Open Library of Humanities

Open Library of Humanities

Part of the Ubiquity
Partner Network

Interview

How to Cite: Oates, G and Whitelaw, M 2018 George Oates: Making and Remaking Collections Online. *Open Library of Humanities*, 4(1): 32, pp. 1–22, DOI: https://doi.org/10.16995/olh.325

Published: 07 June 2018

Peer Review:

This article has been peer reviewed through the double-blind process of *Open Library of Humanities*, which is a journal published by the Open Library of Humanities.

Copyright:

© 2018 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See http://creativecommons.org/licenses/by/4.0/.

Open Access:

Open Library of Humanities is a peer-reviewed open access journal.

Digital Preservation:

The Open Library of Humanities and all its journals are digitally preserved in the CLOCKSS scholarly archive service.

Open Library of Humanities

INTERVIEW

George Oates: Making and Remaking Collections Online

George Oates[1] and Mitchell Whitelaw[2]

[1] Good, Form & Spectacle, GB

[2] School of Art and Design, Australian National University, AU

Corresponding author: Mitchell Whitelaw (Mitchell.Whitelaw@anu.edu.au)

George Oates is a web designer, producer and developer who has been working with digital collections since 2003. She was one the founding team of Flickr, and the originator of Flickr Commons. She currently heads heritage design studio Good, Form & Spectacle, working with institutions including the British Museum, the Wellcome Library and the V&A. In this interview with Mitchell Whitelaw, Oates discusses her developing practice in making and remaking digital collections, spanning web-scale services, unsolicited interfaces, messy metadata and gentle interventions in institutional practice.

Interview

MW: George, you are currently director of Good, Form & Spectacle, a design studio specialising in cultural heritage—but you have been working with digital cultural collections for some time—including for the Internet Archive, and as part of the team that founded Flickr, as well as establishing Flickr Commons. Can you say a bit about this part of your background outside of the 'institutional' sector? What kinds of ideas about digital collections were forming there, and are these still influential in your work now?

GO: Yes, indeed… I've been working in web software design and development for about twenty years now—hard to believe, I'm sure! After running my own small web design business and cutting my teeth designing intranets and things in Adelaide, I left Australia in 2003, the day before the Iraq War was declared, actually. At that time, I created a goal for myself: to travel the world, and learn all I could for the next ten years, then return home to start my own shop. That's gone pretty much to plan, except I'm in London now, not Adelaide.

I joined the Flickr team in late 2003, before it was even Flickr. They were making an online, community-oriented game called GameNeverending (GNE), centred around instant messaging, exploring a nodal environment, and exchange of objects for mutual benefit and advancement in the game (**Figure 1**). There was a happy community of ten thousand users or so, who ended up forming the nascent community at Flickr, after we decided to change direction to photo sharing when we realised it was going to be too hard to make GNE a viable business. When we made the transition, the game mechanics were left largely untouched, and the game objects were replaced with images, which people were able to pass to one another. I can still vividly recall the day when [Flickr developer] Eric Costello passed us a photograph of his kids playing at the beach from New York (to Vancouver, where the company was based). It was intensely instant and personal. Early sharing of imagery was often as a punchline to a joke, but interestingly started becoming much more personal more broadly, as internet friends shared their actual lives (and, well, meals).

Figure 1: Screenshot from GameNeverending taken from a screencast of the game being played. https://www.flickr.com/photos/caseorganic/6382153979. Attribution: Amber Case on Flickr, CC BY-NC 2.0. https://www.flickr.com/photos/caseorganic.

As we continued to develop the web system, we explored the idea that looking at photographs was such a personal and gestural thing. You often tell stories using photographs, maybe spreading them out on a table, or flicking through an album together. You describe and you annotate, and you point out that she is Auntie Madge, and that's her backyard in about 1950. We acknowledged and embraced that photographs are *social objects*, often to be enjoyed in groups, and described both by the photographer and the viewer. We were also on top of the wave of digital photography becoming ubiquitous and instant, and saw pretty early on that this

warranted developing tools to help people describe and organise their photos, *as they were published.* One of the crucial interface decisions we made was that new photographs were marked public by default. This was in stark contrast to the competitors of the day, which felt more like bank vaults where you may not see another human, let alone have them see one of your photographs. The 'collection' of photographs grew and grew, and interestingly, a happy, utterly organic information system blossomed, as photographers described the subjects of their pictures as they saw fit, using tags. Thomas Vander Wal coined the term 'folksonomy' in 2004, the very early days, and even today, this folksonomic structure is present and dynamic at Flickr, in various languages across just about anything you can photograph (Vander Wal, 2007). We were *never* declarative about how people should tag their photos, and yet, structure emerged.

