Achieving Privacy in a Federated Identity Management System

Susan Landau, Hubert Le Van Gong
Sun Microsystems Laboratories
Robin Wilton
Future Identity

Simplifying Identity Management

- Digital identity is split over multiple domains.
- This complicates user interaction;
- Is wasteful of resources,
- As well as being insecure.
- One solution is federated identity management.

Simplifying Identity Management

- Digital identity is split over multiple domains.
- This complicates user interaction;
- Is wasteful of resources,
- As well as being insecure.
- One solution is federated identity management.
- But federated identity management blurs security boundaries, and brings its own security --- and privacy --- issues.

One way to look at federated solutions:

- You surf the web
  - > Authenticated once
  - > Seamlessly logged in to all sites

Single Sign-On

- You purchase airlines tickets and move to another country
  - > Your online calendar is updated
  - > Your magazine subscription follows you to your new address

Delegated Authorization

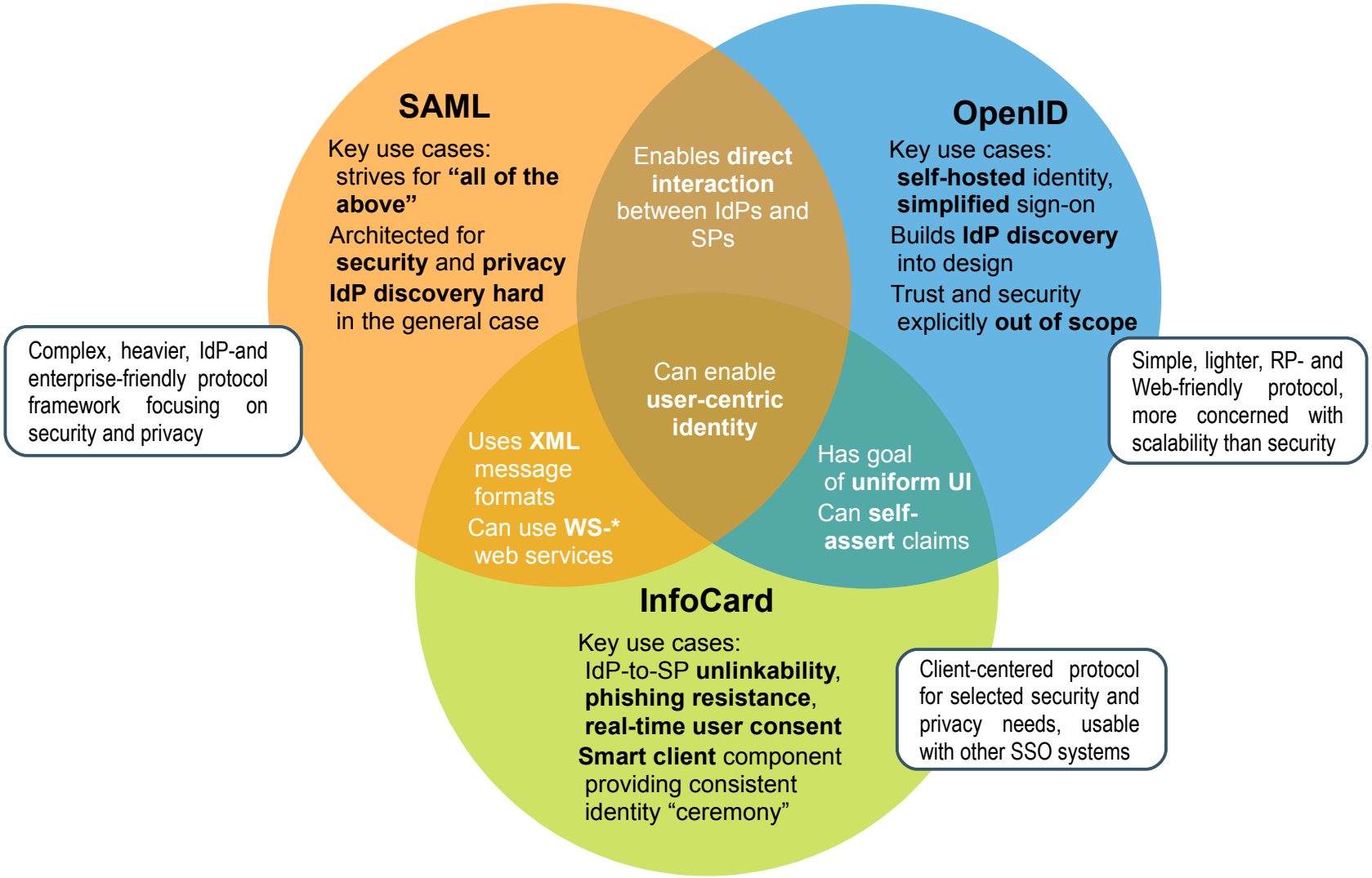
What is a Federated Identity Management System?

