PROCEEDINGS OF SCIENCE

"Social" Science Gateways to e-Infrastructures

Carla Carrubba
Division of Catania of the Italian National Institute of Nuclear Physics
Via S. Sofia, 64 - 95123, Catania, Italy
E-mail: carla.carrubba@ct.infn.it

Marco Fargetta[1]
Division of Catania of the Italian National Institute of Nuclear Physics
ia S. Sofia, 64 - 95123, Catania, Italy
E-mail: marco.fargetta@ct.infn.it

Riccardo Rotondo
Consortium GARR
Via dei Tizii, 6 – 00185 Rome, Italy
E-mail: riccardo.rotondo@garr.it

Roberto Barbera
Department of Physics and Astronomy of the University of Catania and
INFN – National Institute of Nuclear Physics, Division of Catania
Viale A. Doria, 64 – 95123 Catania, Italy
E-mail: roberto.barbera@ct.infn.it

Notwithstanding the existence of a large and highly-performant computing infrastructure, and the effort made to improve the quality of these e-Infrastructures, the user base has remained quite small relative to the potential number of users. On the other hand Social Networks (SNs) have become extremely popular, thanks in large part to their intuitive and easy-to-use interfaces. Recognising the potential to improve the overall usability of e-Infrastructures, recently, a significant effort has been made to develop e-Collaboration environments allowing scientists to access remote computing facilities, referred as "Science Gateways" (SG).

In this paper we present the work done to connect SGs to an existing SN. The main aim of this integration is to allow users to access the tools provided by the SG directly from the SN and to attract a larger audience to these new technologies. The interoperability has been achieved at several levels, from the authentication/authorisation up to the user interface (UI), making the SG accessible from SN user pages.

The International Symposium on Grids and Clouds (ISGC) 2013
March 17-22, 2013
Academia Sinica, Taipei, Taiwan

[1] Speaker

PoS(ISGC 2013)031

1. Introduction

According to the bell-shape curve of the path to technology uptake[1], reported in Figure 1, innovators are responsible for the birth of new technologies but their real spread all over the world takes place only when a chasm happens that convince the vast majority of potential users to adopt them in their day-by-day live.

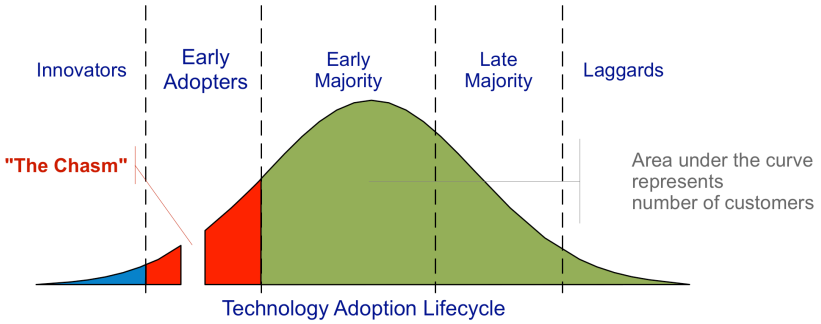

Figure 1 - Roger Bell's curve of the path to technology uptake.

Very often, these chasms coincide with either the development of new models providing easier access to the core technology or with the rapid expansion of the underlying infrastructures these new technologies build upon. Historically, two well recognised cases providing evidence of the above curve were: the World Wide Web and the research and education networks. The former get the chasm when the first graphic web browsers came along around the mid-nineties and the latter when the combined effects of reduction in bandwidth costs and central funding led to the creation of multi-tens-of-Gb/s continental backbones such as the GÉANT network [2] in Europe, largely funded both by the European Commission and the national governments.

In the Information Technology acceptance model [3], the “Perceived Usefulness” and the “Perceived Ease of Use” play key roles for their wider adoption. Two good examples demonstrating their importance in the last years are the tremendous success of smartphones (estimated more than 25% of the about 4 billion mobile phones existing in the world) and the huge uptake of Social Networks. The smartphone adoption got a big impulse after the introduction of a straightforward touch interface by Apple Inc. ™, and the Social Networks have exploited the latest web technologies to made their interface easy. Current reports estimate that about 1 billion people have accounts on the existing Social Networks and that web-based social networking accounts for largely more than 10-15% of the total time spent online in the whole world [26].

