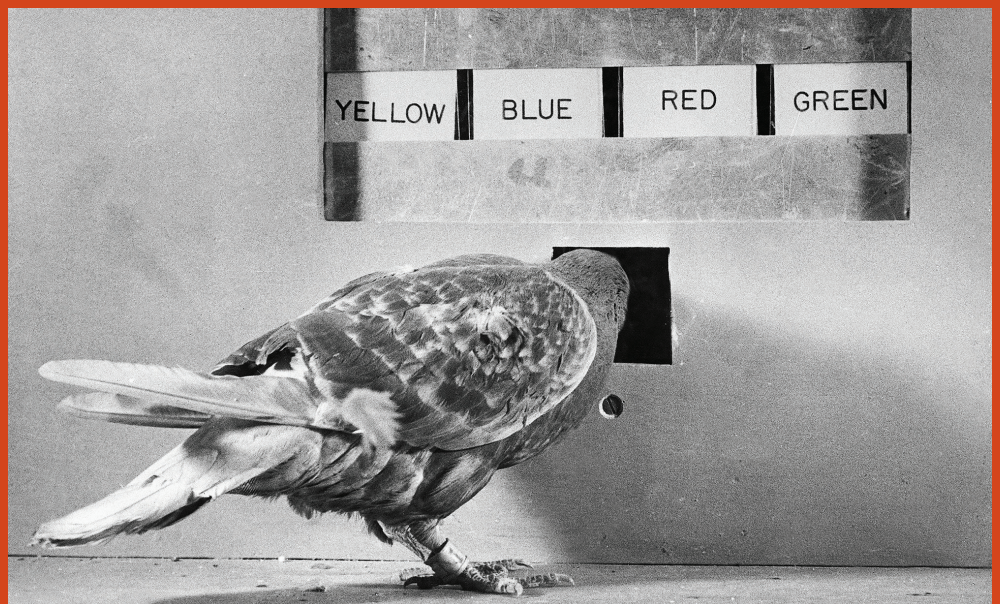


# just what are we archiving?

What kind of  
people will  
we become  
if we keep  
trying to  
archive  
everything?  
**Geof  
Bowker**  
reports from  
inside the  
Skinner Box.



**■ WAS BORN AT THE DAWNING** of the Age of Aquarius; I have graduated to the dawn of the era of Big Data. It seems as if all aspects of our lives are being tracked, monitored, and stored for future use, and by “us” here I include vast swathes of the nonhuman as well as human world. We leave traces everywhere, often without realizing it, and these are potentially stored forever; collectively, they build a picture of ourselves that can be exploited by commercial companies (by way of Google, Facebook), governments, and aspiring political candidates.

I grew up thinking that archives were dusty, dry places that only aspiring historians such as myself could find exciting...and I still treasure the peace of roaming through a nineteenth-century set of police reports on a political group (the National Union of the Working Classes) in the 1830s in England, as well as continue to feel the anguish of what I found there. I have graduated to seeing archives as performative: they constitute the present as much as document the past.

I felt I had entered the world of the Skinner Box when I read the chilling directive by Barack Obama on September 15, 2015: “Executive Order—Using Behavioral Science Insights to Better Serve the American People” (White House 2015). For example, Section 1.b.(iii) called for scientists to: “identify programs that offer choices and carefully consider how the presentation and structure of those choices, including the order, number, and arrangement of options, can most effectively promote public welfare, as appropriate, giving particular consideration to the selection and setting of default options.” At the time, I was reading Natasha Dow Schull’s (2012) *Addiction by Design*, which described exactly the same logic in use by the casinos in Las Vegas to gather as much information as they could about their clients: tracking their movements, taking videos of them as they were robbed by one-armed bandits, predicting when they are about to leave so that they could be given a quick filip to keep them going. And then as I was writing this piece, I learned that Ted Cruz had deployed Cambridge Analytica to harvest data on tens of millions of Facebook users (through the

neat trick of getting Amazon Turkers to give access to their profiles, an access that extended exponentially to all their “friends” in the business) (Davies 2015). What interested me was that the government, the gambling industry, and the hopeful candidate were deploying exactly the same archival practice: find as many traces as you can of a person, track how they have responded in the past to particular messaging, then tailor your messaging so they will act appropriately in the future.

These techniques were not developed for such lofty purposes. They were part of the advertising empires built by such titans of our times as Google and Facebook. The performativity of this kind of archive is the sell: a pair of trousers, a government program, a gambling addiction, or a candidate. There is a certain beauty in the flattening of the spheres: it’s all about the same archival technique. I note the resonance with James Beniger’s (1986) observation in *The Control Revolution*: control is about feedback loops, which is about gleaning, siloing, and consuming information. He notes that his period, the late nineteenth century, saw the birth of mass advertising.

