

Terms of Service

understanding our role in
the world of Big Data

by

Michael Keller and Josh Neufeld

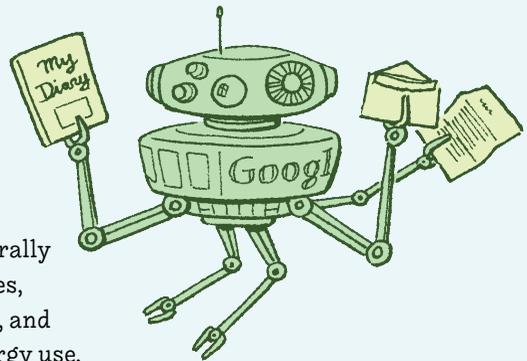


Terms of Service: *Understanding Our Role in the World of Big Data*

By
MICHAEL KELLER and **JOSH NEUFELD**

Edited by
RHYNE PIGGOTT and **LAM THUY VO**

Big Data powers the modern world. With tools like FitBit tracking our every step and supercomputers like IBM's Watson helping Memorial Sloan Kettering treat cancer patients, we literally live it. Between our social media profiles, browsing histories, discount programs, and new tools like Nest controlling our energy use, there's no escape.



It's true humans have always generated, collected and analyzed data, from knowing which time of year to fish the rivers to realizing they should probably stop eating so much takeout. What's different today is the power of everyday devices to gather and analyze that data in real time. Unless one pays with cash, now it's possible, with very little effort, to go from saying "I eat a lot of takeout" to knowing with certainty "I spend 23 percent of my weekly paycheck on takeout food."

Big companies are collecting and using this information, too. This can be good or bad. Maybe we start making better choices about the food we eat and the money we spend. Or maybe an insurance company decides to increase its rate or even terminate a policy because it projects we'll have diabetes in 18 months.

We believe many folks want to learn more about these issues but are turned off by often dense and jargon-laden coverage.

So we made a comic!

What do we gain from Big Data? What do we lose? Co-produced by cartoonist **Josh Neufeld** and reporter **Michael Keller**, we hope this graphic novella will give you a good foundation to start asking your own questions.

Have any comments or stories you'd like to share? Please get in touch:
terms-of-service@aljazeera.net.

— The Editors, October 2014

Spring 2004, San Francisco.

What I've drawn here is a portrait of fear.

The squiggly line is the penumbra of fear, and the dot in the middle is REAL fear.

You're in the penumbra of fear.

Al Gore

Google executives--and the former vice president--meet with a California state senator at the Ritz-Carlton to talk about... Gmail.



When Gmail debuted that year, privacy advocates were immediately up in arms.

We write to your office to investigate Google's "Gmail" service. We believe that Gmail violates California Penal Code 631, which governs eavesdropping on confidential communications.

Gmail made money by scanning the contents of your email for KEYWORDS that it used to serve targeted ads.

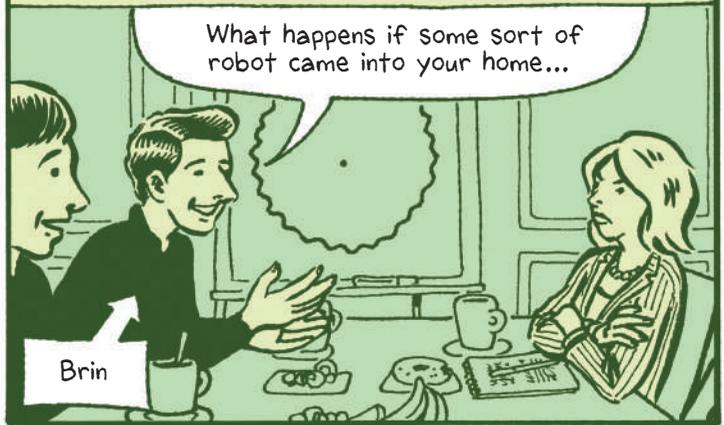
California state Sen. LIZ FIGUEROA was upset that people who weren't using Gmail--but were corresponding with people who were--would also have THEIR emails scanned, effectively being forced into the keyword-scanning system.



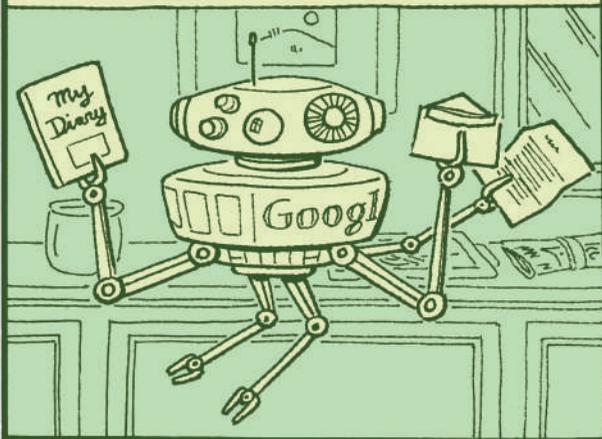
Claiming that Gmail's system ran afoul of the state's Wiretap Act, Figueroa put forward a bill² requiring Google to obtain users' consent before any scanning could happen.

Google, and any other company which seeks to use a similar system, to obtain the sender's consent for outgoing messages, and the consent of the sender and the recipient for incoming messages, prior to scanning the email's contents for advertising purposes.

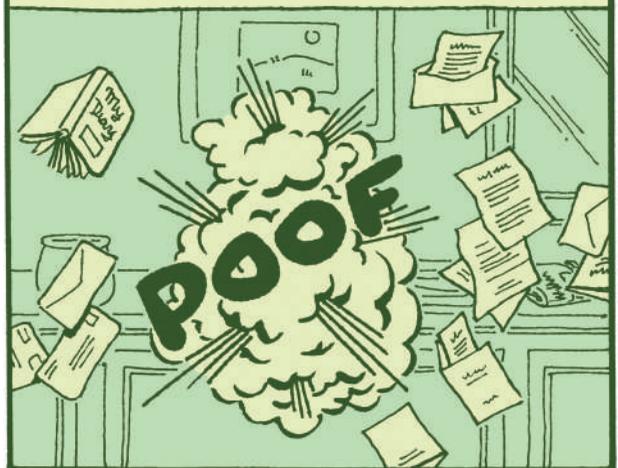
Google co-founders Sergey Brin and Larry Page scoffed at Figueroa's fears that the company was retaining data over time.



"...and started reading your private materials--like your IRS statements, your diary or your private correspondence..."

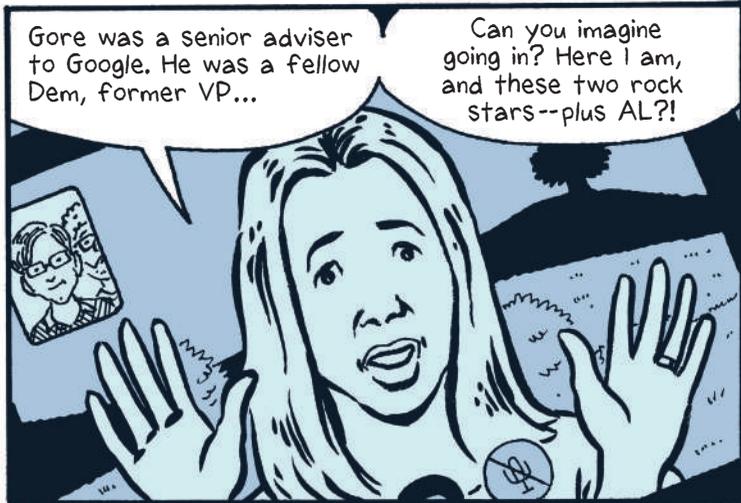
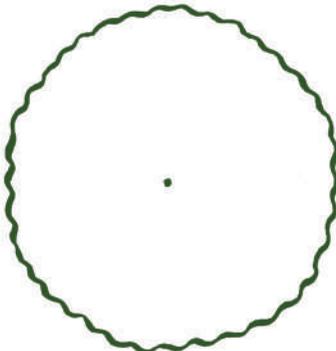


"...but then imploded--and all that went away."



