

The Invention of Internet Identity

Ken Klingenstein
Director, Internet2 Middleware and Security



Topics

- A bit of personal context
- A brief history of Internet identity
- What we have learned from the practice
- What needs to be invented next
- Conclusions

1. Context

- My community
 - R&E federations globally – refeds.org
 - InCommon in the US
- My role and background
 - Internet2 Middleware
 - Shibboleth and SAML
 - Former CIO and Professor at the University of Colorado
 - Original Internet developer/deployer
- Invention

Federations

- A set of organizations that share identity and attributes using SAML based software
 - Anchored by a common set of business practices
 - Has a shared schema
 - Federated operator registers enterprises, aggregates and distributes organizational metadata
- Bi-lateral federations occur between organizations and outsourced cloud software service provider
- Multilateral federations occur in verticals, beginning with R&E, but now extending to government, medical, real estate, financials, etc.

International R&E federations

- > 100M users across >30 countries
- Coverage in several countries is 100%, and extensive in many others.
- Generally part of a national network but associated with another org or independent in a few
- Frequently linked to several government activities, in research, education, governance, health, etc.
- Some interfederation activities, including the Kalmar2 union and eduGAIN.
- www.refeds.org

Grab File Edit Capture Window Help

Atlases - High Resolution Pathology Images

https://atlases.muni.cz/en/index.html

Most Visited - Getting Started - Latest Headlines

OneView Internet Login CNN.com - Breaking News, U.S. Washington Dulles International Program - 10th Symposium on Dashboard - Internet2 Wiki Atlases - High Resolution Path

Atlases - PATHOLOGY IMAGES

Collection of high resolution histological images

Lang:

Registered users: 16229

In order to have an access to the high resolution images you have to LOGIN below.

If you have an account in one of the following Identity Federation, click on the logo.

- eduID.cz
- Log in with MAYF
- SIR
- SVITCH
- DFN
- Geduh
- GakuNin
- Routin
- arnes
- cafe
- Log in KALMAR

If you are not member of any listed identity federation, click on the button below:

[Local account](#)

[Contact us](#) | [Privacy](#) | [How to cite Atlases](#)


CESNET

Done

Select your identity provider

SELECT YOUR IDENTITY PROVIDER

English | Български | Nyorsk | Sámegiella | Dansk | Deutsch | Español | Svenska | Suomi | Français | Italiano | Nederlands | Luxembourgish | Czech | Slovenščina | Hrvatski | Magyar | Język polski | Português | Ρυθμιζόμενες | ភាសាខ្មែរ | 日本語 | 中文 | አማርኛ | አማርኛ | አማርኛ | አማርኛ | አማርኛ | አማርኛ



You have previously chosen to authenticate at Internet2

[Login at Internet2](#)

All | Nordic countries | Spain | UK | eduGAIN | Guest providers | Miscellaneous

808 entries

Incremental search...

- Internet2
- AAIEduHr - Croatian Research and Education Federation
- Aberdeen College
- Aberdeen College Staff
- Aberystwyth University
- Abingdon and Witrey College
- Accrington & Rossendale College
- Adam Smith College
- AES11
- ALBA - CELLS
- Anglia Ruskin (Old System)
- Anglia Ruskin University Login



InCommon today

- 250+universities, 450+total participants, growth continues rapid
- > 10 M users
- Traditional uses continue to grow:
 - Outsourced services, government applications, access to software, access to licensed content, etc.
- New uses bloom:
 - Access to wikis, shared services, cloud services, calendaring, command line apps, medical, etc.
- Certificate services bind the InCommon trust policies to new applications, including signing, encryption, etc.
- FICAM provisionally (privacy to be worked) certified at LOA 1 and 2 (Bronze and Silver).

