PERSONAL DATA JOURNAL

Issue N°2 March, 2012

Feature

Understanding OAuth

by Markus Sabadello, Technical Editor

Selecting the OAuth ("Open Authorization") protocol as the topic for the second feature article of our Personal Data Journal is a logical choice for two reasons. Firstly, the vision of establishing an ecosystem around personal data is intrinsically linked to the topics of authorization and access control. Whether we are talking about giving individuals more privacy and more control over their personal data, or whether we are exploring new economic models to be built around it, the question of who can access what under which permissions and obligations is central to achieving them.

Secondly, OAuth appears to be one of the few basic common denominators (if not the only one) among the different companies and projects that are currently working to realize a user-centric Personal Data Ecosystem (PDE). While there exist intrinsically disparate approaches to expressing data models (XML, relational model, or semantic?) or to offering APIs (REST/JSON or SPARQL?), the use of OAuth for managing authorization to access personal data seems to have achieved complete consensus within the user-centric PDE world.

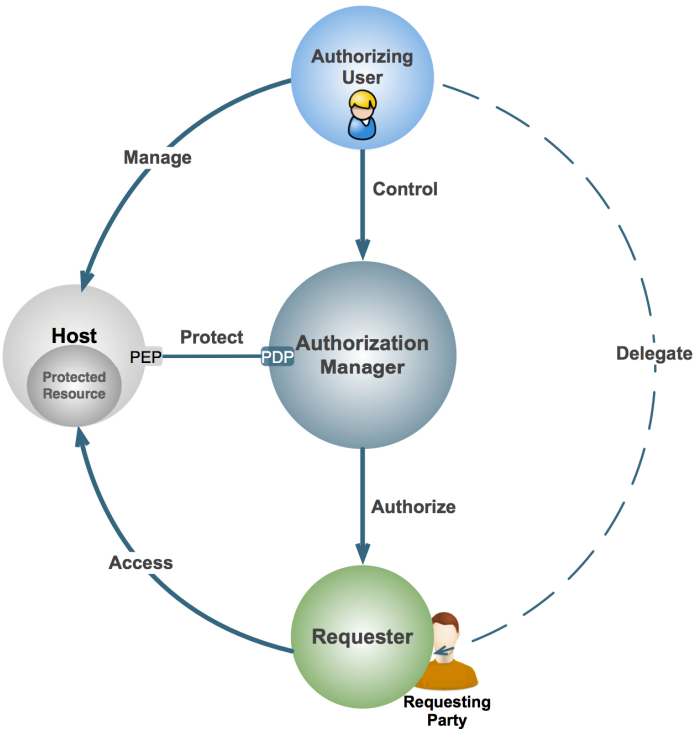


Figure 2: User-Managed Access (UMA)

Contents

This article will take a closer look at the evolution, inner workings and potential future application of the OAuth protocol, which has turned out to be one of the fastest growing, and overall most successful, new components in the open web architecture of the last few years. It will likely increase its importance within a wide variety of fields that are key to the PDE, such as user-centric identity, social networking, and vendor relationship management. Hence, it is essential to get it right.

AuthN/AuthZ on the Internet

Within communities that have been working on user-centric identity for years, it can often be heard that one of the Internet's biggest problems is that no identity layer has been built into it from the start that would support user authentication ("AuthN") and authorization ("AuthZ") in a universal way. While this may be true, certain specific mechanisms for these purposes have existed early on, for example various HTTP status codes such as "401 Unauthorized" or "403 Forbidden", or the HTTP Basic Authentication and HTTP Digest Authentication methods that make it possible to provide a username and password to a website when making a request. However, in a highly dynamic and interconnected Web 2.0 world, these approaches were not suitable for many use cases that emerged over time.

OAuth 1.0

The development of the original OAuth protocol began as a community effort in late 2006 by several companies out of a need to develop new models that could authorize access to their APIs. For example, proprietary technologies such as Flickr's API Auth, Yahoo's BBAuth, Google's AuthSub and Ma.gnolia's Dashboard Widgets all served as inspirations for this effort to develop an open authorization protocol. During this process, the Internet Identity Workshop has served as an important forum where much of the technical design and consensus–finding took place. By the end of 2007, the OAuth Core 1.0 specification was completed, and several years later, it was eventually published in 2010 as RFC 5849 within the Internet Engineering Task Force (IETF).

For anyone who is completely new to OAuth, the way it is usually explained is as follows: Imagine yourself as a user with an account at a photo hosting service A, where you uploaded all pictures from your last vacation. Now imagine furthermore that you would like to order high-quality hard copies of these pictures at a photo printing service B. In the worst case (if A and B were completely agnostic about each other), you would have to manually upload your photos again to B, even though you already uploaded them to A. (*Continued on Page 21)*

Personal Data Journal

Personal Data Journal is published by the Personal Data Ecosystem Consortium. Our goal is to create and support a diverse community of companies, small and large around the world building a thriving personal data ecosystem.

Personal Data Journal Staff

Publisher: Kaliya Hamlin
Associate Publisher: Patrick Reilly
Editor: Kelly Mackin
Technical Editor: Markus Sabadello
Researcher: Joseph Boyle

Sales

pat@pde.cc

Subscriptions:
Subscriptions to Personal Data Journal are available at http://personaldataecosystem.org/Frontier for $3,000 per year.

Personal Data Journal is published by **Personal Data Ecosystem Consortium** a non-profit 501(c)6 trade association.

Executive Director:
Kaliya Hamlin

Board Members:
Clay Shirky
Phillip J. Windley, Ph.D.
Tony Fish
Aldo Castaneda

http://personaldataecosystem.org

News

Address Book Theft Storm Engulfs Web

Twitter users around the globe can thank Arun Thampi of Singapore for noticing that his contacts were copied without his permission by a social network app called "Path." Technology bloggers, according to *TechTrend, discovered that iPhone apps like Facebook, Twitter, Foursquare and Foodspotting uploaded user data without permission in some cases*, said an article in the U.K.'s *Daily Mail*.

The CEO of Path, a startup focused on sharing your life with just close contacts, has apologized after a hacker showed that Path's iOS app uploaded extremely personal data to its servers without permission. CEO Dave Morin said the company is "sorry."

http://venturebeat.com/2012/02/08/path-sorry-about-that-whole-data-stealing-thing/

Doing personalization, finding friends etc. without data leaving your personal data store would avoid uploading and revealing data. Why reveal more data than needed? Any sharing should be parsimonious in scope. - Ed.

WEF Personal Data Meeting in Davos

Eighty C-level executives and invited experts from around the globe gathered at the World Economic forum in Davos. The main event was the reading of a description of the new EU Privacy law under consideration.

William Hoffman wrote, in an email to PDJ: *"Throughout a number of sessions at this year's Annual Meeting in Davos, there was a growing recognition of the tremendous opportunity for establishing a balanced personal data ecosystem. The task now is to gain consensus on the core principles for using and protecting personal data and identifying the pragmatic solutions to achieve these goals."*

They are planning on issuing a report and we'll cover that when it comes out. - Ed.

Just Forget Me on BBC

The BBC's report Do you have the right to be forgotten online *gives a really good overview of the new European Data Protection legislation, its aims to put people back in control of their personal data, and what it means for consumers. "Companies can't go foraging for data in the wild and pretend that what they find is theirs. The big idea at the heart of the new directive is that personal data is... personal'. - From Ctrl-Shift:*

http://www.ctrl-shift.co.uk/about_us/news/2012/02/15/bbc-report-on-new-eu-legislation/

Mexico's Data Protection Law

Over the next 16 months, Mexico will phase in its well-drafted but still basic privacy protections for "natural persons." According to the blog *Inside Privacy,* the law brings into force:

"the Law's provisions dealing with data subjects' rights to access, correct and delete personal information relating to them, which individuals have been able to exercise since January 2012. Failure to comply with individuals' requests to exercise these rights are actionable by the Federal Institute of Access to Information and Personal Data and may lead to civil penalties. The regulations also deal with security and breach notification, cloud computing, consent and notice requirements, as well as data transfers. "

http://www.insideprivacy.com/international/mexicos-data-protection-law-fully-in-force/

New York Papers Jump into Covering Personal Data Space

The New York Times and Wall Street Journal are charging into the Personal Data Space. The WSJ recently held a product design meeting for engineers and product managers. And The Times is increasing coverage on the Personal Data space in the last month. (see sidebar)

The ***Wall Street Journal*** *created an interesting interactive page covering "What they Know."* http://blogs.wsj.com/wtk-mobile/ *(It's a great "Personal Data for Dummies" and scary at the same time. -Ed.)*

Clouds in Google's Coffee

Google is facing more privacy problems in Europe in light of a recent decision in Norway to prohibit public sector organizations from using Google Apps due to concerns about where in the cloud the data were being stored, the conditions under

Photo credit: Ty Downing -Ed.

The New York Times Covers PD

Personal Data's Value? Facebook Is Set to Find Out
Personal Data's Value? Facebook Is Set to Find Out. Robert Galbraith/Reuters. An employee at Facebook's headquarters in Menlo Park, Calif. February 1, 2012 - By SOMINI SENGUPTA and EVELYN M. RUSLI

Should Personal Data Be Personal?
Personal data is the oil that greases the Internet. Each one of us sits on our own vast reserves. The data that we share every day — names, **...** February 5, 2012 - By SOMINI SENGUPTA

Start-Ups Aim to Help Users Put a Price on Their Personal Data ...
2 days ago **... Personal data** management has none of the obvious appeal of social networks or smartphones, but concerns about privacy may be changing **...** February 13, 2012 - By JOSHUA BRUSTEIN

What Story Does Your Personal Data Tell?
What rules, if any, do you think should govern how companies collect and use personal and aggregate data? February 7, 2012 - Learning Blog

Facebook Is Using You
Facebook's inventory consists of **personal data** — yours and mine. **...** News Analysis: Should**Personal Data** Be Personal? (February 5, 2012) **...**February 5, 2012 - By LORI ANDREWS

which they can be accessed, and by whom.

The decision arose from a test case in the town of Narvik, where a local council chose to use Google Apps for their email. Although Norway is part of the EU federation, its linkage is somewhat looser than some other countries. This latest case comes on the heels of a decision early last year where the town of Odense was prevented from using Google Apps in its schools for similar reasons.

