Purpose of this white paper:

- Describe the problem; set the stage for framing the solution (when we know what it is)
- Clarify terminology

Target audiences:

- IdP Operators / Campus IT
    - they need to be aware of the implications of the choices they make in their metadata
    - they need to know that they might not be alone wanting to use their institution's name in WAYF services
- Librarian IT
    - they need to be aware of the implications of the choices they make in their metadata (using the Metadata Explorer Tool (met.refeds.org) may help the librarian see what their institution's metadata looks like to others)
    - they need to be in close conversation with Campus IT
    - they need to know how to handle user feedback when there is confusion
- Federation Operators
    - need to consider how they're handling metadata from the same institution
- SP Operators
    - they are the ones sending their users to WAYF
    - and when they are building their own WAYF, they need to be aware
    - they need to know how to handle user feedback when there is confusion
- Vendors of IdP Software
    - encourage vendors to be consistent with what they include in metadata such as including DisplayName
    - support communication with academic identity federations

Challenges in Federated Where-Are-You-From (WAYF) Services

Work Product of the SeamlessAccess WAYF Entry Disambiguation Working Group

Abstract

There are many ways to manage access to remote resources. From the use of IP addresses via VPN or proxy services to federated identity, the goal of enabling access to remote content is both easier and more complicated than ever before. This article focuses on the user experience with federated identity at the point where users have to select the institution they want to use to authorize their access. In particular, we look into the challenges that exist in the Where-Are-You-From (WAYF) identity provider discovery process when a user is presented with nearly identical choices for the same institution.

Introduction

What is WAYF and why is it important? A user wants to access a resource and needs to tell it which institution they are from. The user finds themselves on a WAYF discovery page where they must find their institution. When the institutional name is clear, they have no problem selecting the right one from a multitude of other names. When they find a number of similar names, they are not sure which one to select, they are likely to select the one that is not their institution and they will not get access to the resource. The resources get the names they display from federation metadata.

Federated identity management permits extending Single Sign-On (SSO) above the enterprise level, creating a trusted authority for digital identities across multiple organizations. In a federated system, participating institutions share identity attributes based on agreed-upon standards, facilitating authentication from other members of the federation and granting appropriate access to online resources. This approach streamlines access to digital assets while protecting restricted resources.

Terminology

One of the more immediate challenges of federated identity is understanding the terminology commonly used to describe the different components that make up the federated identity ecosystem. In this paper, we use the following terms:

Identity Provider - A member or a customer of an identity federation that assigns access credentials to qualifying individual natural persons. Typically, Identity

Providers are universities, schools, and research institutes. This is commonly abbreviated as “IdP.”

- **Display Name** (identifying the IdP for the user) - The user-friendly name registered in federation metadata with an Entity ID.
- **Entity ID** (identifying the IdP for the system) - The name the IdP includes in their metadata to identify itself to a federated SSO service.

Service Provider - A provider of resources such as information resources, scholarly collaboration tools, or shared research infrastructure, by means of an online platform that requires end users accessing those resources to be reliably authenticated, and that is a member or a customer in good standing of an identity federation and is as such part of the trust framework provided by the identity federation.

- ***Where Are You From (WAYF) service*** (aka, IdP Discovery Service) (helps the user to select IdP to login to this Service Provider) - A WAYF service allows users to identify the Identity Provider (IdP) they will use to authenticate. To minimize the number of times a user must use a WAYF service, a federated authentication system may store the selected IdP in the user’s browser and use it as a default for subsequent WAYF prompts, including another provider’s site, or encode it into a WAYFless URL.

Federation (aka, Identity Federation) - An organization that operates a trust framework for the exchange of authentication assertions between its members and/or customers, that are Identity Providers and Service Providers. (Examples: InCommon, OpenAthens, Australian Access Federation)

eduGAIN - An interconnection of identity federations around the world, simplifying access to content, services, and resources for the global research and education community. eduGAIN enables the trustworthy exchange of information related to identity, authentication, and authorization.

IdP Proxy - An IdP Proxy is a bridge or gateway between a federation of SAML IdPs and a federation of SAML SPs. To an SP, an IdP Proxy looks like an ordinary IdP. Likewise, to an IdP, an IdP Proxy looks like an SP. Thus an IdP Proxy has the combined capability of both an IdP and SP. A proxy is often used by research collaboration infrastructures.

