A Decade In:
Assessing the Impacts and Futures of Internet Identity
Topics

• Forging the federating elements (1999-2004+)
• A decade of deployments (2004-2014+)
• Where We Are Now
• What We Got Right
• What We Got Wrong
• Where Do We Need to Go
• Observations
The forging of key federating elements

• The force of the academy and three driving use cases
  – Libraries, science, and collaboration
• A bunch of usual suspects began developing key elements
  – Schema – eduPerson et al
  – Federating software – SAML and Shibboleth
  – Metadata and federation operations
  – Alignment and influence with others – international R&E, OASIS, US Gov, IETF, etc
• Work funded by Internet2, NSF, international federations, OASIS, etc
• Major contributors - MIT, CMU, Wisconsin, Brown, Common Solutions Group (StoneSoup), Duke, Ohio State, Berkeley, Virginia, Texas, Washington, Stanford, etc.
• Special *mumble* to RL “Bob” Morgan
A Decade of Development and Deployment

- InCommon started in 2004 and federations began in Switzerland, the UK, Norway, Sweden etc around the same time
  - Need for alignment of federation practices, schema, etc
  - Finding footing – business models, application and relying party adoption, IdP adoption, changing the paradigm
- Generally exponential growth, US and international, over the next ten years, constant reinvention of key elements to meet scale
- Adoption by other verticals, including real estate, financial industries, law enforcement, pharma, etc.
- Introduction of social identity (Gmail, Yahoo, MSN, AOL, Facebook, etc.) around 2009 and their emerging business model
  - The service is free and the user is the product
  - Almost ubiquitous coverage now
Where We Are Now

- Internet identity is a Big Deal, to consumer companies, governments, academia, enterprises, etc.
- Identity, identifiers and attributes
- Technologies and their integration
- Governance, rules of the road, etc.
- Challenges in scaling privacy
- Limited impacts on scholarly identity
Internet identity as a Big Deal

- R&E federations within academia globally, spanning hundreds of millions of users and thousands of service providers
- Google and Yahoo and their ilk as dominant new industries
- Federal government efforts in several areas, from rules of the road to tool development
- Given proliferation of phishing, privacy breaches, etc, its time to do this right
  - MFA federated authentication
  - Attribute based access controls with consent management
R&E federations world-wide
R&E Europe
Entities in InCommon Metadata

- Num IdPs: 328
- Num SPs: 1435
Social identity

- Google, Yahoo, MSN, Paypal, Amazon, ...
- Widespread, almost ubiquitous but now slowing in the US, still accelerating in other countries
- Adapts well to mobile devices
- Working hard on better security through user and device authentication
- Not working on privacy; the user is their product
- Do statistical identity binding to know who you are
- All attributes are self-asserted by the user
- Ubiquity and promiscuous attribute release are powerful appeals to application developers and services.
Identity, identifiers and attributes

- Identity is you and your account
- Identifiers are unique values tied to you, but often offering privacy instead of identity
  - Different identifiers give different type of privacy
    - (opaque but stateful, opaque and non-correlating, etc.)
- Attributes provide privacy, access control and scale
  - Attributes fall into two rough categories
    - Verified – by the identity provider, an attribute provider, a third party verifier, etc. e.g. Legal name, legal date of birth, over legal age, citizenship, student status, role in organization, is in Class X, walk-in-library-user, is PI of a NIH grant in oncology, etc.
    - Self-asserted – e.g. displayname, friends, interests, preferred language and many others that might better be verified
  - Attributes can be used instead of identity for privacy and simplify scalable access control
Technologies and their integration

- There are several technologies to transport identifiers and attributes
  - SAML
  - OpenId Connect
  - Others
- Convertors and gateways
  - SAML2Social, Social2SAML
- The importance of payloads, not protocols
- The importance of metadata
  - Signing keys for organizations, contact points, attributes required or provided, logo for discovery, etc
  - Provides scale
Governance and the Rules of the Road

- US Gov efforts over the years
  - Generally led by GSA or NIST
  - Have always had little leverage with agencies
    - HSPD 12 the one exception
  - C2G has evolved to a federated model
  - Authority over the general marketplace is uncertain
- International dimension compounds all dimensions
  - European policy affected by local laws
  - Snowden et al has affected marketplace
- The big dogs do what big dogs do... 
  - Repurpose personal data, modify search results, change the assignment of email addresses, institute risk-based measures, retain histories, etc
Limited impacts on scholarly works

