Balancing Privacy and Opportunity in the Internet Age

A Background Paper for the Joint Informational Hearing of the
Assembly Judiciary Committee, the Assembly Business, Professions & Consumer Protection Committee, and the Assembly Select Committee on Privacy

December 12, 2013
Santa Clara University

By the Staff Counsel of the Assembly Judiciary Committee

"On the Internet, nobody knows you're a dog."

Introduction

In 1993, when the above cartoon ran in the *New Yorker*,[1] the anonymous nature of the Internet was widely assumed. While anonymity may have been the norm in 1993, in the current age of "Big Data" Internet anonymity may be a comforting but misplaced illusion. Despite its name, "Big Data" –a term increasingly used in the press and scholarly literature – does not simply refer to the *size* of online and digital databases. Just as importantly, it refers to the powerful analytical tools that reveal hidden connections, patterns, and correlations within that data, also known as "data mining". "Big Data" now creates the possibility that even the most disparate pieces of information, when analyzed with other disparate pieces, can be used to identify a specific Internet user. And yet, even as anonymity and privacy fade into memory, the opportunities and progress spurred by California's constantly innovating technology companies have been nothing short of revolutionary.

This background paper seeks to touch, at least at the surface, on some of the risks and benefits of online collection, sharing, tracking, and marketing of personal information by Internet web sites and related businesses in the age of "Big Data." Although there has been a robust debate regarding

tracking and collection of personal data by federal, state and governmental authorities – a debate that unites many privacy advocates and technology companies who otherwise disagree about the private use of personal data – those equally important issues are well beyond the scope of this paper. The first part of this paper will briefly explain how the online market in personal information operates; how personal information is typically collected, shared, and ultimately used; and what kinds of entities are often involved. The paper will next consider some of the approaches that California and governments around the world have taken to address privacy concerns related to Internet marketing and some of the apparent limitations of those regulatory schemes. Some of the impressive recent voluntary initiatives being undertaken by thought-leaders within the tech sector are briefly highlighted as well. The paper concludes by raising some important policy questions on the tracking, collection and use of consumer data.

A. The Extraordinary Reach and Impact of the Internet is Difficult to Overstate

As the Legislature considers challenging questions about how the reasonable privacy interests of Californians can best be protected, it is important to remember just how rapid and far reaching the impact of the Internet has been. Former president Bill Clinton often notes that when he became president just 20 years ago, there were literally only 50 websites. When he left eight years later, there were already 350 million. In the 12 years since, that number has grown exponentially. The Internet now drives the hottest stocks on Wall Street, shapes and reshapes technological innovation on a seemingly hourly basis, provides the foundation for global communication, advances health care in breathtaking ways every day, and makes information and entertainment more accessible for literally billions of people on the planet. Indeed, as Bill Gates has stated, "The Internet is becoming the town square for the global village of tomorrow." So what does this mean for society, government, commerce, and other institutions? No part of our society, or the world, is now untouched by the Internet. It has spurred a true revolution comparable to, if not greater than, the Industrial Revolution, and its impacts continue to compound. Thus any undertaking by government to regulate this marvel must be considered in this context.

B. The Growing Challenges to Traditional Notions of Privacy in the Internet Age

As was discussed in the Assembly's first informational hearing on Internet privacy on March 19, 2013, California has been ground zero in the development of the Internet and so many other advances in human society in the past half century. At the same time, this information revolution has created new and unprecedented challenges to our traditional notions of privacy.

The recent *Apple v Superior Court of Los Angeles County (Krescent)*[2] decision this past February highlighted the need for California privacy law to be updated from the "brick and mortar" world to an online world reflective of a new "e-commerce" business models that foster innovation while providing access to content and services, often for free. Lawmakers now appropriately strive to determine how best to strike the balance between a robust and innovative Internet and one that adequately protects individual privacy and maintains consumer trust. This is clearly not an easy task, but it is a critically important one.

Indeed, legal scholars, journalists, and other commentators are increasingly drawing policymakers' attention at all levels of government to how new technologies and business methods are posing new

threats to our privacy and taking advantage of consumers' lack of understanding about how data about them is collected and shared.

In a 2010 series entitled, "What They Know," the *Wall Street Journal* (WSJ) published more than a dozen articles regarding on-line data gathering– and what they found was unsettling. Remarkably, according to WSJ, on-line data gathering is the fastest-growing business in America. Its pervasiveness and growth have been astounding. The inaugural WSJ article found that "the nation's 50 top websites on average installed 64 pieces of tracking technology onto the computers of visitors, usually with no warning."[3] This tracking technology, usually referred to as "cookies" (more on this later), consists of small files that are downloaded onto the user's browser and has the capacity to track subsequent websites visited by that user. While users of these websites may voluntarily disclose personal information to use the websites they visit, the series noted, they often do not know that the majority of those websites permit third parties, including advertising networks, to install cookies on the user's computer as well.

The use of cookies exposes a major paradox about the Internet today: the gathering of personal data about Internet users' preferences and behaviors is the critical building block for building a seemingly "free" Internet where so much remarkable content and so many amazing services are instantly at our fingertips. Yet it is that very practice of mining personal information that, without adequate controls and oversight, may inadvertently subject all Internet users to the unknown and potentially unwanted sharing of their personal information – and, perhaps more importantly, may be devaluing our basic and traditional notions of privacy. Curiously, relatively few Internet web sites are known to allow users to pay a fee for access in order to retain their privacy, rather than exchanging personal information in return for free services.