Problem Space / Use Cases

During the IdP Discovery journey, an end user may encounter what appear to be identical IdPs in a WAYF service, but in actuality are not. There are a variety of reasons this might occur, reasons that are technically correct with understandable business decisions behind them.

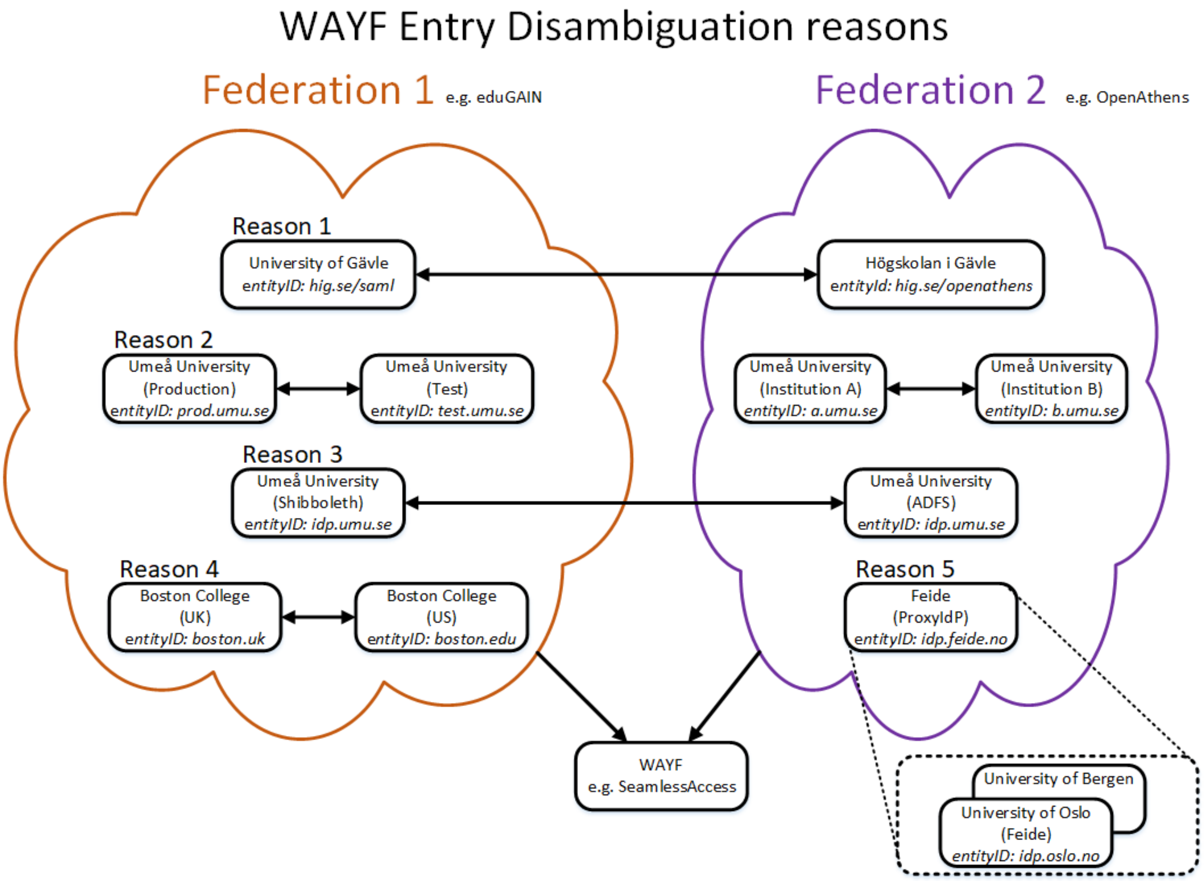
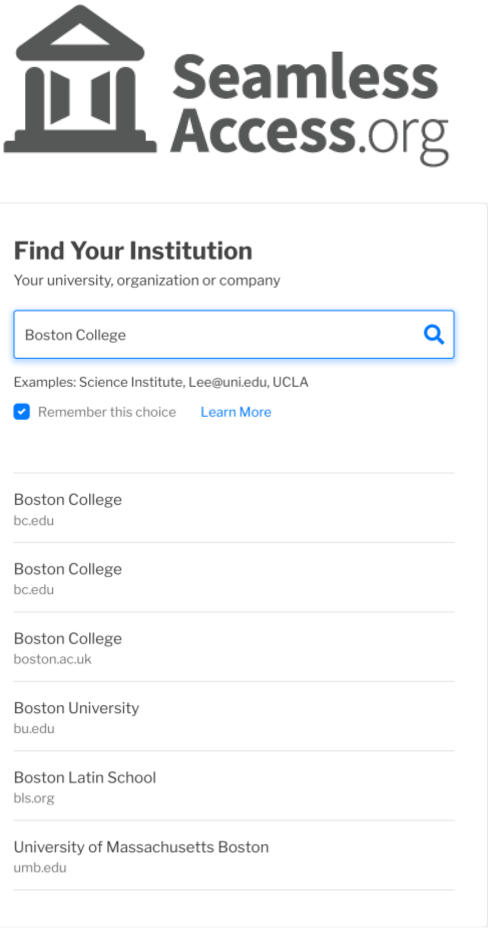


