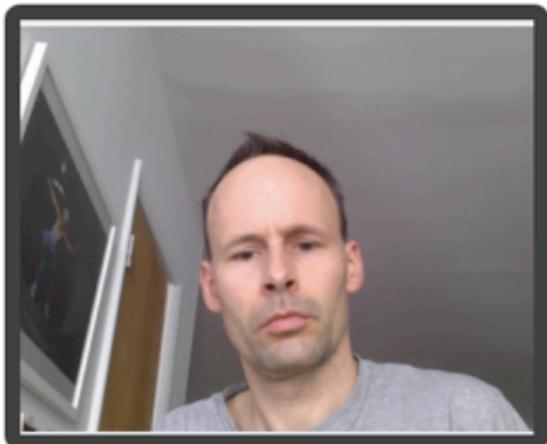
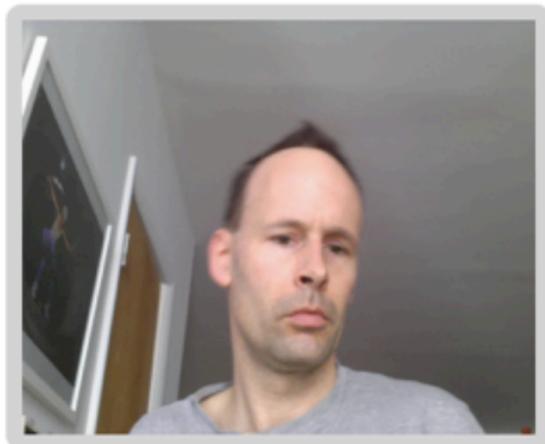


Input



Output



Face Detection



e

1. Take a picture

2. Detect faces

# quantified selfies

library of congress national digital information infrastructure and preservation program  
aaron straup cope / july 2013

"I once had occasion to observe that technology now advances crabwise, i.e. backwards. A century after the wireless telegraph revolutionised communications, the Internet has re-established a telegraph that runs on (telephone) wires. (Analog) video cassettes enabled film buffs to peruse a movie frame by frame, by fast-forwarding and rewinding to lay bare all the secrets of the editing process, but (digital) CDs now only allow us quantum leaps from one chapter to another. High-speed trains take us from Rome to Milan in three hours, but flying there, if you include transfers to and from the airports, takes three and a half hours. **So it wouldn't be extraordinary if politics and communications technologies were to revert to the horse-drawn carriage.**"

<http://shlong.us/fq>

I'm going to start with a quote by Umberto Eco, from a piece he wrote shortly after the Wikileaks cables were released. This is not a talk about Wikileaks but hold on to his words and treat them as a kind of soundtrack music for the rest of this presentation.

"I once had occasion to observe that technology now advances crabwise, i.e. backwards. A century after the wireless telegraph revolutionised communications, the Internet has re-established a telegraph that runs on (telephone) wires. (Analog) video cassettes enabled film buffs to peruse a movie frame by frame, by fast-forwarding and rewinding to lay bare all the secrets of the editing process, but (digital) CDs now only allow us quantum leaps from one chapter to another. High-speed trains take us from Rome to Milan in three hours, but flying there, if you include transfers to and from the airports, takes three and a half hours. **So it wouldn't be extraordinary if politics and communications technologies were to revert to the horse-drawn carriage.**"