In this very rapidly changing world where borders are losing significance, scientific research has not been left too much behind and the more computationally intensive science has exploited the opportunity provided by the connection of billions of computers around the word, creating the e-Science, which has emerged as the most important paradigm of the first decade

of the 21st Century. The development in the last decade focused in creating e-Infrastructure for the e-Science but the total number of users has not met the expectation, considering the size of the involved communities.

The European Commission has promoted a study [4] to investigate the reasons for this limited adoption. Among the others, the study has identified the complexity of the Grid Security Infrastructure (i.e. based on a Public Key Infrastructure) for non IT expert users, the little adoption of standards to let different middleware be interoperable among each other, and the lack of general frameworks to build easily customisable and easy to use high-level user interfaces.

In the recent past, interesting developments have been independently carried out by the Grid community with the Science Gateways [5] [6] and by the National Research and Education Networks with the Identity Federations [7] to ease, from one side, the access and use of Grid infrastructures and, from the other side, to increase the number of users authorised to access network-based services.

In this paper, we present a recent work attempting to integrate Science Gateways and Social Networks, both at the level of cross-authentication and user interface, with the ultimate goal of widening as much as possible the potential number of e-Infrastructure users, both on desktop/laptop and mobile appliances.

The paper is organised as follows: section 2 introduces the framework adopted for the authentication and authorisation in the developed Science Gateway, section 3 reports on the integration done with Facebook, and section 4 shows concrete use cases. Session 5 draws the conclusions.

2. Authentication and Authorisation framework for Science Gateway

Science Gateway (SG) implementations usually consist of a mixing of different technologies integrated to create a uniform environment that allows the mash-up of many services.

The Science Gateways we will refer in this paper is based on a specific framework, named the Catania Science Gateway Framework [8][9]. These Science Gateways are based on Liferay [10], an award winning enterprise ready portal framework fully Java based and supporting many of the standards used in modern web-based applications such as, to mention the most used, JSR 286 [11] (also known as “portlet 2.0”).

The authentication adopted for the Science Gateways uses version 2 of the SAML standard [12][13] and its Shibboleth [14] and SimpleSAMLphp [15] implementations. Additionally, a Liferay plugin has been developed to make the portal able to authenticate users with the Shibboleth token.

People willing to access the Science Gateway need first to apply providing personal information such as their organization and email address(es). These integrate the data coming from the Shibboleth tokens and used for the authorisation.

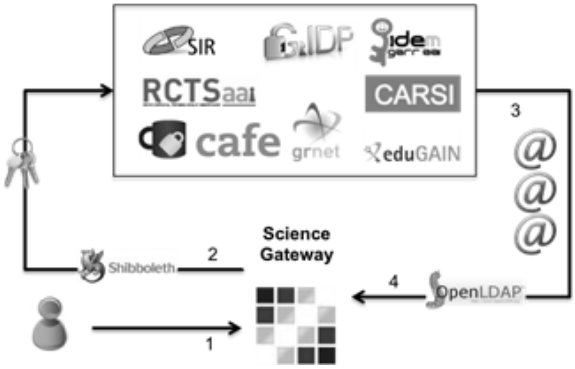

Figure 2 - Authentication and authorisation framework.

Among the recognised Federations, the Grid IDentity Pool (GrIDP) [17] has been created at INFN Catania (a division of the National Institute for Nuclear Physics) to act as a catch-all Identity Federation. GrIDP aims at gathering all the IdPs that do not already belong to any supported federations and IdPs for home-less users. This is useful whenever it is necessary to authenticate the so-called "citizen scientist" (i.e., people belonging to the general public) and let users access the e-Infrastructure for dissemination and self-learning purposes.

Once a Science Gateway administrator authorises the users, these are able to sign in. Figure 2 sketches this two-step procedure. On the right side there is the authentication phase demanded to the IdP user belong to. Authorisation, operated by our LDAP (Lightweight Directory Access Protocol) is pictured on the left side where is underlined the exchange or mail address(es) between the IdP and the LDAP to allow users in, as explained more in details in this section.

In step 1, when users ask to be signed in, they are automatically redirected to the Identity Federation discovery service, shown in Figure 3. Once selected the federation, users are redirected to the Identity Provider discovery service, shown in Figure 4 for the GrIDP federation, where select the belonging Identity Provider.