- Three actors:
  Principal, or user, with particular digital identity;
  Identity Provider (IdP), who authenticates user **once**, then issues authentication credentials;
  Service Provider (SP), who provides services to authenticated users.
- Many forms of authentication, as strong as IdP wants/needs.
- Current variants: OpenID, CardSpace, SAML (which underlies Liberty protocols).

Variety of Standards

- OpenID 1.0
  - Meant for “open” Internet.
  - Principal identifiers are typically URLs.
  - IdPs are OpenID Providers; you can choose any one you want.
  - SPs --- called RPs --- have no protocol basis for trust.
- CardSpace
  - IdPs are managed card providers or self issued (managed solely on Principal's device).
  - Various models: a card might be accepted only by the issuer or by other RPs.
- SAML
  - Accepts existence of limited federations --- Circles of Trust.
  - Accepts need for business contracts to govern CoTs.

Venn of Federated identity Systems



*Contributed by Eve Maler

Why do we use Federated Identity Management?

- Convenience.
- Efficiency.
- Security.
- Compliance with data protection requirements.

The Social Contract of Data Sharing

- Handing data over isn't free.

The Social Contract of Data Sharing

- Handing data over isn't free.
- There's a social contract: I give you this, I get that in return.

This constrains the IdP and the SP in principle. In practice the constraints can be difficult to express and enforce – whether technically or contractually.

What is the System?

- Technology (e.g., OpenID, CardSpace, SAML).
- Business contracts.
- Legal, regulatory, commercial, and technical implementation factors.

What is the System?

- Technology (e.g., OpenID, CardSpace, SAML).
- Business contracts.
- Legal, regulatory, commercial, and technical implementation factors.

Privacy Driver	Incentive
Best Practices Industry Code of Conduct Legal and/or regulatory controls	Improve User Trust Industry Sanctions Avoid prosecution and/or liability

The Risks in Data Collection Systems

- Data disclosure (inappropriate, excessive, without consent).
- Metadata disclosure (allowing linkages between actions, allowing inferences, etc.).
- Regulatory exposure.

The Risks in Data Collection Systems

- Data disclosure (inappropriate, excessive, without consent).
- Metadata disclosure (allowing linkages between actions, allowing inferences, etc.).
- Regulatory exposure.

- Therefore what you want to do is minimize data collection while collecting what you need to do the job.

Approaches to Minimal Data Collection

- Federation.
- Privacy-preserving methods for supplying disparate pieces of information to an SP (Gevers 2007, Sun solution in 2005, etc.).
- Idemix, a zero-knowledge based system for creating credentials that protect user privacy (IBM). One problem with Idemix: no ability to share the credential (if shared once, usable everywhere by shared party).

Crucial Issue: what's the use case?

- Where is the credential being used?
- Is the issuer willing to have the credential used this way? If not, what are the costs to the user of doing so?
- Plumbing the depths of the usage case illuminates the privacy pressure points.

Crucial Issue: what's the use case?

