



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Digital curiosities

Citation for published version:

Terras, M 2010, 'Digital curiosities: Resource creation via amateur digitization', *Literary and Linguistic Computing*, vol. 25, no. 4, pp. 425-438. <https://doi.org/10.1093/lc/fqq019>

Digital Object Identifier (DOI):

[10.1093/lc/fqq019](https://doi.org/10.1093/lc/fqq019)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Literary and Linguistic Computing

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Digital Curiosities: Resource Creation Via Amateur Digitisation

Melissa Terras

Senior Lecturer in Electronic Communication, Department of Information Studies,
University College London.

m.terras@ucl.ac.uk

Abstract

Most memory institutions are now engaging with digitising holdings to provide online access. Although recent developments in technology have allowed users to create high quality digital resources out with institutional boundaries, little consideration has been given to the potential contribution that the general public can make to digitising our cultural heritage. This paper seeks to scope the growing trend of the creation of amateur online museums, archives, and collections, and demonstrates that the best examples of this endeavour can teach best practice to traditional memory institutions in how to make their collections useful, interesting, and used by online communities.

Introduction

Digitisation, "the conversion of an analog signal or code into a digital signal or code" (Lee 2002, 3) is now commonplace in most memory institutions, as digital representations of cultural and historical documents, artefacts, and images are created and delivered to users, generally online. The exponential growth in digitisation projects towards the close of the 20th Century, along with the establishment of guides to good practice and technical guidelines, has meant that "Countless millions of pounds, dollars, francs and marks [have been] ploughed into digital projects that have involved the conversion of library, museum and archive collections" (Lee 2002, 160). Much of the early academic debates regarding the purpose, merit, and scope of digitisation are now resolved as institutions create high quality resources for the general user and academic researcher alike (Hughes 2002, Deegan and Tanner 2002). As a result

Digitisation is not a per-se research issue but is part of a wider context related to the information society and the effective use of the digital content by cultural institutions (Minerva 2003, xxiii).

However, an area seldom considered in academic literature is the creation of digital resources by amateurs. Recent developments in Web 2.0 technologies (those that facilitate and encourage creativity, information sharing, and collaboration, see O'Reilly 2005) means that museums, libraries and archives are now re-considering their relationship with users and the general public, both in the use of digital collections and how users can contribute to an increasingly rich digital resource environment. This paper assesses the scope of online resources created outside institutional boundaries by keen individuals who wish to participate in digitising

Terras, M. "Digital curiosities: resource creation via amateur digitization". Literary and Linguistic Computing (2010) 25(4): 425-438.

cultural heritage, providing an overview and conceptualising the potential contribution that can be made by amateur digitisation.

Through integration with communities and individuals creating their own cultural and heritage content, this paper demonstrates that the uncharted territory of digital resources created outside traditional memory institution boundaries can provide a rich source of materials for both the general public and academic researchers. Additionally, those creating such online materials are generally more successful in interacting with their relevant online communities than memory institutions are. As a result, instead of being viewed as mere digital "cabinets of curiosities", the best digital resources created by enthusiasts, in their own time and at their own expense, can inform the library, archive, and cultural heritage community about best practice in constructing online resources, and reaching relevant audiences in the process.

Context

The rise of online "museums" created by amateur enthusiasts, generally containing digital images of holdings and artefacts, is a seldom considered but growing phenomenon. Amateur online collections have appropriated a variety of technologies, from static HTML, to the hosting opportunities afforded by online, new media, social networking sites such as www.flickr.com. In addition, with memory institutions appropriating Web 2.0 technologies and approaches themselves – such as tagging, and encouraging user feedback and involvement – amateur enthusiasts are now being encouraged to contribute to the online presence of established institutions. Online "museum" material resulting from amateur digitisation projects can provide a rich source of primary resources for both scholars and the general public, and although this has been all but ignored until recently by the Library, Archive, Cultural Heritage, and Arts and Humanities communities, its democratising nature is worthy of further consideration:

Libraries have always been far more selective than in generally acknowledged when it comes to their collections. I am not talking now of selection within formats (books, records, videos, etc), but of ruling out, consciously or unconsciously, vast areas of recorded information. Much of the stuff that we used to ignore now shows up on the Internet and the Web. To demonstrate this, just do a search on any subject and review the few thousand "hits" with a view to imagining their tangible analogues.... On and on it goes – acres of the cyberworld full of ephemera. What else is out there? (Gorman, 2003, p. 11).

It is acknowledged that "cyberspace is littered with the productions of ignorant, semi-literate, and/or crazed individuals", (Gorman 2003, p. 14) and in many cases, these online collections function as 21st Century cabinets of curiosities. They can be viewed as amusing, eccentric, or even worrying obsessions with a particular type of ephemera which the rest of the world has chosen to leave undocumented, providing a "an individual, a "netizen" ...[with the] means of expression for anyone with minimal technical skills but abundant passion and dedication" (Harden 1998). The Guardian newspaper described the Museum of Online Museums (<http://coudal.com/moom/>) thus: "The internet has brought advancements, but nowhere has it been more successful than in the field of meaningless rubbish. Here, vast swathes of tat are housed in one handy place for easy navigation" (2007, 31). Just because the creator

Terras, M. "Digital curiosities: resource creation via amateur digitization". Literary and Linguistic Computing (2010) 25(4): 425-438.

describes their collection as a museum does not mean to say it functions as we expect of a memory institution, whatever that may be.

