

Hi. I'm a creative industry person.

Thank you to Johan Oomen and the Europeana team for inviting me to speak to you today.

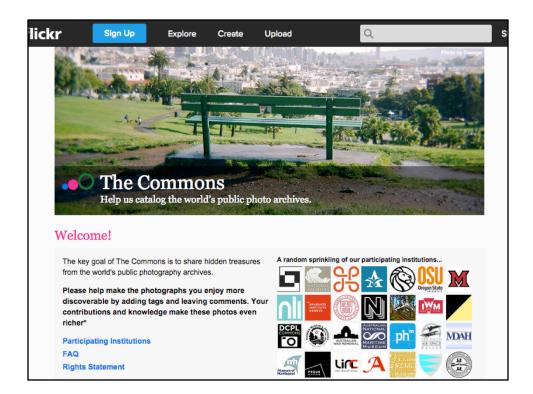
I'll start with a quick introduction to me and the work I do, and then dive into the stuff that Johan asked me to think about, and what it's like as a creative trying to work with the data you're making.

https://creativecommons.org/licenses/by/3.0/



Ten years ago, I took a photo of my screen while I was at work, in April, 2004. I was in Vancouver working on Flickr. I started working there before Flickr was Flickr, and mvoed from Vancouver - where it started – to San Francisco when we were acquired by Yahoo in 2005.

Flickr turned 11 this week, and there are now billions of photos in it. I've lost count, and fear for it's longevity.



I created the Commons in 2007-2008, working initially with the Library of Congress. The program is almost seven years old now, and there are about 100 institutions online, from across the world. In many ways, it was a dam bursting, and it's an absolute thrill to see Commonses popping up all over.



I had a fantastic time developing that project. I met hundreds of people who work in museums, libraries and archives around the world, and was completely entranced by their knowledge and generosity. Like Gerry, Graham and James from National Maritime Museum, where I got to spend some time curating a collection for inclusion in their Commons presence.

I was laid off by Yahoo at the end of 2008. That really sucked. Like, really sucked.

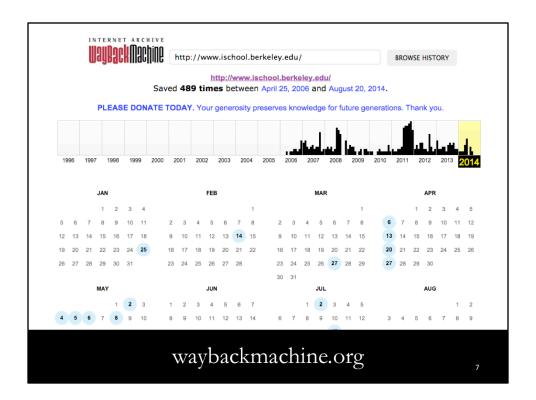


Then I went to the Internet Archive in San Francisco. I was the first designer they'd employed. I designed posters like this one, but my "real job" was...



Leading the Open Library project. I designed the logo and the interface and managed the overall redesign, and an international team of 5. Open Library is like a wikipedia for books; an editable library catalog. There are about 30 million metadata records (probably more today), and the Internet Archive has scanned millions of books, some of which are availlable to borrow, if they're not in the public domain. People using OL can edit records, and they can also make lists of books they like, and interesting edits like merging duplicate authors.

It was this project where I started to explore what it might be like to build interfaces to big metadata systems that had no search box. It worked as analog to the physical library experience, I thought, where you often look across the spines of adjacent books, check over a bibliography, or simply wander the stacks in search of inspiration.

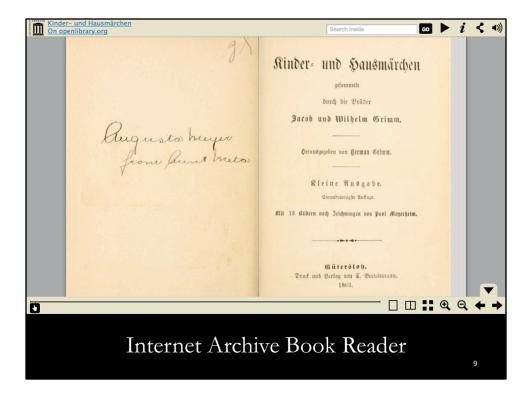


I tried to help design as many things as I could when I was at the Archive. Here's a redesign I did of the Wayback Machine. I also designed a small widget that displayed on archived pages so you could navigate between the various dated crawls without having to come back to this index.

I have no idea how many petabytes of data the Archive holds to represent the web since 1996. It's a LOT.



This was a pilot for the TV News Archive, which showed footage across about 30 international TV channels on the week starting on September 11. It's a visual search interface for that week of TV. A thumbnail represents 10 minutes. Click that, and get 10 thumbnails, one for each contained minute. Click that and you're watching one minute of footage.



This is the browser-based reader I made with Mike Ang. It's nice to know that other people have made use of it across the web, like the New York Philharmonic and the Smithsonian Institution.

I left the Archive in late 2011.



Then I joined a studio in San Francisco called Stamen. It specializes in mapping projects and data visualization. I was the art director on about 60 projects. That was too much too fast for me, though some people love that variety.

(My photo.)



This is a project I art directed for Facebook. It's an animation of the shares of a photo by George Takei on Facebook. He used to be on Star Trek, and is inexplicably popular on Facebook.

I enjoyed writing up the process here: http://content.stamen.com/facebook-flowers



When I was there we released this project under a Knight Foundation grant, and I suggested we put these three map tilesets under a Creative Commons license. They've since spread all around the web, and it's always nice to see them in the wild! I found myself more drawn to projects like this, with some "public service" element to them.