As part of its "What They Know" series, the WSJ published an exchange of editorials by Nicholas Carr, an author and privacy rights advocate, and Jim Harper, the director of information policy studies at the Cato Institute. Carr argued that the harvesting of our personal information without our knowledge, much less our consent, was nothing less than an "assault on liberty." In addition to the danger that our personal information will end up in the wrong hands, or that advertisers will manipulate us and our information in order "to influence our behavior and even our thoughts in ways that are invisible to us," Carr saw an even greater danger: that "*continuing erosion of personal privacy . . . may lead us as a society to devalue the concept of privacy, to see it as outdated and unimportant*" (emphasis added).[4]

In stark contrast, Harper emphasized the many benefits that consumers gain from tracking and information sharing. It is not simply that targeted ads are more useful to us. More importantly, Harper contends, it provides users with more free online services: "The reason why a company like Google can spend millions and millions of dollars on free services like its search engine, Gmail, mapping tools, Google Groups, and more is because of online advertising that trades in personal information."[5]

When it comes to the collection, sharing, and tracking of personal information, these contrasting views go to the heart of the matter. As Harper notes, companies like Google and most other commercial websites make their money by selling advertisement space based on the user's profile or by permitting third parties to install cookies on their websites. Online services, like Google maps, are very expensive to produce and maintain, yet they are often free to the user. Without the advertising

revenue, Google and other companies would need to charge a user fee. Of course, some sites– for example, the popular Ancestry.com – charge users a monthly fee, but this is not the historical norm. Given the advertising-driven business model that dominates the Internet, what policy strikes the appropriate balance between providing consumers with free access to the kinds of online services they clearly desire, while at the same time protecting consumers' reasonable interest in privacy? Is it the responsibility of each consumer to strike that balance for himself or herself? More importantly, can consumers find that balance if they lack sufficient knowledge about the kinds of information that is tracked, with whom it is shared, and all of the purposes for which it might ultimately be used, not only by the website that collects the information but by the third parties to whom it is sold? Can consumers strike that balance if they do not have notice and control over third-party use of their personal information?

C. Self-Regulate or Government Mandate? How Should Lawmakers Strike The Right Balance Between A Robust And Innovative Internet And Adequate Protection of Individual Privacy?

Perhaps the key policy question in the debate over online tracking and behavioral advertising concerns the extent to which privacy protections should come from industry self-regulation or government-mandated regulation, or a collaboration of efforts by both working together. According to a recent report by the *Washington Post*, browser manufacturers – including Google, Apple, Microsoft, and Mozilla – are considering browser controls that would limit the ability of third-party advertisers to install tracking cookies on a user's computer browser. At present, some browser manufacturers, including Microsoft, have implemented privacy controls in their browsers that allow a user to transmit a "request" not to track their behavior across websites. However, neither existing controls or options, nor existing law, appear to require an advertising network or commercial website to *honor* those requests. According to the *Post* report, the new devices under consideration would not simply send a request but actually *block* cookies. Just what the effect of this development will be is uncertain. Some privacy groups applaud the idea as a meaningful control. Other privacy advocates contend that it will lead to an "arms race" as the advertising industry develops new technologies that counter the new controls. Some advertisers, on the other hand, contend that this will destroy the Internet by undermining the business model that provides users with free online content and services.[6]

Thus the challenge facing policymakers today: determining the right balance in protecting Internet users' reasonable privacy expectations without killing (or even injuring) the proverbial goose that continues to lay so many golden eggs here in California and across the world. This effort needs to be made in light of the increasingly robust voluntary initiatives being undertaken by the technology industry itself, which has expressed a preference for self-regulation as a means to protect consumer privacy rather than increased government regulation.

PART ONE
HOW IT WORKS:
THE ECONOMICS AND MECHANICS OF ONLINE MARKETING

A. The Economics of Online Marketing.

As noted above, while a small number of Internet sites charge users for access, it appears that most operate for free or for minimal up-front fees, instead making their revenue from advertisers. Every search engine query and website visit is reportedly logged, collected, shared, and analyzed by multiple entities including the "publisher" of the visited website and an array of third-party market researchers, "analytics" providers, and advertising networks that track browsing behavior and collect various user information, all generally unbeknownst to the user.[7]

Although the mechanics of this process can be fairly complex, the economics are fairly simple: an abundance of free online content and services is paid for by "targeted" advertising that in turn relies upon "online behavioral tracking." Advertisers can infer a great deal about us based on the kinds of websites that we visit and the queries that we make on search engines. This information allows the advertiser to appeal to us – or at least to anyone using our computer or mobile device – with much more finely tailored, and presumably more effective, targeted advertisements. As a result, consumers enjoy seemingly free online services – educational content, video games, photo-sharing, and social networking among others – in exchange for inadvertently providing valuable marketing information. Because of the unique non-subscription evolution of the Internet, it remains unclear whether consumers would ever have been willing to pay "up front" for these remarkable online services if advertisers stopped funding the Internet.

And of course targeted advertising is nothing new. Long before the advent of the Internet, television advertisers were making assumptions about people who watched certain kinds of programming at certain times of the day, and they likewise tailored their advertisements accordingly. Daytime television dramas are still referred to as "soap operas," a term dating to the early days of commercial radio. The laundry detergent commercials that are the staple of daytime dramas gave way to beer and car commercials during Sunday football games. Online behavioral advertising is based on similar kinds of stereotyping, but the much more interactive nature of the Internet allows advertisers to engage in much more refined targeting. Visiting a particular website, like watching a particular television show, allows the advertiser – or the service assisting advertisers -- to draw many detailed inferences about potential consumers. But compared to web surfing, watching television is very passive. The television viewer does not give the advertiser any more information than the fact that she is watching a particular show. But when the same person visits a webpage, she may voluntarily enter additional information, or click on other links that reveal more about herself and her interests. From the advertiser's perspective, this more interactive process permits more accurate inferences. While the principles of targeted advertising remain the same, the tools have become remarkably – and some would say frighteningly – powerful, pervasive, and potentially intrusive.

