Prepared by:
Rhys Smith (smith@cardiff.ac.uk)
Information Services, Cardiff University
Executive Summary

- A wide variety of access methods exist to establish a link between a user wishing to access protected content of a publisher and an existing subscription/license agreement that covers that usage. IP Authentication is a heavily used access method despite its major failings since it represents an extremely easy to use method from the end user perspective.

- One of those up-and-coming access methods is Federated Access. This method has many advantages to user, institution, and publisher, but has one major issue from the user perspective: the discovery problem.

- This issue means that many publishers are reluctant to adopt federated access since it presents a method of access that is far more complex than IP Authentication – and are therefore worried that the increased barrier to access of their resources may result in a loss of custom. Many publishers indicated an willingness to follow and advice, guidance, or instructions around how to present the discovery process to users if it would improve their experience and therefore mitigate against their worries.

- When it comes to improving the end user experience of discovery; users consistently indicate one major desire - consistency. An end user study showed that they generally do not understand any of the terminology used, and have no wish to – they simply want the terminology and design to be similar across all sites so that they can learn what to look for.

- Thus, a brand is needed for Academic Federated Access with a simple name and a strong logo is needed, to provide the consistency that users have expressed a wish for. To some extent, the actual details of the name and logo are irrelevant, as long as there is only one and it is easily recognisable.

Recommendation 1 – A brand should be created for academic federated access. For this brand to be successful, it needs widespread adoption worldwide. The brand should include a short name and a logo; these need not mean anything but simply provide a familiar point of reference.

- Such a brand can only be successful (i.e. provide the consistency that users desire) if widely adopted by federations worldwide: publishers will only readily following such branding guidelines if this is the case.

- Along with this brand, a set of guidelines as to how best to implement a WAYF service are needed. This WAYF service should employ a dynamic search interface that allows easy and fast choice of identity provider (and possibly removes the need for asking the user which is their relevant federation).

Recommendation 2 – A “style guide” should be created for publishers to follow around implementing discovery using the brand created.

- The JISC need to take these recommendations forward by a process of establishing community consensus (where the community in this case is the collection of federations worldwide) around the brand and design guidelines in order to help its adoption within the community. Brand consultants should also possibly play a part in the process in order that the brand created is strong and effective and therefore enhances the end user experience to the maximum extent.
Table of Contents

Executive Summary........................................................................................................2

1 Introduction................................................................................................................2
  1.1 Intended Audience.................................................................................................5
  1.2 Structure of Report.................................................................................................5

2 Accessing Subscription Resources............................................................................6
  2.1 Access Methods.....................................................................................................6
    2.1.1 IP Authentication............................................................................................7
    2.1.2 URL-rewriting Web Proxy (& IP Authentication)............................................8
    2.1.3 VPN/Proxy Server (& IP Authentication).........................................................9
    2.1.4 Institutional Credentials................................................................................10
    2.1.5 Service Credentials.......................................................................................11
    2.1.6 Centralised Access Management....................................................................12
    2.1.7 Federated Access Management......................................................................13
  2.2 Finding Content.......................................................................................................14
  2.3 Conclusions............................................................................................................14

3 The Discovery Problem...............................................................................................15
  3.1 Federated Access and Discovery............................................................................15
  3.2 Discovery Solutions................................................................................................16
    3.2.1 Avoiding Discovery..........................................................................................16
    3.2.2 Client-less Discovery........................................................................................17
    3.2.3 Client-mediated Discovery................................................................................19
  3.3 Conclusions............................................................................................................19

4 Existing Discovery Solutions......................................................................................20
  4.1 Methodology...........................................................................................................20
  4.2 Federated Login Steps............................................................................................20
    4.3 Results................................................................................................................21
      4.3.1 Login Steps.....................................................................................................21
      4.3.2 Finding the Login Area...................................................................................22
      4.3.3 Access Method...............................................................................................23
      4.3.4 Choosing Federated Login..............................................................................24
      4.3.5 Choosing the Federation................................................................................25
      4.3.6 Choosing the Institution................................................................................26
    4.4 Conclusions...........................................................................................................28

5 End-user Evaluation of Existing Discovery Solutions...............................................29
  5.1 Methodology...........................................................................................................29
  5.2 Overall Results.......................................................................................................30
  5.3 Participant Analysis.................................................................................................31
  5.4 Detailed Results.....................................................................................................31
    5.4.1 Login Steps.....................................................................................................36
    5.4.2 Login Method..................................................................................................37
    5.4.3 Choosing Federated Access Type.....................................................................38
    5.4.4 Choosing Correct Federation(DS/WAYF of WAYFs).......................................39
    5.4.5 Choosing Correct Institution (DS/WAYF Service)..........................................40

6 Provider Views on Discovery......................................................................................45
  6.1 Access to resources.................................................................................................45
  6.2 Adopting Federated Access...................................................................................45
  6.3 Discovery................................................................................................................46

7 Conclusions and Recommendations..........................................................................47
  7.1 Recommendations..................................................................................................47
Table of Figures

Figure 1: Example of Avoiding Discovery (Cardiff University Database Portal) ........................................... 16
Figure 2: Example of In-built WAYF/DS (www.sciencedirect.com) ................................................................. 17
Figure 3: Example of Centrally Provided WAYF/DS (UK federation) ............................................................... 18
Figure 4: Number of Login Steps .................................................................................................................. 21
Figure 5: Method used to start the login process .......................................................................................... 22
Figure 6: Terminology used to start the login process ............................................................................... 22
Figure 7: Access methods available ............................................................................................................. 23
Figure 8: Terminology around Choosing Federated Login ........................................................................ 24
Figure 9: Terminology around choosing the federation ............................................................................. 25
Figure 10: Technology used to choose federation ...................................................................................... 25
Figure 11: Discovery Service choice ........................................................................................................... 26
Figure 12: Terminology used on in-built WAYF .......................................................................................... 27
Figure 13: Technology used on in-built WAYF ......................................................................................... 28
Figure 14: Remembering selection on in-built WAYF .............................................................................. 28
Figure 15: Options available to remember selection on in-built WAYF .................................................... 28
Figure 16: End User Study - Participants by category ............................................................................... 32
Figure 17: End User Study - Participants by gender .................................................................................... 33
Figure 18: End User Study - Participants by length in sector ..................................................................... 33
Figure 19: End User Study - participants by e-resource usage ................................................................. 34
Figure 20: End User Study - On/off campus e-resources usage ............................................................... 34
Figure 21: End User Study - Existing awareness of logos ........................................................................... 35
Figure 22: End User Study - Existing awareness of terminology ............................................................... 35
Figure 23: Average Rating vs Number of Login Steps .............................................................................. 36
Figure 24: Finding Login - Rating of clarity of terminology .................................................................... 37
Figure 25: Finding Login - Rating of ease of use of method ................................................................. 37
Figure 26: Choosing Federated Login - Rating of clarity of terminology ................................................. 38
Figure 27: Choosing Correct Federation - rating of clarity of terminology .............................................. 39
Figure 28: Choosing Correct Federation - Rating of ease of use of method ........................................... 39
Figure 29: Choosing Correct Institution - In-built vs Central WAYF ....................................................... 40
Figure 30: Choosing Correct Institution - Rating of clarity of terminology .............................................. 41
Figure 31: Choosing Correct Institution - Rating of ease of use of method (distribution) ...................... 42
Figure 32: Choosing Correct Institution - Rating of ease of use of method (average) ............................ 42
1 Introduction

In early 2009, JISC Collections commissioned a study to explore the approaches taken by providers of electronic resources to the implementation of federated access management, and its effect on the end user experience. The study reviews end user scenarios for accessing protected resources and the general problem of "Discovery" inherent when using Federated access, examining current methods and terminology in place that attempt to solve the problem. End user testing and input from various stakeholders informed a process of examining the pros and cons of each, culminating in a series of recommendations and guidance as to how to improve the Discovery process.