What is interesting about our new archival practices is that they don’t want to just slot me into a category (white, elderly, middle class): they just want *me*, the traces that I have left and am leaving in the present. What is the message that works for me, Geoffrey Bowker? Let my traces tell you.... I’ve sure bought a lot of stuff and arguments lately I never meant to originally.

Helen Verran warned against the “hardening of the categories” (1998) what we are dealing with here is an advanced case of the softening of the categories. Lest I sound too cynical, let me say that this is an ontologically beautiful vision. A field I have found fascinating for years is that of mereology (the science of parts and wholes) because it does not assume the identity of a thing with itself over time. I share few cells and social associations with the Bowker that my mother birthed, and an ontological commitment that does not seek to put me into a finite set of fixed pigeonholes is, I think, a Good Thing. Transience, along many dimensions, is the rule: fixed categories are, to borrow a phrase from

Michel Tournier, the vitriol of the soul. Designing a fluid archival trace is a wonderful thing. The strong vision—shared by the Tardean in Latour and by Chris Anderson in *Wired Magazine*—is that our new archival practices will permit an infinitely rich view of the world that does not need to go through the annoying filter of totalizing theory.

This is a vision based on our archival traces being so complete that we can capture all that is needed to describe and act in the world. And yet the data

don't even know we should be monitoring. In the world of biodiversity policy, almost all of our traces come back to the genome of a particular species. However, we are all obligate symbionts, housing within our bodies a vast array of flora and fauna and in our homes another extensive array. Yet we don't build our archives around relationships, we build them around things (if there is one fundamental flaw in our generic archival practices, it is this).

## **What is interesting about our new archival practices is that they don't want to just slot me into a category ... they just want me, the traces that I have left and am leaving in the present.**

plenum is not yet there: it is theoretically impossible. Where we are left is inhabiting an uneasy zone between a proximate future, always just around the corner—about five years away, when we will have all the data—and a set of archival practices in the present that perpetuate certain kinds of invisibility: things we cannot or will not see.

So what kinds of things are *not* being archived in this world of total data? One kind of trace Google does not have (though it has access to all of my email, courtesy of my university) is quite simply the things I choose not to write. When I leave written traces I—along with major companies—choose to leave a record which (I hope) cannot come back to bite me. A number of years ago, a university dean advised me that an angry email I had sent was okay in sentiment, but it was not the kind of thing one said in email; one said it face to face. In the halcyon early days of the Internet, companies such as Microsoft got into trouble because they left traces of their internal strategy to scupper Apple (we saw how well that worked...). We are active contributors to the archival traces we leave, and while we may be tracked along myriad dimensions, we are not tracked along all of them. The archival traces we choose to leave provide a picture of ourselves as we would have acted had we been model citizens. A second kind of trace not being left are the traces that powerful actors work to keep out by choosing not to measure them. Thus it has taken a wonderful yearlong campaign by the *Guardian* newspaper to get the FBI to agree to put together some kind of centralized database of police killings in the United States. Things that are not being measured cannot come back to bite you: they are left scattered in unconnected databases that, in that ever-proximate future, will be linked. Despite massive data collection efforts after the Chernobyl incident, many of the effects of the disaster were not captured (Kuchinskaya, 2014), though I realize as I write this that I can only assert it, since there are no data: estimates range between a low of about five thousand and a high of mid-six hundreds of thousands. A third kind of trace is the trace that we

The dark vision of the future is that the performative archive we are creating is one destined to make Skinner right. I do indeed respond to many of the stimuli I am being prodded with, and better the data analytics on the archive, the more I respond accordingly. The bright vision (as a child of the Age of Aquarius, after all) begins with the recognition that in the world we are building—a child of the era of governmentality discussed by the early Foucault—our archives are central cultural, social, and economic constructions, and should be analyzed as rich, complex creations as much as stark, useful facts about the world. If our archives are to constitute our being, let's make them interesting. ■

---

**GEOFFREY C. BOWKER** is Professor of Informatics at University of California, Irvine and is working on the nature of the Archive.

### **BIBLIOGRAPHY**

- Beniger, J. R. 1986. *The Control Revolution: Technological and Economic Origins of the Information Society*. Cambridge, MA: Harvard University Press.
- Davies, Harry. 2015. "Ted Cruz using firm that harvested data on millions of unwitting Facebook users." *The Guardian*, December 11. Available at <http://www.theguardian.com/us-news/2015/dec/11/senator-ted-cruz-president-campaign-facebook-user-data>.
- Dow Schull, N. 2012. *Addiction by Design*. Princeton, NJ: Princeton University Press.
- Verran, H. 1998. 'Re-Imagining Land Ownership in Australia'. *Postcolonial Studies*, Vol 1, No 2, pp 237-254.
- White House. 2015. Executive Order—Using Behavioral Science Insights to Better Serve the American People. September 15. Available at <https://www.whitehouse.gov/the-press-office/2015/09/15/executive-order-using-behavioral-science-insights-better-serve-american>.