In related news, Germany is now working on stricter data protection rules that would specify where geographically personal data could be stored. Cloud protection has now become a competitive issue for Google. France Telecom and Thales are intending to promote French cloud services over Google rivals.

The *Financial Times* Tech Hub reported that "unease is exacerbated by the "PATRIOT" Act, which requires US companies to hand data over to US authorities, *when asked, even if that data is stored in Europe."*

http://blogs.ft.com/fttechhub/2012/01/google-faces-norwegian-public-sector-ban/#axzz1mOkBs8BR

Do Consumers Need "Data Lockers" for their Personal Information?

Patricio Robles, Ecoconsultantcy

"On paper, this sounds appealing, but there are obvious challenges. Getting consumers to use these "data lockers" probably won't be easy. Besides the huge issue of trust, inputing data and managing who it's shared with could be quite time-consuming, confusing and inconvenient. And it won't necessarily guarantee that once that data is provided, it won't be used improperly. On the business side of the equation, it's not clear that businesses and consumers would see eye-to-eye on what their data is worth once the value is measured exclusively in dollars and cents.

But more practically, such services seem to ignore the fact that consumers are already engaged in exchanges of value around their personal data. … With this in mind, the introduction of a new set of middlemen to 'help' consumers manage and sell their personal data seems entirely unnecessary."

(We sigh when we see uninformed expertise. The monetization of personal information without end user management is the greatest land grab in human history. Information is power and when it is controlled by others, the outcome is unknown. -Ed.)

Do Consumers Need "Data Lockers" for their Personal Information?

Sites are Accused of Privacy Failings in Wall Street Journal

Getting personal information removed from websites that collect it can feel a lot like playing "Whac-a-Mole." "It's very difficult to get out and stay out," says Michael Fertik, chief executive of Reputation.com. CPO Jim Adler of people search website operator Intelius cited shared common names and name and address changes as difficulties in finding all data to remove.

Attorney Sarah Downey of privacy startup Abine said BeenVerified was the only people search site where records consistently popped back up, saying when she called to complain, a customer-service rep told her: "Oh, you should have asked them to remove you permanently."

Also in the article: the White House is expected to release a "Privacy Bill of Rights" this year with disclosure and opportunity-to-correct requirements similar to last year's Kerry-McCain bill. (*PDJ Covers this next issue -Ed.)*

FTC is increasing scrutiny of background-check providers' compliance with the Fair Credit Reporting Act, and commissioner Julie Brill asked data brokers to improve consumers' visibility of their data. Also busting directory and ad companies for misrepresentation.

Sites Are Accused of Privacy Failings

EC Proposal on Data Protection

On January 25 European commissioner Viviane Reding presented a proposal to control data protection in Europe more strictly. Major criticism is that two standards are proposed: one for businesses and citizens, and a weaker standard for law enforcement.

Businesses and citizens will have a binding regulation, which is much stricter and clearer than the current directive. The key elements: a single 'privacy' point of contact for multinationals, a reporting obligation in the event of data leaks, binding corporate rules and heavy fines for violations. This set of measures protects the individual in his relationship with commercial companies. He gets more control over what's happening with his private data.

Criticism focuses on the fact that the individual in their relationship with the government is at the mercy of local legislation. In some European countries, local legislation is much less fair.

The Dutch Data Protection Authority (CBP) calls it "a plan with a low level of ambition". Dutch liberal Sophie in 't Veld finds it unacceptable that the government receives an exception. "The decisions taken by the government are based on this information. These decisions often have a far larger effect on individuals than those of companies." The government will have to be addressed at least as stringently as businesses. 'The real snag: Data is only under the new European rules when it is stored within the EU. **Qiy** fully complies with new European rules, including the fact that data is stored regional."

EU Personal Data Protection Law:

* http://ec.europa.eu/justice/data-protection/document/review2012/com_2012_11_en.pdf
* http://ec.europa.eu/justice/data-protection/minisite/
* http://ec.europa.eu/justice/data-protection/document/review2012/factsheets/3_en.pdf

Chinese Bloggers Say 'No Thanks' to New Real Name Rules

The Chinese real names announcement has sent a shudder through the Chinese blogging community. In jurisdictions around the globe these kinds of activities are taking place. But only in environments like China is the full meaning of the loss of benign anonymity really understood as a threat to speech.

Chinese intellectuals quit microblogging in response to real-name registration requirement
http://the-diplomat.com/china-power/2012/01/31/intellectual-microblog-exodus/

Tribal Settlement Objectors see Personal Data Published

Attorneys who negotiated a $3.4 billion settlement over misspent Native American land royalties published the phone numbers and addresses of the four people objecting to the deal. The plaintiffs' attorneys wrote in their letter that the "hopes and wishes of 500,000 individual Indians" had been delayed by those four people. If it weren't for them, the first payments would have been made before Thanksgiving, the letter said.

http://azstarnet.com/news/national/tribal-settlement-objectors-see-personal-data-published/article_221da6cd-5e21-58fc-b085-a462c64f5148.html

EVENTS

Bolded events were not listed in previous issues.

Personal Digital Archiving

February 22–23, 2012
San Francisco, CA
http://www.personalarchiving.com/
PDEC staff are attending - highlights will be in the March Issue. - Ed.

O'Reilly's Strata Conference

February 28 - March 1, 2012
Santa Clara, CA
http://strataconf.com/strata2012
We will be covering highlights in the March Issue. -Ed.

OpenID Connect Interop

March 2, 2012
San Francisco, CA
http://groups.google.com/group/openid-connect-interop
On ongoing interoperability testing between implementations of OpenID Connect Implementers' Draft.

SXSW Interactive

March 9-13, 2012
Austin, TX
http://sxsw.com/interactive
Michael Schwartz CEO of Gluu, (a PDEC Startup Circle company based in Austin) is hosting Personal Data Ecosystem Community day on Saturday March 10th. Kaliya Hamlin will be at these events and if you or colleagues are attending this will be one of the best places to meet up with others in the field.

Gartner Identity and Access Management Summit

March 12–13, 2012
London, UK
http://www.gartner.com/technology/summits/emea/identity-access/
Kaliya is giving the keynote address.

NIST–IDtrust Workshop

March 13–14, 2012
Gaithersburg, MD
http://www.nist.gov/itl/csd/ct/nstic_idtrust-2012.cfm
"Technologies and Standards Enabling the Identity Ecosystem"
The workshop will focus on how technologies and standards can help the framework of the identity ecosystem coalesce.

NSTIC Governance RFP Meeting

March 15, 2012
Washington, D.C.
Link and Location TBD - see the publisher's note to learn more about NSTIC.

Digital Identity World Australia

March 20, 2012
Sydney, Australia
http://www.terrapinn.com/conference/digital-id-world-australia/index.stm

Structure: Data by GigaOM

March 21–22
New York
http://event.gigaom.com/structuredata/
Cost: $896
Lots on data analysis, no talks say they are from user-centric perspective, all are business or unstated perspective. But it is happening and it does affect Personal Data.

Trust Management Symposium

March 22–23, 2012
Botsdam, Germany
http://www.hpi.uni-potsdam.de/meinel/Trust_Symposium_2012
The symposium brings industry to meet academia to establish connections, discuss research and present problems in the industry that may have solutions in the academic world.

New Digital Economics Silicon Valley

March 27–28, 2012
San Francisco, CA
www.newdigitaleconomics.com/SiliconValley_2012/index.php

Data 2.0 Summit

April 3, 2012
San Francisco, CA
http://data2summit.com/
PDEC staff are attending - highlights will be in the April Issue. - Ed.

WSJ Hosted Data Transparency Weekend

April 13–15, 2012
New York
datatransparency.wsj.com/
Hackathon to develop tools.

Data Usage Management on the Web at WWW 2012

April 16, 2012
Lyon, France
dig.csail.mit.edu/2012/WWW-DUMW/
Data usage control generalizes access control in order to address what happens to data in the future and after it has been shared or accessed. Spanning the domains of privacy, the protection of intellectual property and compliance.

European Identity & Cloud Conference

April 17–20, 2012
Munich, Germany
www.id-conf.com/
This is the premier spring conference in Europe covering these issues.

Internet Identity Workshop

May 1–3, 2012
Mountain View, CA
www.internetidentityworkshop.com/
This is also PDEC's main convening opportunity and it is global in nature. Key European innovation and thought leaders in the space and they are planning to attend the event. We strongly encourage all those interested in making the ecosystem real attend.

IPSI SmartData International Symposium

May 14–16, 2012

Toronto, Ontario, Canada

www.ipsi.utoronto.ca/sdis/

This event was brought to our attention by Ann Cavokian the Privacy Commissioner of Ontario who has been leading the Privacy by Design movement. -Ed.

The future of privacy, and, in turn, our freedoms, may well depend on the ability of individuals to reclaim personal control of their information and identities online. In failing to protect personal data, the liberating potential of the Internet may be compromised, as various organizations (public or private sector, not to mention criminal interests) may use this free flow of information to exert control over, and potentially harm individuals. SmartData is a vision to create Internet-based virtual agents which will act as an individual's online proxy to securely store their personal information and disclose it based upon the context of the data request and instructions authorized by the data subject.

Web 2.0 Security and Privacy Workshop

May 24, 2012

San Francisco, CA

www.w2spconf.com/2012/

This workshop is co-located with the IEEE Symposium on Security and Privacy (below). The goal of this one-day workshop is to bring together researchers and practitioners from academia and industry to focus on understanding Web 2.0 security and privacy issues, and to establish new collaborations in these areas.

IEEE CS Security and Privacy Workshop

May 24-25

San Francisco, CA

http://www.ieee-security.org/TC/SPW2012

Conference on Web Privacy Measurement

May 31– June 1, 2012

Berkeley, CA

www.law.berkeley.edu/12633.htm

Hosted by the Berkeley Center for Law & Technology. Studying tracking technologies.

European e-Identity Management Conference

June 12-13, 2012

Paris, France

Cost: €220-€770

www.revolution1.plus.com/eema/index.htm

Business, public sector and government who are involved in policy, security, systems and processes.

Cloud Identity Summit

July 17-21, 2012

Keystone, Colorado (near Denver)

http://www.cloudidentitysummit.com

This event hosted by Ping Identity and lead by its CEO Andre Durand is unique for its high quality of presentations and attendees along with its family atmosphere. There were over 100 families in attendance - Andre's wife organizes a whole series of family activities in the day time and evening meals are with everyone together. The event leans towards an enterprise focus but will cover topics around identity and personal data.

