NatSCA News

Title: Google me a penguin: Natural History Collections and the web

Author(s): Mark Carnall

Source: Carnall, M. (2010). Google me a penguin: Natural History Collections and the web. *NatSCA News*, *Issue 19*, 60 - 64.

URL: http://www.natsca.org/article/122

Google Me a Penguin: Natural History Collections and the Web

Mark Carnall

Grant Museum of Zoology and Comparative Anatomy, University College London
Email: mark.carnall@ucl.ac.uk

A lot of useful work has been undertaken in recent years with natural history collections and the web. Many online databases have been launched and there is a huge network of subject specialist web forums offering professional curatorial advice from topics as diverse as setting entomological specimens through to sourcing fossils for museum galleries. Nonetheless, although online databases may have been created through funding with the idea of making collections accessible to 'everyone'; online museums do not, in fact, *engage* with the 'web public' at all. Museum websites do not support the kind of curiosity-based learning and wonder that natural history galleries inspire, or provide access to infectiously enthusiastic and knowledgeable natural history curators, conservators, educators, volunteers and managers.

Introduction
It appears to take between five and ten years for technology, be it social networking sites like Facebook (2004), Twitter (2006) or web practices like metadata tagging through Delicious (2003) or news aggregating on Reddit (2008) or Digg (2004), to filter down from the early adopters and technophiles to museums. In that ten-year gap, young people will typically adopt and then drop the technology, marketers will find a way to exploit the technology and even libraries will have used the technology before museums eventually start to experiment (around the same time a number of extortionate 'How to…' training courses start to appear in the museum specialist press). Often this lag between launch and museum use is beneficial because the technology will have been critically tried and tested and is therefore more likely to persist for a number of years. Occasionally this means museums are jumping into new technologies late and often struggle to stand out. Furthermore, museum spaces and their wonderful objects don not naturally translate to the web and museums don't appear to have a comfortable niche. Compare natural history museums websites to the websites of charitable conservation organisations, natural history 'glossy mags' and natural history documentary makers. These other organisations have used the web to complement their 'meatspace' entities; as will be seen, natural history museums struggle to stand out or offer a consistent service. Of higher concern is that even a potential virtual visitor who knows that the museum exists may not even make it to the museum website. Generally, museums are very good at assessing and addressing issues with physical, intellectual or psychological access to museums and museum displays, but as will be shown there has been little thought put into the web accessibility.

Methodology
In order to demonstrate how accessible natural history museums are to a virtual visitor, a simple experiment was designed to model a web-based search undertaken by a hypothetical virtual visitor interested in finding out more about penguins. The virtual visitor has a homework assignment and would like to find out more reliable information about penguins (see **Notes** for discussion about reliable information on the web). This virtual visitor is the ideal individual that a museum may hope to inspire or at least assist in the search for information.

Penguins were chosen as the topic under research because penguins are almost universally recognisable, in contrast to examples like Northern Hairy-nosed Wombats or The Tissue. The word penguin is also less likely to return confusing results from search term matches. Penguins were also selected because they are not 'super-Hollywood' animals like lions, tigers or pandas which we might expect to be very well covered by museum websites. Lastly, penguins were desirable because almost every zoological natural history collection will have penguin specimens.

In order to assess the breadth and quality of information, our theoretical web visitor has the following six questions to answer about penguins. How many kinds of penguin are there? What do penguins eat? What kind of animal is a penguin? Where can you find penguins? Do penguins fly? What is the biggest and smallest kind of penguin?

Museums were given a slight favourable bias in the web search. For each museum the starting point for the information hunt would be to type the word penguin into the first obvious search box found on the museum homepage. In reality, it is unlikely that a visitor would run this query through a museum website; more likely a search engine would be used. In order to model typical web-browsing behaviors, a number of constraints were put in place during the hunt for the relevant information. If no results for searching the word *penguin* then the search was recorded as a fail. If there were results listed then the visitor was given a rather generous bounce time (the amount of time somebody browses for relevant information before giving up and going elsewhere), of 40 seconds and a maximum of 20 clicks were allowed subsequently to find as many of the answers as possible. Again, the odds are slightly stacked in the favour of the museum, considering that the bounce time is an order of magnitude more than it is estimated the average website visitor will tolerate (*Every Second Counts: How Website Performance Impacts Shopper Behaviour).*

The Sample Set
The sample set of museums was quite simply all the natural history collections that were in attendance at the 2009 NatSCA meeting, a list of 24 museums and collections. This sample set was chosen as aptly, the analysis would be presented at the 2010 NatSCA meeting.