Figure 3 - Identity Federation Discovery Service.

Figure 4 - Identity Provider Discovery Service

In step 2, upon selection of the IdP, the users are forwarded to the corresponding login page and, if the authentication is successful, they are redirected back to the Science Gateway.

In step 3, the user is authenticated and starts the authorisation check. The Science Gateway gets the data included in the Shibboleth token providing user identity and, among the available attributes, extracts the email address used as unique identifier inside the Science Gateway. The developed Liferay plug-in matches the user email address(es) against an LDAP server connected with the Science Gateway. The LDAP is populated with the information provided in the registration phase and integrated with the authoritative groups the user is linked to.

In step 4, if the performed match is successful, the SAML token is converted in a Science Gateway user having the related groups and privileges associated with the authorisation level inside the given Science Gateway. Otherwise, a page warning about the missed information will appear before to forward the user to the registration form.

Using the email address, which has a many-to-one relation with the user, instead of a one-to-one identifier has an important side effect: a single account of a given Science Gateway can be associated with many different credentials corresponding to different IdPs but identifying the same person. In fact, for the authorisation to be granted, it is enough to match at least one email address coming from an IdP.

3. Social Network Integration

Since many Social Networks provide their authentication framework to external resources, Science Gateways can easily integrate these Social Networks as special credential providers for

the authentication. Actually, the integration is not part of the Science Gateway but a special Identity Provider called Social Networks' Bridge Identity Provider [18] has been created using SimpleSAMLphp and included in the federation. This IdP acts as a bridge towards the Social Networks' credential providers so the Science Gateway uses only a single mechanism based on SAML for user authentication and leaves the communication with the Social Networks to external tools. Figure 5 shows the Social Network selection by which users can chose the credential provider they wish to use for their authentication.

As mentioned above, the user authorisation is performed using an LDAP registry. To be registered in the LDAP registry, users must be identified and this is not easy for identities coming from Social Networks. Credentials generated by Social Networks can be associated with a real user and authorised to access the Science Gateway only if at least one of the two following conditions are satisfied:

- the user want associate credential from Social Networks with valid credential released by an official identity provider of a university or research organization;
- the user is known in person by the administrator(s) of the Science Gateway.

These limitations reduce the security risk of blindly accepting social credentials. The theft of credentials is still a risk for this approach but it is present also in the case of credentials released by institutional identity providers.

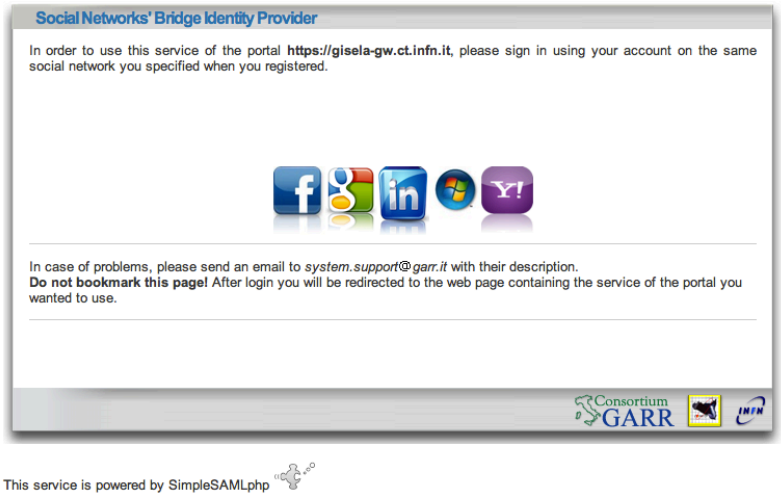

Figure 5 - The Social Network selection of the Social Networks' Bridge Identity Provider

Nowadays, Facebook is the most used Social Network with more than 1 billion users in the world, with the majority in the age range 23-49. We can suppose that many researchers which could use e-Infrastructures are Facebook users.

Furthermore, Facebook allows to create both user and community pages and to integrate external applications in them through a rather complete set of API. The possibility to extend the

basic functionalities of its user interface with those provided by external applications make Facebook a very powerful platform to share contents and applications.

For all the above reasons, and not only, we targeted to integrate our Science Gateways with Facebook in order to increase the number of users. Additionally, Liferay provides a complete set of API to interact with Facebook.