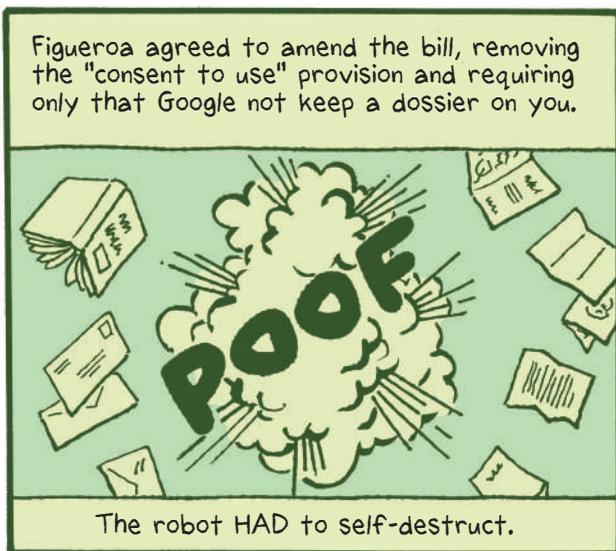
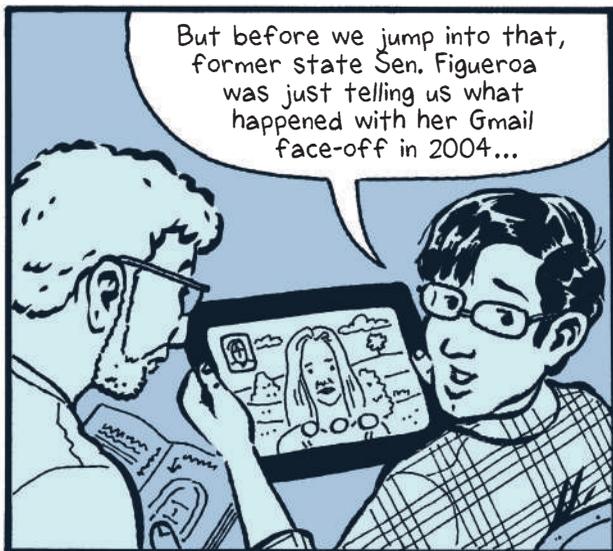


Figuroa was in the "penumbra of fear" because she was afraid of something the Google execs said they WEREN'T doing -- collecting a comprehensive digital dossier.



*Al Gore did not respond to repeated requests for comment. Google declined to comment for this story.





It was a moment when legislators could have changed how the robots operate--to prevent a version of the Internet where advertisements read and track everything you do.



You mean like the Internet is now?

Exactly.



Today, much more sophisticated robots follow us around.

NOT * AS * FAST *
AS * LAST * WEEK



And it feels like people have no choice but to accept it.

And all because of Gmail?



What Gmail offered-- on a mass scale--was that for a benefit, people would give up new kinds of personal information.

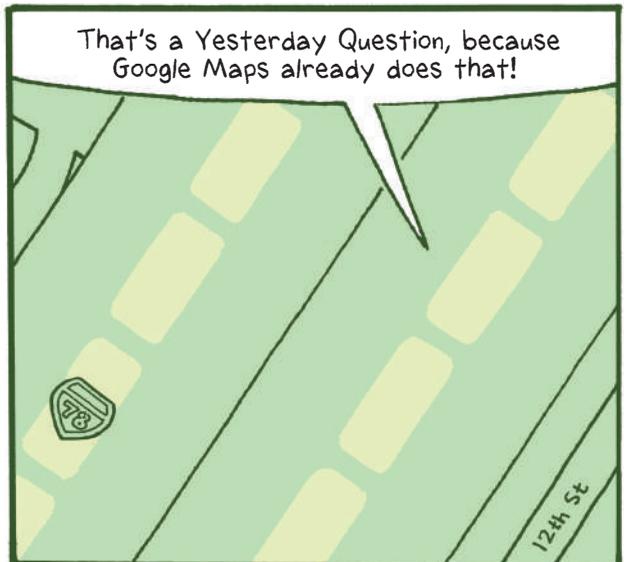
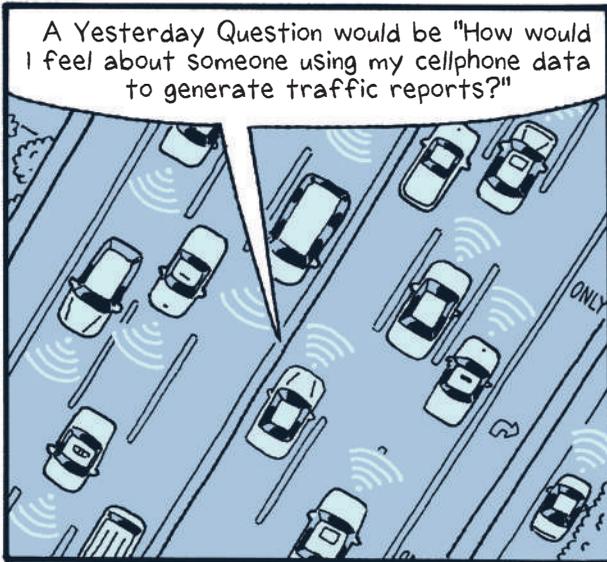


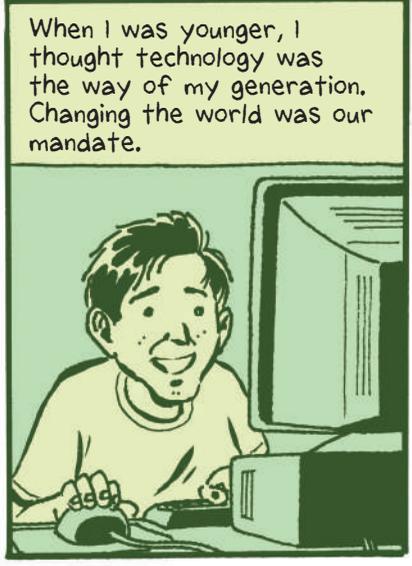
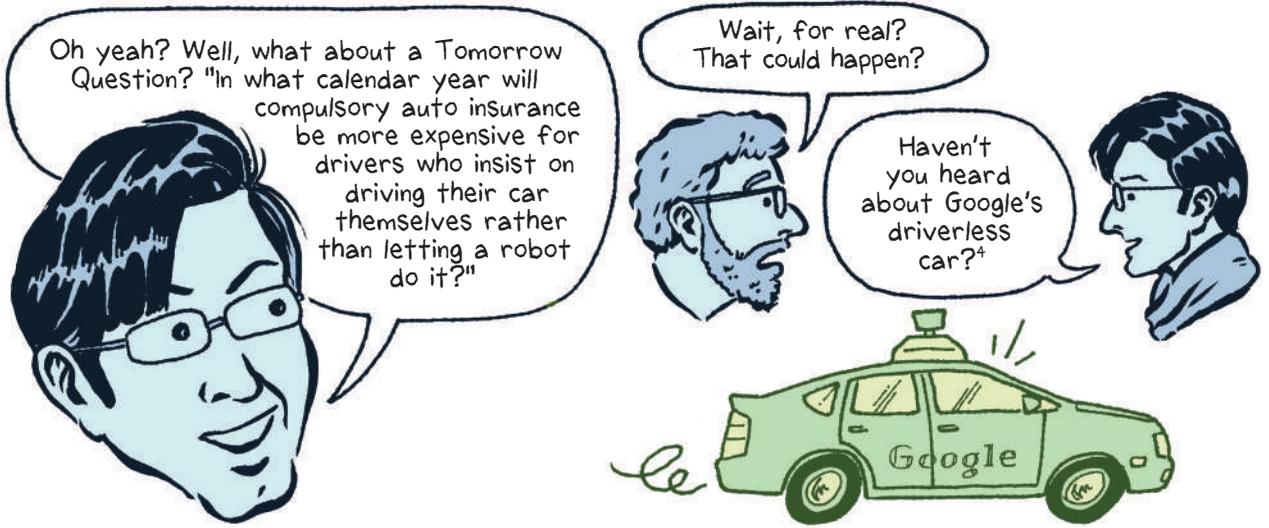
There are a couple of professors in Colorado who have done a lot of thinking on these issues...