1.1 Intended Audience

This report is aimed at various categories of reader:

- Developers of federated access technologies wishing to understand how real end users perceive discovery for federated login
- Staff from providers of electronic resources looking for guidance based on real end user feedback on how to implement discovery for federated login in a way that offers the best user experience
- Staff from academic institutions who manage access to electronic resources looking for information around discovery in federated login
- Staff and students of academic institutions who wish to know more about how access to electronic resources works and why they need to indicate where there are from when using a federated login

1.2 Structure of Report

Firstly, Section 2 looks at the myriad of ways that a person can find content and then prove they belong to an institution with a valid subscription so that they can access that content.

One of these methods is FAM, which has the discovery problem. Section 3 looks at exactly what this discovery problem is, and what solutions exist.

Given a lack of guidance, many service providers have done discovery differently – different terminology, design, etc. Section 4 looks at how much each kind has been implemented on the UK federation.

Section 5 then gets presents the detail of an end user study conducted to gather feedback on various aspects of discovery as currently implemented.

Next, Section 6 presents an overview of thoughts around discovery gathered at a workshop for providers of electronic resources hosted by the project.

Finally, Section 7 presents various recommendations around how it is best to move the area of discovery forward in a way that enhances the end user experience in a manner acceptable to providers of electronic resources.
2 Accessing Subscription Resources

Academic institutions (called identity providers for the remainder of this document) commonly purchase access to many subscription e-resources on behalf of their members (or of groups of their members). In order for these members to gain access to e-resources that they are entitled to access, a method is required that allows the provider of a resource (called a service provider for the remainder of this document) to associate an individual requesting access to an existing valid subscription agreement.

A wide range of methods for a provider to establish this link exists today; this section examines the landscape of available methods, briefly discussing the advantages and disadvantages of each. Given that a user can make use of these methods to gain access to resources, two related areas are the problems of how users find content in the first place and of how institutions can gather statistics of how much each resource is being used. This section also briefly examines these areas.

2.1 Access Methods

To enable a service provider to establish a link between an individual and an existing subscription agreement, an individual needs to be able to somehow “prove” that they are associated with a particular institution who has a subscription agreement: in this context this step is usually called Authentication. Additionally, an Authorisation step may be required: this is a step that provides information about the individual to the service provider, allowing them to make more fine grained decisions about whether that particular individual should be given access.

In all cases, the user will belong to an academic institution, will be using an electronic device (such as a computer, smartphone, etc.), and will be attempting to access a web-based electronic resource.
### 2.1.1 IP Authentication

Every electronic device on the internet that is capable of communicating with an electronic resource will have a unique IP address assigned to it while connected to the internet. Academic institutions typically have a limited set of IP addresses that they “own” and can give out to electronic devices when they are present on the campus network.

IP Authentication essentially performs the authentication step by assuming that if a user is coming from an IP known to be “owned” by a particular institution, then that user is a member of that institution. There is no authorisation step possible when using this access method.

The major advantage of using IP Authentication is that it is conceptually simple for users to understand and fairly simple to implement for both the academic institution and service provider. It also usually gives seamless access to e-resources for users, since the authentication will occur without the user interface being taken over – as far as the user is concerned, it happens completely transparently. The major failing of IP Authentication, however, is that it technically does not do what is required – that is, to establish a definite connection between the user and an institution. All IP Authentication “proves” is that a particular user is currently connected to the campus network of a particular institution – and nothing more. Additionally, since the individual is never uniquely identified, personalisation cannot be performed – at least, without the service provider deploying another access method on top.

IP Authentication is very widely deployed - despite the major failing outlined above - because of the ease of use it represents for members of institutions when on their campus network.

<table>
<thead>
<tr>
<th>IP Authentication</th>
<th>User</th>
<th>Institution</th>
<th>Provider</th>
</tr>
</thead>
</table>
| **Pros**          | * Easy to use  
* Allows anonymous access  
* Can access resources directly - no requirement to go via a portal page  
* No work required to enable access  
* No credentials to remember | * Allows library walk-in use  
* No technical work required | * Easy to implement  
* When members of an institution leave, they lose access automatically  
* No single point of failure |
| **Cons**          | * No off-campus access  
* No personalisation (without combining with another access method) | * Administration work in keeping IP address lists up to date at all service providers  
* Cannot control access to e-resources for specific users  
* No dissemination of information to users required | * Only approximates the requirement of being able to “prove” an individual is a member of a particular institution  
* Administration work in keeping IP address lists up to date for all customers  
* Cannot uniquely identify users |
Many academic institutions have implemented a URL-rewriting web proxy such as Proxy; this is a technology that sits on the campus network and acts as a go-between for traffic from off-campus users and e-resources. Instead of a user directly requesting a web page from the service provider hosting it, they instead request it from specially constructed URL on the web proxy which fetches the real page and shows it to the user. Thus, all requests appear to be coming from the web proxy and not the user's device.

Using a URL-rewriting web proxy in conjunction with IP authentication (i.e. the IP address of the proxy server is one of those known to be “owned” by an institution) provides many of the advantages of IP Authentication whilst alleviating one of its major issues – off-campus use is possible. No authorisation step is possible with this access method. URL-rewriting web proxies are widely deployed since they extend the ease of use of IP authentication to allow off-campus access. Institutions usually provide a library portal page showing all resources that are available to users; the links on these pages will point to the specially constructed links on the proxy server.