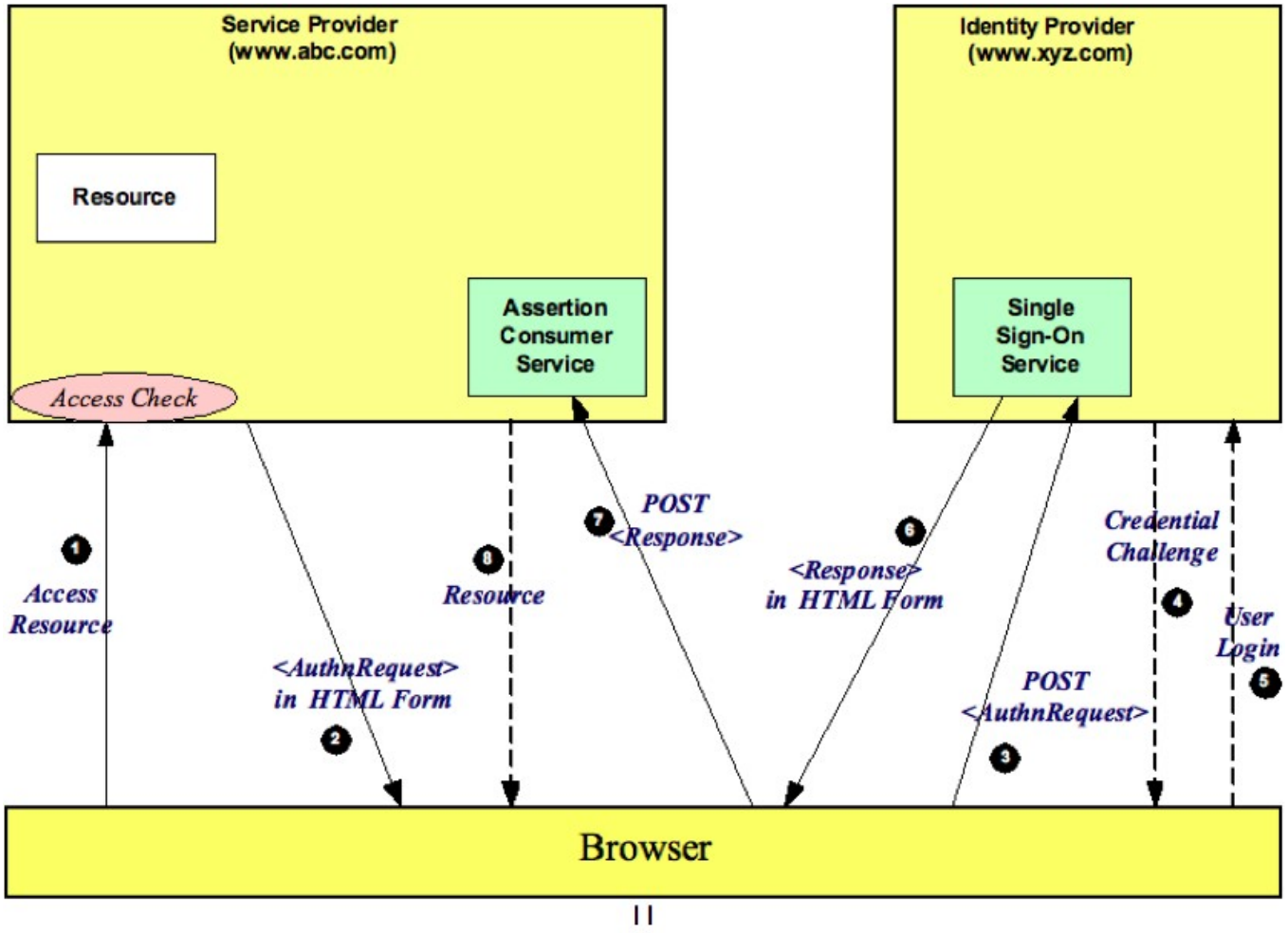
- Where is the credential being used?
- Is the issuer willing to have the credential used this way? If not, what are the costs to the user of doing so?
- Plumbing the depths of the usage case illuminates the privacy pressure points.

- Then you can use the various tools society provides: contracts, law, regulation, liability.

