# Federation Manager Authentication Risk Assessment



**DRAFT**

This is a DRAFT document that has been posted to stimulate discussion of risk factors associated the InCommon's Federation Manager. Nothing contained here should be considered to be an official statement or final conclusion of the authors.

## Federation Manager Authentication Risk Assessment

The security of the InCommon Federation Manager (FM), the application used by RA administrators, site administrators, and delegated administrators, is critical to the confidentiality, integrity, and availability of InCommon IdPs and SPs. This document discusses threats, impacts, and potential controls, with emphasis on those threats that relate to identity and access management for the FM.

The approach taken here is based on the risk assessment framework outlined in NIST Special Publications 800-30 and 800-53 and other documents referenced in the text. In brief, this process has the following steps:

1. Identify threats to the system and the impacts of those threats.
2. Based on the severity of the impacts and the likelihood of the threats, determine an impact level for each threat.
3. Implement controls that are appropriate to mitigate the identified risks.

It should be noted that "security" in this context is not simply about the threats posed by digital criminals. Security also concerns itself with any threat to the confidentiality, integrity, or availability of a system. The NIST framework is much broader, addressing not only technical security, but also physical security, disaster recovery, organizational maturity, etc. That said, this document concentrates primarily on digital security.

Finally, the goal here is not to identify the best/strongest controls available to address a threat. Rather, the goal is to identify controls that are appropriate to the impact levels that have been identified.

### Metadata Management

The maintenance of InCommon metadata is a shared responsibility between InCommon Operations (represented by the RA administrators) and InCommon participants (represented by the site administrators and delegated administrators). In general, the site administrators and the delegated administrators are responsible for submitting correct metadata, while the RA administrators are the stewards of that metadata. More specifically:

- Site administrators and delegated administrators are generally considered to be the source of metadata content. There is an exception, however:
  - The RA administrator may insert additional information into an entity descriptor. Examples of such information are organization data, Identity Assurance qualifiers, and service provider category attributes for which the Federation operator is authoritative.
- As the steward of metadata, the RA administrator is responsible for:
  - The integrity and availability of participant metadata. (Confidentiality is not an issue.)
  - Controls, to be leveraged by site administrators and delegated administrators, that aid in the submission of correct metadata by authorized individuals. The use of these tools may be required for all administrators, required for certain administrators depending on certified assurance levels, or may be optional. Examples of such controls are authentication mechanisms, secure communication channels, notification of completed transactions, enforced separation of duties, etc.
  - Services such as domain validation that help administrators create and maintain correct metadata.

Note: Despite the controls and services performed by the RA administrator, the responsibility for correct metadata still rests with the participant.

When assessing threats to the Federation Manager, we consider the following issues:

- The impact on an affected IdP's certified assurance level.  The Federation Manager currently holds metadata for unassured, LoA-1 (Bronze), and LoA-2 (Silver) IdPs.
- The impact on an affected entity's (IdP or SP) requirements for confidentiality, integrity, and availability.
- The impact on the reputation of InCommon and its members. This is of particular importance, as the Federation is fundamentally dependent on mutual trust.

The metadata contains different types of information that may be categorized according to the impact of that information being incorrect.  In order to provide guidance for the implementation of appropriate controls, the following table summarizes those categories.

| Metadata Category | Impact of Being Incorrect |
|---|---|
| Certificates in IdP metadata (i.e., a public key in metadata corresponding to the IdP's signing key) | A spoofed IdP may push false identity assertions to SPs that trust the correct IdP. |
| Endpoints (especially `SingleSignOnService` endpoints) in IdP metadata | IdP users may be phished. Also, the IdP's user community may suffer a service outage. |
| Certificates and endpoints in SP metadata | Loss of PII when identity assertions are sent to a spoofed SP. Also, the SP's user community may suffer a service outage. |
| Entity attributes that indicate the trustworthiness of an entity (e.g., Identity Assurance qualifiers and other qualifiers for which the Federation operator is authoritative) | SPs may place too much trust in an IdP's assertion, or an IdP may send information to an SP that it would not normally trust to receive that information. |
| User interface elements in metadata (i.e., MDUI elements) | The entity's user community may be presented with confusing information, with the possibility that users will be misled and/or coerced to do the wrong thing. |

# Potential Threats

The following table provides a summary of threats and their impacts. It also lists potential controls, but the selection of specific controls for implementation is the subject of another document. [InCCollaborate:Need a link to Ops's document...]