Liferay provides means to integrate either the Science Gateway as a whole or its single portlets, as applications in Facebook. We decide to start integrating the whole portal to address the initial need of expand the community.

Users need special authorisations to be able to deploy applications into Facebook. These users, referred in Facebook as developers, must be identified through a special procedure which requires the use of trusted mechanisms. It is important to underline that Facebook application developers do not need to be developers of the Science Gateway, they simply need to communicate several information to the administrators of the Liferay server where the Science Gateway is running. This convenient decoupling brings two different benefits: it allows Science Gateway administrators to keep full control on the decision to open or not their services to the huge community of Social Network users; on the other hand, Social Network application developers can concentrate on the social part without the need to know the Science Gateway implementation in detail. Moreover, Science Gateway developers can work on their portals without worrying about how they will be exposed to the large audience of Social Network users.

Figure 6 depicts the procedure followed in order to integrate a Science Gateway in Facebook as an application.

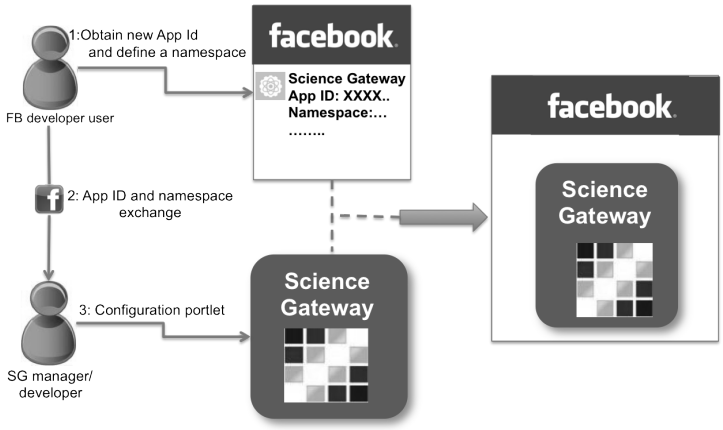

Figure 6 - Procedure to integrate a Science Gateway in Facebook

The Facebook developer requests, by filling a web form, an App ID and defines a namespace for the new Facebook application as first step. During the creation of the new application, the Facebook application developer can also create a community page that would contain the application launcher and could be used by all members of the community as unique access point (see more in the next section).

In the second step, developers send the App ID and the namespace of the application to the administrator of the Science Gateway so that they, as third step, can configure accordingly the whole portal or just a single component of it.

When the three steps are successfully completed, the Science Gateway is exposed as a Facebook application and can be accessed by the members of the Social Network. As shown in Figure 7, Facebook users can access the new application in two different, but equivalent, ways.

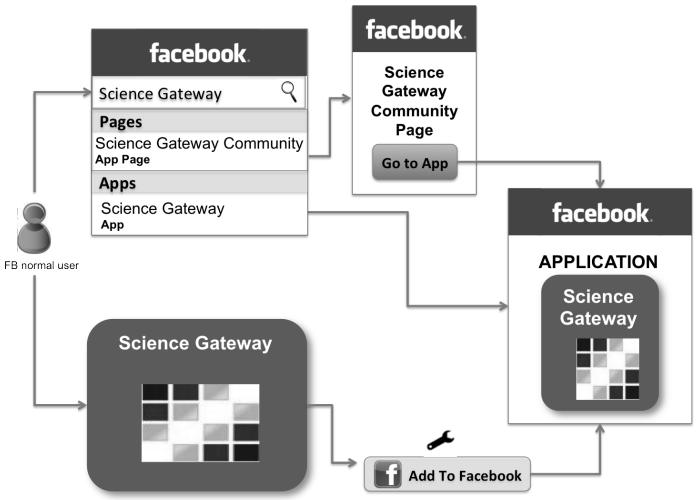

Figure 7 - Ways to access a Science Gateway from within Facebook

In the first way (upper part of the figure), they can use the Facebook search box to look for the Science Gateway community page they are interested in, open it, and then click on "Go to App".

In the second way (lower part of the figure), users visiting the home page of the Science Gateway website click on the spanner icon and select "Add To Facebook".