Figure 1: WAYF Entry Conflicts

This may occur for a variety of reasons

Reason 1 - Similar Display Names, Multiple Federations

Example: The library has an IdP service registered in OpenAthens with a similar display name as an IdP service from the campus that is registered in InCommon; metadata from both federations is included in the SeamlessAccess metadata aggregate. It is impossible for the user to distinguish between the Library/OpenAthens and the Campus/InCommon IdP in the discovery service.



Seamless Access.org

Find Your Institution

Your university, organization or company

Boston College

Examples: Science Institute, Lee@uni.edu, UCLA

☑ Remember this choice   Learn More

Boston College
bc.edu

Boston College
bc.edu

Boston College
boston.ac.uk

Boston University
bu.edu

Boston Latin School
bls.org

University of Massachusetts Boston
umb.edu

Figure 2: WAYF results example, similar Display Names across multiple federations

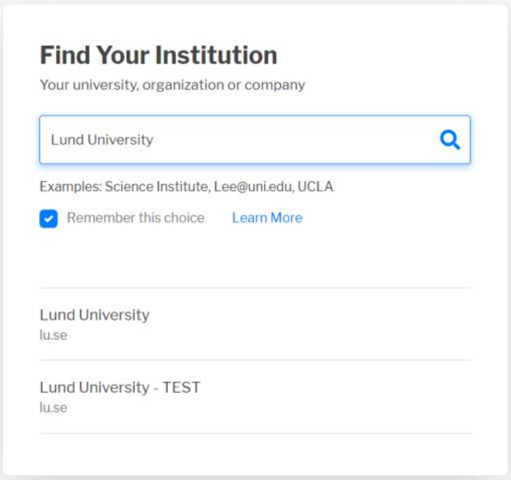
One of the more common use cases involves a scenario where a department within an organization signs up to one federation, while the organization itself signs up for another federation. This results in a situation where an institution has two separate IdPs registered with similar Display Names but different Entity IDs in two separate federations. When an SP or WAYF service aggregates the metadata from multiple federations, then both IdPs will be shown in that aggregated WAYF service.

In some cases, the similarity between the display names is only obvious to a human. This is particularly the case when the pattern being matched is a name in English and its direct translation in another language.

While one obvious response is to disallow the aggregation of metadata, this is not always possible. SPs are often members of more than one federation, and so will want to display all the IdPs possible from any federation the SPs have joined. For services like SeamlessAccess, the more IdPs are represented in the metadata, the more choices are available to the user. Users are confused when they cannot find their organization in a WAYF list.

Reason 2 - Similar Display Names, Same Federation

Example: A campus has a test IdP and a production IdP with the same entityID (in the same federation). Alternatively, two departments on a single campus have registered separate IdPs with a federation with similar Display Names.