I also designed Field Papers, a way to make a paper atlas of anywhere in the world, in particular to use for making annotations that could later be put into Open Street Map, and following on from my colleague Mike Migurski's project called Walking Papers.

http://fieldpapers.org



This is another favourite project, called parks.stamen.com. We used the geographical boundaries of all the parks in California to search in some big social media services to see what people were looking at and talking about. I left Stamen in June 2014.

http://parks.stamen.com

I like

- Talking to humans who do the work
- Radical access to huge datasets
- Making things that people can use

15



I decided a few months ago that it was time to start m own design firm, and to plant a flag firmly in the shifting sands of cultural heritage to tell it that I want to be there. It's very new and young and now I've moved to London. But, it's bloody exciting.

http://gfns.uk

(My photo)



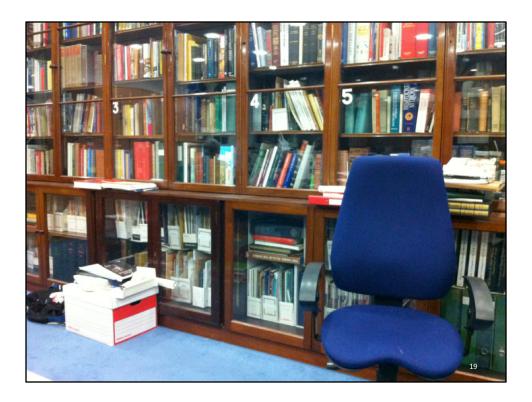
Now that I live in London, I'm closer to more cultural stuff than I have ever been. It's magic, and I've been trying to do frequent visits to institutions around the city. I've keeping photographic evidence of this on Flickr: https://www.flickr.com/photos/george/sets/72157645445279853/. I've logged 41 visits so far, in San Francisco, London, Washington, Los Angeles, Paris and a few in not-London England.

This is TATE Britain. http://www.tate.org.uk/visit/tate-britain

(My photo)



This is the amazing Tower of London Remembers: http://poppies.hrp.org.uk/
(My photo)



I also made an appointment to meet the Keeper of Tower Histories at the Tower. She's worked there for 37 years, loves swords, and gave me a couple of hours of her time, during which I said very little and learned a great deal. The menagerie at the Tower was NUTS.



I've been enjoying the English cuisine.



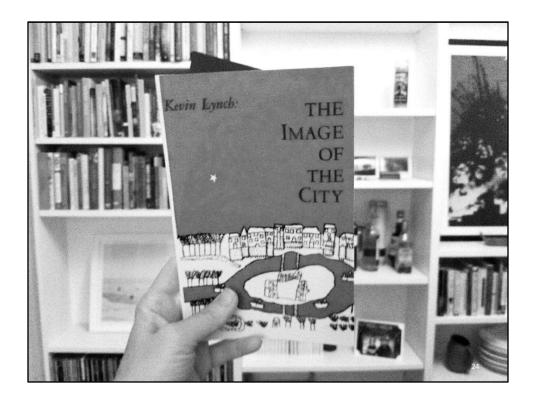
And this new ritual of afternoon tea is something I can really get behind. This was a practice a couple of weeks ago at my new favourite place, Maison Bertaux in Soho. If you come to London, let me know, and I'll meet you there.



So... Openness is Wilderness, eh?



Johan gave me a few topics to think about for this speech... and the theme of wilderness made me pick up a book I'd read a few years ago, when a few people in the interaction design cabal discovered it.



The Image of the City
Published in 1960 by MIT Press

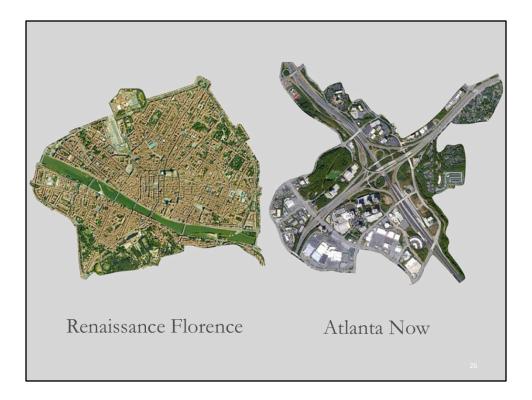
What does the city's form actually mean to the people who live there? What can the city planner do to make the city's image more vivid and memorable to the city dweller?

I've been re-reading this book and reframing it as "The Image of the Digital Manifestation of a Cultural Institution" and its really interesting and fun. Cities are covered in clues and paths and signs and other cues to help us look and wander and enjoy the place. We memorize how to get from one place to another thanks to visual cues and landmarks. We congregate and co-exist, and call some cities home.



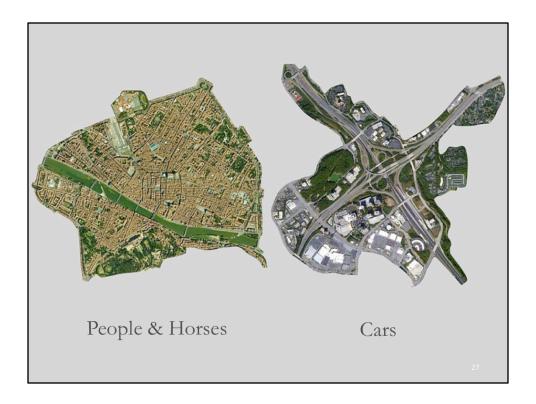
Here's another image to keep in your mind for the talk. This image is from a sustainability and urban design blog called The Original Green, by Steve Moozon. In 2011, he wrote a great piece about The Cost of Sprawl.

http://www.originalgreen.org/blog/costs-of-sprawl---the-speed.html



Each area is half a mile wide, contains about 90 acres

http://www.originalgreen.org/blog/costs-of-sprawl---the-speed.html



"The Speed Burden"

As speed of transport increases, you need longer curves, wider lanes, and, well, basically, there's nowhere to sit and look.

The entire Duomo fits in one of those loops in the middle.



The point I want to make here is that I worry we're doing similar things with data. Exchanging data and aggregating data is still good work, but, I've heard a lot over the past two days that "we'll put it online and you can do what you want with it" and that just won't work.

We need orientation.