National Institute of Health

The screenshot shows a web browser window titled "Federation" with the URL <https://federation.nih.gov/FederationGateway>. The page features a yellow header with the text "NIH Federated Login". Below the header, there is a section for "Account Type" set to "Research Organizations". A dropdown menu is open, displaying a list of institutions with checkboxes next to them. The "University of Chicago" is currently selected. To the left of the dropdown, there is a "Warning Notice" section. In the center of the page, there is a disclaimer about unauthorized access and a link to "Interact with NIH". At the bottom of the page, there are logos for CIT and other entities, along with contact information for the NIH Helpdesk.

Account Type: Research Organizations

Interact with NIH

- Dartmouth College
- Duke University
- Florida State University
- Georgetown University
- ICommon LLC
- Indiana University
- Iowa State
- James Madison University
- Johns Hopkins University
- Lafayette College
- Laromus Berkeley National Laboratory
- Louisiana State University
- Medical University of South Carolina
- Miami University
- Michigan State University
- New Landring Marine Laboratories
- New York University
- Northwestern University
- Ohio University Main Campus
- OhioLink
- Old Dominion University
- Oregon Health & Science University
- Pennsylvania State University
- Purdue University Main Campus
- Rice University
- Rutgers, The State University of New Jersey
- Stanford University
- The Ohio State University of Columbus
- Stony Brook University
- Texas A & M University
- The Ohio State University
- University at Buffalo, The State University of New York
- University of Alabama at Birmingham
- University of Alaska Seward System
- University of Arizona
- University of Arkansas for Medical Sciences
- University of California - Office of the President
- University of California-Davis
- University of California-Los Angeles
- University of California-San Diego
- University of California, Berkeley
- University of California, Davis
- University of California, Merced
- University of California, Riverside
- University of California, San Francisco
- University of Chicago
- University of Cincinnati
- University of Dayton
- University of Florida

Warning Notice
This is a U.S. Government computer system, which is not intended for release to the public. All information on this computer system may be disclosed to the public under the provisions of the Freedom of Information Act. There is no right to privacy.

business by authorized personnel. Unauthorized access or use of this computer system may subject violators to criminal, civil, and/or administrative action.

authorized personnel for official purposes, including criminal investigations. Such information includes sensitive data exempted to comply with confidentiality and privacy requirements. Access or use of this computer system by any person, whether authorized or unauthorized, constitutes

If you need assistance - Please call the NIH Helpdesk 301-496-4337 (5-9E,F); 866-319-4337 (toll-free) or [Submit a Help Desk Ticket](#)

CIT

INTERNET®

Research.gov

The screenshot shows the Research.gov homepage in a Firefox browser window. The browser's address bar displays the URL: http://www.research.gov/research-portal/appmanager/base/desktop?_rfpb=true&pagelabel=research_home_page. The page features a dark blue header with the Research.gov logo and the tagline "POWERING KNOWLEDGE AND INNOVATION". A navigation menu includes links for Home, Contact Us, Site Map, and Help, along with the date November 23, 2011. A search bar is located in the top right corner. The main content area is divided into several sections: a "LOGIN AS" dropdown menu with options like NSF Visitor, NSF User, NSF Staff, USDA User, and InCommon; a large banner for "Now PIs & Co-PIs can access FastLane via Research.gov Using Single Sign-On!"; an "Alerts" section with a maintenance notice; an "Our Services" section with links to Research Spending & Results, Science, Engineering, and Education Innovation (SEE Innovation), Policy Library, and Grants Application Status; and a "RECOVERY.gov" section with a "Learn More" link. A weather widget on the right shows 49°F, Partly Cloudy, in Arlington, VA, on Wednesday, November 23, 2011, at 7:51 pm EST. The browser's status bar at the bottom right shows the Windows logo and a registered trademark symbol.