<table>
<thead>
<tr>
<th>URL-rewriting Web Proxy</th>
<th>Institution</th>
<th>Provider</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pros</strong></td>
<td>* Can control off-campus access to e-resources by enacting access control policies on proxy server</td>
<td>* Easy to implement</td>
</tr>
<tr>
<td>* Easy to use</td>
<td></td>
<td>* When members of an institution leave, they lose access automatically</td>
</tr>
<tr>
<td>* Off-campus access</td>
<td></td>
<td>* No single point of failure</td>
</tr>
<tr>
<td>* Allows anonymous access</td>
<td></td>
<td></td>
</tr>
<tr>
<td>* No work required to enable access</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cons</strong></td>
<td>* Have to implement the URL-rewriting web proxy</td>
<td>* Only approximates requirement of being able to “prove” a user is a member of a particular institution</td>
</tr>
<tr>
<td>* Have to access resources via library portal page</td>
<td>* No administrative work</td>
<td>* Administration work in keeping IP address lists up to date for all customers</td>
</tr>
<tr>
<td>* No personalisation possible</td>
<td>* No dissemination of information to users required</td>
<td>* Cannot uniquely identify users</td>
</tr>
<tr>
<td>(without combining with another access method)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>* Credentials to remember to log in to proxy</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2.1.3 VPN/Proxy Server (& IP Authentication)

This access control method is conceptually similar to the above method, except in a more general way. Many academic institutions have implemented a proxy server or VPN service for its users. Proxy servers are a technology that sit on the campus network and acts as a go-between for all traffic from off-campus users and various services, including internet access. A VPN service allows off-campus users to establish a secure connection onto the campus network, meaning their device is essentially on-campus. Thus, for both technologies, all requests appear to come from on-campus. Using these technologies in conjunction with IP authentication provides many of the advantages of IP Authentication whilst alleviating one of its major issues – off-campus use is possible. No authorisation step is possible with this access method.

Proxy Servers and VPN services are fairly widely deployed for various reasons, and if they exist can be used to help give off-campus access via IP authentication to e-resources. However, neither provide a simple means of access for users, since a user has to configure their web browser to use a proxy server or has to establish a VPN connection which often requires they install a VPN client on their device.

<table>
<thead>
<tr>
<th>VPN &amp; IP Authentication</th>
<th>User</th>
<th>Institution</th>
<th>Provider</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pros</strong></td>
<td>* Easy to use once configured</td>
<td>* Can control off-campus access to e-resources by enacting access control policies on proxy server or VPN service</td>
<td>* Easy to implement</td>
</tr>
<tr>
<td></td>
<td>* Off-campus access</td>
<td>* When members of an institution leave, they lose access automatically</td>
<td>* No single point of failure</td>
</tr>
<tr>
<td></td>
<td>* Allows anonymous access</td>
<td>* No administrative work</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>* No dissemination of information to users required</td>
<td></td>
</tr>
<tr>
<td><strong>Cons</strong></td>
<td>* Have to perform an initial configuration step</td>
<td>* Have to implement a proxy server or VPN service</td>
<td>* Only approximates requirement of being able to “prove” a user is a member of a particular institution</td>
</tr>
<tr>
<td></td>
<td>* No personalisation possible (without combining with another access method)</td>
<td>* No administrative work</td>
<td>* Administration work in keeping IP address lists up to date for all customers</td>
</tr>
<tr>
<td></td>
<td>* Credentials to remember to log in to proxy/VPN</td>
<td>* No dissemination of information to users required</td>
<td>* Cannot uniquely identify users</td>
</tr>
</tbody>
</table>
2.1.4 Institutional Credentials

A service provider creates a single set of credentials for every institution that subscribes to their resource. These credentials are given out to the relevant institutions, who can then give it out to its members as necessary. When users wish to access a restricted resource, the service provider asks them to provide the institutional credential, which is checked against their local credential store. Institutional credentials offer a method of authentication that works by assuming that if a person knows the credential for a particular institution, then that user is a member of that institution. No authorisation step is possible.

The major advantage of using institutional credentials is that it is practically simple to implement, and offers a greater level of control over a resource than IP Authentication. However, it is not particularly secure: ex-members of institutions will still be able to access the resource (unless the service provider recreates the credentials are regular intervals), members of an institution can share the credentials with other people, etc.

<table>
<thead>
<tr>
<th>Institutional Credentials</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pros</strong></td>
<td><strong>Institution</strong></td>
<td><strong>Provider</strong></td>
</tr>
<tr>
<td>* Off-campus access</td>
<td>* Can control access to e-resources</td>
<td>* Easy to implement</td>
</tr>
<tr>
<td>* Allows anonymous access</td>
<td>* No technical work required</td>
<td>* “Proves” a person is a member of a particular institution</td>
</tr>
<tr>
<td>* Can access resources directly - no requirement to go via a portal page</td>
<td>* No library walk-in use</td>
<td></td>
</tr>
<tr>
<td>* No work required to enable access</td>
<td>* Administration work in managing all institutional credentials</td>
<td></td>
</tr>
<tr>
<td></td>
<td>* Have to disseminate all institutional credentials to members</td>
<td></td>
</tr>
<tr>
<td>* Have to remember institutional credential for each resource</td>
<td></td>
<td></td>
</tr>
<tr>
<td>* No personalisation possible (without combining with another access method)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Cons</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>* The “proof” is not particularly verifiable (usernames can be shared)</td>
</tr>
<tr>
<td>* Administration work in managing credentials</td>
</tr>
<tr>
<td>* When members of an institution leave, they still know the institutional credentials</td>
</tr>
<tr>
<td>* Cannot uniquely identify users</td>
</tr>
</tbody>
</table>
2.1.5 Service Credentials

A step up in complexity from institutional credentials are service credentials – every single user has their own set of credentials. A common way in which this is implemented sees a user of a resource create a set of credentials at the website of the service provider; they would at this point indicate somehow which is their home organisation (e.g. directly, or by entering their institutional email address). This request is sent to the home organisation, who can vet all requests and mark the account as belonging to one of their members.

Service credentials offer a method of access control that combines authentication and authorisation into one; the very fact that the user has a set of valid credentials that has previously been marked as owned by a member of an organisation both authenticates and authorises them.

While this method allows off-campus use of a resource, and the service provider to uniquely identify users, a great deal of administrative overhead is required. The service provider needs to manage a potentially very large credential store (with all of the overhead that comes with this – e.g. password resets), while the institution has to vet all requests and authorise the accounts of its members. There is also an element of insecurity inherent in this approach – given that the credential gives access only to a specific resource, there is no particular reason for users to not share their credentials with other people.

<table>
<thead>
<tr>
<th>Service Credentials</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>User</strong></td>
</tr>
<tr>
<td><strong>Pros</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Cons</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
2.1.6 Centralised Access Management

A centralised access management infrastructure is a more scalable approach to providing users with credentials for use with e-resources. A single organisation (independent of academic institutions and service providers) acts as an access control mediator: they maintain a central credential store for all users of all institutions and the service provider configures their resource to hand access control decisions to this mediator. The central credential store could be managed centrally, or be devolved to the institutions themselves. Credentials are usually provided to users by their institution.

The major advantage of using centralised access management for the user is they have only one extra set of credentials to remember to access all service providers who make use of the system. This also reduces the administrative overhead of accounts as compared to Service Credentials – there is only one set of credentials to manage. However, the access control mediator has to ensure their service is as resilient as possible since it is a single point of failure – if it fails, access for all users to all service providers fails.

A highly successful example of a centralised access management system is the Classic Athens system in the United Kingdom, managed by Eduserv since 1996.

