



## e-Framework Service Usage Model Name

- Name: Grid Enabling MIMAS Services (GEMS) Service Usage Model

## Version

- 1.2

## Version History

Version	Date	Author	Description	Organization / Project
1.2	Nov 2008	M. Argüello Casteleiro	Initial Draft	NCeSS, University of Manchester
		Pascal Ekin		NCeSS, University of Manchester

## Rationale

Establishing a social science Data Grid is a key component of the wider e-Social Science strategy. Current social science data infrastructure (academic and non-academic) needs to be Grid enabled in a standards compliant and sustainable way. Data service infrastructures need to be able to support multiple forms of access (i.e. single database approach) to minimise duplication of effort.

The SUM presented here has a specific focus on Grid Enabling MIMAS Services (GEMS) project that has the purpose of providing Grid enabled access to the aggregate statistics from the 2001 Census via OGSA-DAI on the National Grid Service (NGS). The advantage of the approach followed by GEMS is that it will maximise and build upon the ESRC/JISC investment in the establishment of existing data infrastructure, avoid having to maintain multiple database systems to support different forms of access and also provide a more rapid and flexible method of building Data Grids via the NGS.

## Classification<sup>1</sup>

<i>To be provided by the submitter:</i>				
<b>SUM Type</b>	<input checked="" type="checkbox"/> Domain	<input type="checkbox"/> CORE (a commonly recurring SUM; designation requires e-Framework Integrity Group approval)		
<b>Domain(s)</b>	<input type="checkbox"/> Learning & Teaching	<input checked="" type="checkbox"/> Research <input type="checkbox"/> Libraries	<input type="checkbox"/> Administration <input type="checkbox"/> IT Services	<input type="checkbox"/> Common
<b>Maturity</b>	<input type="checkbox"/> Immature	<input type="checkbox"/> Mature		
<b>Purpose(s)</b>	<input type="checkbox"/> Exemplar	<input checked="" type="checkbox"/> Application	<input type="checkbox"/> Modelling	<input type="checkbox"/> Toolkit
<b>XOR (exclusive "or")</b>	<input checked="" type="checkbox"/> Service Genres	<input type="checkbox"/> Service Expressions		
<b>Development Status</b>	<input type="checkbox"/> Proposed	<input type="checkbox"/> Developmental	<input type="checkbox"/> Prototype	<input type="checkbox"/> Production
<b>Deployment Scale</b>	<input type="checkbox"/> Isolated	<input type="checkbox"/> Ubiquitous		
<b>State Behaviour</b>	<input type="checkbox"/> Stateful	<input type="checkbox"/> Stateless		

<sup>1</sup> See definitions of the Service Usage Model Classification Scheme categories and their allowable choices at: <http://www.e-framework.org/Services/ServiceClassificationScheme/ClassificationSchemeForSUMs/tabid/817/Default.aspx>

Transactional Behaviour	<input type="checkbox"/> Transactional and ACID	<input type="checkbox"/> Transactional but Non ACID	<input type="checkbox"/> Non-Transactional
Batch Behaviour(s)	<input type="checkbox"/> Individual	<input type="checkbox"/> Batch	
Time-Constraint Behaviour	<input type="checkbox"/> Hard Real Time	<input type="checkbox"/> Soft Real Time	<input type="checkbox"/> None
Service End Point	<input type="checkbox"/> Provider	<input type="checkbox"/> Requestor	<input type="checkbox"/> Transcoder (both requests and provides)
Authentication/ Authorization Dependency	<input type="checkbox"/> Auth-Dependent	<input type="checkbox"/> Auth-Independent	
Protocol Binding(s) (only applies to service expression-based SUMs)	<input type="checkbox"/> Web Service <input type="checkbox"/> SOAP	<input type="checkbox"/> REST <input type="checkbox"/> HTTP	<input type="checkbox"/> Other
<b>To be determined by the e-Framework:</b>			
<b>Status</b>	<input type="checkbox"/> Approved	<input type="checkbox"/> Placeholder <input type="checkbox"/> Unapproved	<input type="checkbox"/> Superseded <input type="checkbox"/> Withdrawn
<b>Confidence Level</b>	<input type="checkbox"/> High	<input type="checkbox"/> Medium	<input type="checkbox"/> Low

## Description

There already is an established demand for grid enabled access to 2001 Census aggregate data and GEMS means to demonstrate how an existing data service infrastructure can be Grid enabled in a standards compliant and sustainable way by providing an OGSA-DAI grid service, well as a user friendly web interface, to the Census data set.

