Mon, 30 Jan 2017 13:49:11 -0500

DRAFT NIST Special Publication 800-63C

Digital Identity Guidelines

Federation and Assertions

Paul A. Grassi
Justin P. Richer
Sarah K. Squire
James L. Fenton

Privacy Authors:
Naomi B. Lefkovitz
Jamie M. Danker

Usability Authors:
Yee-Yin Choong
Kristen K. Greene
Mary F. Theofanos
Ellen M. Nadeau

COMPUTER SECURITY

National Institute of Standards and Technology
U.S. Department of Commerce

DRAFT NIST Special Publication 800-63C

Digital Identity Guidelines

Federation and Assertions

Paul A. Grassi
Applied Cybersecurity Division
Information Technology Laboratory

Justin P. Richer
Bespoke Engineering
Billerica, MA

Sarah K. Squire
Engage Identity
Seattle, WA

James L. Fenton
Altmode Networks
Los Altos, CA

Privacy Authors:

Naomi B. Lefkovitz
Applied Cybersecurity Division
Information Technology Laboratory

Jamie M. Danker
National Protection and Programs Directorate
Department of Homeland Security

Usability Authors:

Yee-Yin Choong
Kristen K. Greene
Mary F. Theofanos
Information Access Division
Information Technology Laboratory

Ellen M. Nadeau
Applied Cybersecurity Division
Information Technology Laboratory

Month TBD 2017

National Institute of Standards and Technology
Kent Rochford, Acting Under Secretary of Commerce for Standards and Technology and Director

Authority

This publication has been developed by NIST in accordance with its statutory responsibilities under the Federal Information Security Modernization Act (FISMA) of 2014, 44 U.S.C. § 3541 et seq., Public Law (P.L.) 113-283. NIST is responsible for developing information security standards and guidelines, including minimum requirements for federal information systems, but such standards and guidelines shall not apply to national security systems without the express approval of appropriate federal officials exercising policy authority over such systems. This guideline is consistent with the requirements of the Office of Management and Budget (OMB) Circular A-130.

Nothing in this publication should be taken to contradict the standards and guidelines made mandatory and binding on Federal agencies by the Secretary of Commerce under statutory authority. Nor should these guidelines be interpreted as altering or superseding the existing authorities of the Secretary of Commerce, Director of the OMB, or any other Federal official. This publication may be used by nongovernmental organizations on a voluntary basis and is not subject to copyright in the United States. Attribution would, however, be appreciated by NIST.

National Institute of Standards and Technology Special Publication 800-63C
Natl. Inst. Stand. Technol. Spec. Publ. 800-63C, xxx pages (MonthTBD 2017)
CODEN: NSPUE2

Certain commercial entities, equipment, or materials may be identified in this document in order to describe an experimental procedure or concept adequately. Such identification is not intended to imply recommendation or endorsement by NIST, nor is it intended to imply that the entities, materials, or equipment are necessarily the best available for the purpose.

There may be references in this publication to other publications currently under development by NIST in accordance with its assigned statutory responsibilities. The information in this publication, including concepts and methodologies, may be used by federal agencies even before the completion of such companion publications. Thus, until each publication is completed, current requirements, guidelines, and procedures, where they exist, remain operative. For planning and transition purposes, federal agencies may wish to closely follow the development of these new publications by NIST.

Organizations are encouraged to review all draft publications during public comment periods and provide feedback to NIST. Many NIST cybersecurity publications, other than the ones noted above, are available at http://csrc.nist.gov/publications (http://csrc.nist.gov/publications).

Reports on Computer Systems Technology

The Information Technology Laboratory (ITL) at the National Institute of Standards and Technology (NIST) promotes the U.S. economy and public welfare by providing technical leadership for the Nation's measurement and standards infrastructure. ITL develops tests, test methods, reference data, proof of concept implementations, and technical analyses to advance the development and productive use of information technology. ITL's responsibilities include the development of management, administrative, technical, and physical standards and guidelines for the cost-effective security and privacy of other than national security-related information in Federal information systems. The Special Publication 800-series reports on ITL's research, guidelines, and outreach efforts in information system security, and its collaborative activities with industry, government, and academic organizations.

Abstract

This document and its companion documents, SP 800-63-3, SP 800-63A, and SP 800-63B, provide technical and procedural guidelines to agencies for the implementation of federated identity systems and for assertions used by federations. This publication supersedes corresponding sections of NIST SP 800-63-1 and SP 800-63-2.

These guidelines provide technical requirements for Federal agencies implementing digital identity services and are not intended to constrain the development or use of standards outside of this purpose. This guideline focuses on the use of use of federated identity, and the use of assertions to implement identity federations. Federation allows a given credential service provider to provide authentication and (optionally) subscriber attributes to a number of separately administered relying parties. Similarly, relying parties may use more than one credential service provider.

Keywords

assertions; authentication; credential service provider; digital authentication; electronic authentication; electronic credentials; federations.

Acknowledgements

The authors would like to acknowledge the thought leadership and innovation of the original authors: Donna F. Dodson, Elaine M. Newton, Ray A. Perlner, W. Timothy Polk, Sarbari Gupta, and Emad A. Nabbus. Without their tireless efforts, we would not have had the incredible baseline from which to evolve 800-63 to the document it is today. In addition, special thanks to the Federal Privacy Council's Digital Authentication Task Force for the contributions to the development of privacy requirements and considerations.

Audience

Compliance with NIST Standards and Guidelines

Conformance Testing

Trademark Information

Requirements Notation and Conventions

The terms "SHALL" and "SHALL NOT" indicate requirements to be followed strictly in order to conform to the publication and from which no deviation is permitted.

The terms "SHOULD" and "SHOULD NOT" indicate that among several possibilities one is recommended as particularly suitable, without mentioning or excluding others, or that a certain course of action is preferred but not necessarily required, or that (in the negative form) a certain possibility or course of action is discouraged but not prohibited.

The terms "MAY" and "NEED NOT" indicate a course of action permissible within the limits of the publication.

The terms "CAN" and "CANNOT" indicate a possibility and capability, whether material, physical or causal or, in the negative, the absence of that possibility or capability.

Table of Contents

1. Purpose

This section is informative.

This recommendation and its companion documents, [SP 800-63-3] (sp800-63-3.html), [SP 800-63A] (sp800-63a.html), and [SP 800-63B] (sp800-63b.html), provide technical guidelines to credential service providers (CSPs) for the implementation of remote authentication.

This document, SP 800-63C, provides requirements to CSPs and relying parties (RPs) of federated identity systems. Federation allows a given CSP to provide authentication and (optionally) subscriber attributes to a number of separately administered RPs. Similarly, RPs may use more than one CSP.

2. Introduction

This section is informative.

Federation is a process that allows for the conveyance of authentication and subscriber attribute information across a set of networked systems. In a federation scenario, the verifier or CSP is referred to as an identity provider, or IdP. The RP is the party that receives and uses the information provided by the IdP.

To accomplish this task, federated identity systems use assertions. Assertions are statements from an IdP to an RP that contain information about a subscriber. Federation technology is generally used when the RP and the IdP are not a single entity or are not under common administration. The RP uses the information in the assertion to identify the subscriber and make authorization decisions about their access to resources controlled by the RP. An assertion typically includes an identifier for the subscriber, allowing association of the subscriber with their previous interactions with the RP. Assertions may additionally include attribute values or attribute claims that further characterize the subscriber and support the authorization decision at the RP. Additional attributes may also be available outside of the assertion as part of the larger federation protocol. These attribute values and attribute claims are often used in determining access privileges for Attribute Based Access Control (ABAC) or facilitating a transaction (e.g., shipping address).

In a federated identity scenario, the subscriber does not authenticate directly to the RP. Instead, the federation protocol defines a mechanism for an IdP to generate an assertion for the identifier associated with a subscriber, usually in response to a request from the RP. The IdP is responsible for authenticating the subscriber (though it may use session management as described in [Section 7 of SP 800-63B] (sp800-63b.html#sec7)). This process allows the subscriber to obtain services with multiple RPs without the need to hold or maintain separate credentials at each. This process can also be used to support *single sign on*, where subscribers authenticate once to an IdP and subsequently obtain services from multiple RPs.

Federation requires relatively complex multiparty protocols that have subtle security and privacy requirements requiring careful consideration. When evaluating a particular federation structure, it may be instructive to break it down into its component interactions. Generally speaking, authentication between the subscriber and the IdP will be based on the authentication mechanisms presented in SP 800-63B, while interactions between the IdP and RP will convey attributes established using procedures in SP 800-63A and other self-asserted attributes. Many of the requirements presented in this document, therefore, have some relationship with corresponding requirements in those two documents.

The following table states which sections of the document are normative and which are informative:

Section Name	**Normative/Informative**
1. Purpose	Informative
2. Introduction	Informative
3. Definitions and Abbreviations	Informative
4. Federation	Normative
5. Assertion	Normative
6. Assertion Presentation	Normative
7. Federation Assurance Level (FAL)	Normative
8. Security	Informative
9. Privacy Considerations	Informative
10. Usability Considerations	Informative
11. Assertion Examples	Informative
12. References	Informative

3. Definitions and Abbreviations

This section is informative.

