Leaving the Teenage Years Behind:
Internet Identity Comes of Age
Topics

• A wild and reckless youth
  – The start
  – Successes and changes
• Challenges of the teenage years
• Coming of age
  – Dynamic metadata and Discovery
  – Integration of identities
  – GDPR and its needs
  – Front-end and back-end privacy
  – Consent activities
    • A Side CAR
  – Baseline expectations
  – IoT – can we manage what manages us
• Fearless and foolish predictions
Internet Identity: A Wild and Reckless Youth

• Emerged around 2000 from an R&E driven need for interrealm authentication and authorization to support collaboration, growing into commercial to commercial relationships and then into a consumer and social identity foundation
• From a hierarchical PKI to federated SAML
  – PKI was globally scalable but not locally deployable; Federation was locally deployable but is it globally scalable?
  – Federated identity: Authenticate locally; act globally
  – The importance of attributes for privacy and scalable access control
• From Club Shib to InQueue to InCommon
• From the US to global and the start of Refeds
• From the biz of custom big dogs to ubiquity and the cloud
• From SAML to JSON
Challenges of the teenage years

• Too much success
  – Size of Metadata
  – Difficulty of IdP discovery

• A big and diverse world to serve
  – Other initiatives now need identity
  – Integration with other forms of identity

• An increasing spectrum of skills and interests to serve
  – From the cutting edge down to increasing use of commercial cloud service
  – Internationally, from the original big feds to a growing set of new countries

• Moving from skinny to fat trust profiles
  – With the number of tunable knobs in protocols, profiles become key to interop
  – Skinny profiles specify little – easier to meet, harder to trust
  – Fat profiles specify everything – harder to meet but perhaps more trustworthy
Internet identity: Coming of age

- Dynamic metadata and Discovery
- Integration of identities
- GDPR and its needs
- Attribute release
- Consent and the Kantara WG
- Baseline expectations
- Front-end and back-end privacy threats
- IoT – can we manage what manages us
Dynamic metadata and discovery

• The federation metadata files are growing too large
• Dynamic metadata
  – Similar to moving from /etc/hosts to DNS 25 years ago
  – Protocols and code now exists; some experimentation starting
  – Federations will likely construct some common bundles for ease of deployments
• Discovery
  – The process of helping a user find their identity provider
  – Active work in RA21
    • Options include browser assists, central discovery service, pull downs
• The interplay between the two
  – Dynamic metadata confounds building traditional pull-down or NASCAR discovery services
Integration of identities

• Federated identity has gotten deep traction
  – MFA is growing rapidly

• Integration of social identity and federated identity is happening
  – Social 2 SAML and SAML to Social gateways exist
  – Still tricky management of self-asserted versus vetted (enterprise) attribute values

• Sovereign identity gaining niche traction
  – Long-held desires for fully autonomous identity now enabled by blockchains
  – A game-changer, for some definition of game
  – Not yet integrated, if only by the nature of the developers
  – sovrin.org as a gathering place for the crowd

• Role of special identities evolving – e.g. ORCID
GDPR (General Data Protection Regulation)

- Created by EU to manage data protection uniformly across the EU
  - Is binding for every member EU nation
  - With many global impacts
- Covers a vast waterfront of issues from tracking to attribute release to right to be forgotten to data breaches to . . .
- Consists of a set of rules (Articles) and then example interpretations of the rules in key areas (Recitations)
- Penalties of up to 4% of global revenue
- Identifies six reasons for attribute release, including contract, consent, national security, legal actions, etc.
  - Specifies when consent is not to be used, when it should be used, the quality of the consent, etc.
- It affects many, perhaps most, US institutions.
**GDPR**

**ENFORCEMENT**
- **Fines**
  - Up to 20 million euros or 4% of total annual worldwide turnover.
  - Less serious violations: Up to 10 million euros or 2% of total annual worldwide turnover.

**INTERNATIONAL DATA TRANSFER**
- **Adequate Level of Data Protection**
- **Model Contractual Clauses**
- **Binding Corporate Rules (BCRs)**
- **Privacy Shield**

**LEGAL BASIS**
- Performance of a contract
- Compliance with a legal obligation
- To protect a person’s vital interests
- Task in the public interest
- Legitimate interests

**RESPONSIBILITIES OF DATA CONTROLLERS AND PROCESSORS**
- **Data Protection Officer (DPO)**
- Designate DPO if core activity involves regular monitoring or processing large quantities of personal data.
- For high risk situations
  - Notify supervisory authorities no later than 72 hours after discovery.