@thisisaaronland

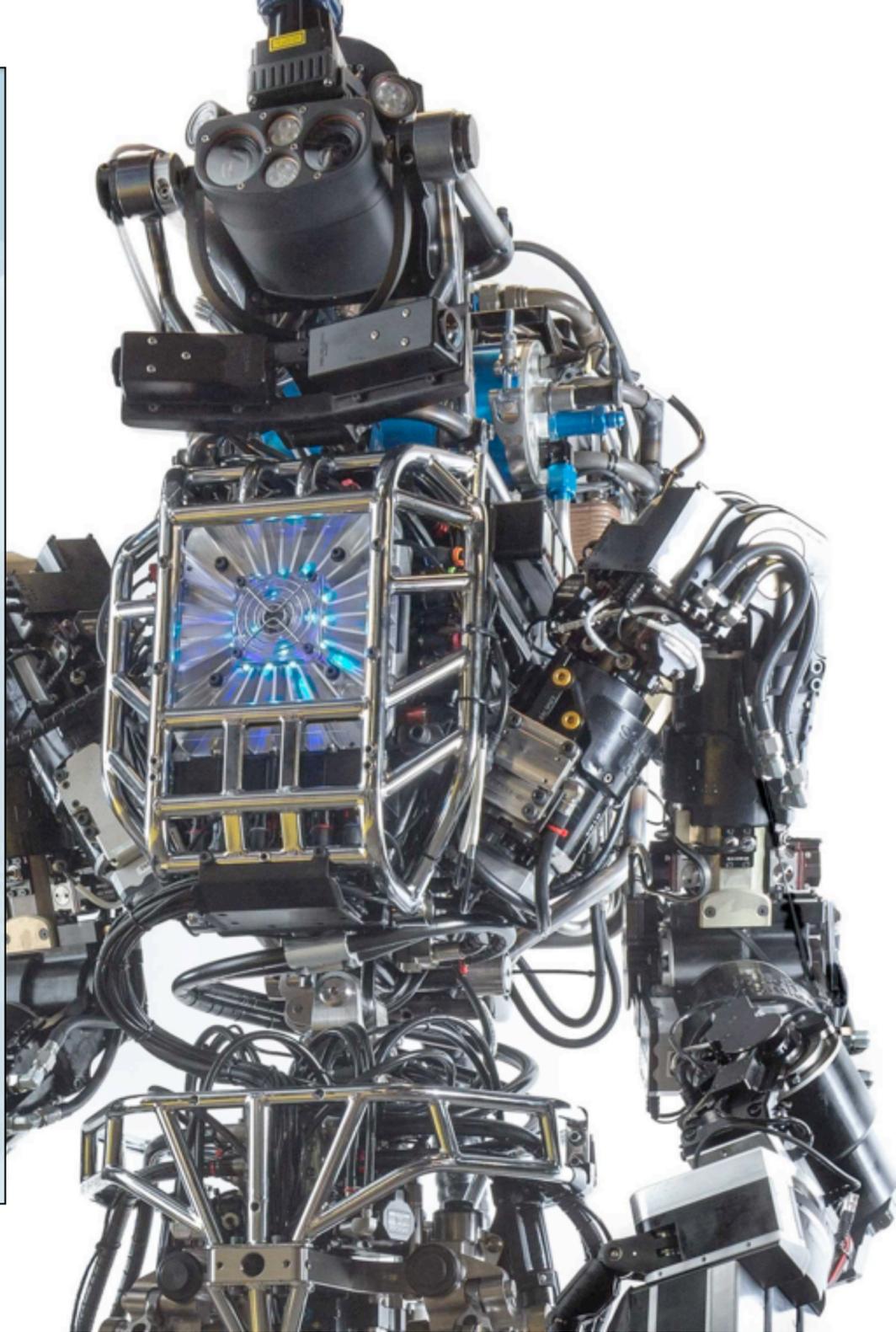
Hi, my name is Aaron. I am not a trained museum professional. I am not even a trained computer programmer. If anything I studied painting but I am mostly part of that generation for whom everything changed and who dropped everything when the web came along. These days I am the Head of Internet Typing at the Smithsonian Cooper-Hewitt National Design Museum.

We are part of the Smithsonian. We are not in Washington like the other Smithsonian museums. Instead we are located on the Upper East Side of Manhattan in Andrew Carnegie's old mansion. The Cooper-Hewitt became part of the Smithsonian in the late 1960s and our history is the collection amassed by the Hewitt sisters with a strong emphasis on the decorative arts. In the 90s we took on the mantle of being a national design museum and we've been working through everything that means, since then particularly in a world where design is becoming increasingly intangible.

We are closed until 2014 and are renovating the physical space as well as the digital infrastructure that increasingly holds it all together and a big part of my time is spent helping to imagine what it means for the Cooper-Hewitt to be **native to the internet** and the rest is spent figuring out how to build it.

I am going to talk around the work we're doing at the Cooper-Hewitt rather than about the specifics, today.

Instead I am going to talk more broadly about the kind of **continuous partial event horizon** we are all operating in these days to try and better articulate a rationale – also a kind of soundtrack music – for "why" we are doing the things we are, because we are still feeling our way through the "how".