How do technical, legal, and policy protections play out?

- If the technical protection is simple or foolproof or the only solution, it works.
  Examples: crypto protections for data at rest, SSL for web transactions.
- Otherwise, the solutions will be a combination of legal, regulatory, and technical.
- And --- minimizing data collection.

SAML v2.0


Service Provider
(www.abc.com)
Resource
Assertion
Consumer
Service
Access Check
Identity Provider
(www.xyz.com)
Single
Sign-On
Service
1
Access
Resource
2
<AuthnRequest>
in HTML Form
3
POST
<AuthnRequest>
4
Credential
Challenge
5
User
Login
6
<Response>
in HTML Form
7
POST
<Response>
8
Resource
Browser

Cross-Domain Federation

- Per the Liberty protocols, user's data is only released with the user's consent and based on the user-defined policies

Rent-A-Kayak.com*
Service Provider

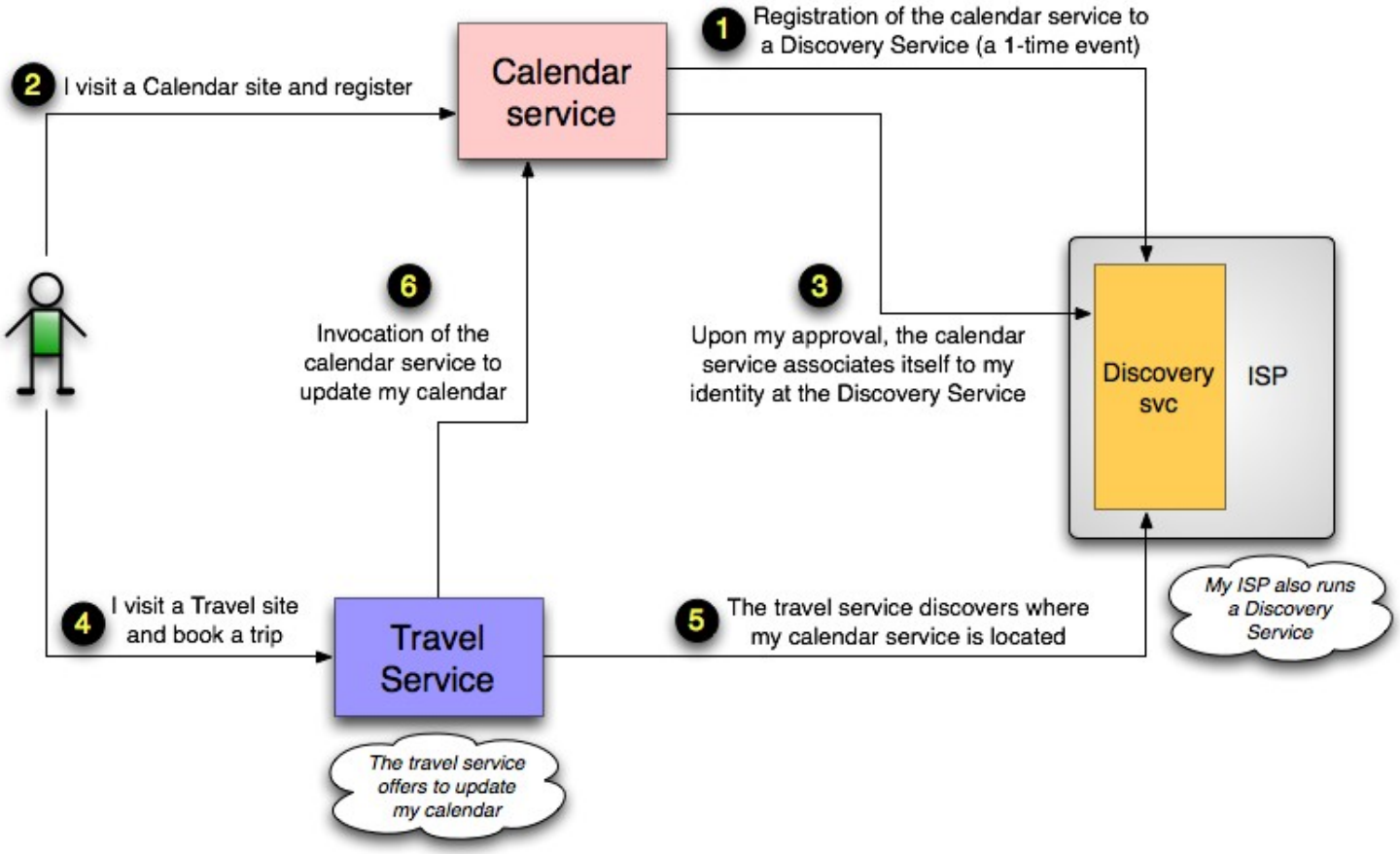
1. service provider requests user attributes from identity provider