It was some time in the summer of 2007 when an email trickled into my inbox from the Library of Congress (LoC). They were writing to us to enquire about Flickr becoming their 'Web 2.0' partner. Even after having a huge photo archive online at loc.gov for over 10 years, they were looking for another place to share their treasures, and engage with new audiences, and thought of us. It was tricky to just say yes, mostly because the photographs in their collection were earmarked as 'no known restrictions', but at the time, Flickr offered either fully copyrighted, or the selection of Creative Commons licenses. Being a library (and not explicitly a collecting institution), the LoC wasn't able to match these licenses. So, I set to work with Yahoo! Legal to see about adding this additional classification. It wasn't a license as such, but more of an assertion, where in the case that a photograph may be found to be copyrighted somehow, it could be removed. It was then that I realised that probably every institution grappled with this challenge; that it's not always possible to declare rights to share imagery easily, and indeed, it's often a *total nightmare* to work out provenance and share things freely. That's when I had the idea to broaden the work we were doing with LoC into a larger program, which I called The Commons on Flickr (Flickr, n.d.). We launched with LoC in January of 2008 (Oates, 2008a), to a huge spray of help and annotations from the Flickr community

in the first 24 hours and beyond (Oates, 2008b), and in that first year, I think I helped about 35 institutions from around the world join the program, resulting in exponentially more views for their collections, and some really detailed, factual and rich research and reporting done by fans of the initiative. As far as I know, it's still going strong today, with millions of 'no known copyright restrictions' resources, and loads of activity and engagement. It was also during 2008 that I was bitten by the cultural heritage bug, mostly because I got to meet lots of folks who work at museums, libraries and archives around the world, and, well, I fell in love. Sadly, in its ultimate wisdom, Yahoo! decided to lay me off at the end of 2008 (Johnson, 2008). That made me even more interested and concerned about the nature of this enormous photographic archive that basically describes the planet from 2004 until today, and the fact that it's owned by corporate interest. That's possibly for another article.

Knowing that I was now interested in the cultural heritage/digital crossover, I didn't want to work in a commercial setting, and it seemed like the best place for me in San Francisco was the Internet Archive, a big non-profit whose mission is to provide universal access to all knowledge. No biggie. Somehow I convinced Brewster Kahle to give me a job as the director of the Open Library project (Open Library, n.d.). It's like Wikipedia, but for book metadata. Anyone can edit, fix or add to the 30 million or so bibliographic records. So, I set about designing and building version two (after the initial design and build done largely by Aaron Schwartz et al. some years before I turned up). I worked to incorporate much of what I'd learned designing Flickr: show activity, design for exploration (and not just search), develop better communicated developer tools and APIs, and look for ways to automatically improve the data we had (for example, by implementing the existence of what are called works in book-land, where a work is the primary book, of which there are many editions, per a mouthful of a standard called Functional Requirements for Bibliographic Records, or FRBR). It was a baptism of fire in the world of library metadata too, which I still find profoundly interesting, mostly because cataloguers in libraries need to be able to describe a *copy* of something, and that requires some

kind of standardised description of things, to help make sure that your 1934 Russian edition of *War and Peace* is the same as the one in that other library. I like that that's a different challenge to describing a unique object, which you find a lot more in museums, and especially in archives.

I got to work on some other projects at the IA too, which I also enjoyed, namely: the Understanding 9/11 TV News Archive (**Figure 2**), where I designed a really pragmatic interface to browsing the week of TV news surrounding the terrible events in New York (Internet Archive, 2011); design sketches of the broader TV news archive; and a revamp of the Wayback Machine, where I was proud to institute the little browser navigation thingy that let you easily skip across versions of sites from throughout that site's history (Internet Archive, n.d.). I also enjoyed the sheer scale of the collections at the IA... with millions of things of various kinds, it's a huge archive, and I was sorry I wasn't able to get my hands on the main site at archive.org before I moved on.

From there I moved to the award-winning design shop called Stamen Design, still in San Francisco, joining the small team as art director. Back in the
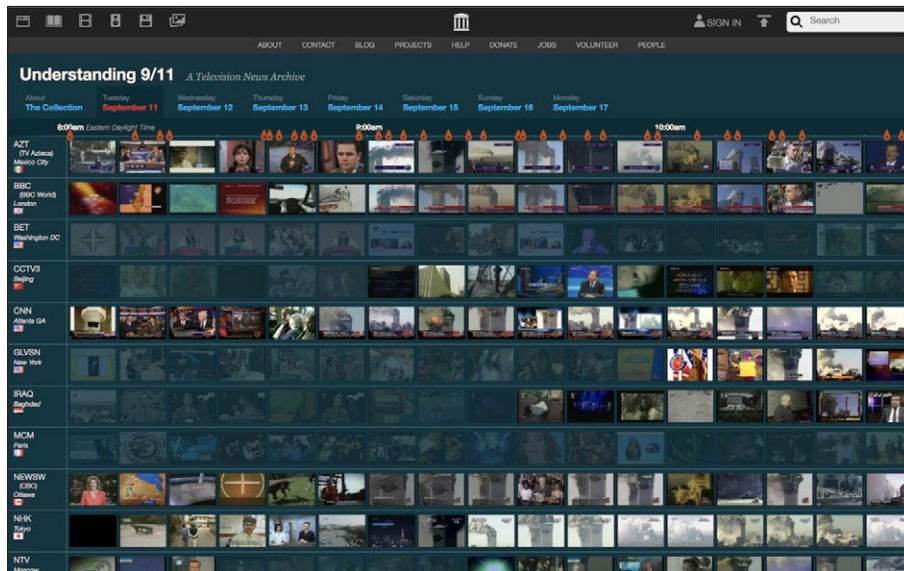

Figure 2: Understanding 9/11 TV News Archive interface.