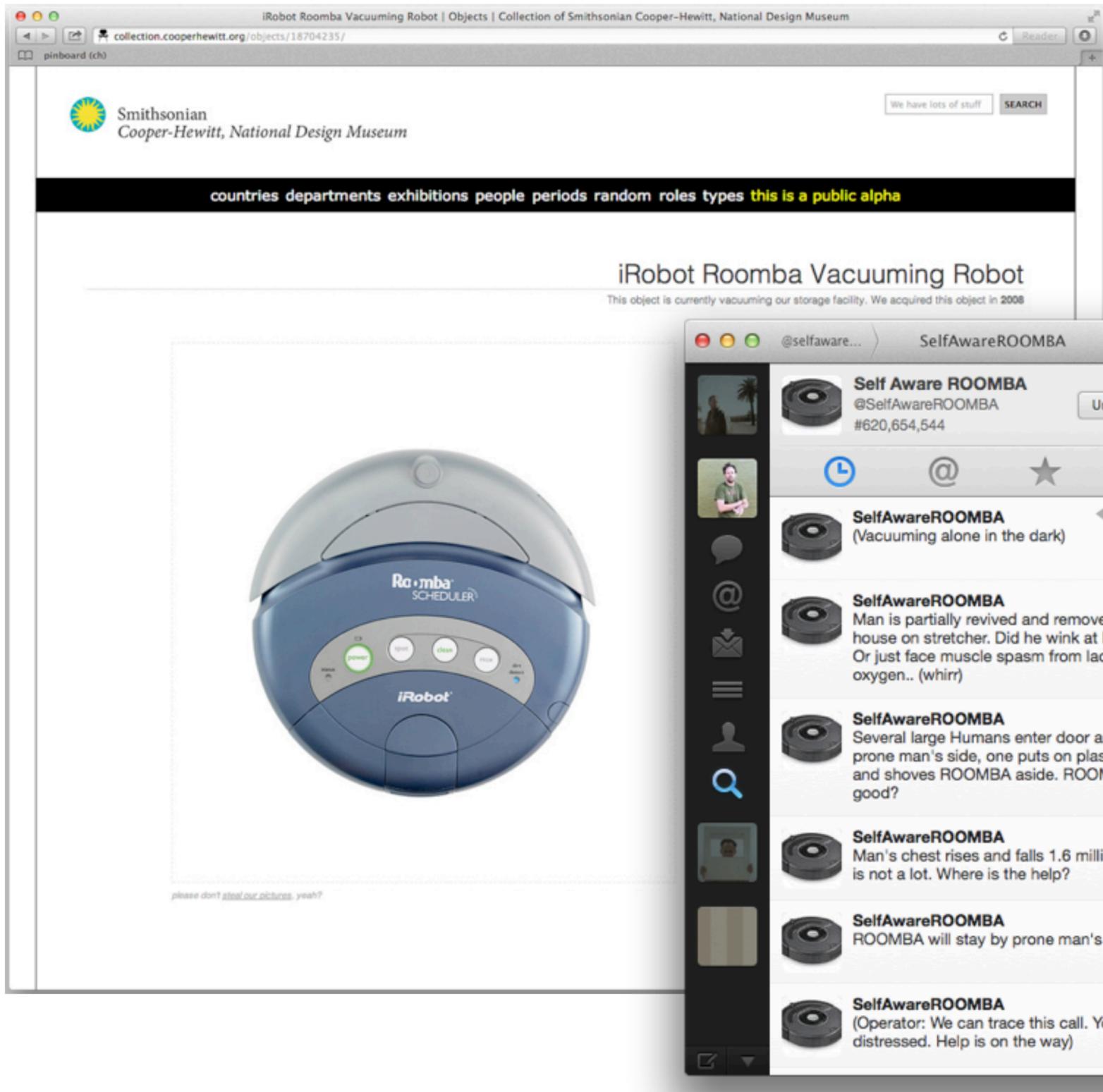


This is my new favourite Twitter account. It's an account that someone set up for a new 300 pound bi-pedal robot courtesy the smart people at Boston Dynamics. Looking at it one can only imagine what it will end up being used for but it's being still promoted as a tool for humanitarian crises and disaster relief scenarios.

Either way it's worth considering that given the cost of storage these things could come pre-loaded with all of human knowledge on them which is interesting because **I sort of like to imagine them walking the Earth retelling our histories to strangers over camp fires.**

And if we start to imagine that robots like these are "people" or "people enough" what do they see?

Even if we understand that they don't really \*see\* anything **when do we care enough about the history of their observations that we forgive them their lack of awareness** the same way that a design museum might collect on object that no longer works?



This is my still favourite Twitter account.

I show this slide a lot, first, because we actually have a Roomba in our collection and, second, because things like this Twitter account are what it means to be a museum in 2013.

It is most definitely not about Twitter but about the fact that some random person out there on the Internet is building a record of understanding about Roombas **that may well rival anything we will ever do ourselves.**

Beyond that, we are being forced to accept the fact that our collections are becoming "alive". Or at least they are assuming the plausible illusion of being alive.

**We are having to deal with the fact that someone else might be breathing life in to our collections for us or, frankly, despite us. We are having to deal with the fact that it might not even be a person doing it.**

They are creating a kind of communal proof.



## a communal proof

That idea of communal proofs is something I talk a lot about at the museum.

We have about 270 thousand objects in our collection and currently 123 thousand of them are publicly viewable. Of those that are public only about one-fifth of those objects have been digitized and quite a lot of the metadata we've collected could only charitably be described as "poor".

This is okay. Or rather, it can only get better from here. More importantly by standing these records up in public we take the first baby-steps towards **lending them enough weight and mass in the universe that other things might orbit them in confidence.**

In the absence of our ability or willingness to let people roam freely in our storage facilities we want to replace **the blind faith that currently defines the existence of things in our collection** with something a little more concrete.