However, the content of these online sources ranges from the amusing, to serious attempts at providing information resources to both scholarly and amateur researchers which are just not available anywhere else, being useful even if they lack the institutional backing and guidance of their official online counterparts. These "museums" can vary from the ramshackle and quirky to the glossy and complete documentation of ephemera which established institutions are either not interested, able, or willing to catalogue, digitise, and provide online: "one librarian's ephemeron is another's invaluable cultural resource" (Gorman 2003, p. 14). The Museum of Online Museums maintains a registry of such creations, including Devil's Rope: The Barbed Wire Museum (<http://www.barbwiremuseum.com/index.htm>), the Museum of Menstruation and Women's Health (www.mum.org), and Total Rewind, "the virtual museum of vintage vcrrs" (<http://www.totalrewind.org/>): all award winning, and featuring exhaustive documentation and digitised source material not available anywhere else. Additionally, many amateur digitisers are creating "pools" of digitised objects utilising image-hosting sites such as www.flickr.com as a platform, creating exhaustive documentation of, say, vintage dressmaking patterns (<http://www.flickr.com/groups/vintagepatterns/>), or book cover artwork of cheap paperbacks from the mid twentieth century (<http://www.flickr.com/groups/paperbacks/>).

Memory institutions themselves are beginning to experiment with Web 2.0 environments, asking the general public to interact with their digitised material through social tagging, bookmarking, and commenting (<http://www.steve.museum/>). A forward-thinking project at Oxford University, the First World War Digital Poetry Archive (<http://www.oucs.ox.ac.uk/ww1lit/>), has taken this one step further by successfully asking the general public to come forward with their ephemera to include in the archive. Harnessing the energy, passion and interest of amateur digitisation is of clear interest to the cultural and heritage sector. However, we know little about both the creators and users of amateur digitisation projects.

This paper surveys the hitherto ignored phenomenon of virtual and online museums and digitised ephemera created by amateur enthusiasts, to ascertain the motivation, scope, implementation, perception, and usefulness of such activity. Often viewed as predominantly meaningless "tat", this paper demonstrates that virtual collections created by amateurs are used, useful, and worthy of further consideration. Memory institutions can learn from the techniques employed in creating these online resources, to improve their own online presence.

Methodology

First, the literature on digitisation was reviewed to ascertain whether amateur contributions had been studied. Although digitisation is a well documented and well considered enterprise, there is a paucity of information within traditional library, archive and information resources regarding the contribution that amateurs can make to our digitised cultural heritage. Most digitisation guidelines and guides to good practice are focussed squarely on memory institutions such as libraries, archives and

museums (Lee 2001, Hughes 2004). Additionally, due to the time it takes to publish academic material, research regarding Web 2.0 technologies and how they can contribute to the dissemination and sharing of cultural heritage is only starting to become available. Raising questions about the digitisation of private collections also touches on issues of archiving, and collecting, and, taking a longer viewpoint, the contribution that amateurs can make to society. Any relevant literature discovered is referred to in the relevant discussion session.

Second, a hundred stand alone, self-confessed "virtual museums" and online collections of cultural or historical material were reviewed to indicate the coverage, scope, and purpose of their collections¹. These were chosen at random, via circular web links and pools such as "Unusual Museums of the Internet" (<http://www.ringsurf.com/ring/museum/>), portals such as the Museum of Online Museums (<http://www.coudal.com/moom/>), and the following of links presented within the online resources listed there. Likewise, groups and pools on Flickr were reviewed. The whole span of quality was represented in the sample – from ghost pages left hanging in cyberspace from 1994, to regularly updated contemporary resources, and pool of photographs added to daily, however, it became obvious that the usual rubric used to judge quality of websites applied to this selection of sites. As was to be expected, poor quality resources were in the majority, with many being abandoned online, demonstrating the passing fads of web design. The best were accurate, authoritative, objective, current, and gave coverage of a collection to provide a unique information source (Tate 2009).

Six creators of high quality websites identified through this survey were interviewed to gain their insight into the purpose, coverage and use of their material, and their view of the role of non institutional collections. (A further four who were approached did not respond). These resources were

1. Total Rewind: The Virtual Museum of Vintage VCRs (<http://www.totalrewind.org/mainhall.htm>)
2. Cover Browser: "View 450,000+ covers of comics, books, and more" (<http://www.coverbrowser.com/>)
3. Record Envelope: A little library of Factory Sleeves (<http://crossedcombs.typepad.com/recordenvelope/>)
4. Today's Inspiration: "A place for those with an interest in illustration from the 40's and 50's to share their knowledge, views and opinions." <http://todaysinspiration.blogspot.com/>
5. JonWilliamson.com: "Vintage Ephemera, Photography And More" (<http://www.jonwilliamson.com/>)
6. The Museum of Menstruation and Women's Health: "Discover the rich history of menstruation and women's health" (<http://www.mum.org/>).

Useful comments from these discussions are interspersed with the findings, below. In all cases the anonymity of the respondent is respected.

In addition, four "pools" on the photograph sharing site Flickr (www.flickr.com) were studied, to gain an insight into how Flickr is now being used as a host for

¹ Although this was a fairly large sample, one hundred museums cannot be taken to be statistically relevant for the whole Internet, and therefore the results presented here are qualitative rather than quantitative.