B. Cookies and Beyond: The Mechanics of Online Tracking.[8]

Describing the technology of online tracking and the workings of the Internet is far beyond the scope of this paper. Moreover, the technology is constantly changing and its capacities constantly expanding. With that caveat in mind, it may be generally said that online "behavioral" advertising relies on mechanisms that permit advertisers to track a user's browsing behavior. Among the critical components of this tracking under current technology is the "cookie." According to Google's privacy policy, a cookie "is a small file containing a string of characters that is sent to your computer when you visit a website."[9] Cookies generally store a user's preferences and other information about the user's browsing behavior. Some consumers may reset or delete cookies to avoid being tracked, but

many website features or services will not function properly without cookies being enabled. Cookies also come in different forms.

First-Party Cookies Used Within a Single Website. Scholars and commentators on Internet commerce and privacy typically distinguish between "first-party cookies" and "third-party cookies." First-party cookies allow the website to collect certain information from the user's browser and then recall that information whenever that browser revisits the website. This can provide advantages to the website. For example, because of first-party cookies, Amazon reportedly knows what books or goods a customer has purchased or viewed on a prior visit and can therefore create more targeted and effective advertisements when the consumer returns.

Third-Party Cookies Used Within A Single Website. Websites may also use third-party cookies to provide "analytics" services for the website. For example, according to its privacy policy, Amazon.com employs a third-party company to provide its advertising "analytics." This third-party entity tracks the user as he or she moves through different pages *within* the website. The third party analyzes this behavior, makes certain assumptions about the user based upon it, and then advises the website operator on the best ways to increase traffic and make more sales to that customer.[10]

Third-Party Cookies That Track Users Across The Internet. Of greater concern may be third-party cookies that aid data brokers and aggregators.[11] These types of cookies are not operated for the benefit of the particular website visited. Instead, they are used to track a browser across multiple websites, and thus can draw many more inferences about the particular user. For example, so-called "ad networks" apparently place cookies on several of the most popular websites and, as such, have the capacity to track a single browser as it visits the many websites within the "network." The ad network does not necessarily know the user's personal identity, but more likely recognizes the browser or IP address. Advertisers generally do not need to know a consumer's name; it is sufficient to know that a person using a particular computer or device visited a certain combination of websites, and this allows them to make more refined inferences and send more finely tailored ads to those computers or devices. While consumers may or may not know whether a website uses first-party cookies, at least they know that they are dealing with that particular website. Consumers may be less aware of third-party tracking because a consumer usually has not established any direct relationship with the third-party tracker. Most website privacy policies disclose the use of third-party cookies with varying degrees of clarity for consumers who read closely and know how to interpret the policy. However the consumer may not be able to discover the identity of those third-party trackers. And of course most consumers likely never read the privacy policies, at least not closely.

C. Looking Ahead to the Post-Cookie Era.

According to recent press reports, just as policymakers are educating themselves to try to understand how “cookies” work, the Internet may be moving into a "post-cookie era."[12] According to representatives of the advertising industry, cookies have become less effective and less efficient, partly because users can delete cookies, though apparently the industry has sought to mitigate this issue by developing "flash cookies" which, because they may be implanted in multiple directories, can "re-spawn" even where the user thinks that he or she has deleted them. Cookies are also becoming less favored due to the increasing use of mobile devices, which apparently cannot effectively or efficiently support cookies.

There is not yet, however, a single technology that has replaced cookies. According to a recent report in the *San Jose Mercury*, Apple assigns an "Advertising Identifier" to each iPhone. This apparently allows advertisers to collect information from apps or services that operate on that particular device. Other major players in the online world are trying to develop cookie alternatives, but for the time being cookies are still the dominant key to online tracking.[13]

D. Mobile Apps and Connectivity.

As noted, the decreasing usefulness of cookies is closely related to the increasing use of mobile applications. With or without cookies, mobile applications can provide incredibly valuable – and potentially quite sensitive – locational information in addition to many other forms of personal information. In addition, the Federal Trade Commission recently invited comments on privacy issues raised by the growing "connectivity" between multiple devices, mobile or otherwise, that may allow tracking across devices. The FTC acknowledges that these connected devices "can provide important benefits to consumers . . . [Yet] at the same time, the data collection and sharing that smart devices and greater connectivity enable pose privacy and security risks."[14]

The rise and potential fall of cookies, the still-developing alternatives to cookies, and the rapid pace of technological change all suggest that many difficulties lie ahead for both consumers and policymakers in attempting to safeguard informational privacy online or in the mobile ecosystem. For the most part, as discussed below, policymakers have aimed at least for greater "transparency" in the hope that by providing meaningful disclosures about how information is collected, used, and shared, consumers will be able to make reasonably informed and rational choices that correspond with their privacy comfort zone. Unfortunately it has become increasingly evident that despite policymakers' good intentions, the current "disclosure and consent" legislative approaches to privacy protection on the Internet may not be working nearly as effectively as had been initially hoped. Prior to reviewing the American approach to privacy protection, we shall first turn to the European approache to consumer privacy protection – at least to their aspirations.

PART TWO
THE PREVAILING DISCLOSURE-BASED APPROACHES

A. The European Union (EU) Approach: "Notice and Consent"

The EU Directive: Adopted in 1996, the EU Data Privacy Directive generally requires EU member states to enact data protection laws.[15] The EU Directive does not impose any specific requirements on entities that collect, share, and use data (known as "data controller" in EU parlance). Rather, the Directive is aimed at member states, which are expected to adopt legislation that implements the Directive's broad principles. The EU Directive requires member states to adopt legislation that permits the processing of personal data *only if* the "data subject" (i.e. the person to whom the data applies) has unambiguously given his consent, or one of five other instances, including the processing is necessary for the purposes of the legitimate interests pursued by the controller or by the third party or parties to whom the data are disclosed, except where such interests are overridden by the fundamental interests or rights of the data subject.