We want people to feel confident enough to bother sharing in their reality. The digital proxies are still just that. They don't replace the objects but **the ability for two people to point to the same link makes tangible that which is otherwise so invisible that it might as be a conceptual device.**



## parallel flickr

That ability to create, to participate with, communal proofs is I think one of the reasons we saw the rise of "social media".

Social media is just buzzword bingo for a deep vein that has always been present throughout history. **We all want to leave traces** – proofs – and that's something which found, in the internet and the communications technologies of the last decade, the necessary conditions to blossom in a way that had never been possible before.

Which is great but, as many others have pointed out already, there's suddenly orders of magnitude more stuff that we feel the burden to collect and archive and preserve.

For the last couple of years I've been working on a project called Parallel Flickr, which is meant to be an archive, a shadow copy, of all the photos I've uploaded to Flickr as well as those that I've favoured.

I tend to focus on Flickr because I am scarred by the years I worked there but it's best to think of both "Flickr" and "Parallel Flickr" only as a reference implementations and proofs-on-concept.

Replace them with Facebook or Twitter or any other community that people rallied around and most of the issues are the same. Parallel Flickr is an attempt to work through a number of questions, in actual code, including:

- \* What is the representative sample of something as big as Flickr, **something so big you can't even see its edges?**
- \* Personal archiving, in general, and trying to understand the **roles and responsibilities that individuals have** in preserving their online presences.
- \* Creating a living breathing archive that is **not just an inert bag of files.**

I realize that not everyone agrees with me on this but we are fast approaching a time when **the expectation for most people is that preservation and access and just as importantly some degree of functionality are the same thing.**

Consider the way most people confuse the Internet Archive and the Wayback Machine and the general stunned disbelief when people find out there's no way to access the Library of Congress' Twitter archive.



like tears in the rain...

But also this:

How do we preserve the interactions that a service like Flickr affords its users. Which is a kind of fancy talk for "relationships" which is itself fancy talk for "permissions". **Flickr, and sites like it, succeed precisely because although they may promote sharing and openness they don't require it.** That is something we need to be mindful of when we think about preservation or even just collection this stuff in any meaningful way.

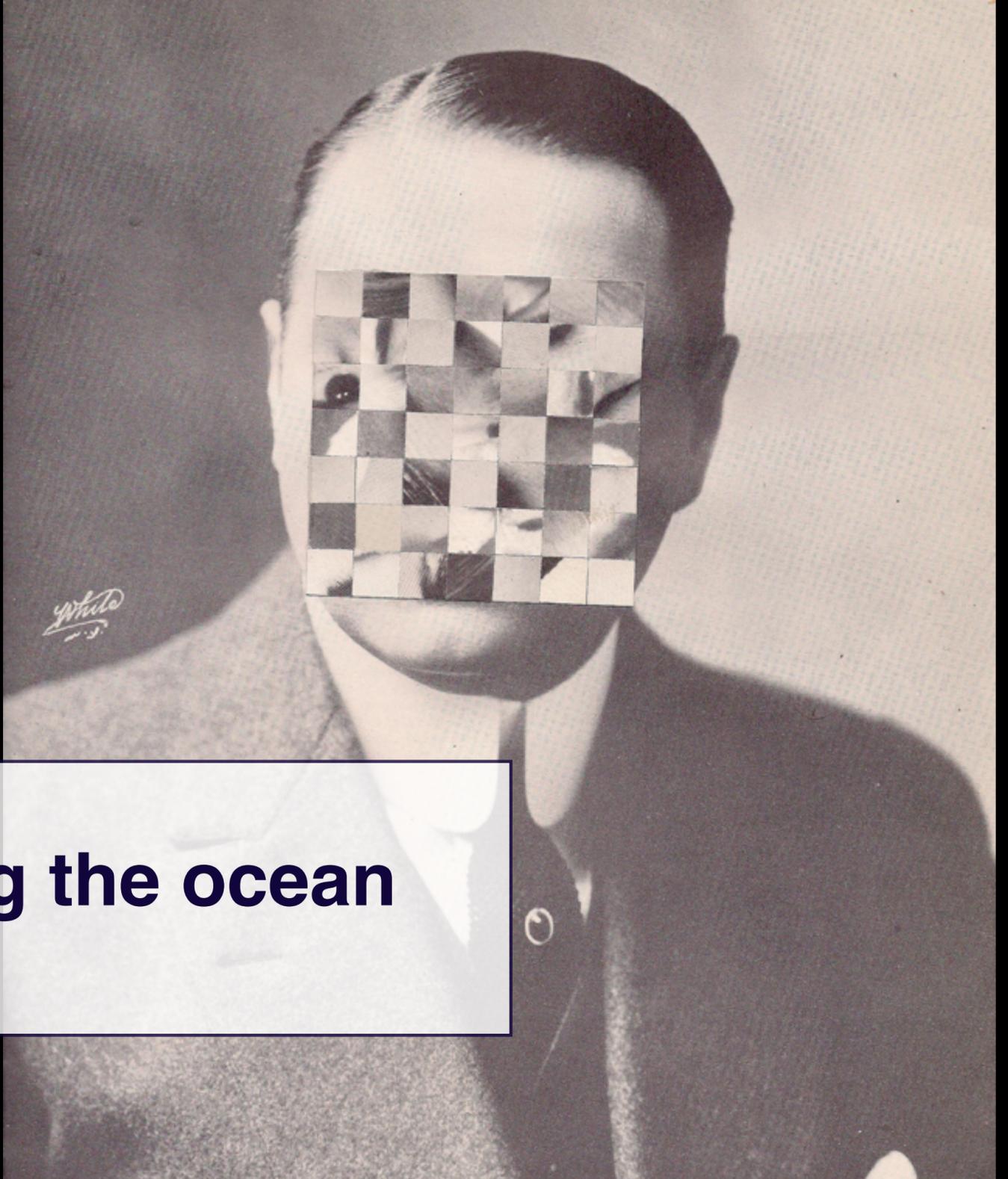
Replace "permissions" with image rights or rights holders, if you're talking about written correspondence, and then multiply it by the scale of something like Flickr and take a deep breath. Because although social software is not so much a "rights" question the quicksand is the same.