**DATA BREACH NOTIFICATION**
- A personal data breach is "a breach of security leading to the accidental or unlawful destruction, loss, alteration, unauthorized disclosure of, or access to, personal data transmitted, stored or otherwise processed."
- If likely to result in a high privacy risk → notify data subjects
- Notify supervisory authorities no later than 72 hours after discovery.

**PERSONAL DATA**
- **Identified**
- **Identifiable**
- **Sensitive data**
  - Religious or Philosophical Beliefs
  - Trade Union Membership
  - Political Opinions
  - Racial or Ethnic Origin
  - Health

**SPECIAL PROTECTIONS**
- **Racial or Ethnic Origin**
- **Religious or Philosophical Beliefs**
- **Health**
- **Trade Union Membership**
- **Sex**
- **Life**
- **Political Opinions**
- **Biometric Data**
- **Genetic Data**

**Effective Judicial Remedies**
- Compensation for material and non-material harm.

**TERRITORIAL SCOPE**
- **EU Establishments**
- **Non-EU Established Organizations**
  - Offer goods or services or engaging in monitoring within the EU.

**LAWFUL PROCESSING**
- Collection and processing of personal data must be for “specified, explicit and legitimate purposes”
- With consent of data subject or necessary for:
  - Performance of a contract
  - Compliance with a legal obligation
  - To protect a person’s vital interests
  - Task in the public interest
  - Legitimate interests

**CONSENT**
- Consent must be freely given, specific, informed, and unambiguous.

**RIGHTS OF DATA SUBJECTS**
- Right to Information
- Access and Rectification
- Purpose Specification and Minimization
- Right to Portability
- Right to Erasure
- Automated Decision Making
- Transparency

**INTERNATIONAL DATA TRANSFER**
- **Adequate Level of Data Protection**
- **Model Contractual Clauses**
- **Privacy Shield**

**THE PLAYERS**
- **Data Subects**
- **Data Controllers**
- **Data Processors**
- **Supervisory Authorities**

**SENSITIVE DATA**
- **Health**
- **Trade Union Membership**
- **Sex**
- **Life**
- **Political Opinions**
- **Biometric Data**
- **Genetic Data**

**DATA PROTECTION BY DESIGN**
- Built in starting at the beginning of the design process.

**TEACHPRIVACY**
- **www.teachprivacy.com**

Workforce awareness training by Prof. Daniel J. Solove
Some gnarly details

• PII and Sensitive PII
  – Almost everything is PII – from IP address to persistent identifiers
    • Some identifiers are not e.g. ePTID
  – Sensitive PII
    • Religious, ethnic, sexual, health, trade-union membership, etc.
    • Requires special handling in everything from protection to presentation

• Research data use
• Right to be forgotten
  – Cloud based backups
• “This call may be recorded…”
• Data breach notifications
  – 72 hours
• Data protection officer and individual data protection training
Attribute release

• Institutional policies limit default attribute release
  – Perhaps the greatest barrier to inter-institutional collaborations
  – Driven by interpretations of FERPA, HIPPA, etc

• Options such as explicit end-user consent not widely deployed

• Use of R&S (Research and Scholarship) to address release of a core set of scholarly attributes
  – Adoption growing, albeit slowly
  – International implications, at least perceived
Consent Activities

- **Kantara WG**
  - Consent receipts v1.1 released
  - Best Practices in Consent Management starting up

- **Interactive Advertising Bureau**
  - Making cookies and consent compliant with GDPR

- **CAR – Consent Informed Attribute Release**
  - Originally funded by NIST as part of NSTIC
  - Multi-protocol consent as a service, integrating institutional and individual preferences
  - Emphasis on informed content for effective user decisions
CAR Architecture
Review what you are releasing to CIlogon

CIlogon is requesting information about you from your TIER record.

What would you like to release to CIlogon?

- **Email Address**: [kjg@internet2.edu]
  - **Permit**: Yes
  - **Deny**: No

- **Legal Name - Last/Family**: [Klingenstein]
  - **Permit**: Yes
  - **Deny**: No

- **Name - First/Given/Legal/Other**
  - **Permit**: Yes
  - **Deny**: No

- **Name - Full (Preferred)**
  - **Permit**: Yes
  - **Deny**: No

- **Scoped NetID**: [kjg1@tier.internet2.edu]
  - **Permit**: Yes
  - **Deny**: No

Hide:

Choose one:

- Save my choices: don't show me this screen again unless necessary.
- Save my choices, but show me this screen next time.
- Don't save the choices I made just now. Show me this screen next time.