Terras, M. "Digital curiosities: resource creation via amateur digitization". *Literary and Linguistic Computing* (2010) 25(4): 425-438.

digitised content. Many of those posting items were contacted to ask about their motivation, and use of Flickr, and questions were posted to the groups discussions. The following groups were studied:

7. Pulp Fiction: <http://www.flickr.com/groups/pulpfiction/pool/>
8. Old-Timey Paperback Book Covers:
<http://www.flickr.com/groups/paperbacks/>
9. Vintage Patterns: <http://www.flickr.com/groups/vintagepatterns/>
10. The Great War Archive Flickr Group:
<http://www.flickr.com/groups/greatwararchive/pool/>

Finally, memory institutions currently encouraging user interaction via Web 2.0 technologies were surveyed to ascertain the extent of user involvement. These are detailed in the discussion session, below.

Findings

Coverage and Topics

As the survey of standalone virtual amateur museums progressed, it became obvious that most presented novel, detailed, and niche content with a very specific scope. Ephemera which had not been collected – or even noticed – elsewhere was documented, stored, presented, and catalogued. Various themes emerged: the history of specific, and now defunct, technologies was popular, such as VCRs, boomboxes and ghetto blasters (<http://www.pocketcalculatorshow.com/boombox/>), C90 cassette tapes (<http://c-90.nm.ru/tapes.htm>), and reel to reel recorders (<http://reel2reeltexas.com/catIndex.html>). Comic books were well featured, either collated in an overview website (<http://comics.org>, <http://www.comicbookdb.com>, <http://comiccovers.com>), or a single comic digitised in its entirety, such as the complete backrun of the 1980s UK girl's comic "Misty" (<http://mistycomic.co.uk/>). Personal and embarrassing collections which tend to be viewed as being socially taboo also featured regularly, such as 1930s condom wrappers (<http://www.ep.tc/condom-envelopes/>), and the Museum of Menstruation and Women's Health. However, by far the most popular theme was ephemera which dealt with graphic design, particularly in advertising, packaging, and nostalgia, such as The Museum of Vintage Advertising Displays (<http://www.shinegallery.com/stdipa.html>), The Vintage Christmas Catalogue Archive (<http://www.wishbookweb.com/>), The Vintage California Sardine Can Label Archive (<http://www.sardineking.com/>), and The Gallery of Vintage Poison Labels (<http://www.spookshows.com/poison/poison.htm>). Documenting the changing graphical design of book covers is also a focus of online activity.

Another popular area is the digitisation of family history and genealogical material – this is deliberately not covered here as there is already much focus on online genealogy elsewhere (see Terras 2009), and it concerns the digitisation and creation of records, rather than the recording of objects.

Terras, M. *"Digital curiosities: resource creation via amateur digitization"*. *Literary and Linguistic Computing* (2010) 25(4): 425-438.

A narrow, delineated focus, and very specific remit, seems to govern most of these online galleries, collections, and archives (the terms are used interchangeably and do not map to traditional information studies definitions). The purview, or obsession, of each individual site is always narrow, and beyond the scope of traditional memory institutions. Additionally, the collections tend to be completionist: with a passion to make sure every single variant is collected and documented (this can be opposed to collections management policies in memory institutions, where often a representative sample is viewed as being sufficient):

It's a bit of a digital collector's fever, I guess, as I really enjoy pushing up the number of images available on the site (Creator A).

As the collection and the website expanded, I started to see it as a story that could be told, rather than just an assortment of old [things] and a few interesting facts about each. I tried to get hold of everything that seemed significant in this story; the "firsts" of course, but also those with a unique technological feature, or the last in a line that became extinct. Certainly the website became an important part of the collection's growth, as the process of organising and writing the story often showed up "holes" that needed to be filled in order to properly illustrate the history (Creator F).

Interestingly, many of the sites and collections are not concerned with actually owning the items in question offline: there is often no physical manifestation of the archive to complement the online collection, as an image of the book cover, comic, or bottle top is often enough for it to be added to the collection:

The images were not all scanned by me, far from it: most images come from a variety of other digital sources, like CDs, websites, search results, eBay auctions, Amazon product listings, or big collections people have sent to me. As for the (images) I scanned, I mostly ordered books, and then had a scanning company rip apart the pages and scan the images. The main aim ... is to be a bit of a "meta" collector, collecting ... from all across the web and the real world and putting them into one single place for easy reference and for the joy of looking at them (Creator A).

Many of the collectors interviewed commented on their own reasons for focussing on their specific collections, and often the period they were interested in related very closely to when they were children:

Being born in 1966, my sense of nostalgia for this era (mid 40's to early 60's) is likely due to several factors. As a child I fondly remember many of these signs, building, ads, etc. firsthand as they were still around and in good shape when I was very young. I also remember reading the magazines at my grandparents' house as they had decades worth of old *Life* and *Saturday Evening Post*'s they saved. (Unfortunately those are long gone) (Creator B).

It would therefore be interesting, as further research, to consult literature on the psychology of collecting, as these archives differ greatly to traditional archives and collections within memory institutions.

Various themes emerge as to why creators wished to make their collections public:

I thought it would be nice to share my small collection. The graphics are amazing, showing what designers can do with limited funds, format and a lot of imagination. (Creator C).

Terras, M. "Digital curiosities: resource creation via amateur digitization". Literary and Linguistic Computing (2010) 25(4): 425-438.

It is worth remembering that these collections would otherwise not be available unless they were displayed online:

It's impossible for one person (or at least me) to run a physical museum and especially without funding; I tried it once for four years and it almost killed me! The online museum reaches around the world and updating can be done when I want to. It's open 24/7. It's cheaper online, especially if ad revenue covers costs. I express myself about the same way - I didn't hold back in the museum in my house - but once I say it online I don't have to repeat myself; it's there until I take it off... (Creator E).