<table>
<thead>
<tr>
<th>Centralised Access Management</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>User</strong></td>
</tr>
<tr>
<td><strong>Pros</strong></td>
</tr>
<tr>
<td><em>Off-campus access</em></td>
</tr>
<tr>
<td><em>Allows personalisation</em></td>
</tr>
<tr>
<td><em>A single set of credentials for multiple services</em></td>
</tr>
<tr>
<td><em>Can access resources directly - no requirement to go via a portal page</em></td>
</tr>
<tr>
<td><em>No work required to enable access</em></td>
</tr>
<tr>
<td><strong>Cons</strong></td>
</tr>
<tr>
<td><em>No anonymous access</em></td>
</tr>
<tr>
<td><em>The single set of credentials is (usually) another set of credentials to remember</em></td>
</tr>
</tbody>
</table>
2.1.7 Federated Access Management

Federated Access makes use of a trust relationship between institution and service provider, devolving the responsibility of authentication and authorisation to the institution. A user authenticates at their home institution, proving to them which member of that institution they are, and the institution essentially “vouches” to the resource for the individual attempting to access it.

In Federated Access, both authentication and authorisation happens at the the institution end -- the results are passed to the service provider in a trustworthy verifiable manner. To authenticate to their home institution to enable this, the institution implements FAM software which links to an existing credential store – the user authenticates with credentials they already own.

A major advantage of this type of access management are that all access decisions made at the service provider end are based on information provided by the home institution of the user – the entity which is most likely to have accurate and up-to-date information about the user. Another advantage (the most important from the user perspective) is that they can use existing credentials to gain access to a resource – they do not have to remember another username and password.

The major disadvantages, however, are that federated access technology needs implementing at both institution and service provider, a scalable method of managing trust relationships between all parties is needed, and the problem of "discovery" is encountered (see Section 3).

### Federated Access Management

<table>
<thead>
<tr>
<th>Pros</th>
<th>Institution</th>
<th>Provider</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Off-campus access</em></td>
<td><em>Administrative overhead already exists</em></td>
<td><em>Can uniquely identify users</em></td>
</tr>
<tr>
<td><em>Allows personalisation</em></td>
<td><em>Can control off-campus access to e-resources</em></td>
<td><em>No administrative work required</em></td>
</tr>
<tr>
<td><em>Allow anonymous access</em></td>
<td><em>Technical implementation of FAM software</em></td>
<td><em>Can identify individuals with help of institution</em></td>
</tr>
<tr>
<td><em>Use existing credentials to access resources</em></td>
<td><em>No work required to enable access</em></td>
<td><em>“Proves” a person is a member of a particular institution</em></td>
</tr>
<tr>
<td><em>Can access resources directly - no requirement to go via a portal page</em></td>
<td><em>Technical implementation of FAM software</em></td>
<td><em>No library walk-in use</em></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cons</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Discovery</em></td>
<td><em>Technical implementation of FAM software</em></td>
</tr>
</tbody>
</table>
2.2 Finding Content

Given that a resource has implemented one or more of these access methods, how do users find their content?

Note that the use-case examined here is that of a user who has found a specific citation and they wish to gain access to the actual article the citation refers to. The use-case of a user wishing to search for articles within a specific area is another problem entirely and not explored in this study.

There are several main methods that a user can use, including:

• Direct access - if the user, looking at the citation, can identify the provider of that article and knows the web address of that provider's website, they may go directly to the source. At this point, they will need to use one of the access methods presented above to prove their entitlement to gain access to that article.

• Direct access via library portal – if the user, looking at the citation, can identify the provider of that article, they can usually go to their institution's library portal; this is a service many institutions offer than lists all electronic resources the institution subscribes to, with direct links to each resource.

• Link resolvers – if the user has found the citation on an OpenURL resolver enabled article database, a link will be shown next to the citation that allows the user to access the actual article (presuming the institution has a license agreement with the provider). The link will connect the user to the institution's link resolver, which automatically calculates whether the institution has access to the full article and, if it does, sends the user directly to the article in at the provider's website.

• Aggregators – these are web sites that do not publish content of their own but instead collect content in a single place. The user, citation in hand, could access an aggregator and search for the article given information in the citation. If the aggregator finds the article , the user can then access it (presuming the institution has a license agreement with the provider).

• Search engine – the least targeted manner of finding an article referred to in a citation is for the user to use a standard search engine, such as Google or Yahoo! to search for the article given information in the citation. This will potentially lead the user to the provider of that article's website, at which point, they will need to use one of the access methods presented above to prove their entitlement to gain access to that article.

2.3 Conclusions

Providers of electronic resources can make use of one of a number of access methods that enable an individual wishing to access their protected resource to show that they are covered by an existing subscription agreement and should be given access. Of these, IP Authentication is widely deployed due to the ease of use it represents from an end user perspective – despite its shortcomings. Federated Access is the latest in this long line of these access methods to have been developed, and represents an access method that could potentially provide a good balance between ease of use from an end user perspective and security and verifiability from the resource provider perspective. The one major usability issue surrounding federated access, however, is the issue of discovery. This is examined in detail in the next section.
3 The Discovery Problem

An important part of federated access is Identity Provider discovery; known as the Discovery Problem. This section reviews what actually this problem actually is, and the solutions to the problem that exist.

3.1 Federated Access and Discovery

The typical use-case of a user interacting with a service provider and wishing to gain access to restricted content occurs along the following lines:

1. The user, using their web browser, connects to a service provider and requests to view restricted content.
2. The service provider receives this request. To ascertain whether this person should be granted or denied access, they need to know some information about that person.
   - In the federated world, this means that the user needs to be sent to their home organisation's identity provider, which will “vouch” for that person and pass across information about them to the resource provider.
3. The service provider “discovers” which is the user's home institution
4. The service provider redirects the user to their home institution's identity provider.
5. The user authenticates at their identity provider, which responds to the service provider, letting them know that this user authenticated successfully, and often providing some information about that user.
6. The service provider receives this information, and then either grants or denies access based upon the information it received.

In the logic flow presented above, step 3 is the crux of the discovery problem - how does the service provider figure out which is the home institution/identity provider of the user?
3.2 Discovery Solutions

There are several methods currently in use that perform this discovery step. These can be split into three categories – avoiding discovery, client-less discovery, and client-mediated discovery.

3.2.1 Avoiding Discovery

The first category of solution to the discovery problem is not in fact a solution at all, but a way of avoiding the problem in the first place for certain use cases.

To achieve this, each institution can configure a page (usually their existing library portal page) to list all resources available to their users along with links to these resources. These links can be carefully constructed such that they send the user to that institution's identity provider, and once the user has successfully authenticated, directly onto that resource. Thus, the service provider never has to "discover" which institution the user is from, since the first time they see the user the user has already authenticated. This approach is known as IdP-initiated SSO.

While this approach can be very effective for a given use case – where the user first goes to the institution's library portal page to access resources – it doesn't actually solve the problem of the main use case described – a user starting out at the service they wish to access. That particular use case requires another approach, known as SP-initiated SSO, which comes in two forms: discovery that works without the need of a client, and discovery mediated by a client. Each of these is explored next.
3.2.2 Client-less Discovery

SP-initiated discovery without the use of a client is conceptually very simple – the service provider asks the user to manually tell them which is their home institution.

This method of discovery comes in two forms:

- The user tells the service provider directly; or
- The service provider sends the user to a centrally provided service; the user tells this service.