OSCON (Open Source Convention)

July 17-21

Portland, Oregon

http://www.oscon.com/oscon2012

This O'Reilly event is the heart of the open source world and draws people from around the world. Open Standards are a key aspect of the event Federated Social Web get work done in F2F meetings during this event. There are several open source projects in PDEC I (Kaliya) expect they will present/be covered at this event.

New Digital Economics London

June 12-13, 2012

London, UK

www.newdigitaleconomics.com/EMEA_June2012/

(SOUPS) Symposium on Usable Privacy and Security

Date: July 12–13, 2012

Washington, D.C.

cups.cs.cmu.edu/soups/

Paper deadline March 9.

Cost: $100–$400

Chip-to-Cloud Security Forum

September 19–20, 2012

Nice, France

http://www.chip-to-cloud.com/

"From smart cards to trusted mobile and Internet of Things"

Abstract deadline March 23.

SIBOS

September 19–23, 2012

Osaka, Japan

http://www.sibos.com/osaka.page

€950/day, €2800/week

This is the annual gathering of SWIFT the international bank messaging cooperative. Kaliya has presented to them a number of times and they are proactively involved in understanding the way traditional banks and banking networks can play a role in the emerging ecosystem.

Event Overwhelm?

If you are a Frontier subscriber one of the perks you get is the the ability to request personal event recommendations from Kaliya Hamlin about which events will bring you the most value based on your learning and business goals.

Standards

W3C: WebID Community Group

A WebID Community Group has been created at the W3C. The group is a continuation of the WebID Incubator Group and it will continue development of a specification for the WebID protocol, build test suites, document use case, issues, and to grow the community of implementations. Joining the group requires registering a W3C account, as well as signing an intellectual property agreement.
http://www.w3.org/community/webid/

IETF: JOSE Drafts

Initial Internet Drafts of specifications from the IETF JOSE (Javascript Object Signing and Encryption) Working Group have been submitted. These specifications are:

- JSON Web Signature (JWS) – Digital signature/HMAC specification
  http://www.ietf.org/internet-drafts/draft-ietf-jose-json-web-signature-00.txt
- JSON Web Encryption (JWE) – Encryption specification
  http://www.ietf.org/internet-drafts/draft-ietf-jose-json-web-encryption-00.txt
- JSON Web Key (JWK) – Public key specification
  http://www.ietf.org/internet-drafts/draft-ietf-jose-json-web-key-00.txt
- JSON Web Algorithms (JWA) – Algorithms and identifiers specification
  http://www.ietf.org/internet-drafts/draft-ietf-jose-json-web-algorithms-00.txt

Although suitable for a wide variety of uses, these Javascript technologies for signing and encryption have been developed in the course of the OpenID Connect community's efforts, and will form an important component of the upcoming OpenID Connect specifications. Together, this set of technologies has also been called "The Emerging JSON-Based Identity Protocol Suite".

News about OpenID Connect

Nat Sakimura, one of the most active contributors to the OpenID Connect effort, has posted a simple introduction to how OpenID Connect works on the technology level, covering the topics of making an OpenID Connect request to a server, receiving a response, accessing user information, and discovering service endpoints. For anyone interested in learning how OpenID Connect will work (without having to read the actual specifications), this post is the perfect way to get started. In the emerging PDE, OpenID Connect may become an important component for expressing user identity, as well as for sharing and federating personal data.

http://nat.sakimura.org/2012/01/20/openid-connect-nutshell/

Also, Axel Nennker of Deutsche Telekom, who has recently been elected as a Community Board Member of the OpenID Foundation, has deployed a number of OpenID Connect test servers for use by implementers.

http://ignisvulpis.blogspot.com/2012/01/openid-connect-test-servers.html

The following paper explains how a profile of OpenID Connect can be used to build a decentralized claims architecture, in which different authorities can interoperate with each other.
http://www.aicit.org/IJIPM/ppl/007_IJIPM1-195IP.pdf

IETF: OAuth 2.0 Draft considered as Proposed Standard

A new Internet Draft (version 23) of the OAuth 2.0 Authorization Protocol has been submitted by the IETF Web Authorization Protocol Working Group. In addition, the IETF's Internet Engineering Steering Group (IESG) has received a request to consider this draft as a Proposed Standard. A decision is expected within the next few weeks.
http://www.ietf.org/internet-drafts/draft-ietf-oauth-v2-23.txt

IETF: The Atom "deleted-entry" Element

A new Internet Draft related to the Atom Syndication Format has been submitted to IETF. This draft adds mechanisms to Atom which publishers of Atom Feed and Entry documents can use to explicitly identify Atom entries that have been removed. This pattern of marking data items as deleted is also known as "tombstones". The Atom Syndication Format is an important component for Activity Streams, for the Federated Social Web, and potentially also for PDE–related projects that decide to use feeds for publishing and subscribing to personal data.

In light of numerous debates and legislation about a "right to be forgotten", this small addition to the well-established Atom standard is highly interesting, and can potentially even inspire other standards to adopt similar patterns for deleting personal data which an individual once shared and no longer wants to be visible.

http://www.ietf.org/internet-drafts/draft-snell-atompub-tombstones-14.txt

DMARC standard to coordinate authentication of emails

The largest email providers along with major financial, social media, and email security providers announced a working group developing a Domain-based Message Authentication, Reporting & Conformance (DMARC) standard incorporating the longstanding SPF and DKIM validation methods and adding policies and feedback (superseding ADSP) to enable real-time coordination against constantly mutating spamming and phishing. DMARC.org is planning IETF submission of the draft standard, and a discussion list is now available.

W3C Tracking Protection Working Group is working on design of HTTP-header based Do Not Track as suggested in last year's FTC report, with active discussion of a large number of issues. The group hopes to bring a draft to the final recommendation stage by June 2012.

IETF: New UMA Draft

A new Internet Draft for the User-Managed Access (UMA) Core Protocol has been submitted to IETF. UMA builds on OAuth 2.0 and may turn out to play a role within the emerging PDE for access control to personal data. It is already used by several members of the PDEC Startup Circle.
Title: User-Managed Access (UMA) Core Protocol

Author(s): Thomas Hardjono
Filename: draft-hardjono-oauth-umacore-03.txt
http://www.ietf.org/internet-drafts/draft-hardjono-oauth-umacore-03.txt

NSTIC Governance Recommendations

NIST (National Institute of Standards and Technology) issued a report entitled Recommendations for Establishing an Identity Ecosystem Governance Structure laying out a roadmap to establish a privately-led steering group for the complex policy and technical issues in creating the Identity Ecosystem envisioned in NSTIC (National Strategy for Trusted Identity in Cyberspace).

This follows the previous week's Federal Funding Opportunity for pilot programs for NSTIC.

NIST intends to issue a Federal Funding Opportunity (FFO) for an organization to convene the Identity Ecosystem Steering Group and provide it with ongoing secretarial, administrative and logistical support. This will enable the group to convene its first formal meeting late this spring.

OASIS: New Privacy Management Reference Model Draft

The OASIS Privacy Management Reference Model (PMRM) TC has published a new draft version (Working Draft 01) of their main deliverable. This TC works to provide a standards-based framework that will help business process engineers, IT analysts, architects, and developers implement privacy and security policies in their operations.

OASIS: New ID-Cloud Draft

The OASIS Identity in the Cloud (ID–Cloud) TC has published a new draft version (Working Draft 01a) of their gap analysis. This TC identifies gaps in existing identity management standards and investigates the need for profiles to achieve interoperability within current standards. This effort is highly relevant to PDEC, since it looks at a wide range of technologies and identifies differences between them. The output of this TC could be a valuable basis for working on an interoperable PDE.

twitter

@PersonalDataJ

Startup Circle News

Personal:

Part of US Department of Education Initiative to support student data export.

from Josh Galpers's, Chief Council of Personal' post:

http://blog.personal.com/2012/01/grade-a-idea-my-data-button-brings-ed-records-to-you/

When you hear the words "we're from the government, and we're here to help you," skeptics advise fleeing in the other direction.

However When, U.S. Chief Technology Officer Aneesh Chopra and U.S. Education Secretary Arne Duncan announced the "My Data Button" initiative last Thursday for freeing individual education records from the government and giving control over them to the individual, the expression rang true.

My Data Button would give individuals the power to access and import their federal education data for their own use. The rationale is simple: It's our data, and we should be able to have a copy to use however we want.

Three private sector companies stepped up to help make this announcement become a reality. As Chopra announced, Personal – in addition to Microsoft and Parchment– committed to offering services to help individuals upload, access and control this information. We are proud to join this effort.

Imagine having all of your education records in a Gem [Personal's term for its object types] housed in your own data vault, conveniently at your fingertips and ready for reuse in your private, personal network.

> "*New York Times* reporter Joshua Brustein provides a great introduction to the model that Personal and companies like us are developing. However, a central question remains unresolved: what is the true economic value of personal data?"
>
> **- Personal's Blog**

Links to the Administration's Fact Sheet.

Aneeh's the US CIO's Post about it.

Connect.me:

Private Beta & Respect Trust Framework

by Drummond Reed

Last week we passed 1/2 million vouches.

We've passed 15K active users, with a waiting list of another 60K (we can only let in so many new users every day until we have our API-based back end ready in late March).

In mid-January we introduced the Founding Trust Anchors and started a limited period of special Trust Anchor vouching. As of this writing we have 569 Trust Anchors, with over 1000 additional nominees. Our goal is to grow to 1000 Trust Anchors by March 1, when we will turn on full Trust Anchor vouching, where each Trust Anchor will have a lifetime limit of 150 active Trust Anchor vouches.

Qiy:

Coverage in Ernst & Young Magazine

http://www.ey.com/NL/nl/Services/Strategic-Growth-Markets/Exceptional-January-June-2012---Qiy

His vision was to "turn the information world upside down" by reversing the traditional relationship between individuals and organizations in the online sphere. "We have been conditioned to believe that individuals need to adapt to systems implemented by organizations," he explains. "Qiy's philosophy is that an organization's systems should adapt to the individual."

Marcel Bullinga Author of *Welcome to the Future of the Cloud* highlights Qiy's personal dashboard on Singularity Hub

http://singularityhub.com/2012/01/18/qa-with-dutch-futurist-marcel-bullinga-as-his-latest-book-looks-to-2025/

Which technology (or branch of science) do you feel will have the biggest impact in the next fifteen years? Who do you see as the leader in the development of that technology?

