusing on the actual usage of The Commons by the institutions and by the Flickr community. Vaughan (2010) conducted an online survey in which he questioned all of the participating institutions (27 at the time) over a one-month period in 2009. The survey was structured into five sections: background, institutional staff involvement, social interactions, statistics, and assessment. Question types were a mix of open-ended response questions and multiple-choice questions. The results showed that the participating institutions had positive experiences with The Commons emphasizing that exposure of the collections to a broader audience was a very important reason in joining The Commons. Using Flickr in order to “utilize Web 2.0 features” to engage user involvement and discussion was the second most important reason for joining.

Lemelin (2010) conducted a similar study to evaluate the experiences, in terms of usage statistics, goal achievement, and increased user awareness. The used methodology for this study was also an online survey based on Vaughan earlier work. In contrast to Vaughan, Lemelin focused purely on institutions that used Flickr but were not a part of The Commons. In total she obtained data from 89 institutions and compared that to the data Vaughan collected in 2009. The data showed that institutions affiliated with The Commons received

much more engagement and visits compared to institutions that were not part of The Commons. While many of these institutions somehow feel that their goals have been met, they report low usage statistics. A more peculiar finding of this study was that a large part of these 89 institutions disabled the functionalities in which community members could interact with the photos.

A substantial body of publications, with a focal point on institutional backgrounds, staff involvement, user-generated content, statistics, and assessment, consists out of blog posts, interviews and published reports and papers from the institutions themselves. These various institutional publications describe the first experiences with The Commons as well as general statistics demonstrating the total amount of user-generated content (see Chan, 2008; LibraryTechNZ, 2008; Springer et al., 2008; Vershbow, 2008; Johnston, 2009; Kalfatovic et al., 2009; Moortgat, 2009; Theimer, 2009). Table 1 shows an overview of the quantitative statistics as reported by the institutions. The table is chronologically ordered on the period the report covers.

*Table 1.* Quantitative Statistics As Reported By The Institutions Themselves Based On Chan (2008) LibraryTechNZ (2008), Springer et al. (2008), Vershbow (2008), Johnston (2009), Kalfatovic et al. (2009), Moortgat (2009) And Theimer (2009).

<b>Institution</b>	<b>Period</b>	<b>Photos</b>	<b>Views</b>	<b>Comments</b>	<b>Tags</b>	<b>Favorites</b>	<b>Contacts</b>
The New York Public Library	After the first day	1,300	53,220	123	380	1,112	–
Oregon State University	After the first five days	116	13,000	50	–	–	200
The Powerhouse museum	After the first three months	600	103,000	–	2,433	–	–
Nationaal Archief	After the first six months	771	1,078,000	1,916	6,852	771	2,148
Smithsonian	A period of seven months	–	627,259	513	–	559	3,000+
Library of Congress	After the first ten months	4,615	1,400,000	7,166	67,176	3,645	15,000
George Eastman House	After one year	813	1,971,520	3,961	9,885	4,615	15,000
National Library of New Zealand	After one year	600	56,926	–	–	–	365

The attention did not limit itself within the social space of Flickr. Joining The Commons

also had a positive effect on visits to the institutional website. For the Library of Congress visits (to their pages) originating from Flickr, rose over 2,000% in the first month. The Flickr pilot had a positive effect on the traffic to their Prints and Photographs Online Catalog (Springer et al., 2008). Moortgat (2009) reports a doubling in visits, from 200,000 to more than 440,000, to the Dutch website of The National Archives.

#### 4.1 Types of user-generated content found on The Commons

A considerable amount of user-generated content was added to photos in the form of tags and comments. Reviewing the nature of the tags, The Library of Congress observed that the majority of these tags were of a descriptive nature purely describing items seen on the photos. Less than 1% were deemed as emotional tags (Springer et al., 2008). Leitão (2010) witnessed the same with regard to Biblioteca de Arte da Fundação Calouste Gulbenkian. Community members tend to display a strong preference to describe specific image details, to identify characteristics, to label colors and other aspects of the photo. The same nature of tags were found in studies focusing on social tagging (see Hekman & Van Vliet, 2012; Van Vliet & Hekman, 2012).

As for comments, the institutions received a whole range from basic enthusiasm ('beautiful! 'stunning!'), corrections, illuminations, geotagging, technical tips, questions about high resolution reproductions, and questions in general (LibraryTechNZ, 2008; Vershbow, 2008). A popular question was that of members wanting to add institution photos to their groups (Edmunson-Morton, 2009). Not participating in The Commons, the Royal Commission on the Ancient Historical Monuments of Scotland (RCAHMS) observed the same type of comment contributions. The majority of community members interpreted the 'add a comment' functionality as an opportunity to contribute new information to the digital archive in the form of data, references to other sources or corrections to existing records. Most notable is that entries are often written in a descriptive, impersonal style, and focus on factual rather than anecdotal aspects such as personal experiences, memories, observations or direct references to family history (Clari & Graham, 2012).

