Personal Data2.0

Igniting the Personal Data Ecosystem

AMERICAN CONFERENCE

April 7, 2011

Palo Alto

This day was part of New Digital Economics AMERICAS Executive Brainstorm and Developer Forum. It was a participant-led workshop builds on the output from the World Economic Forum's 'Rethinking Personal Data' project.

> ***End User-Centricity*** *refers to the concept of organising the rules and policies of the* ***personal data ecosystem*** *around the key principles that end users value:* ***transparency*** *into what data is captured,****control*** *over how it is shared, trust in how others use it and* ***value*** *attributable because of it.*
>
> ***Trust Networks*** *are a formalised specification of policies and rules to which all participants must conform to be trusted. These include requirements around identity, security, privacy, data protection, technical profiles and assessor qualifications. This trust may be subject to different levels of assurance or protection, which are explicitly made clear to all parties.*
>
> Marc Davis, Ron Martinez & Chris Kalaboukis.
> "Rethinking Personal Information - Workshop Pre-read" Invention Arts and World Economic Forum, June 2010

STL Partners Logo World Economic Forum Logo

personal data ECOSYSTEM

personaldataecosystem.org

The purpose of this day was to make progress on making the vision of a Personal Data Ecosystem real. Oriented around participant driven mutli-disciplinary sessions focused on putting theory into practice. Below are aspect of the invitation an then a table of contents with all the notes from all the sessions that happened.

The Hosts

Simon Torrance CEO of Telco 2.0 and New Digital Economics and Kaliya Hamlin a leader of the Internet Identity Workshop (IIW) & Personal Data Ecosystem Consortium are teaming up to co-lead the day.

Building on earlier work

The Telco 2.0 Initiative, run by STL Partners, has been putting forward key thought leadership on how telecommunications companies can develop two-sided business models and provide new services. Acting as 'custodians' of consumers personal data is one such potential area.

The World Economic Forum's Rethinking Personal Information project generated considerable interest at Davos in January 2011, with the mandate to create more use cases on personal data ecosystems. The project's concluding recommendations are in five areas:

1. Innovation around user-centricity & trust
2. Global principles,
3. Regulation,
4. Interoperability/Open standards
5. Knowledge sharing.

The workshop deliverables will be mapped back to these five areas.

The Internet Identity Workshop has been at the forefront of the user-centric identity movement since 2005. At its last event key questions addressed included:

What are potential business models including those for 3rd and 4th party providers? How should we be defining key elements of the emerging ecosystem such as the functions of a Personal Data Store?
What is the role of Vendor Relationship Management tools?
What are the standards available to foster interoperability?

Participants from the following Industries and Disciplines were invited

Industries

- Web companies (big ones and startups)
  - Mobile carriers
  - Advertising
  - Browser vendors
  - Cable companies
  - Entertainment companies
  - Device manufactures
  - Back-end infrastructure vendors
  - Banking groups
  - Insurance providers

Disciplines

- Executives
  - Product managers & designers
  - End-user advocates
  - Software architects
  - Lawyers
  - Business strategists
  - Change agents
  - Privacy experts
  - Cyber security professionals

This is how agenda creation was described on the website.

Agenda

This will be created live the day of the event. So now you are wondering - how do I “know” what will happen? What will be covered? Will it be worth my time? All participants/attendees have an opportunity at the beginning of the day to propose sessions - these can take the form of a presentation, discussion or a question that needs an answer. Because we are inviting the leading thinkers in the space along with the startups working on the ground, it will be dynamic and engaging covering the cutting edge real time issues and challenges emerging. With the real time dynamism includes real time documentation - all the notes from all the sessions will be collected and distributed to attendees.

Potential Topic Areas *identified before the conference*

The broad objective of the workshop will be on defining Global (user-centric) Internet Identity that connects the telco stack with web stack. The specific topics will be chosen by participants on the day. The topics outlined below are therefore provided for guidance only.

- What are the global principles (and exceptions) that apply?
- By focusing on specific data elements (e.g. identity and location) Can we map-out elements of interoperability?
- Who controls what data and how?
- How does data get to be used? (e.g. Location)
    - Standalone / in combination with other data sets (potentially from a different source e.g. MOSAIC) / by a service (e.g. web search)
    - Individual level / aggregate level
    - Granularity (exact, neighbourhood, city, state, country, continent)
    - Modality/Geo-fencing (within/outside a predefined area)
    - Timeliness (current, historic precise, historic average)
- How are social networking sites and social media tools applying user-centric identity?
- What are the open standards to make it work? (identity and semantic)
- What are technical implementations of those standards?
- How do different standards and technical implementations interoperate?
- What are the new social norms and legal constructs needed to make it work?
- What tools are needed to make it useably secure for end-users?
- What are the businesses cases / models that drive all this?

The topics covered were put forward by attendees the day of the event. This book contains the notes from all the sessions. At the end of the day attendees were invited to put forward next actions AND these along the topics were contextualized relative to the 5 recommendations of the WEF report.

TABLE OF CONTENTS

Innovate around User-Centricity and Trust

Chained Identity

Session: 1A

Convener: Wendell Baker

Notes-taker(s): Greg Goodwin

URL: http://hub.personaldataecosystem.org/wagn/Chained_Identity_Ecosystem

Discussion notes:
Two players in online entertainment & advertising system

1. Audience: login is traditional method to determine identity. Phone & tablet inherently have identity.
2. Advertising engines like stock markets combining specific user id and 3rd party data to serve best ad.

Advertisers, publishers, merchant data

Publisher data leakage

- Publisher = web site
- Some places estimate the audiences for various websites – demographics of ESPN vs NYTimes
- Expensive to advertise at NYTimes, people try to figure out who those folks are and advertise to them when they visit a cheaper site
- Individual users have no say today, how do you compensate at <1cent per transaction

Challenge: how do we link chosen identity system with mobile device identity

Individual data and behaviour – error rate is huge. Not a bank.

Login:

- Identify yourself to the site
- Set of interests – mostly declared interest, inferred is emerging
- Actual geographic location is permission-based.
- Need to disassociate shared device from user, reconnect same user using multiple devices
- Yahoo doesn't fingerprint, reassociate IDs etc. Some companies do. Usually done in anonymous look-alike user groups.

Interest Manager

- Specifies kinds of data that you will allow to be used

Yahoo provides data back to publishers, back to advertisers on campaign performance. Data not available to users.

Concern – chaining identity together across web, mobile etc pushing boundary of transparency, trust

Nobody has figured out identity in communal TV viewing context yet.

Cross Service Data Portability

Session: 1F

Presenter: Rajat Shroff & Eric Sachs

URL: http://hub.personaldataecosystem.org/wagn/Cross_Service_Data_Portability

15 attendees

Identity Solo
Who is who, a specific type of data

Data Solo
What data can I use
How do users understand the data they are giving in OAuth flows

There are no standard terms
It works better when users are making decisions about data associated with a specific brand, like Blogger or Netflix. IT works worse for broad brands like Google and Facebook that are not as well defined for users.
Also need to explain the value proposition that the user is getting for giving access to their data. The data provider cannot do that, the client service requesting the data must explain it. That causes the pages with explanation and data to be in different places which is a problem.

Personalized Ads
Even if they can be made more personalized, the context of the page tends to be of more importance. In particular, there are types of pages where the user has very little interest in clicking ads.

Benefits of cloud vs. On my phone
Where you store information can make users more comfortable
Android is cloud devices, iphone is based on storing data locally

What is the ROI of sharing?
local cache is more controlled

Is the discussions about technoogy or use-cases

Assuming users are sharing data between services, then

If you are starting an interaction with a commerce site
..start with an anonymized id
Discussed experience of Checkout vs Paypal
The anonymization of checkout might be good for buying at porn site, but paypal is more popular because it does not try to over-anonymize

Online credit card entry pages tend to have verification seals, or credit card companies control your financial exposure. What is the equivalent for data sharing pages?
May not be feasible for data sharing because there is no well quantified financial amount to give back to the user if their data leaks

Will there be a few large data oligarchires or lots of small

Kaliya wants to whiteboard the Data bank/store/lockers - data verticals - services

Tripit example of importing your entire mailbox
- how are users making that trust decision
Why not look at all your email for Amazon purchases

Is that good for users?
Theoretically, but maybe it would undercut the economics of some websites
FB started open, but now locking the data

Offermatic example for credit card monitoring and coupons
What is the economic incentive for banks to provide APIS to the data they store about their customers
Is the data the users? Is the data the banks?

Economics of Apps Marketplaces to help justify the creation of a platform for data/service providers like banks

Exchange of user data without their involvement

What about startups that fail, and their data is no longer accessible to them

Maybe no single owner of data, but more like water rights that are shared

Mint.com got her permission to share the data
Other companies sometimes share data about same users without opt-in

Portable Context

Session: 2G

Convener: Joe Andrieu (from SwitchBook)

Notes-taker(s): Evan Welbourne

URL: http://hub.personaldataecosystem.org/wagn/Portable_Context_Session

Tags:
Context, SwitchBook, personal store, organizing context, user control

Discussion notes:

Early on, an open question and light discussion of the definition of "context"

- How does is it distinct from identity or persona and how is it related?

Introduction by Joe Andrieu of SwitchBook

- Great YouTube video overview at: http://www.youtube.com/watch?v=WmtrkYh8PR0
- See also www.switchbook.com
- Describes a system that collects your web browsing activities and stores them locally in a file
- This context file can then be used by participating websites to personalize your experience
- The end-user can also edit the contents of the file, create multiple files

Discussion of the user's role in curating the collected context:

- Currently in SwitchBook the user can edit the context file with a sidebar interface
- In SwitchBook separate files can be created that capture different categories of context

- Q: In general, how can end-users organize, filter, and control context? Evernote is an example of a system that collected personal data but takes a lot of effort to use. Is there a better way?
- Response: Use an event-oriented architecture like Kynetx (http://apps.kynetx.com/), this would support automatic filtering of context.
- Response: In 10-15 years from now it may be that all this is accomplished automatically by data mining and machine learning similar to how Atigeo's (http://www.atigeo.com/) technology works.
- Follow-up response: Even with machine learning there is still the problem of *intelligibility*: how can an end-user control the contents of a personal data store if they have no way of understanding its contents or how the content is created (e.g., misc clusters and weights produced by a complex learning or data mining algorithm).
- Apps written by third parties could be responsible for organizing the collected context into relevant files or data stores.