There are a variety of definitions used in the area of authentication. While many terms are consistent with earlier revisions version of SP 800-63, some have changed in this revision. Since there is no single, consistent definition of many of these terms, careful attention to how the terms are defined here is warranted.

The definitions in this section are primarily those that are referenced in this document. Refer to the other documents in the SP 800-63 document family for additional definitions and abbreviations specific to their content.

Assertion

A statement from a verifier to an RP that contains identity information about a subscriber. Assertions may also contain verified attributes.

Assertion Reference

A data object, created in conjunction with an assertion, which identifies the verifier and includes a pointer to the full assertion held by the verifier.

Attribute

A quality or characteristic ascribed to someone or something.

Attribute Claim

A statement asserting a property of a subscriber without revealing all of the information in one or more attributes, independent of format. For example, for the attribute ‘birth date’, a claim could be ‘older than 18’ or ‘born in December’.

Attribute Value

A complete statement asserting a property of a subscriber, independent of format. For example, for the attribute ‘birthday’, a value could be ‘12/1/1980’ or ‘December 1, 1980’.

Authenticated Protected Channel

A communication channel that uses approved encryption where the initiator of the connection (client) has authenticated the recipient (server). Authenticated protected channels provide confidentiality and man-in-the-middle protection and are frequently used in the user authentication process. TLS [BCP 195] is an example of an authenticated protected channel when the certificate presented by the recipient is verified by the initiator.

Authentication

The process of establishing confidence in the identity of users or information systems.

Authentication Protocol

A defined sequence of messages between a claimant and a verifier that demonstrates that the claimant has possession and control of one or more valid authenticators to establish his/her identity. Secure authentication protocols also demonstrate to the claimant that he or she is communicating with the intended verifier.

Back-Channel Communication

Communication between two systems that relies on a direct connection (allowing for standard protocol-level proxies), without using redirects through an intermediary such as a browser. This can be accomplished using HTTP requests and responses.

Federation

A process that allows for the conveyance of identity and authentication information across a set of networked systems. These systems are often run and controlled by disparate parties in different network and security domains.

Federation Proxy

A component that acts as a logical RP to a set of IdPs and a logical IdP to a set of RPs, bridging the two systems with a single component. These are sometimes referred to as “brokers”.

Front-Channel Communication

Communication between two systems that relies on redirects through an intermediary such as a browser. This is normally accomplished by appending HTTP query parameters to URLs hosted by the receiver of the message.

Identity Provider (IdP)

The party that manages the subscriber’s primary authentication credentials and issues assertions derived from those credentials. This is commonly the CSP as discussed within this document suite.

Pairwise Pseudonymous Identifier

An opaque unguessable subscriber identifier generated by a CSP for use at a specific individual RP. This identifier is only known to and only used by one CSP-RP pair.

Relying Party (RP)

In this document, the party that receives and processes the assertion identifying the subscriber.

4. Federation

This section is normative.

In a federation protocol, a three-party relationship is formed between the subscriber, the IdP, and the RP as shown in Figure 4-1. Depending on the specifics of the protocol, different information passes among the participants at different times. The subscriber communicates with both the IdP and the RP, usually through a browser. The RP and the IdP communicate with each other in two ways:

- The *front channel*, through redirects involving the subscriber; or
- The *back channel*, through a direct connection between the RP and IdP, not involving the subscriber.

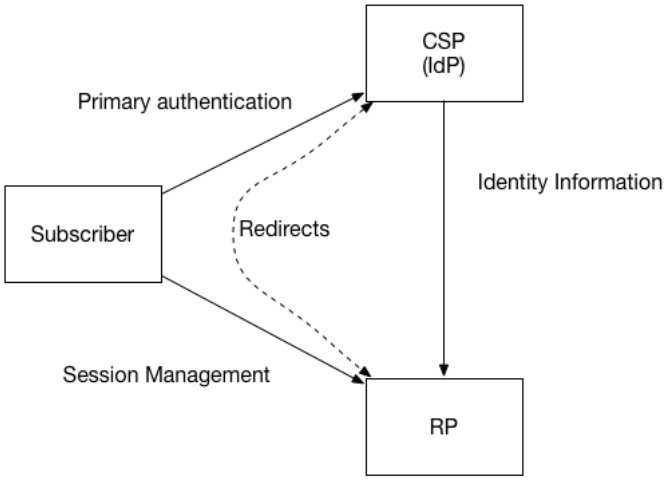

Figure 4-1. Federation

The subscriber authenticates to the IdP as described in [SP 800-63B] (sp800-63b.html), and then the result of that authentication event is asserted to the RP across the network. The IdP can also make attribute statements about the subscriber as part of this process. These attributes and authentication event information are carried to the RP through the use of an assertion, described in Section 5.

4.1. Federation Models

This section provides an overview of and requirements for common identity federation models currently in use. In each model, relationships are established between members of the federation in several different ways.

4.1.1. Manual Registration

In the manual registration model, the IdP and RP manually provision configuration information about parties with which they expect to interoperate. IdPs MAY configure RPs using an explicit whitelist, allowing services to transfer information as part of the authentication transaction. In such cases where an RP is not whitelisted, the IdP SHALL require appropriate runtime decisions to be made by an authorized party, such as the subscriber, before releasing user information.

IdPs and RPs MAY act as their own authorities of who to federate with or MAY externalize those authority decisions to an external party as in Section 4.1.3.

Protocols requiring the transfer of keying information SHALL use a secure method to establish such keying information needed to operate the federated relationship during the registration process, including any shared secrets or public keys. Any symmetric keys used in this relationship SHALL be unique to a pair of federation participants.

Federation relationships SHALL establish parameters regarding expected and acceptable identity assurance level (IAL) and authentication assurance level (AAL) in connection with the federated relationship.

4.1.2. Dynamic Registration

In the dynamic registration model of federation, it is possible for relationships between members of the federation to be negotiated at the time of a transaction. This process allows components to be connected together without manually establishing a connection between components. IdPs that support dynamic registration SHALL make their configuration information (such as dynamic registration endpoints) available in such a way as to minimize system administrator involvement.

Protocols requiring the transfer of keying information SHALL use a secure method to establish such keying information needed to operate the federated relationship during the registration process, including any shared secrets or public keys. Any symmetric keys used in this relationship SHALL be unique to a pair of federation participants.

IdPs SHALL require appropriate runtime decisions to be made by an authorized party, such as the subscriber, before releasing user information. An IdP accepting dynamically registered RPs MAY limit the types of attributes and other information made available to such RPs. An RP capable of dynamically registering MAY limit which IdPs it is willing to accept identity information from.

Frequently, parties in a dynamic registration model do not know each other ahead of time. Where possible, this SHOULD be augmented by *software statements*, which allow federated parties to cryptographically verify some attributes of the parties involved in dynamic registration. Software statements are lists of attributes describing the RP software, cryptographically signed by an authority (either the IdP itself, a federation authority as in Section 4.1.3, or another trusted party). This cryptographically verifiable statement allows the connection to be established or elevated between the federating parties without relying solely on self-asserted attributes. (See [RFC 7591] Section 2.3 for more information on one protocol's implementation of software statements.)

4.1.3. Federation Authorities

Some federated parties defer to an authority known as a *federation authority* to assist in making federation decisions and to establish the working relationship between parties. In this model, the federation authority generally conducts some level of vetting on each party in the federation to verify compliance with predetermined security and integrity standards.

Federation authorities SHALL establish parameters regarding expected and acceptable IAL and AAL in connection with the federated relationships they enable. Federation authorities SHALL individually vet each participant in the federation to determine that they adhere to their expected security, identity, and privacy standards. Vetting of IdPs and RPs SHALL establish, as a minimum, that:

- Assertions generated by IdPs adhere to the requirements in Section 5.
- RPs adhere to IdP requirements for the handling of subscriber attribute data, such as retention, aggregation, and disclosure to third parties.
- RP and IdP systems use approved profiles of federation protocols.

Federation authorities MAY assist the technical connection and configuration process between members, such as by publishing configuration data for IdPs or issuing software statements for RPs.

Most federations managed through authorities have a simple membership model: either parties are in the federation or they are not. However, more sophisticated federations MAY have multiple tiers of membership which can be used by federated parties to tell whether other parties in the federation have been more thoroughly vetted. IdPs MAY decide that certain subscriber information is only releasable to RPs in higher tiers, and RPs MAY decide to accept certain information only from IdPs in higher tiers. The nature and structure of such a multi-tiered system is outside the scope of this document.

4.1.4. Proxied Federation

In a proxied federation, communication between the IdP and the RP is intermediated in a way that prevents direct communication between the two parties. There are multiple methods to achieve this effect; common configurations include:

- A third party that acts as a federation proxy (or *broker*)
- A network of nodes that distributes the communications

Where proxies are used, they function as an IdP on one side and an RP on the other side. Therefore, all normative requirements that apply to IdPs and RPs SHALL apply to proxies in their respective roles.