Additionally, the use of online technologies affords the creators both freedom and autonomy:

While there are limitations associated with individual management of an extensive site such as this, there are countervailing advantages. Most importantly, this consists of complete freedom in terms of content. Given time, I can put online any off-the-wall idea I think up, without any corporate or bureaucratic constraint. Hopefully, this broad-ranging approach does not result in a haphazard site, but rather a variety of content which different visitors can sample according to their personal taste. This freedom of content provides, I believe, an idiosyncratic flavor to the site which makes it unique (Harden 1998).

All of the enthusiasts enjoyed finding an audience for their collections, sharing values with likeminded individuals in the online community, and gaining a sense of recognition and pride through creating their resources. The digitisation is carried out as a not-for-profit hobby: the interaction with other enthusiasts and viewers afforded by using Internet technologies gives a sense of camaraderie and often encourages rigorous debate between enthusiasts keen on properly documenting their chosen topic, and interest from individuals willing to contribute to the collection:

I'm amazed at the generosity of people. I have often been given [things] for my collection, often at some expense to the donator (I've had [items] sent from Ireland and Germany, and not been permitted to refund the postage costs) -- simply because they recognise that I'm attempting to preserve something that could otherwise be lost. Again, without the exposure of the internet, that would never have happened (Creator F).

This brings with it a sense of pride and accomplishment:

I'm proud to have connected many artists of the last 'golden age' of illustration (or in the case where the artists have passed away, their surviving families) with a new appreciative audience that wants as much as I do to keep the memory of their work alive. This is something I'm passionate about and gladly make time for every day (Creator B).

A means of Production

Various technologies were employed as a host platform by the amateur museums surveyed, all embracing the "one-to-many" Internet publishing model. Those established in the early days of the web tended to still use static HTML, such as Total Rewind (established 1996) and MUM (established online in 1998). Those with large collections, such as CoverBrowser, often utilised online databases and server side

Terras, M. "Digital curiosities: resource creation via amateur digitization". *Literary and Linguistic Computing* (2010) 25(4): 425-438.

scripting to allow users to search their vast collections. However, the creation of most of the successful amateur online resources coincides with

the large scale shift toward user-generated web content – a movement defined by the related practices of (to use the argot of the field) 'generating' and 'browsing', 'tagging' and 'feeds', 'commenting' and 'noting', 'reviewing' and 'rating', 'mashing-up' and making 'friends' (Beer and Burrows 2007).

These Web 2.0 activities are aided by the availability of easy to use platforms for hosting web content. Although, in the early days of the web, much was made of the potential democratisation of information through the creation of HTML based, hand-coded websites, viewable online by anyone, putting a website online involved significant investment in time and energy in learning rudimentary aspects of web design, programming, site hosting, etc, and also depended on having access to a personal computer, internet connection, and a scanner or other digitisation device. Web 2.0 technologies, with their easy to use platforms (such as blogs, wikis, and photo hosting and sharing) enable users to create professional looking resources without much technological or temporal investment, which has led to

changing relations between the *production and consumption* of content; the mainstreaming of *private information posted to the public domain*; and the emergence of a *new rhetoric of 'democratisation'*.... This is defined by stories and images of 'the people' reclaiming the Internet and taking control of its content; a kind of 'people's internet'... This, we are led to believe, has led to a new collaborative, participatory or open culture, where anyone can get involved, and everyone has the potential to be seen or heard (Beer and Burrows 2007).

As a result "Users with a personal interest in a technology can collectively and collaboratively, in a distributed and largely undirected environment, provide useful public resources" (Brady 2005, p. 225). Technologies such as blogs, Photo sharing websites, and wikis

enable people to accumulate knowledge as well as share and manage it... Knowledge is synthesised by communication between people sharing objectives. ... This increased ability to communicate helps to foster collaboration... as well as forming diverse types of community that can engage in activities such as research and problem solving... bringing resources together and sharing knowledge to jointly further their common understanding of a topic... suggesting that the new tools may be encouraging a cultural shift (Brady 2005, p. 226).

Many of the amateur museums sites surveyed now use blogging software as their base, such as Today's Inspiration, (established in 2005), Record Envelope (established in 2008) and Jonwilliamson.com, (established in 2008). Many host their predominantly image based content on Flickr, an online photo sharing website launched in 2005, and stream it to their blogs. Additionally, many of the collections surveyed have no other site rather than the Flickr group: it has rapidly become the dominant source for all aspects of image based culture, including pictures of specific items (for example, book cover design, or graphic design in general), and pertaining to specific topics (such as ephemera connected to the second world war). This has many advantages: first, it is a low cost way of hosting image based material. Secondly, by engaging with the already present Flickr community, the collections can be presented to an already active online audience:

I use Flickr for image storage and as part of an online community of archivists of vintage imagery and my blog as a means of documenting my ongoing research/self-education, as well as a way of connecting with kindred spirits and to "fish" for new information (Creator B).

I like looking at what other folks are posting, and the fact that the group already had a built-in audience are reasons for me joining versus creating a new group (Creator C).

However, all creators interviewed were not aware of any guidelines or procedures in creating traditional archival metadata regarding their collections, or any guidelines regarding image quality and veracity:

This is not a formal venue either. If it weren't being done as a hobby, but as a real exhibit then I would fully disclose any modifications. But Flickr is a dumping ground with few rules or guidelines and certainly no regulation or quality control for most groups (Creator D).