Although the EU Directive is often described as prohibiting the processing of personal data without the data subject's "consent," obtaining consent is just one of six ways of legitimizing the processing of

personal data. Moreover, according to some scholars who have studied how member states have implemented the Directive, there is little consensus on how "consent" is obtained. The EU Directive also requires the data controller, upon request, to provide the data subject with information about the identity of the controller, the purposes for which data is being processed, and the recipients or categories of recipients of the data. Finally, subject to certain restrictions, the individual has the right to object to sharing of personal data for direct marketing purposes.

EU Proposed Regulation of 2012 and "Right to be Forgotten": Partly due to the ambiguity of the EU Directive and the fact that critical details of implementation are left to member states, the European Commission adopted a seemingly stronger, more binding, Proposed Regulation (Regulation) in January of 2012. At the time of this writing, it appears that this Regulation is still just a proposal. One of the more controversial elements in the new Regulation is the "right to be forgotten." The proposed Regulation would declare that personal data belongs to the data subject and not to the data controller or the data processor. If and when the proposed Regulation is adopted, the "right to be forgotten" will mean that the data subject will have the right to delete any personal data relating to the data subject and to prevent any further dissemination of that data. However, there are some exceptions to the "right to be forgotten."

B. The U.S. Approach: "Notice and Choice"

Compared to the EU, the U.S. approach – both at the federal and state level – is much more *ad hoc* and, on paper at least, much weaker than the EU approach. To begin with, neither the U.S. government nor any of the fifty states have adopted anything as comprehensive as the EU Directive.[16] Instead, the most substantive federal and state laws in the U.S. tend to target specific categories or uses of personal data, such as legislation regulating the use and disclosure of financial[17], medical,[18] or educational information.[19]

Outside of these specific special protections, a number of states – including California – have opted almost exclusively for "disclosure-based" approaches. Unlike the EU Directive, state laws in the U.S. have not, with a few exceptions, required data collectors to obtain a consumer's consent or even to allow consumers to obtain, correct or delete personal data. [20]

Thus, while the EU Directive is best described as a "notice *and consent*" approach, the U.S. policy is best described as merely a "notice" or "disclosure" approach. Indeed, although some commentators refer to the U.S. as having a "notice and choice" approach, it appears that the only "choice" is to not use the website or online service if one knows and then actually objects to the way that the site collects, shares, or uses information. And given the ubiquitous nature of the Internet in our lives today, any approach that simply offers users the "choice" of not using the Internet may not be a seen as a reasonable option.

FIPPs and the Current U.S. Emphasis on Industry Self-Regulation: Perhaps the most significant difference between the U.S. and EU approach is that the U.S. largely continues to rely primarily on industry self-regulation. While sector specific laws like the Health Insurance Portability and Accountability Act of 1996 (HIPAA) impose limitations and generally require consent before sharing or disclosing the information, in almost all other areas federal policy merely sets forth guiding principles – such as the Fair Information Practices Principles – for businesses to follow voluntarily rather than mandating specific requirements by statute.

The Fair Information Practices Principles, or FIPPs, are a widely accepted set of principles that formed the core of the Privacy Act of 1974 and have been the basis, more or less, of many other federal laws, including the Fair Credit Reporting Act, the Right to Financial Privacy Act, and the Children's Online Privacy and Protection Act.[21] Specifically, FIPPs fall into four general categories:[22]

- **Notice and Awareness (Transparency):** Individuals should receive notice of an entity's privacy practices – especially the type of data collected, how it is collected, and with whom it is shared – *prior to* the collection of personally identifiable information and be allowed to make informed choices regarding certain uses of their personal information, either through an "opt in" or "opt out" mechanism.
- **Access and Participation:** An individual must be able to view the data an entity has on record and be allowed to correct incomplete or false information in the entity's possession.
- **Integrity and Security:** Data must be accurate, up-to-date, complete, and not stored longer than necessary.
- **Enforcement and Redress:** An individual must be able to file complaints with the entity to have their issues addressed. There should be a mechanism to ensure compliance with the above standards, either through industry self-regulation or government regulation.

While FIPPs principles have informed a handful of federal statutes, they have never been codified as such. Instead, agencies like the Federal Trade Commission (FTC) have encouraged relevant industry associations to adopt FIPPs – or something similar to them – as a means of self-regulation. Recent proposals from the White House, discussed below, and the FTC set forth recommended privacy frameworks that closely track FIPPs.

The White House Consumer Data Privacy Framework: The White House Consumer Data Privacy Framework released by the Obama Administration in February 2012 offers a modest restatement of FIPPs, calls for industry self-regulation through the development of enforceable codes of conducts, seeks to strengthen the enforcement power of the FTC, and seeks cooperation with international entities.

The first part of the Privacy Framework is the Consumer Privacy Bill of Rights, which sets forth seven comprehensive principles that closely follow FIPPs. These principles are not enforceable rules; rather, they are intended to provide businesses with both "guidance" and the "flexibility" to implement the principles in a manner that is appropriate to their particular business model and the nature of the information that they collect, use, and share. Specifically, the guiding principles are:

- **Individual Control:** Consumers have a right to exercise control over what personal data companies collect from them and how they use it.
- **Transparency:** Consumers have a right to easily understandable and accessible information about privacy and security practices.
- **Respect for Context:** Consumers have a right to expect that companies will collect, use, and disclose personal data in ways that are consistent with the context in which consumers provide the data.
- **Security:** Consumers have a right to have their personal data securely handled and protected from unauthorized access.

- **Access and Accuracy:** Consumers have a right to access and correct their personal data, in a manner that is appropriate to the sensitivity of the data and the risk of adverse consequences if the data is inaccurate.
- **Focused Collection:** Consumers have a right to reasonable limits on the personal data that companies collect and retain. In other words, a company should not collect more information than it needs for legitimate business purposes.
- **Accountability:** Consumers have a right to know that companies will abide by stated policies and adhere to the Consumer Privacy Bill of Rights.