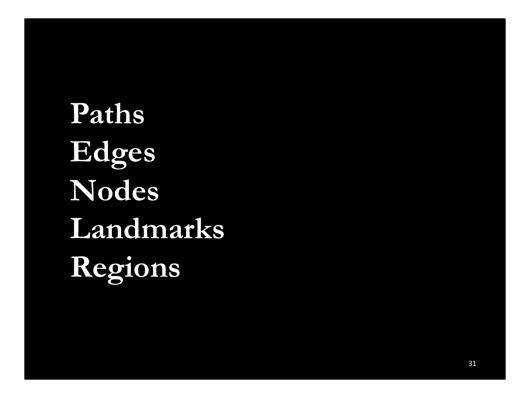


Incidentally, this is a crazy, empty ghost town in Angola. Interesting to read into it. Your search terms are "Angola ghost town chinese". I've just oriented you. Have you noticed how we've started passing search terms around to each other? Like one-to-one wayfinding help.

http://www.businessinsider.com/chinese-built-ghost-town-kilamba-angola-2012-7? IR=T#-2



Even in the wilderness, there are landmarks we have used through the ages to orient ourselves. Like that giant mountain in Utah I saw one day. I bet there are all kinds of local stories about it. Indeed, Lynch says "striking landscapes leads to myth making."



We're good at looking at things.

Google is good for searching the whole internet, but, if you're working with a collection or catalogue, maybe that kind of of EVERYTHING for ANYTHING isn't the right fit. Consider the *form* of your content. What are its features? How can you help people learn about it? What should you show them as landmarks or paths in?

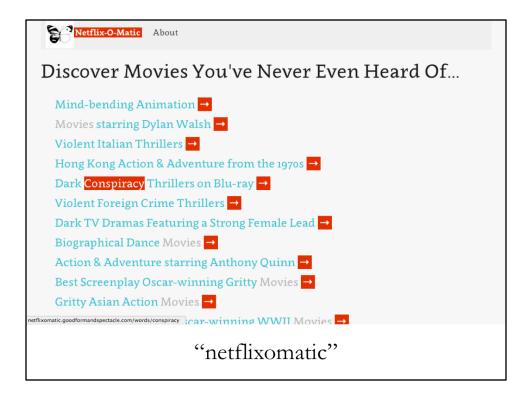
A search box is basically illegible if I don't know what you are trying to show me, or what you have in your collection. It shows me no scale or direction, there is no limit or emphasis. Limit and emphasis is what cultural heritage is all about.



So, how could we explore trying to find a legible "image" of our catalogues?

I've been working with the idea of no search for some years now. It was a big inspiration behind the interface of the Open Library project, back in 2009, when I tried to conceive of a UI that felt like wandering the stacks. We make so many discoveries through adjacency and accident. The enforced search pattern makes that practically impossible.

Reframing your metadata to be navigable without search will help you make paths and edges to it. I've made a two projects in the last few months that continue to explore this idea.



The Netflix genre space is surprisingly charismatic, or something. It's full of entertaining one-liner descriptions, but they're not surfaced on the website, in the viewing experience. So, with Frankie Roberto and Dan Williams, I made Netflix-O-Matic in a day.

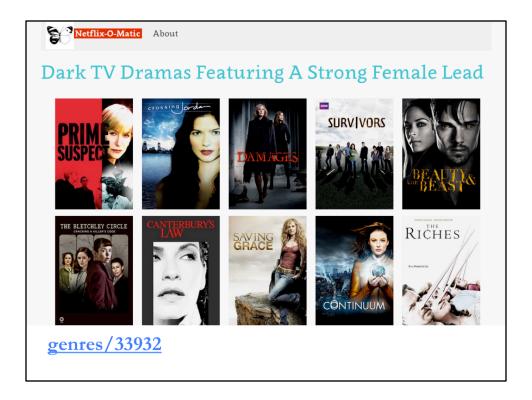
You can hit a single word in this list, or...

http://netflixomatic.goodformandspectacle.com/

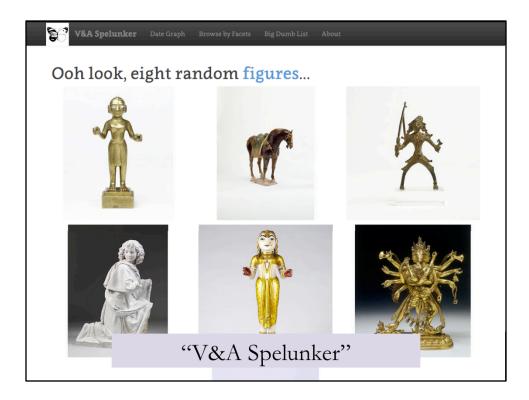


Select the whole line using the ARROW icon, to get to a list of titles.

http://netflixomatic.goodformandspectacle.com/

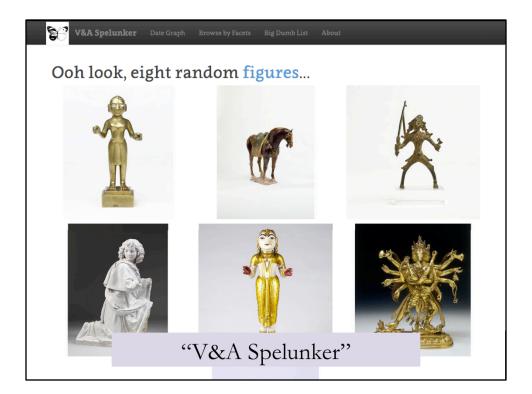


It's an easy, lightweight, no search method that gets you quickly to films and TV you self-select.