Instead, an "intuitive metadata" seems to have emerged with most collections: creators save all available information regarding the object, book, item, etc in a form which users can see, and usually search. Within the Flickr interface, this is done through comments, notes, groups, sets, and tags. This gives a final benefit to using Flickr: the inbuilt functions provide basic "collections management" tools, allowing related information about the resources to be stored alongside the images, and found and searched by users. (However, although this use of Flickr as a rudimentary "Image Management Database System" supplies the enthusiastic creators of non-institutional digitisation projects with a robust, connected platform, their lack of formal knowledge of technical and management standards will always preclude these resources from being eligible for any research funding for further digitisation: "digitization is happening in communities that do not explicitly adhere to principles increasingly acknowledged as central to the success of publicly funded efforts" (Heath forthcoming). There is also often a hanging question of who owns the copyright to various images, adverts, and objects digitised and hosted by these amateur sites).

The creators of the content are not shy in encouraging others to use their resources. They frequently use Twitter (www.twitter.com), the short messaging service, and Facebook (www.facebook.com), the social networking website, to post links to highlights or updates they feature on their resource. Blogs are updated regularly and routinely, and they encourage users to sign up to RSS feeds from their blogs to spot any new information posted there. They interact with their user communities through these interfaces, being interested in any approach that is made to them. A large amount of time is spent in updating and maintaining the online presence of the collection:

It's a bit like the work of a gardener, I suppose, moving plants, planting new ones, sprinkling water here and there, removing old plants, checking which plants turn out to be the most beautiful and planting more of those, and so on. As I'm a work-from-home or work from cafe with laptop freelancer, I don't keep precise track of the exact hours spent, though. Growing this is my hobby and work at the same time, a lot of fun (Creator A).

My wife calls it my "second job" so too much, I suppose. I suspect it works out to two or three hours per day (Creator B).

Terras, M. "Digital curiosities: resource creation via amateur digitization". Literary and Linguistic Computing (2010) 25(4): 425-438.

As a result of these activities, they often appear towards the top of the list on a search for their topic in a search engine, or are linked to from Wikipedia as a trusted source, thus encouraging more traffic to their site.

This can be contrasted with the static "scan and dump" digitisation approach undertaken by many projects with the cultural and heritage industries: once an institutional website is created, it is often left to its own devices, with little sustainability funding made available to allow the regular upkeep and maintenance, and lack of the type of interaction with user communities necessary to attract and keep visitors which were described by these passionate amateurs.

Usage Statistics

Cultural and heritage institutions are notoriously poor at reporting "evidence of use" of their digitised resources (Warwick et al 2008). It is suspected – and anecdotal evidence suggests - that many digital resources are seldom, if at all, used. In contrast, many of the websites visited in this survey were proud of the usage statistics, showing counters on pages which detailed thousands of hits, or hosting images on Flickr which had been viewed hundreds, if not thousands, of times. When queried, the creators interviewed were quick to produce stunning evidence of use of their digital resources which many traditional memory institutions would cut off their Left Wing to achieve:

I get a steady weekday visit average of 1,200 to 1,500 visitors per day to my blog. On weekends it drops to about 6 - 8 hundred per day. I don't really bother tracking stats on Flickr, but I have had 3 and a half million views of my "photos" on Flickr since I started by resource at the end of 2005 (Creator B).

Most resource creators can provide examples where specific, detailed queries from interested researchers have been answered through use of their unique collections. Specifically, many of those working with image based and graphic design material cited that the bulk of their audience were interested in graphic design, ephemera, and working within advertising and illustration. Academic researchers are happy to turn to these amateur websites when they are the only – or best - source of information about the given topic:

I know it has been [used] on many occasions. I have been contacted regularly by students writing research papers, for instance (Creator C).

Absolutely; in fact yours is the second academic project into the use of media that I've helped with this year. I have also been quoted in a government white paper, and contributed to quite a few academic papers and books. I've also noticed text from my site appearing in the "official history" pages of at least one technology company... (Creator F).

As far as I know most of the material is found nowhere else... I try to provide all the information I have about the objects... Many people have asked permission to use information and pictures from the site for projects and I always say yes. I suspect many people use the pictures and information without asking. Teachers from high school and college have told me they use information for classes; someone even said it was us that started the academic

Terras, M. "Digital curiosities: resource creation via amateur digitization". Literary and Linguistic Computing (2010) 25(4): 425-438.

interest in [this], nice to hear but I kinda doubt that. Researchers have asked to use pictures for articles, a Professor at Duke University asked for a copy of one, etc., etc. and many publications also have used material from the archives... I don't check the number of users but am aware of page views, which have increased in the past few years. I can tell when something like a magazine article mentions us because the numbers temporarily increase (Creator E).

It has been noted that

"one consequence of the choices made by academic sources is the ceding of important territory in some of the most dynamic and visible parts of the internet... private individuals are engaging in practices that are open and that do promote access, and they are reaping the benefit of their decisions" (Heath, Forthcoming).

Many online resources created by cultural and heritage institutions do not foster the relationship between users and resource in the same way that these successful amateur websites do. Resource creators are generally aware of usage statistics, and are aware of their specific audiences, interacting with them efficiently via social media, including twitter, blog comments, and email. By tapping into the ready made audiences provided by a platform such as Flickr, many of these resources are much more visible and much more used than stand alone digital collections hosted in isolation on institutional servers:

"While we may not have our collections displayed in the virtual equivalent of the Smithsonian, we do have them in the virtual equivalent of Grand Central Station" (Creator B).