Find Your Institution

Your university, organization or company

Lund University

Examples: Science Institute, Lee@uni.edu, UCLA

☑ Remember this choice   Learn More

Lund University
lu.se

Lund University - TEST
lu.se

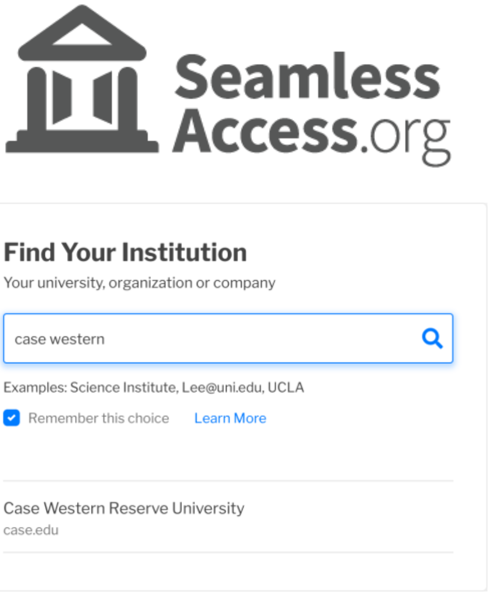
Figure 3: WAYF results example, similar Display Names across a single federation

The apparent name conflict does not necessarily have to cross federation boundaries. Even within a federation, duplications may appear if the display names are similar but the entity ID is different. This might happen if two departments register their own IdPs with names like “University of the Very Long Name - Physics” and “University of the Very Long Name - Library.” While the example above is clear, the length of the name may result in the last part being cut off, leaving the user without enough information to distinguish between the two options.

A similar case can happen when an organization registers a test IdP as well as a production IdP. If the display names are not obviously dissimilar, then the user may not be able to recognize which system they should be using when authenticating.

Reason 3 - Same EntityID, Multiple Federations

Example: a university has a Shibboleth IdP in InCommon and an OpenAthens IdP with the same entityID (in different federations). This is often a part of a planned migration from one IdP to another, but sometimes the institution does not update the metadata to remove the old IdP. Which IdP (old or new) is presented to the user will depend on what the SP has in their metadata; the user will be unable to determine or control if that is the current IdP for their institution. There are no visible distinctions when presenting the IdP to the user, and the SP may not be able to control which IdP is presented to the user. In the example below, the user will see only one IdP, but it is the older one and will not work as intended.



Find Your Institution

Your university, organization or company

case western

Examples: Science Institute, Lee@uni.edu, UCLA

☑ Remember this choice Learn More

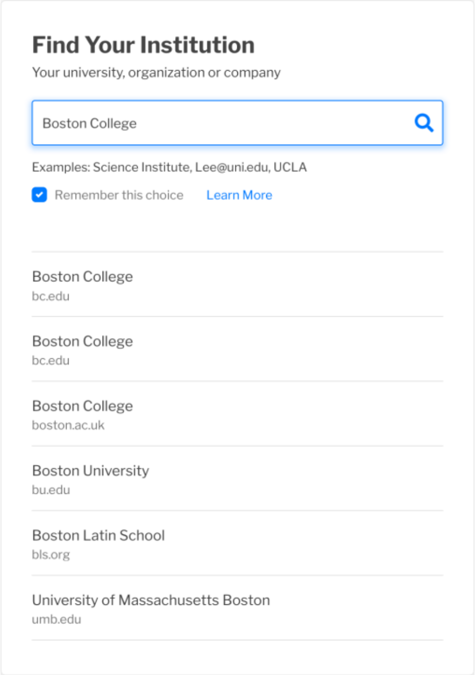
Case Western Reserve University
case.edu

Figure 4: WAYF entry example, same EntityID across multiple federations

In previous scenarios, the entity ID was the technically distinguishable characteristic in all examples. There are situations, however, where the entity ID remains the same even though the backend system is different. This may happen when an entity ID registered in one federation points to a campus IdP, but in another federation, that same entity ID points to something entirely different. Unfortunately, which IdP is listed in a WAYF will depend on the SP, which IdP they have in their metadata, and what order it happens to be displayed to the user. This is often outside of the SP’s control. The user will also have an inconsistent login experience because they may see the “wrong” IdP for that particular SP and not gain access to the content they are entitled to.

Reason 4 - Similar Display Names, Unrelated Institutions

Example: Boston College in the US, Boston College in the UK.



