After Action Review: Introduction of `&#13;` characters into metadata 2020-10-09

Date: 2020-10-16
Prepared by: Nic Roy, Internet2 Trust and Identity Services

Summary

InCommon deployed a change to federation metadata on Tuesday, October 6th, to introduce a new feature into metadata. On Friday, October 9th, InCommon operations was notified that there were instances of an XML-encoded carriage return character, encoded as `&#13;` in InCommon metadata, which was negatively affecting the ability of ADFSTookit to verify the signature on any metadata aggregate which contained it. Since some of the InCommon metadata which was affected was exported to eduGAIN, failure to consume metadata by ADFSToolkit deployments in multiple federations resulted.

Timeline

`2020-10-06 16:14 MT` InCommon publishes updated metadata containing a new ‘mailto:’ scheme prefixed to existing contacts. This was a planned change, but it also resulted in the publication of a number of Service Provider (SP) and Identity Provider (IdP) entity descriptors with the XML-encoded carriage return `&#13;` in metadata. This metadata was schema-valid and conformed with SAML metadata standards.

`2020-10-09 09:31 MT` Federation operators in other eduGAIN member federations begin reporting failure of ADFSToolkit to consume metadata, and these failures are correlated with the InCommon metadata containing `&#13;` characters.

`2020-10-09 10:54 MT` Nic Roy sees the email from international federation operators and begins to evaluate the situation in InCommon.

`2020-10-09 11:10 MT` Nic Roy notifies InCommon service manager and leadership about the issue and notifies them of plans to perform an emergency fix involving modification of participant metadata, re-signing and publishing of metadata, and notification of affected site administrators.

`2020-10-09 11:50 MT` Nic Roy in collaboration with the InCommon operations team identifies the following entity descriptors containing the offending character:

https://analytics.test.uchealth.edu/sp
https://ventiv.ucop.edu
https://www.coral.washington.edu/
https://www-test.coral.washington.edu/
https://testshib.msacademicverify.com/shibboleth-sp
https://reta.med.umich.edu/shibboleth
https://wiki.osris.org
https://esyllabus.pharmacy.uiowa.edu/saml
https://redcap.uncg.edu/shibboleth
https://dl.acm.org/shibboleth
https://wwu.sclintra.com/AuthServices-1
https://uwisc.hosted.ethosce.com
https://uwisc.hosted.test.cloud.ethosce.com/sites/all/libraries/simplesaml/www/module.php/saml/sp/metadata.php/default-sp
https://www.peoplegrove.com/saml
https://demo.portal.iontuition.com
https://fortlewis.photoshelter.com/sso/SAML2

`2020-10-09 12:29 MT` Nic Roy finishes the fix for metadata exported to eduGAIN, and notifies international federation operators.

2020-10-09 13:00 MT Nic Roy fixes metadata for non-exported entity descriptors and begins to notify all InCommon site administrators whose metadata was fixed.

2020-10-10 17:39 MT Confirmation is received from the reporting federation that their ADFSTookit issues have been resolved.

2020-10-13 07:28 MT American Museum of Natural History informs InCommon Operations that their ADFSTookit deployment still won't consume InCommon metadata. Another instance of the issue is discovered in some non-exported metadata:

https://analytics.uchealth.edu/sp

2020-10-13 08:25 MT Metadata is fixed for this entity descriptor, and it's confirmed that no further entity descriptors exist with this character in their metadata. Nic contacts the AMNH site admins to let them know their issue is resolved, and contacts UCOP site admins to let them know another entity descriptor of theirs was modified with the fix.

Scope/impact of the outage

ADFS instances using ADFSToolkit, in global federations which consumed eduGAIN-sourced InCommon metadata, were unable to refresh metadata between approximately 2020-10-06 16:14 MT and 2020-10-09 12:29 MT. ADFS instances in InCommon using ADFSToolkit were further unable to refresh metadata until 2020-10-13 08:25 MT. Since these instances were still able to use existing metadata from before this period of time, it's likely that there was a small amount of impact to users of these IdPs, and there are work-arounds which could be used to load new metadata during this time.

Results of the investigation

The InCommon operations team and InCommon software engineers investigated the root cause of this issue, and developed the following timeline:

1) Over the course of the lifetime of the InCommon federation, a number of carriage return characters were introduced into participant metadata, likely through copying and pasting of descriptions of SPs and IdPs into the mdui:Description and ServiceDescription fields, from Microsoft Word documents.
2) In May, 2019, the InCommon federation changed its metadata rendering engine to use a better XML library. One of the differences is that it correctly converts carriage returns to `&#13;`

3) InCommon participants running ADFSToolkit started to notice issues with this character in metadata and reported it to InCommon. InCommon developed a fix to prevent this character from being submitted in new metadata.
4) A number of carriage returns persisted in metadata, but since they hadn't been passed through the new renderer since InCommon introduced it (no one had changed the offending metadata since before the release of the new renderer on May 29, 2019) they weren't converted to `&#13;`, and since they are whitespace, operations did not see them when inspecting metadata.
5) On 2020-10-06, operations resubmitted every piece of metadata that was not in the process of being edited by an InCommon Site Administrator (so almost all our metadata) in order to add the 'mailto:' scheme to contacts. This caused the metadata to pass through the new renderer, but not through the front-end input cleanup, thus converting the old CRs into `&#13;`. While operations performed quality assurance on this change, the volume of change was so high that seeing the few instances of `&#13;` in the diff was difficult, and they were not noticed. Since the metadata was schema valid, standards-conformant and parseable by software such as Shibboleth, we promoted it into production.

Since the initial cleanup, software engineers checked the federation manager database for any further carriage returns or anything that starts with `&#` and ends with `;` and found a couple in some inactive organizations with not-currently-published metadata, which have now also had a fix applied retroactively.

Recommendations

1) Search for and clean up all remaining instances of this issue in the federation manager database - this is done.
2) Introduce additional input validation which will run all metadata submissions through InCommon's eduGAIN import rules, which check for this and a number of other issues, and reject metadata which does not pass this additional validation. This recommendation will have to wait until after full implementation of Baseline Expectations version 2, for a number of technical reasons related to non-SSL endpoints in InCommon metadata which will be cleaned up during Baseline Expectations version 2.
3) Microsoft should fix its XML canonicalization implementation to correctly handle the presence of XML-encoded whitespace such as `&#13;`.