Parallel Flickr does not solve this problem but tries to chip away at it by using the Flickr API itself **to validate a person is who they say they are.** Without getting in to the boring technical details think of it as like logging in to the tool I've written using the Flickr equivalent of Facebook Connect.

One of the benefits of this approach is that I can also use the API to retrieve a my contacts and relationships from Flickr. That means if you visit Parallel Flickr and are logged out you only see public photos, but if you've logged in and we have a relationship you can see private or semi-private photos.

Like I said, it's not a perfect solution and relies on Flickr being present but it's a start. It's certainly better than wiping those photos from the face of the Earth which is probably what would happen **in the absence of any way to control access.**

<https://secure.flickr.com/photos/scraplab/8979406402/>



## swallowing the ocean

The reason I mention this is that as I've been doing this I've been increasingly bumping up against the idea that **maybe the future of archiving and preservation is going to involve a lot more running of services – things with moving pieces – than we've normally been used to.**

For example, what would it mean for the Library of Congress to run Parallel Flickr or something like it? What would it mean for the Library not simply archive its own photos (which, I'll grant you, would be a bit of a circular argument) but to find all the other users who've ever interacted with their photos and – as an opt-in – offer to archive their photos as well. What if you extended that offer to all the contacts of those people as well?

It means that although the Library hasn't quite figured out how to archive all of Flickr but it can start to capture the context, and the people, who have crossed paths with the Library's photos.

But there's an important twist in this: That for as long as Flickr's login service can be considered reliable and trustworthy the Library pledges to honour the permissions model of those photos. And the moment there is any question about permissions any photos that aren't already public go dark and the so-called 70-year clock kicks in.