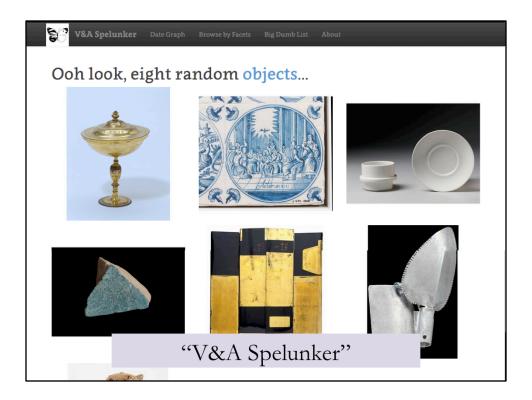
This is the V&A Spelunker, which I developed at the end of 2014 with Tom Armitage. Andrew Lewis at the V&A gave me a USB with the catalogue on it after I met him for lunch one day.

Tom and I are both huge fans of a technique I basically always use when I'm working with a new dataset is to SHOW EVERYTHING. That's what we did.

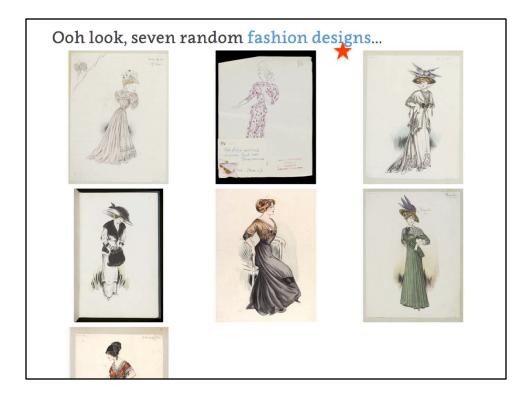


It's a quick way to explore the contours of a dataset. Link everything to everything, show big pictures.

You can see up the top there, there's a single facet grouping present at the top, which lets you dive into figures if you're curious.



At the bottom of the page there's a big link that says "AGAIN!" which just redraws the page with another random sample from another type of thing.



Andrew, who works at the V&A, told me that even just that single filter addition was interesting... They've had a "random sample" UI for ages on the official V&A site, but just this little filter does a couple of things... it gives visitors a glance at the TYPES of things that the V&A has, and a path to help them see more of them if they want to.



Within about two seconds I'm looking at this delightful furnishing fabric of acanthus scrolls and stars...

http://va.goodformandspectacle.com/things/9037



From there I can look at all the stuff from England http://va.goodformandspectacle.com/things/877



And then all the clasps



And now I'm in Dalmatia, and can see that the V&A has 23 things from Dalmatia in the collection.

http://va.goodformandspectacle.com/things/13724

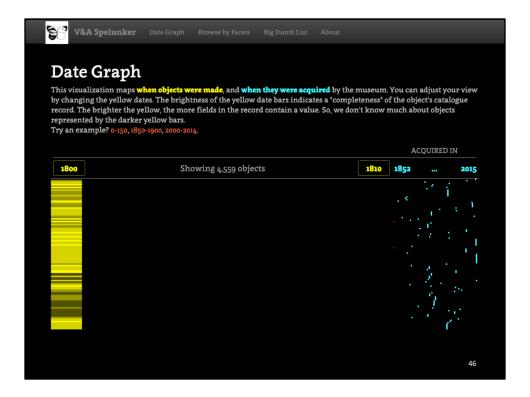
```
/things/9037
/facets/place/England
/things/877
/facets/object/clasp
/things/116234
/facets/place/Dalmatia
/things/13274
```

What can we do with this? This path, this USE. This seems like the sort of data which we might be able to use to show things that people like or think are interesting.

I'll come back to that later.

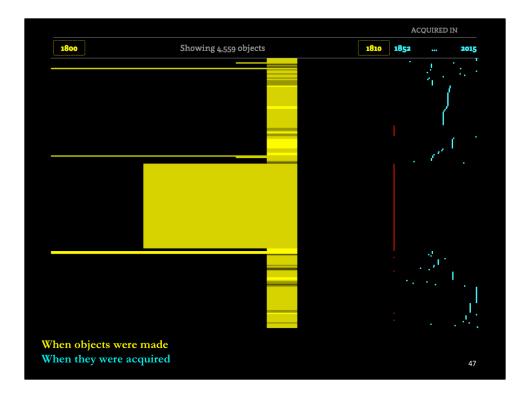


What happens when we try to be explicit about the overarching contours of a data set? Does it reveal peaks and rivers we can use to explore and navigate?



This Date Graph was part of the V&A Spelunker.

I was very pleased to write it up for the V&A blog too: http://www.vam.ac.uk/blog/digital-media/the-va-spelunker

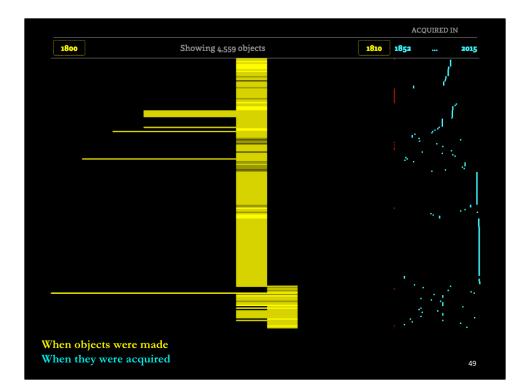


It's actually pretty simple, I think. It takes 3 inputs: the dates when objects were made (can be a range, start – end), and the year something was acquired.

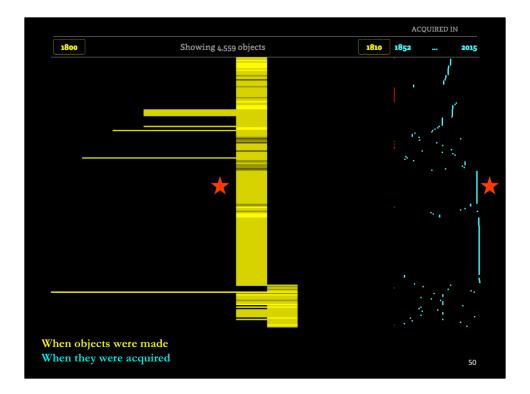
This snapshot is looking at about 100 objects – one object per row – and things that were made between 1800-1810. You'll also notice there's different yellows. Brighter means that the records are more "complete" where complete just means there's something in a database field. Darker yellows mean less fields.