Comments on user control of context (not necessarily specific to SwitchBook):

- It could be a nice UI feature to have a slider that lets you set how much context (i.e., the last minute, the last 5 minutes, the last day or year) is available to websites that use the context
- Q: In SwitchBook, when a website uses the context file, do they get some select part of it or do they get the whole thing?
- Response: The whole thing, finer-grained control is achieved with multiple files.
- There is a lot of value in having users specify their intent directly. Users could even specify the conceptual link between low-level context log entries and high-level activities.

Comments on ownership of personal data:

- "Ownership" is perhaps an outdated term
- The vast, dynamic set of potential legal terms that would need to be articulated and negotiated for individual cases can't be practically encoded into a piece of software
- Note that NELIC 2011 is coming up: http://www.robotandhwang.com/?page_id=189 – a conference on new and emerging legal infrastructures

- A related question is in where the data lives physically: on the user's device, on a server, in the cloud?

User centric, personal data store roundtable

Session: 3C

Convener & Notes Taker: CLOUD CEO Kathy Tavitian

URL: http://hub.personaldataecosystem.org/wagn/User_Centric_Data_Store_round_table

Net is broken, focused on pages, lacks connection between people, including value creation segment, e.g. if I have cancer, could give my info to all doctors working on that type of cancer, maybe anonymously. Create a new markup language CTML Contextual Markup Language

Is it like XACML or XDI? "Don't those assume the current paradigm?"

Idea: you keep your data, let others present / manifest it in different ways. Sounds similar to personal data store so far

Q (Simon from sing.ly): Isn't the idea of a PDS orthogonal to language it's stored in? (then introduces sing.ly PDS idea)

Locker code is open source so others could provide PDS based on it. sing.ly will be a provider itself.

Right now Locker is "just a bunch of connectors"

Locker will store your "data exhaust". e.g. FitBit sending your heart rate data

CLOUD interested in Locker

Q: (CLOUD) How to authorize releasing only part of info to someone?

Q: What format does Locker use? JSON. Other than that try to keep original format.

How do you relate data from different sources? Still working on how exactly. Do you have any ideas? Again, we think JSON is powerful. There's not enough to solve the problem now, we need actual data in order to decide it or let the market decide.

John Carney of CableLabs: There is lots of Semantic Web work already on understanding/relating data.

Cameron Lewis of Statz: Does anyone here work with health data? "Patient owns the writing, doctor owns the stationery" Number of imaging protocols alone is huge. Data involved is terabytes. Authentication another major problem. "3 trillion dollar economy that doesn't want your ass anywhere near it" HIPAA doesn't allow you to restrict redistribution of de-identified data.

Q: Has sing.ly dealt with de-identified data? Not yet. Haven't even decided it is a necessary piece, though would be nice.

Hard to enforce revocation.

CableLabs: We have all this data but it's highly regulated, PDS will help us find and use data that user has given permission for.

Statz: Have you done research on where equilibrium value is between doing something + cost of doing something about privacy vs not doing anything?

Cable: Ads are $1.5b for Comcast but this is trivial.

Affiliate fees are 50% of revenue (rest advertising) so they're afraid to go on net.

If premium customers opt out of ads, your ad value dives

Cable: Consumers want free and no ads => only cheap crap content like reality shows will be around 5 years from now

sing.ly: Thinking a lot about how to certify how steps in data pipeline are trustworthy, or that data has arrived intact

CLOUD: How can Locker promote connections? A: Need to be able to exchange data directly between people's lockers. TeleHash will provide P2P

sing.ly will provide each user with fully encrypted virtual machine

Having all your data in one place will make innovation possible. yodl.ly and Mint do less than will be possible with locker

CLOUD: If access to locker is hacked, then they get all your info

Statz: We take broad view of personal data, not just traditional PII

Statz: We don't care as much about existing ad-based data and economy. Mobility and ubiquitous sensors are going to change everything in next few years. E.g. behavior in driving is a BFD, has a huge economy in itself. Or bike biz is $20b, but now no way to measure it - sensors will allow this

Statz: there is so much info that you don't even have to get into conditions, pharmacology
There is no more value to be had from corporate-centered, the place to go is user-centered. E.g. cable has all this data, can't do shit with it.

Xavier (Orange): There is confusion between big data and personal data

Statz: Example from gas station biz. Try to find something with higher price elasticity than gas. Tobacco does, cheap cigs got young males to stop at our station.

Define global principles for using and sharing personal data

Trust Frameworks

Session: 5A

Convener: Drummond Reed & Scott David

Notes Taker: Drummond Reed

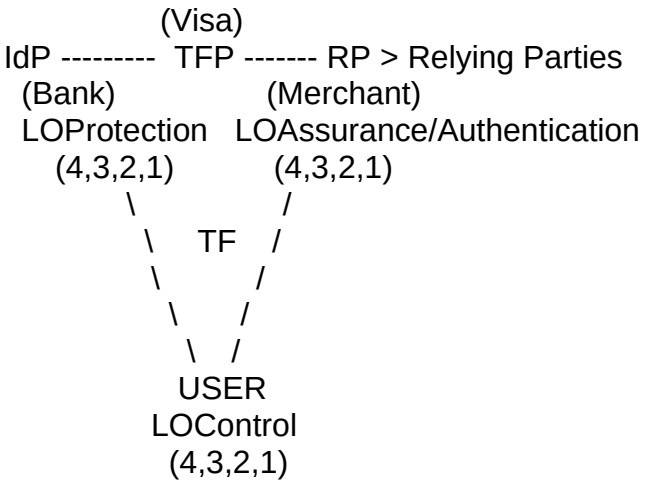
URL: http://hub.personaldataecosystem.org/wagn/Trust_Frameworks_Connect_me

Discussion Notes:

Trust Frameworks - set of contracts
- everyone agrees to act in a reliable way across systems

Open Identity Exchange
- Trust frameworks are socialized
- Listed on Open Identity Exchange

Matching of Trust Frameworks and Services



Example: using a gov ID to enter a bar = framework with stability

Tools and Rules
Define: 4 levels of risk (damage to property or person, how severe, how likely)
Resolve: 4 levels of credentials to match levels of risk

We need 1,2,3,4, levels for all players
Unpack privacy and align needs/mutual dependencies (IdP, RP, User)

4th Party VRM
- Exchanges of information (first party - personal data, 2nd Party RP wants data, 3rd Party wants to help with exchange)
- 4th Party Trust Framework

Level of authentication for the user

IdP is receiving assurance from RP
Data flow with LOP elements that promise both transfer data or not - the "Duty Infrastructure" which keeps people acts reliably

Bubble of risk reduction
Base of rule set of all communities

Functional control that a person can deliver their own privacy
Global competitive issue for Government (who is going to be the Swiss Bank of Data?)

Privacy policies are site specific, the Trust Framework brings this to a higher plain -> SupraPrivacy
Level of agreements that make people comfortable

USER - LOControl
Innovation
look at least bad old stuff
took the word "privacy" out, using "control" instead
unpack privacy (OIX.org website, append 7, fair information practice comparison tool - 1973)
- no collection without person's knowledge
- be able to access and correct info
- notice and consent
- FIPPS tool re-sorted by topic (Scott David) 40 yrs of data on what users think is fair
- next step to cross with Bills of Rights
- Brandeis Privacy - use of name or image without your permission = invasion of privacy (tort)
- Torts
- Constitutions

4 Torts - Civil Harms (So we can categorize harms)
1. Misappropriation - infringement (commercial use)
2. Trespass - Intrusion (peeping Tom)
3. False light - Defamation (libel, slander)

---

4. Big Honking => Publication of private facts

These all cover a lot of territory which is an intrusion to person
First Amendment
Fourth Amendment - intrusion of home

Information crossing the boundaries
- API becomes digital membrane
- The membrane is no longer a "place"
- "Duty" membrane
- Within this bubble, the risk is reduced
- Technology makes the software reliable
- Legal VRM - wrapper

3 Phases
- Entropy confusion
- Creative commons
- Negotiating relationships - give people tools to make things more reliable to make negotiations
-
- Create a Trustable Process

- Clear roles for all parties, then 2 legs of triangles are covered, then RPs might see the value > Trillions of transactions

Scott David is not boring

W.E.F. Personal Data Project Overview and Engagement

Session: 4C

Convener: William Hoffman and Rico Oyola

URL:

http://hub.personaldataecosystem.org/wagn/W_E_F_Personal_Data_Project_Overview_and_Engagement

Rethinking Personal Data Phase II

Build a community of partners interested in addressing the interrelated and complex cultural, business, technology and policy trends shaping the personal data ecosystem. Develop a user-centric set of recommendations for individuals, private enterprise and policy-makers:

1. Innovate around user-centricity and trust
2. Define global principles
3. Strengthen the dialogue between regulators and the private sector
4. Focus on interoperability and open standards
5. Continually share knowledge

April 6-7 (Palo Alto) Personal Data 2.0

STL Partners, Deep-Dive Workshop
Mix industry (mostly telecom) experts
Discuss Role of Telco as IsP
Ignite/build steering committee
Establish 9 month research collection methodology

May 3-5 (Mountain View) 12th Internet Identity Workshop

Mix industry identity experts (internet co.)
Data Principles
Interoperability & Standards
2) Define Global Principles -develop draft for circulation
4) Focus on Interoperability & Open Standards

May 13 (London) New Digital Economics

Mix industry (mostly telecom) experts
Discuss Role of Telco as IsP
4) Focus on Interoperability & Open Standards

June 9 (Vienna) World Economic Forum Annual Meeting on Europe & Central Asia Personal Data Day Workshop

Mix industry CEOs, Sr Executives & Ministers
Support Global trust frameworks
Support pilot-field trials

Sept16 (Dalia) World Economic Forum

Innovation, Cyber-security (personal data)
Identify Chinese data/identity themes
Oct (New York) World Economic Forum Private Session
Sr strategy, Innovation, Policy executives.
1) Innovate around user-centricity and trust
2) Define global principles
3) Strengthen the dialogue between regulators and the private sector
4) Focus on interoperability and open standards

December

World Economic Forum, STL Partners, MIT
Harvard Berkman Center, IIW,
Industry Partnership experts, Knowledge partner, issue experts
Draft Finalize Report
Impact Disseminate Plan

January 28, 2011 (Davos) World Economic Forum

Mix industry CEOs, Sr Executives & Ministers
Rethinking Personal Data CEO Session
Finalize report launch
Annual Meeting Private & Public Sessions

Who Do You Trust?

