

Shibboleth as a framework to enable access to distributed PDP material

Francis Lowry, NTU March 2006

The main aim of this project was to explore how to use Shibboleth as a mechanism to allow access to PDP data, building on the existing work of the RIPPLL project using UK LeaP to model PDP data to support transfers.

The model explored was based on a set of web services protected by Shibboleth, which enabled the PDP data to remain distributed but allowed the collation of remote data sets into one single presentation view. The datasets used were the same as for the main RIPPLL transition data, with the addition of some dummy qualification records from the Background Qualifications section within the NTU VLP.

One of the areas not considered at this point is the referencing of PDP Material, i.e. a learner's link to his or her own content hosted on a remote site. We hope to explore this in a subsequent project.

Core Assumptions:

- Functionality needs to match the UKLeaP specification in terms of privacy, i.e. separation of data into Public and Private
- Shibboleth is core to providing access to resources providing the data
- Data is distributed and is referenced dynamically at runtime
- The learner has full control over data both in terms of management and publication
- Security and access is in-built
- The model proposed would be as generic and extensible as possible.

Rather than the 'pass the parcel' route, whereby the PDP data physically migrates from institution to institution along with the learner, this model assumes that PDP data, or aspects of PDP data, remains at the original institutions and that each institution will provide access to the material on demand via web services.

Within each institution there is a catalogue of PDP elements which are available for 'publication'. This catalogue forms part of the authorisation and provisioning process, with the learner having control over which components of his or her PDP portfolio are accessed from which remote location. This catalogue provides the basis for the data provisioning where, for a specific learner, a web service will return the PDP data relating to the learner to the consumer web service.

The institution at which the learner is aggregating PDP data provides him or her with a registration process which allows them to register a remote web service and specify what data from that web service can be accessed both privately and publicly using some form of identity attribute. A consumer web service then uses this local registry to access data on behalf of the learner and presents it in the format requested.

There are two formats explored here. The first allows learners to review all the PDP data, marked both private and public, recorded for them at each of the provisioning institutions. The second provides a public view, e.g. a mini CV, to any person who has access to the appropriate URL i.e. this URL provides a non-Shibboleth authenticated presentation of some of the public data that the learner has decided should be made available.

Shibboleth and user-specific resource availability

One of the fundamental precepts of Shibboleth is that the specific identity of the person accessing a resource is not required for authorisation: a trust relationship with the originating identity provider and specific combinations of attributes provide the

appropriate information to allow resources to be made accessible. To enable user-specific release of data such as PDP requires us to request some form of a unique identifiable attribute for the learner to enable provision of a user specific service. While Shibboleth does not preclude this functionality, in order to implement this or similar feature sets institutions must include appropriate attributes to allow user-specific access to resources.

The developmental model comprises of several components:

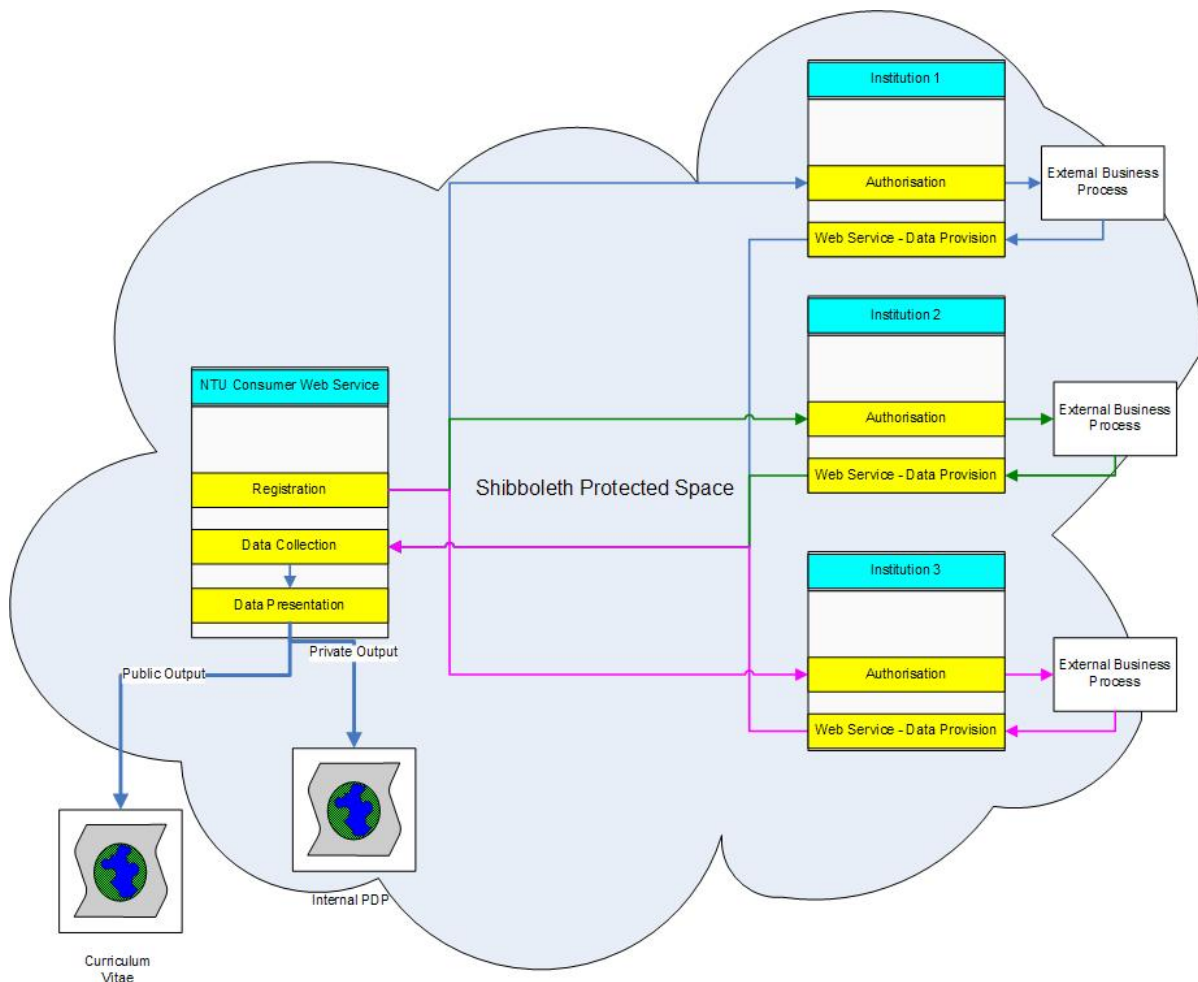
Provider (multiple)