3. user accepts or rejects exceptions to existing policies and preferences

MyTravelAgent.com*
Identity Provider

2. attributes released per user's policies and preferences

Liberty ID-WSF



Privacy Protections in Liberty/SAML

- Liberty protocols address **only data exchange.**
- Engineers design protocols, but contracts are key.
- What are best practices for data management?

Privacy and Security Best Practices

- IdP should safeguard Principal's credentials; mechanism in place to require authentication domain to use credentials properly.
- SPs should inform Principals of their data practices, provide Principals with choices regarding secondary use of data, maintain security of Principals data within their control, and not use or share Principal's data except in accordance with SP privacy policy and/or consent and directive of Principal.
- Attribute Provider has same responsibilities.

Identity Governance Framework

- Goal: break identity silos and fully leverage identity networking by decoupling applications from infrastructure.
- The Identity Governance Framework defines:

a set of declarative policies that document and govern exchange of identity-related data between data consumers and data providers.

- In other words:
  - >Why should information be transferred, collected or updated?
  - >Who gets to do what?
  - >Where will the data be used or held?
- Identity Governance covers:
  - >CARML – Schemas for data & transaction definitions
  - >WS-Policy – Privacy assertions
  - >AAPML – Attribute Authority Policy (XACML)

So does this work? Attacks on Liberty privacy protections

- Pfitzmann 2003: found some ambiguities in Liberty protocols; fixed.
- Pfitzmann and Waidner 2003: found a man-in-the-middle attack on original protocols (also found by Jonathan Sergent of Sun/Liberty).
- Alsaleh and Adams 2006: consumer privacy issues.
- Josang et al. 2007 and Bhargav-Spantzel et al. also have concerns, but ...

Alsaleh and Adams: Liberty doesn't protect consumer privacy

- An IdP introduces Principal to Circle of Trust.
- Account linkage between IdP and SP.
- An SP might retain user address.
- Two SPs could collude and share information about a Principal.
- An SP could amass information about a Principal, leading to identification of the Principal.

- An SP knowing a user's preferred IdP violates user's privacy.
- An SP could determine Principal's most recently used IdP.
- An SP can request reauthentication of a Principal at any time.

Alsaleh and Adams: Liberty doesn't protect consumer privacy

The risks are being considered through the wrong lens:

- privacy issues resolved through legal means;
- privacy issues resolved in favor of usability;
- privacy issues previously resolved.

The Real Points

- Digital identity has to balance competing sets of needs from users, IdPs, and SPs.
- Digital identity will only succeed with the consent of the user (coercion possible depending on circumstance).
- Federation enables identity-management systems that are secure, minimize data exchange, and are thus inherently privacy protective.

The Real Points

- Digital identity has to balance competing sets of needs from users, IdPs, and SPs.
- Digital identity will only succeed with the consent of the user (coercion possible depending on circumstance).
- Federation enables identity-management systems that are secure, minimize data exchange, and are thus inherently privacy protective.
- Implementers must weigh risks and balance accordingly. Federation gives them one way to do so.

The Real Point

- "The identity challenge is both technical, business, and policy oriented."
  - *Liberty Alliance*

susan.landau@sun.com
hubert.levangong@sun.com
futureidentity@fastmail.fm