Community members have the ability to add notes to photos. Adding notes offers the option of annotating a specific part of the photo by selecting an area on the photo and add text to that designated area. These notes appear when the mouse moves over the designated part of a photo. Although useful conversations have taken place between members in these notes, community members hardly used this functionality. The nature of the notes were to highlight something in the image, to ask a question or make a joke about what it is (Springer et al., 2008).

Beside tags, comments and notes containing information, opinions or questions, community members contributed in a more creative way. In the form of 'then and now' debates, community members purposely reshot pictures of old locations in the present or adding links to Mapquest and Google Earth providing the exact address. Contributions also

consisted of information on how the location has changed or remained the same over time (Chan, 2008; Springer et al., 2008).

While one of the more apparent fears of user-generated content is its credibility, it should be noted that many of the added tags, comments and notes were based on elaborate information sources lending credibility (Leitão, 2010). The Library of Congress witnessed a group of 20 “power commenters” who regularly returned and provided corrections such as place names, more precise dates, event names, and fuller names for individuals previously identified only by surname. They supported their contribution by adding links to e.g. the New York Times archive, Wikipedia, and highly specialized Web sites devoted to these specific relevant areas of interest such as military aviation, railroads, and sports history (Springer et al., 2008).

It would seem that user-generated content has led to some increase in knowledge (Leitão, 2010). The National Archives of the Netherlands were able to use 3% of the comments to modify 23 of their records (Moortgat, 2009). However, it is largely unclear whether the user-generated content is actually used by the institutions. Institutions claim they do check out the content, review it with their curators, and fill in gaps, or updates the current information stored away in their formal catalogues (Chan, 2008; Springer et al., 2008; Graf, 2009).

## 5. Methodology

In order to extract a rich dataset for our study we used Flickr’s open Application Programming Interfaces (the Flickr API). The ideology of the Flickr API is that anyone can write their own program to present public Flickr data such as photos, video, tags, profiles or groups, in new and different ways (Flickr, 2012b). In order to access the API we needed to register our application and apply for an API key. Besides registering and obtaining a valid API key, our application needed to abide the community guidelines and API Terms of Use. The rate limit of 3600 queries per hour across the whole key was one of the crucial terms we had to abide to during the data harvesting. If we would abuse or overtax the key, Flickr could expire or turn off the key in order to preserve the Flickr API functionality for others.

PHP was chosen as a server-side scripting language to access the API methods and to communicate with the MySQL database. The phpFlickr class 3.0, written by Dan Coulter, was used to access the API methods (procedures or function associated with a class) using PHP. In total there are 193 methods that return various information on photos, users or groups. For instance, to retrieve the complete list of participating institutions we used the flickr.commons.getInstitutions method. This method returns the following XML response:

```
<institutions>
  ...
  <institution nsid="29998366@N02" date_launch="1224571148">
    <name>Nationaal Archief</name>
```

```

    <urls>
      <url type="site">http://www.nationaalarchief.nl/</url>
      <url type="license">...</url>
      <url type="flickr">...</url>
    </urls>
  </institution>
  ...
</institutions>

```

Besides XML the API is capable of outputting other formats such as JSON, JSONP and PHP Serial. In total we used nine methods to retrieve the data needed for our study (see Table 2).

*Table 2. An Overview Of Used API Methods.*

Method	Description
flickr.commons.getInstitutions	Returns a list of the current Commons institutions.
flickr.people.getPhotosOf	Returns a list of photos containing a particular Flickr member.
flickr.photos.getInfo	Returns information about a photo
flickr.photos.getFavorites	Returns the list of people who have favorited a given photo.
flickr.photos.getAllContexts	Returns all visible sets and pools the photo belongs to.
flickr.photos.comments.getList	Returns the comments for a photo
flickr.photosets.getList	Returns the photosets belonging to the specified user.
flickr.photosets.getPhotos	Returns a list of photos in a set.
flickr.people.getInfo	Returns information about a user

Results from the Flickr API were stored in a MySQL database. We did this for three reasons: (1) not to strain the rate limit of 3600 calls by accessing data on the fly; (2) easy access to the datasets for future analyses and (3) to add detailed timestamps to certain types of information (e.g. tags and notes) that contain no timestamps. In the case of tags, the API does not return the creation date of a specific tag. A creation date is relevant to know in order to evaluate and monitor the growth of tags.

### 5.1 Harvesting the data

We accessed the Flickr API for a total of 14 weeks, from week 46 of 2011 through week 7 of 2012. A total of 12 weeks was deemed necessary to witness e.g. the growth in visits to the photo and tags. Our first step in the data harvesting was to retrieve a list of institutions participating in The Commons. This was done in week 46 of 2011. The second step was to see how many photos were shared by these institutions. The third step consisted of retrieving detailed information about the photos, such as a description, photo url, date of upload and number of views and information about the user-generated content such as the user id, number of comments, tags and notes. Every five minutes, 300 photos were accessed through the API starting on Monday morning until Wednesday evening. Photos can be added to other groups and be organized in photosets. From Wednesday evening until Saturday evening this

data was harvested.