In the first case, the service provider's website itself asks the user the question “Where are you from?” This is known as an in-built WAYF. In the case of general federated access, this could involve the service provider showing the user a list of logos (e.g., the Google OpenID identity provider, the Yahoo OpenID identity provider, etc.). In the specific case of Shibboleth, this will usually involve the service provider showing a list of subscribing institutions. Either way, the user has to indicate which is their identity provider by manually clicking on the correct link/entry.

In second case, the service provider sends the user to a centrally provided service (that is, provided by the federation operator) called the Discovery Service (DS), or Where Are You From (WAYF) service (while there are technical differences between a DS and a WAYF, conceptually they are similar enough that the terms are used interchangeably throughout this document). This service will ask the user to indicate where they are from in exactly the same way as the first case.
In-built WAYFs offer a potentially "friendlier" end user experience, since the user stays within the realm of the service provider's website at all times; whereas sending a user to a WAYF service disparate from the service provider's website introduces a potential element of confusion into the end user experience. Additionally, a centrally provided WAYF service will normally list all members of a federation (potentially many options for the user to choose from) while an in-built WAYF can be configured to only show the subset of those options that actually subscribe to that service provider's service.

However, both styles of client-less discovery, while conceptually fairly simple to understand and implement, suffer from two main problems: a scaling issue, and a multi-federation issue.

The first of these issues is simple – as more and more institutions adopt Federated Access technologies around the world, the list of institutions that a user has to choose from continually grows. The UK federation alone, for example, has already reached the point where users have to find their home institution from many hundreds of options. This is not a good experience for the end user at the best of times; given the current increase in usage of mobile computing platforms such as smartphones and netbooks where this issue becomes a greater problem, this issue will only grow steadily worse.

The second of these issues is seen on those service providers that provide their service to customers worldwide, and therefore have to exist on multiple federations (since federations are currently largely geographically based). In this case, before the service provider can ask the user to indicate which institution they are from, the usually first narrow down the list of options by first having to ask them which federation they belong to. This is also not a particularly good experience for the end user; not least because the vast majority of users have no idea what a federation is, let along which one is the correct choice for them. While federations are largely geographically based, however, this issue can be somewhat mitigated by asking the user what country they are from in order to deduce
the correct federation – a question that makes infinitely more sense to the average user.

### 3.2.3 Client-mediated Discovery

Client-mediated SP-initiated discovery solves the main issue of client-less discovery – that fact that a user has to manually indicate where they are from – through the use of a client that essentially answers that question on behalf of the user or removes the need to ask the question in the first place. The obvious main caveat of this approach is that the user needs to have a client installed and configured to do this for them.

This method of discovery comes in three main forms:

- The user's client tells the service provider where the person is from;
- The user's client is the identity provider;
- The user's client proxies the identity provider.

The first of these methods requires a plug-in to the user's web browser be implemented. This plug-in will be listening for WAYF requests coming from a service provider and will automatically answer on behalf of the user. An example of this is provided in the SAML2 specification, which includes the idea of ECP (Enhanced Client or Proxy) which formally specifies how this can happen. No real implementation of the ECP protocol yet exists, however.

The second of these methods – the client is the identity provider - is less relevant to the main use case described in this document – that of a user's institution “vouching” for them that they are covered by an existing license agreement – and is included here for completeness sake only. This is because this method centres around the idea of self-issued identity. The service provider could ask the client “who are you”, and the client would respond with self-asserted information configured by the user.

In the final of these methods – the client proxies the identity provider – the service provider asks the client to provide authentication and authorisation information. The client then sends this request on to the user's identity provider (it proxies it), receives the response, and in turn sends this response back to the service provider. Thus, the service provider never actually needs to know which identity provider it is indirectly talking to. An example of this approach is the idea of a “managed card” in the Windows Cardspace system.

Client-mediated discovery offers a very good end user experience, since it is a solution to the discovery problem that is far more intuitive (and potentially seamless) as compared to client-less discovery.

### 3.3 Conclusions

Discovery is a necessary process in the world of federated access. IdP-initiated SSO as a method of avoiding discovery is a good method of bypassing the need for discovery given specific circumstances, but is not applicable the general use case described in this study.

Client-mediated discovery is undoubtedly the best actual solution to the discovery problem since it is fairly intuitive and effortless process (other than initial setup of the client) for the user, thus providing a good end user experience. However, the user needs to have a client installed and configured correctly. At the present time, such clients are relatively uncommon, and thus client-less discovery will remain important for at least the short to medium term future. Even long term, client-less discovery may have to remain an option for those cases where users wish to gain access to restricted content via federated means when not using a client managed by them, or using a device which does not support such a client.
4 Existing Discovery Solutions

Section 2 discussed the wide range of access methods available for users to establish their affiliation to an academic institution, and Section 3 examined the possibilities available to perform the Discovery step for the Federated Access Management access method. This section examines how prevalent each of these methods and possibilities actually are.

4.1 Methodology

To gather statistics the project team surveyed all of the major resources on the UK federation, as well as a number of its smaller resources. This consisted of 81 sites. For each of these, various aspects of the federated login process were recorded; including the terminology and methods in use for each of the four possible login steps (as discussed below).

4.2 Federated Login Steps

The process of SP-initiated client-less federated login will involve the user working their way through one or more of the following steps in the journey that takes place between a user deciding they wish to view restricted content, and logging in at their institution's Identity Provider:

1. Finding the Login Area
   • Firstly, the user will have a find a “login” link on the page they are currently viewing. This might sometimes be a simple login link that takes them to a dedicated login page; sometimes it might be a dedicated login area on the page itself that contains login options; while other times a login request may only be shown to the user when attempting to access content itself.

2. Choosing to login via Federated means
   • Next, if the resource provides several access methods, the user may have to select which one they wish to use.

3. Choosing the correct Federation
   • Once the user has found their way to a page which starts the Federated Access Management option, if the resource exists on multiple federations the user is often asked to choose the correct federation for this institution.

4. Choosing the correct Institution
   • Once the user has indicated which federation they are from, the user will be presented with a choice of institutions that they could be from, and will be asked to select theirs.

Note that not all of the above steps may be present at a particular service provider – those that only sell their resource to members of a single federation may skip step 3, for example.
4.3 Results

4.3.1 Login Steps

The first area examined was a simple user interface metric designed to assess the effort required from the user to perform federated login: the amount of steps necessary between a user deciding to login and finding themselves at their institution's Identity Provider was recorded. In this context, a “step” means a movement from one web page to another – so clicking a login link to move to a login page, then clicking a Federated login option to move to a page offering a list of all institutions on the UK federation, then choosing an institution from a drop-down menu to move to that institution's Identity Provider would involve a total of three steps.

Figure 4 shows the results of this investigation. The smallest amount of steps seen was one – a list of institutions was available directly on the main page; while the largest was four – each of the login steps was given its own webpage. The most common, by quite a margin, was two steps for those resources which existed on the UK federation only (a federated login link on the main page led to a list of UK federation institutions. The second most common was three steps – mostly seen at those sites which exist on multiple federations – where the extra step performed the “choosing the correct federation” action.
4.3.2 Finding the Login Area

In the sites surveyed, 28% had a login link that took a user to a login technology page, 62% had the Shibboleth login choice embedded on main page (either on front page or an automatically displayed “please login” page), 2% asked the user to login when attempting to access content.