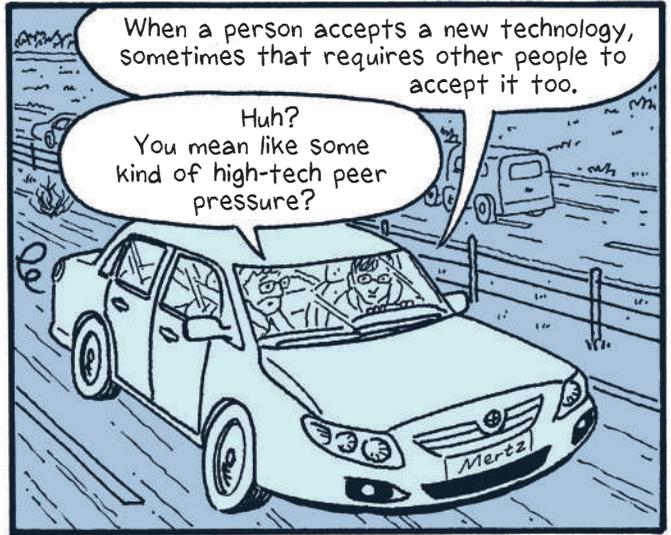
Let's zip over to Boulder to talk to them. You can explain more in the car on the way to the airport.











The UNRAVELING THEORY



Scott Peppet,
University of
Colorado at
Boulder.

The key idea is that once enough people reveal their information, then NOT revealing your information becomes a stigma --even for those with less-than-ideal characteristics.⁶



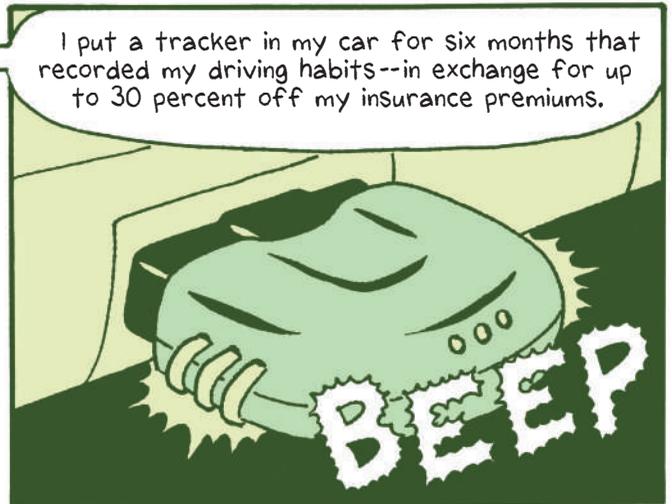
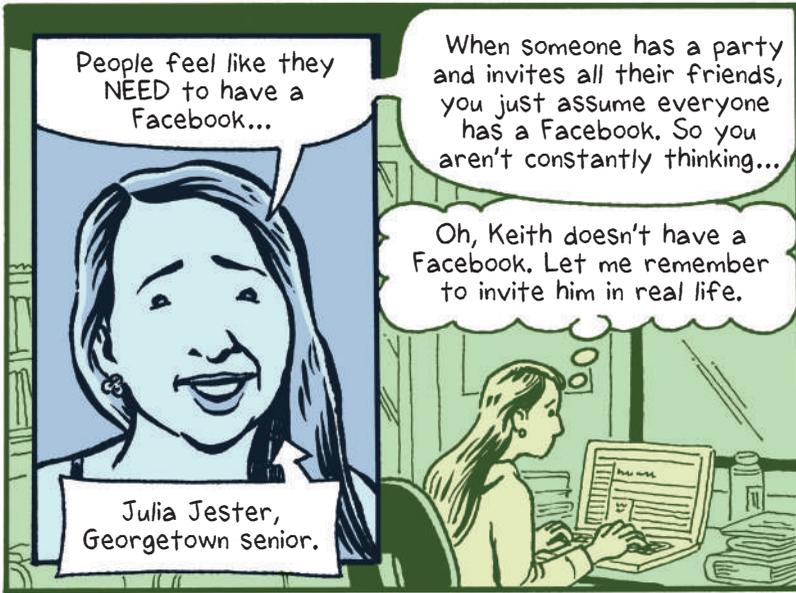
Like with Foursquare. Let's say all my friends use it. At what point do I look like I'm HIDING something if I don't use it?



Hmmm. I wonder if the cafeteria we passed downstairs gives you a deal for checking in...

As a part of our research, we've been speaking to college students to see how unraveling might play out with social media...

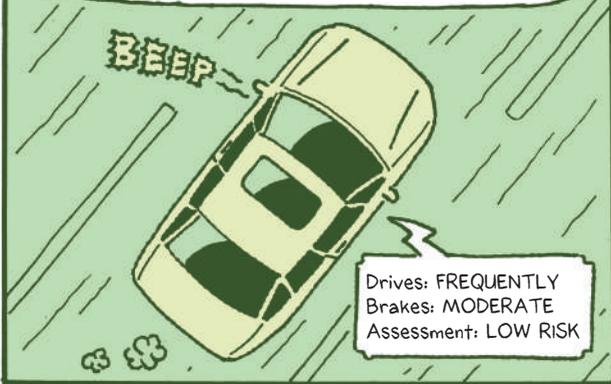






It would take a while in a market as big as car insurance for the unraveling stigma to happen.

That said, it's really a Big Data question also. Progressive has REALLY accurate information about these million drivers who have signed up.



Drives: FREQUENTLY
Brakes: MODERATE
Assessment: LOW RISK

It can [also] learn a lot about the drivers who HAVEN'T signed up.



Maybe it will just be GREAT that we all have trackers in our cars. But I don't think so necessarily.



Insurance has always worked under this idea of shared risk.



Right, people at different risk levels average out one another's costs in pools.



"If everybody's insurance were priced individually, perfectly..."

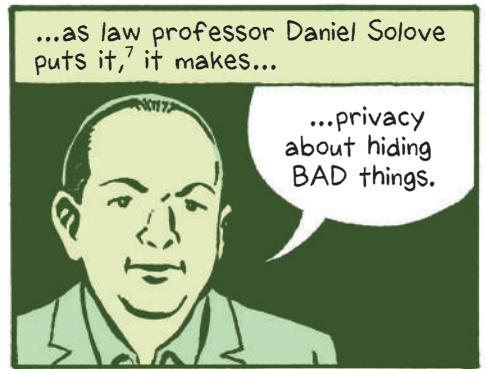
- HIGH-RISK: \$436.11/mo.
- LOW-RISK: \$32.56/mo.
- HIGH-RISK: \$211.45/mo.
- LOW-RISK: \$19.33/mo.

"...you're kind of eviscerating the idea of insurance pools."

And we haven't lived in that world before.

I think that's going to be a very strange world.



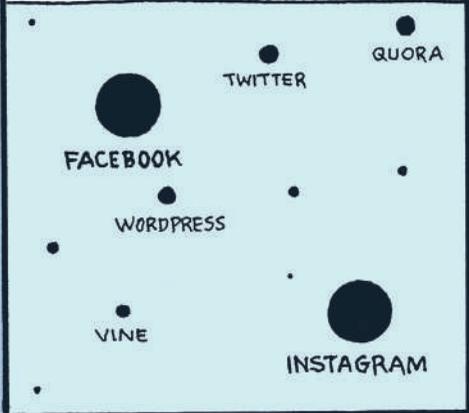


WHO Gets to Tell a STORY About YOU?



Whether it's your car insurance company deciding how good a driver you are...

Or an employer judging how good a fit you'll be based on your social media profile...



