Response to the Notice of Inquiry on Information Privacy and Innovation in the Internet Economy by the US Department of Commerce

The Department of Commerce's Internet Policy Task Force is conducting a review of the nexus between privacy policy and innovation in the Internet Economy. This document answers two questions posted in the Notice of Inquiry, based on 13 years of experience at the World Wide Web Consortium (W3C) related to privacy on the Web.

I. Introduction

W3C is an international community where Member organizations, a full-time staff, and the public work together to develop Web standards. Led by Web inventor Tim Berners-Lee and CEO Jeffrey Jaffe, W3C's mission is to lead the Web to its full potential.

W3C efforts related to privacy on the Web began in 1997, when development started on the widely known Platform for Privacy Preferences (P3P), published as a Web Standard in 2002.

The W3C staff have been part of the broader privacy conversation throughout the last decade, and have participated in many different research projects on Privacy in the United States and in Europe, including the Transparent Accountable Datamining Initiative, Policy Aware Web, Theory and Practice of Accountable Systems, PRIME and PrimeLife. One important vehicle for making connections from research work to other work is the W3C Policy Languages Interest Group (PLING), which also helps to bridge communities fragmented around policy languages and access control. Findings from the research influenced the work carried out in other W3C Working Groups, but not to the extend we had hoped for.

The role of the standards W3C builds is increasingly broad: W3C is no longer tied to the document mindset of the early Web; instead, we build the standardized underpinnings for what looks increasingly like a Web operating system: General purpose data formats, general purpose communications frameworks, general purpose APIs that make device features accessible to the Web that had previously been outside the sandbox.

As we build and design advanced APIs that permit access to risky features, topics like the transparency of the data collection itself, limiting the scope of user errors, or the user's ability to recover from erroneously granted consent take center stage. These factors at times influence the design of APIs (does the user pass a selection of cards from his address book to a web site, or

does he grant the web site access to the address book). At other times, all we might be able to do in specifications is to sketch basic requirements, as the distinction between a privacy friendly and a dangerous implementation may be entirely dependent on the details of user interfaces and interactions, beyond the scope of what can be reasonably specified.

II. Notice, choice, use-based models & accountability systems

The Notice of Inquiry puts forward questions about **use-based privacy protection models** (including accountability systems) to overcome limitations of the **notice & choice model**.

Answer

II.a Existing technologies

For the comparison of these two models, it is useful identify the following phases in the "life-cycle" of personal information:

- collection
- primary use ("the primary reason personal data was requested")
- secondary ("opportunistic") use
- deletion

Notice & choice approaches involve the collection step. Use-based approaches limit primary and opportunistic uses of personal data by emphasizing technologies and promises that address the "back end" of commercial or benevolent endeavors.

While P3P was initially designed to help the notice & choice model by giving clear information to the user, it was later used to enable back end systems and middleware to help manage the promises made to consumers or business partners. This can be seen as a first attempt to provide technical support for use-based approaches. Since 2002, this notion has been pursued again and again in research:

- In the PRIME project, the notion of Sticky Policy appeared and was shown to work.
- In the TAMI project, researchers showed that manual and automatic re-use of personal information can be efficiently monitored and audited to determine whether such uses were reasonable and within the pre-defined boundaries.
- In the PrimeLife project, participants explored the efficient downstream data usage control if personal data is handed on to third parties.

The technologies discussed above primarily address privacy needs under the assumption that all parties are acting in good faith. Even without considering enforcement in the presence of

malicious actors, the technologies are often seen as complex, costly and expensive to implement. Data models have to be adapted, new business processes have to be designed, staff has to be trained. An investment of such order of magnitude is a challenge and has to be backed by potential benefits.

But while implementing these technologies may be expensive, that cost must be weighed against the cost (both risks and actual implementation cost) of doing nothing. A fair observation of the last ten years suggests that doing nothing has often won in this weighing. Doing nothing costs nothing to implement, has only a moderate impact on driving customers away, and does not constrain further opportunistic use by freely given commitments. The risk of non-compliance with regulations is often mitigated by weak enforcement of that regulation: Even given the strong European regulation implemented by data commissioners, the German statistics show 2.2 inspectors per 100.000 companies which results in an average control every 39,000 years. Given this low risk, investment in data usage control is improbable or will be cosmetic at best. We will see more incomprehensible statements on notice & choice augmented by further complicated statements on use-based permissions.