Session: 1B

Convener:

Notes Taker: Kathy Tavitian

URL:http://hub.personaldataecosystem.org/wagn/Who_Do_You_Trust

Discussion Notes:

- There is a need for a transparent infrastructure to manage personal data.
- Who/what do people trust to manage their personal data? Who should be the "point of integration" for personal data?
- There was discussion on the idea of a "custodian" with fiduciary responsibility to an individual. In this context, the term "fiduciary" means the custodian would put your interest before their own with respect to your data.
- There was discussion on the idea of trust generally versus "contracted" trust.
- Trust includes confidence in the reliability of the data custodian and in the willingness of the data custodian to protect your data.
- It is important for any personal data management systems to have scalability.
- It may be that there are different types of entities that are appropriate for different kinds of data; however, that still does not eliminate the silo problem.
- Some of the rights we might want for third-party management of personal data include: auditability, portability, ability to correct errors, ability to get fair value for any use of the data, transparency, and control over any disclosure or transfer of the personal data.
- Today multiple entities are collecting personal data with little to no regulation.
- One challenge with respect to monetizing your personal data is that the current paradigm of privacy law does not extend to any concept of property rights.
- Issues with respect to personal data collection: privacy rights, human rights, property rights.

There was discussion on VRM and the four-party system.

- First party is the customer.
- Second party is the vendor.
- Third party is vendor-driven and on the vendor's side.
- Fourth party is customer-driven and on the customer's side (e.g., personal data custodian)

There was reference to a book by John Hagel (The Power of Pull)

In many cultures, there would be very little trust of government in the personal data custodian role.

There is also very little trust of entities like Facebook, Google, etc. who are not regulated with respect to use of personal data.

There was discussion about whether telcos would be a good fourth party for personal data management. One advantage is that they are already heavily regulated, and they already manage a large amount of personal data.

It is hard to trust any company to manage your personal data unless they are regulated in a way that protects your data. It helps to have trustworthy people on the Board (e.g, Oprah), but management and the Board can change.

One person said he would not trust any company to manage his personal data. He argued for a open source distributed platform for personal data management – your own curated version of you.

18-24 year-old individuals, when polled by Intel, said they trust Facebook, Google, PayPal, etc. However, it is not clear that they understand the lack of regulation governing these entities, because, when specifically asked, these young people indicated that they want control over their data (although they are often willing to trade off this control for things of value to them). Part of the reason for this willingness to trade off control for value is due to behavioral economics: people underestimate the risks and overestimate the value of offers made to them.

The group leader noted that the discussion was focusing on two different areas: the rights sphere (what are the risks and protections needed?) and the commerce sphere (what are the opportunities to monetize data?)

Mark Davis noted that "data has to flow." That is a given, it is just how and who gets control and benefits from that data flow when it comes to personal data.

It was noted that the Internet is essentially a giant copying machine (digital data can be reproduced ad infinitum).

One individual noted that there will be market power and political power in harnessing personal data and enabling connections between people with respect to this data. He speculated that this would be an issue in the next presidential election.

In order to have this collective market and political power from connecting individuals via their data, people need knowledge and understanding of what can happen to their data. Today most people don't have a good understanding of how their personal data is collected and used.

Companies have been collecting personal data for a long time, and they consider it proprietary information.

In order to raise awareness about the issues, which would be more compelling: a fear-based message on the risks of not controlling your own data, or a message focusing on the advantages of user centricity and control?

A big issue is that the legal system has not adopted the concept of digital personhood. Could each individual incorporate in order to formalize their digital personhood?

Could everybody have their own server and domain name that they own?

Human "API" could be proxied in the Cloud.

A fourth party can simplify the personal data management for individuals (like a stock broker). Would people pay for a personal data bank exchange?

There was discussion on the best way to store personal data. Is a data vault the right answer? There are risks with this approach because a breach could be catastrophic.

In any case there needs to be some uniformity in the rules and process in order to attain interoperability. Examples were given of SMS and railroads in how this can evolve. This can evolve through regulation or market forces (resulting in an economic incentive to cooperate), or a combination of both. A regulated industry increases trust in the process.

Often no action is taken until a crisis has occurred (e.g., Securities Acts of 1933 and 1934)

Technology has advanced a lot, and regulation has not kept up with the changes. For example, ECPA (the Electronic Communications Privacy Act) was passed before the web existed and has not been updated.

There was discussion about the fact that this is an international issue, not just a U.S. issue. U.S. regulation alone, therefore, will not be enough. As an example, it was noted that the recent Epsilon data hacker was probably not a U.S. citizen and may not be governed by U.S. laws and regulations.

Strengthen the dialogue between regulators and the private sector

Modeling and Monitoring the Emerging Personal Data Ecosystem

Session: 4-G

Convener: Verna Alee, Stephen Young

Session Notes:

There was a discussion about different models around seeing the system including Value Network Mapping. They invited others to share models and Analysis.

NSTIC National Strategy for Trusted Identities in Cyberspace

Session: 5G

Convener: Kaliya Hamlin

Notes-taker: Natalie Fonseca

URL: http://hub.personaldataecosystem.org/wagn/NSTIC

Other Attendees : Eric Sachs, Phil Wolff, Didn't get other names

Discussion notes:

NSTIC stands for National Strategy on Trusted Identities in Cyberspace. The website is: http://www.nist.gov/nstic.

Kaliya started with the problem that govt is trying to solve with NSTIC…many govt departments are responsible for and provide services to millions and millions of citizens. Right now, there's almost no way to allow the citizen to log in to the govt website to view their information. For example, you'd like to know the IRS is going to show you your tax record only to you and not to someone else.

She said there are two choices to solve the problem that are both bad:

- Govt issues some kind of token to citizen and all their accounts get linked to this one government issued ID. National digital identifier. Culturally unacceptable in the U.S.

- Each one of the agencies issues its own credential to the user. Too expensive.

If these approaches are bad, why not leverage fact that people has verified credentials from private sector entities? So part of building ecosystem is I have a choice about whom I might ask for a token to prove who I am and that this might be done in a way that it's an open system. Examples would be banks, schools, utilities, etc.

Questions that arise:

- What about liability? What does safe harbor look like?
- Where does trust come from between govt agencies and institutions

Open Identity Exchange (OIX) creates a policy repository, aka "trust frameworks." Examples of trust framework providers:– ICAM, Email Assertions, PBS kids, OCLC (library network).

Identity Providers and Relying Parties (where you're hosting data) – example of Google as an IdP and IRS as a RP. Creates a need for auditors.

Eric Sachs pointed out that NSTIC was moved to the Department of Commerce (from the Department of Defense) to make it more clear that the goal is actually to spur more online commerce. He said there has been confusion about that.

The OIX helps connect those in the marketplace: trust framework, auditor, IdP and RP. Kaliya pointed out that Scott David of OIX wants to add LOC (Level of Control).

It's an interesting model in that it allows a multitude of orgs to create trust frameworks. Don Thibeau is chair of OIX. Province of British Columbia is looking at using OIX.

Eric mentioned that there will be a NSTIC meeting on April 15th where there will be demos by different vendors, and they are all smart card vendors except Google (he'll be participating).

The Obama administration looked into the biggest threats on the Internet that are preventing more ecommerce – and they found out the biggest threats were malware, password reuse, phishing. Govt put a lot of pressure on industries to improve password reuse, to avoid regulation.

Where some of the collision is happening now is b/c there are a lot of laws about how people interact with govt agencies (a lot of website operators start to read documentation that was designed for govt websites; e.g. govt websites can't use cookies).

Also discussed different levels of assurance with ICAM and two-factor authentication.

What's the role of the Telco's in "Personal Data Economy"

Session: 2B

Convener: Owen Tripp, COO and Co-Founder, Reputation.com

Notes taker: Simon Torrance, CEO, STL Partners/Telco 2.0 Initiative

URL:
http://hub.personaldataecosystem.org/wagn/What_s_the_role_of_the_Telco_s_in_Personal_Data_Economy

Discussion Notes:
Question 1: Are telcos in a good position to act as custodians of consumers' personal data? Answer: yes.
Question 2: Should the main motive be new revenues or brand enhancement and churn reduction? Answer: both (potentially).

Notes:

- Telcos are held to very high standards of probity by regulators (Google and Facebook aren't). This 'brand value' can be leveraged. Recent consumer surveys show that telcos are trusted to look after consumers' data more highly than other organisations (even government). Telcos have a bigger opportunity to become 'trust brokers' in the digital economy.
- Important to understand what we mean by consumers' personal data: where it comes from, how it's used, what the potential applications can be (cf. Phil Laidler's framework presented at the Telco 2.0 session on the first day of the event)
- Important to offer tangible benefits to consumers in return for accessing their data or exploitation of it (cf. the Promo Tonos case study, Orange Spain).
- Important to work in spaces that people already understand and accept – eg. 'better direct mail'.
- Combining data from different sources (eg. credit ratings + location) can be more valuable, but creates problems around policy.
- Different consumer demographic groups have different needs and attitudes: some people are more interested in protecting their data, others are interested in cheaper services (that result from sharing their data).
- Digital persistence: how long should consumers' personal data be kept? An issue to bear in mind.
- Laws significantly lag both rapid technological development and consumer attitudes.

Focus on interoperability and open standards

Practical Personalization

Session: 3G
Convener: Phil Windley
Notes Taker: Phil Windley

Phil presented slides along with talking about and decided to share the first chapter of his forthcoming book The Live Web.

The Live Web

The Web is an amazing place. In less than two decades it has completely transformed commerce, banking, travel, the news business, and even social interactions. Whole industries have been created, some have been destroyed, and others have been changed to where they are unrecognizable. No one, of course, could have predicted this in 1993. From static Web pages, through simple forms, to today's beautiful, interactive Web sites, the evolution of the Web has been dramatic.

Recently some of the most compelling parts of the Web aren't even sites at all, but services like Twitter or Foursquare. Certainly, these services *have* Web sites, but the site isn't their only or even their primary means of interaction. Developments like these are the vanguard of a completely new way to use the Web. In fact, some quip that with social networking, the Web has finally achieved its purpose in being.

But with all the mind-boggling changes that have happened in the last 15 years, I believe that the biggest changes are yet to come. The pace of change is accelerating. The World Wide Web is still growing and as we look around, we will find the seeds of technology that, if nurtured, will make the Web useful in ways we can barely imagine.

The Web of the future will link together the devices you use, your online personas, and even things we don't think of as being "on the Web" now like your car and your house. This cooperating, loosely coupled mesh will work to achieve results important to you—your purpose—in context and at the right time. The current world of Web sites and services is inadequate for building this future Web—what's being called "the Live Web."