Results

	How many species	What do they eat?	What kind of animal?	Where do you find penguins?	Do penguins fly?	Biggest Smallest	Notes
Birmingham Science museum	0	0	0	0	0	0	No results
Cliffe Castle	0	0	0	0	0	0	No results
Dundee University Museums							Timed out
Grant Museum of Zoology							Timed out
Glasgow Museums	0	0	0	0	0	0	No results
Hampshire County Museums Service	0	0	0	0	0	0	No results
Horniman Museums & Gardens	0	0	1	0	1	0	Gallery trail
Hunterian Museum Glasgow	0	0	0	0	0	0	No results
Hunterian Museum London	0	0	0	0	0	0	No results
Kelvingrove Art Gallery and Museum	0	0	0	0	0	0	No results
Museums in Leicestershire							Timed out
Leeds City Museum	0	0	0	0	0	0	No results
University of Manchester	0	0	1	0	0	0	Gallery guide
National Museum Wales	0	0	0	0	0	0	No results
Norfolk Museums & Archaeology Service	0	0	0	0	0	0	No results
National Museum of Ireland	0	0	0	0	0	0	No results
Natural History Museum London						1	Clicked out
Oxford University Museum	0	0	0	0	0	0	No results
Plymouth City museum	0	0	0	0	0	0	No results
Cole Museum							Timed out
Royal Albert Memorial Museum	0	0	0	0	0	0	No results
Warwickshire Museum Service							Timed out
World Museum Liverpool							Timed out
York Castle Museum	0	0	0	0	0	0	No results

Fig. 1. Table summarising the results of searching the word *penguin* on museum websites and the results that matched the six pre-selected questions.

As can be seen from Fig. 1, there was a relatively low success rate. Of the museums sampled, fifteen museums returned no results for the search term *penguin*. Six museums returned some results for the word *penguin* but none of the sample questions could be answered. Answers to the questions could only be found on three museum websites. An online gallery guide from the Horniman Museum detailed penguins as flightless birds. The Manchester Museum specified that penguins were birds as part of a gallery guide and The Natural History Museum listed somewhere on the site information about the biggest and smallest penguin, however, the rest of the search timed out after 20 clicks. Not only was the return rate relatively poor but the results that did come back ranged from classic penguin-based jokes through to results related to *Lady Chatterley's Lover* (presumably the Penguin edition).

Regardless, these results from the rudimentary analysis might be entirely moot because in addition to the above test, when the word penguin was searched on Google, museums did not even feature at least as far as the thirteenth page result before the author gave up. Incidentally, the Wikipedia page for penguins was the third returned result which also had 100% of the answers to the sample questions on one page.

Conclusion

As was demonstrated by this quick test, natural history museums are failing the web public. Although slightly better results may have been anticipated, the overall result should not come as a surprise. Museums have nowhere near the resources to compete with sites like Wikipedia and even if they did, every museum producing exhaustative resources on every taxon would be unfeasible and the duplication of work would be wasteful in a sector which already struggles for adequate resources. Rather than compete with Wikipedia, a website which receives billions of hits per year, natural history museums should look to exploit Wikipedia so that any relevant information about their collection will be brought to the attention of the people searching for it.

So for example, if your museum holds the biggest, smallest, oldest, fattest, localist, fakist or ugliest penguin then make sure that is on the Wikipedia page for penguins. If Darwin owned a penguin in your museum collections, then make sure that information is on the page for Charles Darwin and penguins. This could be achieved by employing a Wikipedian, recruiting a Wikipedian volunteer or train staff to add relevant information when it comes up. This will at least ensure that individuals looking for this information are significantly more likely to find it. Putting empirical information online is only one way of reaching the web public - but reaching is very distinct from engaging the web public which requires a completely different tactic.