My pick: a small startup at www.qiy.com. It is the closest thing to my vision of a personal dashboard that I have discovered so far. I met the owner, Marcel van Galen, and he convinced me that in his business model the individual owner will stay in control. This will sweep aside the Google and Facebook attitude of "company owning". It is vital, by the way, that neither Google nor Facebook will ever buy Qiy.

Resources

WEF Report

Big Data, Big Impact: New Possibilities for International Development

www3.weforum.org/docs/WEF_TC_MFS_BigDataBigImpact_Briefing_2012.pdf

Given the flood of data generated by digital devices around the globe.

> *Researchers and policymakers are beginning to realise the potential for channelling these torrents of data into actionable information that can be used to identify needs, provide services, and predict and prevent crises for the benefit of low-income populations. Concerted action is needed by governments, development organizations, and companies to ensure that this data helps the individuals and communities who create it.*

The report highlights a few key areas where this data could make the most difference:

- Financial Services
- Education
- Health
- Agriculture

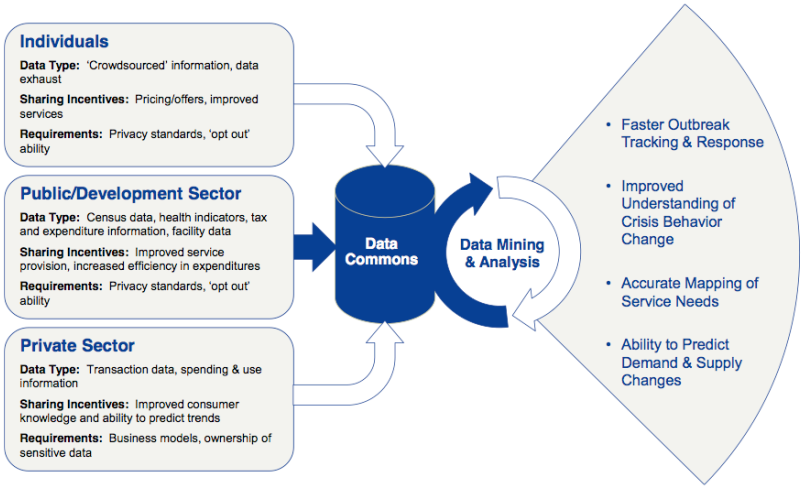
They name user-centric solutions offering compelling possibilities.

Data ecosystem dynamics are highlighted as a future focus to consider different types of data, actors and their incentives along with proposing the development of a data commons.

Obstacle that are named include:

- Privacy and Security
- Data Personalization
- Data Sharing Incentives
- Human Capital

Novel approaches to overcoming these obstacles are named such as "data philanthropy" and of course Governments have a catalytic role to play.



Boston Consulting Group: The Evolution of Online-User Data

by Ed Busby, Tawfik Hammoud, John Rose, and Ravi Prashad

https://www.bcgperspectives.com/content/articles/marketing_technology_evolution_of_online_user_data/

The gathering of online-user data is among the most exciting and controversial business issues of our time. It often brings up concerns about privacy, but it also presents extraordinary opportunities for personalized, one-to-one advertising.

The report highlights trends that are driving the supply side from advertisers and demand side from Advertiser each with a growing in building user profiles. It names trends that hamper grown including spending shifts to closed platforms like Facebook, concerns about accuracy, the proliferation of low-cost remnant inventory and a reluctance to share PII because of fears around regulation and public backlash. The article goes on to outline the different types of data that are collected and their view into the current marketplace for User Data with 6 distinct layers. They conclude with an their analysis of the implications for ecosystem companies. They mention a persona-data ecosystem but it does not outline a future with user-centric tools and systems.

Control-Shift Report on Privacy

Ctrl-Shift is at it again with a new report on privacy from an e-retailer perspective.

The report is free with registration: http://www.ctrl-shift.co.uk/shop/product/60

The data set is for sale. http://www.ctrl-shift.co.uk/shop/product/61

Ctrl-Shift scored the privacy policies of the IMRG Hitwise TopShop list of 100 online retailers against ten key questions including how clearly the privacy policy is written, how easy it is for the customers to express and change their preferences, whether their data is used for marketing purposes and how they treat cookies and behavioral targeting.

This 21 page report covers the following topics:

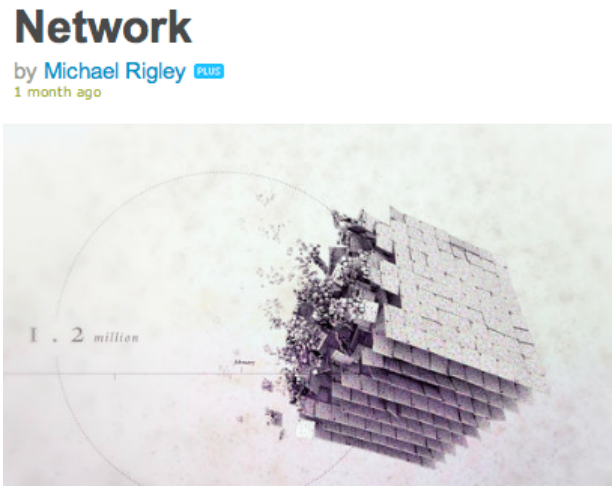
- Summary findings and conclusions
- Introduction: Why research privacy policies?
- The new data sharing relationship with customers
- The Results: Overview
- Survey results: The details
- How is the privacy statement written?
- Preference Management
- Options for receiving electronic communications
- Data sharing for marketing purposes
- Providing individuals with access to their data
- Cookies, pixel tracking and related mechanisms
- Behavioral targeting
- Data retention
- Policy changes
- The scope of the contract

VIDEO

Network

by Michael Rigley

http://vimeo.com/34750078



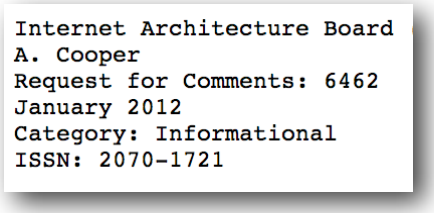
According to Michael Rigley, the average user has 736 pieces of personal data collected every day and service providers store this information for one to five years.

The video explores the "secret life of our MMS data and the tradeoffs we inadvertently face as we choose convenience of communication over privacy and control of personal data," writes Maria Popova at BrainPickings.org.

We recommend this video because it explains in plane english the types of data people generate as they use their mobile devices, what metadata is and how it is then used to make meaning from the data and what is crucial is that most people don't know about how long the information is stored and how it is used by the phone company.

Report on the Internet Privacy Workshop

Review by Markus Sabadello


Internet Architecture Board
A. Cooper
Request for Comments: 6462
January 2012
Category: Informational
ISSN: 2070-1721

Report text is at: www.rfc-editor.org/rfc/rfc6462.txt

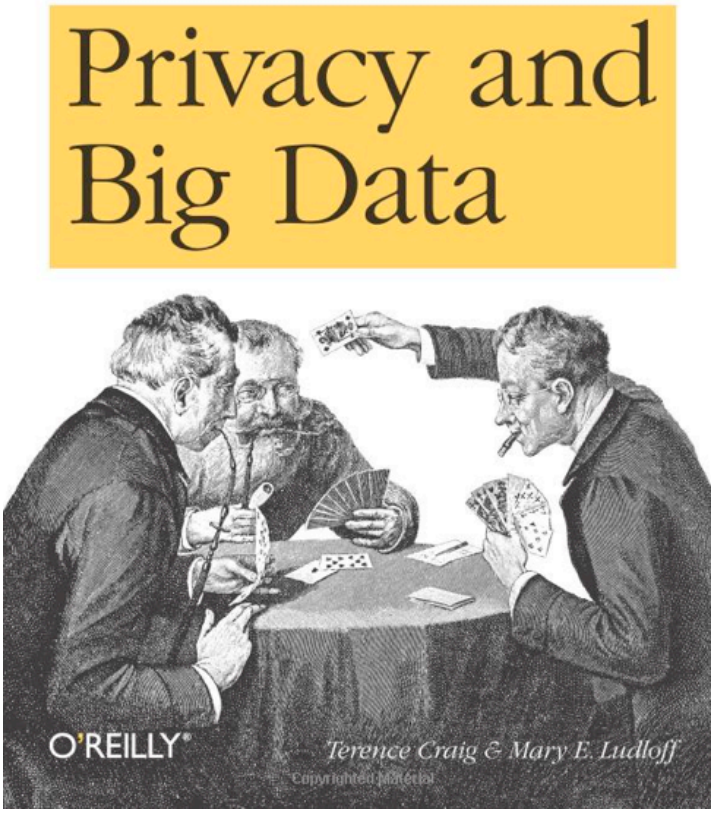
The report from the event was released at the end of January 2012 the event was December 8–9, 2010, the IETF's Internet Architecture Board (IAB) co-hosted an Internet privacy workshop with the World Wide Web Consortium (W3C), the Internet Society (ISOC), and MIT's Computer Science and Artificial Intelligence Laboratory (CSAIL).

The objective was to discuss some of the fundamental challenges in designing, deploying, and analyzing privacy-protective Internet protocols and systems, and to find ways to address such challenges in a systematic way. One of the key assumptions was that the topic of privacy is not an issue that can be looked at from an isolated perspective, but rather one that touches on many other standards development efforts. This vision of treating privacy as an overarching principle has since then be partially realized, for example by the establishment of the W3C Privacy Interest Group (itself part of the W3C Privacy Activity), or the IETF Privacy Program.

Topics of the workshop included the increasing ease of user/device/application fingerprinting (try the Panopticlick tool), difficulties in distinguishing first parties from third parties when making web requests, unforeseen information leakage, and complications arising from system dependencies. Some of the concrete technologies that were discussed were the W3C's early P3P standard, HTTP cookies, HTTP referrer headers, private browsing modes in web browsers, Do Not Track (DNT) technologies, the Tor onion router, the Geolocation API, and the OAuth protocol. Beyond the technological level, the workshop also addressed problems with transparency and user awareness, the difficulty of achieving balance between business, legal, and individual incentives, and the role of regulation in pushing for this balance. The tension between privacy protection and usability was also a major topic. For example, using Tor protects you from network surveillance, but it decreases browsing speed. Disabling cookies can protect you from being tracked by websites, but it impacts personalization.