largely commercial realm, I think I worked on about 50 different projects for as many different clients in my time there. That pace was refreshing and tiring all at the same time, and it gave me unique visibility into the datasets of those various organisations, and a nuanced understanding of what it means to design visualisations for data comprehension. We worked on a ton of cartographic projects too, which I grew to love, as placing things in or on the world can also lead to a different understanding of the data to hand, and is often a quickly understandable technique to introduce a new viewer to a new dataset, because you can ask 'what's it like where I live'. Some of the projects I particularly enjoyed there were:

- maps.stamen.com, where we delivered openly usable map tiles for the world in a few different styles (Stamen Design, n.d.), and then its accompanying child project Map Stack, which was sort of like Photoshop, but for maps (Stamen Design, 2013)
- Field Papers, which was a handy tool to gather and print out an 'atlas' for any bounding box in the world, so it could be annotated physically and then imported into the fabulous OpenStreetMap project (Stamen Design, 2012a)
- Surging Seas, a tool to show the effect of sea-level rise overlaid with demographic data in that area (Stamen Design, 2014)
- The City from the Valley, a visualisation of data we collected ourselves showing how many of the tech slaves of Silicon Valley commuted to their work from San Francisco in private corporate shuttles, instead of using public transport (Stamen Design, 2012b).

Sometimes you can only recognise a path in hindsight, and looking back, there's certainly continuity in my work from its beginnings, even in Adelaide as a web designer and information architect, to these gigantic social content systems, to visualising data, to arcane institutional information systems, to now.

Ideas that formed through these years of designing and being part of huge content communities include:

· See the whole from its parts
· Look outwards from a single object along all its connective tissue
· Show activity around objects
· Get to know the grain of a dataset before you represent it; be aware of design's editorial power here
· Relinquishing control can be liberating
· Leverage hyperlinks to make networked objects
· Clear constraints breed creative responses
· Benefits of cross-collection aggregation remind me of the classical typological arrangements in museums, apparently invented by Pitt-Rivers back in the day.

Then, I moved to London to seek my fortune, and started my own shop that is designed to blend my two main interests: software design and cultural heritage, and here we are.

MW: In your recent work with G,F&S you've created some powerful and provocative demonstrations of how digital collections can be 'remade'. Your 'spelunkers' are large-scale collection browsers that provide quite radical redesigns of institutional collections. Not only in their appearance, but in what is shown: how collections are sliced and faceted. Can you say a bit about how this line of work originated?

GO: The very first one we made, Netflix-o-matic, was a direct result of a journalist called Alexis Madrigal exploring Netflix genres (**Figure 3**) (Good, Form & Spectacle, 2014). He was trying to reconstitute them, which I personally found quite weird, especially when I just wanted to explore them!

Interesting responses from the good folk of Metafilter, too, actually *wanting* some randomness:
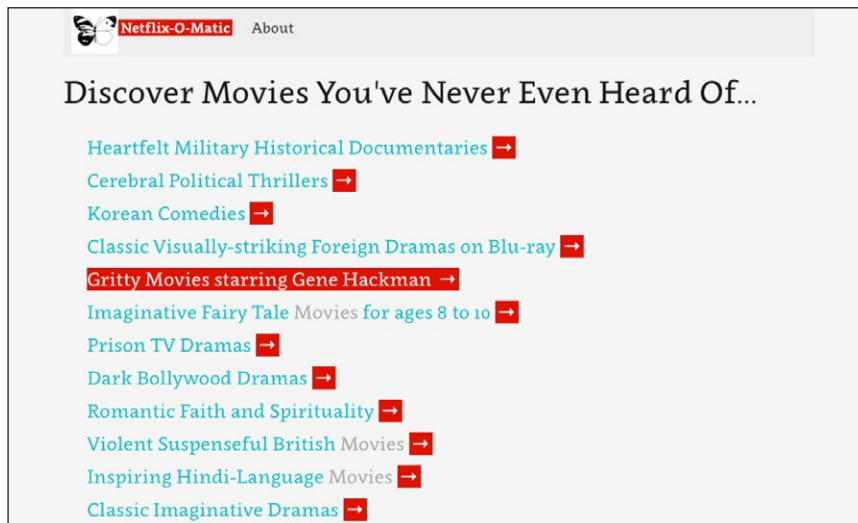
Figure 3: Netflix-o-matic homepage.

> I wish my Netflix recommendations were made up of 50% random stuff like
> this. It would be way better than seeing the same 40 recommendations in
> 20 different categories. Yo Netflix; I am never watching *Full House* so you can
> stop trying to convince me (Metafilter, 2015).

I think these interfaces are powerful because they're not directive. They don't say
'look at the things we think are most important' but instead say 'follow your own
path'. I love the idea that every single person using these has a different path through
the collections. That's something I grew to love at Flickr.

It's shocking to me that online cultural experiences still often only show tiny
thumbnails of things or maybe even no image at all without several clicks. Sure, if
there's no digitised version, there'd be nothing to show, but if there is, *why make it
tiny*? So all the work I've made in this area provides big images, which is a very easy
way to make the exploration more satisfying.

The design decision to prioritise object descriptions over all other fields on Two
Way Street led to more interesting metadata being seen more quickly (**Figures 4,
5** and **6**) (Good, Form & Spectacle, 2015). The stuff in the description field is really
informative and interesting, and shows the might of the British Museum experts
who look after all this stuff for us.