National Science Foundation (NSF)

The screenshot shows a Firefox browser window displaying the Research.gov InCommon Federation login page. The browser's address bar shows the URL: `https://identity.research.gov/sso/fedauth.jsp?env=prvw&app=portal`. The page header includes the Research.gov logo, navigation links for Home and Help, the date November 23, 2011, and a font size adjustment tool. The main content area is titled "InCommon Federation" and contains a "System Use Notification" and an "Access to Sensitive Information" section. Below the text is a dropdown menu for selecting an institution, with a "Log In" button and a "Cancel" button. The dropdown menu is open, showing a list of institutions including California Institute of Technology, Carnegie Mellon University, Colorado State University, Cornell University, Indiana University, Indiana University of Pennsylvania, Internet2, Johns Hopkins, Louisiana State University, Oklahoma State University Main Campus, Penn State, Texas A & M University, The University of Arizona, The University of Memphis, University of Baltimore, University of California, Davis, University of Central Florida, University of Chicago, and University of Cincinnati Main Campus. A green box at the bottom of the page contains links for Privacy Policy, FOIA, and No F, and a logo for "Led by The National Science Foundation".

Research.gov
POWERING KNOWLEDGE AND INNOVATION
Home Help
November 23, 2011
Adjust Font Size: A A A

InCommon Federation

System Use Notification
This is a National Science Foundation (NSF) Federal government computer system. Unauthorized attempts to modify any information stored on this system, to defeat or circumvent security features, or to use this system for other than its intended purposes are illegal and may result in disciplinary action, criminal prosecution, or both.

Access to Sensitive Information
The InCommon Federation provides NSF's research and education community easier access to online services using Research.gov. InCommon leverages technology developed under an NSF-funded grant that enables researchers and sponsored programs offices to securely access Research.gov using the user ID and password issued by their institution. Select your institution from the drop down menu below, to be taken to the InCommon login page for your institution.

Make your selection here Log In Cancel

Make your selection here

- California Institute of Technology
- Carnegie Mellon University
- Colorado State University
- Cornell University
- Indiana University
- Indiana University of Pennsylvania
- Internet2
- Johns Hopkins
- Louisiana State University
- Oklahoma State University Main Campus
- Penn State
- Texas A & M University
- The University of Arizona
- The University of Memphis
- University of Baltimore
- University of California, Davis
- University of Central Florida
- University of Chicago
- University of Cincinnati Main Campus

Privacy Policy | FOIA | No F
Led by The National Science Foundation

h, Virginia 22230, USA

Growth areas

- A growing awareness of multi-lateral federation in other verticals – Wall Street, Financials, Real Estate,
- Government interactions
- Content providers
- Medical community
- Outsourced service providers
- Demand aggregation built on federation

Invention

- Some inventions work
- Some inventions don't
- A very, very few transform the world
 - Movable type
 - The Internet