The current Web certainly won't go away; it's too valuable. But the Live Web will be built along side of it. This book will examine the concepts and techniques that will take the Web forward into a purpose-based future, discuss gaps and ways to fill them, and introduce architectures and systems for building tomorrow's Live Web applications now.

Bad News from Your Radiologist

To glimpse the potential of the Live Web, imagine you've just received bad news from your radiologist: your recent MRI reveals you have two ruptured disks. Think of all that has to be done:

1. Choose a good orthopedic surgeon who is nearby and works with your insurance provider.
2. Schedule an appointment for an initial consultation and follow-on visits.
3. Arrange to transfer medical history information to the new doctor.
4. Make payments for services.

Today we'd be forced to approach this task in an ad hoc, out-of-context way. Sure, we would use online resources to make this easier, but we'd still be the primary organizer adding structure and meaning to the tasks.

Instead, imagine the following scenario:

1. An email arrives from the patient management system used by your radiologist telling you the results of your MRI.
2. An application that you control processes the email, extracts relevant tasks, and places them on your todo list.
3. When you're ready to complete the first task, picking a surgeon, a service provider selector application is engaged to show surgeons meeting criteria you have established (e.g. must be nearby and within my insurance plan) as well as those set by the radiologist (e.g. practice type). You use the selector application to review and choose a surgeon.
4. Your calendar begins the task of scheduling an appointment, giving you a selection of times that are free on both your and the surgeon's calendars.
5. You pick an appointment time and receive confirmation from the doctor's office that your appointment has been scheduled.
6. The surgeon's office sends a request for medical history and an application that manages your personal data gives you the opportunity to authorize the transfer and select what information will be sent.

Having even this small set of tasks fulfilled automatically seems like science fiction to most people, but we'll construct prototype versions of all the services necessary for this scenario in Chapter 14. Creating an architecture that is flexible enough to carry out these tasks without being purpose-built and brittle isn't impossible but it does require some new thinking. You'll see that while we'll make a few simplifying assumptions, nothing about the applications we write will be outside the bounds of what could be done today, using tools and techniques available right now.

The Live Web

The "Live Web" is a recognition that time matters online. The term "live" was applied to search early on to describe search results that included recent, rather than simply relevant, results. Real-time search results have recently gained significant traction with Google, Twitter, Facebook, and others providing the ability to search the most recent tweets or status updates.

I'm convinced that the Live Web goes well beyond search—as important as that is—to inform a completely new way to viewing online interactions. The Live Web doesn't replace the old static Web but will grow and exist in symbiosis with it.

The Live Web isn't just about speed and timeliness—it's about context. On the static Web, each site comes with its own context. As we move from site to site, the context changes and the only continuity is what we can maintain in our heads. On the Live Web, we bring our context with us and applications weave the sites, data, and services from the static Web into a single, purpose-based, living tapestry. The service becomes personal, not in some superficial "Hello Phillip Windley!" way, but with respect to what I'm interested in right now—my context.

The Live Web uses end points that represent my context. Endpoints maintain the continuity of context as I move from place to place. Events at those end points can trigger rules on my behalf that can be written by anybody and executed by engines that can be located anywhere. The Live Web connects with the static Web anywhere anybody likes, but is not a subset of the static Web. Nor is it subject to the limitations of the client-server model.

In the static web, the default original state is one in which all sites are independent and all visitors are dependent on them leading to its classic silo-like nature. This is a function of the old client-server model; individual users have no independent status. One symptom of this is boilerplate "agreements" in which the dominant party (the site owner) can say or change what they please, while the submissive party (the visitor) has to submit completely.

In contrast, on the Live Web, the default original state for end points is independence. Choices about dependence start there. Endpoints work on behalf of individuals. When the individual user is capable of bringing her own context to an engagement, interaction is richer with greater possibilities. On the Live Web, many more things can be done, and much more wealth can be created, than on the old static Web alone.

While the old world of individual sites and services won't go away—after all, we need their infrastructure—the Live Web will live along side them, making them richer and more useful to all of us.

Live Web Examples

The health-care scenario above is a good example of the Live Web in action. Here are a few more simple examples:

Figure 1: Hoverme shows your friends social profiles.

Hoverme (see Figure 1) is an application that is installed as a browser plug-in. Once installed, as you hover over a friend's name on Twitter or Facebook, a pane pops up to show that same friend on other social networks. Data from multiple sites is being mashed up in real-time as the user needs it and is placed in context where the user will see it.

Who Am I Talking To is an application that runs on Salesforce.com. The application uses data from your Salesforce account to mashup email, phones, and the Web so that people can only call in if they've sent their request first via email. Once the call does come in, the application brings up their profile data on Salesforce.com automatically.

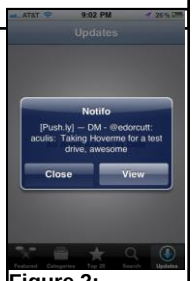


Figure 2: Notification on an iPhone

Notifio and Push.ly (see Figure 2) are applications that combine to push instant notifications of tweets that contain certain words to your smartphone's notification system.

In each of these cases data is being used in real-time and in a contextually relevant way to bring value to the user. None of these examples are Web sites, but are rather applications that mashup various Web data, sites, and services in a way that is meaningful to the user.

Components of the Live Web

Realizing the Live Web requires several important pieces that we can think of as its foundation (see Figure 3):

1. **Programmatic access to data.** As we'll see shortly, access to data from various Web sites happens through something called an API. This is one of the hottest online trends of the last several years and provides real-time information that makes the Web exciting and interactive.
2. **Endpoints that work on behalf of users.** Endpoints are any application under the user's control that can manage context, signal relevant events, and respond to directives from applications.
3. **Event processors that respond to events.** These special applications work in concert with endpoints and APIs to orchestrate and automate the actions that are appropriate for a given circumstance.
4. **Context**. Context arises from the interactions of endpoints, event processors, and programmatic data. Context is generated and regenerated, continually refreshed as things change. Context is supported by strong user identity, user-control, and privacy controls. Useful systems of APIs, endpoints, and application will not only be capable of managing and using this context information, but also allow developers to easily incorporate relevant context information into applications. Context provides the means of filtering the vast amount of data and events to those that are relevant to a given person.

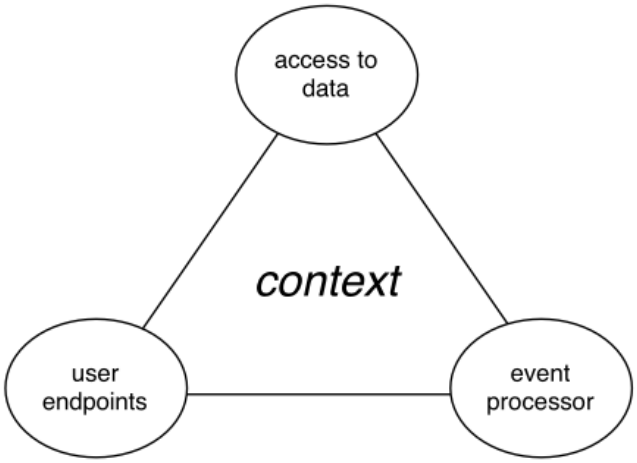


Figure 3: The Live Web Ecosystem

The Live Web comprises these four elements. Building applications in the Live Web requires that we correlate events from multiple endpoints and data from multiple services using context that is meaningful to the user. Methods, architectures, and technologies for contextually correlated events and data will be the primary topic of this book.

The Evolution of the Web

Since the Web was introduced the world in 1993, it has undergone many significant changes. Table 1 gives a view of the evolution of the Web from the standpoint of several important characteristics:

- how identity information is managed,
- how data is represented,
- who controls the data, and
- what essential services are offered.

Document Web

The "document-based Web" was the very first Web, where people put up Web sites that were really nothing more than their company brochures reformatted in HTML. Identity was non-existent, leading to the famous New Yorker cartoon with the caption "On the Internet, no one knows you're a dog." Any data that made it's way out of the brochures and into HTML wasn't structured in a way that made it useful. The site owner controlled the data, such as it was, and the primary service was returning a document for rendering by the user's browser.

		Identity	Data	Control	Services
Static Web	Document Web	you're a dog	Stuck in unstructured documents	site owner	document return
	Web 1.0	cookies & siloed credentials	unstructured documents & proprietary DB with dynamic queries	site owner	document return & siloed CRUD
	Web 2.0	cookies & shared credentials via Oauth	structured documents, proprietary DBs & APIs with dynamic queries	site, app owners, or users	document return, siloed CRUD, apps, & API response
Live Web		personal data under the control of the user	semantic documents, proprietary DBs & APIs, public/private data streams, static & dynamic queries	data originator (usually users) with granular control	document return, CRUD, apps, API response, semantic mappings, & event handling

Table 1: The evolution of the Web

Web 1.0

We didn't call it Web 1.0 at the time, but the document-based Web quickly gave way to one based on primitive forms and pages with a few dynamic features. Web sites were still primarily collections of static pages with a sprinkling of dynamically generated content. Querying the site's database was the primary mans of interaction. Logins, supported by cookie-backed sessions first began to appear so that a site could be "customized." The site owner still controlled the data, although the concept of privacy was becoming an issue. The primary service is still document returns, but some site-specific database operations are appearing.

Web 2.0

Web 2.0 is the era of the interactive Web application. Although the underlying basis of the Web hasn't changed, developers are creating increasingly sophisticated Web applications rather than mere Web sites. Identity is starting to use cross-site credentials and data sharing is possible through protocols like OAuth. API[1] calls return structured data and the user might use the Web or a purpose-built client application. The emergence of more sophisticated identity regimens has started to put users in charge of their own data and how it is shared.

Live Web

I've placed each of these previous eras in the evolution of the Web in what I call the "static" Web. Some might bridle at the term since the rate of evolution and the levels of interactivity seem anything but static. But all these eras have something in common: they are based on Web servers that hold relatively static collections of documents or data. When users interact, they query, add, or update this data.

In contrast, the Live Web is characterized by dynamic streams of data, tied closely to the user's identity and typically under their control. The data comes in a variety of forms. The static Web is still there, underlying the Live Web. Data is accessible programmatically and in real-time. The following sections will describe these characteristics in more detail.

The Cloud and Programmatic Access to Data

More and more companies are putting the data and services that drive their Web sites online using an API. Euphemistically the result of this move has been called "the cloud." There are good reasons why cloud-based data and services is gaining traction: Cloud-based services are more accessible, more convenient, and cheaper than equivalent services delivered using more traditional means.