Photos can be added to a users favorite list. The Flickr API returns data about who added a certain photo as favorite photo. Unlike tags and notes, it was not necessary to continuously monitor this since the returned data contains an exact timestamp of this action. The harvesting of this was done in week 6 of 2012. The same was done for comments that were added to photos or photosets.

To get a better understanding who contributes to The Commons a list was computed containing all the accounts of users that actively participated (e.g. tagged, commented or added a photo as favorite). In week 7 these accounts were retrieved and contained information such as username, the user's real name, general description, location, number of photos shared and if the user has a pro account. Due to API restrictions or a possible flaw in our authentication process we are yet unable to retrieve the user's gender.

## 6. Results

In this section we will discuss our preliminary findings by analyzing the retrieved data. Table 3 gives an overview of harvested data in the first and last week.

*Table 3. Growth Of Data In A Three-Month Period.*

<b>Data</b>	<b>Week 46 2011</b>	<b>Week 5 2012</b>	<b>Growth (%)</b>
<b>Photos</b>			
Total number of photos	191,683	196,548	2.5
Total number of views	96,961,701	105,310,076	9
Total number of photos with more than one view	191,318	196,446	3
Total number of photos with more than 100 views	80,085	84,544	6
Total number of photos with more than 1,000 views	18,579	20,385	10
Total number of photos without a view	359	102	-70
<b>Tags</b>			
Total number of tags	1,281,910	1,348,952	5
Total number of unique tags	137,789	142,361	3
Total number of tags by the institutions	876,880	916,465	5
Total number of tags by the users	405,030	432,487	7
Total number of users that added a tag*	6,598	6,655	1
Total number of photos with one or more tags	182,588	185,804	2
<b>Comments</b>			
Total number of comments	120,620	128,191	6
Total number of comments by the institutions	4,958	5,881	19
Total number of comments by the users	115,662	122,310	6
Total number of users that added a comment *	40,600	42,145	4
Total number of photos with one or more comments	32,547	34,360	6
<b>Notes</b>			

Total number of notes	21,719	22,053	2
Total number of notes by the institutions	963	983	2
Total number of notes by the users	20,756	21,070	2
Total number of users that added a note *	6,068	6,044	-0.4
Total number of photos with one or more notes	8,402	8,553	2

#### **Favorites**

Total number of favorites	626,170	672,449	7
Total number of users that added a favorite	129,733	136,361	5
Total number of photos with one or more favorites	54,780	59,513	9

#### **Photosets \*\***

Total number of photosets	1,805	1,826	1
Total number of photoset views	10,012,583	10,505,774	5
Total number of photoset comments	1,120	1,134	1
Total number of users that added a comments	850	879	3

#### **Groups \*\*\***

Total number of shares to groups	27,660	27,766	0.3
Total number of groups	3,354	3,373	1
Total number of photos added to a group	15,269	15,485	1

\*also includes institution accounts, \*\* was harvested in week 47, \*\*\* was harvested in week 48.

### 6.1 Photo uploading and account management

In the first week of data harvesting 191,683 photos were retrieved. In week 5 we retrieved 196,548 photos indicating that institutions still upload content to The Commons. The average increase of photos was 2.5% over a twelve-week period. We found that 27 of the 56 institutions did not upload new content during our monitoring period. While looking at the institutions that upload content we found that some institutions upload small portions of content on a weekly base e.g. The Library of Congress uploads around 50 new photos every week. We also witnessed institutions that upload different sized portions of content at a time e.g. San Diego Air & Space Museum uploads between 15 and 2500 photos per week. The number of photos per institutions was unevenly distributed (see Table 4), ranging from a modest 31 photos from Museum of Reykjavík to a massive number of 109,029 photos from San Diego Air & Space Museum. The majority of the institutions, a total of 51, share less than 5,000 photos on their Flickr The Commons accounts.

*Table 4. Distribution Of Total Number Of Photos Amongst Institutions.*

Number of photos	0 - 2,500	2,500 - 5,000	5,000 - 7,500	7,500 - 10,000	> 10,000
Number of institutions	46	5	1	2	2

We also observed that only 21 new photosets were created indicating that there was hardly any activity in organizing new or existing photos. Institutions have the possibility to

share their photos to other, mostly community member created, groups. In total 216 photos were shared to existing groups.

## 6.2 Views

Looking at the number of views photos received (see Table 3), we witnessed a steady growth rate in views indicating that the photos are being looked at. In the first week the total number of views was 97 million and in the final week it had risen to 10.5 million. Table 5 illustrates that the majority of the institutions has 10,000 or less views per week. Interesting to witness is that 20 thousand of the 196 thousand photos (10%) were viewed more than a thousand times. From those 20 thousand photos, we found that three-quarters originate from ten institutions (see Table 6). Even though the photo collection from the San Diego Air & Space Museum embodies more than the half of The Commons, it ranks 36st when it comes to photos being viewed more than a thousand times. In total 65 of the 109 thousand photos were viewed more than a thousand times.