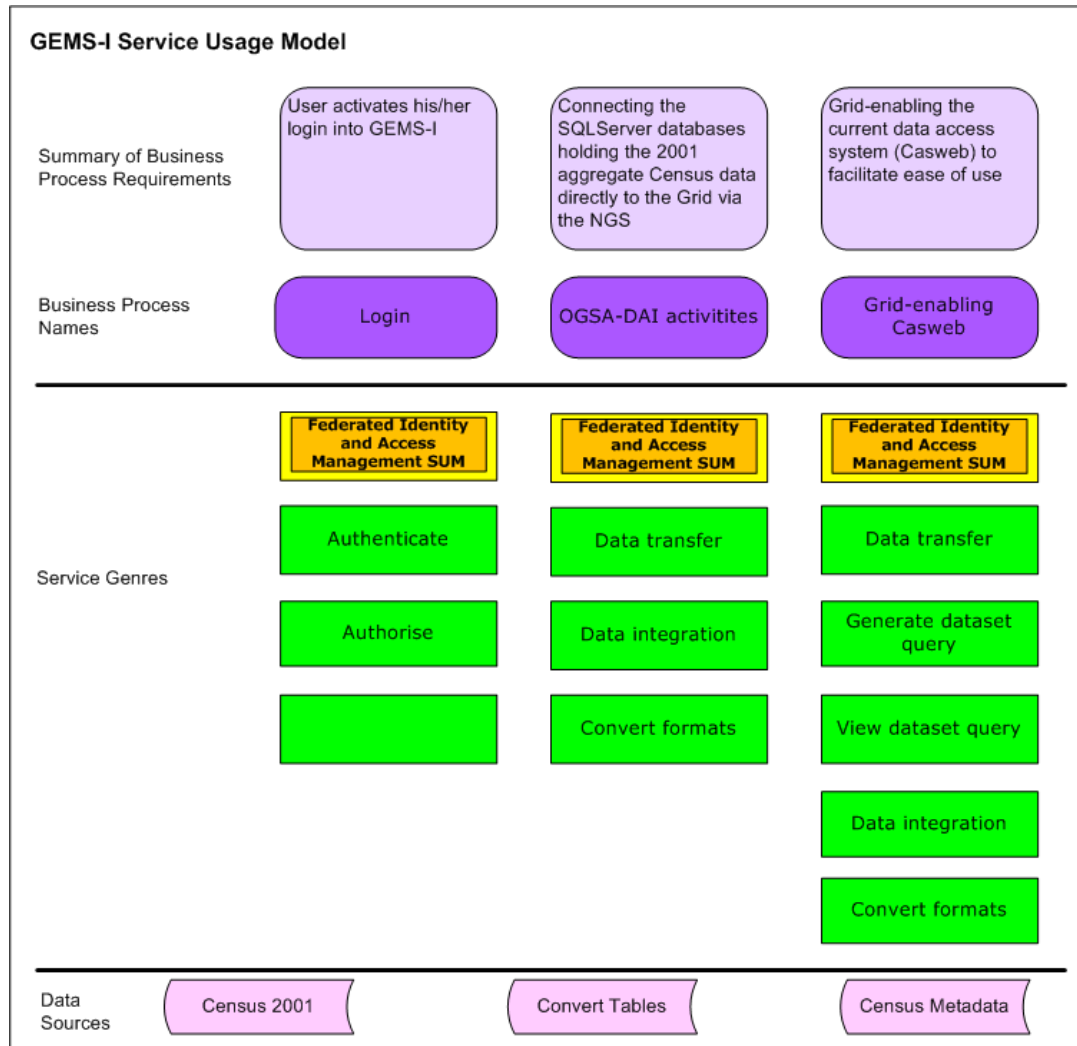
The purpose of the Grid Enabling MIMAS Services (GEMS) project is to provide Grid enabled access to the aggregate statistics from the 2001 Census via OGSA-DAI on the National Grid Service (NGS).