As I write this in late 2010, Programmableweb.com, a directory of online APIs, lists over 2300 APIs in its index. This number is certain to grow. The list includes APIs for searching, financial services, blogging, ad networks, dating, email, government, security shopping, and so on. Some are free and others charge money. Some are personalized (like my Twitter friend feed) and others are

[1] API originally stood for "Application Programming Interface" and applied to services provided by operating systems to applications, but hardly anyone thinks of that anymore. The acronym has more generally come to apply to any interface to programs, data, or services that a program can use to interact with them.

general information (like the Google news feed). APIs are the unit of programming on the Web—similar to libraries in traditional applications.

The move to APIs and the programmatic access to data that they represent is nothing short of a tectonic shift in how the Web works, realigning the forces that drive the Web and making new applications possible. As with most shifts, the move to APIs seems sudden, but it's been building for years as developers struggled to create programs that "mashed up" information and services from multiple Web sites.

At first glance, you may wonder why companies would be willing to give programmatic access to their data through APIs. Further analysis shows that there's a good business reason for it. Consider the diagram (adapted from Dion Hitchcliffe[2]) in Figure 4.

More and more traffic to Web applications is coming through the API because APIs let other programs use your Web application. This extends your online reach significantly. Sam Ramji said, "For successful companies, 80% of traffic will be coming from beyond the browser."[3] Giving your Web application an API allows other applications to carry your service to their users.

APIs are one way to get other smart people to help you build your product and service into something people find valuable. Rather than "build it and they will come," APIs are a way to have others "come and help build it." As the creator of a Web application, we can never come up with all the interesting ways people may find for using it. Providing an API allows them to extend your application in ways that meets their needs and, in the process, do the same thing for others. Later we'll see how concepts like webhooks extend this idea even further.

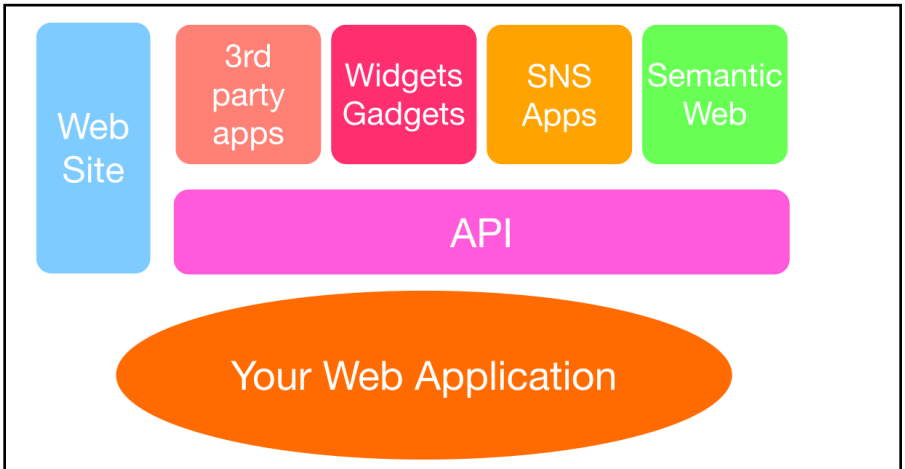


Figure 4: APIs give your Web application reach beyond your Web site

Indirect models will allow the market for Web applications to scale well beyond what is possible with just a Web site. The hoopla over cloud computing isn't really about Amazon, Rackspace and the infrastructure plays you've heard about. What makes cloud computing interesting is the move to the API by thousands of companies. This wealth of APIs is a tectonic shift in how the online world operates and portends the changes that will comprise the Real-Time Web.

The Real-Time Web

But mere APIs aren't enough. To see why, imagine two scenarios. In the first, your teenage daughter is out with friends. Her curfew is midnight but it's 12:20am and you haven't heard a word. You're worried, imagining an accident on the freeway, or worse. You're calling her cell, her friends, and trying to keep calm.

In the second scenario, your daughter is again out with friends, but this time, a few minutes before twelve, you get a call that goes something like "Hi Mom. I'm going to be 15 minutes late…the movie ran long."

One scenario is filled with hassle and anxiety—the other with convenience and tranquility. There's no doubt which scenario we'd rather be in. And yet, online, we're rarely in a situation where a service anticipates our needs and meets them without prodding on our part. The location metaphor of the static Web puts in the mode of "seek and find" or "go and get."

When an API merely responds to requests, it's like a program that only accepts input, but can never send its output to another application unless it's asked first. We call such APIs *half-duplex APIs* and readers familiar with the Web and the underlying client-server model will recognize the roots of half-duplex APIs in the foundational technologies of the Web.

In the early days of the Web, all that mattered were domain names and Web pages—brochure-ware, I called it. Later the name Web 2.0 embodied the idea of interactive Web sites where users could actually do something beyond filling out simple forms. The earliest interactive Web services were ecommerce tools in the late 1990's. Later the idea of interactive Web sites extended to all kinds of services from finance to document editing.

---

[2] Dion Hinchcliffe, Open APIs reach new high water mark as the Web evolves. November 4, 2008 http://www.zdnet.com/blog/hinchcliffe/open-apis-reach-new-high-water-mark-as-the-web-evolves/215

[3] Sam Ramji, Darwin's Finches, 20th Century Business, and APIs. May 10, 2010 http://www.readwriteweb.com/cloud/2010/05/how-darwins-finches-and-apis-a.php

Interactive Web services have several problems. First, they tend to be silos that interact with other sites and services in only the most rudimentary ways. If the service doesn't do everything that the user wants, there's almost never a way to combine two services together to solve a unique problem. Second, and more problematically, as we've seen, they are built on a request-response model that requires the user to initiate each interaction.

In contrast, some web applications are beginning to push information to users. This is not only more convenient for users, but creates data streams that other tools can combine to solve problems that require mashing up information from multiple sources. We call the set of technologies and practices that enable users to receive information as soon as it is published by its authors the real-time Web.[4]

You don't have go any further than the Facebook, Twitter, or Foursquare apps on your iPhone to see the real-time Web in action. These services aren't just interactive Web sites, but are creating streams of data about the people you follow. The stream of tweets from my friends is available to me in a variety of places without me needing to visit any particular Web site.

The real-time Web won't replace the interactive Web—we'll always need Web sites—but the real-time Web will augment it in important ways.

Information Overload & Context

At first glance, having services push information to you rather than having to go from place to place gathering it sounds good. But anyone who's tried to follow hundreds of people with an online service like Facebook or Twitter knows that you can't see the constant stream of updates and get anything else done. We won't make people happy if the real-time Web merely fills up their inbox with hundreds of notifications per day.

New York University Professor Clay Shirky has famously quipped that there's no such thing as information overload only filter failure[5]. Even small increases in spam volume, for example, can lead to the perception that volume has doubled or tripled if they cause our spam filters to fail. Technology allows things to happen at scales that overwhelm the social filters we have previously had in place. The failure of filters that have previously worked well cause us to feel we're drowning in information when in fact, we need more and better filters.

This information overload causes real anxiety. Marry Hodder, who's been working on the real-time search problem for years has conducted a study that shows why this is so. She says[6]:

> "I did a study in 2002, which I repeated in 2004 and again…in 2008. I asked users to track their online information intake for one week. There were only 30 people in each study, chosen randomly from Craiglist ads, but what I found across each group of 30 was that the average time spent online with news and information sites was 1.25 hours in 2002, 1.85 hours in 2004 and 2.45 hours in 2008. These people are not in Silicon Valley, but they do all have broadband at home and live in the US. Every one of them reported some level anxiety over the amount of data they felt they needed to take in in order to feel informed. They often dealt with it by increasing the time they took to stay informed. They didn't know that better filters might actually reduce their anxiety."

Shirky and Hodders's insights tell us why we need more than just real-time APIs. A real-time, API-driven Web threatens to overwhelm us with information, and just filtering, as people traditionally think about it, won't be enough. Instead, the future will require that we react and respond to all this data in ways that are meaningful to people—in context.

Context provides the means of making sense of all this streaming data. When we put data in context, we can understand what it means and react to it more intelligently. Don't mistake this for some Utopian, artificial intelligence-laden future Web. The technologies for doing this are within our reach now.

Turning the Web Inside Out

The Live Web is a radical shift in how people use the Internet—rather than simply viewing static pages, or even interacting with a Web site, the Live Web uses dynamic streams of information to present contextual, relevant experiences to user.

These dynamic streams of information include such diverse data flows as Twitter streams and Facebook news to RSS feeds of product recalls. Many of these information streams are enabled and supported by APIs since programmatic access to the data is critical to its reuse in various guises.

Information reuse is a major premise of the Live Web since much of the information available is not nearly as interesting by itself as it is in combination with other data streams. Further, this reuse is often highly personal since what you want from the information stream and how you want it mixed with other data is different from what I want.

---

[4] Real-time web, Wikipedia, http://en.wikipedia.org/wiki/Real-time_web, referenced December 17, 2010.

[5] Clay Shirky, It's Not Information Overload. It's Filter Failure. Web 2.0 Expo NY. http://web2expo.blip.tv/file/1277460/

[6] Marry Hodder. Live Web, Real Time . . . Call It What You Will, It's Gonna Take A While To Get It. http://techcrunch.com/2009/06/30/live-web-real-time-call-it-what-you-will-its-gonna-take-a-while-to-get-it/

Making streams "real-time" skews design to one favoring interrupts over polling. Protocols like PubSubHubBub and services like push.ly are created with the explicit purpose of taking polling-based technologies like RSS and giving them an interrupt-driven façade.

Dynamic Queries, Static Data

In the mid 1990's the Web had developed enough that people were starting to build form-based, data-driven Web sites. MySQL was on the rise, but connecting to databases and building Web pages from the results was still a complex programming task. Out of this environment, PHP emerged as the tool of choice for building interactive Web sites. Later Ruby on Rails took this model to the extreme and became the de facto standard for building interactive Web applications.

The modern Web is built on the notion of a Web browser making a request of a Web server and getting a response that it renders and shows to the user as shown in Figure 5.

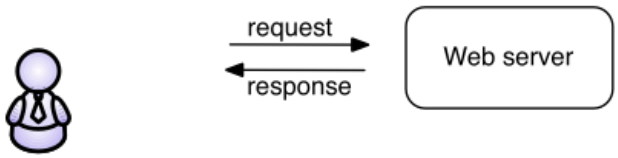


Figure 5: Web browsers and servers

For interactive, data-driven Web applications we can expand the Web server box to show a standard architecture.