*Table 5.* Distribution Of Total Number Of Average Weekly Views Amongst Institutions.

Total number of views	0 - 10,000	10,000 - 20,000	20,000 – 30,000	30,000 - 40,000	> 40,000
Number of institutions	33	13	5	2	3

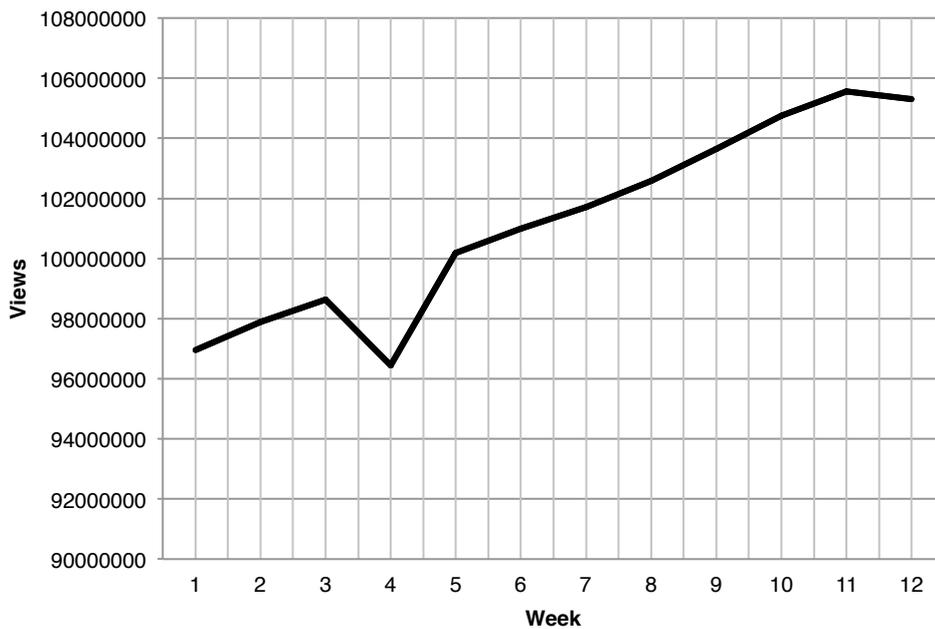
*Table 6.* Top 10 Of Institutions With Photos Viewed More Than A Thousand Times.

Institution	Number of photos
The Library of Congress	6808
Smithsonian Institution	1373
Nationaal Archief	1240
State Library of New South Wales collection	957
George Eastman House	870
The U.S. National Archives	858
Brooklyn Museum	846
New York Public Library	828
Powerhouse Museum Collection	781
State Library and Archives of Florida	389

The total number of photos from this top 10 consists out of 14,950 photos.

Figure 1 shows the growth in views over time. Two noticeable dips can be witnessed in week 49 and week 5. This is due to the removal of photos by the institutions and due to data that ‘disappeared’ while accessing the API (see discussion).

Figure 1. Growth Of Total Views In A Three-Month Period.



### 6.3 Tags and comments

In the final week, a total of 1.34 million tags were retrieved. The overall growth of tags of 5% was substantial starting from 1.28 million in the first week. We found that on average 96% of the institution's uploaded content received one or more tags. The 1.34 million tags consisted out of 142 thousand unique tags (11%).

Looking at the user-generated content we could make a clear distinction between the content added by the institutions and the content added by the community members. One of our findings was that the institutions themselves added 68% of the roughly 1.3 million tags.

The total number of tags added by community members rose from 405 thousand to 432 thousand tags in a three-month period. These 27,457 tags were added by 57 community members. In total 6,599 community members tagged 42,508 of the 196,548 photos (21%) in week 5. We discovered that 75,699 of the 432,585 community members assigned tags comprised out of unique words. Drilling deeper in the data we discovered that 32% of the community assigned tags, 65% were added by 10 "power taggers". These 10 power taggers added 280 thousand tags in total. Community member that ranked number one on the list assigned an astonishing 158 thousand tags to 14 thousand photos. Reviewing the content of these tags we saw that his top three mostly used tags were 'SDASM' (94 thousand), 'aviation' (93 thousand) and 'aeronautics' (87 thousand). By assigning a lot of the same tags the user in question was able to achieve such a high amount of contribution. The other nine members showed similar tagging behaviour. Most commonly community members do not add more than 10 tags (see Table 7).