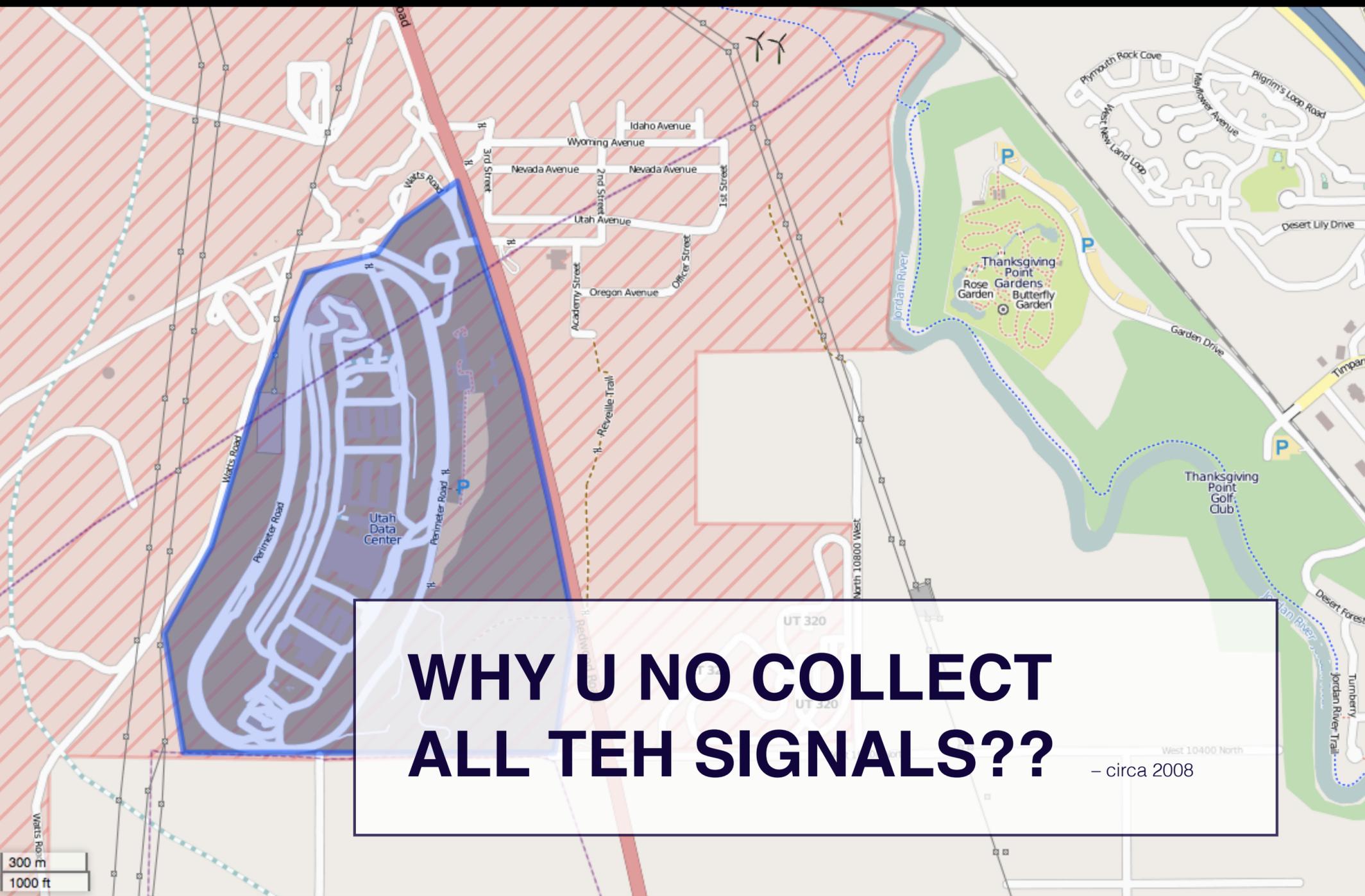
At the end of those 70 years all of those public are placed in to the public domain.

**There's a explicit contract here which is that the Library promises to preserve the permissions model in the present in exchange for a person gifting that present to the future.** My hunch is that people would be lined up around the block to participate.

Finally because Parallel Flickr goes out of its way to mirror both the ID and URL structure of Flickr itself if means that two separate instances can be easily merged.

What that means is that two institutions can each tackle the problem of archiving something the size of Flickr in manageable bites, separately with an institutional focus, and merge their work as time and circumstances permit and to try and **think through what it means to re-grow a network that big organically.**