Pro-Amateur Creators

The contribution that amateurs can make to established culture is often derided (Keen 2007). Indeed, the information amateurs can contribute to the Internet is often perceived as containing "all the evils of the cultural world – plagiarism, lack of transparency, misleading or inaccurate information, even outright fraudulence" (Isaksen, 2009). When deciding on the title of this paper – before carrying out the in depth interviews with creators of resources - it seemed obvious to use the term "amateur" to describe those outwith institutional boundaries who were creating their own digitised cultural and heritage content. However the term amateur holds negative connotations, and does not fully recognise the range of expertise and the amount of knowledge that the creators of these digital resources often had, or respect the community aspect of their enterprises and interests:

subaltern groups of amateur media producers will continue to thrive in the shadows of mainstream media production, pushing back at their dominant logics and narratives. For those shut out of dominant and professionalized ... subjectivities by factors such as age or class or identity, these shadow exchanges and amateur reputation systems offer an alternative economy of value and productive participation... Cultural content is being exchanged and engaged with at a scale that is larger than intimate and personal communication, but not at the scale of mass media transmission... Although lacking access to professional media networks, these amateur networks are

Terras, M. "Digital curiosities: resource creation via amateur digitization". Literary and Linguistic Computing (2010) 25(4): 425-438.

viable subeconomies where... people gain a sense of expertise, deep knowledge and validation from knowledgeable peers. In other words, these are expert communities, although not professionalized ones (Ito 2006, p.64).

Many of the creators producing these diverse amateur resources were extremely self motivated, enthusiastic, and dedicated, testing boundaries between definitions of amateur and professional, work and hobby, independent and institutional, and production and consumption. The contribution, and dedication, of those shrugged off as "amateurs" has been previously discussed: Leadbetter and Miller (2004) recommended the term "Pro-am", meaning someone who "pursues an activity as an amateur, mainly for the love of it, but sets a professional standard". Robert Stebbins uses "serious leisure" (1992, 2001) in his seminal studies to describe committed amateurs working across various fields, listing their perseverance, endurance over time, personal effort based on specially acquired knowledge, training or skills, personal experience, unique ethos, and strong identification with their chosen pursuit (1992, p 6ff).

It is also worth considering the role that amateurs have played, and still play, in research. Many academic disciplines started outside established curricula, including classics, maths, black studies, astrophysics, oral history, women's studies, and contemporary history. Many scientific developments that occurred in the 19th Century, in areas such as meteorology, biology, and astronomy, were dependent on ranks of highly skilled amateurs prepared to carry out detailed observations and experiments (Finnegan 2005, p. 10-11). Museums and memory institutions themselves were founded by keen amateurs and collectors, morphing into establishments over time. The social history of knowledge

is a history of the interaction between outsiders and establishments, between amateurs and professionals, intellectual entrepreneurs and intellectual rentiers. There is also interplay between innovation and routine, fluidity, and fixity, 'thawing and freezing trends', official and unofficial knowledge. On the one side we see open circles or networks, on the other institutions with fixed membership and officially defined spheres of competence, constructing and maintaining barriers which separate them from their rivals and also from laymen and laywomen (Burke 2000, 51-2).

Often, it is those outside established institutions that have taken the lead in exploiting new technologies: metal detectors reinvigorated amateur archaeology, telescopes enabled astronomical research, binoculars transformed ornithological fieldwork, cassette recorders encouraged language and dialect studies and the creation of oral and life histories (Finnegan 2005). Is it really beyond the reach of amateur enthusiasts who create online resources to inform institutions of best practice in doing so?:

Might the definitions and practices of knowledge be in any way reshaped in the technologies now deployed by many independent researchers – multi-modal as well as verbal, electronic, broadcast, print? (Finnegan 2005, p. 9).

Might the democratising nature of web 2.0 technologies allow a platform for us to reconsider collective and personal cultural histories? Given the groundswell of interest in creating online content, might we reconsider the potential quality and potential coverage of digital resources created by pro-amateurs (Shirky 2008)? Might we reframe how we make digitised cultural heritage accessible, in the light of the

Terras, M. "Digital curiosities: resource creation via amateur digitization". Literary and Linguistic Computing (2010) 25(4): 425-438.

success of the pro-amateur digitisation movement? Can institutions adopt their successful tactics, to increase the use and usefulness of their collections?

Memory Institutions and Web 2.0 Technologies

It would be wrong to suggest that memory institutions were not aware of web 2.0 technologies, and the fact that they can reach and interact with audiences in this manner. Indeed, many institutions are looking to websites such as Flickr and Twitter, and setting up interactions which cross the institutional boundaries. For example Smithsonian has joined the Library of Congress, the Powerhouse Museum, and the Brooklyn Museum in releasing hundreds of photos from their archives online – free of copyright restrictions – to The Commons on Flickr

The key goals of The Commons are to firstly give you a taste of the hidden treasures in the world's public photography archives, and secondly to show how your input and knowledge can help make these collections even richer. You're invited to help describe the photographs you discover in The Commons on Flickr, either by adding tags or leaving comments. (<http://www.flickr.com/commons>).

Other institutions using Flickr include the Victoria and Albert Museum (http://www.vam.ac.uk/activ_events/do_online/flickr_group/index.html), The Tate (<http://www.tate.org.uk/britain/exhibitions/howweare/>) and Culture 24's Museums at Night programme (<http://www.flickr.com/groups/museumsatnight09/>). Many of these encourage users to upload their own content, such as pictures of the existing collections, or photography competitions around a given theme, and to interact with the institution's material online. However, it is very seldom considered that the objects and collections that people have themselves may be treated as an extension of institutional collections, and drawn together via web 2.0 technologies.