4. Use Cases

Making Science Gateways available as Facebook applications allows users to have a unique access to Social Networks and Science Gateways, meshing up e-socialization, e-collaboration and e-science services. Indeed, users can exploit very heterogeneous services exposed through the same interface and they do not need any more to switch among different websites to talk to friends, to collaborate with colleagues, and to run jobs or manage data on the e-Infrastructures they are authorised to use.

As we mentioned in the previous section, a Facebook application is available in a community page. Therefore, we have created a separate community page for each different Science Gateway we have implemented[2]. In each of them, the visitor can access the

[2] Currently, we have created 13 Science Gateways and likewise community pages. An updated list is available at http://www.catania-science-gateways.it/science-gateways

corresponding Science Gateway from within the Facebook environment just clicking on the "Go to App" button.

Creating a community page on Facebook has several benefits for the Science Gateway that should be considered. Social Networks allow an easier and faster way to spread information. More people can get aware of the Science Gateway activities and be involved in case of common interests for the same scientific domain. As a consequence, the community is not only an access point for the Science Gateway services but also a virtual meeting point where people can interact and communicate with all other members of the community, exchanging ideas and information about the scientific research carried out by means of the Science Gateway.

In the remaining part of this section we will describe in more detail two of the Science Gateways developed, namely the agINFRA Science Gateway [19] and the GARR Science Gateway [20].

4.1 agINFRA Science Gateway Community

agINFRA [22] is a project funded by the European Commission within its Seventh Framework Program. It started in October 2011 and will last three years. agINFRA aims at creating a data infrastructure on top of the European Grid Infrastructure for the agricultural science community.

The agINFRA Science Gateway meshes-up Grid Computing and Cloud Computing enabled applications and lets users access them in an easy way. The web page of agINFRA Science Gateway Community on Facebook appears as shown in Figure 8.

Figure 8 - The agINFRA Science Gateway Community page on Facebook.

It contains general information on the community as well as on the agINFRA project and, well in evidence on the right side, the button “Go to App” to launch the agINFRA Science

Gateway as a Facebook application. Clicking on the "Go to App" button, users can then access the agINFRA Science Gateway.

Such tight integration allows users to work on the Science Gateway and monitor, at the same time, all their social activities such as notifications, messages, requests from colleagues and friends, etc.

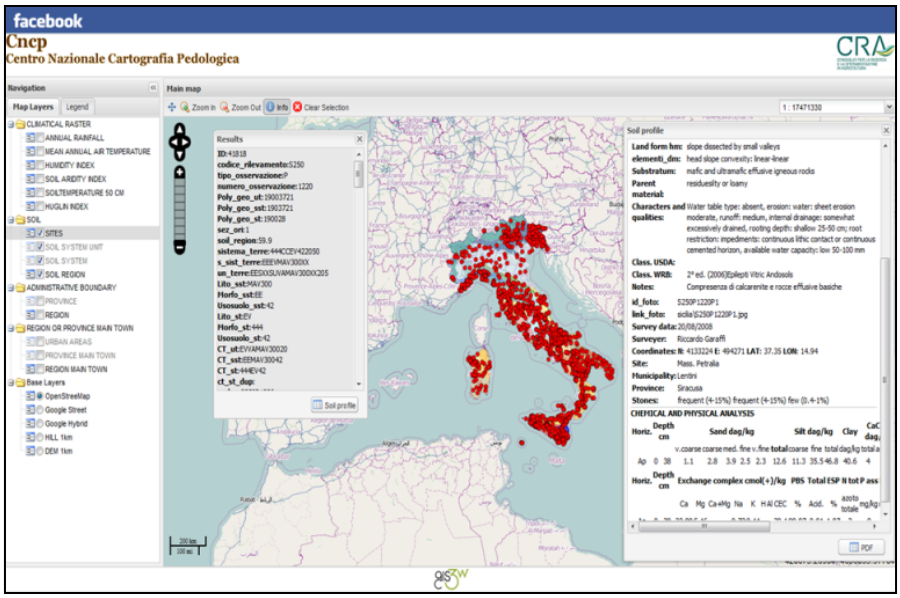

Figure 9 - The Italian Soil Information System service of the agINFRA Science Gateway integrated as a Facebook application

One of the main applications in agINFRA is the Italian Soil Information System (ISIS) [22] developed at the Agrobiology and Pedology Research Centre [23] of the Italian Agricultural Research Council [24]. ISIS is a WebGIS-enabled Cloud Computing based application allowing online consultation of the database of the Italian soils.