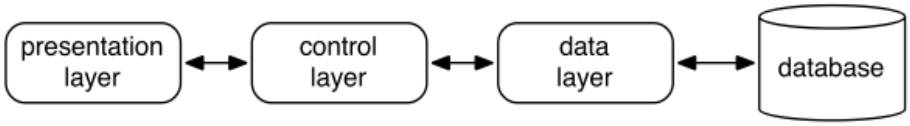


Figure 6: Data-driven Web application architecture

This architecture shown in Figure 6, variously called an n-tier or MVC (for model-view-controller) architecture, is at the heart of all well-designed interactive Web applications. The architecture is performant, while providing good separation of concerns to ease the program construction and maintenance efforts.

One of the hallmarks of this architecture is the database. The Web application is written to create and modify entries in the database. The Web application is driven through queries against the database. PHP, Ruby on Rails, and other Web application languages are designed to make the task of managing the data and dynamically constructing queries against it as easy as possible.

We can characterize this pattern as "dynamic queries against static data" or DQSD.

Dynamic Data, Static Queries

The Live Web demands the dual of this model. In the Live Web, the data is dynamic. We view it as streams or rivers of information flowing past us, rather than as a pool of data that we collect and control.

Making use of these information flows requires that we recognize patterns in the flows as well as between them. Here are some examples:

> *Tell me when someone mentions my company on Twitter.*
>
> *Tell me when someone checks into the same place I am.*
>
> *Tell me when one of my Facebook friends posts an update on their blog.*
>
> *Tell me when someone calls my customer service line after visiting the customer service page on my company's Web site.*

These simple patterns represent static queries against the dynamic information streams. We can characterize this pattern as "dynamic data against static queries," or DDSQ. In the DDSQ model, a data source produces a stream of data that is continually changing. The query is run against that stream and only items passing the query are processed further.

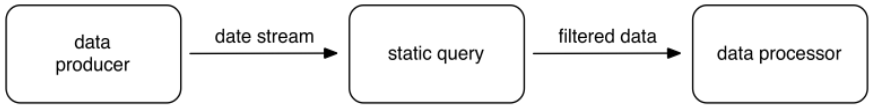


Figure 7: Data producers stream data and queries are run against these streams.

The DDSQ model doesn't replace, but rather complements the DQSD model. After all, the data streams start somewhere and most of those starting points are pools of data managed through traditional DQSD Web applications.

There are several well-known examples of this concept:

- One great example is stock alerts that trigger buy and sell orders when certain events occur. The triggers continuously monitor the stream of data about stocks you care about.
- Another example is Google alerts. When you create a Google alert you're giving Google a standing, static query that it continuously runs over the stream of new stuff that it sees. When there's match, you get an email.
- Twitter searches are similar. Once you've done the search, Twitter continues to apply it to the stream of tweets looking for matches.

On each of these examples static queries established once and infrequently changed are continuously run against a stream of dynamic data.

The example queries given above all started with "Tell me…" but we'll see that we can do much more than merely notify people of certain patterns and that by programmatically reacting to patterns in dynamic information streams, we can create compelling, sophisticated applications.

The request-response model of a typical Web site puts the focus on the server and location for a simple reason: to make the request, your client has to know where to go. On the other hand, forgetting about location and moving to streams of events frees us from server silos. Going a step further and envisioning these streams of events as being individualized for a given person or entity allows us to concentrate on the user's context and purpose.

Putting static queries over dynamic data streams to work for the individual will require better abstractions for modeling these streams of data and writing programs that work against them in context. Modeling them as events gives us a way to watch for scenarios that matter to the user. Without an event-like abstraction, this is hard and expensive to pull off.

Events

Events are a natural way to think about online interactions. When someone sends a tweet, updates their Facebook page, or checks-into a location on Foursquare, this generates an event. Further, they're a natural way to pull systems into the online world. Consider, for example, a point-of-sale system that could generate events when a product is rung-up, a sale is totaled, or payment is tendered.

An event is a formal notification of an occurrence of interest. A given event happens at a particular time, contains information about what happened, and might contain parameters to further distinguish the event.

Viewing updates on the real-time Web as events gives us a model for classifying and processing them. For example, a tweet can be seen as a "Twitter event" that occurs at a given time and has required parameters like "message" and "user " with optional parameters such as "client used."

Each event is a member of a certain event type that defines certain required and optional parameters through an event schema. While a Twitter event has a required parameter representing the message of the tweet, a Foursquare event might have a required location parameter with an optional message parameter.

But an event model goes beyond the darlings of the real-time Web like Twitter, Facebook, and Foursquare to encompass other online activities. Updates in an Atom or RSS feed, for example, can be thought of as events—especially when viewed through the lens of protocol extensions like PubSubHubBub.

The Evented Web

If you crack open a book on complex event processing[7], you'll find descriptions of enterprise-class systems built on message-oriented middleware busses. We will move events out of the enterprise and onto the Web giving as powerful tool for working with real-time data and creating contextual, Live Web applications.

Events, at least abstractly, contrast with the client-server model used by many Web applications in being message-oriented and asynchronous. An event is a signal that something has happened whereas a request is a signal that the requestor wants something to happen. When one system raises an event, it does so by sending some kind of message and doesn't necessarily wait for responses before continuing.

Enterprise event systems use special message busses for transport based on expensive middleware, but that's not realistic on the Web. Consequently, if we're going to use events online, we need to build them on top of the protocol of the Web, HTTP. Fortunately, that work has already begun. The following sections will discuss some of these developments and their importance to building the evented web. In the next chapter, we'll introduce a system for using events to build applications for the Live Web.

---

[7] See Appendix B, Resources, for a list of books and articles on event processing.

Please, Interrupt Me

As we've already seem, extensions to common protocols, such as PubSubHubBub for RSS and Atom, can accomplish the task of turning a formerly request-response-style protocol into one that makes calls to subscribers when relevant updates are posted. This means that the task of determining when a feed has changed can be done with interrupts rather than polling.

PubSubHubBub works using a server, called a *hub*, running software that implements the PubSubHubBub protocol. When a publisher changes it's feed, it pings the hub. The hub signals it's subscribers using an HTTP callback URL (or *webhook*) they've provided. They can then retrieve the changed feed in the usual manner. RSS and Atom feeds can signal that they support PubSubHubBub by putting a `link` tag with the `rel` attribute set to "hub" and an `href` attribute pointing to the hub.

Webhooks

As useful as PubSubHubBub is, it only works for Atom or RSS feeds. The embedded pattern of using URLs to define the interface that one system should use to call another when an event of interest happens—like a new item being placed in a feed—is called a "*webhook*."

Webhooks are HTTP callbacks[8]. The listener or subscriber gives the publisher a URL. When the publisher wishes to signal the subscriber, it makes an HTTP post to the URL. Webhook-enabled Web applications can be extended by the system responding to the webhook call.

For an example of how webhooks work and the problem they solve, consider the Twilio API. Twilio is a cloud-based telephony service. Twilio applications can answer and make phone calls as well as send and receive SMS messages. Twilio issues phone numbers on their service. When someone calls that number, Twilio has to do something. This is where webhooks come in.

Twilio developers have to be able to tell Twilio what to do when it receives a call. Obviously, the more flexible and configurable Twilio is in responding to calls, the more valuable it will be to developers and end-users alike. Twilio could have just let developers configure actions via a menu but menu-based systems can't provide as much flexibility as developer would like. The kind of flexibility Twilio wanted demands the ability to script the system.

Rather than deploying a scripting language on their system that developers use to create applications, Twilio uses webhooks. When developers configure a phone number, they supply a URL that Twilio will call using an HTTP POST whenever the phone number is called. Twilio expects to receive, as a result of the call, instructions (in Twilio's XML command language) about what to do next. Those instructions usually include another webhook to call when Twilio is finished and has received a response from the caller.

In this way, the developer controls Twilio using programs they create on their own Web server rather than Twilio's. Without webhooks Twilio would either be much less powerful or much more complex. With webhooks, it's simple, easy, and powerful. Webhooks show how the evented Web can be created on top of HTTP using a simple pattern. This pattern is critical to building the Live Web.

Building the Live Web

The static and interactive Webs use a model that places servers in the center of the interaction and make people mere appendages to the server through the use of what end up being second-class clients.

The Live Web doesn't displace this model, but envisions a Web that works for and in behalf of people and what matters to them. The Live Web will prove to be more valuable to all of us and more profitable for companies that adapt to it than what has gone before.

Building the Live Web requires a new model that allows us to conceptualize people using the Web to work for them. Events provide a mental abstraction for describing the Live Web and building applications that make it real. Context allows events to be correlated in meaningful ways. The Live Web will be built upon contextually correlated events.

As developers begin to use events in parallel with the interactive Web development techniques they have already mastered, they will build new, exciting applications.

This book presents a new architecture and language for a Web-based event-processing network that is already being used to create the applications of the Live Web. The architecture, called the Kynetx Event Processing Network, is supported by an open-source event processing system[9] that allows contextually aware event processors and handlers to be built that easily take user context into account.

The Live Web links together all our devices and all the information that matters to us into a cooperating, loosely coupled mesh that works to achieve our purpose. The tools, techniques, and models necessary to build the Live Web are available now. Let's get started.

---

[8] Webhooks, Jeff Lindsay. http://wiki.webhooks.org/w/page/13385124/FrontPage referenced December 21, 2010.

[9] See Appendix C, Getting Started with KRL, for information about use either the hosted service or downloading and installing your own node.

Replacing Passwords with Smartphones: protecting personal data

Session: 4A

Convener: Eric Sachs, Google,

Notes-taker(s): Nitin J. Shah, Broadband Mobile Wireless Innovations,

URL:
http://hub.personaldataecosystem.org/wagn/Replacing_Passwords_with_Smartphones_protecting_personal_data

Total 23 attendees:

Tags: Identity, Passwords, Email, Smartphone, Authentication, Regulation

Discussion notes:

SUMMARY: PASSWORDS AND SMARTPHONES: CAN SMARTPHONES OVERCOME THE LIMITATIONS OF PASSWORDS FOR EMAIL?
(CAN TELCO'S AND WEB COMPANIES COOPERATE ON INDENTITY TO MAXIMIZE VALUE? HOW CAN THE GOVERNMENT HELP?)