There are some projects that have encouraged outreach to interact with the general public. For example. Oxford University's Great War Archive (<http://www.oucs.ox.ac.uk/ww1lit/gwa/>) contains over 6,500 items contributed by the general public between March and June 2008. Contributions were received via a special website and also through a series of open days at libraries and museums throughout the country, helping and aiding the general public in scanning and submitting their personal artefacts and recollections. A Flickr group (<http://www.flickr.com/groups/greatwararchive/>) has extended the collection beyond that of the funded phase.

Likewise, the National Library of Wales (<http://www.ourwales.org.uk/>) worked with groups to create digital community archives, to build up accessible collections of material collated and interpreted by community groups, to allow them to present their own history and give the communities their own voice online. The BBC's WW2 People's War project (<http://www.bbc.co.uk/ww2peopleswar/>) invited members of the public to submit their recollections and images of World War Two to their digital archive, collecting 47,000 stories and 15,000 images between June 2003 and January 2006. A national story gathering campaign was launched, depending on volunteer assistance, and depending on many libraries, museums, and other institutions to act as 'associate centres' where stories could be gathered. Both these initiatives have had

Terras, M. "Digital curiosities: resource creation via amateur digitization". Literary and Linguistic Computing (2010) 25(4): 425-438.

the opportunity to reach out to areas of society which may not have the IT literacy or expertise to create their own digital resources.

Such interaction with the general public is still rare, however. Additionally, little has been done to bridge the gap between non-institutional pro-amateurs, their private collections of ephemera, and institutional collections and their online presence. Yet, as demonstrated by the resources described here, there is vast expertise, dedication, and willingness to provide digital resources of specific topics, and the work of these pro-amateurs is creating new and novel information, and research, which others can use and refer to:

Most 'learning age' analysts talk in terms of 'learning' but a few now bring out its 'inquiry' dimensions and in doing so challenge the mystique and exclusivity associated with 'research'. This may well be reinforced by the developing technological opportunities for public and private collaboration on the web, even perhaps for new involvements ... through web interactions or interlinked computing. In short there are growing numbers of people with the opportunities, experience, and confidence to actively participate in knowledge through their own researching (Finnegan, 2005, p. 15).

Identifying how memory institutions can capitalise on the interest and dedication of online pro-amateurs, and how they can extend the reach, use and usefulness of their own collections in the light of the successful audience statistics these pro-amateur collections garner, should be a priority of research in digitisation for the near future. Additionally, it is interesting that in the successful outreach projects listed above, the institution's buildings and resources were used to encourage visitors to interact with and create their own digital resources (such as hosting scanning days, or advice drop ins), and the role that memory institutions can play in this regard needs to be further examined.

This has been reinforced by a recently commissioned report for the UK Government's Joint Information Systems Committee, regarding "Digitisation, Curation and Two-Way Engagement" (Chris Barratt Consulting 2009). The report concluded that there is "a very wide range of projects, organisations and communities involved in activity... However much of this activity is uncoordinated." (p.4). The use of sites such as Flickr "means that in some project rapid progress can be made to build and enrich collections. This can support sustainability and also innovative flows of new knowledge" (ibid). Small scale projects appear to be flourishing, although there is "no mechanism" for bringing together the "body of practice and experience that has developed within the HE/FE sector and beyond.... To share experiences and to develop common approaches" (ibid) between institutions and the general public. The report calls for co-ordinated work by JISC between its Digitisation Programme (which funds the digitisation of institutional collections) and its Business and Community Engagement Programme, and stresses the role that community-related institutions such as libraries and archives can have in acting as conduits and enablers to the general public, communities, Further, and Higher Education institutions.

Future Areas of Research

Terras, M. "Digital curiosities: resource creation via amateur digitization". Literary and Linguistic Computing (2010) 25(4): 425-438.

Given the qualitative nature of this research, and the scoping nature of the study, there are many further areas of research which would be fruitful to pursue. Firstly, further investigation regarding the psychology of collecting, as opposed to the collection policy of institutions, may illustrate some useful insights into the different approaches personal and institutional collections take. Secondly, further dialogues with both creators and users of these pro-amateur resources may ascertain why they were so popular with users, and the features and characteristics which could be adopted successfully by larger institutional websites. It would then be useful to carry out a controlled experiment with various institutions which had not yet made use of web 2.0 technologies, to ascertain how integrating their digitised collections with resources such as a blog, Flickr, Twitter, Facebook, etc affected the quantity and type of users of their digital resources. Statistics regarding traffic could be collated and analysed, demonstrating whether adopting these popular technologies could encourage use of digitised collections. It should be noted that we are currently at a transformative time, when many institutions are considering the switch from static to more dynamic modes of interacting with their potential audiences. We should also be aware that status updates, tweets, blog postings, and blog comments are digital ephemera in themselves, and unlikely to survive even in the short term: by gathering and studying these exercises in community interaction, we can study, analyse, and preserve the type of interactions institutions attempted at this turning point in web technologies.

Conclusion

Enthusiastic digitisation by amateurs, a phenomenon previously ignored by information professionals, is providing a rich source of online cultural heritage content which often documents areas not covered via traditional institutions. Indeed, ephemera and popular culture materials are often better served by the pro-amateur community than memory institutions. The energy and zeal displayed by amateur digitisers is worthy of further consideration, as amateur collections often complement existing collections, providing an alternative free discussion space for enthusiasts. The pro-amateur community is much better at interacting with online audiences than memory institutions are, and the user statistics generated by these collections indicate that frequent updates and interactions are the key to establishing regular users of online content. As a result, memory institutions should be aware that it is no longer acceptable to "scan and dump" their digitised collections into online databases and expect them to be used. Outreach efforts using Web 2.0 websites such as Flickr, Twitter, and Facebook, can encourage and build an audience online. Indeed, linking stand alone institutional websites into websites such as Flickr, which have an inbuilt audience, may provide a way to increase the use of digitised heritage content. Web 2.0 technologies present great potential in linking the amateur with the institution, extending the reach and scope of digitised cultural heritage. By acknowledging the contribution that pro-amateurs can make to online content, institutions may be able to reinvigorate their online presence.