Figure 4: Two Way Street: here's the key visualisation of the system. Every square is a decade in the life of the institution, shaded by number of acquisitions that decade. Also lists where the acquisitions came from, a subject rarely publicised in museum land.
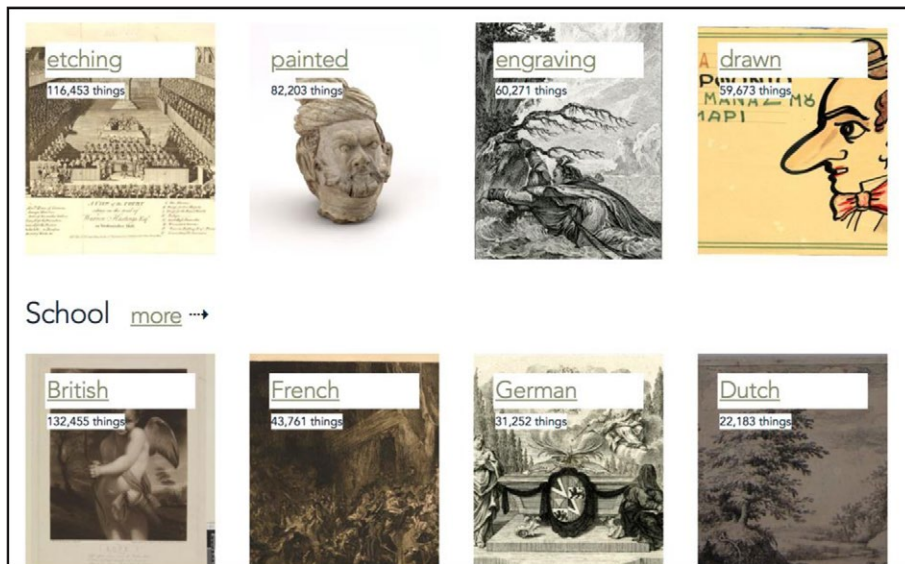

Figure 5: Two Way Street: Show Everything. Here we see ordered lists of the Type of Thing and School facets displayed visually.
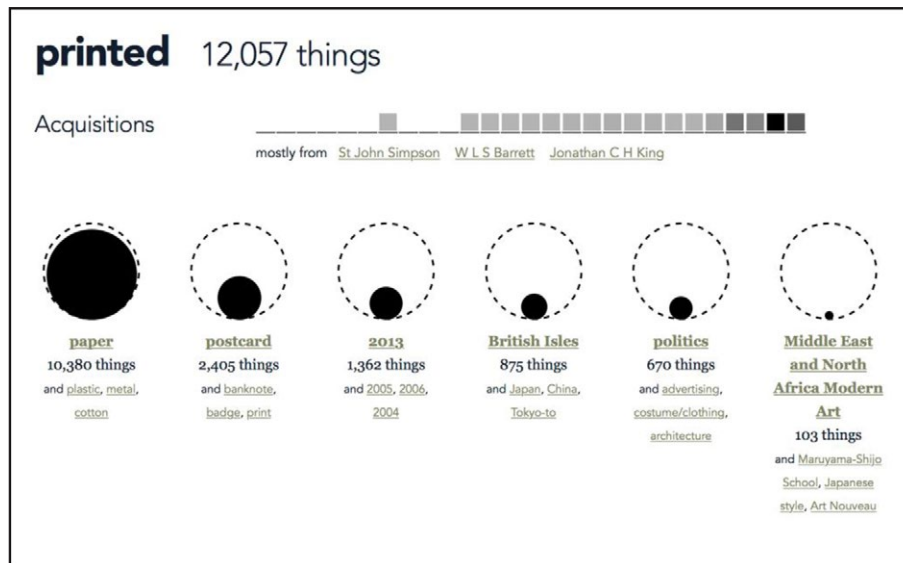
Figure 6: Two Way Street: a facet view, showing acquisition trends and visualising composition of the facet.

MW: Here you're starting outside the museum sector, and recognising the commonalities between the British Museum and Netflix. This is a key insight, but one that might challenge some institutions. How do you see this relationship between heritage collections and the wider world of digital content collections, from iTunes to art.sy?

GO: Companies born after the web came along are in a luxurious position compared with those that came before. There's far less of a struggle to account for and display your wares if you've been digital from the outset. Even though there are certainly still some data issues with newer organisations, if your metadata is 'born digital' you already have a huge advantage over organisations who have to operate with a blend of records on paper, index cards, and any number of other formats in their catalogue management. For example, from day one at Flickr, every item in the 'collection' was immediately put into the digital infrastructure. There was never a need to translate from physical to digital.

If you look at a company like Amazon, or any online retailer, they know where everything is within an inch of its life. They know when stock is low, or items aren't moving, so can control the flow of 'objects' because they know precisely how many of everything they have. It's easy to take for granted these days, but the insight everyone has into global supply chains, whether we're consumers or suppliers, is radical compared to what we knew even twenty years ago. It makes perfect sense to know where all your goods are because you're trying to *sell* them and make money, not just *keep* them for posterity, or *lend* them. You might say the commercial imperative has an influence over the efficacy of collections management.