- Authorisation process
- Data provision

Consumer

- Registration
- Collection/Aggregation.
- Presentation

with Shibboleth as the framework for authentication and access:



Overall Process

Registration and authorisation:

1. Learner accesses current institution's PDP registration web service
2. Learner requests authorisation from provider institution
3. Once authorisation is approved, learner registers specific PDP attributes to be utilised from the provider institution
4. Process is repeated until learner has registered all data required from all providing institutions.

Collection and provisioning:

1. PDP data is requested based on the attributes registered with the consumer
2. Each provider returns the data to the consumer
3. Data is formatted and presented depending on the output type.

Development framework

The following tools and infrastructure were used in the development and testing of this model. (Additional work needs to be done to extend the testing across multiple Shibboleth sites and to refine the data model used within the web services, but the basic model and concepts function as planned.)

- Shibboleth 1.3 Identity provider
 - Windows 200x server
 - CAS
 - Tomcat 5.5
 - iDP pointing to an MS Active Directory for identification details and attributes
- Shibboleth 1.3 Service provider
 - Windows 2003 Server
 - Microsoft IIS (versions 5 & 6)
- Development tools
 - Microsoft Visual Studio 2003
 - C#
- Multiple Windows XP workstations with IIS deploying the provider web services
- Original RIPPLL project UK LeaP data stored in relational form in a SQL Server database.

For testing purposes, the consumer web service is the only component protected by Shibboleth at the moment. This will be extended in future phases to allow for each provider to sit behind a different Shibboleth service provider.

This is not a major security concern as, depending on how the keys and the registration process outlined below are implemented at each provider, and as the aggregation and presentation of the data at the consumer institution end is protected with Shibboleth, it is not essential from a technical perspective to protect the provider institution's web services.

PDP registration

This functionality resides at the consumer institution. Once the learner has logged into Shibboleth and appropriate attributes have been released to allow for individual user identification, the registration service allows the learner to record which components of PDP data are to be utilised, both from a private and public perspective, and from where they should be sourced. A simple design based on the UK LeaP specification was used to identify the individual PDP components used:

```
- <learnerinformation>
  - <identification>
    + <name>
    + <demographics>
    + <address>
    + <contactinfo>
    + <contactinfo>
  + <reflexion>
</learnerinformation>
```

The data mapping to these elements was stored in a relational database with each provider web service returning different data.

Authorisation process

This component has not been developed to any degree as it is anticipated that it will reside at the provider institution. The assumption is that there will be some supporting business process which will authorise the release of the PDP data to the requestor and allow an appropriate unique identifier to be returned to the PDP registration. This will act as a unique reference for that learner's PDP data sourced from the provider. At present, the assumption has been made that the unique identifier will be the primary key record for the source database.

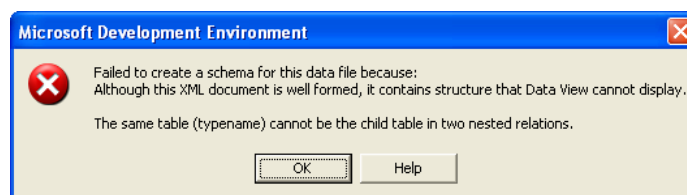
The value stored in the registration component is in itself meaningless. With the exception of the core identity data, which is assumed to be provided from the learner's current institution and does not have to be provided through this model, none of the data streams have any identifiable attributes for the learner.

In addition to this, by allowing the local authorisation process to define what the unique key is, the option has been left open to allow for a private/public key shared arrangement to either encrypt the data stream back to the consumer web service, or even to use the shared keys to sign the returned data, i.e. to allow for assessments to have an additional guarantee of authenticity. The consumer service just needs to provide a unique key for the learner originally provided in the registration process, generally the *HTTP_SHIB_USERNAME* attribute, along with the key provided by the provider resource to allow the provider to confirm authenticity. Wrapping the communications between the consumer web service and the provider within SSL will allow a fuller, more complete security model.

Data provision

Once the consumer institution has registered a provider institution's web service and obtained authorisation to access the data, then the aggregation component can request the data from the service provider.