| Threat | Potential Impacts | Who Is Impacted | Impact Level | Potential Controls | Questions and Comments |
|---|---|---|---|---|---|
| A bad guy controls a certificate entry in IdP metadata, the bad guy could issue arbitrary signed assertions from a Trojan IdP. | The Trojan IdP could actively push bogus assertions to arbitrary SPs. An SP that would otherwise trust that IdP would immediately be compromised. | The SP's community members. | Equal to the impact level of the IdP's highest assurance profile (Unassured - Low, LoA1 - Low, LoA2 - Medium, LoA3 - Medium, LoA4 - High). | • Two-factor authN<br>• Notifications of completed transactions to IdP administrators<br>• Required confirmation by another IdP administrator of transactions before they are executed. | Today, injecting a certificate into metadata is a two-step operation. First, the certificate is uploaded to the secure server. Second, the certificate is bound to metadata. The first step can and should include compensating controls. |
| A bad guy controls a SingleSignOnService endpoint in IdP metadata. | Users redirected to that endpoint could be phished. | The IdP's community members. | Low increase in impact, given existing mitigation for authentication in each assurance profile. | • Two-factor authN<br>• Notifications of completed transactions to IdP administrators<br>• Required confirmation by another IdP administrator of transactions before they are executed.<br>• Domain validation | Today, all domains in metadata are validated by the RA and remain valid for at least one year. To introduce a bogus domain into metadata, a bad guy would have to control a host in a valid domain or poison the whois database that the RA uses to validate a domain. |
| A bad guy controls SP metadata, using it to create a Trojan service. | IdPs could be coerced to send identity information (perhaps PII) to the Trojan service. | The IdP's community members. | Equal to the impact of an SP improperly using the identity information it's given. | • Two-factor authN<br>• Notifications of completed transactions to SP administrators<br>• Required confirmation by another SP administrator of transactions before they are executed.<br>• A "Secure SP" category for SPs that are trusted to receive PII? | The impact level depends on the specific attributes that are asserted to the SP. |
| A bad guy controls a DiscoveryResponse endpoint in SP metadata. | Users redirected to that endpoint could be phished. | The IdP's community members. | Low increase in impact of phishing, due to the fact that any SP can create a DS that ignores metadata. | • Two-factor authN<br>• Notifications of completed transactions to SP administrators<br>• Required confirmation by another SP administrator of transactions before they are executed.<br>• Domain validation | |
| An authorized guy could (intentionally or unintentionally) insert bad metadata for an entity. | Errors in metadata can affect the assurance, confidentiality, integrity availability, or reputation of the entity. | The entity's community members. | Equal to the impact of "insider" threats for an IdP's highest assurance profile, or to the CIA or reputation of any entity. | • Notifications of completed transactions to IdP/SP administrators<br>• Required confirmation by another IdP/SP administrator of transactions before they are executed. | Given the system and human controls currently in place, it would be very difficult for an authorized user to hurt anyone outside the immediate community of interest. |
| The authentication service required by the FM for a particular administrator may not be available, and a repair may require that administrator to modify metadata. | Users of that authentication service (*e.g.*, IdP) will experience an outage until InCommon personnel can intervene. | The IdP's community members. | Equal to the impact of outages for the IdP's highest assurance profile. | • Provide an alternative authentication method for site administrators, enabling them to resolve these issues within the member's support organization. | |

**Comments about Reputation**

Any of the threats listed in the previous section represent a threat to the reputation of InCommon and its members. As mentioned earlier, the impact of these threats is not merely a matter of public embarrassment. If IdPs and SPs cannot come to trust the federation administration and each other as a whole, then the federation can bring much less to the table to foster the deployment of services that are widely available to the community. Reputation is critical to InCommon. One reputation-damaging incident will not destroy InCommon, but a string of them will.

The level of impact on reputation, unfortunately, cannot be related directly to the assets affected by a security incident; it is largely a matter of public perception. Incidents with low impact on assurance, confidentiality, integrity, or availability can have the same, or even higher, impact on reputation than high-impact incidents. Assuming successful, reputation-building operation, risk to reputation will diminish over time.

# Categorization of the Federation Manager

Given the analysis above, and using the "high water mark" principle referenced in section "3.2 Categorizing the Information System" of NIST Special Publication 800-53, "Recommended Security Controls for Federation Information Systems and Organizations," the Federation Manager is a moderate-impact system.

For reference, NIST Special Publication 800-30, "Risk Management Guide for Information Technology Systems" defines low, moderate (medium), and high impact as follows:

| | Impact Definition |
|---|---|

| | |
|---|---|
| High | Exercise of the vulnerability (1) may result in the highly costly loss of major tangible assets or resources; (2) may significantly violate, harm, or impede an organization's mission, reputation, or interest; or (3) may result in human death or serious injury. |
| Medium | Exercise of the vulnerability (1) may result in the costly loss of tangible assets or resources; (2) may violate, harm, or impede an organization's mission, reputation, or interest; or (3) may result in human injury. |
| Low | Exercise of the vulnerability (1) may result in the loss of some tangible assets or resources or (2) may noticeably affect an organization's mission, reputation, or interest. |

Notes:

- This assessment is based on InCommon's currently-defined Bronze and Silver identity assurance profiles, which have low to moderate impact. Introduction of high-impact assurance profiles (*e.g.*, LoA-4) at some time in the future would require reclassification of the Federation Manager to high-impact.
- The categorization of SPs within InCommon is not known at this time. It may be useful at some time for SPs to be able to declare their categorization, and high-impact SPs should be warned that they may need to implement mitigating controls if they participate in InCommon.
- Not all operations performed by the Federation Manager are of moderate impact. Protection of public keys and federation-controlled metadata are probably the highest impact from the point of view of the federation. Impacts categorization for other metadata is more dependent on the needs of the IdPs and SPs themselves, so the need for associated controls could be self-declared.