- Interface the Census database to NGS hosted OGSA-DAI instance
- Modify the existing the Casweb web interface to work with GEMS grid service & OGSA-DAI
- Implement a bridge between Athens/Shibboleth and grid middleware certificate authentication/authorization
- Offer a value added service for Census data clients through an OGSA-DAI grid service activity
- Integrate the existing ConvertGrid geographic data conversion service into the OGSA-DAI grid service activity

## Business Process Modelling

- Single sign-on login
- Connecting the MS SQLServer databases holding the 2001 Census aggregate data directly to the Grid via the NGS
- Grid enabling the current data access system (CASWEB)

## SUM Diagram



Visio® template for SUM diagram, revised 20070822  
 Template © Copyright 2007, e-Framework Partners

### Usage Scenarios [optional]

Paul is a researcher who needs to download some aggregate data compiled from the 2001 UK Census, before linking it to UK border data for visualisation purposes. In order to join the two datasets, he needs to have the aggregate data converted to a geography common to UK border data.

#### 1. Login

Paul logs in as normal through the University Athens entry portal by way of his browser configured to work with Paul's personal UK e-Science certificate.

#### 2. Browse through data

Paul navigates and selects relevant data through the powerful Casweb light client (modified to work with OGSA-DAI instead of on top of MS SQL Server).

### 3. Construct workflow

In order to simplify data integration, Paul decides he wants his data formatted to XML before having the data converted to a target geography.

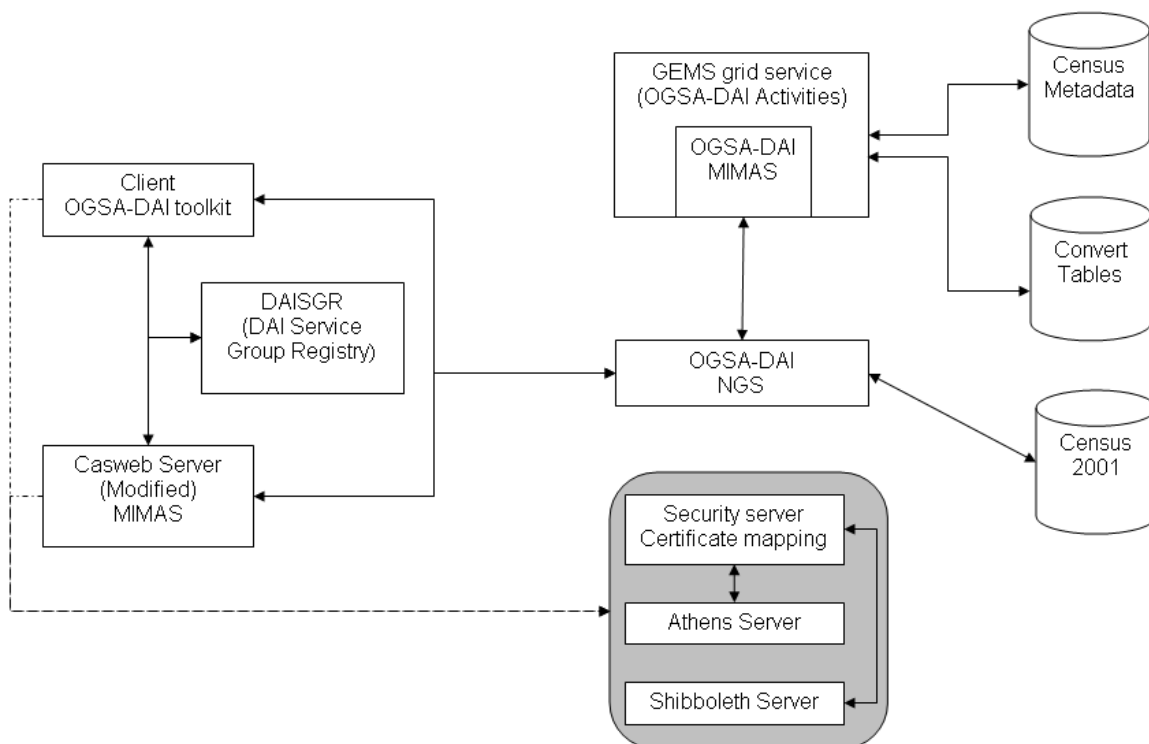
### 4. Download data

Paul downloads data onto his desktop, unaware of the fact his action triggered a series of business processes i.e. grid workflow: generation of SQL, querying of RDB through OGSA-DAI middle-ware, conversion to a user chosen data format stream, inclusion of metadata in the result stream, conversion of data results to a common target geography through the activation a ConvertGrid service.