The initial assumption was that the data to be parsed to the provider web service would be UK LeaP formatted XML. Unfortunately, although the XML is well formed, the development tools used were not able to parse them due to an error with C# in .NET.



Despite both providing a DTD and reverse generating a valid schema for the XML, we were still unable to use the UK LeaP data as a data source to the provider web service (see [Appendix 1](#) for sample XML).

The error message relates to an acknowledged limitation with .NET. Once all the components are functioning, we plan to return to exploring using UK LeaP as the mechanism for modelling XML data from the web services, and hope to find an appropriate work-around.

Rather than get involved with technical issues with the UK LeaP XML at this stage, we decided to push the datasets derived from the SQL tables. These were replicated as temporary tables within the .aspx application with the following structure:

PERSON

person_key	INTEGER
person_title_code	VARCHAR
surname	VARCHAR
forename	VARCHAR
gender_code	VARCHAR
date_of_birth	DATE
nationality_code	VARCHAR

BACKGROUND_QUALIFICATIONS

person_id	INTEGER
qual_achieved_date	DATE
qual_level	VARCHAR
qual_description	VARCHAR
qual_result	VARCHAR

PERSONAL_STATEMENT

person_key	INTEGER
statement_date	DATE
statement_title	VARCHAR
statement	VARCHAR

All the web service providers were designed to return the data to the consumer in these structures. As long as the web service provider structured its output data in this format, the data could be sourced from any file format, database or remote location.

Data collection/aggregation

After the requests for the data have been made, the web service starts to collect the returned objects and aggregate the data prior to presentation. To allow for appropriate debugging and testing, the current collection service is not threaded: it waits for a specific period for the data to be returned from each provider service. This does mean that the performance has not been optimised; however this drawback is acceptable at this phase of development. Threading the consumer service will be explored at a later phase when the developments are moved to a multi-Shibboleth model.

Presentation

By separating the collection phase and defining it as a stand-alone component, it is very simple to define different presentational/functional views of the PDP data. Two basic views have been developed to demonstrate the integration of these web services with Shibboleth. The first is a very simple page with a drop-down allowing the learner to see all the data that is held in the remote systems irrespective of the security settings on the data, i.e. the learner can preview both public and private data. The second is a very

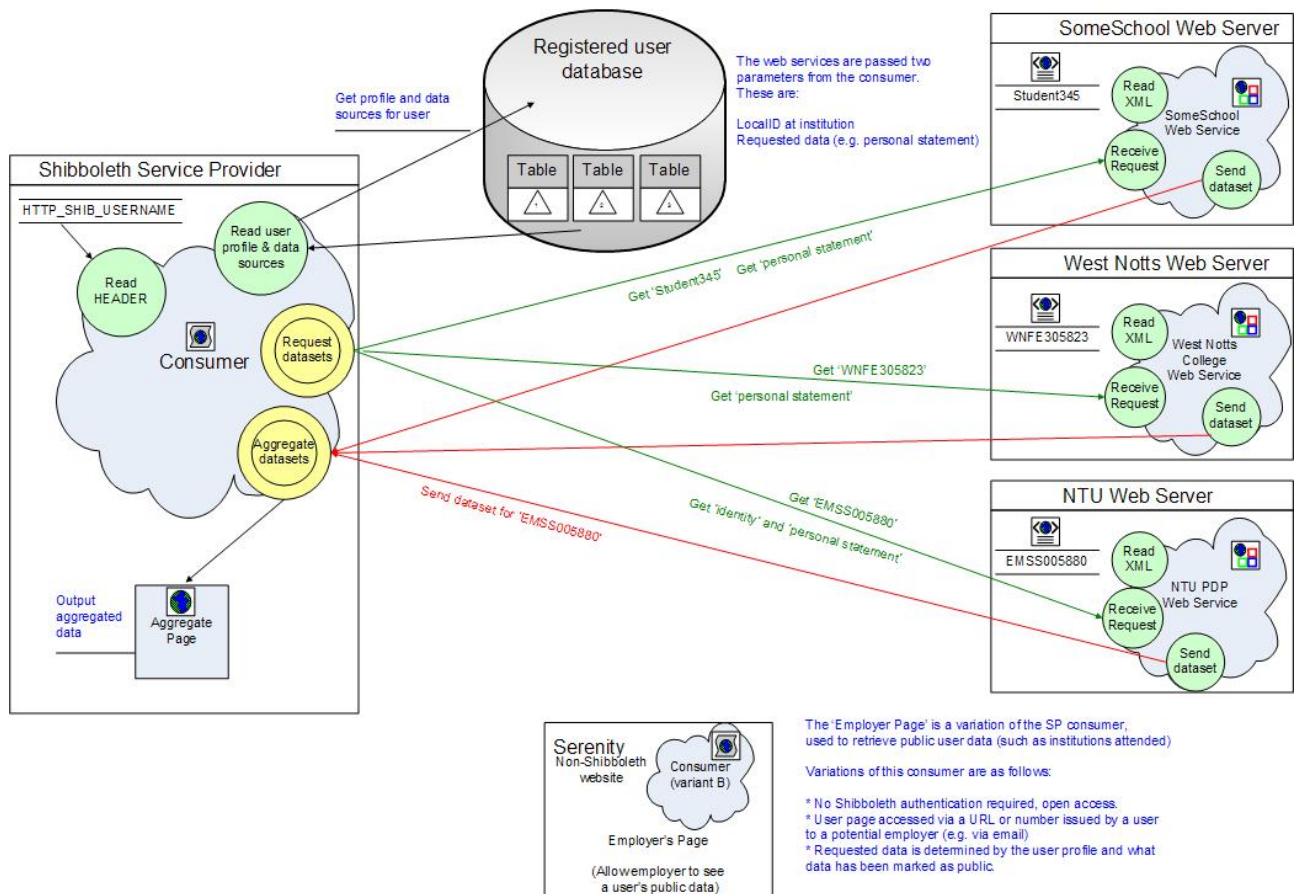
simple CV presentation of just the learner's public data. Additional qualification elements have been added to flesh out the presentation page.