Invention

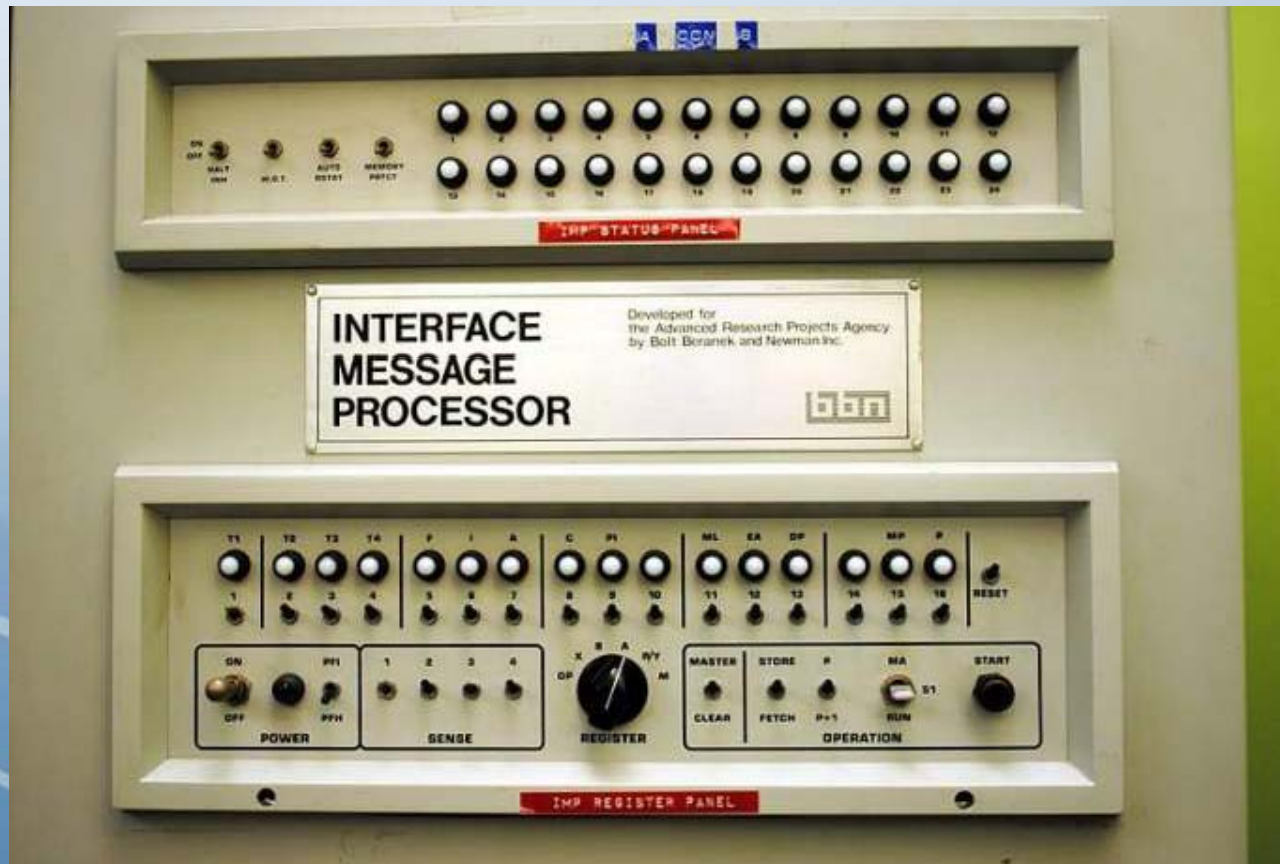
- "Everything that can be invented has been invented."
 - Charles H. Duell, Commissioner, U.S. Office of Patents, 1899.
- "I think there is a world market for maybe five computers."
 - Thomas Watson, chairman of IBM, 1943





kjk@internet2.edu









2. A brief history of Internet identity

- Classic PKI
- The development of federation
- The rise of social identity
- The need to support, and integrate, across the mix

Classical PKI

- Strong identity, no privacy
- Scalable, but not easily deployed
 - Challenges in technology, libraries, policy, economics
- Still required for some capabilities (signed email, document signing, etc)
- Has received some modifications
- Usage today slowly growing
 - US Gov PIV cards
 - Various national identity systems in Europe

The development of federation

- Beginning in 2000, coordinated activities in R&E sector and OASIS standards org
- SAML, Shibboleth, etc.
- Deployments beginning in 2004
- Very large government and corporate federations emerging (>6000 nodes)
- Multilateral in R&E; initially bilateral in corporate world, but changing