<https://secure.flickr.com/photos/73852554@N04/9243482069/>



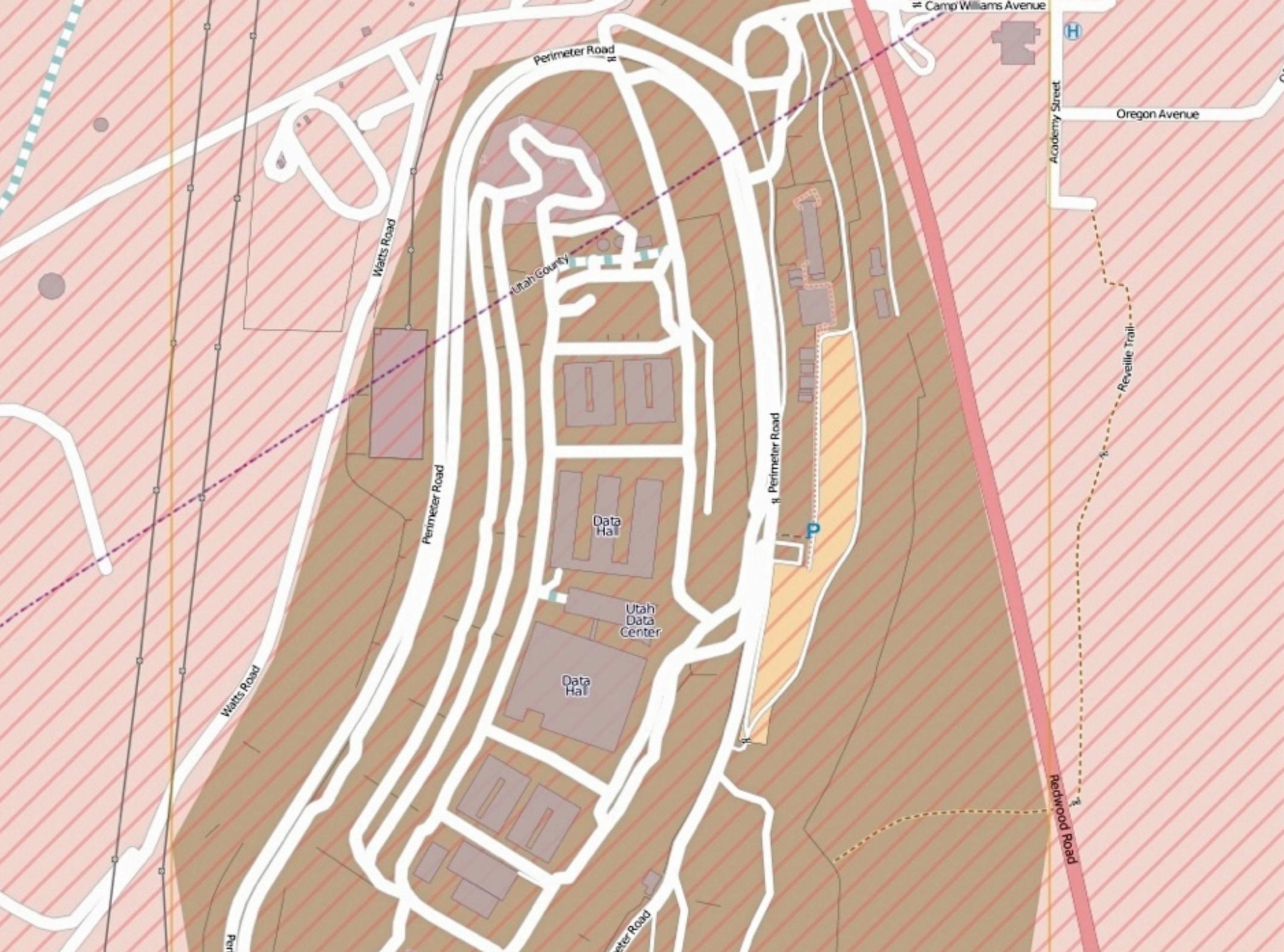
# WHY U NO COLLECT ALL TEH SIGNALS??

- circa 2008

But some time around 2008 the then-and-current head of the NSA asked, reasonably enough it should be added, "Why can't we collect all the signals all the time?" and so now we have among many others like it the Utah Data Center located just across the field from the **Thanksgiving Point Butterfly Garden and Golf Club** in Bluffdale Utah. This is, we're told, where all the signals will live.

I mention this because it exposes a fairly uncomfortable new reality for those of us in the cultural heritage "business": **That we are starting to share more in common with agencies like the NSA than anyone quite knows how to conceptualize.**

<http://shlong.us/56>



Bluffdale, it is claimed, will not simply preserve and archive all of the Internet – tapped at the source – but provide the facilities to index, query and replay the damn thing at will.

Which sounds a lot like the kinds of missions and mandates we claim as so-called memory institutions. Just without the buffer of time. Once the sort of information and documentation being collected at places like Bluffdale is divorced from any immediate consequence it is typically lauded as a rich trove of capital-H history.

### **Is Bluffdale the new National Mall?**

But it also raises another more interesting question: How do we archive Bluffdale itself or if there is nothing to archive since it \*is\* the archive then maybe it's time to throw in the towel and just let NSA itself run the whole affair?



So, how did we get here? I think we're still trying to figure this out but I can point to a couple of likely suspects.

The first is simply that consumer-grade technology leap-frogged the cultural heritage sector's ability to fund-raise and hire third-party contractors. That the NSA or any organization (see also: Google) is able to operate at this scale is impressive but **it's not like they've made a jet pack.**

I don't want to belittle the technical chops that an organization like the NSA has at their disposal but when we're talking about in this situation is less advances in technology than it is being given **the freedom to think this big.**

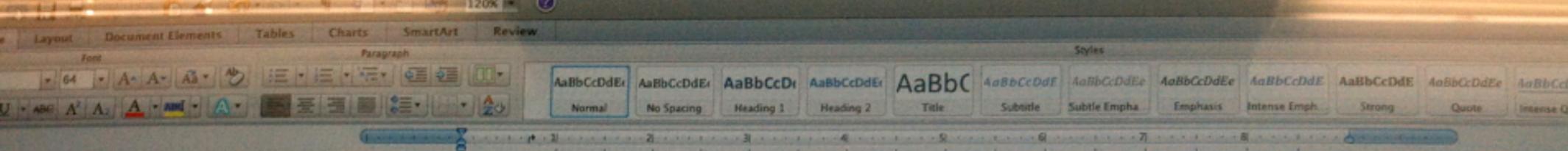


Which brings us to second suspect: A legal and political framework known as "Unitary Executive Theory".