Looking at the terminology in use on either the login link itself, or that is displayed on the dedicated login page, 78% used terminology based around “login”, 11% around “sign in”, and 2% used “click here to access/enter”.

Figure 5: Method used to start the login process

Figure 6: Terminology used to start the login process
4.3.3 Access Method

Looking at what methods of access were available on the surveyed sites, all offered a Federated Access option (since we were surveying FAM enabled sites). 63% offered access via Athens, 62% with service credentials, and 7% via public library card. Note that we did not gather information for IP Authentication or institutional login since this information was not readily available.

Figure 7: Access methods available
4.3.4 Choosing Federated Login

When choosing the Federated Access login, 24 different variations of terminology on 5 major themes were seen. The major themes were based around the following words:

- Academic
- Federation
- Institution
- Organisation
- Shibboleth

Of these, institution was the most common (41%), followed closely by federation (33%). Shibboleth featured fairly regularly (16%), then organisation (7%), then academic (3%). Some resources used combinations of multiple terminology – e.g. “Institutional Log-in (Shibboleth).

One area of potential confusion for users is that some resources use the terminology “institutional login” to describe federated login, while some use it to describe “institutional credentials” (as defined in Section 2.1.4).

For completeness, the 24 variations in this terminology were:

- Academic Sign in
- Academic Users
- Host Institution
- Home Institution
- Institutional access
- Institutional username and password
- Institution login
- Institutional Login
- Athens/Institutional Login
- Institutional Log-in (Shibboleth)
- Participating Institution
- Your Institution
- Home Organisation
- Shibboleth
- Shibboleth Log
- Shibboleth Users
- Shibboleth/Open Athens
- UK Federation
- UK Federation (Shibboleth)
- UK Federation WAYF
- UK Federation authentication
- UK Federated Access Management
- UK Institution Users
- UK Access Management Federation
4.3.5 Choosing the Federation

30% of the resources surveyed allowed federated access from multiple federations. Of those, all but one (the Internet2 Shibboleth wiki), used their own in-built WAYF. Of those 30% of resources, 36% presented it to the user by requesting they choose their “Federation”, then their institution, while 64% presented it by requesting the user choose their region or country, then their institution.

Three main methods of allowing a user to select their federation/region were in use – a scroll menu, a drop down menu, and a list of links. Of these, the drop down menu was by far the most prevalent (seen in 72% of cases), followed by scroll menu (20%), and finally the list of links (4%)
4.3.6 Choosing the Institution

When employing a Discovery/WAYF Service, the distribution seen between those using the centrally provided UK federation WAYF (which uses the Internet2 provided software), and those that deployed their own in-built WAYF, was roughly half-half. All resources that allow federated access via multiple federations deployed their own in-built WAYF.

![Figure 11: Discovery Service choice](image)

The most common terminology in use on the in-built WAYF service deployed when prompting the user was based around “institution” (e.g. “please select your institution”). The terms “Federation”, “Identity Provider”, “Organisation” and “Shibboleth” were all seen, but far less often.

![Figure 12: Terminology used on in-built WAYF](image)
The three most common methods in use on the in-built WAYF service deployed were (in order of prevalence): displaying a list of institutions (where each entry was a clickable link), using a drop down list of institutions, and using a scroll menu of institutions. Search interfaces were deployed on a very small number of resources.

![Figure 13: Technology used on in-built WAYF](image)

Of the in-built WAYF services surveyed, roughly one third gave the user the option of remembering their choice of institution to speed up the process the next time the user visits the resource. Of those, a wide range of options were available to the user; the most common was a simple “remember” option which remembered the selection indefinitely (as long as the cookie persisted).

![Figure 14: Remembering selection on in-built WAYF](image)

![Figure 15: Options available to remember selection on in-built WAYF](image)
4.4 Conclusions

The main conclusion to be drawn from the information gathered is that there is a severe lack of consistency in the way that every single step of this process is presented to users.
5 End-user Evaluation of Existing Discovery Solutions

Given the wide range of terminology and technology/methods in place for all steps of the discovery process of federated login, the obvious next step is to find out what real users think of each of the options. This section presents this project's investigation of this area.

5.1 Methodology

Given the data presented in Section 4, a set of ten resources were chosen that were representative of the most commonly seen terminology/methods for each of the steps of the discovery process. Several sessions of around 15 people each were run at Cardiff University and Kidderminster College, soliciting input from different user groups. The user groups in question were:

- HE Undergraduate Students (Science subject)
- HE Undergraduate Students (Humanities subject)
- HE Undergraduate Students (Medical subject)
- HE Taught Postgraduate Students
- HE Research Postgraduate Students
- FE Students
- HE Library Staff
- HE Staff
- FE Staff

During each session, the following process occurred:

- First, a general welcome and description of the purpose and procedure of the study was given to participants. This included describing the four separate steps of discovery that participants were going to encounter.
- Secondly, each participant was given a questionnaire to fill out. This questionnaire first aimed to gather some background information about the participant (such as gender, time in sector, general e-resource use, existing awareness of logos and terminology, etc.). Next, it asked the participant to work their way through each of the ten resources (presented to each participant in a random order so as not to create any bias in the results), rating what they thought of each of the login steps in terms of clarity of terminology used, ease of use of the technology/method in place, and various other aspects (mostly using a 5 point Likert scale). Finally, it asked the participants to rank the 'best' three resources in terms of clarity of terminology and ease of use.
- Finally, a short debriefing session took place where, as a group, the participants were asked what they thought of what they had seen, what they thought was particularly good or bad, and how they thought the user experience could be improved.

Following these sessions, all data was entered into a database to allow for analysis. A large collection of results thus exists; the most interesting results are presented next.
5.2 Overall Results

The main findings of this end user study are as follows:

- The biggest problem is the complete lack of consistency in how federated login is presented, in terms of both terminology and design;
- Participants generally do not understand any of the terminology in use – and do not want to;
- Participants generally “learn” what to look for to accomplish federated login (without knowing what that actually is) and on subsequent visits to new resources simply look for something that matches what they are used to seeing;
- In terms of how to present the discovery process to the user, less is more – users preferred those resources than managed to combine multiple steps on one page;
- An in-built WAYF service has the advantage of integration (in terms of design) with the resource employing it, while a centrally provided WAYF service has the advantage of consistency – once the user has used it once, they become familiar with it. Producing a set of design and terminology guidelines for in-built WAYF deployment which providers follow would combine both of these advantages, resulting in a better user experience.
- Participants indicated (during debriefing) that a well implemented dynamic search facility would be their preferred method of finding their Identity Provider;
- There was no significant variation on any of these findings for any of the participant categories (apart from Library Staff – since they understand the technology behind the scenes a lot better than the average user).
5.3 Participant Analysis

128 participants took part in these sessions; between them they made 891 individual ratings of resources.