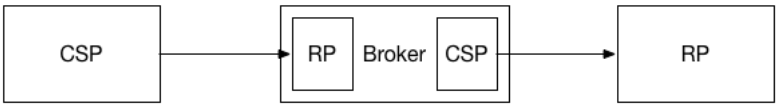

Figure 4-2. Federation Proxy

A proxied federation model can provide several benefits. Federation proxies can simplify technical integration between the RP and IdP by providing a common interface for integration. Additionally, to the extent a proxy effectively blinds the RP and IdP from each other, it can provide some business confidentiality for organizations that want to guard their subscriber lists from each other. Proxies can also mitigate some of the privacy risks of federation described in Section 4.2 below.

See Section 9.5 for further information on blinding techniques, their uses, and limitations.

4.1.5. Runtime Decisions

The fact that parties have federated SHALL NOT be interpreted as permission to pass information. Federated parties MAY establish whitelists of other federated parties who authenticate subscribers or pass information about them without runtime authorization from the subscriber. Federated parties MAY also establish blacklists of other federated parties who are not allowed to pass information about subscribers at all. Every party that is not on a whitelist or a blacklist SHALL be placed by default in a gray area where runtime authorization decisions will be made by an authorized party, often the subscriber.

To mitigate the risk of unauthorized exposure of sensitive information (e.g., shoulder surfing), the IdP SHALL, by default, mask sensitive information displayed to the subscriber. The IdP SHALL provide mechanisms for the subscriber to temporarily unmask such information in order for the subscriber to view full values. The IdP SHALL provide effective mechanisms for redress of applicant complaints or problems (e.g., subscriber identifies an inaccurate attribute value). For more details on masking and redress, please see Section 10 on usability considerations.

The subscriber SHALL receive explicit notice and be able to provide positive confirmation before any attributes about the subscriber are transmitted to any RP. At a minimum, the notice SHOULD be provided by the party in the position to provide the most effective notice and obtain confirmation, consistent with Section 9.2. If the protocol in use allows for optional attributes, the subscriber SHALL be given the option to decide whether to transmit those attributes to the RP. An IdP MAY employ mechanisms to remember and re-transmit the exact attribute bundle to the same RP.

4.2. Privacy Requirements

Federation involves the transfer of personal attributes from a third party, the IdP, that is not otherwise involved in a transaction. Federation also potentially gives the IdP broad visibility into subscriber activities. Accordingly, there are specific privacy requirements associated with federation.

Communication between the RP and the IdP could reveal to the IdP where the subscriber is conducting a transaction. Communication with multiple RPs allows the IdP to build a profile of subscriber transactions that would not have existed without federation. This aggregation could enable new opportunities for subscriber tracking and use of profile information that do not always align with the privacy interests of subscribers.

The IdP SHALL NOT disclose information on subscriber activities at an RP to any party, nor use the information for any purpose other than federated authentication, to comply with law or legal process, or in the case of a specific user request for the information. The IdP SHOULD employ technical measures, such as the use of pairwise pseudonymous identifiers described in Section 5.2.5 or privacy-enhancing cryptographic protocols, to provide unlinkability and discourage subscriber activity tracking and profiling.

An IdP MAY disclose information on subscriber activities to other RPs within the federation for security purposes such as communication of compromised subscriber accounts.

The following requirements apply specifically to agencies:

1. The agency SHALL consult with their Senior Agency Official for Privacy (SAOP) to conduct and analysis to determine whether the agency that is acting as either an IdP or an RP in an identity federation triggers the requirements of the Privacy Act.
2. The agency SHALL publish or identify coverage by a System of Records Notice (SORN) as applicable.
3. The agency SHALL consult with their SAOP to conduct an analysis to determine whether the agency that is acting as either an IdP or an RP in an identity federation triggers the requirements of the E-Government Act.
4. The agency SHALL publish or identify coverage by a Privacy Impact Assessment (PIA) as applicable.

4.3. Reauthentication and Session Requirements in Federated Environments

In a federated environment, the RP manages its sessions separately from any sessions at the IdP. The session at the RP starts when the RP processes the federation protocol from the IdP. At the time of a federated login, the subscriber MAY have an existing session at the IdP which MAY be used as part of the authentication process to the RP. The IdP SHALL communicate any information it has regarding the time of the latest authentication event at the IdP, and the RP MAY use this information in determining its access policies. Depending on the capabilities of the federation protocol in use, the IdP SHOULD allow the RP to request that the subscriber re-authenticate at the IdP as part of a federation request.

Due to the distributed nature of a federated system, the subscriber is capable of terminating sessions with the IdP and RP independently of one another. The RP SHALL NOT assume that the subscriber has an active session at the IdP past the establishment of the federated log in. The IdP SHALL NOT assume that termination of the subscriber's session at the IdP will propagate to any sessions that subscriber would have at downstream RPs.

See Section 7 of 800-63B (sp800-63b.html#sec7) for more information about session management requirements.

5. Assertions

This section is normative.

An assertion is a packaged set of attribute values or attribute claims about or associated with an authenticated subscriber that is passed from the IdP to the RP in a federated identity system. Assertions contain a variety of information, including assertion metadata, attribute values and attribute claims about the subscriber, and other information that the RP can leverage, such as restrictions, and expiration time.

Assertions MAY represent only an authentication event, or MAY also represent attribute values and attribute claims regarding the subscriber.

All assertions SHALL include the following assertion metadata:

- Subject - An identifier for the party that the assertion is about (the subscriber), usually within the namespace control of the issuer (the IdP).
- Issuer - An identifier for the IdP that issued the assertion.
- Audience - An identifier for the party intended to consume the assertion (the RP).
- Issuance - A timestamp indicating when the assertion was issued by the IdP.
- Expiration - A timestamp indicating when the assertion expires and SHALL no longer be accepted as valid by the RP (note that this is the expiration of the assertion and not the expiration of the session at the RP).
- Identifier - A value uniquely identifying this assertion, used to prevent attackers from replaying prior assertions.
- Signature - Digital signature or message authentication code (MAC), including key identifier or public key associated with the IdP, for the entire assertion.
- Authentication Time - A timestamp indicating when the IdP last verified the presence of the subscriber at the IdP through a primary authentication event (if available).

Assertions MAY also include the following information:

- Key binding - Public key or key identifier of a key held by the subscriber to demonstrate their binding with the assertion.
- Attribute values and attribute claims - Information about the subscriber.
- Attribute metadata - Additional information about one or more subscriber attributes, such as that described in [NISTIR 8112].

Assertions SHOULD specify the AAL when an authentication event is being asserted and IAL when identity proofed attributes or claims based thereon are being asserted. The IAL and AAL MAY be specified in an alternate form, such as a composite level of assurance. If not specified, the RP SHALL NOT assign any specific IAL or AAL to the assertion.

Assertions MAY include additional attributes. Refer to Section 6 for privacy requirements on presenting attributes in assertions. The RP MAY fetch additional identity attributes from the IdP in one or more separate transactions using an authorization credential issued alongside the original assertion. The ability to successfully fetch such additional attributes SHALL NOT be treated as equivalent to processing of the assertion.

Although details vary based on the exact federation protocol in use, an assertion SHOULD be used only to represent a single login event at the RP. After the RP consumes the assertion, session management (sp800-63b.html#sec7) by the RP comes into play; the assertion SHALL NOT be used past the expiration time contained therein. However, the expiration of the session at the RP MAY occur prior to the expiration of the assertion. See Section 4.3 for more information.

5.1. Assertion Binding

An assertion can be classified based on whether presentation by a claimant of an assertion reference or the assertion itself is sufficient for establishing the binding between the subscriber and the assertion, or if a stronger binding is required.

5.1.1. Bearer Assertions

A bearer assertion can be presented by any party as proof of the bearer’s identity. If an attacker is able to capture or manufacture a valid assertion or assertion reference representing a subscriber, and that attacker is able to successfully present that assertion or reference to the RP, then the attacker MAY be able to impersonate the subscriber at that RP.

Note that mere possession of a bearer assertion or reference is not always enough to impersonate a subscriber. For example, if an assertion is presented in the back-channel federation model (described in Section 6.1), additional controls MAY be placed on the transaction (such as identification of the RP and assertion injection protections) that help to further protect the RP from fraudulent activity.

5.1.2. Holder-of-Key Assertions

A holder-of-key assertion contains a reference to a key possessed by and representing the subscriber, and the subscriber SHALL prove possession of that key in addition to presentation of the assertion itself. The key MAY be a symmetric key or a public key that corresponds to a private key. An assertion containing a reference to a key held by the subscriber for which key possession has not been proven SHALL be considered a bearer assertion by the RP.

The key referenced in a holder-of-key represents the subscriber, not any other party in the system including the browser, IdP, or RP. This key MAY be distinct from any key used by the subscriber to authenticate to the IdP.

In proving possession of the subscriber's secret, the claimant also proves with a certain degree of assurance that they are the rightful subject of the assertion. It is more difficult for an attacker to use a stolen holder-of-key assertion issued to a subscriber, since the attacker would need to steal the referenced key material as well.

Note that the reference to the key is asserted (and signed) by the issuer of the assertion; reference to a given key SHALL be trusted at the same level as all other information within the assertion.

The assertion SHALL NOT include an unencrypted private or symmetric key to be used with holder-of-key presentation. The RP MAY verify the claimant's possession of the key in conjunction with the IdP, for example, by requesting that the IdP verify a signature or MAC calculated by the claimant in response to a cryptographic challenge.