C. California's Disclosure-Based Approach

While Congress continues to struggle with political gridlock on the issue, and the White House Framework has set forth aspirational principles as a starting point, California has enacted two disclosure-based statutes: The California Online Privacy Protection Act (B&P Code Section 22575 *et seq*.) and the "Shine the Light" law (Civil Code Section 1798.83.) These statutes rely on a disclosure-based approach to privacy management.

California Online Privacy Protection Act. Cal OPPA (B&P Code Section 22575) requires the operator of a commercial website or online service that collects "personally identifiable information," as defined, to post a privacy policy on its website. According to the California Attorney General, the "online services" covered by Cal OPPA include applications for mobile devices.[23] The law does not impose any specific security requirements or restrict how data can be collected, used, or shared. Instead, Cal OPPA merely requires that a privacy policy be posted and that it disclose the general categories of personal information collected and the kinds of third parties with whom that information might be shared. As recently amended by AB 370 (Ch. 390, Stats. 2013), Cal OPPA will also require the privacy policy to tell users whether or not it honors a "Do Not Track" request made by the consumer's browser. Nothing requires a website to honor those requests; the privacy policy need only state whether or not the website honors Do Not Track requests.

Pending Legislation: AB 242: There appears to be a broad consensus by commentators and scholars that Internet-based privacy policies are not achieving the objective of greater consumer understanding about how a particular Internet site will use and transmit their personal information. Even the FTC, which seeks to ensure that companies live up to their privacy policies, has conceded that "the notice-and-choice model, as implemented, has led to long, incomprehensible privacy policies that consumers do not read, let alone understand."[24]

AB 242 (Chau), expected to be heard in early 2014 by the Assembly Judiciary Committee and then the Assembly Business, Professions and Consumer Protection Committee, seeks to address this problem – namely, that so-called privacy policies are not really about privacy but rather about how personal information is shared and sold. The measure in its current form seeks to address the problem that such policies are rarely read or understood by consumers before they click "accept" for the services provided by seeking to limit their length and require that they be written at a reasonable grade level (rather than in legalese) so as to be comprehensible to an average consumer.

As the Legislature proceeds with its consideration of AB 242, and considers how best to format privacy information in a way that is effective, useful and utilized by consumers, it is important to

keep in mind what Professor Helen Nissenbaum, a privacy expert at New York University, calls the "transparency paradox":

> Achieving transparency means conveying information-handling practices in ways that are relevant and meaningful to the choices individuals must make. If notice (in the form of a privacy policy) finely details every flow, condition, qualification, and exception, we know that is unlikely to be understood, let alone read. But summarizing practices in the style of, say, nutrition labels is no more helpful because it drains away important details . . . An abbreviated, plain-language policy would be quick and easy to read, but it is the hidden details that carry the significance. Thus the transparency paradox: transparency of textual meaning and transparency of practice conflict in all but rare instances. We seem unable to achieve one without giving up on the other, yet both are essential for notice-and-consent to work.[25]

Some industry representatives are aware of this conundrum and of the Legislature's concern that the current privacy policy regime may not be working as had been hoped, and are proactively convening experts from their organizations to try to develop new approaches, such as icons or other means, to better educate consumers about the use of their personal information.

"Shine the Light": While Cal OPPA requires every online business to post a "privacy policy," California's "Shine the Light" law requires any business that collects "personal information" (as defined) to respond to a user's request for what personal information about the user has been collected and with whom it has been shared, subject to certain exceptions. As with Cal OPPA, "Shine the Light" does not impose any restrictions on the collection and sharing of personal information; it simply requires any business subject to the law to provide the user, upon request, with information about the collection and sharing of personal information.

"Shine the Light," applies only if the third party uses the information for "marketing purposes." A business that shares personal information with a third party for any non-marketing purpose – without regard to risk of harm to the consumer – does *not* need to be disclosed under "Shine the Light." Moreover, a business is not required to honor a consumer request if it does not affirmatively know that the third party intends to use the information for marketing purposes, and nothing in the statute appears to require a business to inquire as to these third-party uses. In addition, a business is not required to respond to a "Shine the Light" request if it has posted a privacy policy that provides consumers with a cost-free means of opting out of having personal information collected.

According to one study, Americans visit, on average, about 100 websites per month. Each of these websites, in turn, may share information with multiple third parties, either directly or by allowing third-party cookies. The third parties, in turn, likely share that information with other entities. A consumer who wanted to know what might happen to information shared with a given website (or all 100 of them) would need to make requests of each third party with whom information was shared, and then again with the third parties of the third parties, *ad infinitum*. Perhaps not surprisingly, only a handful of researchers at U.C. Berkeley have apparently made a request under this statute.[26]

Pending Legislation: AB 1291: AB 1291, the "Right to Know Act," is currently being held in the Assembly Judiciary Committee for further study. It would repeal and entirely recast the existing Shine the Light law.

Recent Litigation Under California Unfair Competition Act and Consumer Legal Remedies Act: Just a few weeks ago, Apple won a summary judgment motion against plaintiffs in a privacy class action brought by purchasers of iPhones, iPads and iPods who alleged that those devices allowed for disclosure of their personal information, including address, current whereabouts, unique device identifier, gender, age, and zip code, through apps they downloaded. Plaintiffs claimed that they would not have paid so much for their Apple devices if they had known how their private information would be used. Apple countered that plaintiffs did not consider privacy issues when deciding what device to purchase. The case, *In re iPhone Application*, 11-2250 (N.D. Cal.), was recently dismissed by District Court Judge Lucy Koh, who ruled that plaintiffs could not prove that they had even seen Apple's privacy policy, let alone read or relied on it, and thus could not prove their case. Contrary to what some might argue, this demonstrates that the inaccessibility of privacy policies can potentially have direct, negative effects for consumers.