*Table 7. Distribution Of Total Added Tags By Community Members.*

Total number of added tags	0 - 10	10 - 20	20 – 30	30 - 40	40 - 50	> 50
Number of community members	5,253	513	198	137	78	476

Total number of community members = 6,655

In week 5, 128,191 comments were added to 34,360 photos by 42,145 community members. Of the 196,548 photos, 17.5% had one or more comments. In a three-month period, 1,545 community members added 6,648 new comments. Most community members did not comment more than ten times (Table 8).

*Table 8. Distribution Of Total Added Comments By Community Members.*

Total number of added comments	0 - 10	10 - 20	20 – 30	30 - 40	40 - 50	> 50
Number of community members	40,939	653	232	91	39	191

Total number of community members = 42,145

In contrast to tagging, institutions do not comment a lot. A total of 5,881 comments were added by 53 institutions. The photosets created by the institutions received a total of 1,134 comments. Of these comments the majority (1,043 comments) came from the community members supporting the prior observation that institutions do not comment a lot. Grouping and analyzing the content, we could conclude that the majority of the comments consist out of a positive opinion towards the photo. Also, we found a lot of request from group owners asking the institutions to add the photo to their group. Table 9 shows the top 20 comments and their occurrences.

Table 9. Top 20 Of Comments Most Used By Community Members.

Comment	Occurrence
Hi, I'm an admin for a group called [...] and we'd love to have this added to the group!	473
Hi, I'm an admin for a group called [...] and we'd love to have this added to the group!	420
Congratulations! This is a wonderful shot! You are invited to post it to: [...]	390
Hi, I'm an admin for a group called [...] and we'd love to have this added to the group!	343
nice	310
wow	251
All Kenya.	239
Hi, I'm an admin for a group called [...] and we'd love to have this added to the group!	225
Great shot	218
Beautiful	211
wow!	190
beautiful!	187
Amazing!	179
Amazing	176
Hi, I'm an admin for a group called [...] and we'd love to have this added to the group!	169
Hi, I'm an admin for a group called [...] and we'd love to have this added to the group!	152
Great	135
article on the process of photochrom:[...] and more commonly from 10 to 15 tint stones.	135
cool	133
best shot	132

#### 6.4 Favorites

The possibility for community members to add a photo to their favorite list is one of the most used Flickr functionality. In week 5, 136,361 community members added a total of 59,513 photos to their favorite list. In total, this functionality was used 672,449 times in the three-month period.

### 7. Conclusion

Various institutions report that The Commons has had a positive impact in the sense of reach and user-generated content. However, in our literature study we found no empirical research to backup this statement or even research that gives a general understanding of how the Flickr community interacts with the participating institutions. By extracting a rich dataset from The Commons we created this general understanding on The Commons in regard to institutional and community participation. While our rich dataset holds many 'hidden treasures', for this explorative paper we primarily focused on looking at the first results in regard to institution activity, reach and user-generated content.

We found during our study that institutions barely add new content or organize existing content. A minimal number of new photosets were created and half of the institutions continuous to upload new content to Flickr on a regular base. The other half was, during our

study, non-active on the image-sharing site. We also noticed that institutions seldom add photos to other groups in order to further promote their photos within Flickr. A reason could be that this is a labor-intensive task.

A minority of the photos uploaded by the institutions were viewed more than a 1,000 times (10%) indicating that only a small portion of The Commons reaches the community. Uploading a large number of photos does not automatically guarantee more views as we observed for the San Diego Air & Space Museum. Though the photos shared by the institutions are being viewed, the majority of the institutions receive an average of 10,000 or less views per week. There was no drastic decline or a drastic increase in the average number of views per institution per week. A reasonable explanation for this is that Flickr is designed for showing the most recent photos. Contributions by the institutions have to 'compete with other photos in a domain where 4.5 million new photos are added every day. Seeing no drastic decline in average views per institutions means that the institutions somehow survive the competition for attention.

A total of 1.34 million tags were retrieved during our study. We could discover 142 thousand unique tags indicating that community members and institutions use the same tags for different photos. While the number on itself is quite impressive we found that the number tags contributed by community members is quite marginal. Community members added a total of 432 thousand tags. The majority of these tags (65%) were added by 10 "power taggers". We witnessed a growth of 27 thousand in three months, which is quite impressive. However, the number of community members contributing these tags is quite low, 57 in total. Overall we found that 6,655 users tagged photos added to The Commons and the majority did not add more than ten tags. It would seem that tagging itself is not a popular activity amongst community members and the diversity of used words is quite low.

In a three-month period 1,545 community members added 6,648 new comments. In total 42,145 community members added one or more tags. We witnessed that the majority of community members added less than 10 comments. Further analyzing the comments we found that the majority of comments consist out of a positive opinion towards a photo or a request for that photo to be added to a group. In future research (see discussion) we want to see if we can make a distinction between comments based on its content. Institutions do not comment a lot. In total 5,881 comments were added indicating that institutions sparsely partake in open dialogue with community members.