Figure 9 shows the ISIS application integrated in the Facebook web environment. The application shows the result of a query to the ISIS geo-databases to retrieve quantitative information about the soil in one of the registered sites.

4.1.1 GARR Science Gateway Community

GARR [25] is the Italian Research and Education Network (NREN). Consortium members can rely on a valid portal as the GARR Science Gateway [20] to access high level services regarding GRID and CLOUD e-Infrastructure thanks to the easy tools offered by the Science Gateway to its community.

An example of the Science Gateway functionalities available through Facebook is shown in Figure 10. A user, waiting for some analysis to terminate, can communicate with others community members for other activities. It is easier to monitor the status of the GRID

submissions and download them without the needs of moving among different web pages and sites.

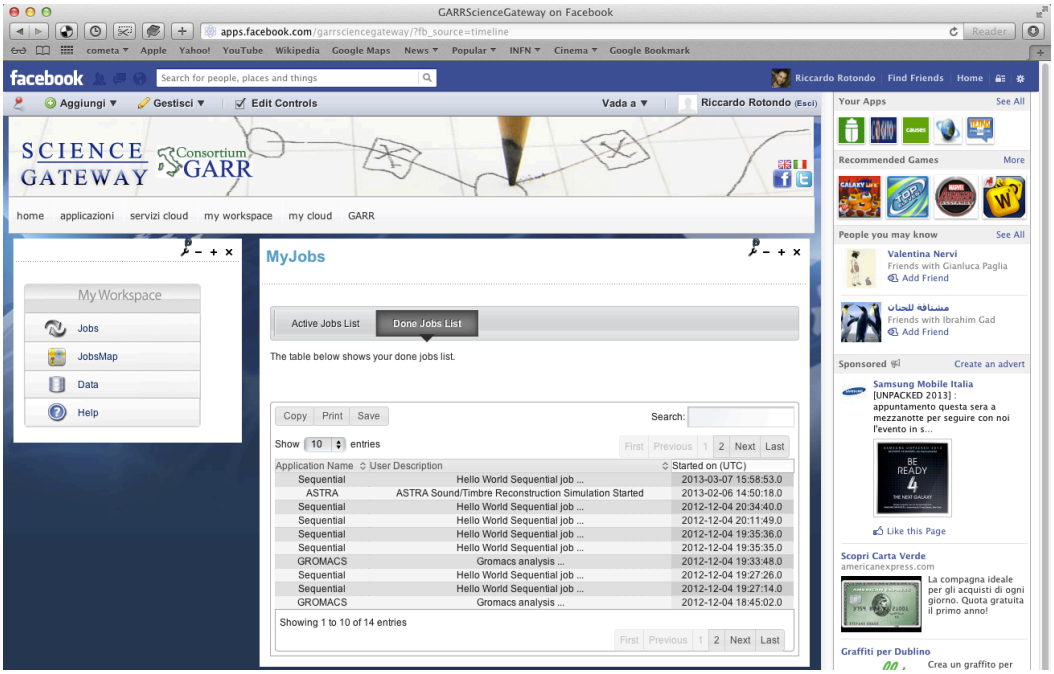

Figure 10 - Monitoring Job submission through GARR Science Gateway Community Page on Facebook

5. Summary and Conclusion

This paper reports on the techniques and the tools we have implemented to integrate Science Gateways with Social Networks. It also shows two use cases: the EU funded agINFRA project and the GARR Science Gateway.

The work starts from the experience made with the Social Networks' Bridge Identity Provider. This aimed at authenticating users on Science Gateways with the same credentials they have with the most known Social Networks.

The integration of Science Gateways with Social Networks at the level of user interface and collaborative environment represents a milestone in the path to realise our vision: make e-Infrastructure easy to be accessed and used by the largest possible number of potential users and tightly integrated with social services.

Nevertheless, this integration activity is still at an early stage. The mesh up of Social Network services and Science Gateway applications is not tight enough and needs extra work to create a unique e-Collaboration environment. As a result, the community pages have mainly been used to spread information. Actually, the access with social credential has been granted to selected users for test only.