5.2. Assertion Protection

Independent of the binding mechanism (discussed above) or the federation model used to obtain them (described in Section 4), assertions SHALL include an appropriate set of protections to prevent attackers from manufacturing valid assertions or reusing captured assertions at disparate RPs.

5.2.1. Assertion Identifier

Assertions SHALL be sufficiently unique to permit unique identification by the target RP. Assertions MAY accomplish this by use of an embedded nonce, issuance timestamp, assertion identifier, or a combination of these or other techniques.

5.2.2. Signed Assertion

Assertions SHALL be cryptographically signed by the issuer (IdP). The RP SHALL validate the digital signature or MAC of each such assertion based on the issuer's key. This signature SHALL cover all vital fields of the assertion, including its identifier, issuer, audience, subject, and expiration.

The assertion signature SHALL either be a digital signature using asymmetric keys or a MAC using a symmetric key shared between the RP and issuer. Shared symmetric keys used for this purpose by the IdP SHALL be independent for each RP to which they send assertions, and are normally established during registration of the RP. Public keys for verification of digital signatures MAY be fetched by the RP in a secure fashion at runtime, such as through an HTTPS URL hosted by the IdP. Approved cryptography SHALL be used.

5.2.3. Encrypted Assertion

Assertions MAY be encrypted so as to allow only the intended audience to decrypt the contents. The IdP SHALL encrypt the contents of the assertion using either the RP's public key or a shared symmetric key. Shared symmetric keys used for this purpose by the IdP SHALL be independent for each RP to which they send assertions, and are normally established during registration of the RP. Public keys for encryption MAY be fetched by the IdP in a secure fashion at runtime, such as through an HTTPS URL hosted by the RP.

All encryption of assertions SHALL use approved cryptography.

5.2.4. Audience Restriction

Assertions SHALL use audience restriction techniques to allow an RP to recognize whether or not it is the intended target of an issued assertion. All RPs SHALL check the audience of an assertion, if provided, to prevent the injection and replay of an assertion generated for one RP at another RP.

5.2.5. Pairwise Pseudonymous Identifiers

In some circumstances, it is desirable to prevent the subscriber's account at the IdP from being easily linked at multiple RPs through use of a common identifier. When using pairwise pseudonymous subject identifiers within the assertions generated by the IdP for the RP, the IdP SHALL generate a different identifier for each RP as described in Section 5.2.6 below.

When pairwise pseudonymous identifiers are used with RPs alongside attributes, it may still be possible for multiple colluding RPs to reidentify a subscriber by correlation across systems using these identity attributes. For example, if two independent RPs each see the same subscriber identified with different pairwise pseudonymous identifiers, they could still determine that the subscriber is the same person by comparing the name, email address, physical address, or other identifying attributes carried alongside the pairwise pseudonymous identifier in the respective assertions. Privacy policies MAY prohibit such correlation, and pairwise pseudonymous identifiers can increase effectiveness of these policies by increasing the administrative effort in managing the attribute correlation.

Note that in a proxied federation model, the initial IdP may be unable to generate a pairwise pseudonymous identifier for the ultimate RP, since the proxy could blind the IdP from knowing which RP is being accessed by the subscriber. In such situations, the pairwise pseudonymous identifier is generally established between the IdP and the federation proxy itself. The proxy, acting as an IdP, can itself provide pairwise pseudonymous identifiers to downstream RPs. Depending on the protocol, the federation proxy may need to map the pairwise pseudonymous identifiers back to the associated identifiers from upstream IdPs in order to allow the identity protocol to function. In such cases, the proxy will be able to track and determine which pairwise pseudonymous identifiers represent the same subscriber at different RPs. The proxy SHALL NOT disclose the mapping between the pairwise pseudonymous identifier and any other identifiers to a third party or use the information for any purpose other than federated authentication, to comply with law or legal process, or in the case of a specific user request for the information.

5.2.6. Pairwise Pseudonymous Identifier Generation

Pairwise pseudonymous identifiers SHALL be opaque, containing no identifying information about the subscriber. They SHALL also be unguessable by a party having access to some information identifying the subscriber. Pairwise pseudonymous identifiers MAY be generated randomly and assigned to subscribers by the IdP or MAY be derived from other subscriber information if the derivation is done in an irreversible, unguessable manner (e.g., using a keyed hash function with a secret key). Normally, the identifiers SHALL only be known by and used by one pair of endpoints (e.g., IdP-RP). However, an IdP MAY generate the same identifier for a subscriber at multiple RPs at the request of those RPs, provided:

- Those RPs have a demonstrable relationship that justifies an operational need for the correlation, such as a shared security domain or shared legal ownership; and
- All RPs sharing an identifier consent to being correlated in such a manner.

The RPs SHALL conduct a privacy risk assessment to consider the privacy risks associated with requesting a common identifier. The IdP SHALL ensure that only intended RPs are correlated; otherwise, a rogue RP could learn of the pseudonymous identifier for a correlation by fraudulently posing as part of that correlation.

6. Assertion Presentation

This section is normative.

Assertions MAY be presented in either a *back-channel* or *front-channel* manner from the IdP to the RP. There are tradeoffs with each model, but both require the proper validation of the assertion. Assertions MAY also be proxied to facilitate federation between IdPs and RPs under specific circumstances, as discussed in Section 4.1.4.

The IdP SHALL transmit only those attributes that were explicitly requested by the RP. RPs SHALL conduct a privacy risk assessment when determining which attributes to request.

6.1. Back-channel Presentation

In the *back-channel* model, the subscriber is given an assertion reference to present to the RP, generally through the front channel. The assertion reference itself contains no information about the subscriber and SHALL be resistant to tampering and fabrication by an attacker. The RP presents the assertion reference to the IdP, usually along with authentication of the RP itself, to fetch the assertion.

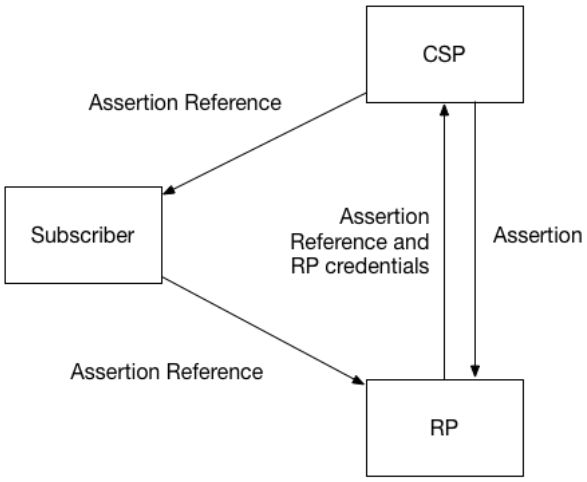

Figure 6-1. Back-channel Presentation

The assertion reference:

- SHALL be limited to use by a single RP.
- SHALL be single-use.
- SHOULD be time limited with a short lifetime of seconds or minutes.
- SHOULD be presented along with authentication of the RP.

In this model, the assertion itself is requested directly from the IdP to the RP, minimizing chances of interception and manipulation by a third party (including the subscriber themselves).

This method also allows the RP to query the IdP for additional attributes about the subscriber not included in the assertion itself, since back-channel communication can continue to occur after the initial authentication transaction has completed.

In the back-channel method, there are more network transactions required, but the information is limited to the parties that need it. Since an RP is expecting to get an assertion only from the IdP directly, the attack surface is reduced it is more difficult to inject assertions directly into the RP.

The RP SHALL protect itself against injection of manufactured or captured assertion references by use of cross-site scripting protection or other accepted techniques.

Claims within the assertion SHALL be validated including issuer verification, signature validation, and audience restriction.

Conveyance of the assertion reference from the IdP to the subscriber as well as from the subscriber to the RP SHALL be made over an authenticated protected channel. Conveyance of the assertion reference from the RP to the IdP as well as the assertion from the IdP to the RP SHALL be made over an authenticated protected channel.

Presentation of the assertion reference at the IdP SHOULD require authentication of the RP before an assertion is issued.

6.2. Front-channel Presentation

In the *front-channel* model, the IdP creates an assertion and sends it to the subscriber after successful authentication. The assertion is used by the subscriber to authenticate to the RP, often through mechanisms within the subscriber's browser.

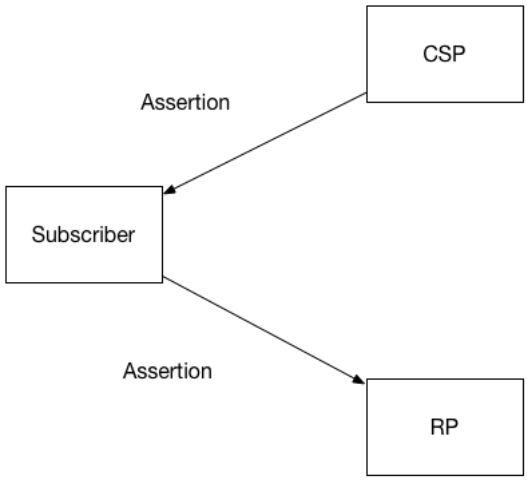

Figure 6-2. Front-channel Presentation

In the front-channel method, an assertion is visible to the subscriber, which could potentially cause leakage of system information included in the assertion. Furthermore, in this model it is more difficult for the RP to query the IdP for additional attributes after the presentation of the assertion.