D. Private Sector Initiatives

The technology industry generally opposes public regulation of Internet privacy in favor of voluntary self-regulation by the industry with respect to commercial purposes. This approach has caused some observers to opine that the industry employs a double standard when it later argues against governmental collection and analysis of Internet user data, such as the recent revealed activity by the National Security Agency.[27]

Of course, companies that operate or develop online services and applications have, at their own initiative, taken a number of steps to protect consumer privacy. Some companies contend that they understand that privacy protections are in their best interests, as consumers will be reluctant to engage in online commerce if they genuinely fear identify theft, reputational harm, or an erosion of privacy. These private, industry self-regulation initiatives adopt a disclosure-based approach that is not altogether different than what has been proposed by legislators and government regulators. Three of these initiatives – by no means representing an exhaustive list – are briefly summarized below.

DAA Program: In 2009, several leading marketing and advertising industry associations developed the Digital Advertising Alliance (DAA) Program. At the core of the Program is a set of guidelines known as the "Self-Regulatory Principles for Online Behavioral Advertising (Principles)." The Principles are very similar to FIPPs and the proposed White House Consumer Privacy Bill of Rights. It offers an icon that participating businesses can display on their websites.

TRUSTe Privacy Seal: TRUSTe is a global "Data Privacy Management" (DMP) company that awards a TRUSTe Certified Privacy Seal to all businesses who successfully complete a Privacy Assessment and implement recommended changes. A company that adopts the recommended "best practices" can display the Privacy Seal on all certified websites, apps, and platforms. TRUSTe first conducts a "privacy assessment" of a business's existing privacy practices, recommends changes, and then offers the Seal to businesses that adopt those changes. TRUSTe also cooperates with DAA so that TRUSTe companies may also display the DAA icon.[28]

Mobile Application "Short Form Notices": The United States Department of Commerce recently convened a Multi-Stakeholder Process on Application Transparency to develop a voluntary short-

form privacy notice for mobile applications. The stakeholders included privacy groups and application developers, application platforms, and other entities within the mobile app "ecosystem." This process led to the development of a "Short Form Notice" and a "voluntary code of conduct" that is designed to promote transparency in the sale, promotion, and use of mobile applications. The short form notice describes the types of data collected, including biometrics, browser history, phone or text log, contacts, financial, health or medical information, location information, and information about user files that are stored on that device and that contain content, as well as the third parties with whom user-specific data is shared. A short form notice is *not* required if the data is de-identified and reasonable steps are taken to prevent the data from being re-identified.

PART THREE
CAN DISCLOSURE-BASED APPROACHES WORK IN A DATA AGGREGATED WORLD?
THE POSSIBLE LIMITS OF "PRIVACY SELF-MANAGEMENT"

Despite their differences, privacy protection approaches taken by the EU, the United States and several states start with the assumption that if consumers know enough about how a website, online service, or mobile application collects, shares, and uses their personal information, they will be empowered to make rational choices and thus "self-manage" the privacy of their information.

Some experts argue that such approaches are fundamentally flawed, largely because, they do not provide most consumers with sufficient information to make a truly rational or informed decision about how to manage their personal information online.[29] This is not necessarily because privacy policies and notices are too long and complicated for the average consumer (though as noted this appears to be very often the case). Rather, disclosure-based approaches arguably often fail, these experts contend, because no disclosure, no matter how clear and concise, may be able to adequately account for the multiple players involved in Internet marketing, the almost endless and unpredictable possibilities of "downstream" uses of information, or how people make real-time decisions in the online world. The ecosystem of Internet marketing and online tracking is, they contend, simply too complicated and unpredictable to allow people to make truly rational choices, no matter how much information they are given at the point when the consumer visits a website or when information is initially collected. Particular problems these experts cite include:

The Problem of Defining "Personally Identifiable Information": "Personally identifiable information" (PII) is the most central concept in existing privacy legislation at the international, national and state levels in that it triggers whatever requirements the legislation sets forth. However there is as of yet no clearly agreed-upon definition of that term. Even in California, statutes do not define the term consistently. More to the point, experts are concerned that information is usually defined as either PII or non-PII *at the point of collection*, which fails to take into account whether "downstream" data collectors, brokers, and aggregators have the ability to combine disparate pieces of non-PII and link them to a specific person.[30] Many argue that the EU definition of PII is too broad and "expansionist" in that it includes any information that is *potentially* identifiable, which may be unlimited. The U.S. approach, on the other hand, is often criticized as too narrow and "reductionist" because it fails to take adequate account that non-PII can be combined to create PII.

The Problem of "Re-Identification" and "Downstream" Uses: The problem of defining PII is made more difficult by, and is closely related to, the problem of "re-identification" and "downstream" uses of personal information. UCLA law professor Paul Ohm succinctly states the problem:

> Reidentification science disrupts the privacy policy landscape by undermining the faith we have placed in anonymization. This is no small faith, for technologists rely on it to justify sharing data indiscriminately and storing data perpetually, while promising users (and the world) that they are protecting privacy. Advances in reidentification expose these promises as too often illusory.[31]

Prof. Ohm contends that it is precisely these problems of "reidentification" and "downstream" uses that render notice and consent approaches ineffective and make the definition of "personally identifiable information" so problematic. While the initial collector of the data can make all required disclosures, the notice will not be particularly useful to the consumer if the unwanted use occurs somewhere further downstream and the downstream collectors have the capacity to re-identify what was once de-identified or anonymous data.

PART FOUR
WHERE SHOULD POLICY-MAKERS FOCUS THEIR REVIEW REGARDING INTERNET PRIVACY PROTECTION?

Many scholars who have criticized disclosure-based "notice and consent" approaches do not necessarily think that they should be abandoned. Rather, they contend that such approaches are, by themselves, inadequate because they simply create a false sense of security.[32] The limitations of the existing online privacy protections may require a step back to address broader underlying questions concerning privacy and the Internet, and the possible alternatives to privacy self-management. These include:

Collection, Use, Sharing, and Tracking of PII:

1. Should commercial websites and online services, including mobile application developers and platform providers, be prohibited by law from collecting and sharing a person's personally identifiable information without the affirmative opt-in consent of that person?