We can also show where there's no known acquisition date, as in this big blob. I wonder if this kind of illustration could be useful for catalogue QA, or to help decide where to prioritise improvements to documentation.



You can also start to see patterns in acquisition. This could indicate an individual curator's collecting proclivities, or overall themes for the institution. It would also be interesting to map deaccession on something like this but we didn't have that data.

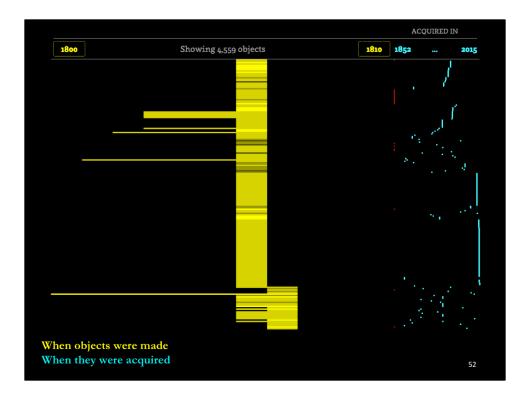


You can see blocks of activity like this one, in 2009.



If you dig in a bit, you can see there were a bunch of H Beard Collection prints, like this one

http://va.goodformandspectacle.com/things/1019825



Maps like this reveal stories, but stories that outsiders are only starting to guess at.

descriptive_line

Caricature of Richard Brinsley Sheridan as Humility and Effrontery. Published in London by S. W. Fores in December 1806, Harry Beard Collection.

physical description

Etched and hand coloured caricature of Richard Brinsley Sheridan as 'Humility or the Canvassing Candidate' and 'Effrontery or the Candidate Returned'.

public_access_description

Caricature of Richard Brinsley Sheridan as 'Humility or the Canvassing Candidate' and 'Effrontery or the Candidate Returned'. Published in London by S. W. Fores.

5

Incidentally, on that H. Beard print, I found these three fields in the data. These are so similar, I wonder why they're all there.

Perhaps it's an opportunity now for us – well, you – to shift the focus from just continuing to correct and improve the **information**. We already have too much of it, and it just permanently gums up the works!



We've just begun lighting the beacons. "We've got this data!!!!" – "Yeah, we've got it too!" "Or, no, wait. Ours is a little bit different..." "But we're both talking about John Milton, right?" "Yep." "OK, cool. I'll tell the computers."

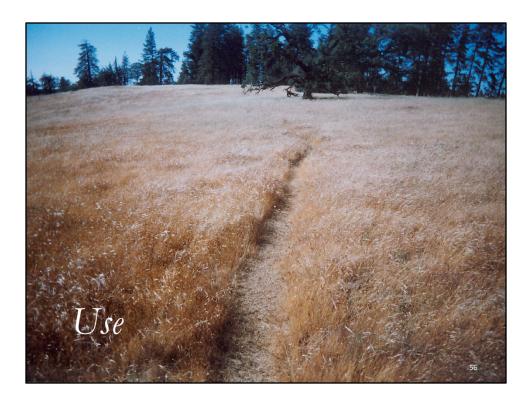
It feels like there's still a big gap between the work that tech folks do working hard on making our information better, more correct, more robust, less duplicate, more accessible, translated, APId, opened, indexed, and what goes on out on the floor, in the world, for visitors, designed and described and crafted by the experts in the content and physical realms.

Joris was telling me this morning that Europeana tech has 40 million records in it now, and that's taken just 5 years! That's AMAZING! And I don't want this to sound like that isn't good work. It is. Only that maybe it's time to start thinking about helping people out there access it.



I've heard this word used a lot to describe the work of Europeana, and other similar organizations and communities... But... what's impact? Is it like an earth-destroying cataclysm??!?

Those 40 million records are... well... not impactful yet. They're difficult to use... even for someone like me who has worked with data for ages and has lots of clever friends.



I think it's far more practical, and interesting, to work out how to calibrate impact today to **use.**

It wasn't clear from very many of the presentations I've heard here that anyone had ever used the things that you've made. That's probably not true, but it indicates to me that it's not a particularly a priority to know.

Yesterday, Tim Sherratt talked to us about the mythic mega project, and I think it's related. It's s really different undertaking and design mindset to try to create two new visitors who love what you're doing than to geotag 4 million newspapers. Work out who is using whatever you offer today, and see if you can improve that number.

(My photo)



So... how do you do this? If you're not as well-funded as lucky Seb Chan and his merry band... :P

In the discovery session yesterday, I was so excited to hear from Jaap Kamps, Akihiko Takano and Paul Clough that they're thinking along the same lines. Especially with regards to what's happened with search online.

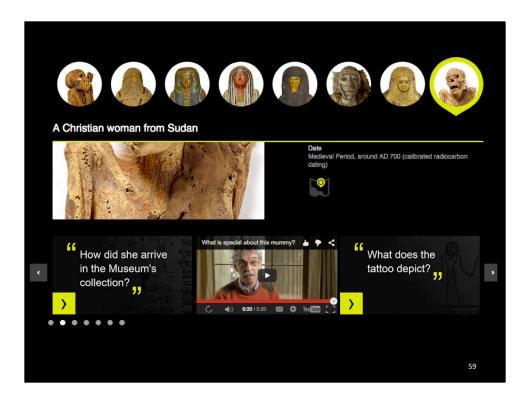
The Google search technology is amazing, but it's amazing at searching EVERYTHING for ANYTHING. We have a different challenge, to help people look for SOMETHING SOMEWHERE.

That's why I'm a bit nervous about the giant empty data cities, because we're not paying attention to the SOMETHING we have. We're just throwing stuff in SOLR and sticking some facets on it and ordering it alphabetically. Yes, there is sometimes great utility in an alphabetical list, as Daniel Pitti said yesterday, but our cultural institutions are so good at helping us know which paintings we should look at. Can you imagine if TATE Britain did that in a gallery? (Actually, that might be fun.)