Since the assertion is under the control of the subscriber, the front-channel presentation method also allows the subscriber to submit a single assertion to unintended parties, perhaps by a browser replaying an assertion at multiple RPs. Even if the assertion is audience restricted and rejected by unintended RPs, its presentation at unintended RPs could lead to leaking information about the subscriber and their online activities. Though it is possible to intentionally create an assertion designed to be presented to multiple RPs, this method can lead to lax audience restriction of the assertion itself, which in turn could lead to privacy and security breaches for the subscriber across these RPs. Such multi-RP use is not recommended. Instead, RPs are encouraged to fetch their own individual assertions.

The RP SHALL protect itself against injection of manufactured or captured assertions by use of cross-site scripting protection or other accepted techniques.

Claims within the assertion SHALL be validated including issuer verification, signature validation, expiration, and audience restriction.

Conveyance of the assertion from the IdP to the subscriber as well as from the subscriber to the RP SHALL be made over an authenticated protected channel.

6.3. Protecting Information

Communications between the IdP and the RP SHALL be protected in transit using an authenticated protected channel. Communications between the subscriber and either the IdP or the RP (usually through a browser) SHALL be made using an authenticated protected channel.

Note that the IdP may have access to information that may be useful to the RP in enforcing security policies, such as device identity, location, system health checks, and configuration management. If so, it may be a good idea to pass this information along to the RP within the bounds of the subscriber's privacy preferences.

Additional attributes about the user MAY be included outside of the assertion itself as part of a separate authorized request from the RP to the IdP. The authorization for access to these attributes MAY be issued alongside the assertion itself. Splitting user information in this manner can aid in protecting user privacy and allow for limited disclosure of identifying attributes on top of the essential information in the authentication assertion itself.

The RP SHALL, where feasible, request attribute claims rather than full attribute values. The IdP SHALL support attribute claims.

7. Federation Assurance Level (FAL)

This section is normative.

This section defines allowable Federation Assurance Levels, or FAL. The FAL describes aspects of the assertion and federation protocol used in a given transaction. These levels can be requested by an RP or required by configuration of both RP and IdP for a given transaction.

The FAL combines aspects of assertion protection and assertion presentation into an ordinal measurement scale applicable across different federation models. All assertions SHALL comply with the requirements in Section 5. While many other combinations of factors are possible, this list is intended to provide clear implementation recommendations representing increasingly secure deployment choices. Combinations of aspects not found in the FAL table are possible but outside the scope of this document.

This table presents different requirements depending on whether the assertion is presented through either the front channel or the back channel (via an assertion reference). Each successive level subsumes and fulfills all requirements of lower levels. Federations presented through a proxy SHALL be represented by the lowest level used during the proxied transaction.

Table 7-1. Federation Assertion Levels

FAL	Requirement
1	Bearer assertion, signed by IdP.
2	Bearer assertion, signed by IdP and encrypted to RP.
3	Holder of key assertion, signed by IdP and encrypted to RP.

For example, FAL1 maps to the OpenID Connect Basic Client profile or SAML (Security Assertion Markup Language) Web SSO Artifact Binding profile, with no additional features. FAL2 additionally requires that the OpenID Connect ID Token or SAML Assertion be encrypted to a public key representing the RP in question. FAL3 requires the presentation of an additional key bound to the assertion (e.g., the use of a cryptographic authenticator) along with all requirements of FAL2. Note that the additional key presented at FAL3 need not be the same key used by the subscriber to authenticate to the IdP.

Regardless of what is requested or required by the protocol, the FAL in use is easily detected by the RP by observing the nature of the assertion as it is presented as part of the federation protocol. Therefore, the RP is responsible for determining which FALs it is willing to accept for a given authentication transaction and ensuring that the transaction meets the requirements of that FAL.

If the RP is using a front-channel presentation mechanism (e.g., the OpenID Connect Implicit Client profile or the SAML Web SSO profile), it SHOULD require FAL2 or greater in order to protect the information in the assertion from the browser or other parties in the transaction.

Table 7-2 lists strict adherence to M-04-04 Level of Assurance, mapping the corresponding FALs.

Table 7-2. Legacy M-04-04 FAL Requirements

M-04-04 Level of Assurance (LOA)	Federation Assurance Level (FAL)
1	1
2	2
3	2
4	3

However, Table 7-3 shows the expanded set of FAL's that are allowable to meet M-04-04 Level of Assurance. Agencies SHALL select the corresponding FAL based on the assessed M-04-04 LOA.

Table 7-3. Recommended M-04-04 FAL Requirements

M-04-04 Level of Assurance	Federation Assurance Level
1	1, 2, or 3
2	2, or 3
3	2, or 3

M-04-04 Level of Assurance	Federation Assurance Level
4	3

7.1. Key Management

At any FAL, the IdP SHALL ensure that an RP is unable to impersonate the IdP at another RP by protecting the assertion with a signature and key using approved cryptography. If the assertion is protected by a digital signature using an asymmetric key, the IdP MAY use the same public and private key pair to sign assertions to multiple RPs. The IdP MAY publish its public key in a verifiable fashion, such as at an HTTPS-protected URL at a well-known location. If the assertion is protected by a MAC using a shared key, the IdP SHALL use a different shared key for each RP.

8. Security

This section is informative.

IdPs, RPs, subscribers, and parties outside of a typical assertions transaction may be malicious or become compromised. An attacker might have an interest in modifying or replacing an assertion to obtain a greater level of access to a resource or service provided by an RP. They might be interested in obtaining or modifying assertions and assertion references to impersonate a subscriber or access unauthorized data or services. Furthermore, it is possible that two or more entities may be colluding to attack another party. An attacker may attempt to subvert assertion protocols by directly compromising the integrity or confidentiality of the assertion data. For the purpose of these types of threats, authorized parties who attempt to exceed their privileges may be considered attackers. This section lists some common attacks against assertion transmission transactions.

- *Assertion manufacture/modification* - An attacker generates a forged assertion or modifies the content of an existing assertion (such as the authentication or attribute statements), causing the RP to grant inappropriate access to the subscriber. For example, an attacker may modify the assertion to extend the validity period and keep using an assertion; or a subscriber may modify the assertion to have access to information that they should not be able to view.
- *Assertion disclosure* - Assertions may contain authentication and attribute statements that include sensitive subscriber information. Disclosure of the assertion contents can make the subscriber vulnerable to other types of attacks.
- *Assertion repudiation by the IdP* - An assertion may be repudiated by an IdP if the proper mechanisms are not in place. For example, if an IdP does not digitally sign an assertion, the IdP can claim that it was not generated through the services of the IdP.
- *Assertion repudiation by the subscriber* - Since it is possible for a compromised or malicious IdP to issue assertions to the wrong party, a subscriber can repudiate any transaction with the RP that was authenticated using only a bearer assertion.
- *Assertion redirect* - An attacker uses the assertion generated for one RP to obtain access to a second RP.
- *Assertion reuse* - An attacker attempts to use an assertion that has already been used once with the intended RP.

In some cases, the subscriber is issued some secret information so that they can be recognized by the RP. The knowledge of this information distinguishes the subscriber from attackers who wish to impersonate them. In the case of holder-of-key assertions, this secret could already have been established with the IdP prior to the initiation of the assertion protocol. In other cases, the IdP will generate a temporary secret and transmit it to the authenticated subscriber for this purpose. When this secret is used to authenticate to the RP, this temporary secret will be referred to as a secondary authenticator. Secondary authenticators include assertions in the front-channel model, session keys in Kerberos, assertion references in the back-channel model, and cookies used for authentication. The threats to the secondary authenticator are as follows:

- *Secondary authenticator manufacture* - An attacker may attempt to generate a valid secondary authenticator and use it to impersonate a subscriber.
- *Secondary authenticator capture* - An attacker may use a session hijacking attack to capture the secondary authenticator when the IdP transmits it to the subscriber after the primary authentication step, or the attacker may use a man-in-the-middle attack to obtain the secondary authenticator as it is being used by the subscriber to authenticate to the RP. If, as in the back-channel model, the RP needs to send the secondary authenticator back to the IdP in order to check its validity or obtain the corresponding assertion data, an attacker may similarly subvert the communication protocol between the IdP and the RP to capture a secondary authenticator. In any of the above scenarios, the secondary authenticator can be used to impersonate the subscriber.

Finally, in order for the subscriber's authentication to the RP to be useful, the binding between the secret used to authenticate to the RP and the assertion data referring to the subscriber needs to be strong.

- *Assertion substitution* - A subscriber may attempt to impersonate a more privileged subscriber by subverting the communication channel between the IdP and RP, for example by reordering the messages, to convince the RP that their secondary authenticator corresponds to assertion data sent on behalf of the more privileged subscriber.

8.1. Threat Mitigation Strategies

Mitigation techniques are described below for each of the threats described in the last subsection.

- *Assertion manufacture/modification* - To mitigate this threat, the following mechanisms are used:
    1. The assertion is digitally signed by the IdP. The RP checks the digital signature to verify that it was issued by a legitimate IdP.