The workshop concluded with a set of recommendations each single one of which is highly relevant for the PDE: The need to develop a privacy terminology and privacy threat models; The responsibility for protecting privacy to be split between protocols, APIs, applications, and services; The minimization of user data; The goal to give users granular control over their privacy; And the challenge to find the right balance between privacy and usability. A press release, meeting minutes, as well as the accepted position papers and slides are available for further information.

Book Review



by Terence Craig and Mary Ludloff

Paperback and eBook: 106 pages
O'Reilly Media (September 29, 2011)
$19.99/9.99 Kindle
Amazon Store O'Rielly Media

by Kaliya Hamlin

"*We*," Yevgeny Zamyatin's blockbuster novel of a dystopian society, was the first book that could be called "political science fiction." It was the inspiration for Orwell's *1984* and the model for Ayn Rand's *Anthem*. Zamyatin wrote of a glass city where everything everyone did was public and no privacy existed. It was a nightmarish vision of total repression. With nearly a thousand pieces of personal data being collected and sold on a daily basis, the specter of endless pilfering of user information has never been more important. That's why *Big Data and Privacy* is such an interesting book.

The very first page in the first chapter, entitled "**A Perfect Storm**" puts it well: the stakes have never been higher. "More businesses are making more money from data generated by users online and when people are told about how these industries work, about 80% disapprove." They make the point that the issue is not about how we are being advertised to, but about "the collection and use of our personal information from a commercial and political standpoint."

Written by the founders of a company called Pattern Builders, whose business is developing sophisticated tools to analyze digital data, they outline in the preface that data is the lifeblood of their industry, and that if they don't self-regulate enough they will lose the trust of their users and/or be over-regulated by government.

The book is indeed an informative narrative of the highlights and inflection points that lead to the internet revolution; and its growing to become a large part of our work and personal lives. They cover the nature of the big data age with a deluge of statistics explaining its size, along with why it is exploding because the costs of storing and analyzing it are dropping exponentially.

Behavioral Advertising is questioned as the "big bad wolf" of privacy; they say that "there is nothing morally wrong with it as long as you, the consumer, are aware of it. If your personal data is collected and solely used for the purpose of advertising, its impact is pretty benign." Instead they point out that "the debate is about how we balance privacy, security, and safety in an increasingly transparent and dangerous world." The words they didn't use were free speech, political, and economic freedom, all of which are impacted when living in "A City of Glass."

The next chapter, **The Right to Privacy in the Digital Age** defines three basic types of privacy:

•Physical: freedom of intrusion on your person •Informational: the expectation of privacy for personal information when collected and stored•Organizational: that organizations be able to keep information secret

It highlights a range of activities that were once considered private and might now be considered public, and how boundaries between physical privacy and informational privacy can feel blurred when discussing the issues. They do a good job of explaining and contrasting the US "right to be let alone" with the EU "honor and dignity" cultural and legal frames around privacy. They conclude the book by touching on *Networked Privacy "I" Versus the Collective "We"* highlighted by Danah Boyd in a talk at the Personal

Democracy Forum in 2011, explaining that because our data interactions are connected, our privacy is connected as well.

The Regulators chapter opens with these two questions:

- Is privacy a commodity that each individual, based on his or her preferences, can sell or rent in return for a service or product?
- Is privacy a basic human right that transcends commoditization, which must be protected at all costs?

They give a brief history of privacy regulation worldwide and evaluate the various privacy regulatory models: Consumer regulation (via their use of tools to protect themselves), Self-regulation, Sectoral Laws, and Comprehensive Laws. They cover all the different sectoral regulation in the US along with covering the context for the Federal Trade Commission and Federal Communications Commission involvement in privacy regulation. This is contrasted with the European top-down approach and a quick tour of other countries privacy laws. But they don't mention its connection to moral, social, political and economic freedom.

The **Players** chapter begins by highlighting 4 different groups, Data Collectors, Data Marketers, Data Users and Data Monitors/Protectors. Each of these groups place different intrinsic value that our personal data represents. They go through a short history of Online Advertising, highlighting key steps in the evolution of the current data ecosystem for targeting people online.

"Here is the rub: the information collected about is is not just used by advertisers to sell stuff to us. It's used for a myriad purpose, none of which we have control over.

They cover how Intellectual Property rights diffuse with Digital Rights Management tools and systems intrude on privacy in ways not possible before the digital era. They highlight the incident when Amazon removed George Orwell's *1984* from users' Kindles after they had purchased the e-book.

Next they turn their attention to the fact that Google, Yahoo, Facebook and others making billions of dollars collecting our data and using it for targeted advertising and that under US laws once you have collected the data you can sell it. I was very surprised by this fact they highlighted that the State of Florida sold DMV information (name, data of birth and type of vehicle driven) for 63 million dollars.

They highlight that businesses that see the data they collect as a possession that they own, can buy it, rent it, sell it and use it and that it is in their best interest to discover as much as they can about their users. They highlight that these players have little interest in comprehensive privacy regulation. They highlight many examples of how data is used without our knowledge. They go on to outline the limited involvement of companies in the self-regulatory efforts.

They highlight both Ann Cavoukian's Privacy by Design initiative and name Personal Data Ecosystem Consortium as "a Silicon Valley trade group... which promotes the idea that individuals control their own data through the use of personal data stores and services"

They begin the final chapter **Making Sense of it All** with a quote from Kennedy: "Like it or not, we live in interesting times" and say:

Every generation faces inflection points where the unknown becomes known. There are moments when the actions we take have unintended consequences; how we deal with those consequences defines us as individuals.

They raise a range of questions and make an important point that "there is a tug of war between all kinds of players who come at privacy from different perspectives, ranging from the utopian to Orwellian views of big data's impact on privacy." They name the heart of the matter for privacy, Commodity Versus Right, and that in the digital world we are all connected. They summarize 4 different bills proposed before the US Congress in 2011, and the FTC introducing a Privacy by Design Framework.

At the end of Zamyatin's book *We*, he describes a moment where the protagonist awakens and realizes that he is an individual, with a free mind. He manages to achieve privacy. He then experiences love for the first time. Without control of our persons and papers, freedom threatens to become an empty proposition; even meaningless. At the end of *Privacy and Big Data*, the authors conclude that this is not the first time that technology has leapfrogged ethics, bringing us to the age–old question of what we can do versus what we should do. In the end, people have to want control of their internet "effects", and it is our belief that once they have the tools to treat this property as their own, they will. Once empowered, they will have the choice to participate and share for their benefit and the benefit of society as a whole.

Opinion

What does a Free Market Look Like?

by Alan Mitchell

Assuming free markets a good thing, what does a free market look like? Is it one where there are no restrictions? Or does the resulting free-for-all create unaccountable concentrations and abuses of power which require legislative checks and balances?

This argument is being played out (once again) with renewed intensity as European nations – and US corporations – join the debate about the EU's draft new data protection regulations.

Unlike the US, since 1998 Europe has had relatively strong legislation designed to protect individuals' privacy and data. These laws are now being updated for an Internet age. The draft legislation – which will be debated for a few more years before finally passing into law – signifies an important mindset shift amongst the Europeans. Fifteen years ago, when Europe's first data protection rules were being formulated, individuals were firmly seen as the largely passive and powerless 'subjects' of corporations' data activities. The implicit assumption behind this new draft is that individuals should be empowered with much greater control over their own data, actively managing how this data is collected and used in various ways.

The two clearest examples of this are a new 'right to portability' and a new 'right to be forgotten'.

The right to portability. This comes in two flavours.

First, it establishes a right to move personal data from one service provider to another "without hindrance from the controller from whom the personal data are withdrawn". This has big implications for services like Facebook.

Second, it also establishes a right for individuals to have a copy of the data gathered by a company released back to them "in an electronic and structured format which is commonly used and allows for further use by the data subject". This effectively enshrines the principles of 'midata' – a voluntary programme of data release to individuals in the UK - into European law. It creates a new environment where individuals are seen as managers and users of their data, moving it around in ways that they can control.

The right to be forgotten. Under the draft, individuals can demand that data held about them by a company can be erased if "the data are no longer necessary in relation to the purposes for which they were collected or otherwise processed". There are many ifs and buts about detailed implementation here – but the trend is clear and again there are significant implications for the likes of Facebook.

Next to these new rights come new wordings which significantly enhance individuals' control over their data. One pivotal change is in the definition of 'consent'.

Under current European law, companies can only gather certain bits of data, or use this data in certain ways (e.g. for electronic marketing communications) if the individual has 'consented' to the company doing so. In practice, 'consent' was rendered close to meaningless by the notion of 'implied consent'. So long as the individual was 'informed' of the company's data policies somewhere (e.g. the small print of a privacy policy or terms and conditions), if they used the service or site they were deemed to have consented to these terms.

The draft regulations turn this on its head, requiring that individuals give "freely" a "specific, informed and explicit indication" of their wishes, with companies gathering this data bearing "the burden of proof for the data subject's consent".

If this shift survives the next two years of lobbying, it could transform the data relationship between individuals and organisations. Its potential impact is heightened further by changes to the definition of 'personal data' which now includes any information which can be used "directly or indirectly" to identify an individual including "reference to an identification number, location data, online identifier". The potential implications for the behavioural targeting industry are significant, especially in the light of other EU legislation requiring individuals consent to cookies being placed on their computer.

The regulations tighten up existing rules in other more subtle ways. Current legislation already enshrines the principle of data minimisation: that organisations should only collect data that is necessary for the purpose stated, and keep it only as long as is necessary for this purpose. The new rules add that this data can only be collected and retained if the purpose 'could not be fulfilled by other means'.

There are, meanwhile, many changes to how these laws should be enforced. The new rules will take the form of a

'regulation' which means that once enacted, they will apply instantly, and uniformly, across the whole of Europe. The previous rules required separate legislation by each member state with some countries ending up with much stronger protections than others. But penalties for serious violations being increased to potentially 2% of the firm's global annual turnover.

Significantly for American companies, the new laws will create jurisdiction over American-based organizations doing business in Europe, whether as government contractors, consumer-facing businesses or Internet-based businesses (including providers of cloud computing).