Blogging the Museum

Aside from just making sure that information is available, curators should, if it is appropriate, be engaging the web public not through the presentation of ideas and facts but through, it will be argued here, an exposition of what it is like to work in a natural history collection. This is most easily achieved through the kinds of content that can be posted on weblogs (blogs). Here are the author's top ten tips for engaging the web public:

I. Audiences

It is important to think of the web audience as its own discrete audience and engaging with the web audience should be thoroughly thought through with the same level of planning as engaging with a new real world audience. Who specifically are you targeting? Is there an opportunity to share an audience? Is this audience already catered for in the competitive online space? How will you measure effectiveness? Putting content online does not automatically equate to being available to everyone. In some museums it may not be appropriate to engage web audiences either because resources do not allow it, existing audiences are sufficient or the web audience is not connected to the museum's mission. In which case a basic website with contact information is sufficient. Not every museum by default should be engaging with their web audience for reasons given below.

II. Frequency of updates

Frequently updated websites (generation of new content) are better than websites which do not feature a weblog at all, which in turn is better than a website where the blog is infrequently updated. Blogs which are not updated regularly will not keep your web audience engaged or keep them coming back. If a blog looks 'dead' then the audience is less likely to engage through leaving comments, recommending your site or tagging various entries for others to read. Some of the most popular blogs are updated hundreds of times every day by a team of full time staff. This keeps the content fresh and gives the audience something new to access. Also it allows a broader range of subjects and topics to be covered, meaning that the potential readership for your museum blog is greater.

III. Cross-media

The wonderful thing about weblogs is that they allow images, videos and audio to be uploaded. Both natural

history specimens and working with natural history specimens is a bit of an odd job, and inherently interesting so use every opportunity to 'show not tell'. Even with the best will in the world your content may well be competing with hundreds of other posts and articles through a feed reader or a site aggregator, so grabbing the attention of readers is as important as what you have to say. Fortunately, the nature of museums objects naturally lend themselves to photography and film.

IV. Sustaining an audience
Museums engaging with the web audience should be prepared to sustain the attempt to build up an audience over a long period of time. Building up a healthy audience (see 'check the statistics' below) can take years rather than months. Do not set up a weblog for a year-long contract funded project for example because more often than not the blog will not have built up an audience before all the staff associated with that project move on and the blog is closed, or worse, is just left un-updated. For institutions where natural history collections make up a larger network of museums, it is a good idea to set up a collaborative weblog which means that temporary projects can feed into a central blog which is sustained after individual projects are over. In addition, keeping the blog updated is easier, although catering to a mixed interest audience needs to be taken into account.

V. Check the statistics
Ideally, before embarking on blogging the museum, care should be taken to define how effectiveness, success, or failure, will be measured. There are a number of software applications like Google analytics, which automatically collate web statistics from counting how many visitors you received, how they got to your site, how long they stayed, how many pages they looked at and even where your visitors are geographically located. In annual reports, will hits to the website be counted on a par with physical museum visits? Is 400 hits a day with an average bounce rate of 3 seconds better or worse than 20 hits a day with an average bounce rate of an hour? Both represent the same amount of virtual dwell time. Another advantage of collecting the statistics is that it is very easy to see which topics or posts attracted the most attention or drew in the most new visitors, one type of automated visitor evaluation. This information is useful for planning future posts as well as optimising weblogging practice.

VI. Find a voice
Traditional interpretation in natural history collections is often based on empirical observations and is apersonalised, however, this voice may not work as well if blog entries are to be entertaining, passionate or even humorous. The institutional voice should be considered. Will staff blog as: themselves e.g. Mark Carnall, represent one aspect of the collection e.g The Curator, The Conservator, The Janitor or will they blog as a historical character e.g Charles Darwin. Establishing a voice will make it easier to keep entries consistent, decide what will and won not be bloggable and will also dictate how to deal with issues arising over blogging contentious subjects (a disclaimer of some sort is normal practice).