We found that favoring a photo is the most popular activity done by community members. In total 59,513 photos were added by 136,361 community members as a favorite. This functionality was used 672,449 times in the three-month period. A reasonable explanation would be that the action itself is easy and does not require time compared to e.g. commenting. In our introduction we quoted a blog post from Miller (2012) who gave some general statistics regarding The Commons. The numbers roughly match apart from the favorites. Miller wrote that they found nearly 7 million favorites while we found 672 thousand.

We have no reasonable explanation for such a large difference.

Looking at the total activities we found that in total 167,871 community members actively did something with the photos since The Commons started in January 2008. With a possibility of 51 million community members this number is quite small. We found that the majority of the photos hardly receive any user-generated content. In total 42,508 photos received one or more tags, 8,553 photos contain one or more notes, 59,513 photos were added as a favorite at least once, and 34,360 photos received one or more comments. This is quite minimal compared to a total of roughly 196 thousand.

Institutions report that they are quite positive in regard to The Commons based on their own findings. Institutions joined The Commons for three reasons: (1) to expose collection to a broader audience, (2) to gain additional knowledge about their collections and, (3) to gain experience in interacting with a large community. Looking at our findings we can conclude that Flickr The Commons is indeed a good alternative for institutions to share the hidden treasures momentarily stored away in their archives or on their own website.

The total amount of user-generated content was quite minimal and the body of community members was also quite small. Tags tend to be re-used and the majority of comments consist out of positive opinions or requests to use the photo. Further qualitative analysis of the content should determine if The Commons is indeed a community to gain additional knowledge about collections. Looking at the sheer numbers we predict this will be quite minimal.

Based on the number of comments and action to share a photo to another group we can conclude that institutions are not very active on The Commons and do not partake or experiment in online discussion. Flickr The Commons is a great initiative but based on our numbers will lose footing if institutions do not participate more actively. A lack of analytical tools could be the reason for minimal institution participation.

## **8. Discussion**

Working with live data has had its downsides. Sporadically photos disappeared and reappeared the next week. To illustrate this: in week 49 the New York Public Library had a decrease of 2,325 photos. This was noted and checked manually multiple times that same week to check if this was indeed the case and not a software error. In week 50 these 2,325 photos reappeared. This explains the dip in Figure 1. We noticed the same for tags. Tags tend to disappear only to reappear the next week. Our only logical conclusion could be that Flickr works with multiple instances of the data scattered over multiple databases. When a database is offline another one takes its place. It could be that these instances do not contain recent versions of the data explaining the disappearance and reappearance.

One of the items we did not monitor in the three-month period was the increase or

decrease in institutions contacts. Community members and institutions have the option to befriend each other. Through these friendships more tacit knowledge could be exchanged, something that is invisible for us. Another thing we did not extract was the gender of the community members. Somehow we were unable to request this information through the API.

During our study we purely focused on the quantitative statistics. This gave us great insight into the actual usage patterns of the institutions and community members. In an upcoming paper we will focus on the geographical patterns in online user behaviour. One of the questions we want to answer in the future is if The Commons actually contributes to the creation of new knowledge. Do community members from South America actually contribute to institutions in Europe? We also want to try to get in contact with the 56 institutions. Doing so we want to gain a better understanding how these institutions use The Commons and what their strategies were in making large collections accessible to a broad audience.

### **Acknowledgements**

Special thanks to Rogier Brussee for helpful and insightful comments on earlier drafts. We also want to thank our colleagues at the Crossmedialab for being a sounding board during our data analysis phase.

### **References**

- Bray, P., Chan, S., Dalton, J., Dietrich, D., Kapsalis, E., Springer, M., & Zinkham, H. (2011). *Rethinking Evaluation Metrics in Light of Flickr Commons*. Paper presented at the Museums and the Web 2011, Toronto, Canada.  
[http://conference.archimuse.com/mw2011/papers/rethinking\\_evaluation\\_metrics](http://conference.archimuse.com/mw2011/papers/rethinking_evaluation_metrics)
- Chan, S. (2010). Why Flickr Commons? (and why Wikimedia Commons is very different). *Fresh & New(er)* Retrieved 4-18-2012, from  
<http://www.freshandnew.org/2010/01/why-flickr-commons-and-why-wikimedia-commons-is-very-different/>
- Clari, M., & Graham, P. (2012). *Learning to Let Go: Changing Patterns of Participation and Learning through the Digital Collections of the Royal Commission on the Ancient and Historical Monuments of Scotland (RCAHMS)*. Paper presented at the Museums and the Web 2012, San Diego, United States.
- Dalton, J. B. (2010). *Can Structured Metadata Play Nice with Tagging Systems? Parsing New Meanings from Classification-Based Descriptions on Flickr Commons*. Paper presented at the Museums and the Web 2010, Denver, United States.  
<http://www.archimuse.com/mw2010/papers/dalton/dalton.html>
- Edmunson-Morton, T. (2009). Talking and Tagging: Using CONTENTdm and Flickr in the Oregon State University Archives Retrieved 4-18-2012, from  
<http://interactivearchivist.archivists.org/case-studies/flickr-at-osu/>