This takes us back to that free market debate. There are already many complaints that the new laws would create significant 'burdens on business', stifle innovation, and perhaps even drive companies away from doing business in Europe. The counter-argument is that the 'wild-west' approach to personal data currently being pursued by many US companies is undermining trust and acting as a long-term break on the growth of online services: many parts of this mooted law are a direct European reaction to the American experience: "we don't want that here!"

Two opposed world views are clashing here. To one, Silicon Valley is the proof of the free-market pudding. Look at the innovation! Look at the wealth creation!

To the other, much of this wealth is being derived from a new form of colonialism. Under the old colonialism, imperial powers gained unfettered extraction and mining rights to mineral resources of untold value – in exchange for a few baubles and trinkets handed to tribal chiefs. The new colonialism is internal – the colonization of individuals' private lives by large corporations; the unfettered extraction and mining of their personal data in exchange for the flimsiest of baubles and trinkets.

European legislators are saying this is not the way forward. The question is, can these clashing world views by reconciled in any way?

Alan Mitchell *is Strategy Director of Ctrl-Shift, who are acting as strategic advisors to the UK Government on its midata project.* www.ctrl-shift.co.uk

Kids and Personal Data: What you need to know about COPPA.

by Denise Tayloe

In the late 1990's, it became clear that children were actively as well as passively disclosing personal information online in the absence of parental consent. The growth of online services and content aimed at children resulted in a wide range of foreseeable problems, including:

- Invasion of children's privacy through solicitation of personal information granted unknowingly by child participants;
- The growing popularity of forums such as chat rooms, email, pen pals, IM and bulletin boards (and more recently, social networking sites, virtual worlds, multiplayer online games, mobile apps and interactive advertising) potentially exposed children to the lures of online pedophiles and generally made children targets for malicious adult behavior;
- An imbalance of power between vulnerable, young computer users and sophisticated new forms of targeted and behavioral advertising; and Absence of a parent's guardianship and consent in a child's online experience.

In October 1999, the FTC issued rules in the Children's Online Privacy Protection Act (COPPA) requiring operators subject to the Act to:

- give parental notice regarding the collection, use and disclosure policies of personal identifiable information;
- obtain "verifiable parental consent" before it collects personal information from a child or before it provides a child access to community tools whereby they could disclose their own personal information;
- allow a parent or guardian to review and/or delete the personal information already collected from his/her child; and establish and maintain reasonable procedures to protect the confidentiality, security, and integrity of personal information collected from children.

Personal information includes: full name, home or school address, e-mail address, phone number, or any other information (identifier across multiple sites, UDID of the cell phone, browser plug-in tied to an account) that would allow the child to be contacted online or offline.

COPPA applies to commercial websites and online services directed to and/or which collect personal information from children Under 13 and to operators of general audience websites, if they have actual knowledge that they are collecting personal information from children. Children-specific websites continue to grow as recording artists, movie producers, celebrities, video game producers, and the producers of children's toys, clothing and other products direct their constituencies to the "company" website or

mobile application. In addition to children-specific sites, there are exponentially more websites that target a teen, general or family audience.

The COPPA requirements can be challenging to adhere to as witnessed by the number of violations identified by the FTC and later fined for noncompliance. As recently as May 2011 the fine of $3m or $7.50 per child for Playdom (recently acquired by Disney) demonstrates that even those with resources may not have the knowledge to comply. Child-centered businesses are acutely aware of the lost opportunity costs of failing to interact with the U13 brand savvy population. Marketers recognize the household spending influence children have and acknowledge the importance of establishing relationships with young consumers because of their purchasing power and the brand influence they possess. However, these marketers are just learning how to use the Internet's power to harness the opportunity for two-way direct dialogue with youth consumers.

Industry Reaction

Since its inception, COPPA created headaches for online content providers seeking to collect children's personal information for marketing and other purposes. Within many companies, legal departments battle marketing departments that seek to collect information from Under 13 children. Websites and online services that attract children and tweens have been slow to embrace interactive solutions that trigger COPPA's guidelines and thus many sites have responded to the law by putting up age-gate restrictions on their site.

"Oops due to COPPA, we cannot allow you to join our site" has become a common message to Under 13 children. Children quickly learned to lie about their age in order to gain access to the interactive features on their favorite sites. As a result, databases have become tainted with inaccurate information, some companies are off the hook for providing the youngest online users with mandated protections and parents remain concerned, confused and tired, often times ignorant and in many cases participatory and ultimately they may even justify and endorse the child lying about their age. Despite the confusion in the industry, the FTC reported to Congress in March 2006 that COPPA was working well.

As a result of the FTC's study, COPPA was extended for the next ten years. The industry appeared to be waiting for COPPA to go away but the renewal of the law has been a signal to the kid's industry that it is time to take the law seriously, and apply methods that legally and responsibly allow Under 13 children the opportunity to join in the activities of their favorite sites with parent's permission.

The FTC is undertaking a full review of COPPA which began in March 2010. In October 2011 the FTC made it clear that mobile marketing is covered by COPPA, and proposed that behavioral advertising will be covered as well, along with the expansion of the scope of what constitutes PII and a discussion of how new methods of verifiable consent will be approved. What remain missing are stricter requirements for reasonable measures to obtain actual knowledge of a user's age. The public submitted comments at the end of

December 2011 and industry expects to hear from the FTC second quarter 2012 regarding the new rules and guidelines that must be followed.

COPPA Fines & Penalties

Under the COPPA organizations may be fined $11 per infraction for violating the regulations. However, this is not the only penalty incurred by the offending organizations. There is the brand or image damage, as mentioned previously, that the violators suffer and can have a measurable impact on offenders. This has been one of the key drivers that cause companies to be very risk averse of the youth demographic. Another lingering impact is a five year consent decree providing for administrative controls over all senior management that hangs over the future operations of the offending company. This watchdog environment serves as a constant reminder that adherence to the regulations is to be taken seriously. Additionally, all of the data that may have been collected from the unauthorized contact with youth must be deleted, thereby losing all of the relationships that the companies expended monetary and human resources to capture and had been hoping to develop. There is also potential for a personal penalty for those individuals in a position of responsibility for the offense.

Conclusion

In the "kid space", ONLY a company that provides privacy controls for the end user will be in a position to ask for permission to know their young customer. Earning trust is hard work and a vibrant identity ecosystem can play an important role in providing a framework for business and consumers to treat each other fairly.

The emerging network of personal data stores and services must consider how to play a role in compliance with COPPA (eg. being a repository for the data sites have collected about their children). My company has a service PrivoConnect that is designed to enable parental rights while giving children access to interact with a world of content and services.

Denise Tayloe *is Founder, President and Chief Executive Officer, dtayloe@privo.com, a PDEC member company.*

Publisher's Note

by Kaliya "Identity Woman" Hamlin

What's NSTIC got to do with Personal Data?

Last June I was flown out to Boston to deliver a presentation about the emerging personal data ecosystem at the [US] National Strategy for Trusted Identities in Cyberspace Privacy Workshop. (You can see my slides here). This event was one of several workshops hosted after the White House released the strategy in April, 2011. While there, the National Program Office (NPO) staff encouraged me to write a response to their Notice of Inquiry about how the Identity Ecosystem then envisioned should be governed and I did indeed write one (you can read it here).

In the last month, the NPO within the National Institute of Standards (NIST) at the Commerce Department issued both its recommendation for the development of a governance body for the Identity Ecosystem and an RFP to give 5-8 grants of $1,250,000-$2,000,000 for pilot projects, with preliminary applications are due March 7th. With this announcement, it now seems like the right time to share more about NSTIC and explain its relevance to the Personal Data Ecosystem.

When the Obama administration came into office in 2008, it initiated a cyber security review. One of the threats they identified was seemingly small within the larger cybersecurity framework, but it was indeed very significant - the re-use of Passwords. Simply put, the fact that people use the same password for multiple separate sites where they also use the same user-name (likely their e-mail address). The fact that people used the same password at a small un-important site with low security where if compromised the same user-name/password combination would work to access people's accounts on more important and secure sites is a huge issue for both privacy, identity theft, and national security.

One result of the process is the development of a National Strategy for Trusted Identities in Cyberspace(NSTIC). The NSTIC document was developed in consultation with industry to catalyze the evolution of an identity ecosystem with stronger credentials and more interoperable digital identities from different industries and sectors.

I just used a buzz word that should be clarified before we go further. Stronger Credential. There are two ways credentials can be stronger:

- How strong is the authentication at the time of login? Do you prove you are the owner of an account/username by just sharing a secret like a password or is there another "factor" such as the device you are logging in from or a one time password from a token or an SMS sent to your phone? The technical term here is "multi-factor authentication."
- How reliable is the assurance that you are who you claim to be? How was the identity you asserted verified during the enrollment process - the issuance of the account?

Different industries have different practices about how they do identity assurance and the standards and practices that are different in different

industries need to be documented and understood so that credentials issued in one industry can be used in another context

If one had an identity issued by CertiPath for example (the trust federation for the defense industry) could it be used when logging into a personal health record stored at a hospital one was just treated at? Both from an assurance of identity perspective but also from a hardware token that provides 2nd/3rd factor authentication.

If one was a contractor with the federal government and one had a HSPD-12 credential issued (Homeland Security Presidential Directive - 12 issued by President Bush to background check and give a standard interoperable digital credential to all federal employees and contractors of which 12 million have now been issued) could one use that to login to one's bank?

If one goes through the know your customer process mandated by US law to be able to have a bank account could that same bank issue a digital identity accepted by the [US] Internal Revenue Service to see one's tax record with the government. The strategy states multiple times that the ecosystem must maintain the current features of pseudonymity and anonymity. Industry seems to be clamoring at the opportunity to provide stronger identity credentials (see the Google LMNOP presentation about is work on verified addresses with Verizon). However it is not clamoring to provide choice and effective tools for pseudonymous and anonymous identities. There are many open questions:

How will such stronger identity credential systems affect "benign anonymity" in low-risk environments?

Is it wise from a political, social, financial, and freedom perspective to encourage the convergence to a few credentials or single credential when systems like this have and can be used for repression?

Can the privatized issuance of identity credentials lead to new free market systems that insure such credentials allowing the free market to bear the risks and rewards?

These types of uses case around interoperability of simple identity credentials is one layer of what NSTIC articulated and is the most pressing problem they are trying to address for several reasons:

For cultural, political, and legal reasons, the US government cannot yet issue a national online ID that governs all one's interactions with government agencies. Nor has it been shown that such a step is even beneficial.

