Shibboleth Enablement of the VOTES Grid Portal

User Documentation

The Virtual Organisations for Trials and Epidemiological Studies (VOTES) project (www.nesc.ac.uk/hub/project/votes) is a three year MRC funded project to explore how Grid technologies can be used to support clinical trials and epidemiological studies. Ensuring that the right people see the right data sets for the right purpose is crucial in this domain. Fine grained security is therefore essential.

This user guide outlines how Shibboleth has been applied to provide user oriented access to and usage of this portal. The prototype portal here uses a variety of representative training data sets and actual software used in the NHS including:

- Scottish Care Information (SCI) Store (http://www.show.scot.nhs.uk/sci/products/store/SCIStore_Product_Description) which provides an aggregated data resource used by most hospitals and trusts across Scotland;
- General Practitioners Administration System for Scotland (GPASS), (http://www.show.scot.nhs.uk/gpass/) which is used by 85% of GPs across Scotland;
- a consent database which allows patients to indicate whether they wish their data to be made available or not, and if so for a specific clinical trial or for a given research area, e.g. cancer research only;
- a representative clinical trials repository at the Robertson Centre for Biostatistics at the University of Glasgow.

This infrastructure is hosted on a test bed at the National e-Science Centre at the University of Glasgow. With this infrastructure we will show how different roles can get access to different clinical data sets. Specifically, there are two roles designed to show how Shibboleth based access and authorization is supported, namely: *Investigator* and *Nurse*. Users with either these roles are able to access to the same sets of databases. However the "Investigator" has access to all the parameters and data fields in the databases, whilst the "Nurse" only has access to limited parameters that cannot be used to identify individual patients, e.g. information that could only be used for generic statistical purposes. Authorization decisions themselves are made by the portal when attributes are delivered by Shibboleth from an Identity Provider (here we use the identity provider for the DyVOSE project).

The VOTES Grid portal and the associated Grid services contained within it are built using the Globus Toolkit 4 (www.globus.org/toolkit/docs/4.0) and OGSA-DAI (www.ogsadai.org.uk) technology, and uses the GridSphere Portal Framework (www.gridsphere.org) to make the front-end as user-friendly as possible to clinical trials researchers wishing to query distributed clinical data sets.

In this user guide we have established three username/password pairs to explore the VOTES functionality and how Shibboleth works.

Username	Password	Role
guest1	guest1	Nurse
guest2	guest2	Nurse
guest3	guest3	Investigator

To test the Shibboleth enabling of the VOTES Grid Portal, please follow the below instruction:

1. Open a browser and enter the URL: http://labpc-2.nesc.gla.ac.uk/gridsphere[1]. The page will redirect user to the WAYF (Where Are You From) page to select the institution that you come from.
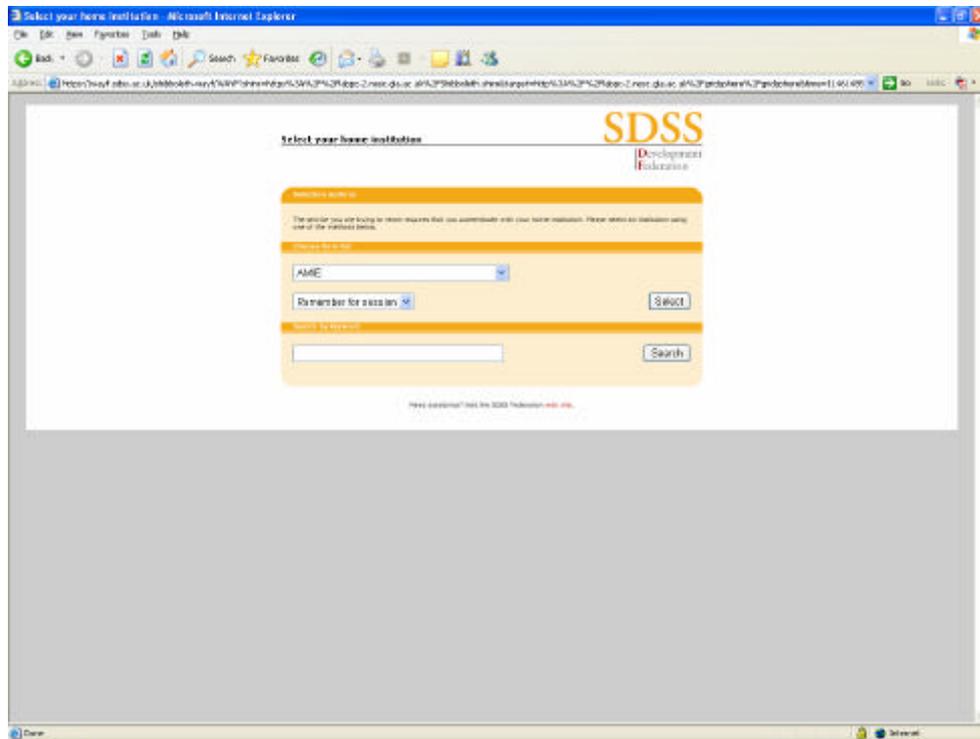

