equivalent of every U.S. citizen writing three tweets per minute for almost 27,000 years.[2] Ninety percent of the world's data, from the beginning of time until now, has been generated in the past two years,[3] and it is estimated that that total will double every two years from now on.[4]

Big data will have important, even transformative uses. No one questions some of the benefits big data analytics can bring. They include increased personalization for daily activities—helping companies determine which ads you see online, which articles a newspaper recommends to you, and which book to recommend you read next. But the potential benefits may also address important societal issues—keeping kids in high school,[5] conserving our natural resources by

[2] Lucas Mearian, World's data will grow by 50X in next decade, IDC study predicts, COMPUTERWORLD , June 28, 2011, available at http://www.computerworld.com/s/ article/9217988/ World_ s_data _will_grow _by_50X_in_next_decade_IDC_study_predicts?pageNumber=1

[3] Science News, Big Data, for Better or Worse: 90% of World's Data Generated over Last Two Years, SCIENCE DAILY , May 22, 2013, available at http://www.sciencedaily.com/releases /2013/05/130522085217.htm

[4] Steve Lohr, The Age of Big Data, N.Y. TIMES, Feb. 11, 2012, available at http://www.nytimes .com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html?pagewanted=all&_r=0

[5] Centre for Information Policy Leadership, Big Data and Analytics: Seeking Foundations for Effective Privacy Guidance, at 6-7 (Feb. 2013), available at http://www.hunton.com/files/Uploads/Documents/News_files/Big_Data_and_Analytics_February_2013.pdf

devices that they buy for one purpose – making coffee, storing food, driving to work – but that collect and use a vast amount of personal information about them. Whether it is a connected car, home appliance, or wearable device, the data that these connected devices generate could be higher in accuracy, quantity, and sensitivity and, if combined with other online and offline data, have the potential to create alarmingly personal consumer profiles.

Will consumers know that connected devices are capable of tracking them in new ways, especially when many of these devices have no user interface? How

Similar questions arise in the ongoing discussion about online tracking. For several years now, regulators and industry standard-setting organizations, among others, have focused on cookie-enabled online tracking, and on providing consumers appropriate choices about such tracking.[13] But in recent months, we have seen industry turn its attention to developing other technologies to track consumers. Fingerprinting, which could uniquely identify a consumer's browser and obviate the need for cookies, would provide consumers with even less control.[14] As consumers turn increasingly to their smartphones and tablets, where cookies do not work, industry has deployed other mechanisms to track consumers.[15] How will these tracking technologies affect consumers'

information being collected, whether it is being shared with third parties, to whom, and for what purpose?

These ques

those involving their sexual orientation, health conditions, financial condition, and race.

Let's look at a well-known, even infamous, example. Before Target made news for a data security breach that may involve 110 million consumers' credit cards and debit cards, the company received a lot of attention for its big-data-driven campaign to identify pregnant customers through an analysis of consumers' purchases at its stores, a so-called "pregnancy prediction" score.[16] Target was able to calculate, not only whether a consumer was pregnant, but also when her baby was due.[17] It used the information to win the expectant mom's loyalty by offering coupons tailored to her stage of pregnancy.[18]

To be clear, I don't have any information indicating that Target sold its pregnancy predictor score or lists of pregnant customers to third

[16] See Charles Duhigg,

company that analyzes innocuous data from social media and the like to predict disease conditions like diabetes, obesity, and arthritis in order to persuade particular consumers to join medical trials.[21] All of this is happening outside of HIPAA—outside any US regulatory scheme to protect this information.

challenge the accuracy of the data. Similarly, we should be concerned about the risk that such sensitive personal information may fall into the wrong hands through a data breach. But more fundamentally, I believe we should be concerned about the damage that is done to our sense of privacy and autonomy in a society in which information about some of the most sensitive aspects of our lives is available for analysts to examine without our knowledge or consent, and for anyone to buy if they are willing to pay the going price.

These concerns, of course, are not limited to the world of commercial data brokers. We don't have to pass judgment on the NSA to acknowledge the recent disclosures have sparked a necessary and overdue debate on how to balance national security and citizens' privacy rights. For those of us who have been looking at the issue of privacy in the Internet age for several years, there is a further benefit: Americans are now more aware than ever of how much their personal data is free-floating in cyberspace, ripe for any data miner – government or otherwise – to collect, use, package, and sell.

But with that knowledge comes power – the power to review, this time with eyes wide open, what privacy means – or should mean – in the age of the Internet. I believe that's what President Obama meant in June, and again last month, when he noted that the "challenges to our privacy do not come from government alone. Corporations of all shapes and sizes track what you buy, store and analyze our data and use it for commercial purposes,"[24] and when he called for a "national conversation…about…the general problem of … big data sets, because this is not going to be restricted to government entities."[25]

During our ongoing discussion about NSA surveillance, national

programs and services built on big data analytics.[26] They urge adoption of enhanced privacy protections as a key part of strengthening this trust.

I agree. While I firmly believe that the national security issues must be addressed separately from the commercial privacy issues, I firmly believe that the promise of big data – the huge benefits that society and individuals may reap from appropriate and careful use of data analytics – will not be reached until we address some of these key consumer privacy concerns stemming from the creation, collection and use of sensitive consumer data and profiles.