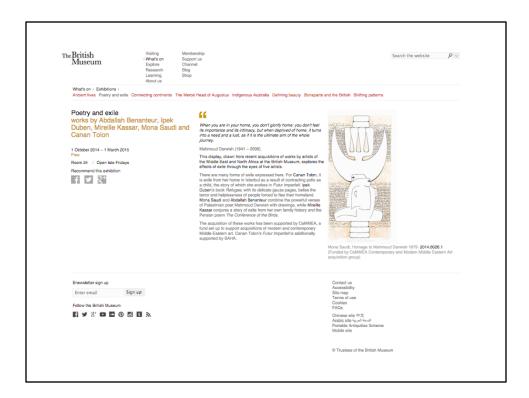
I'm talking about knowing the contours of your collection or catalogue, and even better, understanding its use, and adapting to that. For museums in particular, I think working with exhibitions more closely might be a good way in to this idea. It's surprising how few cultural websites online show the contents of exhibitions, and yet that's where so much effort and knowledge is created.

(My photo.)

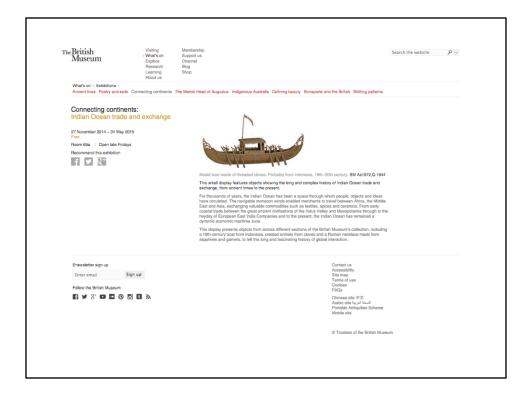


Let's take a look at an example.

This is a current exhibition page for "Eight Mummies". The content is fantastic! There's a ton to look at, and videos that have experts explaining things, which I really like. It's always magic when you get to meet people who work with the stuff.

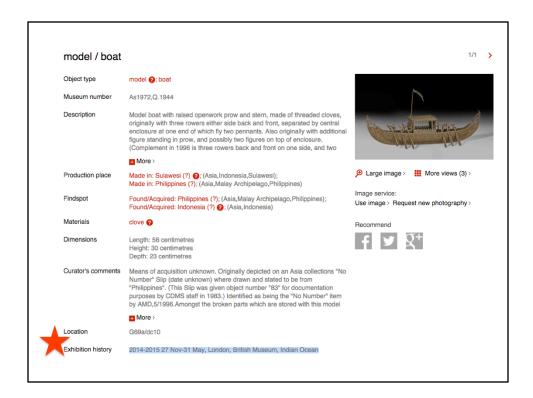


Then I click on the next current exhibition, about Poetry and Exile. There's nothing to explore. It's a dead end.

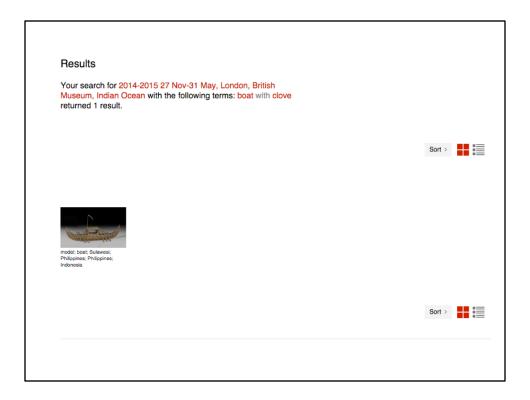


And another, about the sorts of trade and exchange that happened in the Indian Ocean – an incredible subject area.

From the single link on that page, I get to the boat's collection record, and have lost the exhibition.



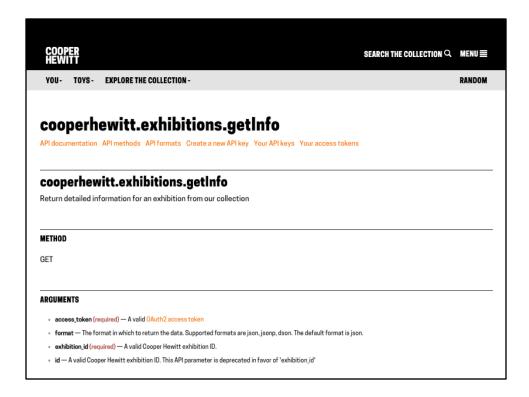
Oooh! There's an exhibition history field! But, it's not linked. There's no way to see more things in the exhibition, not even using the Advanced Search!



So I searched for the string on the exhibition history field in the collections search. (Does that make me a researcher?) \odot

I just saw that same object.

It gives me absolutely nothing to go on to explore what I suspect would be a fantastic exhibition. It's completely inaccessible.



I'm excited to see that the CH has exhibition data available in its API. The BM doesn't have an API.

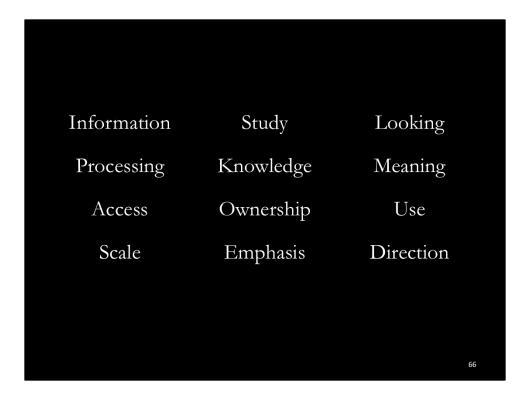
Congratulations to Seb and his team at Cooper Hewitt. Seriously, you should study this work and copy it.

TECH	EXPERTS	HUMANS
Information	Study	Looking
Processing	Knowledge	Meaning
Access	Ownership	Use
Scale	Emphasis	Direction
		65

- I'm wary of these distinctions between types of people
- Researchers OR General Public

I've been thinking about this kind of matrix a lot. I generally prefer to think about system design in terms of active language and things to do instead of personas. Certainly, I've heard lots of people here talk their work either being for Researchers or The General Public, maybe.