One future area to explore is the possibility of signing the qualifications, effectively providing some form of electronic assertion that the data displayed in this electronic CV is actually the confirmed qualification from the originating institution. As stated in the [Authorisation process](#) above, the model has been left open to allow for this future possibility.

Current operational system

The development project has been nicknamed SquirrelIWS. As shown below, the consumer service which will collate all the data from the distributed provider web services is using the HTTP_SHIB_USERNAME as the identifiable attribute for the learner.

This is a key requirement for any user specific resource provision brokered by Shibboleth, and for SquirrelIWS this is the unique key field used within the registration database to map the local learner's ID to the remote provider's web services and PDP data.

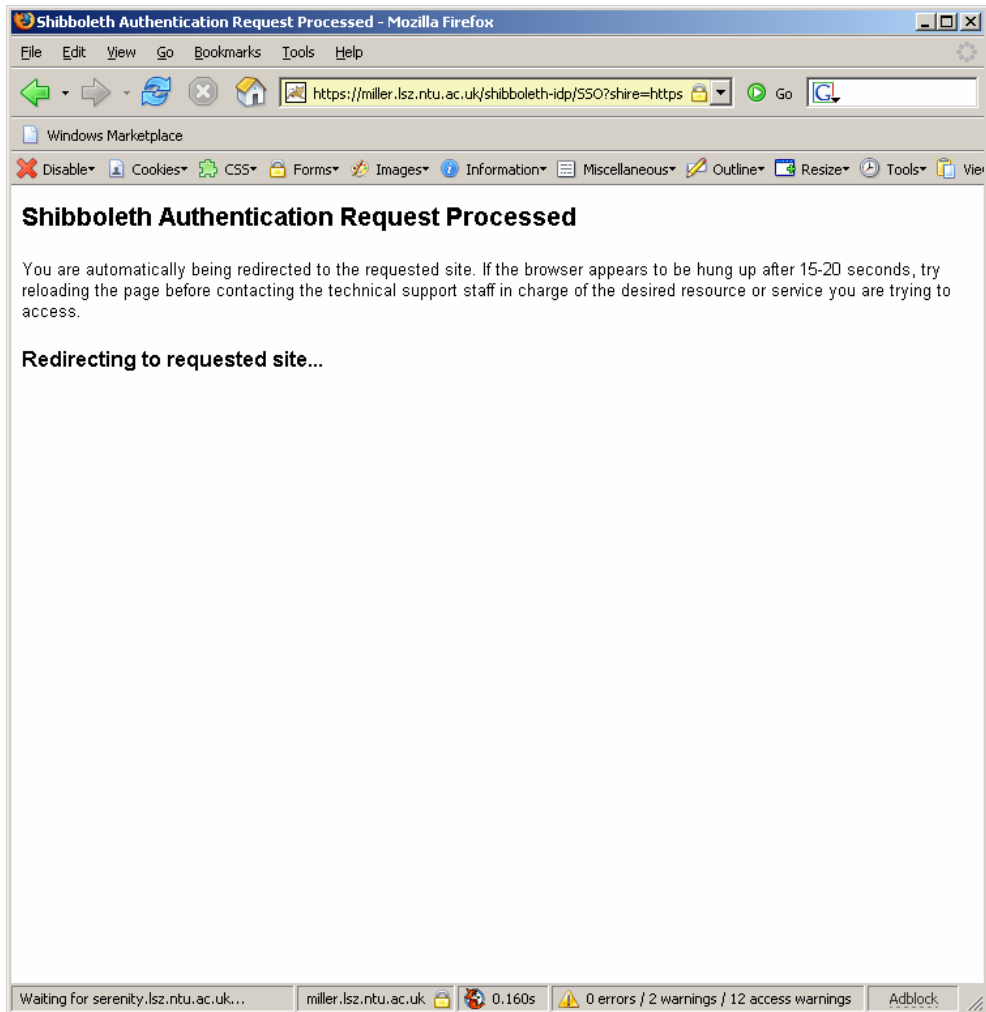


This demonstration model includes one consumer, with three providers all hosted on remote web servers. The model will scale to any number of providers and this is controlled by the entries in the 'Registered User' database.

The steps below walk through the final development model demonstrating a very simple aggregation screen showing all data returned and a mini-CV presentation of public data only. This is not a final polished product, rather sample services to demonstrate the principles. The main component missing is the screen to define which data to obtain from which service.