2. The assertion is sent over a protected session such as TLS. In order to protect the integrity of assertions from malicious attack, the IdP is authenticated.
3. The assertion contains a non-guessable random identifier.

- *Assertion disclosure* - To mitigate this threat, one of the following mechanisms are used:
  1. The assertion is sent over a protected session to an authenticated RP. Note that, in order to protect assertions against both disclosure and manufacture/modification using a protected session, both the RP and the IdP need to be validated.
  2. Assertions are signed by the IdP and encrypted for a specific RP. It should be noted that this provides all the same guarantees as a mutually authenticated protected session, and may therefore be considered equivalent. The general requirement for protecting against both assertion disclosure and assertion manufacture or modification may therefore be described as a mutually authenticated protected session or equivalent between the IdP and the RP.
- *Assertion repudiation by the IdP* - To mitigate this threat, the assertion is digitally signed by the IdP using a key that supports non-repudiation. The RP checks the digital signature to verify that it was issued by a legitimate IdP.
- *Assertion repudiation by the subscriber* - To mitigate this threat, the IdP issues holder-of-key assertions, rather than bearer assertions. The subscriber can then prove possession of the asserted key to the RP. If the asserted key matches the subscriber's presented key, it will be proof to all parties involved that it was the subscriber who authenticated to the RP rather than a compromised IdP impersonating the subscriber.
- *Assertion redirect* - To mitigate this threat, the assertion includes the identity of the RP for which it was generated. The RP verifies that incoming assertions include its identity as the recipient of the assertion.
- *Assertion reuse* - To mitigate this threat, the following mechanisms are used:
  1. The assertion includes a timestamp and has a short lifetime of validity. The RP checks the timestamp and lifetime values to ensure that the assertion is currently valid.
  2. The RP keeps track of assertions that were consumed within a configurable time window to ensure that an assertion is not used more than once within that time window.
- *Secondary authenticator manufacture* - To mitigate this threat, one of the following mechanisms is used:
  1. The secondary authenticator may contain sufficient entropy that an attacker without direct access to the IdP's random number generator cannot guess the value of a valid secondary authenticator.
  2. The secondary authenticator may contain timely assertion data that is signed by the IdP or integrity protected using a key shared between the IdP and the RP.
- *Secondary authenticator capture* - To mitigate this threat, adequate protections are in place throughout the lifetime of any secondary authenticators used in the assertion protocol.
  1. In order to protect the secondary authenticator while it is in transit between the IdP and the subscriber, the secondary authenticator is sent via a protected session established during the primary authentication of the subscriber.
  2. In order to protect the secondary authenticator from capture as it is submitted to the RP, the secondary authenticator is used in an authentication protocol which protects against eavesdropping and man-in-the-middle attacks.
  3. In order to protect the secondary authenticator after it has been used, it is never transmitted over an unprotected session or to an unauthenticated party while it is still valid.
- *Assertion substitution* - To mitigate this threat, one of the following mechanisms is used:
  1. Responses to assertion requests contain the value of the assertion reference used in the request or some other nonce that was cryptographically bound to the request by the RP.
  2. Responses to assertion requests are bound to the corresponding requests by message order, as in HTTP, provided that assertions and requests are protected by a protocol such as TLS that can detect and disallow malicious reordering of packets.

9. Privacy Considerations

This section is informative.

9.1. Minimizing Tracking and Profiling

Federation offers numerous benefits to RPs and subscribers, but requires subscribers to have trust in the IdP. Accordingly, to build subscriber trust in a federated model, it is important that uses of subscriber data are appropriately limited and scoped to the purpose for which it was originally collected. Consult your SAOP if there are questions about whether proposed agency uses fall within the scope of these uses. Sections 4, 4.1.4, and 5.2.5 cover a number of technical requirements the objective for which is to minimize privacy risks arising from increased capabilities to track and profile subscribers. For example, a subscriber using the same IdP to authenticate to multiple RPs allows the IdP to build a profile of subscriber transactions that would not have existed absent federation. The availability of such data makes it vulnerable to uses that may not be anticipated or desired by the subscriber and may inhibit subscriber adoption of federated services.

Section 4.2 also encourages the use of technical measures to provide unlinkability and prevent subscriber activity tracking and profiling. While IdP policies and procedures are important in ensuring adherence to appropriate use limitation and purpose specification principles, technical measures such as outlined in Section 4.1.4 for proxied federation and Section 5.2.5 for pairwise pseudonymous identifiers, can increase the effectiveness of these policies by making it more difficult to track or profile subscribers beyond operational requirements.

9.2. Notice and Consent

To build subscriber trust in federation, transparency must be provided to the subscriber to understand what information will be transmitted, what is required versus optional, and the ability to decide whether to transmit optional attributes to the RP. Accordingly, Section 6 requires that positive confirmation be obtained from the subscriber before any attributes about the subscriber are transmitted to any RP. An effective notice will take into account user experience design standards and research, as well as an assessment of privacy risks that may arise from the collection. There are many factors that should be considered, including the reliability of the assumptions subscribers may have about the collection, whether other information is being collected and appended to the information collected from the subscriber. However, an effective notice is never only a link that leads to a complex, legalistic privacy policy or general terms and conditions that a substantial number of subscribers do not read or understand.

Section 6 does not specify which party should provide the notice. In some cases, a party in a federation may not have a direct connection to the subscriber to provide notice and obtain consent. Although multiple parties may elect to provide notice, it is permissible for parties to determine in advance either contractually or through trust framework policies which party will provide the notice and obtain confirmation, as long as the determination is being based upon factors that center on enabling the subscriber to pay attention to the notice and make an informed choice.

9.3. Data Minimization

Although an IdP may collect additional attributes beyond what the RP requires for its use case, only those attributes that were explicitly requested by the RP are to be transmitted by the IdP. In some instances, an RP does not require a full value of an attribute; for example, an RP may need to know whether the subscriber is over 13 years old, but has no need for the full date of birth. To minimize the collection of potentially sensitive PII, the RP may request an attribute claim (e.g., Question: Is the subscriber over 13 years old? Response: Y/N or Pass/Fail). This minimizes the RPs collection of potentially sensitive and unnecessary PII. Accordingly, Section 6.3 requires the RP to, where feasible, request attribute claims rather than full attribute values. To support this RP requirement, IdPs are in turn, required to support attribute claims.

9.4. Agency Specific Privacy Compliance

Section 4.2 identifies agency requirements to consult their SAOP to determine privacy compliance requirements. It is critical to involve your agency's SAOP in the earliest stages of digital authentication system development to assess and mitigate privacy risks and advise the agency on compliance obligations such as whether the federation triggers the Privacy Act of 1974 or the E-Government Act of 2002 requirement to conduct a PIA. For example, if the Agency is serving as a CSP in a federation, it is likely that the Privacy Act requirements will be triggered and require coverage by either a new or existing Privacy Act system of records since credentials would be maintained on behalf of the agency RP. If, however, the agency is an RP and using a 3rd party IdP, digital authentication may not trigger the requirements of the Privacy Act, depending on what data passed from the RP is maintained by the agency as the RP (in such instances the agency may have a broader programmatic SORN that covers such data).

The SAOP can similarly assist the agency in determining whether a PIA is required. These considerations should not be read as a requirement to develop a Privacy Act SORN or PIA for use of a federated credential alone; in many cases it will make the most sense to draft a PIA and SORN that encompasses the entire digital authentication process or include the digital authentication process as part of a larger programmatic PIA that discusses the program or benefit the agency is establishing online access.

Due to the many components of digital authentication, it is important for the SAOP to have an awareness and understanding of each individual component so as to advise appropriately on what compliance requirements apply. Moreover a thorough understanding of the individual components of digital authentication will enable the SAOP to thoroughly assess and mitigate privacy risks either through compliance processes or by other means.

9.5. Blinding in Proxied Federation

While some proxy structures (typically those that exist primarily to simplify integration) may not offer additional subscriber privacy protection, others offer varying levels of privacy to the subscriber through a range of blinding technologies. Privacy policies may dictate appropriate use by the IdP, RP, and the federation proxy, but technical means such as blinding can increase effectiveness of these policies by making the data more difficult to obtain. As the level of blinding increases, the technical and operational implementation complexity may as well.

Even with the use of blinding technologies, a blinded party may still infer protected subscriber information through released attribute data or metadata such as analysis of timestamps, attribute bundle sizes, or attribute signer information. The IdP could consider additional privacy-enhancing approaches to reduce the risk of revealing identifying information of the entities participating in the federation.

The following table illustrates a spectrum of blinding implementations used in proxied federation. This table is intended to be illustrative, and is neither comprehensive nor technology-specific.

Table 9-1. Federation Proxies

Proxy Type	RP knows IdP	IdP knows RP	Proxy can track subscriptions between RP and IdP	Proxy can see attributes of Subscriber
Non-blinding Proxy with Attributes	Yes	Yes	Yes	Yes
Non-blinding Proxy	Yes	Yes	Yes	N/A
Double Blind Proxy with Attributes	No	No	Yes	Yes
Double Blind Proxy	No	No	Yes	N/A
Triple Blind Proxy with or without Attributes	No	No	No	No

10. Usability Considerations

ISO/IEC 9241-11 defines usability as the “extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.” This definition focuses on users, goals, and context of use as key elements necessary for achieving effectiveness, efficiency and satisfaction. A holistic approach considering these key elements is necessary to achieve usability.