The rise of social identity

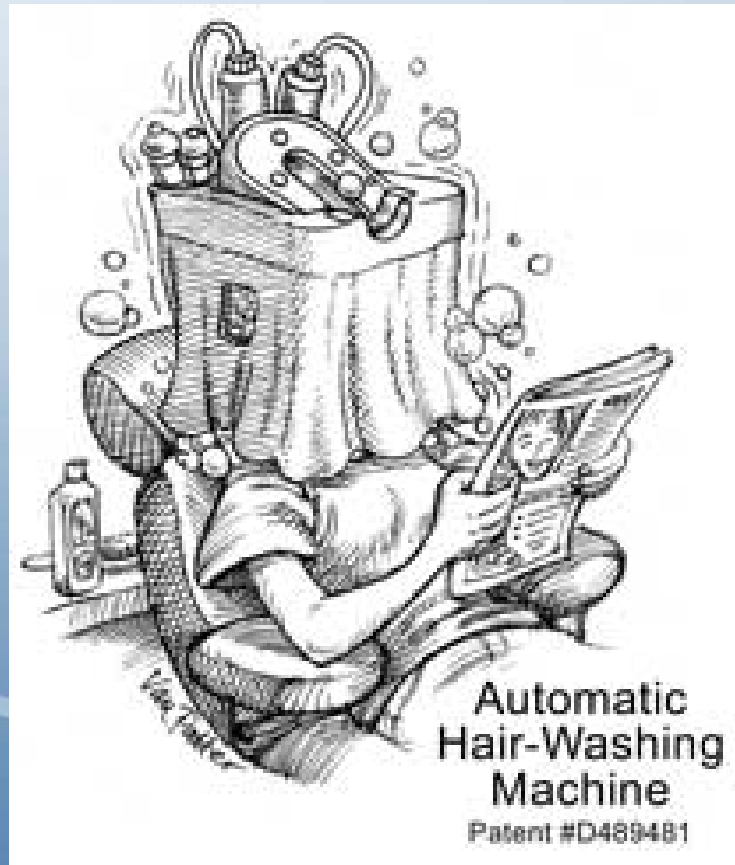
- Beginning with Internet Service providers in the early 2000s, gradually shifting to other social networking, search, and mail providers.
- Initially inconsistent and weak on privacy and security
- Improvements in infrastructure protocols and beginnings of a marketplace now
- Interest in value-added LOA use and an attribute authority service

Integration of Internet identity

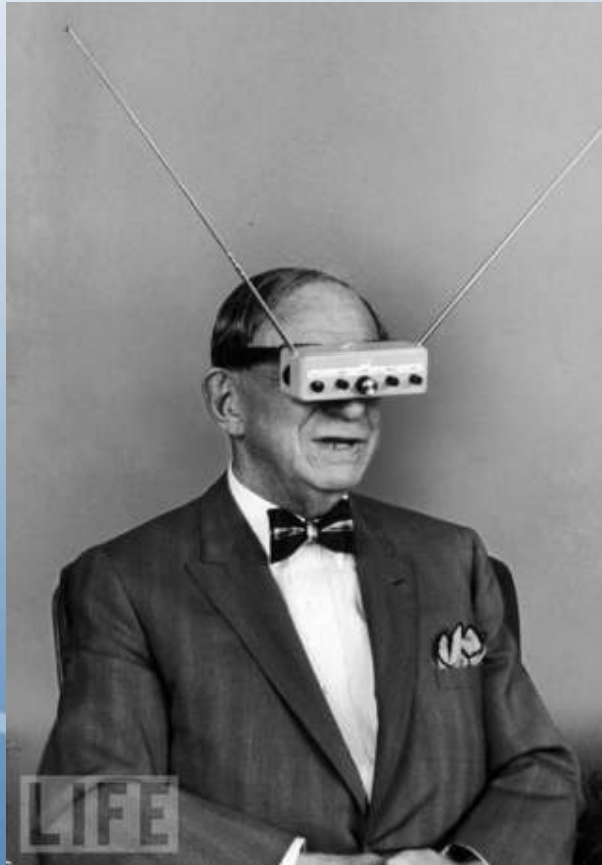
- Ability to deploy a variety of identity types to solve a use case
- Gateways and other approaches to credential conversion
 - PIV in federation
 - Social2SAML gateways
- Integration opportunities increase with OpenId Connect