The scale of the Web is such that only scalable solutions that involve all the participating actors (commerce, consumers, and intermediaries) will work. This will need to be supported by the underlying technology.

Unless some economic incentives are given by the legal framework to invite companies to use the existing technology for privacy aware data management, intelligent data warehouses, and data mining technologies that take into account privacy, substantive change in current business practices is unlikely.

The current economic and legal environment has not provided incentives that would lead to the deployment of privacy enhancing, use-based technologies. As both the technological and legal framework is developed further, further research into the economics of personal information online will be crucial to achieve meaningful privacy. **Therefore, we encourage the DoC to push for further interdisciplinary research on Internet, economy and its relation to privacy in order to find means to encourage deployment of privacy enhancing technologies and have a greater buy-in from the commercial world.**

II.b Challenges to technology enhanced notice, choice and use-based limitations

If controlled by technology, use-based privacy restriction may hinder creative opportunistic re-use of data collections. Many of the inventions on the Internet of the past ten years were made based on creative re-use of existing information. A system — legal or technological — that constrains that creative re-use tends to put brakes on innovation. The challenge, therefore, is to balance privacy values and allowance for data re-use in a fair and reasonable way. In using personal data, how does society promote creativity while remaining responsive to privacy rights and expectations? Neither method, nor content of such reconciliation of interests are on the

horizon. More creativity and research is needed to find new approaches that respect the human right of privacy without blocking to road to more innovation based on data re-use and personalization. While basic rules for privacy may cut off the most blatant abuses, we have to remain careful not to stifle innovation.

III. Usability and code as a new focus of action in the area of privacy

What is the state of development of technologies and business methods aimed at: (1) Improving companies' ability to monitor and audit their compliance with their privacy policy and expressed user preferences; (2) using text analysis or similar technologies to provide privacy notices; and (3) enabling anonymized browsing, communication and authentication? Please describe any other ongoing efforts to develop privacy-enhancing technologies or processes of which the Commerce Department should be aware.

Answer

III.a Investment into user interface research is needed

Research and development of privacy enhanced data management technology is well under way. But progress in deployed technology is slow. The policy languages used so far are still too complex for mainstream consumption. They must be simplified to enable adoption. Given the complexity of the notion of Privacy, simple technology is hard. Therefore, more research is needed.

The theoretical background for descriptive policy languages and associated technologies is by now well-established. We are seeing new ideas emerge, such as splitting data base tables to control the access to knowledge. This can, e.g., be used in cloud scenarios to control privacy and secrecy. The foundations are in place (but not deployed yet) for the privacy-friendly exchange of data, tied to purposes, annotated with notification obligations, and access to one's data. Subject access API standardization may become a hot topic in the future.

Legal requirements on privacy-friendly behavior are high in some parts of the world, and expectations are close behind. But technology does not let us meet those expectations fully, especially within the European context. The nature of requirements tends to make already-complex privacy-enhancing technology even more complicated. Businesses do not want that complexity. They fear that privacy-enhancing technology might in fact drive away customers, instead of building their trust and attracting them.

"Identity" on the Internet and on the Web has been the subject of constant research and development for years. Solving the identity problem on the Web is seen as a major condition for new innovative services. The Internet identity system that takes off promises to bring profits magnified by the economic network effects for those who have pioneered it. Therefore, we observe fierce competition around the notion of identity, with numerous competing technologies

and companies aiming at wide deployment. Complex privacy-enhancing technologies will decrease the chances of wide deployment, and are therefore not found in widely-deployed identity systems.

Users might complain loudly at times, but ultimately use even privacy-unfriendly systems that enable desirable services. As long as there is no "giant oil spill," why should one really opt-out of disruptively useful innovations? There remains, thus, public unease without a compelling technological alternative.

Some services, such as social networking, rely on the users' sharing of personal data and profiles for their business models. Not surprisingly, privacy controls that once existed have eroded over the years. Where such steps crossed out of the public's comfort zone, the outcry was strong enough that "doing nothing" about privacy isn't an option any more. However, the actions taken in response show the tremendous difficulties to create useful user interfaces to privacy controls. Put to the point, privacy user interactions are currently developed through trial and error, where errors are detected through public outcry. How can businesses be encouraged to make a sustained investment into privacy technologies and research?