Find Your Institution

Your university, organization or company

Boston College

Examples: Science Institute, Lee@uni.edu, UCLA

☑ Remember this choice Learn More

Boston College
bc.edu

Boston College
bc.edu

Boston College
boston.ac.uk

Boston University
bu.edu

Boston Latin School
bls.org

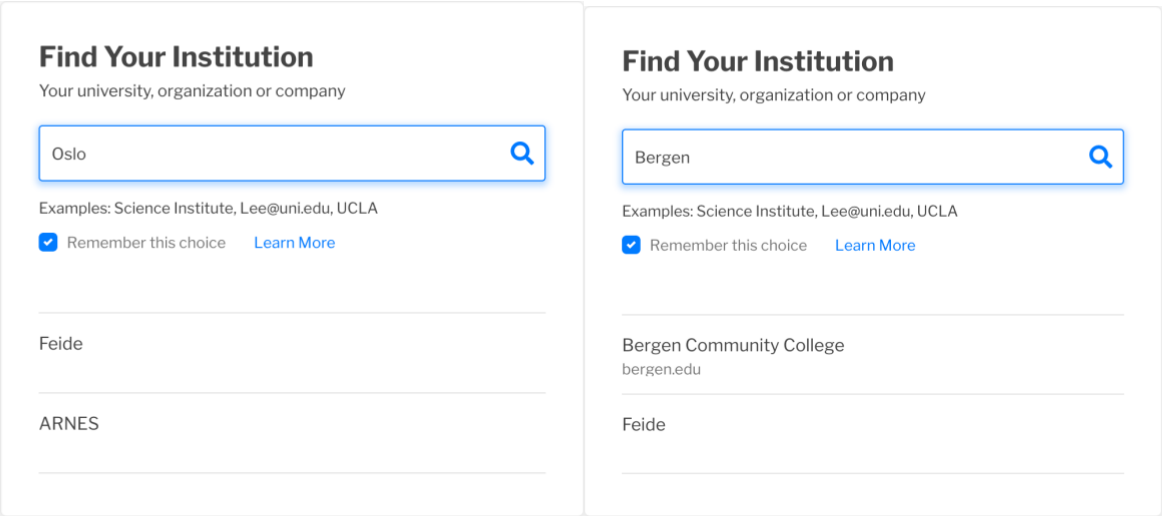
University of Massachusetts Boston
umb.edu

Figure 5: WAYF entry example, similar display names for unrelated institutions

The challenge of visually similar names in a WAYF is not restricted to something that happens within one organization. Two unrelated institutions with similar Display Names may be registered in one or more federations. While not as common as the scenarios involving confusion within a single organization, this is definitely something that occurs today.

Reason 5 - Shared IdP

Example: Feide, the Norwegian identity federation, uses interstitial landing pages to direct the user from the single IdP in eduGAIN to the IdPs at the federation’s member institutions. Searching for the user’s IdP may not result in their institution appearing in the list of options; the only option that will appear is Feide.



Find Your Institution

Your university, organization or company

Oslo

Examples: Science Institute, Lee@uni.edu, UCLA

☑ Remember this choice Learn More

Feide

ARNES

Find Your Institution

Your university, organization or company

Bergen

Examples: Science Institute, Lee@uni.edu, UCLA

☑ Remember this choice Learn More

Bergen Community College
bergen.edu

Feide

Figure 6: WAYF entry example, shared IdP for multiple institutions

In some scenarios, organizations will share a common IdP. This may happen as a result of the complexity of the organization (e.g., a large hospital system such as the National Health Service [NHS] in the United Kingdom). Alternatively, if a national federation is organized to offer a single IdP to all its members, as is the case in Norway, this may display the same characteristic of a common name shared by all participating institutions.

In the screenshots above, “Oslo” and “Bergen” are both part of the name of universities in Norway. The Norwegian federation is called “Feide.” “Bergen Community College” is a community college in the state of New Jersey, US.

Wrap Up

There is no clear best practice for preventing visual name collisions within a WAYF service. These collisions cannot be easily identified by existing technology; human perception plays an enormous part in how users understand these Display Names in a WAYF service. The first step, then, is for the different entities, from IdPs to SPs to federation operators, to be aware that this problem exists. From there, they can make a decision that best meets the needs of their particular use case.

The SeamlessAccess WAYF Entry Disambiguation Working Group is focused on raising awareness of the issues and coming up with short, medium, and long-term solutions that will offer the needed best practice guidance for these issues.