Government agencies that provide citizen's with services can't actually do so without having some identity assurance about who they say they are (a citizen might interact with 4+ agencies) and they don't want to be in the business of verifying citizen identities, issuing digital credentials thus giving citizens 4 different credentials to login to 4 different agencies and the cost to do this is currently at $50 each and well the cost to do this are prohibitive.

Could this be privatized? Leveraging the identity vetting and verification that happens all the time in the private sector for its own business assurance needs to enable citizens to have verified identities from the private sector accepted at government agencies.

Government issued credentials should be able to be accepted by commercial companies.

Supporting thriving commercial activity via online/digital systems because there is confidence in both the overall system and the credentials used by people and companies they do business with.

There are technical interoperability challenges to be solved for this kind of interoperability but there also exist sizable legal liability and political concerns - if a citizen uses a bank issued digital identity credential to login to the IRS - who is liable if something goes wrong. The buzz phrase for these to date has been "trust frameworks" combined technology/policy frameworks that help the parties "trust" (believe they are accurate) the identities that another party in the system has issued. The current term the ABA is considering for this topic in the Lexicon is System Rules: the "agreed-upon" business, legal, and technical rules and standards that will govern an identity system.

The credit card and other industries use them widely to make the networks of parties operate. There are experiments to grow open trust framework systems (where entity that proves they are in compliance with the policies and conforms to the technology standard can have their credentials accepted across other compliant sites).

The American Bar Association Federated Identity Management Task Force is working on defining a lexicon of terms and one of the proposed terms was Trust Framework however at a recent meeting focused on its development it was proposed this be changed to something that was not so broad and confusing in its meaning. I wrote about the issue of using the word to describe the many different types of trust in a system and how overusing it would diminish its meaning and the ability of people to actually trust these systems (you can read that post here).

The issue at it's core is about how trust means different things in different contexts at different scales. Regular citizens who are participating in a "trust framework" are going to think that all the people and entities one encountered within a "trust framework" are trustworthy and that the underlying policies were actually good and in alignment with respecting people's data/identities. All that system rules (trust frameworks) actually do is name the policies for a particular system - these may not be good for users or organizations within them.

The path to a thriving personal data ecosystem coming into being will be through the development of multi-party networks what have at their core system rules (trust & accountability frameworks) that are in alignment with people and respect the individual. Finding ways to manage risk, liability and create accountability with trust-frameworks/ system rules for digital systems in ways that the banking/credit card network exchange valuable information/currency is a logical source of inspiration for this inspiration.

The vision of an Identity Ecosystem articulated in NSTIC goes well beyond identity credentials for verified identities. It envisions a future with both the technical and policy infrastructure needed for people to be able to share all types of attributes - personal data associated with themselves. The US government is keen on seeing market solutions emerge to the privacy and trust challenges posed by today's internet. For this reason Jeremy Grant invited me to present about PDEC at the the NSTIC Privacy workshop. Additionally, I was also invited to present about PDEC and Identity Ecosystem governance on a panel with Jeremy Grant the head of the National Program Office at the RSA Conference in San Francisco February, 29th at 3pm.

To conclude, the NSTIC Pilot grants are an opportunity for companies working on personal data tools and services that extend beyond the first narrow set of use-cases around verified identity login's to collaborate together on the development of an interoperable ecosystem. Identity Commons is hosting a survey for organizations and companies who see themselves as stakeholders in NSTIC and have something to contribute to a pilot proposal.

(OAuth, Continued from Page 2)

Since this is a quite cumbersome process, engineers at B might at some point decide to offer you the option to enter your username and password of A at B's website, for the purpose of retrieving the photos on your behalf. This implementation has a terribly weak design from a security perspective, since your credentials enable B not only to retrieve your photos, but also to perform any other action on your behalf, such as deleting photos or even changing your password and locking you out.

This is where OAuth comes in. It provides a mechanism for B to get permission to retrieve your photos from A, without knowing your username and password, and without receiving more permissions than are necessary to complete the task. In real life, an analogy that has been used by OAuth developer and evangelist Eran Hammer-Lahav to explain this idea is that of a "valet key" that comes with a luxury car and can be given to a parking attendant. Unlike the regular key, which allows full access to the car's functions, the valet key will only allow driving the car for a very limited distance and will not allow access to advanced functions such as the car's trunk or onboard computer.

In OAuth terminology, A is called the "server" (also "resource server" or "service provider"), and B is called the "client" (also "consumer"). The basic OAuth flow involves a series of HTTP redirects between the involved parties, during which the user is authenticated at the server and asked to approve the permissions requested by the client. After that, the server issues an "access token" that can be used by the client to access a "protected resource". The access token has an associated "scope", which determines what exact permissions are made available to the client when accessing the protected resource. Also, it is interesting to note that when making requests, it is not only the user who is authenticated by the server, but the client website authenticates to the server as well, which leads to the familiar "Would you like to allow this app access to your wall/photos/etc.?" screens.

Within a relatively short time, companies all around the web began to realize that the cost of implementing something like OAuth and therefore opening up their users' data to third parties was small compared to the tremendous security risks associated with a web in which users are trained to commonly give away their credentials. This practice has been described as the password anti-pattern, and before the introduction of OAuth, it had been common to use this approach for exporting lists of friends or address book contents.

The lesson that an emerging PDE can learn from this experience is that sufficiently open policies and architectures around personal data are not only desirable from an ethical point of view, but they also prevent the emergence of badly designed workaround techniques that ultimately hurt every actor in the ecosystem.

OAuth 2.0

Following the success of OAuth 1.0, development of OAuth 2.0 began soon after, which – although based on the same high-level idea and on similar design criteria – is a completely new protocol that is not backwards compatible. A

short-lived proposal called OAuth WRAP was also developed but soon became folded into OAuth 2.0, which is currently available as a draft within IETF and expected to become a standard soon. Based on the lessons learned from several years of designing, implementing and using OAuth 1.0, the main improvements of OAuth 2.0 are as follows:

- While OAuth 1.0 had been specifically designed for an architecture involving a user agent (usually a web browser) that enables communication between the client and server via HTTP redirects, OAuth 2.0 adopted a broader perspective, in which the different actors are defined in a more abstract way. Therefore, a more flexible use of credentials and response types, as well as a greater variety of authorization flows becomes possible.
- For example, this more abstract design means that OAuth 2.0 can not only be used by server-side web applications (the so-called "authorization code flow"), but also by user-agent-based applications such as JavaScript within a web page (the so-called "implicit flow"), or even by desktop clients as well as native applications on mobile devices.
- One of the most complex aspects of OAuth 1.0 had been its approach to how a client proves its possession of an access token. This involved two different secrets and a not-so-trivial cryptographic signature procedure. This mechanism has been simplified, and a new cryptography-free option has been added in which a token is simply sent over HTTPS to the resource server.
- Together, the abstraction of flows and the simplification of signatures enable another advantage of OAuth 2.0: The possibility to separate the authorization server from the resource server, which is important for performance and scalability reasons. In other words, this allows for large architectures in which authorization is managed in a centralized way and relied upon by an unlimited number of resource servers.
- OAuth 2.0 can also be deployed and used in a way that minimizes round-trips and the required CPU processing power, which makes it more cost-effective in terms of server resources and therefore more attractive for large sites with millions of users.
- Finally, OAuth 2.0 also introduces the notion of a "refresh token". This makes it possible for the server to issue only short-lived access tokens to a client, which can replace them at any time with fresh ones. This new mechanism is another measure to make the protocol more efficient.

Because of these changes, OAuth 2.0 is on one hand considered generally more easily implemented in its basic form, but on the other hand also more flexible, secure, extensible, and suitable for a wider range of use cases.

What OAuth is Not

It is important to remember that OAuth is a protocol only for authorization, not authentication. This means two things: It means that OAuth does not mandate how the server authenticates a user. This may happen with a traditional username and password, or with OpenID, or with more sophisticated authentication technologies such as biometrics, one-time-passwords, or enrolled mobile devices. It also means that the client should not rely on OAuth for identifying a user and logging her in. It is possible (and actually quite common in practice) to do this if the server allows the client to use OAuth for the purpose of retrieving a user identifier.

For example, this is how the Facebook Graph API allows users to log in to websites with their Facebook account, a process also known as Facebook Connect. This API is based on OAuth 2.0 and gives a client access to a user identifier (in this case, a field named "id" within a JSON object), and therefore makes it possible to log her in. While access to the Facebook Graph API in general is a great example for the kind of use case that OAuth was indeed designed for, there are severe problems associated with using OAuth for authentication purposes. Especially in the implicit flow (called "client-side" by Facebook), this can lead to malicious websites impersonating a user at other Facebook-enabled websites. The authorization code flow (called "server-side" by Facebook) is more secure, but this does not change the fact that OAuth was not made for login purposes. As IETF OAuth working group member John Bradley explains this in a recent post:

> *The problem is that OAuth 2.0 is a Delegated Authorization protocol, and not a Authentication protocol. This leads people to make what turn out to be very bad security decisions around authentication when they follow the basic OAuth flow.*

Another aspect that is sometimes misunderstood about OAuth is that it does not in any way specify how the different actors exchange information beyond what is required for the actual authorization process. OAuth allows access to a protected resource, but how the protected resource is accessed or used depends solely on the application developer. This seemingly simple point might – if overlooked – leads to an incorrect perception that if two different services both supported OAuth, they would automatically (or with limited additional

effort) become interoperable. This is not the case, since for actual interoperability between services the exact APIs and semantics of the protected resource have to be agreed on, which is completely out of scope of the OAuth protocol.

OAuth also does not specify how a server makes the decision of whether to authorize a request or not. Typically this involves asking the user to give explicit consent, but depending on the exact flow the decision could also be made by the server alone based on access control technologies such as the eXtensible Access Control Markup Language (XACML), XDI Link Contracts, or other policy expression languages.

Relation to OpenID Connect

One cannot write a feature article about OAuth without also mentioning OpenID, especially its most recent incarnation OpenID Connect. Firstly, the topics of authentication and authorization are obviously interrelated per se. Secondly, the communities as well as the main proponents that created and promoted these standards overlap considerably. And thirdly, although early OAuth and OpenID specifications were not dependent on each other from a technology point of view, today's OpenID Connect does actually build directly on OAuth 2.0, leveraging some of its extension points.