I get it! It's like the dots on their own don't really mean anything--

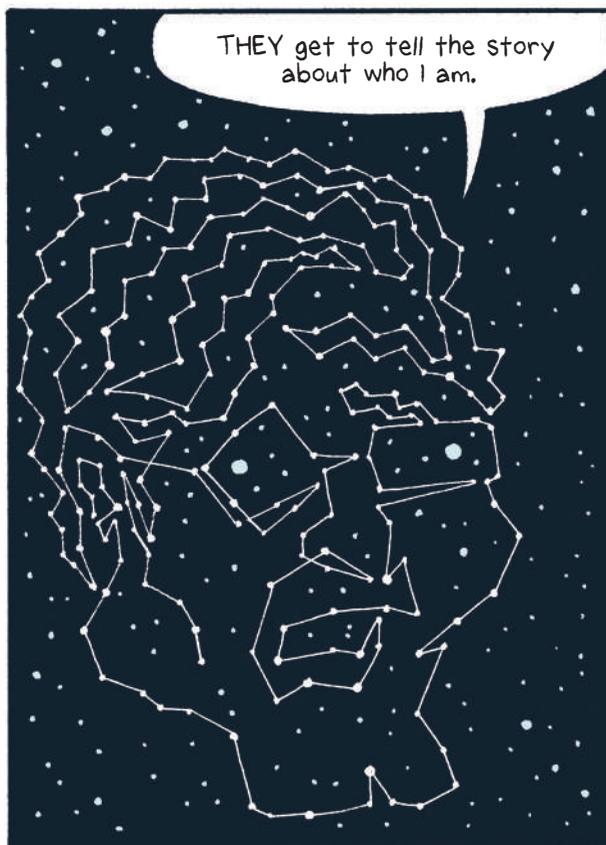


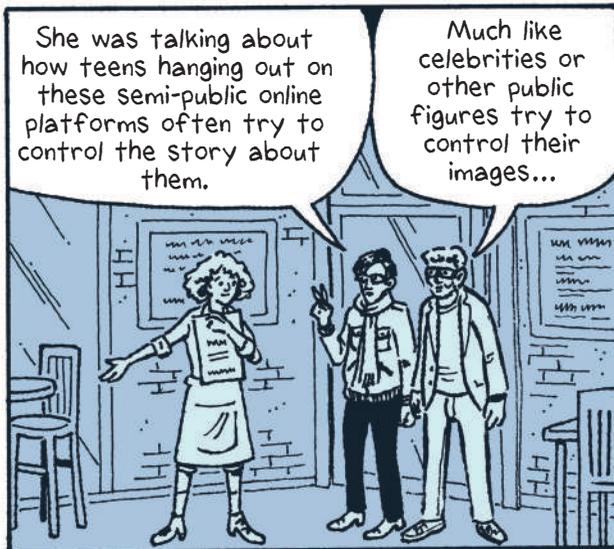
--but then PEOPLE and computers connect the dots to write their own stories.



And with more and more companies collecting more and more dots, who knows HOW they're connecting them?







*danah boyd spells her name in all lower case.

Or I'm going to put up stuff about MYSELF that's really, really cruel --so I can elicit support and feedback.



"When it comes to young people, they're really clear. They'll give up WHATEVER to be able hang out with their friends."



And they don't feel like they have any choice about any of this.



So one way to get that power back is to try to define the story about you in a public way.





--but also when I create an entry for a place, it's like I've created NEW LAND--like an explorer.



That makes total sense. But so does MY story. Because I can make up whatever narrative I want from your data dots!

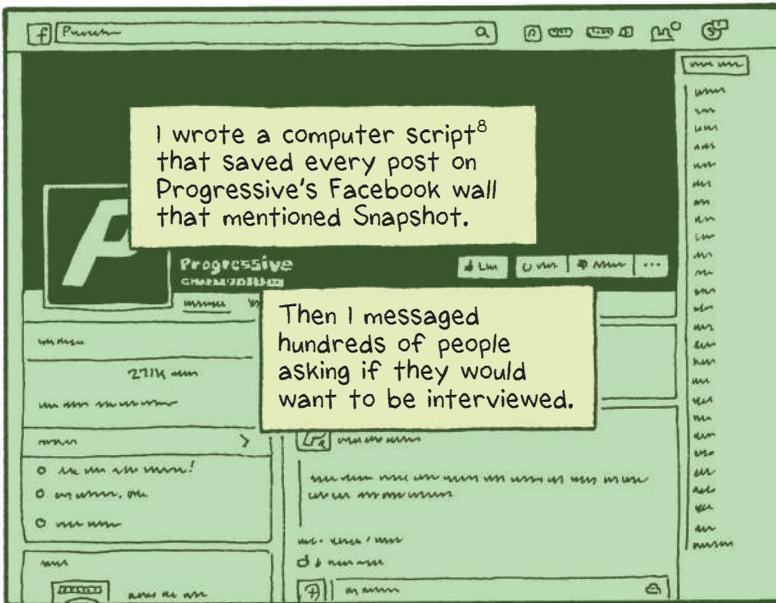


The facts of our daily life become the pillars of someone else's narrative about us.



Yeah, a lot of the issues in this piece make me think about what we as journalists do all the time.

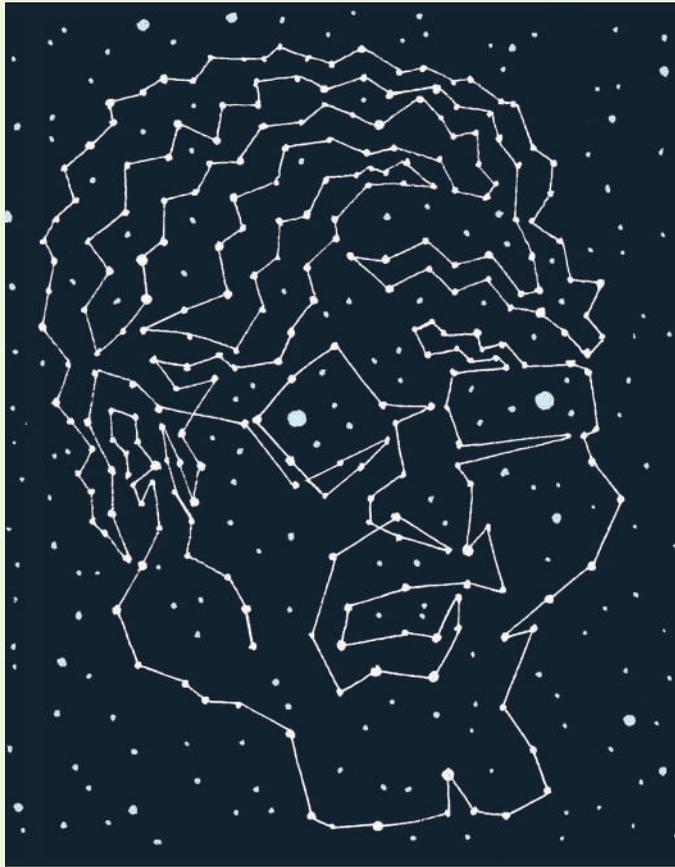
Right. For instance, I found other Snapshot drivers to speak with through Facebook.



I believe what I did was ethical. The comments were public. I identified myself as a reporter. And I didn't pester people who said they didn't want to talk.



But what if I weren't ethical?





We spoke some more about this with Scott Peppet and his colleague Paul Ohm.



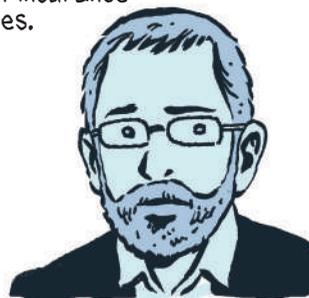
I think a large part of meaningful privacy reform is convincing people that privacy is worth giving some stuff up for.



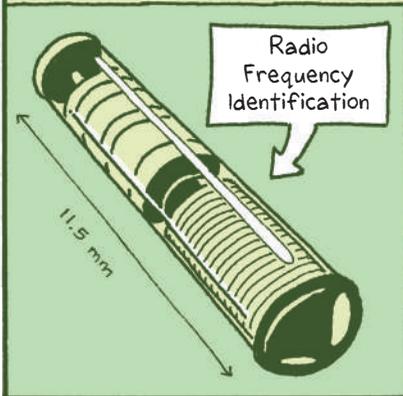
And to date, that argument is rarely heard in D.C.



It sometimes works, though. Peppet has written about how the Genetic Information Nondiscrimination Act--GINA--makes it illegal to use someone's genetic information for certain insurance purposes.



It's also illegal to, say, require employees to implant TRACKING CHIPS in their bodies.



And society penalizes discrimination based on completely public or observable traits like race.



You call these examples "dead body stories" because they show the actual harm so starkly.



The problem with laws that arise from "dead body" stories, they told us, is they cover only very narrow things--like implantable chips.

More comprehensive reform is hard to get through Congress. So what happens is, states pass their own patchwork of laws, looking to one another for workable solutions.