Another massive change that came with the advent of Web 2.0 was the idea that we could edit databases live online. All those early websites we loved were all *alive*, and users in the various systems were writing new data to the database constantly, whether that was a comment on a forum or a photo to see or a crazy rainbow horizontal rule (<hr>) on a webpage somewhere. It was all alive—and *that* is an extreme difference between the pre-web legacy systems and what we have today. I've met quite a few cataloguers who have to deal with a central software client that they don't control and struggle with because the interface is terrible, and they may not even be able to get their own data out when they want it. Curious how that same challenge we as consumers now face—of not being able to get your data out, or delete it—is now held by hugely powerful tech corporations like Google and Facebook. The corporation as collector? (I guess that's my PhD waiting for me.)

There's also something in all this about the origin stories of collections. Lots of museums and libraries began as an individual's own collection, or have incorporated an individual's collection over time. I think there's something hiding in plain sight here about the challenge of organising a mix of idiosyncratic collections into one cohesive whole under one roof. Long before our appreciation of the need for so-called universal identifiers like ISBN or DOI, collectors and their assistants were left to their best devices to describe and organise collections. And even though we've been good at storing and organising things for thousands of years, as David Weinberger describes in his fabulous book we should all read again, *Everything is Miscellaneous*, we've all done it a little differently, and individually (Weinberger, 2007). For example, we're lucky

to have our office in the former home of Sir Hans Sloane, a physician and collector alive around 1700. He lived in this place for almost 50 years, and it's a real thrill to imagine that bits and pieces of his collection would have crossed the same threshold we do when we come to work. His collection was bequeathed to the British nation upon his death in 1753, and indeed, went on to form a core part of the initial British Museum's collection. The fun part is, if you look at Sloane's catalogues, even though he did a great job of at least recording *something* about the things he collected, some of the catalogue entries are scant at best! E.g. *A Roman soldier, A horse galloping, Three chisel heads of different sizes, one very small* (British Museum Collection Database, 2017). But even as he logged many of his acquisitions, however briefly, there are lots of accounts of people visiting Bloomsbury Place to meet him and hear him talk about his collection. What a treat to hear a collector like Sloane describe his treasure! And what a chasm between his effusive, hospitable tours and his catalogues!

I've always enjoyed reading about collectors, and as you can see from this snippet about another prolific, voracious collector, Henry Wellcome, they're a breed not especially known for their organisational skills: 'Wellcome's insatiable appetite for acquiring objects outweighed any inclination to sort, catalogue, or even use the books regularly arriving in his stores' (Wellcome Collection, n.d.). Similarly, in Walter Benjamin's charming essay *Unpacking My Library*, we read about a haphazard, hungry book collector whose collection 'is but a disorder to which habit has accommodated itself to such an extent that it can appear as order' (Benjamin, 1968: 60). Fast forward to the twentieth century and you have a favourite collector of mine, Peggy Guggenheim, prescient and opportunistic, collecting the modern greats before they were great, and certainly recording her purchases and buzz about her exhibitions and collection in scrapbooks and ledgers. (I'm trying to gain access to whatever has been digitised of these so far, currently obscured by Finding Aids online.)

Bring it back to now (and your question), and you have myriad digital systems like Pinterest or Pinboard or eBay that allow thousands of collectors to gather and present things in a fixed, born-digital system however they like. The point is, the system and data structure/container are shared by everyone who uses them from the outset.

MW: I've called your collection projects 'unsolicited interfaces', because in cases like TwoWay.st they were just that, made without the prompting or support of the collection holder. This seems like a powerful and in some way performative creative strategy; can you talk about how this came about, and how these unsolicited interfaces were received, either by their 'target' collections or others?

GO: I like that you're calling this work performative, because the first few we made were certainly that. I was a new business owner in a new city (London), and wanted to make demonstrations of our capability, to show our wares and use them to attract clients.

I was heartened to know that staff at the British Museum enjoyed using Two Way Street, and had even found new things in their departments as a direct result! This is excellent news. The site enjoys about 100 sessions in a day, and those sessions last for an average of four minutes, which isn't bad for something that isn't particularly marketed! About 25% of the visitors are coming back too, which is great. I've also enjoyed the odd random query about one of the objects from people writing to us as if we were actually the British Museum (maybe trying to sell us an object, or report an error, etc.).

Our original idea with Two Way Street was actually to create a space where conversation around objects could happen, hence the name. I suppose we may still do that at some point, but, having experienced huge community at Flickr, I know how much time and attention is required to make a healthy online space, and, given the sometimes controversial nature of the BM's collection and its extreme potential for repatriation, I decided not to pursue the original concept, and stopped at the spelunker.

MW: Since then you've made quite a few 'solicited' spelunkers, including for MoMA and the Wellcome Library. These are now authorised products but they retain a strong G,F&S feel, as well as a very particular voice in the written copy. Can you talk about these collaborations—what do institutions value in the 'remaking' that you are offering?

GO: One of our strengths is that we can operate quickly, and enter a data and presentation challenge with a 'beginner's mind'. We are unbiased consumers of institutional metadata, and this has a bunch of benefits. Perhaps it's even about creating some sort of 'outsider artist' point of view to liberate the organisation from the classic constraints of correctness or towing their official line. I've been catching up on the most recent set of talks recorded at MuseumNext, and I think the keyword used the most is *trust.* Institutions are still struggling to qualify what this actually means, or its potential. When we come in tasked with re-presenting a collection, especially in a design-led process, it's not always clear what the outcome will be when we start, and that requires trust in us. If it's us presenting information, perhaps there's less onus on the institution to bear responsibility?