For each participant:

- Figure 16 shows the distribution of the category;
- Figure 17 shows the distribution of gender;
- Figure 18 shows the distribution of the length of time in the sector;
- Figure 19 shows the usage of e-resources;
- Figure 20 shows an estimation of on-campus versus off-campus access to e-resources;
- Figure 21 shows the existing awareness of relevant logos; and
- Figure 22 shows the existing awareness of relevant terminology.

Figures 16 to 18 show that a good spread of participant category, gender, and length of time in sector were achieved for the end user study.

![Figure 16: End User Study - Participants by category](image-url)
Figure 17: End User Study - Participants by gender

Figure 18: End User Study - Participants by length in sector
Figures 19 and 20 show that the vast majority of all types of participants regularly or occasionally use e-resources as part of their job or their studies; the estimation provided shows a roughly half-half split between accessing these resources on the campus network as opposed to elsewhere.
Figure 21 and 22 shows that:

- Roughly half of participants recognised the Eduserv Athens logo while almost 70% recognised the name;
- Few recognised the Shibboleth “griffin” logo while roughly 40% recognised the name;
- Roughly 40% recognised the UK federation logo while more than half recognised the name;
- Roughly a quarter recognised the “federated access” terminology.

![Figure 21: End User Study - Existing awareness of logos](image1)

![Figure 22: End User Study - Existing awareness of terminology](image2)
5.4 Detailed Results

5.4.1 Login Steps

For each of the ten resources, the number of login steps between the user visiting the resource and navigating to their Identity Provider (presuming the user clicked the correct links each time) was recorded. The “ease of use” ratings recorded per resource, were then averaged, grouped by the number of login steps.

The results show a clear trend of the simpler the login process (i.e. the less the number of steps) the higher participants rated that resource.

The conclusion to draw here is that if service providers wish to improve the user experience of federated login, they should endeavour to minimise the number of steps required.

Figure 23: Average Rating vs Number of Login Steps

The results show a clear trend of the simpler the login process (i.e. the less the number of steps) the higher participants rated that resource.

The conclusion to draw here is that if service providers wish to improve the user experience of federated login, they should endeavour to minimise the number of steps required.
5.4.2 Login Method

For the first stage of achieving federated login – finding the login page/area of a resource - participants were asked to rate the clarity of terminology (“Login”, “Sign In”, or “Click here to access/enter”) and the ease of use of the style of login employed (a login link which moves the user to a separate login page, or an embedded login area on the main page).

The results clearly show that participants viewed the terminology of “Login” (followed by “Sign in”) as the most clear. While resources using terminology such as “Click here to access/enter” may consider that particular terminology as more intuitive than more technical terms such as “login”, any person who uses any web based resource (be it email, social networking, e-resources, or whatever) is exposed to the term “login” extremely regularly – and thus its familiarity overrides any advantage that a non-technical term may have.

Participants rated a login option embedded directly within the page slightly higher than finding a login link; this is possibly explained by the previous conclusion that the more steps are involved, the lower participants ranked “ease of use” - and a login link to a separate page represents one more step than the embedded alternative.
5.4.3 Choosing Federated Access Type

For the second stage of achieving federated login – choosing the access method if more than one option was present - participants were asked to rate the clarity of terminology in use which attempted to make the user understand this was the federated login option (based around either “Academic”, “Institution”, “Organisation”, “Shibboleth”, or “Federation”).

The results show that all terms used received similar ratings – no one particular term stood out to participants as particularly clear or intuitive.

Comments received from participants in the debriefing sessions indicated that this was because none of the terminology made any intuitive sense to them – they were simply looking for terminology that they recognised. Also, a general consensus was seen during the debrief that terminology around “Academic login” or “Educational login” would make the most intuitive sense – “Academic”, however, received the lowest rating. This is likely due to the fact that it is very rarely used so scored low due to it failing the familiarity test.

![Figure 26: Choosing Federated Login - Rating of clarity of terminology](image)
5.4.4 Choosing Correct Federation (DS/WAYF of WAYFs)

For the third stage of achieving federated login – choosing the correct federation if the resource allowed federated access from multiple federations - participants were asked to rate the clarity of terminology in use which attempted to assist the user in understanding what was required. The two styles used by the test resources were those that asked the user to choose the federation and those that asked them to choose the country. They were also asked to rate the ease of use of the style of selection employed (either a drop down list or a scroll list).

The results show the terminology based around country received a slightly higher rating that around the federation: but the difference was not significant. This may, however, be due to the fact that the name of the federation they were looking for (i.e. UK federation) is relatively geographically based already. If participants were from the US, for example, and required to choose either “United States” (or similar) or “InCommon”, this result may have been different.

Looking at the ratings for ease of use of the method of selection employed, both styles of method seen received fairly similar rankings; neither stood out as particularly better.

Comments received from participants in the debriefing sessions indicated that when it comes to terminology here, making the selections geographically based makes the most intuitive sense; however, they are also looking for anything familiar at all times – so a federation name may also be helpful. A general consensus was seen that presenting the choice in the form of “country (federation name)” would provide a user with both sets of information, and therefore the best user experience.
5.4.5 Choosing Correct Institution (DS/WAYF Service)

For the final stage of achieving federated login – choosing the correct Identity Provider - participants were asked to rate the clarity of terminology in use which attempted to assist the user in understanding what was required. They were also asked to rate the ease of use of the style of selection employed (either a drop down list or a scroll list).

The first area looked at was the average rating given for ease of use by the choice of DS/WAYF in use – either an in-built WAYF, or the centrally provided UK federation WAYF. The results seen show no particular difference in rating. However, comments received during the debriefing session indicated that while the in-built WAYFs offer a better user experience in theory – since the user isn't redirected to a completely external web site – use of the centrally provided WAYF service provided a much more consistent user experience – after using the centrally provided WAYF a few times, users learn how to use it and what it does. A general consensus seen in many of the groups indicated that widespread use of in-built WAYF services that follow strict guidelines on terminology and design (thus making in-built WAYF services across multiple resources very similar) would have the advantages of both approaches and would therefore offer the best user experience.
With regards terminology used on a DS/WAYF, the results seen show that the direct approach of simply asking the user “Where are you from?” received the highest rating in terms of clarity. “Shibboleth” received a surprisingly high rating – this is quite possibly due to the fact that it is a fairly memorable word and participants became familiar with it relatively quickly.

Figure 30: Choosing Correct Institution - Rating of clarity of terminology
Looking at the rating of ease of use of the actual method employed to enable the user to select the relevant Identity Provider, the drop down option received the highest rating, followed closely by the scroll menu. Presenting the user with a list of clickable links scored a little lower; while a search interface scored lowest. However, the results seen for the search facility were almost certainly biased by the fact that those few providers who employ a search option mainly employed a very basic search interface.

Comments received during the debriefing session indicated that if a provider employed multiple methods (e.g. a drop down list and a search facility), they generally only noticed and used the first one presented. They also indicated that their preferred option would, despite their ratings given, be a search interface – but only if designed such that it was very easy to use; their definition of “easy to use” generally centred around it being dynamic (i.e. as they start typing they are shown a list of matching institutions which focuses down as more letters are typed). This process should start from the first letter typed (no matter how many items matched) since that would indicate to the user from the first key press that this method was in place.
6 Provider Views on Discovery

As a part of the study, a day long workshop for service providers was held at Cardiff University with the purpose of soliciting their views on the discovery problem. At the same time, their views on federated access in general was gathered.