- Flickr. (2012a). Free Accounts, Upgrading and Gifts. *Help* Retrieved 3-22-2012, from <http://www.flickr.com/help/limits/>
- Flickr. (2012b). What is the App Garden? *Flickr* Retrieved 4-15-2012, from <http://www.flickr.com/services/apps/about/>
- Fysh, S. (2009). Interview: Helena Zinkham, Prints & Photographs Division, Library of Congress. *indicommons*, 4-19-2012, from <http://www.indicommons.org/2009/01/15/interview-helena-zinkham-prints-photographs-division-library-of-congress/>
- Graf, A. (2009). Interview: Meet The Digital Media Team at the National Maritime Museum. *indicommons* Retrieved 19-4-2012, from <http://www.indicommons.org/2009/02/24/interview-meet-the-digital-media-team-at-the-national-maritime-museum/>
- Hekman, E., & Van Vliet, Harry. (2012). *Bringing the Past to the Present - The use of tagging and storytelling for the enrichment of digital cultural heritage*. Paper presented at the CARPE Network conference, Utrecht, The Netherlands.
- Johnston, J. (2009). One year on the Flickr Commons! *George Eastman House Blog* Retrieved 4-17-2012, from <http://blog.eastmanhouse.org/2009/07/17/one-year-on-the-flickr-commons/>
- Kalfatovic, M. R., Kapsalis, E., Spiess, K. P., Camp, A. V., & Edson, M. (2009). Smithsonian Team Flickr: a library, archives, and museums collaboration in web 2.0 space. *Archival Science*, 8(4), 267-277.
- Kopytoff, V. G. (2011). At Flickr, Fending Off Rumors and Facebook, from [http://www.nytimes.com/2011/01/31/technology/31flickr.html?\\_r=2](http://www.nytimes.com/2011/01/31/technology/31flickr.html?_r=2)
- Kurin, R. (2008). Re: Call for Digital Images for Flickr Project Retrieved 3-22-2012, from <https://public.resource.org/scribd/2717297.pdf>
- Leitão, P. J. O. (2010). *Uma Biblioteca nas Redes Sociais: o caso da Biblioteca de Arte da Fundação Calouste Gulbenkian no FLICKR*. Paper presented at the Actas do 10º, Guimarães, Portugal.
- Lemelin, E. (2010). Digital Archival Collections Hosted on Flickr: Institutional Experience Outside of the Commons. *Faculty Publications and Presentations*, from [http://digitalcommons.liberty.edu/lib\\_fac\\_pubs/72](http://digitalcommons.liberty.edu/lib_fac_pubs/72)
- Leung, D. (2012). Say hello to the new Flickr Uploadr. *Flickr Blog* Retrieved 3-22-2012, from <http://blog.flickr.net/en/2012/04/25/say-hello-to-the-new-flickr-uploadr/>
- LibraryTechNZ. (2008). Awesome photo – thanks!! Or, what I've learnt from our Flickr pilot Retrieved 4-18-2012, from <http://librarytechnz.natlib.govt.nz/2008/07/awesome-photo-thanks-or-what-ive-learnt.html>
- Miller, F. (2012). Time Flies! Celebrating 4 Years of The Commons on Flickr. *Flickr Blog* Retrieved 3-23-2012, from <http://blog.flickr.net/en/2012/02/09/time-flies-celebrating-4-years-of-the-commons-on-flickr/>

- Moortgat, J. (2009). Taking Picture to the Pubic - Evaluatieverslag Nationaal Archief & Spaarnestad Photo op Flickr The Commons: Nationaal Archief.
- Neighbors, D. (2008). NAME GAME Help identify the people, events, and locations in the Adolph B. Rice Photograph Collection. *Broadside*, 10-11.
- Oates, G. (2008). *The Commons on Flickr: A Primer*. Paper presented at the Museums and the Web 2008, Montréal, Canada.  
<http://www.archimuse.com/mw2008/papers/oates/oates.html>
- Sheppard, Z. (2010). 5,000,000. *Flickr Blog* Retrieved 3-22-2012, from  
<http://blog.flickr.net/en/2010/09/19/5000000000/>
- Smedley, J. (2008). AWM on Flickr Commons. *Australian War Memorial* Retrieved 4-18-2012, from <http://www.awm.gov.au/blog/2008/11/11/awm-on-flickr-commons/>
- Springer, M., Dulabahn, B., Michel, P., Natanson, B., Reser, D., Woodward, D., & Zinkham, H. (2008). For the Common Good: The Library of Congress Flickr Pilot Project: Library of Congress.
- Stvilia, B., & Jørgensen, C. (2010). Member Activities and Quality of Tags in a Collection of Historical Photographs in Flickr. *Journal of the American Society for Information Science and Technology*, 61(12), 2477-2489.
- The State Library of New South Wales. (2008). NSW State Library launches rare photos of historical 'firsts' on Flickr, from  
[http://www.sl.nsw.gov.au/about/media\\_archive/2008/flickr\\_media\\_release\\_1Oct08.pdf](http://www.sl.nsw.gov.au/about/media_archive/2008/flickr_media_release_1Oct08.pdf)
- Theimer, K. (2009). What does it take to join the Flickr Commons? Find out in an interview with newest member. *ArchivesNext* Retrieved 4-17-2012, from  
<http://www.archivesnext.com/?p=247>
- Van Vliet, Harry, & Hekman, E. (2012). Enhancing user involvement with digital cultural heritage: The usage of social tagging and storytelling. *First Monday*, 17(5).
- Vaughan, J. (2010). *Insights into the Commons on Flickr*. Paper presented at the Faculty Publications (Libraries). [http://digitalcommons.library.unlv.edu/lib\\_articles/123](http://digitalcommons.library.unlv.edu/lib_articles/123)
- Vershbow, B. (2008). NYPL joins Flickr Commons. *NYPL Labs* Retrieved 4-18-2012, from  
<http://labs.nypl.org/2008/12/19/nypl-joins-flickr-commons/>
- Yahoo! (2012). Flickr. *Yahoo! Advertising Solutions* Retrieved 3-23-2012, from  
<http://advertising.yahoo.com/article/flickr.html>