We've now seen more than once that re-presenting big metadata in these exploratory aggregate views actually allows data creators to see their own data in a refreshing way. Often their operations are on single items, and seeing the whole is rare, if not impossible, since often digital administrative tools are unsatisfying, difficult legacy systems. This lack of exploratory tools in the more conventional software systems is pretty outmoded, and certainly doesn't help anyone easily see where their metadata could use some attention. In the Wellcome *What's in the Library?* project, for example, we were able to demonstrate that, of a possible 180 MaRC fields used across the million or so records, only about three were used 100% of the time, and even then, that was for system-generated IDs and such (**Figure 7**). I was sorry we didn't have the chance to work for longer than a week on that step actually, because I was curious to see if we might be able to recognise individual data operators in the visualisations we were making.

Something we're also useful for is to demonstrate how a dataset might be used or exploited: to road-test it. Often, I think, institutions will manage to package up their data for release (which is great!), but are maybe a little myopic when thinking about its usefulness and usability. Why would I, a lone roving developer looking for something fun to do, try to operate on gigabytes of some CIDOC-CRM RDF rats' nest of a dataset that requires me to have really specific schema knowledge and certainly a powerful machine to even move that kind of data around? Having an experienced
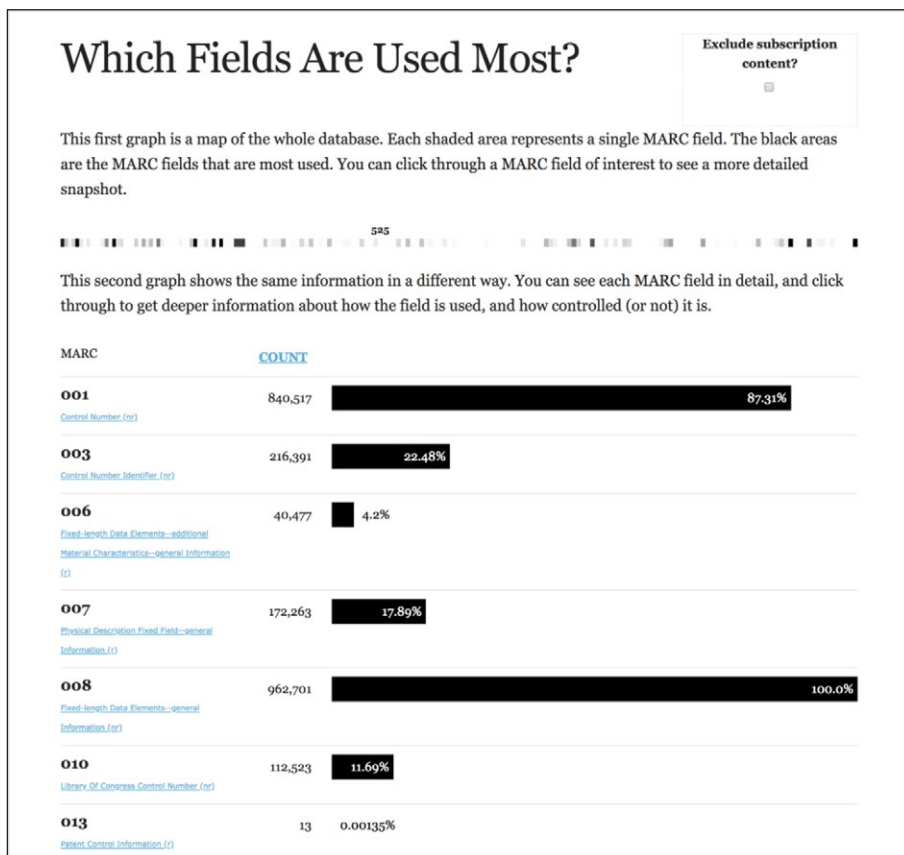
Which Fields Are Used Most?

Exclude subscription content?
☐

This first graph is a map of the whole database. Each shaded area represents a single MARC field. The black areas are the MARC fields that are most used. You can click through a MARC field of interest to see a more detailed snapshot.

525

This second graph shows the same information in a different way. You can see each MARC field in detail, and click through to get deeper information about how the field is used, and how controlled (or not) it is.

MARC	COUNT	
001 Control Number (nr)	840,517	87.31%
003 Control Number Identifier (nr)	216,391	22.48%
006 Fixed-length Data Elements--additional Material Characteristics--general information (r)	40,477	4.2%
007 Physical Description Fixed Field--general information (r)	172,263	17.89%
008 Fixed-length Data Elements--general Information (nr)	962,701	100.0%
010 Library Of Congress Control Number (nr)	112,523	11.69%
013 Patent Control Information (r)	13	0.00135%

Figure 7: What's in the Library? Week 1: Which fields are used most? Available at: http://scope.whatsinthelibrary.com/fields (Last accessed 16 April 2018).

group come in and ask simple questions like 'can you export your data in a variety of formats?' and 'how easy is it to point at digital resources?' and 'are the usage rights crystal clear?' etc. can be really helpful in designing and shaping the usability of the data. It's especially exciting to see some institutions now being much more deliberate and comfortable with putting their metadata out into the world, and also building on each others' work for the greater good.