### Functionality

- Transform query result into a variety of formats (CSV, HTML, etc...) by employing built-in or user uploaded XSL Transform scripts
- Integration of table metadata into query results
- Upload query results to a Grid/FTP server
- View SQL generated by user interface for further integration into an OGSA-DAI client
- Redirect query results to an grid service/OGSA-DAI activity for further processing
- Bulk upload query results to a user specified OGSA-DAI enabled database

### Structure & Arrangement



**Fig. 2 - GEMS architecture from [2]**

### **Security:**

- Athens authentication and authorisation of users accessing the Census 2001 dataset.
- User is authenticated to use NGS resources through his/her UK e-Science certificate deploying inside internet browser.

### **Grid middle-ware:**

- OGSA-DAI is used by the service for selecting and accessing the various datasets. Through its workflow engine, it also coordinates various grid services proposed by ConvertGrid (data format conversion, data geographic conversion) as well as a grid service provided by GEMS (inclusion of metadata into data result streams).
- The light client is in effect a version of the Casweb software - developed in Cold Fusion - modified to work with OGSA-DAI, and authenticate by means of X509 certificates.

### **Data Sources Used**

- Census 2001
- Convert Tables
- Census Metadata

### **Services Used**

#### List of Service Genres:

- Authenticate [it appears in the list of e-framework Service Genres – 9 May 08]
- Authorise [it appears in the list of e-framework Service Genres – 9 May 08]
- Data transfer
- Data integration
- Convert Formats [it appears in the list of e-framework Service Genres – 9 May 08]
- Generate dataset query
- View dataset query

### **References**

This document takes into account existing material, such as conference and journal papers. The most relevant sources are the following:

[1] <http://www.jisc.org.uk/publications/publications/socialsciencesresearchv1.aspx>

[2] <http://www.jisc.ac.uk/whatwedo/programmes/eresearch/gems1.aspx>

[3] <http://pascal.mvc.mcc.ac.uk:9080/gems1>

[4] <http://wiki.ngs.ac.uk/images/9/9d/NGS-Data-20070315-Cole-GEMS.pdf>

[5] M. Argüello, P. Ekin, A. Turner, S. Peters, P. Townend, M. Fraser, P. Halfpenny, R. Procter, A. Voss, and M. Jirotko: Highlighting e-Infrastructure patterns in Grid-based e-Social Science applications. Accepted for Regular Session at UK e-Science All Hands Meeting 2008 (AHM 2008), Edinburgh, UK, September 2008



This SUM is licensed under:  
Creative Commons Attribution-NonCommercial-ShareAlike 2.5 licence  
<http://creativecommons.org/licenses/by-nc-sa/2.5/au/>

### **Submitting the Service Usage Model Description**

For additional guidance in preparing the Service Usage Model description, refer to [Guidelines for Submitting a Service Usage Model to the e-Framework](#)<sup>2</sup> and the technical definitions of the [Service Usage Model Description Elements](#).<sup>3</sup> For further assistance, contact the e-Framework editor at: [editor@e-framework.org](mailto:editor@e-framework.org)

Prior to submitting the description, please read the e-Framework [Intellectual Property Rights statement](#).<sup>4</sup> Also add your information to the copyright statements in the footer of this template.

By submitting this document, you agree to contribute this document under the [Creative Commons Attribution-NonCommercial-ShareAlike 2.5](#)<sup>5</sup> licence.

When you have completed the Service Usage Model description, go to the [Submit SUMs](#)<sup>6</sup> page at [www.e-framework.org](http://www.e-framework.org). Click on “Upload your submission” and follow the directions.

---

<sup>2</sup> Guidance for using this template: <http://www.e-framework.org/SUMs/SubmitSUMs/tabid/715/Default.aspx>

<sup>3</sup> Service Usage Model Element Definitions: <http://www.e-framework.org/Services/SUM/SUMElements/tabid/745/Default.aspx>

<sup>4</sup> Intellectual Property Rights Statement: <http://www.e-framework.org/About/Policies/tabid/611/Default.aspx>

<sup>5</sup> Creative Commons: <http://creativecommons.org/licenses/by-nc-sa/2.5/au/>

<sup>6</sup> Submit Service Usage Model template: <http://www.e-framework.org/SUMs/SubmitSUMs/tabid/715/Default.aspx>