And a patented invention







3. What we have learned from the practice

- Killer apps
- Level of authentication (LOA)
- Attributes
- Metadata

Killer Apps

- Wikis – for research, for education, for collaboration, for ...
- Netmeetings, Chats, etc
- Collaborative calendaring
- Command line, and many non-web, non human-based apps and services

LOA

- Killer Apps for LOA
- Components
 - What's hard in practice
- InCommon Approaches

Killer Apps for Higher LOA

- Official business
- Sensitive data
- High-risk instruments and devices
- Federal agency to federal agency
- Grant management and significant financial transactions

LOA Components

- Identity proofing
- Delivery of initial credential
- Authentication technologies and practices

What's Hard in Practice

- The extended communities to be served
 - Students, and their parents
 - Faculty, and their external colleagues
- Turing the “softness” of policy and law into hard coding
 - In privacy and consent
- Audit
 - Skills, costs

InCommon Approaches

- Basic – current practices, good privacy
- Bronze – LOA 1 – basic, self-asserted compliance, good privacy
- Silver – LOA 2 - rich guidelines, but with local interpretation, audit, good privacy
- Gold – LOA 3 – strong authentication + Silver
- Organizations run at multiple levels, serving different communities of internal users, matching LOA to a variety of external apps
- Audit available either from InCommon or Kantara approved

The importance of Metadata

- It enables the multilateral federation
- It creates scale and allows a marketplace
- It is essential for discovery, privacy management, policy management, etc.
- It needs to evolve from static bundles to a dynamic service

New elements of the metadata

- LOA supported
- Graphical icon for discovery
- Service attributes requirements
- Federated incident handling contacts
- Note that such extensions are set per federation, leading to future issues

Attributes

- Importance
- Types and PII
- Bundles and application categories
- Aggregation
- Their Tao

The importance of attributes

- The need to scale access control
 - Federated identity creates the requirement
 - Number of sites, number of users, etc...
 - The value of roles in regulated verticals
- The need to provide security and protect privacy
 - Put the organization and the user in control
 - The need to provide secrecy
- New and deeply tangled policy area, both nationally and internationally
- Tools, and so issues, are coming...

Types by attribute authorities

- Institutional/enterprise
 - User who has an established, authenticated identity
 - Organizational roles and groups
 - Reassertion of other official credentials (e.g. citizenship, age, etc.)
- Governmental
- Temporal – geolocation, etc.
- Community or collaboration asserted
 - Formal – Virtual organizations, groups
 - Informal – Reputation systems, friends of friends
- Self-asserted – Preferred language, accessibility

R&E basic attributes (eduPerson et al)

- High-level affiliation (eg, member, faculty, staff, student)
- Opaque, persistent and non-correlating identifiers (ePTID)
- A persistent and human-usable identifier (eg, kjk@internet2.edu)
- Name (e.g. Display Name)
- Email address
- An open-ended set of entitlements assigned by the institution, including group membership

PII and One PII taxonomy

Personally Identifiable Information (PII) is a common but abstract term

One Taxonomy trying to make it more solid -

- 0) Attributes that do not identify a unique user (e.g. ePSA)
- 1) Indirect identifiers designed for privacy (e.g. ePTID)
- 2) Indirect identifiers not designed for privacy (e.g. IP address)
- 3) Direct identifiers (e.g. name, address)
- 4) E-mail address & fax number
- 5) Location information (e.g. mobile phone cell)
- 6) Sensitive personal data (health, race, religion, etc.)

Bundles and Application Categories

- Attributes tend to travel in bundles
 - The R&S (research and scholarship) bundle
 - {name, email, authenticated identity, affiliation}
 - Applications are being vetted for minimal use and qualification for R&S
 - Attribute release automatic
- Several bundles are likely, e.g. {opaque-id, affiliation}, {authentication only}

Attribute aggregation

- A very common real world requirement
- Can be done at IdP, SP, or intermediary
 - Code supports it
 - Installations are using it
- Can complicate or simplify trust
- Leads to lots of issues, including reconciling conflicting values, handling incomplete data, etc

The Tao of Attributes 属性之道

- A workshop run by NIH and Internet2 in Sept 2009 that established the existence of a Tao, and not much more
 - <http://middleware.internet2.edu/tao-of-attributes/>
- An ongoing individual sense of the looming importance of the issues, and an occasional sighting of the need
- Must be at one with the environment and the vertical it exists in
- Needs a home...