Terras, M. "Digital curiosities: resource creation via amateur digitization". *Literary and Linguistic Computing* (2010) 25(4): 425-438.

References

Beer, D., and Burrows, R. (2007). "Sociology and, of and in Web 2.0: Some Initial Considerations". *Sociological Research Online*, Volume 12, Issue 5. <http://www.socresonline.org.uk/12/5/17.html>

Brady, M. (2005). "Blogging, Personal Participation in Public Knowledge-Building on the Web". In Finnegan, R. (ed) (2005). "Participating in the Knowledge Society: Research beyond University Walls". Palgrave Macmillan, Houndmills. p. 212-229.

Burke, P. (2000). *A Social History of Knowledge, from Gutenberg to Diderot*. Polity, Oxford.

Chris Batt Consulting (2009). Digitisation, Curation and Two-Way Engagement. Final Report to JISC. http://www.jisc.ac.uk/media/documents/programmes/digitisation/dcatwefinalreport_final.pdf

Deegan, M. and Tanner, S. (2002), *Digital Futures: Strategies for the Information Age*. Digital Futures Series (London, Library Association Publishing).

Finnegan, R. (2005). "Introduction: Looking Beyond the Walls". In Finnegan, R. (ed) (2005). "Participating in the Knowledge Society: Research beyond University Walls". Palgrave Macmillan, Houndmills. p. 1-20.

Gorman, M. (2003), "Cataloguing in an Electronic Age", in Intner, S. S., Tseng, S. C., Larsgaard, M. L. (eds.) (2003), *Electronic Cataloguing: AACR2 and Metadata for Serials and Monographs*, (Binghamton, The Haworth Press), 5- 17.

Guardian (2007), "Web Watch", *The Guardian*, Guide Magazine, October 20 2007, 31.

Heath, S. (Forthcoming). "Diversity and Reuse of Digital Resources for Ancient Mediterranean Material Culture". In Bodard, G. and Mahoney, S. (eds). "Digital Research in the Study of the Classical Antiquity". Ashgate, in preparation.

Harden, M. (1998). "Web Graphics: Art on the Net". Museums and the Web Conference 1998, Toronto, Ontario, Canada, April 22-25, 1998. http://www.archimuse.com/mw98/papers/harden/harden_paper.html

Hughes, L. (2004). *Digitizing collections: strategic issues for the information manager* (London: Facet Publishing).

Isaksen, L. (2009). "Pandora's Box: The Future of Cultural Heritage on the World Wide Web", in Tsipopoulou, M. (2009) (ed.) International Conference of the Hellenic Ministry of Culture on Digital Heritage in the New Knowledge Environment: Shared spaces & open paths to cultural content, Athens, 2008. Athens : Hellenic Ministry of Culture. <http://leifuss.files.wordpress.com/2009/04/pandorasboxrev1.pdf>

Terras, M. "Digital curiosities: resource creation via amateur digitization". *Literary and Linguistic Computing* (2010) 25(4): 425-438.

Ito, M. (2006). "Japanese Media Mixes and Amateur Cultural Exchange". In Buckingham, D and Willet, R. (eds) "Digital Generations: Children, Young People and New Media". Routledge, 49- 66.

Keen, A. (2007) *The Cult of the Amateur: How Today's Internet is Killing Our Culture*. London: Nicholas Brealey Publishing

Leadbetter, C. and Miller, P. (2004). *The Pro-Am Revolution: How Enthusiasts are Changing our Economy and Society*, London: Demos,
www.demos.co.uk/catalogue/proameconomy.

Lee, S. (2002), *Digital Imaging, a Practical Handbook* (London, Facet Publishing).

Minerva (2003), "Summary of Progress". *Coordinating digitisation in Europe*. Progress report of the National Representatives Group, European Commission, The Information Society Directorate-General,
<<http://www.minervaeurope.org/publications/globalreport/globalrep2002.htm>>.

O'Reilly, T. (2005). "What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software". O'Reilly, <http://oreilly.com/web2/archive/what-is-web-20.html>.

Shirky, C. (2008). "Here Comes Everybody: The Power of Organizing Without Organizations". Penguin, London.

Stebbins, R. (1992). *Amateurs, Professionals, and Serious Leisure*. McGill-Queen's University Press, Montreal.

Stebbins, R. (2001). *New Directions in the Theory and Research of Serious Leisure*, Edwin Mellen Press, Lewiston.

Tate, M. A. (2009). "Web Wisdom". CRC Press. (Second edition of *Web Wisdom: How to Evaluate and Create Information Quality in the Web*)

Terras, M. (Forthcoming 2009) "Potentials and Problems in Applying High Performance Computing for Research in the Arts and Humanities: Researching e-Science Analysis of Census Holdings". *Digital Humanities Quarterly*. Accepted.

Warwick, C. Terras, M., Huntington, P., and Pappa, N. (2008). "If You Build It Will They Come? The LAIRAH Study: Quantifying the Use of Online Resources in the Arts and Humanities through Statistical Analysis of User Log Data". *Literary and Linguistic Computing*. 23(1), 85-102.