- *Email providers rely on passwords for login/authentication.*
- *Passwords are not a reliable and robust mechanism: over 1M email accounts per day are hijacked.*
- *SmartPhones present a potentially robust and reliable authentication of a person's ID for email and other uses*
- *Currently OpenID Foundation, OAUTH and other technologies to provide centralized password management and better mechanisms, but require more consumer interaction, and reveal newer issues related to passwords*
- *Can a Smartphone with the right app, and initial set-up process be a more robust and consumer-friendly approach?*
  - *Initial set-up and binding to services requires more effort*
  - *Consumer experience improves*
  - *If the Smartphone is stolen, person will attempt to remedy and react rapidly*
- *Identity: can telcos step up to provide identity?*
- *Will Internet companies bypass the telcos in creating identity frameworks where the Telco is not required (similar to location information)?*
- *Can Web companies and Telcos cooperate to provide a universal identity management framework, across web, mobile?*
- *How can the US Government help?*

Google: webmail providers: hi-jacking of email accounts.
Primary root cause is password hijacking: principally using adult website to persuade consumers to reveal login/password.
Alternative: popular web2.0 website: take over login box at website and try against other sites.

Near 100% hit rate on some other websites as users tend to reuse same passwords across multiple, so consumer behavior and education makes the life of the email hijackers very easy.

One solution: software to help consumers to manage their passwords.
Password managers: onepassword: no password required.
Six weeks between computers, and so if they use onepassword on a new computer they cannot access their email accounts: clumsy: not popular or used.

Webmail providers: cooperated: Open ID Foundation.

Users would (ideally) would all use OpenID, : OpenID foundation: each person would have a single password/ID.

Problem is that now there is a keys to the kingdom problem: concentration of password information in one place. Technical solution actually is OK, but problem changes and gets worse as OpenID gets more popular.

Two-factor authentication: use password and another identifier such as phone number: works,
But makes the consumer experience more compex.

What harm results: 1M/day get their accounts hijacked: < 10% opt in to solve the problem. Also, the same consumers become repeat targets (i.e. consumers do not change behaviro despite apparent harm).

SmartPhones: Google Authenitcator: type password, launch app, and gets a code, and then type that on the site. People with SmartPhones keep service on, so they are live and connected.

Easier to get an SMS or what when traveling internationally? Google authenticator does not use SMS or a browser.

Like an RSA key on a phone (not the same).
Google has a video
Go to google.com/application wesbite: can generate a password that can be used in Outlook, for example so you can create new on-the-fly passwords for when there are non-Google email systems.

Two factor authentication.

Mobile: do not want to deal with a browser. Apps are the better.

Twitter:
Industry has made major strides: plan to develop an approach OpenID industry effort to develop authentication.
OAUTH2, Open ID Connect: lots of technology is available.

Is the password of value and required anymore? Consumer push-back on the use of password.

Scan barcode from your phone: log into the computer, etc.

What about theft of mobile devices?

Levels of assurance: do they need authentication for most of things people people do on the web?

85% of users reuse their passwords across multiple sites, so they are most likely to be compromised.

Bootstrapping problem with two-factor auth:
How to set up the software on a legitimate device: then turn off the protection, and then

Historical perspective:
When websites asked people for email as login:
Or when Google asks for phone numbers: Neither asked for permission from the email or telco providers.

What are options to cooperate with the telcos?
Merging of telco and web and app stacks: sharing of information: who gets paid for what.
Lack of technical discussions between the web and telco companies, despite lots of business discussion.

Phone voicemail app provided by the service provider.
What's in it for the carriers to provide interface to that voicemial to other apps.

Use of Smartphone for identity:
Android:

Initial set-up process of binding the user, phone etc is painful.

How does user tell all their services to let them know they exist.

If you lose your phone: how do you recover? Telco's have done this for years:
If the telco stack allows for the recovery process tied into their identity management, potential for collaboration.

Need thinnes web stack. Should the Sim (TELCO data) card talk to Google (WEB app)?

Could OAUTH handle this? How to send a SAML token and get back an OAUTH token.

If you walk into store, buy phone, can go to a shared computer: send OAUTH to user's email, then create
An OAUTH binding process: phone now has user consent to pass information around.

Need to be sure that the phone can send info to their web properties and services.
Can do per PRPID authentication.

Telco stack: has multiple levels of authentication: hardened processes and procedures.
Telco can see data associated with all this and send you a SAML token, but will not release APIs etc.

Orange etc: SIM, Network, OS, carrier: could integrate all these scenarios: are discussions going on with GSMA? There are a lot of business discussions, but few technical discussions.

NFC: do not need to go thru the network to use the phone as a authentication device, without need for network. Has its own stack: how does the website/Google know to trust that the NFC phone number declared is indeed that device versus

Barcode asks the app on the phone to launch:

How does the barcode declare that it is authenticated?

Application on the phone asks the Google service to make sure that

a) URL and barcode are current and OK: barcode is only valid for < 30 seconds
b) Network can tell the app to launch and to expecting that particular barcode
c) No need for a password

What if someone steals your phone: only 3 to 5% of users put passwords on Android phones

Scale of attack: small for stolen phones, and remedy/mitigation is viable, compared to botnets.

Bruce Snyder: keep passwords written and in your wallet.

One—time bind, secure channel,

First time you get the phone, you first authentication, uses passwords, all new phone, number, etc

Increased adoption of smartphones improves security of accounts.

Two-factor authentication.

Corporate/Personal account could be bound to the same unique identifier.

Value proposition: what's in it for the carrier/telco?

What is globally interesting to the telco's. Identity is a core asset of the telcos: will not reveal the data

Yahoo!/Google: disintermediated location, by doing it on the edge versus:

- Telco's are out of the location business: resistance is useless.
- Telco-based APIs are more robust and viable.
- Argument: location: did telco's miss the boat?

Only one chance to monetize: telco can do expiring tokens: Vs web industry doing QR code and authentication:

- A telco can fundamentally authenticate a user.
- Yahoo!: indentity off the phone: better for the telco's to step up. If the carriers would provide privacy-preserving identification, then there is an efficiency in the economics. Better for the user: need identity that works across the world.

Anonymous Identity:

- Not happening on the web, but happening on my phone.
- It's happening: the analytics are strong enough to know it's you.
- There are cookies on the phone and cookies on the web.
- Yahoo!: The web and mobile internet behavior is tied.

Voice analytics and lead-generation.

Digitally signing a phone call: caller-ID is not good. Phone-call based certification. Caller-ID does work:

Verification can be: cannot spoof a north american

Howard Schmidt: US Government National Cybersecurity coordinator:

Cyber Security # 2/3 issue related to password: What can US government allocate do this: $ 200M available for grants.

Issue: forced two-factor auth on users for email, just like banks?

Money, research and technology? What would we want them to spend money, technology, policy etc.

Potential technical implementation: pilots etc. Red balloons challenge?

Net Neutrality argument: Is identity a utility: pressure on telcos and web to make identity a utility.

User-centric identity is possibly the answer.

Net Neutrality people are the Privacy people and argue against themselves.

Get out of the federated account: need something that works AT SCALE.

We could live in a password-less world: wallets, watches, phones: tend not to lose.

Throw money yes: Regulate no

Users are not willing to pay.

Potential of independent and conflicting decisions.

- FTC do not track:
- FCC new initiatives:
- ANSTEC from Whitehouse:
- Obama, Kerry, McCain etc all related.

Lack of clarity of rights and obligations: Stable digital identity structure.

Ask the government to throw out funds to people to figure out the real problem.

Continually share knowledge

Personal Data Ecosystem

Convener: Kaliya Hamlin

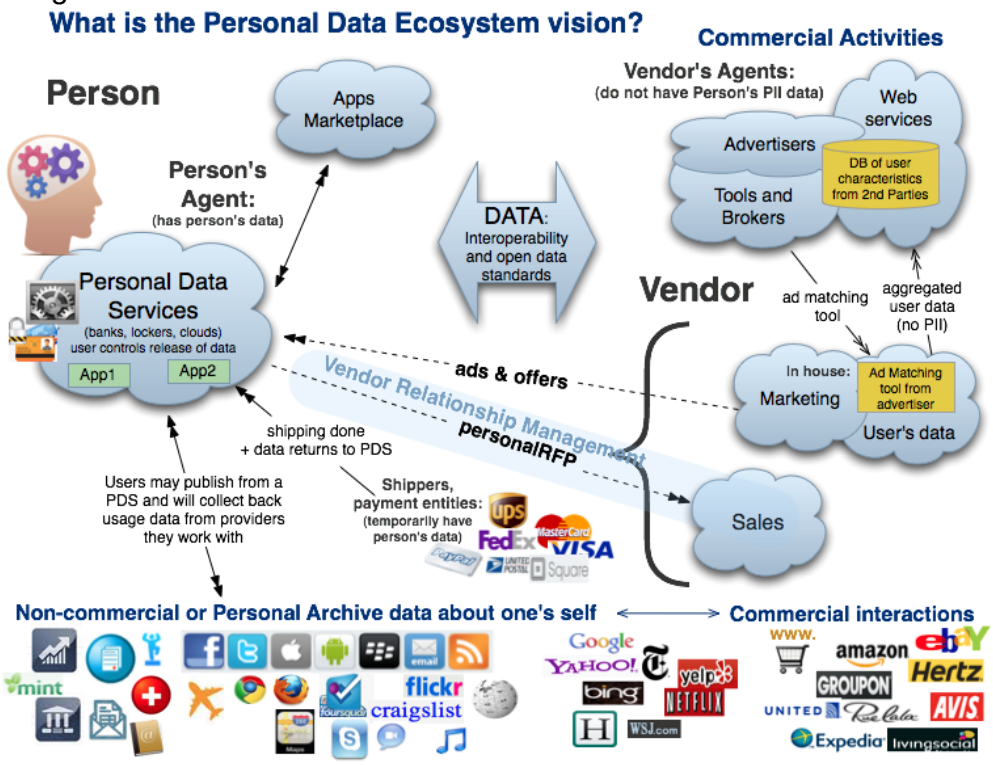
A consortium is forming around the Personal Data Ecosystem.

Personal Data

People generate an enormous amount of data in their every day lives from geo-location data, communications logs, utility bills, health records, financial records, search and browsing logs, social interactions and on and on. Today, systems are not in place to support them collecting, integrating, analyzing and benefiting from their data streams.
Today, large markets trade personal information about people via data aggregators and brokers. The information gleaned in these markets affects people's lives. Everything from life insurance rates, credit worthiness based on who one's friends are, to hotel room upgrades are calculated based on behaviors and affiliations seen in social network and aggregated data about a person.