VII. Be Emotional
Related to the point above, this is only the author's personal preference but blog entries are much more interesting to read if the author is not constrained by writing impassionately or without bias which gives much more freedom for humour and enthusiasm. Furthermore, one of the brilliant aspects of the museum sector is that most individuals are highly enthusiastic about their subject, their areas of interest and their museum. Preventing this enthusiasm to be communicated seems counter intuitive. Without the human element, museum websites can devolve into computerized event listings, which is functional but will not necessarily keep readers coming back.

VIII. Internal Advocacy
Household products, events, historical characters and the trivial thoughts of the average tweeter. These are all things that have their own URLs, cynically indicating that the URL is its own form of communicative media. For large institutions, blogging your activity can be hugely useful for promoting internal advocacy. A successful event or event series will end up as a small part of a series of numbers, maybe with one or two qualitative comments on an end of year report; by linking to a post-event blog post with images or video a better sense of the event can be obtained. CPD and skill-sharing is also more easily shared if blog posts on various topics can simply be forwarded.

IX. Our Hidden Histories
Blog posts can act as a historical record of an event, a process in which the thoughts of museum staff otherwise would not be preserved. Comments from blog readers can also preserve a sense of cultural interaction

which is also often forgotten. Archived annual reports tend to preserve big institutional histories and database entries and logbooks preserve an element of the day-to-day activity of a museum. Ironically, museum curators are typically very bad at recording their own individual histories or the ethnographic history of their profession. Even less of this history is preserved now that electronic mail has all but replaced snail mail correspondence. Blog posts can give an insight into the character of individuals and preserve actions, thoughts and ideas which are too whimsical, brief or specific. Historians of the future would be able to make far more accurate and well-rounded reconstructions of the past if this transient activity is sometimes recorded.

X. Work together

Forming partnerships, communicating with colleagues and sharing skills is potentially made all the easier through blogging if museum workers get into the habit of reading each other's thoughts, sharing relevant links and resources and even debating with each other. This activity does tend to happen on professional web forums and other networks but the discussion tends to be 'behind closed doors'. If these interactions were held in a more public forum then there is a much greater scope for engaging the web public in how museums are work and how they are run. This provides a much greater transparency which is absolutely fundamental for ethically run public museums and could also lead to greater investment and engagement from key stakeholders.

Notes

The reliability of information on the web is a contentious subject. Wikipedia, for example, has been criticised on a number of occasions over the years on issues of accuracy. Anybody, upon registering an account can add information to the millions of pages that make up the electronic encyclopedia. Supporters of Wikipedia claim that the thousands of editors and transparent discussion pages lead to up-to-date, referenced and transparently created information pages. Dissenters argue that the process is too democratic, open to abuse and unreliable. As a counterpoint to the latter stance, readers who are concerned about the accuracy, reliability and transparency of information should compare the Wikipedia page for Human (http://en.wikipedia.org/wiki/Human) to how our species is listed on Mammal Species of the World (http://www.bucknell.edu/msw3/), the Tree of Life Web Project (http://tolweb.org/Homo/16418), the University of Michigan's Animal Diversity Web (http://animaldiversity.ummz.umich.edu/site/index.html), Michael Benton's VERTAPPENDIX (http://palaeo.gly.bris.ac.uk/benton/vertclass.html), Systema Naturae 2000 Taxonomicon (http://taxonomicon.taxonomy.nl/), the Integrated Taxonomic Information System (http://www.itis.gov/) and the Palaeobiology Database (http://www.paleodb.org/cgi-bin/bridge.pl). A number of these online databases, created by universities or groups of specialist professionals, are frozen in time, undated, unreferenced, hugely inconsistent with each other or a mix of two classification systems. For example, the Taxonomicon lists over 30 taxonomic ranks in the *Homo* tree, The Palaeobiology Database lists 4. The Encyclopedia of Life, seemingly pre-empting this analysis, just displays the Wikipedia page for Human (http://www.eol.org/pages/4454124).