It seems like we're in a big waiting game... waiting for someone to get it right.

These experts told us about how California figured out one piece of it--data breach laws requiring that companies notify their customers if certain financial info is hacked.

Now 47 states have similar laws, and we're seeing the effects, like being notified of the recent Target breach.



This stuff is hard. I'm glad we're not lawyers. Let's head back to New York...

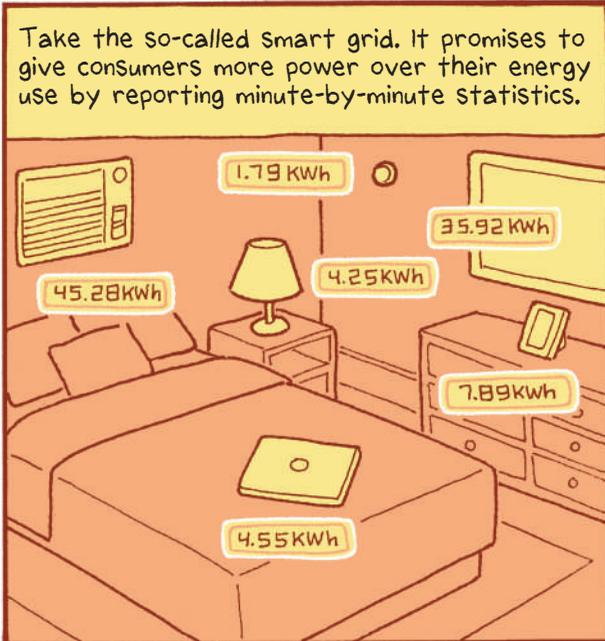


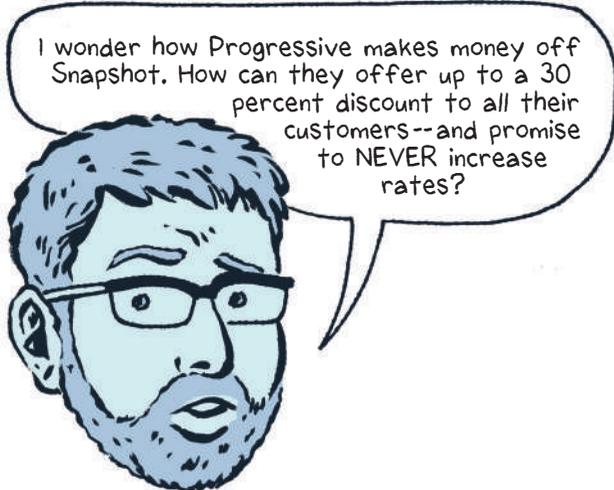


He said studies have shown that these member prices aren't any cheaper when you compare them to inflation--or to prices at stores that don't have member programs.



If you DON'T use the card you're paying MORE than inflation.





I asked Progressive those EXACT questions...

Snapshot
WAAW@Progressive.com www.pr.com/snapshot
 to me.

Offering Snapshot allows us to attract new customers. One in three drivers who are eligible for Snapshot choose to sign up. Our data shows that those drivers who receive a discount stay with us longer. [...] No one pays more as a result of drivers receiving Snapshot discounts.

Is there any way that Progressive benefits from Snapshot other than attracting new customers and having better retention?

After analysis of billions of miles in driving data, Progressive has found that key driving behaviors--like miles driven, braking and time of driving--carry more than twice the predictive power of traditional insurance rating variables, like a driver's demographic profile, income and the year, make and model of the insured vehicle.



A few months ago, I talked to Amanda Caswell, a young woman here in New York, about the idea of her employer collecting health data.

She uses a Fitbit to track her activities, her sleep habits and other personal metrics.

She had worked in marketing and was VERY familiar with data collection and tracking.



If your workplace started offering a program that would use your Fitbit data to help reduce your insurance costs, would you do it?

Sort of like the Progressive Snapshot program, they say, "We'll record your data for six months, at the end of which you might get a health insurance discount-- but we won't RAISE your premiums."



Hell, yeah. I would do it because I'm healthy right now.



I'm awesome. I bike, I walk, I'm healthy, I spin.



But I don't know... Then they have all my information.



See, it's hard... On the one hand, you have the tangible thing--the discount [you're] getting--and on the other hand, you have your personal information.

And that's the thing, you don't know [how] it could affect you, even though it's your personal information. It can't GET more personal.



So here we've been talking about how you find these services so creepy, and you collected similar data on people at your former workplace ...but you don't hesitate to opt in...

That's the pay to play. I have to pay you somehow in order to "play" for free.



Programs like what I discussed with Amanda--that ask for data and reward employees for good health behavior--aren't hypothetical.



"OBAMACARE" increases the incentives possible under these programs, such as weight tracking, quitting smoking, or "health-risk assessment,"¹⁰ like they do at Caterpillar...



A December 2013 market analysis by research firm IBISWorld¹¹ said that workplace wellness incentive programs will see "strong growth" and are currently worth \$2.2 billion in yearly revenue.

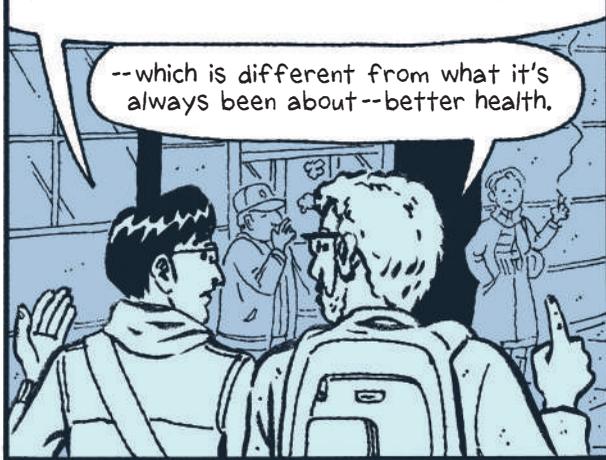


I spoke with a few other experts at Wharton and UCLA who expressed doubts that this information will **NECESSARILY** lead to lower costs.



They took issue with the way companies run these programs or how they structure incentives...

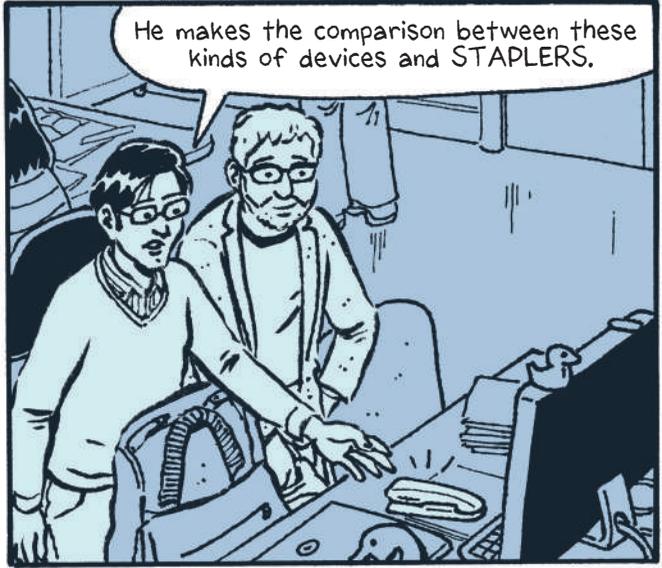
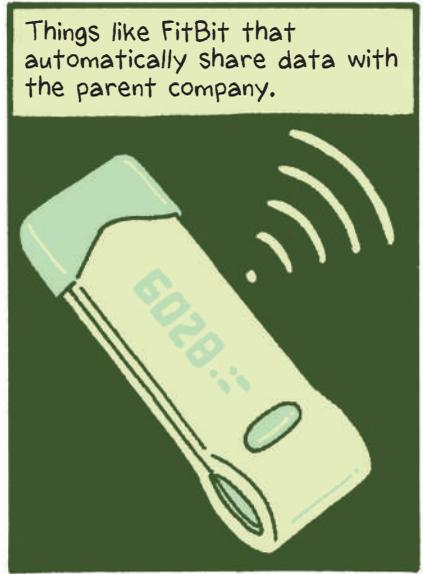
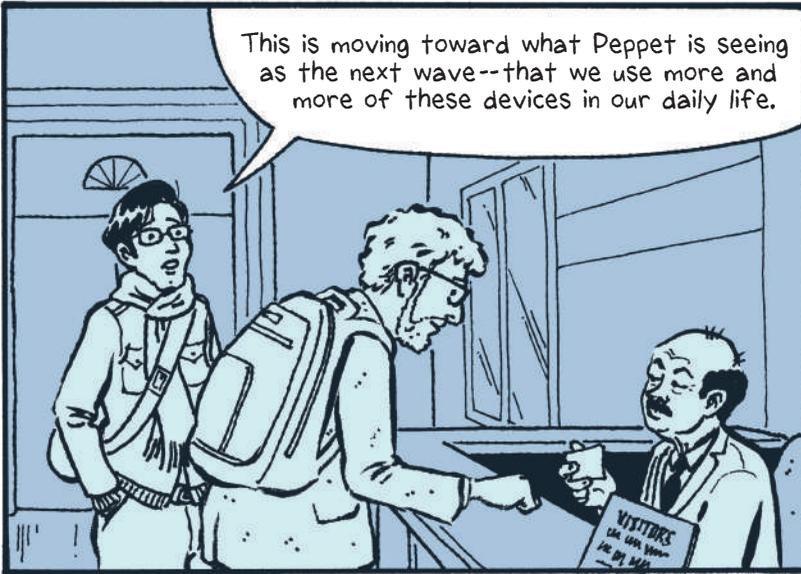
The main problem--the thing that throws the whole incentive business into question--is that it frames medical care in dollar amounts--



