W3C IDIB submission - browser support for large-scale federation

Title: Browser support for identity federation with many identity providers

Abstract: One of the biggest challenges in large-scale federation is IdP discovery. Browser support of data standards for RP and IdP representation, and participation in IdP selection, would help make discovery more scalable, secure, and protocol-independent.

The global higher-education and research community has been using federated signon for several years to collaborate among institutions and projects, and access outsourced services. There are currently about 30 national R&HE federations around the world (see http://refeds.org/resources_list.html) representing several thousand institutions and tens of millions of users. These federations are primarily using SAML technology. In many ways this infrastructure embodies much of the vision of open Internet identity: thousands of peer identity providers (IdPs) representing disparate communities, whose users are able to access myriad resources with their preferred identities. This stands in contrast to the use of open/federated identity in the consumer world, which is dominated by a few very large IdPs with essentially no participation from smaller communities.

One of the significant challenges with this large-scale multi-way federation is IdP discovery, known in some circles as the "NASCAR problem". When a user with a standard web browser visits a relying party (RP) and wants to log in via her preferred IdP, negotiation must happen regarding authentication protocol, characteristics of IdPs, selection and transmission of user attributes (aka claims), how interaction with the IdP (if any) is initiated, etc. As the number of RPs and IdPs in the environment grows, presentations of the selection experience by RPs diverges; choosing the right IdP becomes harder for the user; RPs want to limit possible IdPs based on characteristics such as user attributes (aka claims) offered, cost of service, etc.

A wide variety of methods have been developed and deployed for dealing with the IdP discovery problem. One fairly widely-deployed approach embeds a discovery UI in the RP, with selectable/searchable IdP information derived from SAML federation metadata. Another approach used by some commercial services employs an intermediate (aka proxy) as an authentication endpoint between the RP and a set of IdPs. In both these approaches, as in other web UIs, JavaScript is used to provide a modern user experience with incremental search.

The Universal Login Experience (ULX, http://kantarainitiative.org/confluence/display/ulx/Home) Working Group in the Kantara Initiative has been working on standardizable methods for login initiation (including IdP discovery) that will be usable, work at large scale, be protocol-agnostic, support in-browser mechanisms, and work for many existing use cases.

Principles motivating the ULX work include:

- Protocol-specific UX elements are doomed to failure: As no one authentication protocol is dominant, login initiation must be protocol-independent, so sites can create a consistent experience and evolve without UI disruptions.
- Embrace and integrate traditional login methods with new methods: Local username/password is by far the most commonly used method; open/federated login UX must be able to co-exist with it and extend it.
- IdP selection can happen via many architectural elements: RP, separate IdP Selection Agent (ISA) web service, or in-browser mechanisms. Since any might be chosen at any time by some combination of RP and user, it is optimal to have UX consistency across the set.

Information Card (aka IMI) technology was a key motivator for the ULX WG work, since it necessitated thinking about how smart client capabilities could coexist with unmodified browser approaches in the same UX framework. IMI, it appears, is not going to succeed, but clearly this is not the end of browser enhancements to improve the login UX. A lesson that might be drawn from IMI's demise is that incremental improvements to existing successful approaches are more likely to be widely adopted. Hence ULX is currently focused on in-page login initiation and discovery that could credibly also be supported in browser chrome.

One element of proposed ULX output is a data format representing RP and IdP characteristics. It is similar in function to SAML metadata, but is a JSON format (rather than XML) so as to be more amenable to web UI development. It is also similar in some respects to the Account Management Control Document format proposed as part of the Mozilla Account Manager work. The data in such a document is intended to be input to an ISA (in RP, other web site, or browser) to support the user's IdP selection process. Data includes basic characteristics of the SP including protocol support; claims (aka user attributes) required by the SP; and information about preferred or potential IdPs.

An in-browser ISA permits inclusion of user preferences and stored behavior data in the selection process, and can provide a significant improvement in functionality over website-provided ISAs. This can be done as a usability improvement without changing the basic relationships between RP and IdP (as IMI did), hence can provide the desired incremental improvement.