[26] Brad Smith, Time for an International Convention on Government Access to Data, Microsoft on the Issues (Jan. 20, 2014), available at http://blogssnt6.15 0 6 0 0 1 scn /TT1 1 Tf -0.1f1 Tc 0.001 Tw 9.96 -0 0 9dusrE.9(o)19 toy r on]0 9d D

Here are the steps I believe must be taken by policy makers and industry in the commercial sphere

they should make a public commitment not to try to re-identify the data; and they should contractually prohibit downstream recipients from doing the same.[28]

Robust deidentification efforts along these lines will help address some of the problem. But such robust deidentification will not solve the problem of big data profiling. The entire data broker enterprise seeks to develop greater insight into the activities, status, beliefs, and preferences of individuals. The data the industry employs are therefore about or linkable to individuals – or as a recent trade association's report refers to it – "individual-level consumer data".

2. Create Institutional Ethical Monitoring

Another solution offered to the challenges big data presents to privacy is the creation of entities that monitor the ethical use of data.

[28] See FED TRADE COMM'N, PROTECTING CONSUMER PRIVACY IN AN ERA OF RAPID CHANGE: RECOMMENDATIONS FOR BUSINESSES AND POLICYMAKERS 21

One proposal calls for the creation of "Consumer Subject Review Boards" to determine whether particular projects using consumer data are both legal and ethical.[29] Another proposal calls for individual companies to install the "algorithmist" – a licensed professional with ethical responsibilities for an organization's appropriate handling of consumer data.[30] But the Consumer Subject Review Boards and the algorithmist will only thrive in firms that thoroughly embrace "privacy by design" – from the engineers and programmers all the way up to the C-suite – firms that understand the legal and ethical dimensions of the use of algorithms to make decisions about individuals.

3. Change the Law

Changing the law would help. As some of you have heard me say before, we have pretty good laws in the US governing commercial

[29] See Ryan Calo, Consumer Subject Review Boards, 66 STAN. L. REV. ONLINE 97 (2013), available at http://www.stanfordlawreview.org/online/privacy-and-big-data/consumer-subject-review-boards; Jules Polonetsky, Omer Tene, & Christopher Wolf, How to Solve the President's Big Data Challenge, IAPP Privacy Perspectives, Jan. 31, 2014, available at https://www.privacyassociation.org/privacy_perspectives/post/how_to_solve_the_presidents_big_data_challenge

[30] See VIKTOR MAYER-SCHÖNBERGER & KENNETH CUKIER, BIG DATA: THE REVEOLUTION THAT WILL TRANSFORM HOW WE LIVE, WORK, AND THINK 180–182 (2013).

privacy, and we have excellent enforcement. The FTC—the leading privacy regulator in the United States—has built a robust data protection and privacy enforcement program that focuses on both traditional offline products and services,[31] as well as on the evolving digital and mobile marketplace.[32] The FTC uses its authority to stop unfair or deceptive practices that violate consumers' privacy or place consumers' data at risk.[33] We also enforce laws that protect consumers' financial[34] and health[35] information, information about children,[36] and information used

[31] See, e.g., U.S. v. Check Servs., Inc., No. 1:03-cv-01247 (D.D.C. Aug.ter /Typwum.t6-0.002 Tw 15.96 -0 0 11002 T71.75 385

to make decisions about credit, insurance, employment, and housing.[37]

We engage in rigorous data security enforcement, as was clear when we announced our 50th data security enforcement action earlier this month. And, notably, the FTC vigorously enforces the U.S.-EU Safe Harbor Framework, as demonstrated by our recent actions against thirteen companies with false membership claims.[38] I believe that Safe Harbor is an appropriate data transfer mechanism that gives the FTC an effective tool to protect the privacy of EU citizens.[39]

Yet I believe we need to improve our commercial privacy laws in the US. When I talk about these issues in W-3.857P1o(en)-3m63qt-3.8(t,)3(f)

technological tools to reassert some control over their personal data.[40] Put simply, consumers should have more control over decisions like how much to share, with whom, and for what purpose to reclaim their names.

Here's how it would work. Through creation of consumer friendly online services, Reclaim Your Name would empower the consumer to find out how brokers are collecting and using her data; give her access to information that data brokers have amassed about her; allow her to opt-out if she learns a data broker is selling her information for marketing purposes; and provide her the opportunity to correct errors in information used for substantive decisions.

agree to tailor their data handling and notice and choice tools to the sensitivity of the information at issue. As the data they handle or create becomes more sensitive—relating to health conditions, sexual orientation, and financial condition, for example—data brokers would provide greater transparency and more robust notice and choice to consumers.

The user interface is also critical. It should be user-friendly, and industry should provide a one-stop shop so consumers can learn about

across websites. The Digital Advertising Alliance has deployed an icon-based opt out system – the About Ads Program – and has promised to work collaboratively with browsers so that consumers' choices will be persistent and honored no matter how they are initially exercised.[43] And an international standard-setting organization – the W3C – has convened a working group to create a universal Do Not Track standard through a consensus-based process with representatives from across the spectrum of stakeholders. The State of California's recently enacted law requiring websites that collect personally identifiable information to disclose both how they respond to Do Not Track signals and whether personally identifiable information about a consumer's online activities can be collected when the consumer uses the website as an additional incentive for these various (ed)-3.7(pj3ated)-3.((s)-2(.8(o)-3.4[v2)2[44]

obtain consumers' consent before collecting and sharing consumer data.[45] I urge all of the stakeholders to forge ahead with their work and reach consensus to implement an effective, universal and comprehensive Do Not Track system.

If consensri15.96 -0 0.96 -8_396 -8_3(g)/St( )]TJ 0 15.99(h()-3.7(en)-3.7

action to protect consumer privacy. By implementing the steps we've outlined, industry (and policy makers) can help create an ecosystem that respects consumer privacy and engenders consumer trust, allowing big data to reach its full potential to thrive and benefit us all.