Some of the Tao

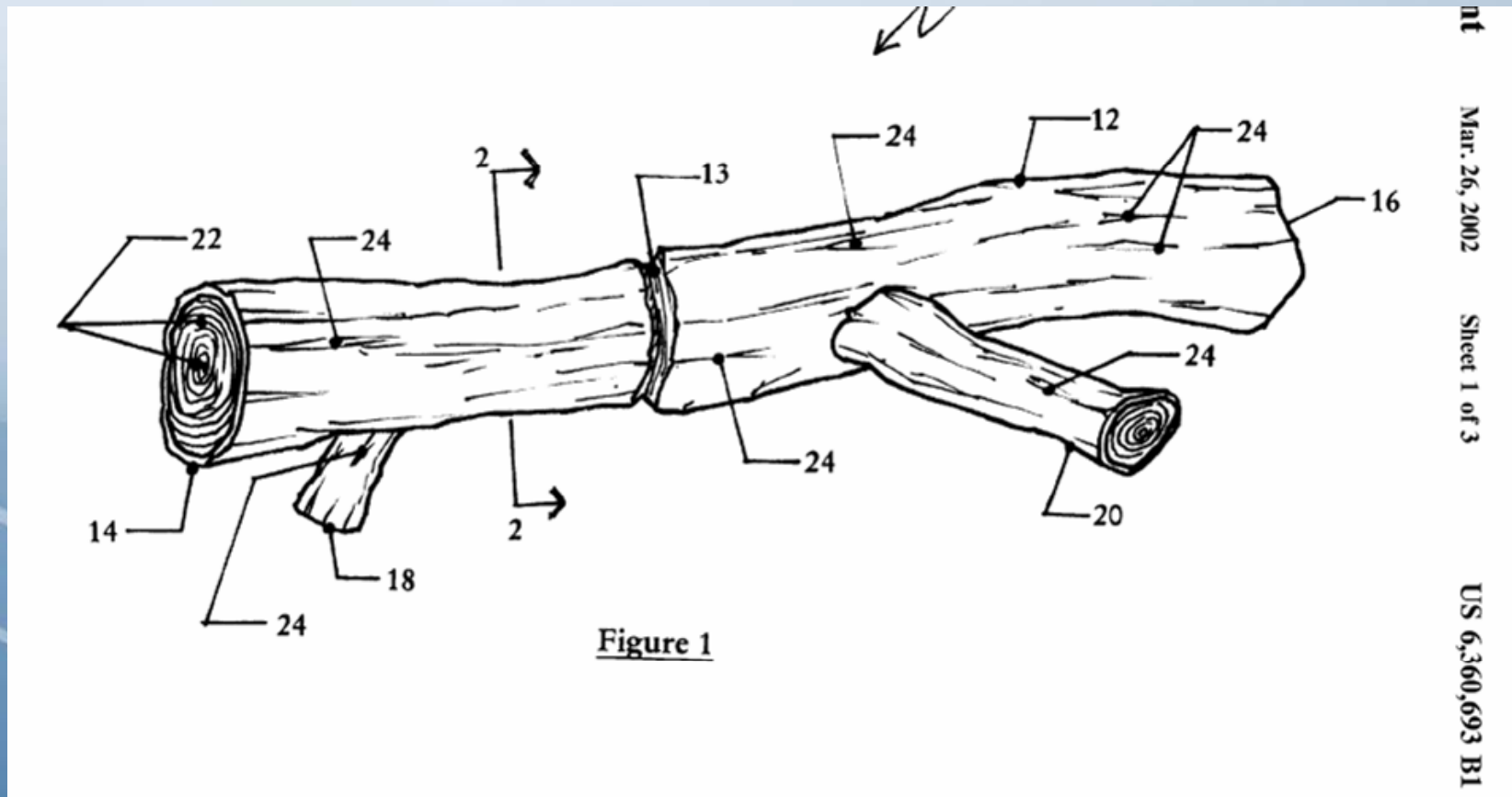
- Schema and Attribute bundles
- Complexity and Extensibility
 - Tagging
 - Complexity vs Metadata
- Attribute flows
 - Creation, transport
 - Attribute authorities and LOA of attributes
- Privacy, user attribute release and consent
 - IdP releasing vs SP asking
 - Query languages
 - The Challenges of Consent
 - Mastering least privilege/minimal release