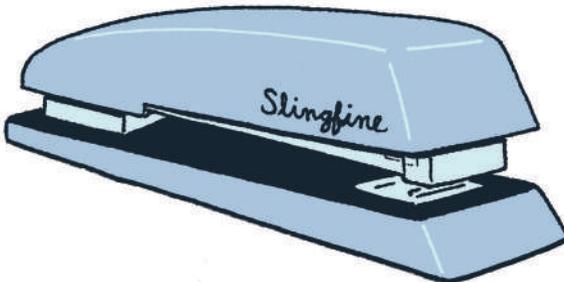
--which is different from what it's always been about--better health.

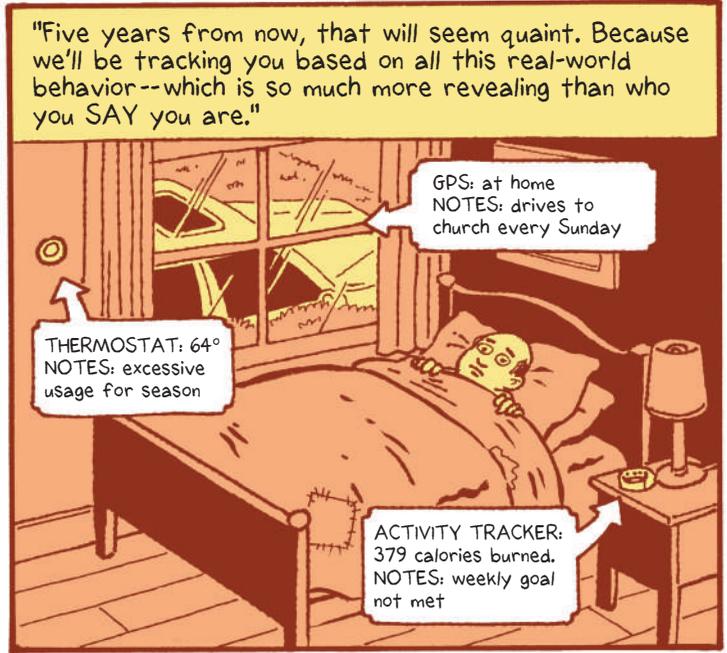
When a patient comes in with a heart attack, you don't want the doctor being concerned about the return on investment.



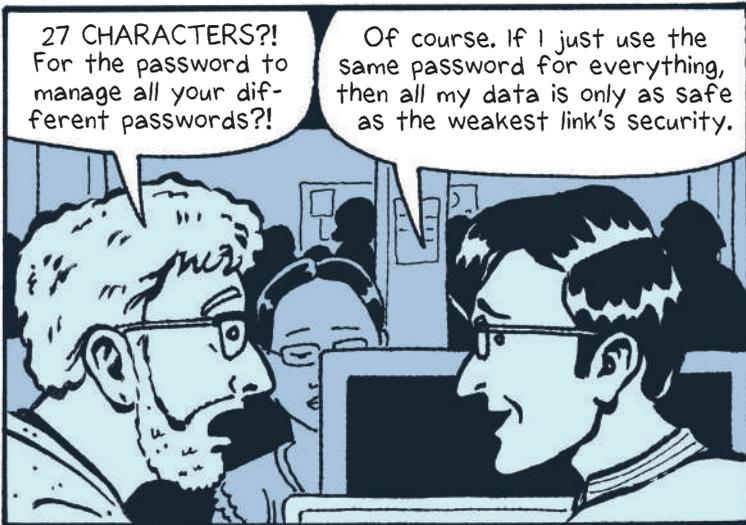
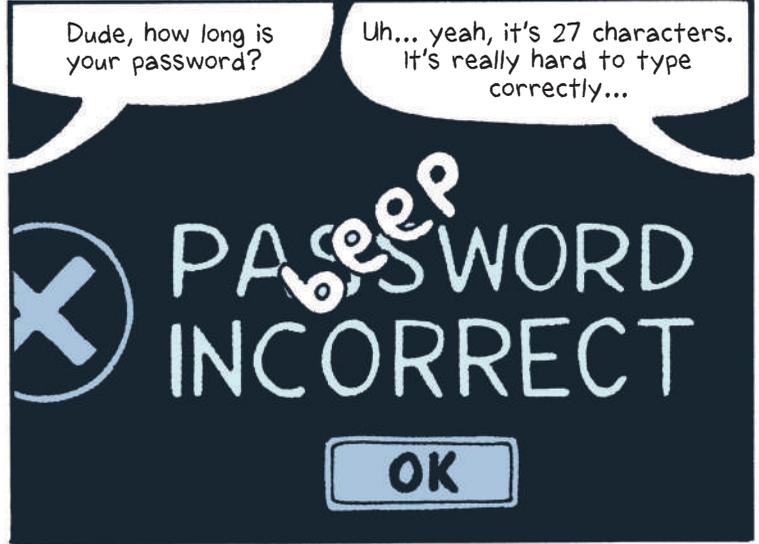
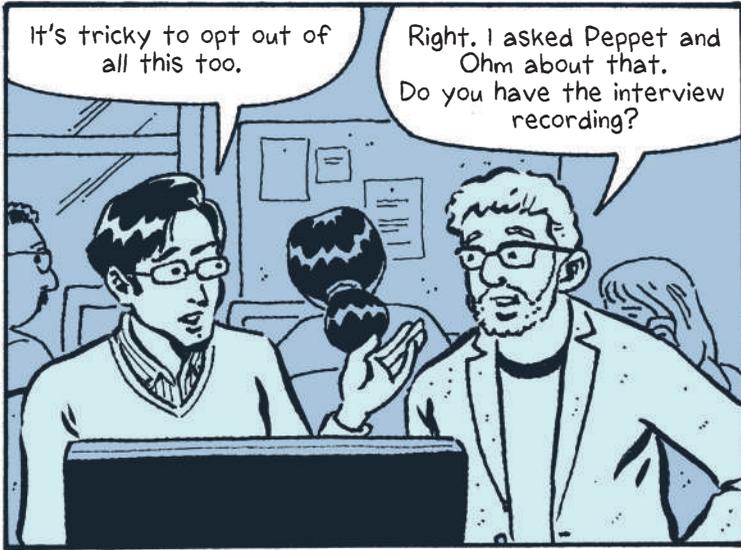


When you buy a stapler, it's just a stapler. You just use it. It doesn't share information--about, say, how many staples you used or the text content of the papers you bound together...