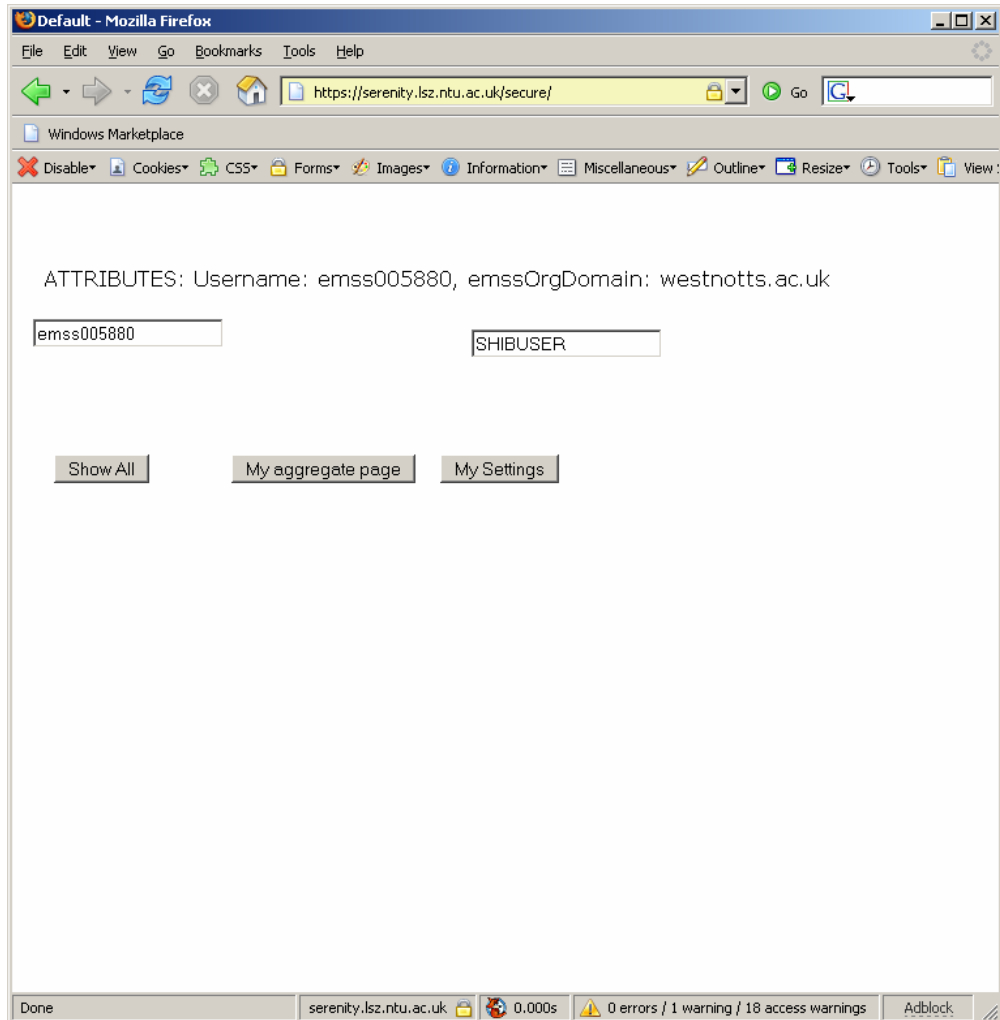
Step 2 – Shibboleth redirection

On successful login to Shibboleth via the IDP (on the 'miller' server at NTU), Shibboleth will redirect back to the Shibboleth PDP Consumer web service.