The workshop included around 15 participants from a range of service providers – from large well known organisations that sell their product worldwide, to smaller providers that sell to a small amount of UK institutions only.

This section describes a small selection of the points raised and views put forward during the day deemed most relevant to this study. These points are categorised functionally and presented as a simple list.

6.1 Access to resources

The changing world:

• The user base that many publishers sell to today is broadening;
• Non academic customers need to supported;
• Publishers provide what the market desires in terms of access methods
• Non academic users need to be supported

Access methods:

• All providers present support IP Authentication – since it is easy to set up (but not as easy as many people think) and represents an access method that is very easy for the user to use. However, its main problems are that it does not provide off-campus access (and therefore the experience between using their resource on- and off-campus is different) and they cannot uniquely identify their users.
• Some providers present employ service credentials. The main issue with them is that users readily share their credentials – a problem particularly bad for businesses customers. However, it represents an easy method of provided access for customers not affiliated with a particular institution or organisation.
• Some providers present employ institutional credentials – but only as a last resort given their major security drawbacks.
• Centralised Access Management was seen as an access method that the majority of publishers were fairly happy with – Athens being a major example. However, customer demand is now shifting towards federated access.
• Publishers are worried that federated access in general, and Shibboleth in particular, is just “the next thing” and that new technologies may come along and replace it. However, they also think that generally it would be “Wise to be more proactive”.

6.2 Adopting Federated Access

Why adopt Federated Access?

• The business case is simple – customers are demanding it. Implementing it, however, requires an effort to convince management, due to the work involved and the potential loss of “control” that comes from adopted a generally privacy-preserving technology. The
demand is slowly increasing, mostly from Europe (and recently, Japan). There is still a lack of demand within the United States.

How is Federated Access adopted?

- Often the design decisions are made by technical staff, rather than designers or management. This is because federated access is seen as a “technical thing”, despite the fact that it presents a major part of the user interface of their web site.

6.3 Discovery

The need for “brand” for federated access:

- A consistent brand (name and logo) for the process of federated access would be useful and would represent a big step forward.

- Such a brand would make the end user experience more acceptable – and anything that removes potential barriers that could drive away customers can only be a positive step forward.

- However, such a brand would need to be adopted at a global level to be of any use. Having many separate brands (e.g. one per federation) would not help matters and therefore publishers would find it difficult to adopt. All participants of the workshop agreed that if a single brand was adopted globally, and the brand made sense to them, they would not hesitate to start using it on their resources.

- A consensus emerged that publishers never collectively decide anything, given that many are direct commercial competitors, and thus either federation operators or institutions themselves (or bodies that represent institutions such as JISC or UCISA in the UK) are the right people to drive forward this process of creating a brand.

- Finally, having separate “federated login” and an “Athens” login is confusing for users and should disappear where possible, given that users can do federated login via OpenAthens. Removing Athens branding would make things simpler for the user.
## 7 Conclusions and Recommendations

This study, informed by research of all major services on the UK federation, end user evaluation, and publisher input has identified a number of key points around the area of Discovery in federated access. The most important of these points are as follows:

- Current solutions to client-less discovery presently employed by service providers are highly inconsistent in terms of terminology, design, and implementation; this represents a very poor end user experience.
- End users generally do not understand any of the terminology in use – and do not wish to need to understand it. Their primary view was that picking a single term to describe federated login – it does not matter what they term is – so that they can quickly become familiar with it and simply look for that term when they encounter a new resource would represent the best end user experience.
- IP Authentication represents a very good end user experience, and is therefore in widespread usage despite its major shortcomings. Publishers are very reluctant to switch to an access method that degrades this end user experience by increasing the barrier to gaining access to subscription resources in case this drives down usage of their resource. Thus, they are highly interested in any improvement in the end user experience of federated access, and given widespread usage of this improvement would be very willing to adopt it themselves.

### 7.1 Recommendations

Recommendation 1 – A brand should be created for academic federated access. For this brand to be successful, it needs widespread adoption worldwide. The brand should include a short name and a logo; these need not mean anything but simply provide a familiar point of reference.

A brand for academic federated access would answer the key points raised at the start of this section; mainly, the issue of consistency. If widely adopted, this would represent a good end user experience – they need only learn the name/logo of the brand once, and for all subsequent usage they can simply look for the recognisable item.

The authors of this report are by no means experts in branding, so the actual name and logo are not specified in this report. However, the following recommendations around the brand are made:

- The brand should include a short name. The name need not mean anything, but could include short words to help identify it as academic federated access. Examples of possibilities include such names as “eduAuth”, “eduID”, “eduLogin”, or “eduAccess” (to pick a naming scheme – i.e. edu* - already prevalent in Europe).
- The brand should also have a small, simple, graphical logo. An example of a successful logo that the brand should aim to follow (in terms of impact, not actual design) is that of OpenID.

The choosing of a name and logo should be performed by community consensus (where the community in question is federation operators worldwide) in order to increase the chance of successful widespread adoption. Thus, it is recommended that a process of consultation should take place (possibly led by the JISC as an interested party) to choose the details of the brand. It may, however, be a good idea to involve professional brand consultants during some part of the process in order to increase the effectiveness of the brand, and therefore its success.
A set of guidelines will allow publishers to implement federated login and discovery within their user interface in a manner consistent with other publishers with a minimum of effort. These guidelines should be created alongside the brand since they represent an aspect of the brand. Given the results gathered during the end user study, however, it can be recommended that a potentially good end user experience would be as follows:

- The brand name/logo should be present directly on the main page of a service provider, rather than on a separate login page.
- This link should take the user to an in-built WAYF service if possible, or a centrally provided WAYF service if not. Which is used, the WAYF service's primary method for a user to find their identity provider should be a dynamic search input box that should display a drop down list of identity providers that focusses and reduces this list as the user enters more characters. This list should start displaying from the first character input (no matter how large the amount of matching identity providers, in order to demonstrate how the search facility works from the very outset. An example technology that could implement this idea readily is the jQuery AJAX library.
- A secondary method of a static drop down menu in order that users using non-javascript enabled devices, et, can still make a choice may also be necessary to include in the WAYF.
- If possible, users should never be asked to identify which federation they are from. This could be implemented as part of the dynamic search facility – the search could match all identity providers across all federations that the service provider is a member of, and present the matching items in the form of “institution name (country/federation)” or similar in order to help the user identify the correct identity provider in case of ambiguity where two institutions in different federations share the same name.
- The WAYF service should offer the user the option of remembering their choice in order to make it easier the next time they visit. The lifetime of this choice should be a large period of time, but offer the user the option to “forget” their choice if necessary.

Implementing the discovery service in the manner described above should reduce the amount of login steps necessary and provide a large measure of consistency, thus enhancing the end user experience.

Recommendation 2 – A “style guide” should be created for publishers to follow around implementing discovery using the brand created.