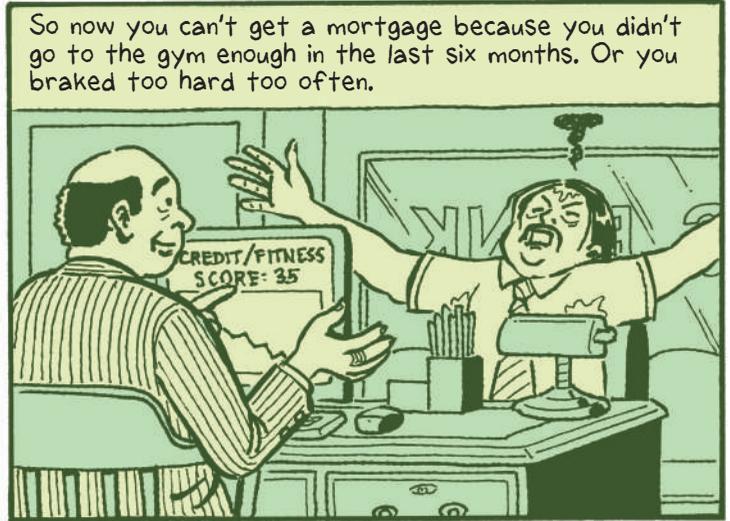






They scan your iris, and they take your fingerprint. And when I walk up to the Clear line and I shove this in the little machine and I put my finger on the thing, it verifies instantly that it's me--and I walk through.





Let's say your social network--who you're friends with, etc., their jobs, etc.--was factored into your credit rating.

Because people who have successful friends are more stable.

Isn't that a saying? "Show me who you walk with, and I'll tell you who you are?" Esmeralda Santiago?

Right. But that's really problematic. You should be able to associate with whomever you want--and not be penalized just for knowing someone.

What kind of society does that look like? Any potential new friend could harm you if they turn out to be anything but a sure thing.

Or what if they record not just how often you drive but WHERE you drive? Which neighborhoods...

Exactly. And how much does the model let you change, let you evolve? How easily can you go from being a high-risk person to a low-risk person?

Yeah. Can some people make those jumps more easily? How is that calculation made, anyhow? Don't we have a right to know how it's made?

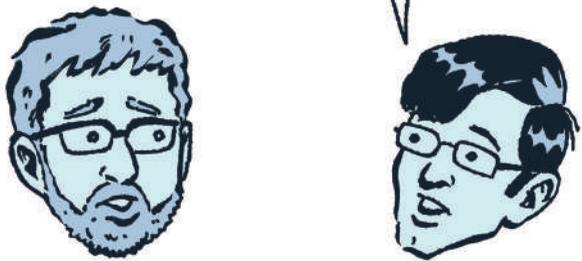
That's the thing. We can see the items in our credit report--but we largely DON'T know the secret formula used to produce the credit score.

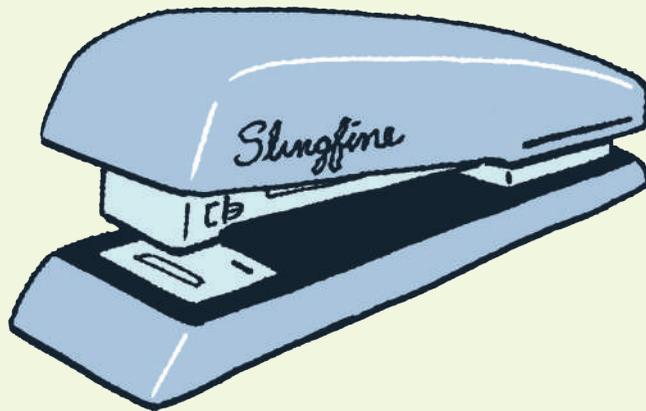
Now multiply that by a hundred--a thousand!--for all the other data out there about you.

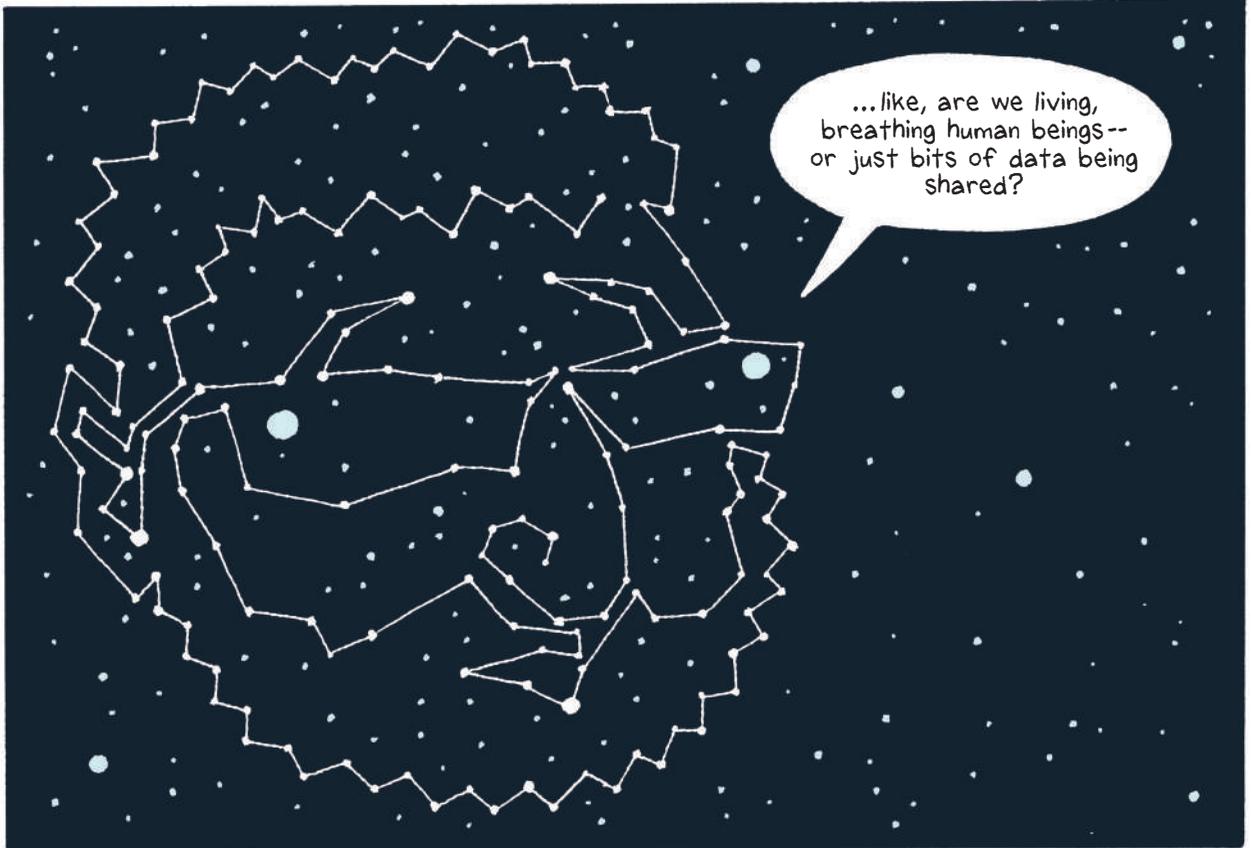
Maybe we've been given a false choice between opting in and giving up control over how that information is used--