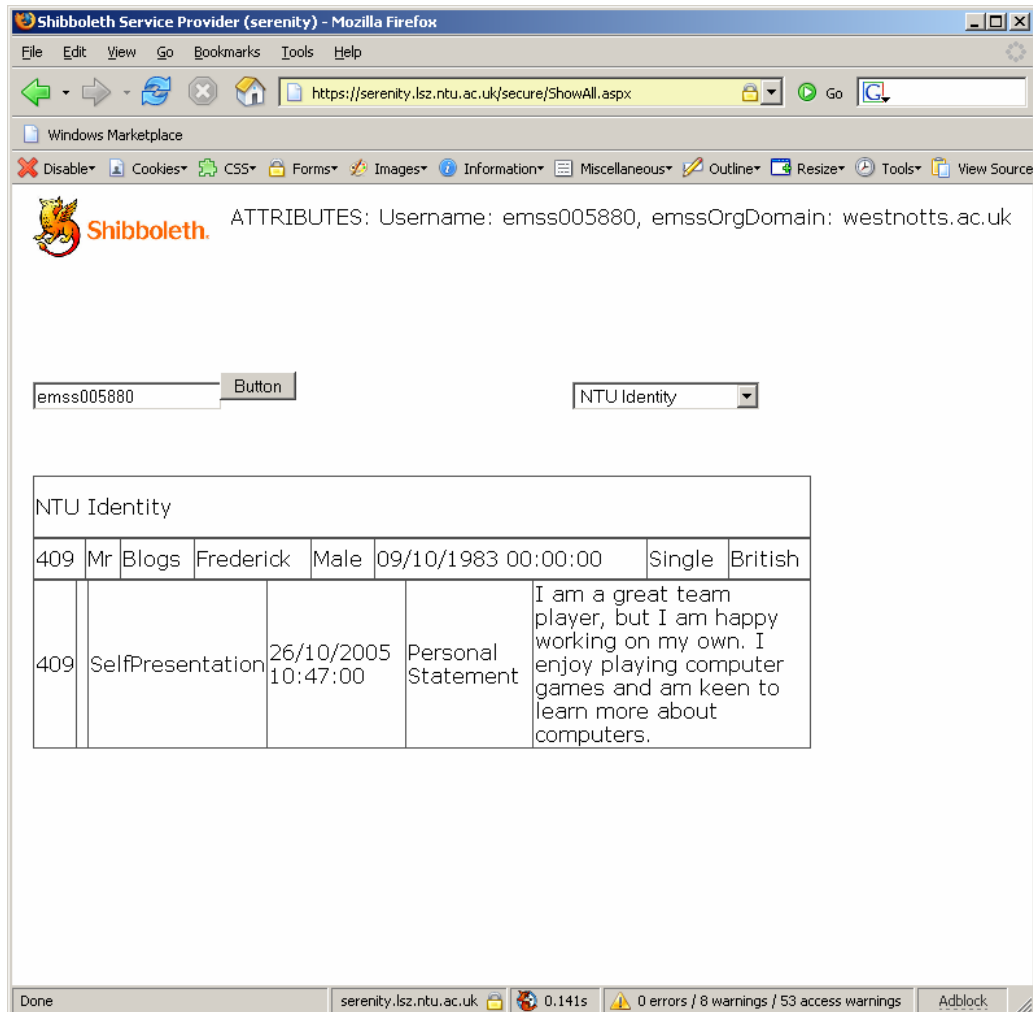
Step 3 – Options screen

Default redirection will present the screen below with three options. The 'Show All' button will display all the data provided from each web service, irrespective of the public/private attribute for the data. 'My Aggregate page' will display a mini-CV which contains very basic identity data, qualifications and some personal statements, effectively combining PDP data from all three sources.



Step 4 a – Identity data and personal statement from NTU

This ShowALL page displays all the data from each service provider along with the Shibboleth attributes requested. Here the data from the NTU web service is displayed:




Shibboleth Service Provider (serenity) - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

https://serenity.lsz.ntu.ac.uk/secure/ShowAll.aspx

Windows Marketplace

Disable Cookies CSS Forms Images Information Miscellaneous Outline Resize Tools View Source

 **Shibboleth.** ATTRIBUTES: Username: emss005880, emssOrgDomain: westnotts.ac.uk

emss005880 Button NTU Identity

NTU Identity							
409	Mr	Blogs	Frederick	Male	09/10/1983 00:00:00	Single	British
409	SelfPresentation		26/10/2005 10:47:00	Personal Statement	I am a great team player, but I am happy working on my own. I enjoy playing computer games and am keen to learn more about computers.		

Done serenity.lsz.ntu.ac.uk 0.141s 0 errors / 8 warnings / 53 access warnings Adblock

Step 4 b – Qualifications and Personal Statement from the West Notts College service provider

Shibboleth Service Provider (serenity) - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

https://serenity.lsz.ntu.ac.uk/secure/ShowAll.aspx

Windows Marketplace

Disable Cookies CSS Forms Images Information Miscellaneous Outline Resize Tools View Source

ATTRIBUTES: Username: emss005880, emssOrgDomain: westnotts.ac.uk

emss005880 Button WestNotts

WestNotts				
4121	June 2000	A Level	Computing	C IT innit
4122	June 2000	A Level	Psychology	C where stuff is
4123	July 2000	A Level	Human Biology	E living things under a microscope
4124	June 2000	A Level	General Studies	D waste of study time
412	SelfPresentation	21/03/2006 12:36:00	Personal Statement	I am a great team player, but I am happy working on my own. I wish to extend my knowledge of DOT NET development.

Done serenity.lsz.ntu.ac.uk 0.125s 0 errors / 8 warnings / 84 access warnings Adblock

Step 4 c – Qualifications from the Secondary Modern web service

Shibboleth Service Provider (serenity) - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

https://serenity.lsz.ntu.ac.uk/secure/Default.aspx

Windows Marketplace

Disable Cookies CSS Forms Images Information Miscellaneous Outline Resize Tools View Source

ATTRIBUTES: Username: emss005880, emssOrgDomain: westnotts.ac.uk

emss005880 Button Secondary Modern

Secondary Modern

411	1	June 1998	GCSE	English	A	reading and writing			
411	2	June 1998	GCSE	Mathematics	B	numbers and stuff			
411	3	June 1998	GCSE	French	E	say what?			
411	4	June 1998	GCSE	Woodwork	C	wood n stuff			
411	5	June 1998	GCSE	Science	B	Physics Biology Chemistry			

Done serenity.lsz.ntu.ac.uk 4.531s 0 errors / 7 warnings / 85 access warnings Adblock

Step 5 – Mini-CV, Aggregation page

Here all the **public** data from the three web services are combined and displayed in a mini-CV page.

Again there is very little polish to the page, but the concept is clearly demonstrated:

emss005880 Button

Here is your aggregate page. These details are flagged as 'public'

Mr Frederick Blogs Male
09/10/83 British

qual_achieved_date	qual_level	qual_description	qual_result
June 1998	GCSE	English	A
June 1998	GCSE	Mathematics	B
June 1998	GCSE	French	E
June 1998	GCSE	Woodwork	C
June 1998	GCSE	Science	B
June 2000	A Level	Computing	C
June 2000	A Level	Psychology	C
July 2000	A Level	Human Biology	E
June 2000	A Level	General Studies	D

statement_date	statement_title	statement
21/03/2006 12:36:00	Personal Statement	I am a great team player, but I am happy working on my own. I wish to extend my knowledge of DOT NET development.
26/10/2005 10:47:00	Personal Statement	I am a great team player, but I am happy working on my own. I enjoy playing computer games and am keen to learn more about computers.