Experience with P3P and its deployment demonstrates some of the obstacles: P3P was a short and simple specification. Nevertheless, businesses were often reluctant to make simple, machine-readable declarations about their data usage. Mostly they preferred vague statements in human readable privacy policies written by lawyers for some perceived additional liability protection. The lesson is that fear of liability may well drive businesses away from the sort of clear and succinct statements that would convey a clear message to the user, whether or not those messages are mediated through a user agent that evaluates a machine-readable privacy policy. Regulation has to take into account the tension between usable privacy experiences on the one hand, and the fear of legal liability on the other hand.

Another element of the P3P experience may shed light on why PETs have not reached a significant market deployment. Considerable investment into the deployment of the P3P led to some modest success on the server side — data suggest that at one time, 28% of the top 1000 sites were using P3P. But P3P relies on efforts from both content providers (who put machine-readable statements on their site) and browsers (which mediate the user's experience based on some processing of privacy policies, e.g. by matching them to preferences). In the case of P3P, major browser vendors did not adopt the technology. Apart from the privacy bird plugin (first AT&T research, then Carnegie Mellon), there was only rudimentary support for P3P in user agents. Browsers mainly combined the compact policy format with a rudimentary user interface, thus increased the fuzziness of the privacy statement resulting in an increase of fear of liability. The resulting privacy messages were barely understandable for the average user.

The incentives for browser implementers are complex: beyond just implementing a policy protocol, they need to work out a meaningful user interaction with the policy, and they have an

interest in popular online services being usable and simple when accessed through their software. User interactions that were tried include:

- writing preferences and warning users when those aren't met – but users won't write preferences
- informing users of a human-readable form of the privacy policy – but users won't read policies
- making canned sets of preferences available for users to choose from – but users won't change defaults

At the same time, implementers will face the pressure to not make interactions appear "scary", even if they involve personal data. As a result, we see a landscape in which client implementations have turned away from implementing policy protocols. Instead, they focus on blacklisting known criminal players, and punting privacy decisions to individual web sites. There has been little further investment in the design of privacy policy related user interfaces in client software.

Sustainable online commerce requires sustained trust by users in their online experiences. A key piece of trust online is confidence that privacy expectations are met. Even when the provider acts in good faith, a consumer who does not understand the provider's effort, will not gain more trust, and might very well walk away. User trust requires user understanding. Privacy-related interactions need to be simple and understandable to everyday users. Unfortunately, today's interfaces tend to display large complex statements or technical jargon that nobody understands, if they say anything about privacy at all. Such incomprehensible messages neither improve privacy, nor increase the trust and confidence required for online transactions.

At this point, research into privacy user interfaces and experiences lags far behind user needs. Research investment is needed into simple, understandable user interfaces and experiences. While research in complex cryptographic primitives can lead to powerful technological enablers, development and deployment of simple user experiences are crucial in order to achieve practical privacy. We have revolutionized interfaces on mobile devices that can change direction if flipped, glow if poked, and so on, but that cannot answer important questions like "who knows where I am?" or "how do I limit who knows where I am?" Once more, what economic incentives would improve this situation and drive innovation? More research investment is urgently needed to develop *simple and helpful* user interfaces and experiences for privacy management.

III.c. Regulation should allow for incremental improvements

From the research projects done for the European Commission, we know how hard it can be to make software that is able to fulfill certain requirements established by law and regulation. The laws were not made with the available technology -- and the evolution of its use in the future --

in mind, but rather in the spirit of describing a desirable end state, based on a given time's culture of technology use. The expectation is that technology will ultimately fall in place. This approach is described with the phrase of technology neutral regulation. While technology neutrality is an important principle, having requirements in law and regulation that are very difficult to achieve with technology today (or that become obsolete in the future) will undermine small, incremental improvements toward better privacy protection, as these won't improve compliance. It would be worthwhile to try an interdisciplinary approach and to confront lawmakers and regulators with technologists to determine what is easily achievable and sufficiently simple to be put into the market. A measure, to be effective, has to be able to address the Web's massive scale. Only simple but intelligent rules and technologies — taking into account the human part of the system — can cope with this requirement.

For questions about W3C or the answers here, please contact Rigo Wenning (rigo@w3.org)