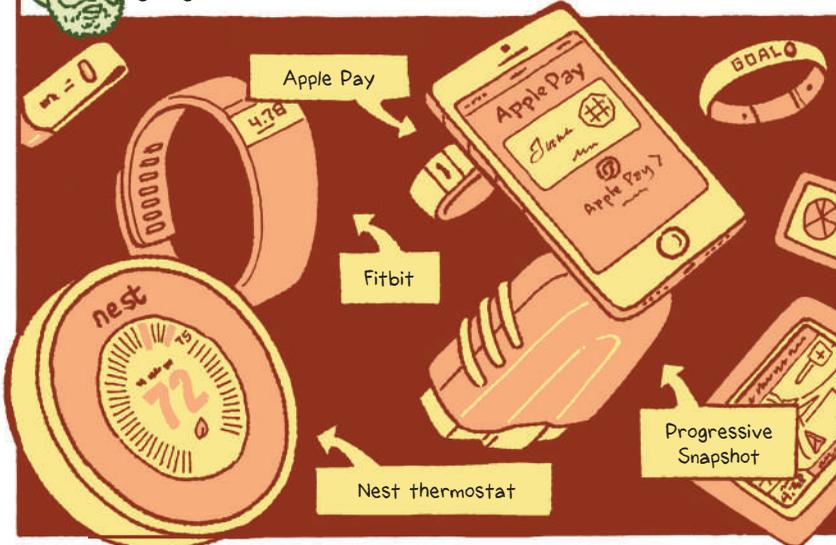
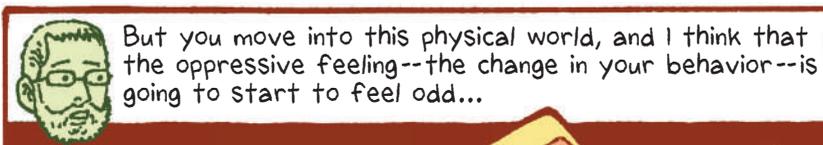
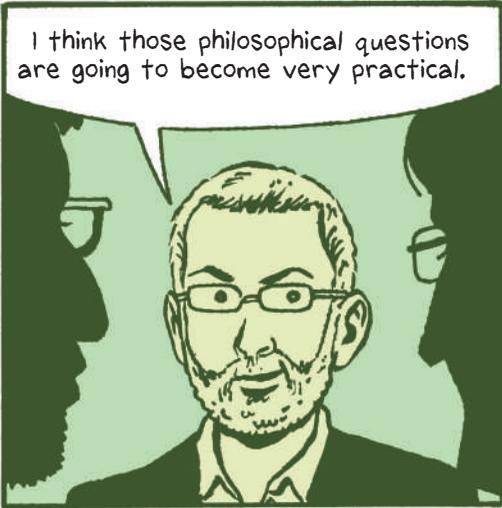
--between sharing everything and being left out.

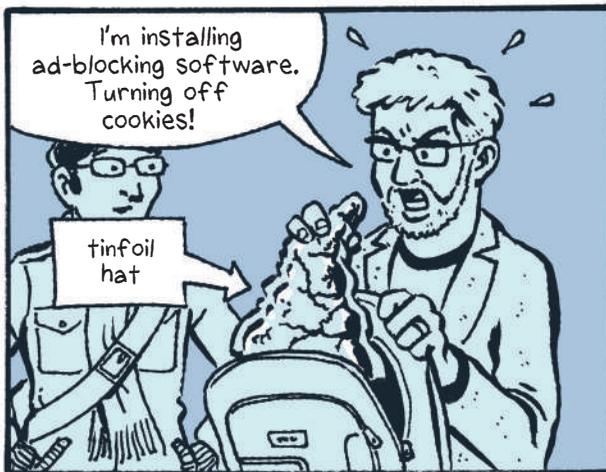
And like data breach laws, if there's a reasonable alternative that changes incentives for companies to NOT use data in uncomfortable ways.



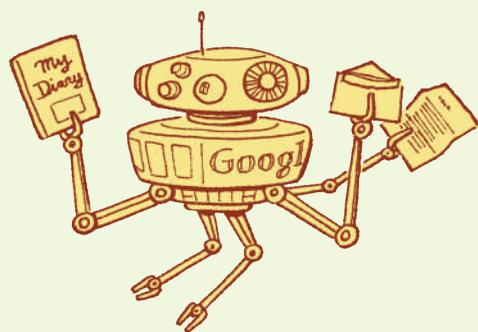












ENDNOTES

- ¹ p. 2, panel 1: The 2004 letter privacy advocates sent to the California Attorney General. <http://epic.org/privacy/gmail/agltr5.3.04.html>
- ² p. 2, panel 3: Senator Figueroa's proposed bill, SB-1822. http://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=200320040SB1822&search_keywords=
- ³ p. 8, panel 2: Geer, Dan. "Tradeoffs in Cyber Security." 2013. *UNCC*. <http://geertinho.net/geeruncc.9x13.txt>
- ⁴ p. 9, panel 2: Jaffe, Eric. "The First Look at How Google's Self-Driving Car Handles City Streets." 2014. *The Atlantic's CityLab*. <http://www.citylab.com/tech/2014/04/first-look-how-googles-self-driving-car-handles-city-streets/8977/>
- ⁵ p. 10, panel 1: Foursquare has since split into two apps, **Foursquare** and **Swarm**, which focus on location reviews and sharing check-ins with your friends, respectively.
- ⁶ p. 11, panel 1: Peppet, Scott. "Unraveling Privacy: The Personal Prospectus and the threat of full disclosure." 2011. *Northwestern University School of Law*. <https://www.documentcloud.org/documents/1293348-unraveling-privacy-the-personal-prospectus-and.html>
- ⁷ p. 14, panel 3: Solove, Daniel. "Why Privacy Matters Even if You Have 'Nothing to Hide.'" 2011. *Chronicle of Higher Education*. <http://chronicle.com/article/Why-Privacy-Matters-Even-if/127461/>
- ⁸ p. 20, panel 4: FaceLift: A library for scraping Facebook posts by keyword. <http://github.com/ajam/facelift>
- ⁹ p. 21, panel 5: Zimmerman, David. "Chick-fil-A Increases Donations to Anti-Gay Groups." 2013. *Boston Spirit*. http://www.boston.com/lifestyle/blogs/bostonspirit/2013/03/chik-fil-a_increases_donations.html
- ¹⁰ p. 29, panel 2: Wiczner, Jen. "Your Company Wants to Make You Healthy: A look at the pros and cons of the most popular wellness programs used by companies." 2013. *Wall Street Journal*. <http://online.wsj.com/news/articles/SB10001424127887323393304578360252284151378>
- ¹¹ p. 29, panel 3: "Corporate Wellness Services in the US: Market Research Report." 2013. *IBISWorld*. <http://www.ibisworld.com/industry/corporate-wellness-services.html>
- ^{12a/b} p. 40, panel 2 & 4: These panels are an imaginary representation of how a typical person might respond to the way these new forms of tracking are changing our daily existence. The words are Peppet's.



Michael Keller is a multimedia journalist at *Al Jazeera America* covering issues at the intersection of technology and civic life. His work has appeared in *The Washington Post*, *The Atlantic*, *Newsweek/Daily Beast*, and others. He graduated from the Columbia University Graduate School of Journalism in 2011 and is a research Fellow at the Tow Center for Digital Journalism at Columbia University. [mhkeller.com]



Josh Neufeld is a nonfiction cartoonist living in Brooklyn. His previous works include *A.D.: New Orleans After the Deluge*, *The Influencing Machine: Brooke Gladstone On the Media*, and the ongoing series *The Vagabonds*. Neufeld was a 2012–2013 Knight-Wallace Fellow in Journalism at the University of Michigan. [JoshComix.com]

*Terms of Service:
Understanding Our Role
in the World of Big Data.*

October 2014. This is a work of journalism. No portion of this comic may be reproduced without written permission, except for review purposes. Copyright ©2014 Al Jazeera America, LLC and Josh Neufeld. All rights reserved.

This is the print companion of the original e-reader edition: projects.aljazeera.com/2014/terms-of-service. Got comments, order inquiries, or stories to share? Drop us a line at terms-of-service@aljazeera.net.

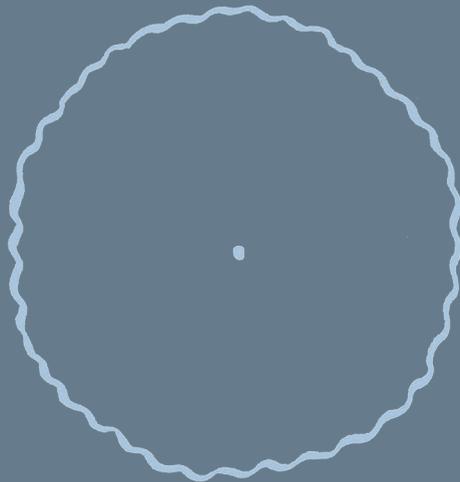
Cover design by Michael Keller and Vaughn Wallace. Publication design & production by Josh Neufeld.

Thanks: Sari Wilson, Ashley van der Grinten.

Second printing November 2014.
Published by Al Jazeera America, LLC.
Printed in the U.S.A.

Terms of Service

Understanding Our Role
in the World of Big Data



ALJAZEERA
AMERICA™

© 2014 Al Jazeera America, LLC and Josh Neufeld.
All rights reserved.