2. Can consumers currently choose to opt-out of having their data collected? Is permitting data collection usually a condition of using a website or online service? Can website operators or online services be required to offer and respect an opt-out?

3. To what extent, if at all, does existing technology permit a user of a website or online service to block the collection and/or sharing of personal information?

4. What is the current status of "Do Not Track" (DNT) mechanisms? How many browser services offer such a mechanism? Are consumers aware of these mechanisms? How user-friendly are they? If such mechanisms are widely and readily available, can websites and online services be required by law to honor a consumer's DNT request?

5. Other than an objection in theory or principle to personal tracking and data collection what if any harms arise from these practices? Do these harm outweigh the benefits?

6. Would laws that limit or prohibit the collecting, sharing, or tracking of personal information reduce the number of free services that are available online?

7. How should the state seek to regulate, if at all, the activities of so-called "data brokers" who aggregate and resell information from a variety of sources? And is it even possible to effectively define who these entities are, let alone properly regulate their activities to ensure proper privacy protocols and consumer protections are in place in this industry?

Privacy Policies:

1. Do existing privacy policies provide consumers with adequate information, in a reasonably comprehensible fashion?

2. Would legislation requiring that privacy policies be clearly written make privacy policies more effective and useful to consumers, who must now navigate sometimes long and legalistic privacy policies?

3. Should privacy policies be required to be more explicit in disclosing the kinds of information that will be collected and identify the specific parties with whom information is shared?

4. Given the multiplicity of players in the Mobile App ecosystem, who should be responsible for complying with existing privacy policy requirements – for example, the app developer or the app platform? What responsibility should the consumer bear?

Who Should Make The Rules?

1. Who should set the rules? Should private industry be encouraged to engage in more self-regulation by the adoption of "best practices," or should these best practices be codified in law to ensure that all businesses, not just the responsible ones, engage in best practices?

2. Given the fact that the Internet does not respect political boundaries – as websites accessed in one state may be owned and operated by a business in another state or even another country – should legislation solely come from the federal government, or do the states have legitimate roles to play in setting policies that work best for them?

3. Should the Legislature make an effort to harmonize its own privacy statutes with President Obama's Consumer Privacy Bill of Rights? What is the proper state role in a framework focused on a federal solution?

4. Should California legislate "out in front" of federal or even international entities, or should the state generally strive to focus on harmonizing with existing federal or international frameworks?

5. How much do Californians really care about online privacy, and how much do they want workable online privacy protections? And can (and will) consumers effectively self-manage online privacy protections?

CONCLUSION
BALANCING PRIVACY AND OPPORTUNITY

As noted throughout this paper, the Internet, and the remarkable technology sector that has developed it, has revolutionized our lives and our methods of conducting business, seeking healthcare, and interacting socially– indeed, nearly every aspect of our existence. Even the most ardent privacy advocates acknowledge that the online collection, sharing, and use of personal data that has fueled the growth of the Internet has produced countless beneficial innovations with more certainly to come. [33]

Because of the way in which the Internet evolved as a largely "free" access point for nearly anyone with a computer and phone line, the collection and sharing of our personal information has to a large extent been the generally unnoticed "price" we pay for those online services.[34] Some fear that if enough consumers refused to consent to the use of their data, the existing Internet business model would be greatly undermined.

Thus balancing the opportunities and risks of "Big Data" requires a candid and thoughtful assessment of what appears to be gained and lost when personal information is collected, shared, and used on the Internet, and how policy-makers, technology leaders and privacy advocates can most effectively collaborate to protect against risks to personal privacy without unduly impinging on technological innovation that has become so central to California's economic and social vitality.

---

[1] *The New Yorker*, July 5, 1993. Cartoon reprinted with permission from Conde Nast.
[2] Apple Inc. v Superior Court of Los Angeles (Krescent) S199384 (February 04, 2013).
[3] *Wall Street Journal* series can be accessed at http://online.wsj.com/public/page/what-they-know-digital-privacy.html
[4] Nicholas Carr, "Tracking is an Assault on Liberty, with Real Dangers," *Wall Street Journal*, August 6, 2010.
[5] Jim Harper, "It's Modern Trade: Web Users Get as Much as They Give," *Wall Street Journal*, August 6, 2010.
[6] "Browser Makers Considering Limits on Tracking Web Users," *Sacramento Bee*, Friday, March 15, 2013.
[7] See for example, Kenneth Cuker, "Data, Data, Everywhere," *The Economist*, Feb. 27, 2010; and Omer Tene and Jules Polonetsky, "To Track or 'Do not Track': Advancing Transparency and Individual Control in Online Behavioral Advertising," 13 *Minnesota J.L. Sci. & Tech* 281 (Winter 2012).
[8] Unless otherwise noted, this description of "cookies" is drawn the definition provided on a variety of Internet privacy policies and the concise summary provided in Polonetsky and Tene, *supra* note 7, at 289-295.
[9] Google Privacy Policy, accessed at
[10] *Ibid.*
[11] See for example Jonathan Mayer and Jon Mitchell, "Third-Party Web Tracking: Policy and Technology," Stanford Cyber Law Center Paper, available at http://cyberlaw.stanford.edu/files/publication/files/trackingsurvey12.pdf
[12] "Google, Apple and Other Tech Giants Look to a Post-Cookie Era," *San Jose Mercury News*, October 28, 2013.
[13] *Ibid.*
[14] FTC Press Release on request for comment at http://www.ftc.gov/news-events/press-releases/2013/04/ftc-seeks-input-privacy-and-security-implications-internet-things
[15] Unless otherwise indicated, the discussion of the EU "Directive" and the subsequent EU "Regulation" is drawn from summaries contained on the EU's Data Protection website, which can be found at http://ec.europa.eu/justice/data-protection/index_en.htm. See also Paul M. Schwartz, "The EU-U.S. Privacy Collision: A Turn to Institutions and Procedures," 126 *Harvard Law Review* 1966 (2013).
[16] Schwartz, *supra* note 15, *passim.*
[17] Gramm-Leach-Bliley prohibits financial institutions from disclosing a customer's personal information, except as specified. (15 USC 6802 et seq.) See also the California Financial Information Privacy Act, California Financial Code Sections 4502-4503. GLB allows states to enact more stringent protections. (15 USC Section 6807)