2. Select the "DyVOSE Project" from the drop down box, and click "Select" button.

---

[1] Note: It was not possible to have the same URL as for the GridBLAST or the search/sort service due to incompatibilities between GT3/GT4, GridSphere and the way in which Shibboleth attributes are delivered and used to make authorization decisions. More information on this will be given in the DyVOSE final report.
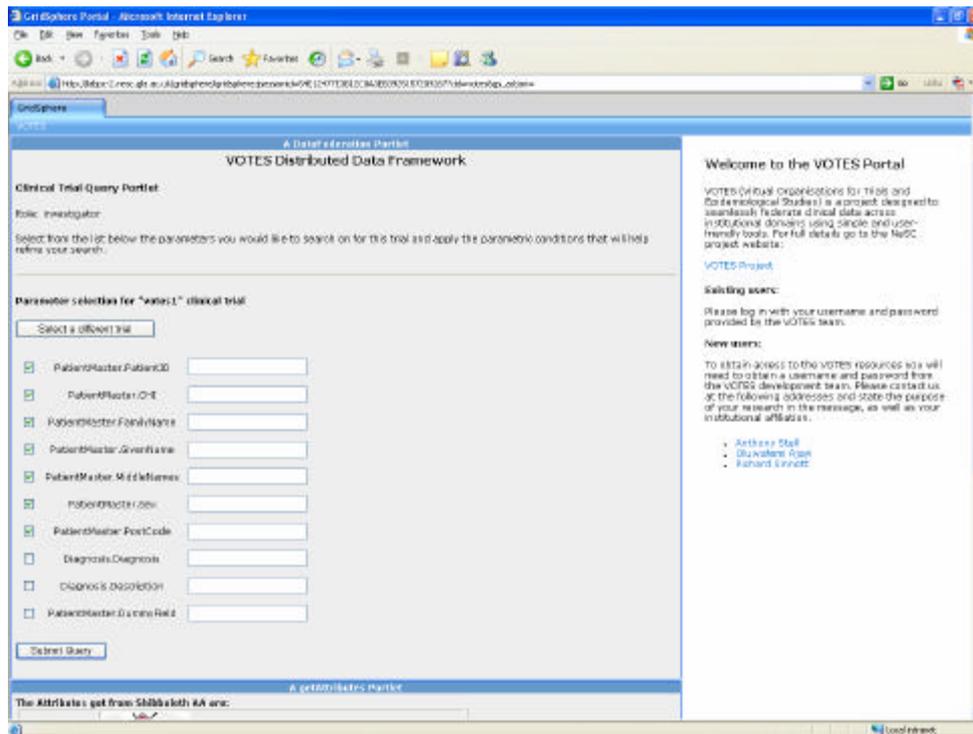
3. First, let's log in as the investigator (guest3/guest3). After authenticating, the VOTES portlet page is displayed as shown here.
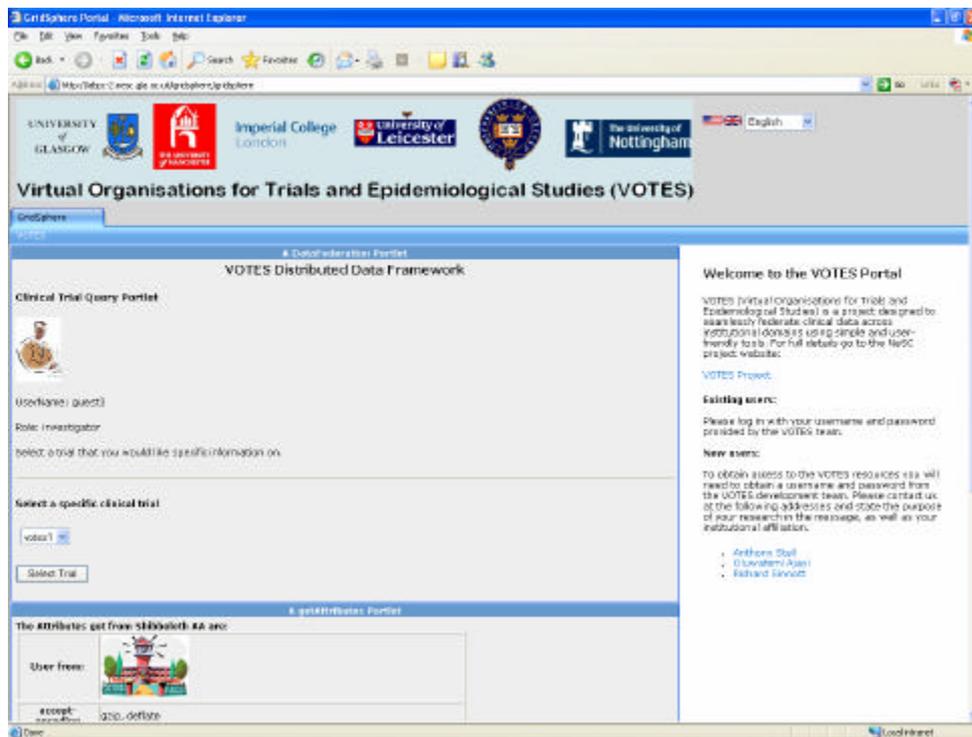