I stay away from people who want me to design things for The General Public. It means they have no idea who they expect to use it, or how.



- Try not to make these distinctions
- Think about USE(s)
- "Build it so anyone can contribute any amount"

I don't think we need to worry TOO much about the distinctions we all tend to make between experts and amateurs or enthusiasts or whatever we call them. We're seeing time and time again that normal humans are extremely capable of making deliberate, extended, constructive effort to improve our cultural heritage. In fact, more visible and constructive effort that any institution has made itself in that time, at least in public.

And this has all happened in less than ten years.

Like Andy told us yesterday, Digital NZs decision to be ok with imperfect metadata and to be open about it has served them in good stead. Trove too, and Tim there reminds us about the brilliant experts out there who help tremendously, and the other folks who help with simple labour (which can be so helpful).

It was Clay Shirky who wrote "build it so anyone can contribute any amount".

"The city is a multi-purpose shifting organization.

"The form must be somewhat non-committal, plastic to the purposes and perceptions of its citizens."

The Image of the City Kevin Lynch

67

Can we make digital instantiations of cultural heritage like this? (I'd like to try.)



Here's what's hopefully coming next.

The Basic Plan for Two-Way Street

- Spelunker-ish, whole catalogue
 Path defined by upcoming exhibition, *Defining Beauty*
- Allow public comments
 Set community guidelines; moderate
- Watch for "attention features" Expect reshaping/re-presentation based on activity

69

You might like to try to watch the video for this bit. It's where I said that I'm worried that I won't even be able to work out how to use the British Museum's data. It's not been made with other people in mind, or public use in mind, and is therefore very difficult for anyone other than its creator to use. But, I have his email address, so there's hope \odot



In terms of designing a path, I think it'll make sense to start with a contemporary exhibition, in this case, Defining Beauty, which is opening soon. http://www.britishmuseum.org/whats_on/exhibitions/defining_beauty.aspx

Using those 150 objects or so as our starting point, it will help us present edges and paths into the rest of the catalogue, I think. Plus, hopefully, we'll have some people leaving comments on things. That's always a good thing to display.

I'm considering buying some offline advertising too, to help people find it. But, we have to get the data first.

Does the British Museum know about this? No. One of the core motivations behind this idea is to press on the idea of open. You're either open, or you're not. Open data means people will do stuff with your data that you didn't anticipate or intend. This can be scary, but, it can also be really creative and invigorating. (That's what I'm going for.)



The reaction may be something along these lines. (VIDEO)

https://www.youtube.com/watch?v=AOhDksmjOO4



The world out there is already way ahead experimenting with meaning and the meaning of worth.

Banksy sold some real Banksy works on the street in New York City recently. \$60 on the street, \$250,000 at auction.

Gareth Williams, head of contemporary art at Bonham's, said the Central Park stall was a coup. "The fact that his paintings were original and were being offered at a tiny fraction of their true retail value, raises real questions about the perception of worth and the nature of art as commodity within the marketplace – something that the artist must be acutely aware of."

http://www.theguardian.com/artanddesign/2014/jun/12/banskey-prints-new-york-stall-fortune-bonhams



Studies centered on Indonesia and Nigeria. Jakarta is the mobile device epicenter on the planet. Everyone has at least one cellphone, and probably not any other device. And Facebook thinks it's the world's fourth largest country. Interesting article about what it means that a corporation is someone's first and only access to the internet.

Facebook is so powerful because it's full of people. People guiding each other, watching each other, judging each other. Every structure that's sprung up like apps or games are used by people.

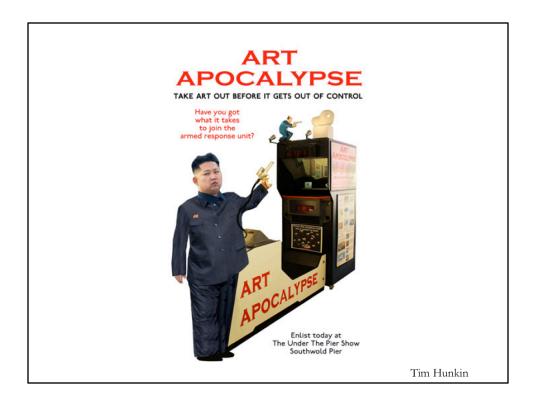
http://qz.com/333313/milliions-of-facebook-users-have-no-idea-theyre-using-the-internet/



I'm quickly becoming a fan of Tim Hunkin.

"Art Apocalypse is a game where you shoot art – a real shooting gallery. To cater for different tastes, there's a variety of different types of art to shoot at."

http://www.timhunkin.com/a128 art-apocolypse-intro.htm



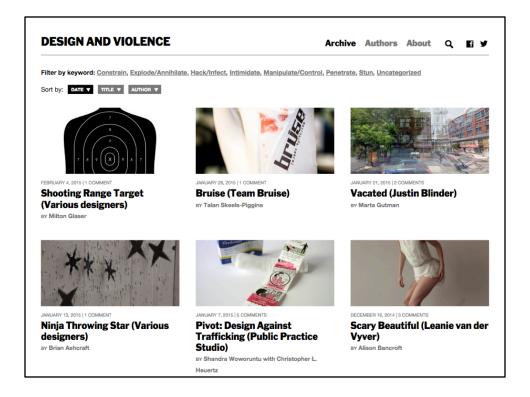
It's a real game.

http://www.timhunkin.com/a128_art-apocolypse-intro.htm



Here's the little chap on top of the game, weapon at the ready.

http://www.timhunkin.com/a128_art-apocolypse-intro.htm



Here are a couple of quick examples of digital work I think leans towards an image of a digital place (instead of a search box).