- Large impacts on our enterprise (cloud services, single sign on, collaboration tools, etc.)
- Less impacts on scholarly activities
  - Large and ambitious federal efforts have taken a long time
  - Scholarly collections have not reengineered access controls
  - Business models are uncertain
  - Attribute retentive institutions hamper progress
  - Trailing edge users hamper progress
- Catalysts for change are not connecting with points of authority
Challenges in Scaling Privacy

• The spectrum of user interest
  – The Westin Scale from privacy fundamentalist to “doesn’t care”

• Populating, releasing and using attributes
  – Attribute retentive institutions
  – International complexities

• Consent managers have different motives
  – Google wants the user to click through
  – PrivacyLens and others in R&E want informed consent

• Federations evolving role in consent management
  – Auditable source of minimum attribute needs for apps
Consent management

• Key dimension of privacy
• Complex set of legal and technical and international and financial and ... issues
  – When and where and how to use is endless discussion
  – Initial and downstream are separate but very related topics
• Requirements list grows – informed, revocable, accessible, etc.
• Worst case are medical information
• The capabilities of the end user are limited
PrivacyLens privacy manager
PrivacyLens as a paradigm

- Enabling effective and informed end-user consent
- Embraces a set of capabilities
  - Hierarchical information, fine grain control, bundling, revocation of consent, flexible notifications, etc.
- Embraces a style of presentation
  - Clear screens and slides
  - Optional display of values being sent
  - Affirmative user actions
- Embraces a variety of platforms and management approaches
  - Protocol-agnostic
  - Enterprise management consoles and management
  - Audit and security logs
What We Got Right

• First and foremost, the basic model
  – Replacing the rigid structures of PKI with flex at key, scalable junctures
  – Using the driver of academic collaboration rather than administrative business processes
  – The right use cases, spanning privacy and security
• The protocol and its standarization process
• A simple but extensible schema that drove interest
• The focus on metadata to drive scale
What We Got Wrong

• A focus on the web
  – Command line crowd neglected early on; the native mobile app neglected now

• Expectations about various US government activities
  – The Federal PKI Bridge, Eauthentication, FICAM, NSTIC, etc.
  – Good people trapped in impossibly complex situations

• That privacy issues could be rationally resolved
  – International issues
  – Attribute retentive institutions
  – Inconsistent user behavior
Where Do We Go From Here

- Scalable and accessible access control
- Expanding the use of metadata
- Integrate security and privacy by design
- Extending the scope of trust
Scalable and accessible access control

- Moving from identity-based (or, gasp, IP address) access control to attribute-based access controls
  - Provides privacy, with audit
  - Provides scale, managed delegation, etc.
  - Requires effective end-user consent management
- Attribute-based accessibility approaches
  - Using attributes to affect the display of content for physical, mental and cognitive disabilities
  - Comprehensive approaches are being developed – see GPII.net for example
  - Allows privacy and accessibility
- Reworking the access controls for restricted materials
Expanding the use of trust-related metadata

- Tags on end-entities in metadata to convey information about services
  - Current uses: R&S tags, Hide-from-discovery IdP’s
  - Future uses may be legion, conveying required attribute bundles, conduct of applications, consent management support, conduct of IdP, terms of use, etc.
  - And the business processes to support the verification of asserted tags
- Moving from static to dynamic metadata services
  - Supports interfederation
  - Drives difficult international community standards processes
Integrate security and privacy by design

- Good privacy begins with good security
- Security by design
  - MFA
  - Federated MFA
- Privacy by design
  - Required and optional sets of attributes
  - Anticipating consent needs for third parties
Extend the scope of trust

- Interfederation
- Federated incident handling
- Require sound operational practices for members
  - The real security, and therefore trust, threats are not in protection of the SAML signing key but in almost everything else the organization does
  - Patches, Heartbleed, OS updates, etc.
- Applying trust to attributes, attribute authorities, etc.
Observations

• It is a true “layer” on the Internet, except it isn’t a layer
• The use cases are remarkably diverse and even more remarkably, can be met with a common infrastructure
• Policy design for access control should follow Protocol Design Principle #3
  – Be conservative in what you send; be liberal in what you receive
• The rapid development of the marketplace overtook policy
• Policies must target outcomes, not technologies