4. The attributes from the Identity Provider (DyVOSE Project) are displayed here in a separated portlet. As you can see, user guest3 is from "area3" with role of "investigator": area3 is a string which can be used to represent the institution the user is from. We note that this information is not used to make authorization decisions within the portal (only the role is important in this example), however it is quite possible to include this and other information returned to make authorization decisions. We also note that other roles are also returned such as *studentteam1* and *studentteam2* since users (guest1,guest2,guest3) are also used to

show how these roles can access and use the GridBLAST service and Search/Sort Grid service (or not as the case may be).

5. Select a specific clinical trial from the drop down box, e.g. *votes1*. Then click the "Submit Trial" button. Here you will see all the parameters and data files that someone with this role is allowed to see. Select the data fields that you would like to see: some of these require only a check, others allow for strings to be searched against, e.g. Diagnosis might look for %cancer%.

6. After clicking the "Submit Query" button, the results are displayed. As an "Investigator" you can see all the information contained in the databases as this role has a high privilege.
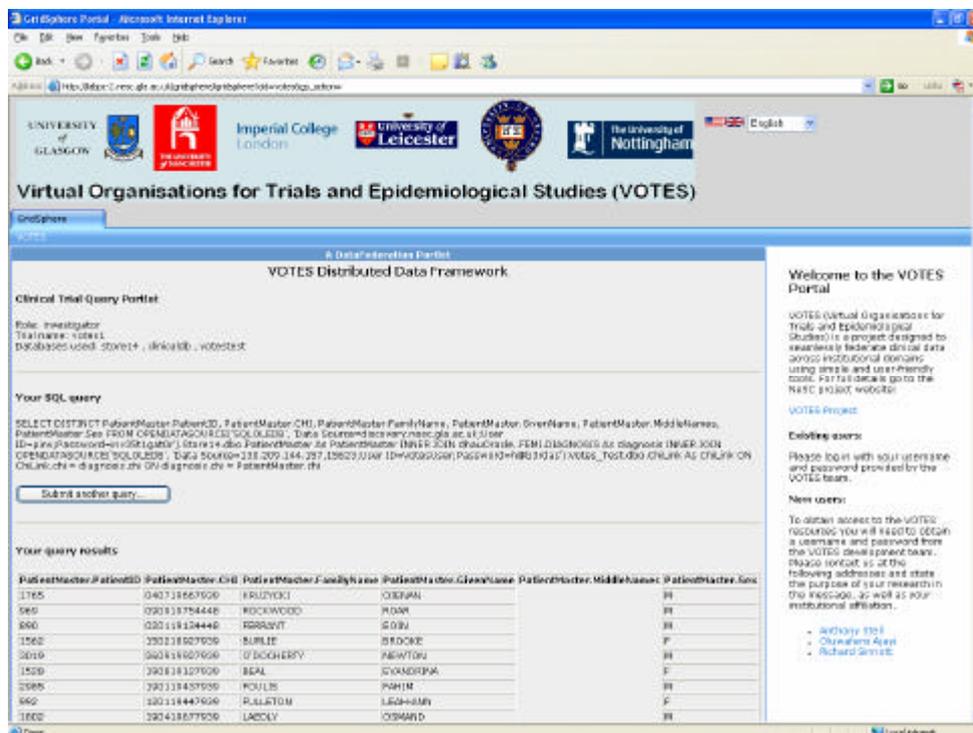
We note that if another trial is selected, e.g. *votes2* or *rcb1*, then the user might have a different set of data (checkboxes) that they can select, i.e. each trial will have their own data sets which different roles are restricted to.