CANCEL  ACCEPT AND CONTINUE
### My Sites

Manage what information will be shared with these sites:

<table>
<thead>
<tr>
<th>Name</th>
<th>URL</th>
<th>Updated</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>CI Logon</td>
<td>citolog.org</td>
<td>05/30/2017</td>
<td>manage</td>
</tr>
<tr>
<td>G7iant Service Provider Proxy</td>
<td>terena.org</td>
<td>05/01/2017</td>
<td>manage</td>
</tr>
<tr>
<td>Internet2 Collaboration Wiki Spaces</td>
<td>spaces.internet2.edu</td>
<td>05/18/2017</td>
<td>manage</td>
</tr>
<tr>
<td>LIGO Wiki</td>
<td>wiki.ligo.org</td>
<td>05/02/2017</td>
<td>manage</td>
</tr>
<tr>
<td>TIER CARMA</td>
<td>carma.testbed.tier.internet2.edu</td>
<td>06/15/2017</td>
<td>manage</td>
</tr>
</tbody>
</table>

### New Site Policy

Manage defaults for what information is shared with new sites.
Manage information sharing for CILogon

Information Requested by CILogon

You can choose whether the following information is shared with CILogon:

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Current Value</th>
<th>Current Choice</th>
<th>Duke Recommends</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name - Full (Preferred)</td>
<td>Ken Klingenstein</td>
<td>permit</td>
<td>permit</td>
</tr>
<tr>
<td>Scoped NetID</td>
<td><a href="mailto:kjk1@tier.internet2.edu">kjk1@tier.internet2.edu</a></td>
<td>permit</td>
<td>permit</td>
</tr>
<tr>
<td>Name - First/Given (Legal)</td>
<td>Ken</td>
<td>permit</td>
<td>permit</td>
</tr>
<tr>
<td>Email Address</td>
<td><a href="mailto:kjk@internet2.edu">kjk@internet2.edu</a></td>
<td>permit</td>
<td>permit</td>
</tr>
<tr>
<td>Legal Name - Last/Family</td>
<td>Klingenstein</td>
<td>permit</td>
<td>permit</td>
</tr>
</tbody>
</table>

Additional Settings

- All other information
  - If CILogon requests information not listed above
    - (any values)  askMe  askMe

- While I'm Away
  - If your choice above is "askMe" but you're not available to answer when CILogon requests information about you
    - (any values)  deny  deny
Baseline Expectations

• Moving from ad hoc trust towards a set of common expectations
• Provide a baseline for trust, make collaboration more predictable and ensure that the InCommon Federation’s strategic value to research and education continues to grow
• Addresses
  – IdP, SP software security and patch maintenance
  – IdM operations and authority
  – Maintaining federated metadata properly, including privacy URL, security contacts, logos, etc.
  – Proper handling of received attributes and appropriate codes of conduct
  – Federated incident handling commitments
• Being advanced in stages, involving consensus building and then deployment
• https://incommon.org/federation/baseline/index.html
Front end and back end privacy threats

• Backend threats:
  – Once the RP has the data
  – Harvesting user’s personal data ala Cambridge Analytica
  – Best managed through legal agreements, limiting the purposes, etc.

• Frontend threats
  – Managing what data the RP gets
  – Opaque consent dialogues
  – Lack of clear guidelines on data minimization
Internet of Things (IoT) and the Enterprise and Alexa

• The security vulnerabilities of IoT are staggering
  – Passwords embedded by manufacturer are well-known and obvious
  – Over the air patches don’t happen
  – Multiply by number of devices for DDOS attacks
  – Ransomware on things just starting

• The privacy vulnerabilities of IoT are staggering
  – What information is being captured by the thing?
  – Who owns the data in the cloud? Samsung? Amazon?
    • Not the individual

• Alexa understanding pronouns
  – Who owns that knowledge of you?
  – What can the government ask?

• Higher ed is a particularly vulnerable enterprise
  – Decentralized acquisitions; no lifecycle management
  – Cutting edge research needs
Some predictions

• Sovereign identity finds a small place in the Internet identity realm
• Blockchain for small circles of trust, e.g. friends, the set of roots of other trust chains
• Consent gets traction and users get better front-end privacy
  – GDPR drives change; pay attention to the purpose field
• IoT security and privacy threats will get more alarming; it’ll be Y2K every day
• Profiles will continue to get fatter, to cover more concerns, but may remain self-asserted
• Convenience will win out over the desire for privacy, sigh.