As has been described earlier in this article, OAuth 2.0 by itself is only meant for authorization purposes, not for logging users into a website. OpenID Connect fills this gap by adding the necessary semantics to enable a client to identify users. To be more precise, it does this by specifying the additional value of "openid" for the scope parameter. Just like with plain OAuth 2.0, a client receives an access token from the server and uses it to access a protected resource – which in the case of OpenID Connect is called UserInfo Endpoint. In doing so, the client obtains a "user_id" field, which is a locally unique and never reassigned identifier for the user. This is what the client uses for login purposes, and this is what turns OpenID Connect into an actual authentication rather than just authorization protocol.

OpenID Connect also specifies an additional value of "id_token" for the response type. This ID token, which is issued in addition to the "regular" access token, makes sure that the server's response is legitimate by providing the client with a set of security measures about the authentication event, such as an audience restriction, a nonce and an expiration time. The ID token is in fact a JSON Web Token (JWT), which together with JSON Web Signature (JWS), JSON Web Encryption (JWE), JSON Web Key (JWK) and JSON Web Algorithms (JWA) is being developed alongside the OpenID Connect effort. Clients can either decode and validate the ID token by themselves, or utilize the server's Check ID endpoint – another extension made by the OpenID Connect specification.

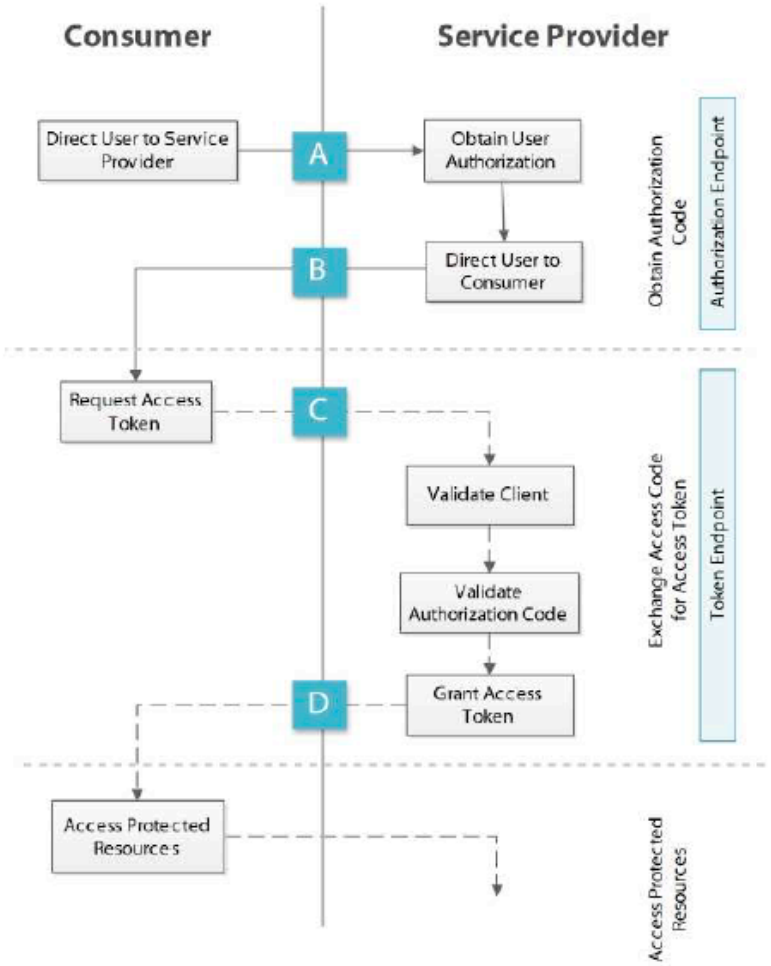


Figure 5: OAuth 2.0 Authorization Code ("Server-Side") Flow

Relation to User-Managed Access

User-Managed Access (UMA), formerly also known as ProtectServe, is another interesting protocol which is closely related to OAuth 2.0. Led by former PayPal engineer and now Forrester analyst Eve Maler, it is being developed at the Kantara Initiative and moving into IETF. It directly builds on OAuth 2.0, even though its different terminology might make it hard at first to realize this ("host" instead of "resource server", "authorization manager" instead of "authorization server", "requester" instead of "client"). The main innovation of UMA is to greatly expand the role of the authorization server, which in plain OAuth is tightly coupled to the resource server. In UMA's model, the user chooses a single authorization manager, which is designed to be responsible for issuing access tokens for all of the user's hosts.

OAuth Example:

Given an OAuth 2.0 access token an HTTP GET request to the Facebook Graph API can retrieve the following information about a user: https://graph.facebook.com/me?

access_token=XXX authorization. The term client does not imply any particular implementation characteristics (e.g. whether the application executes on a server, a desktop, or other devices).

XX


```
{
"id": "588183713",
"name": "Markus Sabadello",
"first_name": "Markus",
"last_name": "Sabadello",
"link":
"http://www.facebook.com/markus.sab
adello",
"username": "markus.sabadello",
"gender": "male",
"timezone": 1,
"locale": "en_US",
"verified": true,
"updated_time": "2012-01-
31T14:10:53+0000"
}
```


In the UMA vision, hosts can include personal data (such as identity attributes), content (such as photos), and services (such as viewing and creating status updates). UMA provides a dashboard-like interface where authorizations to all these hosts can be managed from a unified point, no matter where all those things live. This gives the user more overview and control over who can access what, rather than having to manage permissions in different places scattered around the web.

OAuth and the PDE

Today, OAuth is widely accepted and deployed, and most of the major web companies use either 1.0 or 2.0 for allowing access to their APIs. The way OAuth can be used within a PDE is rather obvious and can be summarized in one simple sentence: It can be the technology of choice for authorizing access to your personal data. Already, several PDEC members are using it to protect their public APIs, through which 3rd party applications gain access to users' personal data stores or similar services.

Using the OAuth terminology, a provider of a personal data store service would be the "server", and any 3rd party wishing to access an individual's personal data would be the "client". The "scope" parameter would be used to specify what kind of access is requested exactly (Write access to your resume? Or read access to one of your entire personas?). And the "access token" would contain the exact permissions that have been granted by the user – you. It could be valid only for a single transaction, e.g. in a situation where your mailing address from your personal data store might be needed for checking out at an e-commerce store. The access token could however also be valid for a long time until you explicitly revoke it, e.g. if you want a company to have an ongoing subscription to some of your personal data. Using UMA as an additional component, this scenario could even be extended to provide you with a dashboard that lets you manage your personal data not only within your personal data store, but in fact all over the web.

One challenge we have to be aware of is that even though OAuth is more secure than asking users to give away their passwords, it can still pose privacy risks if not used responsibly. For example, badly designed click-through OAuth flows and user interfaces might train users not to care enough, and to give away too many permissions to their personal data too easily. Also, it might become difficult for users to keep track of the permissions they have given, and to revoke them when they are no longer necessary or desired. In today's world, such transparency problems have been described as "the dark side of OAuth", and they can range from websites having perpetual access to your Gmail account, to the Twitter API not revoking your OAuth access tokens after you change your password. In a PDE, both the privacy-related user interfaces and the granularity of access control must be designed with great sensitivity.

In conclusion, OAuth seems to be predestined for the ecosystem that PDEC members are envisioning – potentially supporting both increased levels of privacy and control, and very flexible ways of sharing our personal data with companies as well as with other individuals. As has been mentioned in the introduction, out of all the different technologies that are currently being used in the PDE, OAuth might very well emerge as the first piece in the PDE puzzle that a wide consensus can be reached on, even though such a consensus alone does not automatically make different actors compatible. Let us agree that OAuth will be this first baby step, but let us also keep in mind that we have a lot of work ahead of us if we want to achieve an interoperable PDE in which different actors can work together seamlessly.

Editorial

by Kelly Mackin

The Supreme Court's recent action in the U.S. to prohibit police GPS tracking devices on suspect's automobiles without a warrant brings some fresh attention to police activities that affect privacy. While the decision was touted as a decision in support of privacy, the result centered more narrowly on property rights and avoided the more nettlesome and issues of privacy. With property rights under pressure in the United States from enforcement and regulatory agencies, the decision sends a signal that the court is prepared to move the boundary back someways towards the individual; thus observers breathe a sigh of relief.

But the road to better privacy is still long. Recent decisions in the European Zone from Denmark and Norway (reported elsewhere in this publication) show political and public concern for freedom from what Literary Critic Michel Foucault called, The Surveillance Society. With data giants like Google and Facebook, a convergence could be underway that blends business and state interests in an "Orwellian Web."

The Norwegian Decision to prevent public sector entities from using Google Apps arose out of Europe's concern about the broad parameters of the misnamed PATRIOT Act. Specifically mentioned in the reports were the concern that data located in the United States would be compromised.

Events like this make it clear that laws put forward in the name of security end up hurting U.S.-based businesses that intend to compete in a global marketplace. But to stop there would be to miss the sleeping giant. Looking at the reaction in the United States in response to Google's merging of all its privacy policies into one and linking data of users across all their services, one sees a bonafide trend that security and privacy concerns are having a quelling effect on adoption and continued use *inside the United States*. A recent privacy expert on a late night nationwide radio show implored users to avoid Google and choose privacy-protecting search engines such as DuckDuckGo or StartPage, which do not record IP addresses.

Polls in the United States show that a vast majority of people want their privacy restored and this fact bodes well for the evolution of both personal data control and accountable online interactivity. But the Supreme Court did not have a case in front of it that would have allowed for a broader decision to address the elephant in the room. At the least, and for a start, as the above examples demonstrate, it would be good for the United States to recognize that hypersecurity, rule-bending authorities, and draconian laws are bad for business.

I think its important for all participants in this amazing technological process we are in to consider how they can architect their products to better promote the needs and concerns of individuals.

The United States has had a reputation of being the model for an environment in which to nurture and grow new businesses and markets. If the U.S.A, as a society does not tread carefully in the area of security as it intersects the private economy, it could create a situation where the United States becomes a technological island unto itself.

Before I returned to editorial work, I had a great experience in Washington working with The Congress to develop laws that protect people from various types of malware. I can only hope that companies with important business in Washington take the time to impress upon their lawmakers the importance of individual privacy in global business success.

Kelly Mackin is the editor of Personal Data Journal