From the usability perspective, one of the major potential benefits of federated identity systems is to address the problem of user fatigue associated with managing multiple authenticators. While this has historically been a problem with usernames and passwords, as the market evolves, users having to manage many authenticators—whether physical or digital—presents a usability challenge.

While many other approaches to authentication have been researched extensively and have well-established usability guidelines, federated identity is more nascent and therefore lacks the depth and conclusiveness of research findings. As ongoing usability research matures, usability guidelines for federated identity systems will have stronger supporting data.

As stated in the usability sections in 800-63A and 800-63B, overall user experience is critical to the success of any authentication method. This is especially true for federated identity systems since federation is a less familiar user interaction paradigm for many users. Users’ prior authentication experience may influence their expectations.

The overall user experience with federated identity systems should be as smooth and easy as possible, by following usability standards (such as the ISO 25060 series of standards) and established best practices for user interaction design.

ASSUMPTIONS

In this section, the term “users” means “claimants” or “subscribers.”

Guidelines and considerations are described from the users’ perspective.

Accessibility differs from usability and is out of scope for this document. Section 508 was enacted to eliminate barriers in information technology and require federal agencies to make their electronic and information technology public content accessible to people with disabilities. Refer to Section 508 law and standards for accessibility guidance.

10.1. General Usability Considerations

Federated identity systems should:

- Minimize user burden (e.g., frustration, learning curve)
  - Minimize the number of user actions required.
  - Allow users to quickly and easily select among multiple accounts with a single IdP. For example, approaches such as Account Chooser allow users to select from a list of accounts they have accessed in the recent past, rather than start the federation process by selecting their IdP from a list of potential IdP’s.
- Minimize the use of unfamiliar technical jargon and details (e.g., users do not need to know the terms IdP and RP if the basic concepts are clearly explained).
- Strive for a consistent and integrated user experience across the IdP and RP.
- Help users establish an understanding of identity by providing resources to users such as graphics, illustrations, FAQs, tutorials and examples. Resources should explain how users’ information is treated and how transacting parties (e.g., RPs, IdPs, and brokers) relate to each other.
- Provide clear, honest and meaningful communications to users (e.g., communications should be explicit and easy to understand).
- Provide users online services independent of location and device.
- Make trust relationships explicit to users to facilitate informed trust decisions. Trust relationships are often dynamic and context dependent. For example, users may be more likely to trust some IdPs and RPs with certain attributes or transactions more so than others. For example, users may be more hesitant to use federated identity systems on websites that contain valuable personal information (such as financial or health); and depending on the perceived sensitivity of users’ personal data, users may be less comfortable with social network providers as IdPs since people are often concerned with the broadcasting nature of social networking implementations.
- Follow the usability considerations specified in 800-63A, Section 9 for any user-facing information.
- Clearly communicate how and where to acquire technical assistance. For example, provide users with information such as a link to an online self-service feature, chat sessions or a phone number for help desk support. Avoid redirecting users back and forth among transacting parties (e.g., RPs, IdPs, and brokers) to receive technical assistance.
- Perform integrative and continuous usability evaluations with representative users and realistic tasks in an appropriate context to ensure success of federated identity systems from the users’ perspectives.

10.2. Specific Usability Considerations

This section addresses the usability considerations that have been identified with federated identity systems. This section does not attempt to present exhaustive coverage of all usability factors related to federated identity systems; rather it is focused on the larger, more pervasive themes in the usability literature, primarily users' perspectives on identity, user adoption, trust, and perceptions of federated identity space. In some cases, implementation examples are provided. However, specific solutions are not prescribed. The implementations mentioned are examples to encourage innovative technological approaches to address specific usability needs. See standards for system design and coding, specifications, APIs, and current best practices (such as OpenID and OAuth) for additional examples. Implementations are sensitive to many factors that prevent a one-size-fits-all solution.

10.1.1. User Perspectives on Online Identity

Even when users are familiar with federated identity systems, there are different approaches to federated identity, especially in terms of privacy and the sharing of information, that make it necessary to establish reliable expectations for how users' data are treated. Users and implementers have different concepts of identity. Users think of identity as logging in and gaining access to their own private space. Implementers think of identity in terms of authenticators and assertions, assurance levels, and the necessary set of identity attributes to provide a service. Given this disconnect between users' and implementers' concepts of identity, it is essential to help users form an accurate concept of identity as it applies to federated identity systems. A good model of identity provides users a foundation for understanding the benefits and risks of federated systems and encourage user adoption and trust of these systems.

Many properties of identity have implications for how users manage identities both within and among federations. Just as users manage multiple identities based on context outside of cyberspace, users must learn to manage their identity in a federated environment. Therefore, it must be clear to users how identity and context are used.

- Provide users the requisite context and scope in order to distinguish among different user roles. For example, whether the user is acting on their own behalf, or on behalf of another, such as their employer.
- Provide users unique, meaningful, and descriptive identifiers to distinguish among entities.
- Provide users with information on data ownership and those authorized to make changes. Identities, and the data associated with them, can sometimes be updated and changed by multiple actors. For example, some healthcare data is updated and owned the patient, while some is only updated by a hospital or doctor's practice.
- Provide users with the ability to easily verify, view and update attributes. Identities and user roles are dynamic and not static, thus change over time, e.g., age, health, and financial data.
- Provide users means for updating data even if the associated entity no longer exists.
- Provide users means to delete their identities completely, removing all information about the user, to include transaction history.
- Provide users appropriate anonymity and pseudonymity options and the ability to switch among such identity options as desired.
- Provide means for users to manage each IdP to RP connection, to include complete separation as well as the removal of RP access to one or more attributes.

10.1.2. User Perspectives of Trust and Benefits

Many factors can influence user adoption of federated identity systems. As with any technology, users may value some factors more than others. Users often weigh perceived benefits versus risks before making technology adoption decisions. It is critical that IdPs and RPs provide users with sufficient information to enable them to make informed adoption decisions. The concepts of trust, and tiers of trust, fundamental principles in federated identity systems, can drive user adoption. Finally, a positive user experience may also result in increased demand of federation by users, triggering increased adoption by RPs.

This sub-section is focused primarily on user trust and user perceptions of benefits versus risks.

To encourage user adoption, IdPs and RPs need to establish and build trust with users and provide them with an understanding of the benefits and risks of adoption.

- Allow users to control their information disclosure and provide explicit consent through the appropriate use of notifications (see SP 800-63C, Section 9.2, Notice and Consent). Balancing the content, size, and frequency of notifications is necessary to avoid thoughtless user click-through.
    - For attribute sharing, consider the following:
        - Provide a means for users to verify those attributes and attribute values that will be shared. Follow good security practices (see Section 6).
        - Enable users to consent to a partial list of attributes, rather than an all or nothing approach. Allow users some degree of online access, even if the user does not consent to share all information.

- Allow users to update their consent to the list of attributes shared.
- Minimize unnecessary information presented to users. For example, do not display system generated attributes such as pairwise pseudonymous identifiers, even if they are shared with the RP as part of the authentication response.
- Minimize user steps and navigation. For example, build attribute consent into the protocols—so they're not a feature external to the federated transaction. Examples can be found in standards such as OAuth or OpenID Connect.
- Provide effective and efficient redress methods such that a user can recover from invalid attribute information claimed by the IdP (see Section 6).
- Minimize the number of times a user is required to consent to attribute sharing. Balancing the frequency of consent requests avoids user frustration with multiple requests to share the same attribute.

- Collect information for constrained usage only, and minimize information disclosure (see Section 9.3). Unnecessary and superfluous information collection and disclosure or user tracking without explicit user consent erodes user trust. For example, only request attributes from the user that are relevant for the current transaction, not for all possible transactions a user may or may not access at the RP.
- Clearly and honestly communicate potential benefits and risks of using federated identity to users. Benefits that users value include examples such as time savings, ease of use, reduced number of passwords to manage, and increased convenience.

User concern over risk can negatively influence willingness to adopt federated identity systems. Users may have trust concerns, privacy concerns, security concerns, and single-point-of-failure concerns. For example, users may be fearful of losing access to multiple accounts if a single IdP is unavailable, either temporarily or permanently. Additionally, users may be concerned or confused about learning a new authentication paradigm. User perception of benefits must outweigh risk perception in order to foster adoption of federated identity systems.

10.1.3. User Models and Beliefs

Users' beliefs and perceptions predispose them to expect certain results and behave in certain ways. Such beliefs, perceptions, and predispositions are referred to in the social sciences as mental models. For example, people have a mental model of dining out which guides their behavior and expectations at each establishment, such as fast food restaurants, cafeterias, and more formal restaurants. Thus, it is not necessary to be familiar with every establishment to understand how to interact appropriately at each one.

Assisting users in establishing good and complete mental models of federation allows users to generalize beyond a single specific implementation. If federated identity systems are not designed from the users' perspectives, users may form incorrect or incomplete mental models, hence impacting users' willingness for adoption.