This also results in one of our main goals for G,F&S, and that's to help show clearly a type of working to internal folks. It's about showing what can be built and how much agency it's possible to create under your roof when you have good software folks on your team. In some ways, our mission is to put ourselves out of work because our clients will go ahead and hire software people who are *builders* and

not just content producers, to make tools and hack stuff for the institution, instead of it spending vast amounts in fees to some software company in California with slow turnaround and one-size-fits-all products. We show what happens when you have a design-led approach, and ask new questions, producing rough work in a world of marble, using contemporary toolsets and practice.

MW: Many of your spelunkers, such as the Waddeson Bequest Explorer (**Figure 8**), use facets to group collections into usable clusters. Facets need consistent metadata to work well, but collection metadata is notoriously messy and patchy. How do you deal with issues of data quality?

Figure 8: Waddesdon Bequest Explorer: materials facet. Available at: http:// wb.britishmuseum.org/materials (Last accessed 16 April 2018).

GO: Actually, we try to not to alter the data we're working with in any way. We've developed a few different visualisation techniques around data 'strength' for various projects. It's about trying to show, across the whole dataset, which records are 'strong', where strong is a measure of coverage in all possible fields.

Once we can see this dynamic in visual terms, it helps us and the data creators to see the whole dataset in a new way. I hope there's potential in this technique for data creators to get new focus on areas of metadata which would benefit from new or additional attention. I also like the idea that you may be able to recognise individual operators in visuals like this. This kind of work is so often invisible to the naked eye, but where a lot of heavy qualification and intellectual rigour resides. On the flipside though, we've also seen that a lot of metadata is remarkably thin, and that, specifically, contributes to a difficult and unrewarding digital exploration experience.

I also really enjoy seeing the true mess of metadata that's created by humans. I think it's the status quo, and we should all probably stop trying to deny its existence. It's present in every cultural dataset I've ever examined. I love it because it shows the humanness of this work. Humans always find ways to work around rigid things, like water flowing around rocks. If there's a required field in some software somewhere, you can be damn sure there'll be some arcane notation or code or other field springing up around it because the metadata doesn't fit its container.

This mess also often manifests as several different data systems within one institution, and that's a challenge I'm really interested in at the moment. Linked data initiatives are all well and good for institutions with technical capacity, but there's often a more pressing precursor challenge to this, which is to interconnect various data systems that have sprung up internally and organically. It's a delicate thing and usually requires consensus and maintenance.

Related, but not quite the same, are the growing troves of researchers' data, created around or on top of or to circumvent data offerings from institutions. I've certainly done that thing where you crack open your own spreadsheet of data and then start moving things around or adding columns. There's a real opportunity there to supplement, interconnect and enhance institutional metadata here, but it's similarly difficult to ingest or blend.

At Open Library, I designed a tool to help connect different representations of authors too, which I'm pleased to say worked pretty well. It was a basic search feature that looked for variants in author names, and then provided a user interface for people to select the different variants that should probably be collected together to represent the author. Importantly, we didn't throw away the alternates, but stored them in the 'primary' author record as alternates, so records could still be connected to these variants. (I'm surprised this doesn't happen more, actually.) This too, is a reflection of the pretty basic belief I have around digital objects: that you need as many points of entry to them as you can gather. I enjoy thinking of this using the analogy of surface tension. The more connection points you have to each object, the more visible it will be. Compare a Flickr photo with 300 tags in a variety of languages belonging to 40 groups and marked with 1000 favourites to a photograph with a single word title and a maybe a date in a cultural collection somewhere… and see which is easier to find versus which sinks into the digital darkness.

MW: This view of metadata—as inevitably messy and flawed but also human—is at odds with the dominant narratives around digitisation and digital culture, where metadata is clean and consistent, and digital services are functional and transparent. Your approach reminds me of Georgia Lupi's manifesto for Data Humanism (Lupi, 2017). It also suggests a different kind of institutional 'voice'; the tone of the copy in your spelunkers echoes some of the work that Aaron Cope and Seb Chan did with Cooper Hewitt, where interfaces openly exposed missing images or rights restrictions. I don't think it's a coincidence that, like you, Aaron Cope has a background in technology and design outside the museum sector. In summary it seems that your work is intervening (however gently) in some big agendas around technology and cultural heritage. Would you agree?

GO: Did you know Aaron used to work with me at Flickr? We had many a long chat about all this stuff, and we collaborated on lots of Flickr features (like single-sign-on and machine tags). I'd say we've definitely influenced each other over the years. We also both worked at Stamen Design, though we didn't overlap, unfortunately. We used to argue about things like the beauty of RDF.