Done serenity.lsz.ntu.ac.uk 0.484s 0 errors / 11 warnings / 78 access warnings Adblock

Appendices

Appendix 1 - Example source data

West Notts XML

```
<?xml version="1.0" encoding="utf-8"?>
<learnerinformation>
  <identification>
    <name>
      <typename>
        <tysource sourcetype="standard">UKLeaP</tysource>
        <tyvalue>Full</tyvalue>
      </typename>
      <partname>
        <typename>
          <tysource sourcetype="standard">UKLeaP</tysource>
          <tyvalue>Given</tyvalue>
        </typename>
        <text>fred</text>
      </partname>
      <partname>
        <typename>
          <tysource sourcetype="imsdefault" />
          <tyvalue>Surname</tyvalue>
        </typename>
        <text>blogs</text>
      </partname>
    </name>
    <demographics>
      <date>
        <typename>
          <tysource sourcetype="standard">UKLeaP</tysource>
          <tyvalue>Birth</tyvalue>
        </typename>
        <datetime>1983-10-09T00:00:00</datetime>
      </date>
    </demographics>
    <address>
      <typename>
        <tysource sourcetype="standard">UKLeaP</tysource>
        <tyvalue>Private</tyvalue>
      </typename>
      <street>
        <streetnumber>50</streetnumber>
        <streetname>Somewhere St</streetname>
      </street>
      <city>Somewhere</city>
      <region>Somewhereshire</region>
      <postcode>S012 3ME</postcode>
    </address>
    <contactinfo>
      <typename>
        <tysource sourcetype="standard">UKLeaP</tysource>
        <tyvalue>Private</tyvalue>
      </typename>
      <telephone>
        <areacode>01234</areacode>
        <indnumber>567890</indnumber>
      </telephone>
    </contactinfo>
    <contactinfo>
      <typename>
        <tysource sourcetype="standard">UKLeaP</tysource>
        <tyvalue>Private</tyvalue>
      </typename>
      <email>fredblogs1@gmail.com</email>
    </contactinfo>
  </identification>
  <reflexion>
    <typename>
      <tysource sourcetype="standard">UKLeaP</tysource>
      <tyvalue>SelfPresentation</tyvalue>
    </typename>
    <contenttype>
```

```

<referential>
<indexid>reflexion_1</indexid>
</referential>
</contenttype>
<date>
<typename>
<tysource sourcetype="standard">UKLeaP</tysource>
<tyvalue>Create</tyvalue>
</typename>
<datetime>2005-12-11T09:15:00</datetime>
</date>
<description>
<short>Personal Statement</short>
<long>I am a great team player, but I am happy working on my own. I wish to
extend my knowledge of operating systmss and database development.</long>
</description>
</reflexion>
</learnerinformation>

```

West Notts DTD

```

<?xml version="1.0" encoding="UTF-8"?>
<!ELEMENT address (typename, street, city, region, postcode)>
<!ELEMENT areacode (#PCDATA)>
<!ELEMENT city (#PCDATA)>
<!ELEMENT contactinfo (typename, telephone?, email?)>
<!ELEMENT contenttype (referential)>
<!ELEMENT date (typename, datetime)>
<!ELEMENT datetime (#PCDATA)>
<!ELEMENT demographics (date)>
<!ELEMENT description (short, long)>
<!ELEMENT email (#PCDATA)>
<!ELEMENT identification (name, demographics, address, contactinfo+)>
<!ELEMENT indexid (#PCDATA)>
<!ELEMENT indnumber (#PCDATA)>
<!ELEMENT learnerinformation (identification, reflexion)>
<!ELEMENT long (#PCDATA)>
<!ELEMENT name (typename, partname+)>
<!ELEMENT partname (typename, text)>
<!ELEMENT postcode (#PCDATA)>
<!ELEMENT referential (indexid)>
<!ELEMENT reflexion (typename, contenttype, date, description)>
<!ELEMENT region (#PCDATA)>
<!ELEMENT short (#PCDATA)>
<!ELEMENT street (streetnumber, streetname)>
<!ELEMENT streetname (#PCDATA)>
<!ELEMENT streetnumber (#PCDATA)>
<!ELEMENT telephone (areacode, indnumber)>
<!ELEMENT text (#PCDATA)>
<!ELEMENT typename (tysource, tyvalue)>
<!ELEMENT tysource (#PCDATA)>
<!ATTLIST tysource
    sourcetype (imsdefault | standard) #REQUIRED
>
<!ELEMENT tyvalue (#PCDATA)>

```

West Notts XML with W3C schema

```

<?xml version="1.0" encoding="UTF-8"?>
<xs:schema elementFormDefault="qualified"
xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:element name="address">
    <xs:complexType>
      <xs:sequence>
        <xs:element ref="typename"/>
        <xs:element ref="street"/>
        <xs:element ref="city"/>
        <xs:element ref="region"/>
        <xs:element ref="postcode"/>
      </xs:sequence>
    </xs:complexType>
  </xs:element>
  <xs:element name="areacode" type="xs:short"/>
  <xs:element name="city" type="xs:string"/>