People are becoming more aware of the implications of these markets due to news investigations like the Wall Street Journal's "What they Know" series. The World Economic Forum just published a report called "Personal Data: The Emergence of a New Asset Class" articulating the value in data gathered from multiple, diverse sources. This report also recommended that user's should be the center of their own data integration, not outside entities.

Image Source: PDEC



The Personal Data Ecosystem

The Consortium is founded by long time user-centric digital identity advocate, Kaliya Hamlin (@IdentityWoman), to catalyze a thriving ecosystem: advocating for individuals having the tools and rights to access and manage their own data;helping business sectors that depended on and made money in the old personal data ecosystem to transform their practices to make money in the new one;

the entrepreneurs building new businesses around user-centric personal data and trust can thrive.

This market will be very competitive. Cooperation is an essential piece of the puzzle so that people have choice and real data interoperability. PDEC explicitly supports cooperation and collaboration for developing open standards,

interoperability, large and small company collaboration and feedback loops for notice and choice, auditing standards and resource sharing.

Consortium accomplishments to date:

- Responded to the Commerce Department Privacy Green Paper & FTC “Do Not Track” White Paper.
- Website that articulating the ecosystem vision. Aggregated blog, from the key leading thinkers along with PEDC original analysis.
- Publishing a Podcast featuring industry leaders.
- Created a collaborative wiki to document the companies, events, publications and people.
- Ran the Personal Data 2.0 event in Palo Alto April 7, 2011.

Project 1: Value Network Mapping and Analysis

The Value Network Mapping and Analysis project is our first major deliverable. It will help people and businesses understand the current personal data ecosystem and map out a future where people are aggregators of their own data in personal data stores and have services running on their behalf using their data. This business methodology names roles and maps the implicit and explicit value flows in the Ecosystem to show where business opportunities are and how the system can work and still remain in alignment with end user interests and needs.

Constituency Initiatives

The Startup Circle: personal data startups can engage in information sharing and collaboration among leaders, technologists, and communications professionals.

1 Million People for Personal Data: gathering early adopters and other people interested in personal data stores and services.

Industry Collaborative: engaging with technologists and business leaders from banking and finance, telecom, cable, web, advertising, media and other industries seeking to understand opportunities, launch pilot projects and ultimately offer services in the ecosystem.

PDEC Programs

Vision Development & Evangelism

Our vision for the ecosystem is inclusive of a wide range of potential services and business models, while holding true to the core non-negotiable that people are ultimately in control of the sum of their data. We will iterate a cycle of collaborating, producing, and sharing this vision.

Meet the Neighbors

Technologies don’t develop in isolation; they have neighboring technology and business communities who need to know about and understand emerging efforts and trends. PDEC is leveraging the relationships with neighboring communities including attending and speaking at their events.

Standards Engagement and Development

To succeed, an effective personal data ecosystem needs to use open standards to allow many different services to interoperate. We are proactively engaged in a number of standards technical committees and presenting at workshops. We are tracking these and sharing this information with the Startup Circle and Industry Collaborative Members.

Mobile Data Analytics

Session: 3B
Convener: Isaac de la Pena
Notes-taker(s): Evan Welbourne
URL: http://hub.personaldataecosystem.org/wagn/Mobile_Data_Analytics
Tags: MicroStrategy, analytics, mobile data, mobile devices, transparency, control

Discussion notes:

An overview of the Microstrategy analytics platform was presented.

The mobile space has special value for analytics:

- Analytics on data collected from mobile devices
  - Monitor calls, SMS, data traffic, application data, sensors like GPS
  - It's easy to collect data in an automated way
  - Mobile data is difficult to forge, so it's more factual and more valuable

- Analysis of data on mobile devices
  - Companies can use mobile devices (e.g., laptops, iPads) to view analytics data remotely
  - Mobile sensors can filter analytics data. Ex: see just the data for the site I'm currently at.
  - "Gestures" with mobile devices enhance interaction with data. Ex: change iPad orientation to switch from one type of chart to a time series

  - Question: If you allow analytics on mobile devices then you're distributing potentially sensitive corporate information geographically across many devices (most corporations would use a VPN). Isn't that a risk?

  - Response: In general, yes it's a risk – but in many cases the value outweighs the risk. You wouldn't want to use the presented system with military or other highly secured data.

  - Response: In many cases a company will outsource analytics to a third party, so there is already a significant level of trust placed in the analysts.

  - Response: In some ways it can actually be better to have a devoted analytics app like this. A common alternative is for analysts to download massive personal data to their laptop (for analysis) where they stick it in a spreadsheet. Then they walk out the doors with the laptop at the end of the day.

- Various "models for how to use" mobile data analytics:
  - 3rd party optimization
  - Internal optimization
  - An app that shows your usage of personal data
  - Increasing customer intimacy. Ex: They put iPads in every O2 store so that sales associates can enter a consumer's email address and instantly get back that consumer's purchase history.
  - Enhanced advertising and promotions
  - B2B2C and B2B2B models. Ex: fine-grained billing based specifically on what you do in the network

Misc discussion:

Q: Most service providers are concerned with optimizing their network and taking steps to switch to a PDE model is a potentially BIG change in course. Can they really be expected to do that?

Response: It is a big change, but some actually are moving in that direction, O2 for example.

Response: Many of them understand the value of digital data and analytics, they need to be convinced of the value of mobile data and analytics.

Q: When you're aggregating all this personal data from mobile devices, are you also creating a new terms of service agreement for the telcos to use?

R: No, the telcos handle all that. They decide what data to collect and how to collect it, Microstrategy just provides the analytics platform. Also, in many cases the data is being collected from/by employees of the telco which makes things easier (legally).

Discussion about the problem of adoption and trust of mobile analytics by end-users (not employees).

- There is a problem with the current understanding of these systems. It seems that the popular understanding is that there are two, polar-opposite models: "do no track anything" or "go ahead and track everything". This is obviously not the case and it's dangerous to formulate a growing discussion (and potentially legislation!) around that misunderstanding.

- Consumer education is going to be increasingly important. The model should be opt-in and a segment of the consumer base will self-select. The biggest potential misstep here is to assume that all consumers will be OK with some as yet unproven model.

- The Interactive Advertising Bureau (IAB) is a good example – they've done a good job educating users about online advertising in addition to location data.

- One tool for educating users and building trust/understanding could be a web or mobile "dashboard" that shows users the contents of their personal data store and how it is being used. For example, such a dashboard may prominently display the *type* of personal information that has been collected, the companies or entities that personal information has been disclosed to, and what has been gained (e.g., deals, money) in return for disclosure of this information. Users may also be able to drill down and see more specifically what data has been collected.

Mobile Authentication Standards and Case Study

Session: C1
Convener: Alex Harrowell – STL
Notes Taker: Philip Laidler – STL
URL: http://hub.personaldataecosystem.org/wagn/Mobile_Data_Analytics
Other attendees: Russell Worth, Joseph Boyle, Julian Gay

Discussion Notes:

Two main groups of standards

1. EU-originated, network centric and transactional, ID is associated to a device (SIM)
2. US-originated, web-based, identity-centric

For (1) Methods available to authenticate individuals over mobile phones.

1999 ETSI m-com program "mobile signature service" standard. Carrier plays the part of public key infrastructure.

This evolved to EAP-SIM (Extensible Authentication Protocol) which is a 3GPP standard, also adopted by IETF . IEEE 802.11 (WPA & WPA2) have adopted five EAP types.

Strong security as requires something you have (a phone) and something you know (password/PIN)

For (2) IETF/OIDF... OpenID and OAuth used for authentication and authorization. Facebook connected will also move to OAuth2.

Potential vulnerability to spoofing or browser holding passwords.

Use Case – Valimo

- Based in Finland, now part of Gemalt
- Provide 2-factor authentication as a white label product
- Have implemented digital identity as legally valid
    - Turkcell
        - 11 Banks + Stock Exchange
        - Government
    - Finland, Sweden, Telefonica (Spain)

Use case

1. Individual sends request to bank (could be on phone or pc or even in a shop)
2. Bank requests authentication to Valimo
3. Valimo sends challenge to phone
4. User approves back to Valimo
5. Valimo sends approval back to bank
6. Transaction is authorised

Valimo = Secure Identity Management Enabler (white label product for telco)

Hypothesis

Could you come up with something stronger that combines both ?

Action

1. Compelling product driven use cases around combining EAP with OAuth
2. Get Telefonica to propose how this could be done (given that they are involved with both)

Next Actions

Mobilizing the Teclcos - What do we want?

Category: Sharing Knowledge
Convener Nitin J. Shah

Participants:
Alex Rootham
Cameron Lewis
Daniel Gerber

Categorize what the telco's could do - inputs from teclo's and all the entities who rely on them.

I will set an e-mail to distribute to participants: Collate and publish ideas

1) Collect Names/contacts of participants
2) Collate ides from participants and categorize
3) Publish potential vectors and end states for Telco's
4) Collect feedback and share - allow ideas on not only WHAT but HOW

Have 10 Founding Founding Members for the Personal Data Ecosystem Consortium

Kaliya

Portable Context Community Prototype

Category: Open Standards and Interop
Convener: Joe Andrieu,

Desired outcome: Live Testbed of prototype context working with multiple relying parties
Context forum: Switchbook, and IIW, VRM & Information Sharing Workgroup

Action 1: Website Explaining PC and Prototype
Joe Andrew - when In process by April 15th

Action 2: IIW Session, Joe Andrew at IIW May 3-5

Who Else to enroll: Vendors & service providers interested and capable of creating value add on top of portable contexts

Resources needed: $, integrated services, user base

First Steps:
1) Target List 2) Referral Contacts 3) Calls

Produce a voluntary code or manifesto that organizations can subscribe to

Category: Define Global Principles
Convener: Stephen Young:

Context: Knowledgerights.org

What: Small concises voluntary code of practice

Who: Stephen Young, Fall 2011

Personal Data Trust Framework Analysis

Category: Define Global Principles
Convener: Scott David & Drummond Reed
Desired Outcome: To complete an OIX analysis of global FIPPS and map them to PDEC requirements
Context: OIX

What - E-mail input via Scott David
Write to Scott by April 30th

Scott to draft and group will review.
Scott David May 10th at European Identity Conference

Map DataPortability to LOC (Levels of Control)

Convener: Phil Wolff
Others: Dwight Irving, David Riley, Kathy Tavitian, Marc Davis

What is desired outcome: To assure personal data portability promises become auditable in LOPs and LOAs
Forum/Context: Data Portability.org

1) Map of 10 Data Portability Valuew to LOC Language for Trust Frameworks
start now - finish by IIW

2) Complete Test of Data Portability Legal Policy