I hope my work is a gentle intervention. I'm trying to demonstrate ideas and show ways of doing things. I'm especially pleased that the work and thinking around the Alpha I developed for the Wellcome Library (Oates, 2016) has made its way into their major collections redesign (Scott, 2017). I'm also hoping that all of this thinking can be instantiated in the new company I'm building, called Museum in a Box. It plays off all of the concepts I've been developing around networked objects, multiple points of view, dynamic descriptions and digital from the start (Museum in a Box, n.d.). Who knows, maybe it might even become some kind of Web 5.0 collections management system if we can build it into a viable business with happy customers. Or if that doesn't work, perhaps there's a forward-thinking museum or library out there who might like me on the team as a Creative Director. I'd be well up for that, though I think one reason I've been able to be gently interventionist is because I'm on the periphery, and can jab big institutions right in their pudgy bits from afar. I shall leave you with that image.

Competing Interests

George Oates has no competing interests to declare. Mitchell Whitelaw is a Co-Editor of the Special Collection "Remaking Collections," for which this interview was conducted.

References

Benjamin, W 1968 Unpacking My Library. In: *Illuminations*, Trans. Zohn, H and Arendt, H (Ed.). New York: Schocken Books.

British Museum Collection Database 2017 Am,CUPBD2/SH.4. *British Museum*, n. dat. Available at: http://www.britishmuseum.org/research/collection_online/collection_object_details.aspx?assetId=561984001&objectId=3181835&partId=1 (Last accessed 16 April 2018).

Flickr n.d. The Commons. *flickr.com*, n. dat. Available at: https://www.flickr.com/commons (Last accessed 16 April 2018).

Good, Form & Spectacle 2014 Netflix-o-Matic. *Good, Form & Spectacle*, n. dat. Available at: http://netflixomatic.goodformandspectacle.com/ (Last accessed 16 April 2018).

Good, Form & Spectacle 2015 Two Way Street. *Good, Form & Spectacle*, n. dat. Available at: http://twoway.st/ (Last accessed 16 April 2018).

Internet Archive 2011 Understanding 9/11: A Television News Archive, August 2011. *The Internet Archive*, n. dat. Available at: https://archive.org/details/911 (Last accessed 16 April 2018).

Internet Archive n.d. Wayback Machine. *The Internet Archive*, n. dat. Available: https://archive.org/web/ (Last accessed 16 April 2018).

Johnson, B 2008 Now Flickr is Hit by Yahoo Layoffs. *The Guardian*, 12 December. Available at: https://www.theguardian.com/technology/blog/2008/dec/11/yahoo-flickr-layoffs (Last accessed 16 April 2018).

Lupi, G 2017 Data Humanism: My Manifesto for a New Data World. *Georgia Lupi personal website*, n. dat. Available at: http://giorgialupi.com/data-humanism-my-manifesto-for-a-new-data-wold/ (Last accessed 16 April 2018).

Metafilter 2015 Netflix-o-Matic. *Metafilter*, 19 May. Available at: https://www.metafilter.com/149771/Netflix-o-Matic (Last accessed 16 April 2018).

Museum in a Box n.d. Museum in a Box: Object-oriented Experience Design for Museums. *Museum in a Box*, n. dat. Available at: http://www.museuminabox.org/ (Last accessed 16 April 2018).

Oates, G 2008a Many hands make light work. *Flickr blog*, 16 January. Available at: http://blog.flickr.net/en/2008/01/16/many-hands-make-light-work/ (Last accessed 16 April 2018).

Oates, G 2008b Wow. *Flickr blog*, 17 January. Available at: http://blog.flickr.net/en/2008/01/17/wow/ (Last accessed 16 April 2018).

Oates, G 2016 New Work: Wellcome Library Alpha. *Good, Form & Spectacle*, 3 June. Available at: https://goodformandspectacle.wordpress.com/2016/06/03/new-work-wellcome-library-alpha/ (Last accessed 16 April 2018).

Open Library n.d. Welcome to Open Library. *Open Library*, n. dat. Available at: https://openlibrary.org/ (Last accessed 16 April 2018).

Scott, T 2017 Building a Free and Unrestricted Digital Museum and Library. *Wellcome Collection*, 27 October. Available at: https://stacks.wellcomecollection.org/

building-a-free-and-unrestricted-digital-museum-and-library-6cdf28784068 (Last accessed 16 April 2018).

Stamen Design 2012a *Field Papers*. Available at: http://fieldpapers.org/ (Last accessed 16 April 2018).

Stamen Design 2012b The City from the Valley. *Stamen*, n. dat. Available at: https://stamen.com/work/the-city-from-the-valley/ (Last accessed 16 April 2018).

Stamen Design 2013 Map Stack. *Stamen*, n. dat. Available at: https://stamen.com/work/map-stack/ (Last accessed 16 April 2018).

Stamen Design 2014 Surging Seas 2. *Stamen*, n. dat. Available at: https://stamen.com/work/surging-seas-2/ (Last accessed 16 April 2018).

Stamen Design n.d. maps.stamen.com. *Stamen*, n. dat. Available at: http://maps.stamen.com/ (Last accessed 16 April 2018).

Vander Wal, T 2007 Folksonomy. *Thomas Vander Wal personal website*, 2 February. Available at: http://www.vanderwal.net/folksonomy.html (Last accessed 16 April 2018).

Weinberger, D 2007 *Everything is Miscellaneous*. New York: Henry Holt and Company.

Wellcome Collection n.d. The Unceasing Seeker. *Wellcome Collection*, n. dat. Available at: http://digitalstories.wellcomecollection.org/pathways/2-the-collectors/4-unceasing-seeker/ (Last accessed 16 April 2018).