We also note that the federated queries are joined on a unique identifier: the Community Health Index (CHI) number. This is similar to the NHS number in England and Wales.

7. Now we log out (or simply close the browser). In a new browser type in the URL http://labpc-2.nesc.gla.ac.uk/gridsphere. As before enter DyVOSE Project as the Identity Provider, but this time log in as a "nurse" (guest1/guest1).
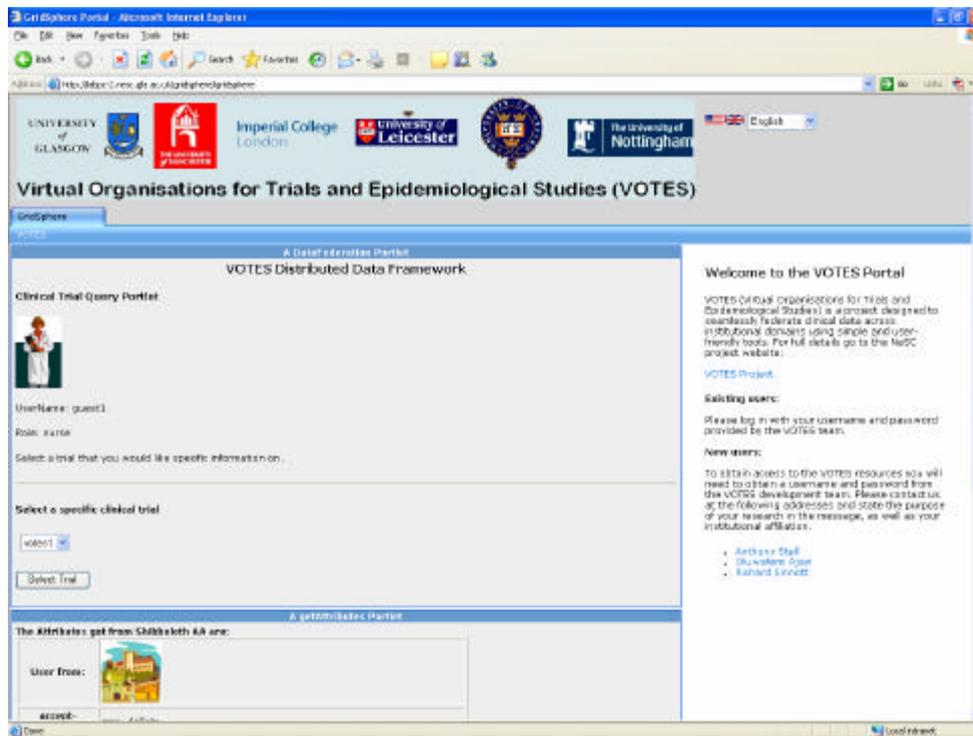
8. With the role of "Nurse", it is now only possible to select from a limited set of data fields. The parameters that include patient specific information are not available to be queried with someone with this role.
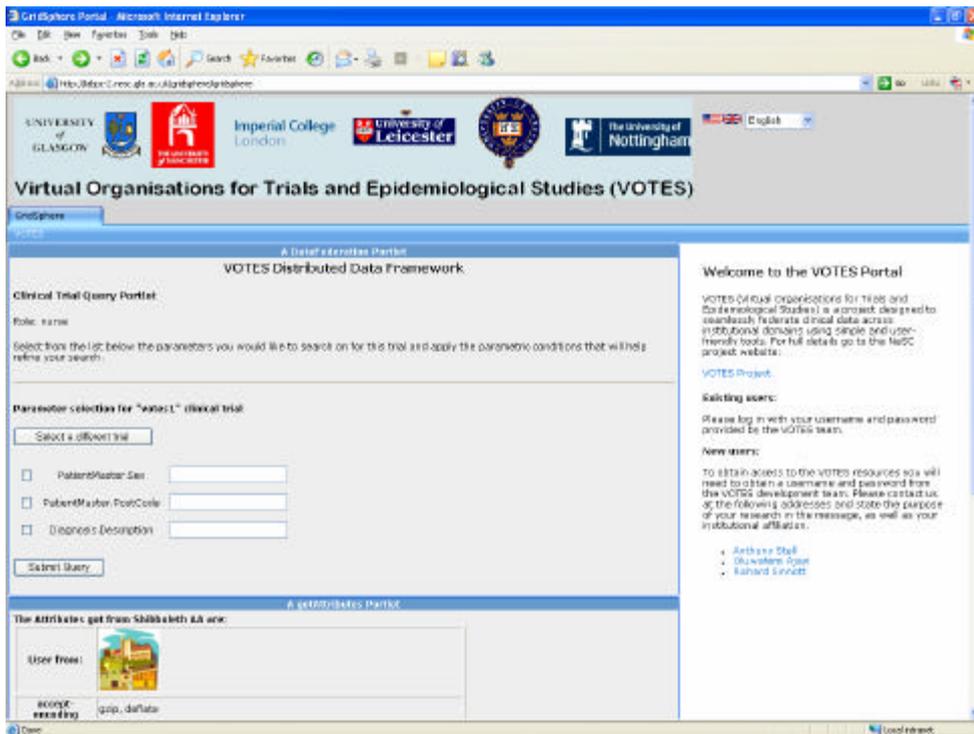
9. Select some fields and click submit query. This is the result page for a "Nurse" user after click the "Submit Query" button.
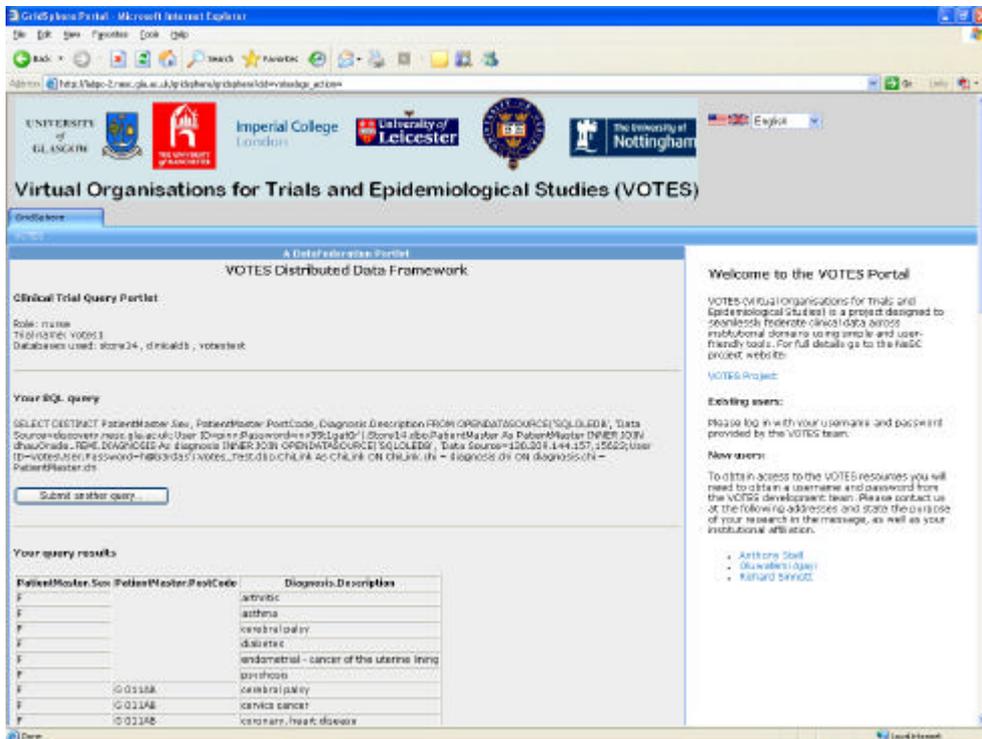
The Shibboleth enabling of the VOTES portal shows how non-computer specialist researchers are likely to access the power of grid computing in an easy and secure way in

the future. Through this infrastructure and use of Shibboleth, access to fine grained (authorised) Grid services is made as easy as using a normal web site and is completely transparent to end users.

We note that the prototype give here is still under a state of development and enhancement. For example, there are 180 different roles identified within the NHS in Greater Glasgow NGS alone with over 33,000 employees. Exploring how best to establish and enforce fine grained authorisation infrastructures with such a multitude of roles is being explored within the on-going VOTES work.

For further information on VOTES or application of Shibboleth technologies, please contact:

- Prof Richard Sinnott (ros@dcs.gla.ac.uk)
- Anthony Stell (ajstell@dcs.gla.ac.uk)
- Oluwafemi Ajayi (ajayio@dcs.gla.ac.uk)