Among the large number of Social Networks available on the web, Facebook was chosen as a primary test-bed because it is the most used in the world and provides a rich set of API and tools that makes it easily bound with other web services. The next step is to extend the Science Gateway to other popular Social Networks such as Google+, LinkedIn, etc., which are already integrated in the Social Networks' Bridge Identity Provider.

Another item in our to-do list is to provide the possibility to create services that could interact directly with the Social Network community page of a given Science Gateway writing messages and/or notifications in a human unattended mode. This will dramatically expand the concepts of job logging and bookkeeping, and job perusal making the day-by-day research work of a scientist a truly social experience for him/her and his/her colleagues.

Acknowledgment

Authors wish to thank Dr. Bruce Becker from Meraka Institute (South Africa) for the long and fruitful discussions that triggered the work described in this paper.

References

[1] E. M. Rogers, "Diffusion of innovations", Glencoe Free Press, 1962.

[2] The GEANT network website: www.geant.net.

[3] F. D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology", MIS Quarterly 13 (3), pp. 319-340, 1989.

[4] eResearch2020 Final Report, "The Role of e-Infrastructures in the Creation of Global Virtual Research Communities", www.eresearch2020.eu/eResearch%20Brochure%20EN.pdf.

[5] N. Wilkins-Diehr, Special Issue: "Science Gateways - Common Community Interfaces to Grid Resources", Concurrency and Computation: Practice and Experience 19(6), pp. 743-749, 2007.

[6] N. Wilkins-Diehr, D. Gannon, G. Klimeck, S. Oster, S. Pamidighantam, "TeraGrid Science Gateways and Their Impact on Science", IEEE Computer 41(11), pp. 32-41, 2008.

[7] The REFEDS website: https://refeds.org.

[8] V. Ardizzone et al., "Science Gateways for Semantic-Web-Based Life Science Applications" in "HealthGrid Applications and Technologies Meet Science Gateways for Life Sciences", Studies in Health Technology and Informatics, volume 175, pp. 119-130, 2012.

[9] V. Ardizzone et al., "The DECIDE Science Gateway", Journal of Grid Computing, pp. 1-19,2012.

[10] The Liferay Portal Framework website: www.liferay.com.

[11] The JSR 286 Standard Specifications: http://download.oracle.com/otndocs/jcp/portlet-2.0-fr-oth-JSpec.

[12] S. Cantor et al., "Assertion and Protocols for the OASIS Security Assertion Markup Language (SAML) V2.0", Tech. Report, OASIS SSTC, http://docs.oasis-open.org/security/saml/v2.0/saml-core-2.0-os.pdf, Document ID: samlcore-2.0-os, 2005.

[13] The SAML specification website: http://saml.xml.org.

[14] The Shibboleth System website: http://shibboleth.internet2.edu.

[15] The SimpleSAMLphp System website: http://simplesamlphp.org.

[16] R. Barbera, M. Fargetta and R. Rotondo, "A Simplified Access to Grid Resources by Science Gateways", Proceedings of Science (ISGC 2011 & OGF 31), pp. 23-33, 2011.

[17] The Grid IDentity Pool Federation: http://gridp.ct.infn.it.

[18] The Social Networks' Bridge Identity Provider website: https://idpsocial.ct.infn.it.

[19] The agINFRA Science Gateway website: http://aginfra-sg.ct.infn.it.

[20] The GARR Science Gateway website: http://sgw.garr.it.

[21] The agINFRA Project website: www.aginfra.eu.

[22] E. A. C. Costantini and G. L'Abate, "A Soil Aridity Index to assess desertification risk for Italy", Advances in Geoecology 40, pp. 231-242, 2009.

[23] The Agrobiology and Pedology Research Centre website: http://abp.entecra.it/index.php?lang=en.

[24] The Italian Agricultural Research Council website: http://sito.entecra.it/portale/index2.php?lingua=EN&id=&id_gruppo=&tipo=&tipo_org=&sezione=&opz_menu=&access_flag=0.

[25] Consortium GARR website: http://www.garr.it.

[26] ComScore™ report: It's a Social World: Top 10 Need-to-Knows About Social Networking and Where It's Headed. http://www.comscore.com/Insights/Presentations_and_Whitepapers/2011/it_is_a_social_world_top_10_need-to-knows_about_social_networking.

PoS(ISGC 2013)031