Attribute Complexity and Extensibility

- Complexity
 - Tagging within attribute vs use of metadata vs context
 - Knowing which IdP to ask for which attributes, especially as we get into aggregation
- Extensibility
 - The ability to add new controlled values, and parameters like validation, date, terms of use
 - How much flat attribute proliferation can be managed through a structured data space?



Another patented vision





- “We won’t have more than 250 machines on the network”
 - Vint Cerf, an inventor of the Internet, 1970’s
- "640K ought to be enough for anybody."
 - Bill Gates, 1981

4. The Invention Ahead

- Working through the Tao
- Interfederation
- Groups and federated Groups
- Privacy managers

Interfederation

- Connecting autonomous federations
- Critical for global scaling, accommodating state and local federations, integration across vertical sectors
- Several operational “instances” – Kalmar2 Union, eduGAIN
- Has technical, financial and policy dimensions
- Key technologies moving forward – PEER, metadata enhancements and tools, discovery
- Has massive impact on the attribute ecosystem

Metadata and ecosystem evolution

- Handling static and dynamic metadata
- LOA of attributes
 - Specifying semantic rules
- Terms of use
- Time limits
- Query languages, etc – legal age

PEER – Public End-entity Registry

- A service instance of the metadata exchange protocol MDX, operated by REfeds with ISOC support
- A publish and subscribe model – end-points or federations publish to PEER and subscribe to custom, aggregated feeds
- Trust model depends on service instance; PEER is currently using DNS ownership
- Can accommodate any kind of federated metadata (including SAML and OpenId endpoints)
- Does limited syntax and semantic validation

Groups and federated groups

- Groups are the primary approach to access control
- In any community, the number of groups $>$ the number of identities
- Group management tools exist; API's are not yet standardized
- Groups with federated members exist, though enrollment may be tricky
- Federated groups (distinct, interacting groups between multiple organizations) the ultimate goal

Privacy, attribute release policies, consent

- Complex, sometimes contradictory requirements from governments around the world
- Whose national policy applies – IdP, SP, user?
- In federated identity, the key focus is around attribute release and consent
 - Some attributes required for transaction; some may be optional
 - Control points at service provider, at identity provider, and with the user
 - Consent at collection of information or release

This is the Digital ID Card to be sent to 'https://aai-demo.switch.ch':

Digital ID Card

Surname	SWITCHaai
Given name	Demouser
Unique ID	234567@example.org
User ID	demouser
Home organization	example.org
Home organization type	other
Affiliation	staff
Entitlement	http://example.org/res/99999 http://publisher-xy.com/e-journals

Don't show me this page again. I agree that my Digital ID Card (possibly including more data than shown above) will be sent automatically in the future.

Cancel

Confirm

5. Conclusions

- Understanding the ecosystem
 - IdP's, SP's, Gateways and middlemen
 - Attribute authorities and data aggregators
 - Virtual organizations and collaborations
- Making a marketplace
 - Establishing the rules of the market
 - Creating competitive service providers



kjk@internet2.edu



An invention to change a lightbulb