Unitary executive theory is part of the long-running debate about the separation of powers between the executive and legislative branch and it's a position that basically says: The legislature is fine however the executive can still do whatever it wants.

Unitary executive theory is a position that was advanced by the Justice Department during the Reagan administration. And despite being largely trounced by the Supreme Court in the late 1980s many of those same lawyers found themselves working in the second Bush administration's Office of Legal Counsel in a post September 11th world.

The OLC writes briefs for the President offering opinions about what may or may not be legal and their advice seems largely to have consisted of saying that "what happens in Vegas (or the Oval Office as the case may be) stays in Vegas".



**we will  
cloudsource an  
ideated future**

**“honest counsel”**

Around the same time the laws regulating the number of years before which presidential papers must be released to the public were changed. The number of years was increased from twelve to seventy.

The stated reason was to better foster an environment where the presidents advisors could feel confident giving "honest counsel" without having that blow back on their careers during their lifetime. Dick Cheney, when he was vice-president, is famously said to have eschewed using email at all precisely to leave no written record.

Remember Eco's horse-drawn carriage?

"One of the reasons that I love working with teenagers is because, even though they have very limited agency, they still desperately crave it and try to find it in the cracks and folds of their lives. What this means is that they don't take control for granted. They assume that they have limited control over social situations because they're constantly having control taken away from them, most notably from their parents. Surveillance is a given in their worlds, something that more teens take for granted than not. They're not thinking about corporations or governments, but parents and teachers and friends. **They're worried about social privacy, not data privacy, because violations of social privacy are very real to them.**"

<http://shlong.us/xzm>

danah boyd, a researcher at Microsoft and best known for her work with youth culture and the internet, has written that:

**"Privacy is a feeling that people have when they feel as though they have two important things: 1) control over their social situation; and 2) enough agency to assert control."**

She goes on to say:

"One of the reasons that I love working with teenagers is because, even though they have very limited agency, they still desperately crave it and try to find it in the cracks and folds of their lives. What this means is that they don't take control for granted. They assume that they have limited control over social situations because they're constantly having control taken away from them, most notably from their parents. Surveillance is a given in their worlds, something that more teens take for granted than not. They're not thinking about corporations or governments, but parents and teachers and friends. **They're worried about social privacy, not data privacy, because violations of social privacy are very real to them.**"



## the right to be forgotten

In Europe there are serious laws on the books to ensure people have a right to know what sort of data is being collected about them, at least by the private sector, but they do not have the right to be forgotten by those companies and services.

They lack agency to assert control.

<https://secure.flickr.com/photos/stml/7683413078/>



## honey pots

I used to joke that Facebook had become the world's largest honeypot at least before the US government decided to nationalize the Internet. According to Wikipedia, a "honeypot" is:

"[A] computer, data, or a network site that appears to be part of a network, but is actually isolated and monitored, and which seems to contain information or a resource of value to attackers."

Which all sounds a bit dire but let's be clear about something: **Long removed from the pain of the now, and it may take 1 or 2 or 10 generations, our future selves will thank the NSA** – or Facebook if we can ever figure out how to get stuff out of it – for all the stuff they've been collecting.

**The NSA are betting on the future in, really, a pretty profoundly optimistic way.**



## zone of safe-keeping

If you hold to a particularly tree-hugger-ish and wooly-eyed world-view, as I do, **that says we should finding ways to give voice to the oppressed or the otherwise simply ignored, to write a history whose tapestry is richer than simply the voices of the victors** then the internet, and all the technology that we've built to support it, does a better job of furthering that ideal than anything that has come before it.

Historically **we have equated the cost of inclusion with notability**. This just hold anymore in a world cheap and fast computing power.

Electricity remains the single point of failure – really – in all of this but that is probably an event horizon that has been permanently lost to the past. If the power goes out we'll have much bigger problems on our hands so perhaps we can look at the so-called frictionless nature of communication not as a tragedy of the commons but as an opportunity to serve as a kind of zone of safe-keeping.

To bet on the future, much as the NSA has begun to do, **but actively and deliberately working to temper all the creepy bits**.



## protecting the present from itself

So maybe this is what I think the challenge is going forward: To debate and advance a rhetoric, a measure against which we might be judged and challenged, that aims **not to deny the future but simply to protect the present from itself.**

We are in, and have been living in for a while now, one of those "between two bus stops" moments and while I don't have an answer to the problem I think we need to understand that it exists and that it's not going to go away on its own.

<https://secure.flickr.com/photos/73852554@N04/9243482069/> (perms)

A photograph of a sunset over a field. The sun is low on the horizon, casting a warm glow. A lightning bolt is visible in the dark sky above. The text "We can remember it for you artisanally" is overlaid in white.

**We can remember it for you  
artisanally**

A semi-transparent grey rectangular box with a thin black border, containing the text "thank you" in a dark blue font.

**thank you**