

THE PROMISE [SIC] OF BIG DATA

22 pages into the White House report on Big Data, this paragraph appears:

Government keeps the peace. It makes sure our food is safe to eat. It keeps our air and water clean. The laws and regulations it promulgates order economic and political life. Big data technology stands to improve nearly all the services the public sector delivers.

It presents several claims that are arguably not at all true:

- Government keeps the peace (where? South Chicago? Iraq? Wall Street?)
- Government makes our food safe to eat (with the few inspectors who inspect factory farms? with federal guidelines that don't combat obesity?)
- Government keeps our air and water clean (I'm more comfortable with this claim, until you consider we're melting the planet with stuff in the air that government doesn't want to regulate)
- Government laws order economic and political life (they may well, but is that order just and good?)

And that, the report says, is all made possible

because of BigData.

Some 15 pages later, after it has reviewed the top secret DHS database analyzing all our public called Cerberus, has admitted the government needs to rethink the meaning of metadata across both intelligence and non-intelligence functions, and explained the new continuous evaluation systems to root out insider threats, the report again proclaims Big Data's good.

When wrestling with the vexing issues big data raises in the public sector, it can be easy to lose sight of the tremendous opportunities these technologies offer to improve public services, grow the economy, and improve the health and safety of our communities. These opportunities are real and must be kept at the center of the conversation about big data.

Meanwhile, the report offers up these other signs of Big Data progress:

- Big data “is also enabling some of the nearly 29 percent of Americans who are ‘unbanked’ or ‘underbanked’ [often because of Big Data] to qualify for a line of credit by using a wider range of non-traditional information—such as rent payments, utilities, mobile-phone subscriptions, insurance, child care, and tuition—to establish creditworthiness.”
- “Home appliances can now tell us when to dim our lights from a thousand miles

away.”

- “Powerful algorithms can unlock value in the vast troves of information available to businesses, and can help empower consumers.”
- “The advertising-supported Internet creates enormous value for consumers by providing access to useful services, news, and entertainment at no financial cost.”

In short, the whole thing is rather breathless about Big Data.

And in spite of the fact that respondents to a totally unscientific (not Big Data) survey said they were most concerned about intelligence (first) and law enforcement (second), the Big Data report avoided much of the discussion about this, relegating it to discussions of local law enforcement’s use of predictive analysis.

And where they do describe surveillance, it’s either to boast about how good the security is on their database, as they do for DHS’ curiously named “Cerberus” database, or to pretend big data doesn’t dominate there, too.

Today, most law enforcement uses of metadata are still rooted in the “small data” world, such as identifying phone numbers called by a criminal suspect. In the future, metadata that is part of the “big data” world will be increasingly relevant to investigations, raising the question of what protections it should be granted. While today, the content of communications, whether written or verbal, generally receives a high level of legal protection, the level of protection afforded to metadata is less

so.

Although the use of big data technologies by the government raises profound issues of how government power should be regulated, big data technologies also hold within them solutions that can enhance accountability, privacy, and the rights of citizens. These include sophisticated methods of tagging data by the authorities under which it was collected or generated; purpose- and user-based access restrictions on this data; tracking which users access what data for what purpose; and algorithms that alert supervisors to possible abuses.

And there are a slew of places in the report – where it talks about HIPAA without talking about using Section 215s to get HIPAA data, where it talks about FCRA without talking about NSLs to get financial data, where it neglects to mention NCTC’s ability to get federal databases, including those of DHS – where it remains silent about the surveillance piggybacking on the issue at hand.

Perhaps the most frustrating part of the report – aside from the fact that it actually had to advance the recommendation that we only use Big Data collected in schools for educational purposes (setting aside how well or poorly Big Data is serving our students) – is the silence about the things we don’t use Big Data for enough, notably solving the financial crisis and regulating banksters (including things like tax havens, inequality, and shadow banking), or really doing something about climate change.

Big Data, as it appears in the report (as presented by a bunch of boosters) is not something we’re going to throw at our most intractable problems. We’re just going to use it to turn the lights off on the other side of the country.

And to spy.