- Clearly explain the working relationship and information flow among the transacting parties (e.g., RPs, IdPs, and brokers) to avoid user misconceptions. Use the actual names of the entities in the explanation rather than using the generic terms IdPs and RPs.
  - Provide prominent visual cues and information so that users understand why seemingly unrelated entities have a working relationship. For example, users may be concerned with mixing online personal activities with government services due to a lack of understanding of the information flow in federated identity systems.
  - Provide prominent visual cues and information to users about redirection when an RP needs to redirect control from their site to an IdP. For example, display RP branding within the IdP user interface to inform users when they are logging in with their IdP for access to the destination RP.
- Provide users with clear and usable ways (e.g., visual assurance) to determine the authenticity of the transacting parties (e.g., RPs, IdPs, and brokers). This will also help to alleviate user concern over leaving one domain for another, especially if the root domain changes (e.g., .gov to .com). For example, display the URL of the IdP so that the user can verify that he or she is not being phished by a malicious site.
- Provide users with clear information, including visual cues, regarding implicit logins and explicit logouts. Depending on the implementation, logging into a RP with an IdP account may authenticate users to both the IdP and RP. Users may not realize that ending their session with the RP will not necessarily end their session with the IdP; users will need to explicitly "logout" of the IdP. Users require clear information to remind them if explicit logouts are required to end their IdP sessions.

11. Assertion Examples

This section is informative.

Three types of assertion technologies will be discussed: SAML assertions, Kerberos tickets, and OpenID Connect tokens.

11.1. Security Assertion Markup Language (SAML)

SAML is an XML-based framework for creating and exchanging authentication and attribute information between trusted entities over the internet. As of this writing, the latest specification for [SAML] is SAML v2.0, issued 15 March 2005.

The building blocks of SAML include:

- The Assertions XML schema which defines the structure of the assertion.
- The SAML Protocols which are used to request assertions and artifacts (the assertion references used in the indirect model described in Section 6.1).
- The Bindings that define the underlying communication protocols (such as HTTP or SOAP) and can be used to transport the SAML assertions.

The three components above define a SAML profile that corresponds to a particular use case such as “Web Browser SSO”.

SAML Assertions are encoded in an XML schema and can carry up to three types of statements:

- *Authentication statements* include information about the assertion issuer, the authenticated subscriber, validity period, and other authentication information. For example, an Authentication Assertion would state the subscriber “John” was authenticated using a password at 10:32pm on 06-06-2004.
- *Attribute statements* contain specific additional characteristics related to the subscriber. For example, subject “John” is associated with attribute “Role” with value “Manager”.
- *Authorization statements* identify the resources the subscriber has permission to access. These resources may include specific devices, files, and information on specific web servers. For example, subject “John” for action “Read” on “Webserver1002” given evidence “Role”.

Authorization statements are beyond the scope of this document and will not be discussed.

11.2. Kerberos Tickets

The Kerberos Network Authentication Service [RFC 4120] was designed to provide strong authentication for client/server applications using symmetric-key cryptography on a local, shared network. Extensions to Kerberos can support the use of public key cryptography for selected steps of the protocol. Kerberos also supports confidentiality and integrity protection of session data between the subscriber and the RP. Even though Kerberos uses assertions, since it is designed for use on shared networks it is not truly a federation protocol.

Kerberos supports authentication of a subscriber over an untrusted, shared local network using one or more IdPs. The subscriber implicitly authenticates to the IdP by demonstrating the ability to decrypt a random session key encrypted for the subscriber by the IdP. (Some Kerberos variants also require the subscriber to explicitly authenticate to the IdP, but this is not universal.) In addition to the encrypted session key, the IdP also generates another encrypted object called a Kerberos ticket. The ticket contains the same session key, the identity of the subscriber to whom the session key was issued, and an expiration time after which the session key is no longer valid. The ticket is confidentiality and integrity protected by a pre-established that is key shared between the IdP and the RP during an explicit setup phase.

To authenticate using the session key, the subscriber sends the ticket to the RP along with encrypted data that proves that the subscriber possesses the session key embedded within the Kerberos ticket. Session keys are either used to generate new tickets, or to encrypt and authenticate communications between the subscriber and the RP.

To begin the process, the subscriber sends an authentication request to the Authentication Server (AS). The AS encrypts a session key for the subscriber using the subscriber’s long term credential. The long term credential may either be a secret key shared between the AS and the subscriber, or in the PKINIT variant of Kerberos, a public key certificate. It should be noted that most variants of Kerberos based on a shared secret key between the subscriber and IdP derive this key from a user generated password. As such, they are vulnerable to offline dictionary attack by a passive eavesdropper.

In addition to delivering the session key to the subscriber, the AS also issues a ticket using a key it shares with the Ticket Granting Server (TGS). This ticket is referred to as a Ticket Granting Ticket (TGT), since the verifier uses the session key in the TGT to issue tickets rather than to explicitly authenticate the verifier. The TGS uses the session key in the TGT to encrypt a new session key for the subscriber and uses a key it shares with the RP

to generate a ticket corresponding to the new session key. The subscriber decrypts the session key and uses the ticket and the new session key together to authenticate to the RP.

11.3. OpenID Connect

OpenID Connect is an internet-scale federated identity and authentication protocol built on top of the OAuth 2.0 authorization framework and the JSON Object Signing and Encryption (JOSE) cryptographic system. As of this writing, the latest specification is version 1.0 with errata, dated November 8, 2014.

OpenID Connect builds on top of the OAuth 2.0 authorization protocol to enable the subscriber to authorize the RP to access the subscriber's identity and authentication information. The RP in both OpenID Connect and OAuth 2.0 is known as the client.

In a successful OpenID Connect transaction, the IdP issues an ID Token, which is a signed assertion in JSON Web Token (JWT) format. The client parses the ID Token to learn about the subscriber and primary authentication event at the IdP. This token contains at minimum the following claims about the subscriber and authentication event:

- `iss` - An HTTPS URL identifying the IdP that issued the assertion.
- `sub` - An IdP-specific subject identifier representing the subscriber.
- `aud` - An IdP-specific audience identifier, equal to the OAuth 2.0 client identifier of the client at the IdP.
- `exp` - The timestamp at which the ID Token expires and after which SHALL NOT be accepted the client.
- `iat` - The timestamp at which the ID Token was issued and before which SHALL NOT be accepted by the client.

In addition to the ID Token, the IdP also issues the client an OAuth 2.0 access token which can be used to access the UserInfo Endpoint at the IdP. This endpoint returns a JSON object representing a set of claims about the subscriber, including but not limited to their name, email address, physical address, phone number, and other profile information. While the information inside the ID Token is reflective of the authentication event, the information in the UserInfo Endpoint is generally more stable and could be more general purpose. Access to different claims from the UserInfo Endpoint is governed by the use of a specially defined set of OAuth scopes, `openid`, `profile`, `email`, `phone`, and `address`. An additional scope, `offline_access`, is used to govern the issuance of refresh tokens, which allow the RP to access the UserInfo Endpoint when the subscriber is not present.

12. References

This section is informative.

[OMB M-03-22] OMB Memorandum M-03-22, *OMB Guidance for Implementing the Privacy Provisions of the E-Government Act of 2002* (September 26, 2003), available at: https://www.whitehouse.gov/omb/memoranda/m03-22.html (https://www.whitehouse.gov/omb/memoranda/m03-22.html).

[RFC 7591] IETF, *OAuth 2.0 Dynamic Client Registration Protocol*, available at: https://tools.ietf.org/html/rfc7591/ (https://tools.ietf.org/html/rfc7591/).

[NISTIR 8112] NIST Internal Report 8112, *Attribute Metadata*, available at: https://pages.nist.gov/NISTIR-8112/NISTIR-8112.html (https://pages.nist.gov/NISTIR-8112/NISTIR-8112.html).

[ISO 9241-11] International Organization for Standardization, *Ergonomic requirements for office work with visual display terminals (VDTs) – Part 11: Guidance on Usability*, ISO 9241-11:1998, ISO: Geneva, Switzerland, 1998.

[Section 508] Section 508 Law and Related Laws and Policies (January 30, 2017), available at: https://www.section508.gov/content/learn/laws-and-policies (https://www.section508.gov/content/learn/laws-and-policies).

[BCP 195] Sheffer, Y., Holz, R., and P. Saint-Andre, *Recommendations for Secure Use of Transport Layer Security (TLS) and Datagram Transport Layer Security (DTLS)*, BCP 195, RFC 7525, May 2015, http://www.rfc-editor.org/info/bcp195 (http://www.rfc-editor.org/info/bcp195).

[RFC 4120] IETF, *The Kerberos Network Authentication Service (V5)*, available at: https://tools.ietf.org/html/rfc4120/ (https://tools.ietf.org/html/rfc4120/).

[SAML] OASIS, *Security Assertion Markup Language (SAML) V2.0 Technical Overview*, available at: http://docs.oasis-open.org/security/saml/Post2.0/sstc-saml-tech-overview-2.0.html (http://docs.oasis-open.org/security/saml/Post2.0/sstc-saml-tech-overview-2.0.html).