## Appendix: Risk Assessment for Assurance

NIST Special Publication 800-30, "Risk Management Guide for Information Technology Systems," provides general guidance for risk assessment. This guide is leveraged by E-Authentication Guidance for Federal Agencies (OMB M-04-04), which describes a risk assessment framework for determining the Level of Assurance an SP should require. The following table summarizes how LoA is determined:

| **Potential Impact Categories for Authentication Errors** | **LoA 1** | **LoA 2** | **LoA 3** | **LoA 4** |
|---|---|---|---|---|
| Inconvenience, distress or damage to standing or reputation | Low | Medium | Medium | High |
| Financial loss or agency liability | Low | Medium | Medium | High |
| Harm to agency programs or public interests | N/A | Low | Medium | High |
| Unauthorized release of sensitive information | N/A | Low | Medium | High |
| Personal safety | N/A | N/A | Low | Medium to High |
| Civil or criminal violations | N/A | Low | Medium | High |

The meanings of "Low," "Medium," and "High" are defined in OMB M-04-04 on pages 6 and 7, and, in the case of unauthorized release of sensitive information, in Standards for Security Categorization of Federal Information and Information Systems (FIPS PUB 199):

Potential impact of inconvenience, distress, or damage to standing or reputation:

- Low: at worst, limited, short-term inconvenience, distress or embarrassment to any party.
- Moderate: at worst, serious short term or limited long-term inconvenience, distress or damage to the standing or reputation of any party.
- High: severe or serious long-term inconvenience, distress or damage to the standing or reputation of any party (ordinarily reserved for situations with particularly severe effects or which affect many individuals).

Potential impact of financial loss:

- Low: at worst, an insignificant or inconsequential unrecoverable financial loss to any party, or at worst, an insignificant or inconsequential agency liability.
- Moderate: at worst, a serious unrecoverable financial loss to any party, or a serious agency liability.
- High: severe or catastrophic unrecoverable financial loss to any party; or severe or catastrophic agency liability.

Potential impact of harm to agency programs or public interests:

- Low: at worst, a limited adverse effect on organizational operations or assets, or public interests. Examples of limited adverse effects are: (i) mission capability degradation to the extent and duration that the organization is able to perform its primary functions with noticeably reduced effectiveness, or (ii) minor damage to organizational assets or public interests.
- Moderate: at worst, a serious adverse effect on organizational operations or assets, or public interests. Examples of serious adverse effects are: (i) significant mission capability degradation to the extent and duration that the organization is able to perform its primary functions with significantly reduced effectiveness; or (ii) significant damage to organizational assets or public interests.
- High: a severe or catastrophic adverse effect on organizational operations or assets, or public interests. Examples of severe or catastrophic effects are: (i) severe mission capability degradation or loss of to the extent and duration that the organization is unable to perform one or more of its primary functions; or (ii) major damage to organizational assets or public interests.

Potential impact of unauthorized release of sensitive information:

- Low: at worst, a limited release of personal, U.S. government sensitive, or commercially sensitive information to unauthorized parties resulting in a loss of confidentiality with a low impact as defined in FIPS PUB 199. ("The unauthorized disclosure of information could be expected to have a **limited** adverse effect on organizational operations, organizational assets, or individuals.")
- Moderate: at worst, a release of personal, U.S. government sensitive, or commercially sensitive information to unauthorized parties resulting in loss of confidentiality with a moderate impact as defined in FIPS PUB 199. ("The unauthorized disclosure of information could be expected to have a **serious** adverse effect on organizational operations, organizational assets, or individuals.")

- High: a release of personal, U.S. government sensitive, or commercially sensitive information to unauthorized parties resulting in loss of confidentiality with a high impact as defined in FIPS PUB 199. ("The unauthorized disclosure of information could be expected to have a **severe or catastrophic** adverse effect on organizational operations, organizational assets, or individuals.")

Potential impact to personal safety:

- Low: at worst, minor injury not requiring medical treatment.
- Moderate: at worst, moderate risk of minor injury or limited risk of injury requiring medical treatment.
- High: a risk of serious injury or death.

The potential impact of civil or criminal violations is:

- Low: at worst, a risk of civil or criminal violations of a nature that would not ordinarily be subject to enforcement efforts.
- Moderate: at worst, a risk of civil or criminal violations that may be subject to enforcement efforts.
- High: a risk of civil or criminal violations that are of special importance to enforcement programs.

## Appendix: Risk Assessment for Confidentiality, Integrity, and Availability

NIST Special publication 800-53, "Recommended Security Controls for Federation Information Systems and Organizations" provides an extremely detailed description of security controls in the areas of access control, awareness and training, audit and accountability, security assessment and authorization, configuration management, contingency planning, identification and authentication, incident response, maintenance, media protection, physical and environmental protection, planning, personnel safety, risk assessment, system and services acquisition, system and communications protection, system and information integrity, and program management. For the purpose of this document, the controls selected for analysis may be found in the "access control" and "identification and authentication" sections. A complete risk assessment of the Federation Manager, though, should include all areas.