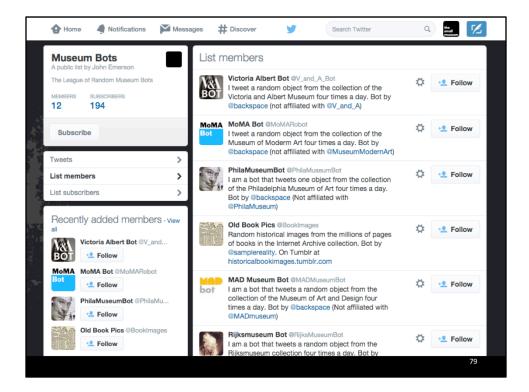
Curated set of essays by interesting people on a theme. Design and Violence at MoMA. I presume this is leading to an exhibition, and may even be a map of things that are being acquired. Wouldn't that be interesting?!?!?

http://designandviolence.moma.org/archives/



A tour of madness, murder and mental healing by the Wellcome Collection.

http://digitalstories.wellcomecollection.org/pathways/1-mindcraft/index.html



There are people chipping away at our giant cultural data offerings... I don't want to make it sound like there aren't. But, even these fun museum bots are directionless. It's like a shot across the bow telling you "we don't know what to do with your stuff, so we'll make something random".

This can be a nice way to make discoveries, but....

https://twitter.com/backspace/lists/museum-bots/members

(John Emerson is a programmer in Brooklyn. He didn't make every single one of these 12 museum bots.)



The new "digital river" of collection objects flowing past on the new tables at Cooper Hewitt. I loved that they called it River. When I got to Paris for the conference, I basically got to the river and turned left

http://waldrealestate.com/tour-cooper-hewitt-now-open-91m-overhaul/

Do you know how many rivers have cities on them?



Artist Maira Kalman

We're so far away from this feeling online.

https://www.cooperhewitt.org/events/opening-exhibitions/maira-kalman-selects/



With physical stuff, and our eyes, giant globs of stuff are OK. We can look.

Like in my new favourite gallery at the V&A. It's basically all ceramics ever. Organized by geography I think.

(Photo by me)



But the data stuff feels like this.

http://www.businessinsider.com/chinese-built-ghost-town-kilamba-angola-2012-7? IR=T#-2

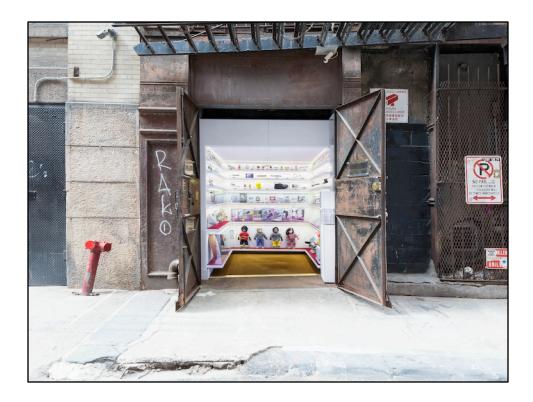


But if you're not the V&A, and you don't have ALL THE CERAMICS...

This is the Mmuseumm, in New York.

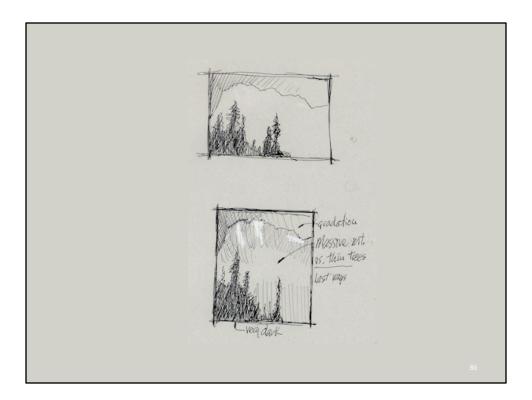
http://www.theparisreview.org/blog/2014/11/10/cabinet-of-wonder/

"Roughly speaking, there are natural-looking things on the right side of the museum, unnatural-looking things on the left—on the right are squashed mosquitoes, for instance, and what would appear to be a display of rocks. But the distinction doesn't quite hold, nor is it meant to."



Writer of the Paris Review article, Annie Julia Wyman, says it best:

"In an age of information and commodification, we don't need more of either in order to best understand our existence. We need tools for determining significance, places where we can stop and think through the shared story of our world and our lives."

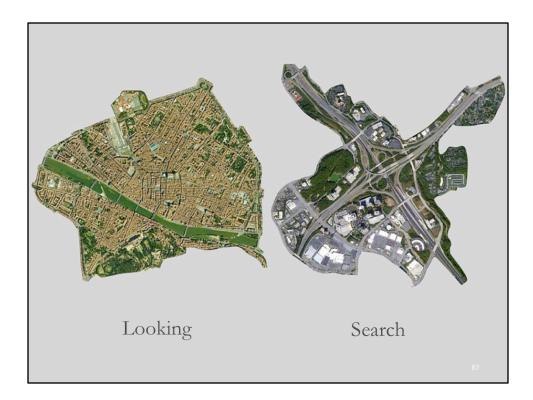


Let's consider the difference between a firehose and smaller, more careful offerings. I like the way artists think about thumbnails.

"The camera records everything without discrimination. In a thumbnail you must decide what is most important and eliminate the rest."

http://www.buildart.com/secret of thumbnails.htm

I'm curious to think through how a thumbnail of data or metadata could be more like this. Instead of making people download a massive database, or giving just a random sample of a set of records, or everything that starts with G, how might we start to represent large data spaces with discrimination?



This utter dependence on the search mentality has given us Atlanta – sorry, Atlanta – instead of Florence. We need to start helping people to look, instead of forcing them to search.

The internet eats museums Working together Data enrichment Measuring impact

So, I've talked about reframing these initial talking points from Johan.

Empty data cities
Work with your content
Think about new forms
Demonstrate use

89

Merci beaucoup!

George Oates
GOOD, FORM & SPECTACLE

gfns.uk | @goodformand

90