[18] Health Insurance Portability and Accountability Act (HIPAA) 42 USC Section 1320d et seq. See also California Confidentiality of Medical Information Act (CMIA), Civil Code Section 56.10 et seq.
[19] Family Educational Rights and Privacy Act (FERPA), 20 USC Section 1232g. California law also addresses confidentiality of student records, California Education Code Section 49060 et seq.
[20] The exceptions include statutes regulating credit reporting (which allow a consumer to correct misinformation) and California's "Shine the Light" law (which, subject to certain exceptions, requires a business that collects personal information from a customer to, upon request, provide the consumer with (1) a list of the categories of information collected and disclosed to any third party for the third party's marketing purposes and (2) the names and addresses of the third parties that received the information.
[21] According to the Federal Trade Commission website, the FIPPs "were first articulated in a comprehensive manner in the United States Department of Health, Education and Welfare's seminal 1973 report entitled *Records, Computers and the Rights of Citizens* (1973) [hereinafter "HEW Report"]. In the twenty-five years that have elapsed since the HEW Report, a canon of fair information practice principles has been developed by a variety of governmental and inter-governmental agencies. In addition to the HEW Report, the major reports setting forth the core fair information practice principles are: The Privacy Protection Study Commission, *Personal Privacy in an Information Society* (1977) [hereinafter "*Privacy Protection Study*"]; Organization for Economic Cooperation and Development, *OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data* (1980) [hereinafter "*OECD Guidelines*"]; Information Infrastructure Task Force, Information Policy Committee, Privacy Working Group, *Privacy and the National Information Infrastructure: Principles for Providing and Using Personal Information* (1995) [hereinafter "*IITF Report*"]; U.S. Dept. of Commerce, *Privacy and the NII: Safeguarding Telecommunications-Related Personal Information* (1995) [hereinafter "*Commerce Report*"]; *The European Union Directive on the Protection of Personal Data* (1995) [hereinafter "*EU Directive*"]; and the Canadian Standards Association, *Model Code for the Protection of Personal Information: A National Standard of Canada* (1996) [hereinafter "*CSA Model Code*"]. Other sources relied upon herein include the *FTC Staff Report* and *FTC Report to Congress/Reference Services.*

[22] In addition to the four categories below, commentators also divide the principles into either "procedural" or "substantive" principles. According to one report, "Procedural principles address how personal information is collected and used by governing the methods by which data collectors and data providers interact. These principles ensure that consumers have notice of, and consent to, an entity's information practices. Substantive principles, by contrast, impose substantive limitations on the collection and use of personal information, regardless of consumer consent, by requiring that only certain information be collected and that such information only be used in certain ways." Most of the principles are clearly procedural in nature. One substantive principle widely adopted by the fair information practice codes is the collection limitation principle, which states that entities should only collect personal information necessary for a legitimate business purpose. *See Privacy Protection Study* at 513-15; *IITF Report* Section II.A.)

[23] In addition to notifying developers and platform providers that Cal OPPA applies to mobile applications, California Attorney General and leading operators of mobile application platforms agreed to in a Joint Statement of Principles in 2012. That statement clarified that "where applicable law so requires, an application ('app') that collects personal data from a user must conspicuously post a privacy policy or other statement describing the app's privacy practices that provides clear and complete information regarding how personal data is collected, used and shared." Signatories included Amazon, Apple, Google, Hewlett-Packard, Microsoft and Research In Motion.

[24] Federal Trade Commission Report, *Protecting Consumer Privacy in an Era of Rapid Change*, March 2012.
[25] Helen Nissenbaum, "A Contextual Approach to Privacy Online," *Daedalus* (Fall 2011).
[26] Lauren Thomas and Chris Hoofnagle, "Exploring Information Sharing through California's 'Shine the Light' Law (August 13, 2009). Available at SSRN: http://ssrn.com/abstract=1448365 or http://dx.doi.org/10.2139/ssrn.1448365
[27] *E.g.*, It's Time to Regulate Consumer Privacy, D. Lazarus, *Los Angeles Times*, Dec. 9, 2013)("AOL, Apple, Facebook, Google, LinkedIn, Microsoft, Twitter and Yahoo said in an open letter to President Obama that in light of recent revelations about the National Security Agency snooping on websites and communications networks, there's an "urgent need to reform government surveillance practices worldwide. That's a pretty bold claim to the moral high ground considering that each of these companies routinely mines customer data for their own purposes (read: profit).")
[28] Information obtained from http//info.truste.com
[29] *See, e.g.*, Solove, "Privacy Self-Management and the Consent Dilemma," *Harvard Law Review*, May 2013.]
[30] Paul Schwarz and Daniel Solove, "The PII Problem: Privacy and a New Concept of Personally Identifiable Information," *NYU L Rev* (2011).
[31] Paul Ohm, "Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization." 57 *UCLA L. Rev.* 1701, 1704 (2010).

[32] As Daniel Solove puts it: "Clinging more tightly to privacy self-management is not the answer. Nor is abandoning privacy self-management or embracing paternalistic regulation."
[33] Omer Tene and Jules Polonetsky, "Big Data for All: Privacy and User Control in the Age of Analytics," 11 *Northwestern Journal of Technology and Intellectual Property* 239 (2013)
[34] Jan Whittington and Chris Hoofnagle, "Unpacking Privacy's Price,´ 90 North Carolina Law Review 1327 (2012