## Appendix A

Table 10. *A Complete Overview Of The Participating Institutions (List Retrieved On 12 October 2011)*

Account name	Country	Joined
Australian National Maritime Museum on The Commons	Australia	May 11, 2010
Australian War Memorial collection	Australia	November 10, 2008
Bergen Public Library	Norway	December 11, 2009
Biblioteca de Arte-Fundação Calouste Gulbenkian	Portugal	August 15, 2008
Bibliothèque de Toulouse	France	June 26, 2008
Brooklyn Museum	United States	May 28, 2008
Center for Jewish History, NYC	United States	April 12, 2010
Cornell University Library	United States	May 19, 2010
DC Public Library Commons	United States	April 14, 2009
Fylkesarkivet i Sogn og Fjordane	Norway	May 25, 2010
Galt Museum & Archives on The Commons	Canada	October 2, 2009
George Eastman House	United States	July 17, 2008
Getty Research Institute	United States	May 15, 2009
Imperial War Museum Collections	United Kingdom	November 11, 2008
Jewish Historical Society of the Upper Midwest	United States	April 7, 2010
JWA Commons	United States	October 15, 2009
Keene and Cheshire County (NH) Historical Photos	United States	June 14, 2010
Ljósmyndasafn Reykjavíkur / Reykjavík Museum of	Iceland	June 14, 2010
LIGC ~ NLW	United Kingdom	May 12, 2009
LSE Library	United Kingdom	October 27, 2009
Musée McCord Museum	Canada	October 14, 2008
Museum of Hartlepool	United Kingdom	April 20, 2011
Museum of Photographic Arts Collections	United States	June 18, 2011
NASA on The Commons	United States	August 30, 2010
Nationaal Archief	The Netherlands	October 21, 2008
National Galleries of Scotland Commons	Scotland	January 14, 2009
National Library NZ on The Commons	New Zealand	November 27, 2008
National Library of Australia Commons	Australia	September 26, 2011
National Library of Ireland on The Commons	Ireland	June 1, 2011
National Library of Scotland	Scotland	June 22, 2010
National Maritime Museum	Australia	September 17, 2008
National Media Museum	United Kingdom	August 27, 2008
New York Public Library	United States	December 16, 2008
nha.library	United States	March 12, 2009
OSU Special Collections & Archives	United States	February 14, 2009
Powerhouse Museum Collection	Australia	April 7, 2008
Riksarkivet (National Archives of Norway)	Norway	September 19, 2011
San Diego Air & Space Museum Archives	United States	May 4, 2011
Smithsonian Institution	United States	June 16, 2008
SMU Central University Libraries	United States	April 15, 2010
State Library and Archives of Florida	United States	February 12, 2009

State Library of New South Wales collection	Wales	September 30, 2008
State Library of Queensland, Australia	Australia	January 26, 2009
Stockholm Transport Museum Commons	Sweden	October 11, 2011
Swedish National Heritage Board	Sweden	March 17, 2009
Texas State Archives	United States	April 15, 2010
The Field Museum Library	United States	April 30, 2009
The Library of Congress	United States	January 16, 2008
The Library of Virginia	United States	October 6, 2008
The National Archives UK	United Kingdom	March 24, 2010
The Royal Library, Denmark	Denmark	July 6, 2011
The U.S. National Archives	United States	February 1, 2010
Tyne & Wear Archives & Museums	United Kingdom	May 18, 2011
UA Archives   Upper Arlington History	United States	April 7, 2010
UW Digital Collections	United States	February 22, 2010
Woodrow Wilson Presidential Library Archives	United States	April 6, 2011

---