```

```

<xs:element name="contactinfo">
  <xs:complexType>
    <xs:sequence>
      <xs:element ref="typename" />
      <xs:element ref="telephone" minOccurs="0" />
      <xs:element ref="email" minOccurs="0" />
    </xs:sequence>
  </xs:complexType>
</xs:element>
<xs:element name="contentype">
  <xs:complexType>
    <xs:sequence>
      <xs:element ref="referential" />
    </xs:sequence>
  </xs:complexType>
</xs:element>
<xs:element name="date">
  <xs:complexType>
    <xs:sequence>
      <xs:element ref="typename" />
      <xs:element ref="datetime" />
    </xs:sequence>
  </xs:complexType>
</xs:element>
<xs:element name="datetime">
  <xs:simpleType>
    <xs:restriction base="xs:dateTime">
      <xs:enumeration value="1983-10-09T00:00:00" />
      <xs:enumeration value="2005-12-11T09:15:00" />
    </xs:restriction>
  </xs:simpleType>
</xs:element>
<xs:element name="demographics">
  <xs:complexType>
    <xs:sequence>
      <xs:element ref="date" />
    </xs:sequence>
  </xs:complexType>
</xs:element>
<xs:element name="description">
  <xs:complexType>
    <xs:sequence>
      <xs:element ref="short" />
      <xs:element ref="long" />
    </xs:sequence>
  </xs:complexType>
</xs:element>
<xs:element name="email" type="xs:string" />
<xs:element name="identification">
  <xs:complexType>
    <xs:sequence>
      <xs:element ref="name" />
      <xs:element ref="demographics" />
      <xs:element ref="address" />
      <xs:element ref="contactinfo"
maxOccurs="unbounded" />
    </xs:sequence>
  </xs:complexType>
</xs:element>
<xs:element name="indexid" type="xs:string" />
<xs:element name="indnumber" type="xs:int" />
<xs:element name="learnerinformation">
  <xs:complexType>
    <xs:sequence>
      <xs:element ref="identification" />
      <xs:element ref="reflexion" />
    </xs:sequence>
  </xs:complexType>
</xs:element>
<xs:element name="long" type="xs:string" />
<xs:element name="name">
  <xs:complexType>
    <xs:sequence>
      <xs:element ref="typename" />
      <xs:element ref="partname" maxOccurs="unbounded" />
    </xs:sequence>
  </xs:complexType>
</xs:element>

```



```

<xs:element name="partname">
  <xs:complexType>
    <xs:sequence>
      <xs:element ref="typename"/>
      <xs:element ref="text"/>
    </xs:sequence>
  </xs:complexType>
</xs:element>
<xs:element name="postcode" type="xs:string"/>
<xs:element name="referential">
  <xs:complexType>
    <xs:sequence>
      <xs:element ref="indexid"/>
    </xs:sequence>
  </xs:complexType>
</xs:element>
<xs:element name="reflexion">
  <xs:complexType>
    <xs:sequence>
      <xs:element ref="typename"/>
      <xs:element ref="contentype"/>
      <xs:element ref="date"/>
      <xs:element ref="description"/>
    </xs:sequence>
  </xs:complexType>
</xs:element>
<xs:element name="region" type="xs:string"/>
<xs:element name="short" type="xs:string"/>
<xs:element name="street">
  <xs:complexType>
    <xs:sequence>
      <xs:element ref="streetnumber"/>
      <xs:element ref="streetname"/>
    </xs:sequence>
  </xs:complexType>
</xs:element>
<xs:element name="streetname" type="xs:string"/>
<xs:element name="streetnumber" type="xs:byte"/>
<xs:element name="telephone">
  <xs:complexType>
    <xs:sequence>
      <xs:element ref="areacode"/>
      <xs:element ref="indnumber"/>
    </xs:sequence>
  </xs:complexType>
</xs:element>
<xs:element name="text">
  <xs:simpleType>
    <xs:restriction base="xs:string">
      <xs:enumeration value="blogs"/>
      <xs:enumeration value="fred"/>
    </xs:restriction>
  </xs:simpleType>
</xs:element>
<xs:element name="typename">
  <xs:complexType>
    <xs:sequence>
      <xs:element ref="tysource"/>
      <xs:element ref="tyvalue"/>
    </xs:sequence>
  </xs:complexType>
</xs:element>
<xs:element name="tysource">
  <xs:complexType>
    <xs:simpleContent>
      <xs:extension base="xs:string">
        <xs:attribute name="sourcetype"
          use="required">
            <xs:simpleType>
              <xs:restriction
                base="xs:NMTOKEN">
                  <xs:enumeration
                    value="imsdefault"/>
                  <xs:enumeration
                    value="standard"/>
                </xs:restriction>
              </xs:simpleType>
            </xs:attribute>
          </xs:restriction>
        </xs:simpleType>
      </xs:extension>
    </xs:simpleContent>
  </xs:complexType>
</xs:element>

```

```
        </xs:extension>
      </xs:simpleContent>
    </xs:complexType>
  </xs:element>
  <xs:element name="tyvalue">
    <xs:simpleType>
      <xs:restriction base="xs:string">
        <xs:enumeration value="Birth"/>
        <xs:enumeration value="Create"/>
        <xs:enumeration value="Full"/>
        <xs:enumeration value="Given"/>
        <xs:enumeration value="Private"/>
        <xs:enumeration value="SelfPresentation"/>
        <xs:enumeration value="Surname"/>
      </xs:restriction>
    </xs:simpleType>
  </xs:element>
</xs:schema>
```