ORCID Integration into Science Gateways

Carrie Arnold

Advanced Computing Interfaces
Texas Advanced Computing Center
Austin Texas
cmarnold@tacc.utexas.edu

Maytal Dahan

Advanced Computing Interfaces
Texas Advanced Computing Center
Austin Texas
maytal@tacc.utexas.edu

David L. Hart

Computational and Information Systems Laboratory
University Corporation for Atmospheric Research
Boulder Colorado
dhart@ucar.edu

Ester Soriano

Integrated Data and Database Services
National Center for Supercomputing Applications
Urbana-Champaign Illinois
esoriano@illinois.edu

ABSTRACT

With science gateways becoming more popular every day, both researchers and gateway developers are facing challenges in collecting, displaying, and verifying researcher career information, especially across multiple projects and organizations users are involved in. ORCID has attempted to simplify this for both researchers and organizations. ORCID is a non-profit organization that provides researchers with a persistent digital identifier, known as the ORCID iD, that links them with their work and allows them to share their activities with participating websites, including science gateways.

KEYWORDS

XSEDE; Science Gateways; ORCID; Unique Identifiers

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org. PEARC '19, July 28-August 1, 2019, Chicago, IL, USA © 2019 Association for Computing Machinery. ACM ISBN 978-1-4503-7227-5/19/07...\$15.00 ACM ISBN 978-1-4503-7227-5/19/07...\$15.00 <https://doi.org/10.1145/3332186.3332227>

INTRODUCTION AND MOTIVATION

With science gateways [1] becoming more popular every day, both researchers and gateway developers are facing challenges in

collecting, displaying, and verifying researcher career information, especially across multiple projects and organizations users are involved in. Users want to be able to easily enter and display their personal and career information across the different gateways they use, and gateway developers want to be able to keep this information as up-to-date as possible.

The Extreme Science and Engineering Discovery Environment (XSEDE)[4] project is the most advanced, powerful, and robust collection of integrated digital resources and services in the world. It is a single virtual system that scientists can use to interactively share computing resources, data, and expertise. An example of a science gateway is the XSEDE User Portal (XUP). The XUP is meant to be a user-focused science gateway into XSEDE services. The XUP provides XSEDE users access to view and manage their accounts and allocations, as well as find information about and access the XSEDE services and resources. Key functionalities of the user portal include the ability to keep an up-to-date user profile and user publications related to XSEDE allocations. For profile information, the XUP has been entirely dependent on user-entered information, from emails to publications, and relies on users to keep this information up to date, which is cumbersome for both users and service providers. In an effort to streamline this process, XUP developers have started work on integrating ORCID iD into XUP.

Figure 1: XSEDE User Portal Homepage

This paper discusses the specific use case of integrating ORCID iD into the XSEDE User Portal (XUP) and XSEDE Resource Allocation System (XRAS). Figure 1, shows a sample of the XSEDE User Portal homepage.

XRAS is a software service that facilitates the process of requesting, reviewing, and managing allocations on XSEDE's network of resources [5]. The XUP provides guided and advanced interfaces for users to submit their allocation requests to XRAS. The XSEDE User Portal, a national collaboration between supercomputing centers aimed at providing a variety of resources to researchers, is a central user-focused science gateway that allows users of XSEDE to manage their allocations, register for training, add publications, and much more. By leveraging the ORCID iD information, XSEDE can expand how it gets and updates user information - from contact info to publications - across projects. The XSEDE use case can be extended to help other science gateways determine whether ORCID integration would be beneficial to them. Our initial integration was primarily focused on ORCID's publications and allocations support. Publications are a valuable metric for success for both researchers and the XSEDE projects. Ensuring users have a reliable way to add and update their publications and associate it with their XSEDE allocations and resources is of vital importance. Although XUP already has functionality for adding publications via DOI search and manual entry, having access to the ORCID API provides us another means to pull and push publications data for our researchers. By enabling ORCID integration, we give users the ability to stay up to date and in-sync with their publication listings which can be very challenging across all the platforms a researcher may use across their career. Allocation information is very site specific, and providing researchers with the ability to sync this information with a central manager like ORCID can prove invaluable for project tracking. This paper will primarily focus on integration of ORCID into the XSEDE User Portal.

WHAT IS ORCID?

ORCID [2] has attempted to simplify data entry for both researchers and organizations. ORCID is a non-profit organization that provides researchers with a persistent digital identifier, known as the ORCID iD, that links them with their work and allows them to share their activities with participating websites, including science gateways. The ORCID iD is an https URI with a 16-digit number that is compatible with the ISO Standard e.g. <https://orcid.org/0000-0001-2345-6789>. No identifying information about a person is encoded in the ORCID iD. The identifiers were designed to be usable in situations where personally-identifiable information should not or cannot be shared. Also, since the ORCID iD is designed to be a career-long identifier, no information that can change over a person's career is embedded in the iD, e.g., country, institution, field of study. After creating an ORCID account, researchers can add their grants, publications, projects, and other important career information and make it available for gateways to display or update. In order to ensure accuracy and reliability, the permissions for both the researchers and organizations is extremely fine-grained. When a researcher submits information, they have the option to set each piece to be visible publicly, for trusted parties, or completely private. The member organization is provided a client key, which it will use to request user tokens with specific scopes, giving them the ability to authenticate, access, and even update the researcher's data. The researcher must log into ORCID from the client organization via OAuth (an open authentication standard [3]) and grant the access the organization is requesting, which can be revoked at any time. The sites store the information, including access token, that is returned and use this to make API calls on behalf of the researchers. By having the ability to read and update ORCID iD information science gateways can better ensure reliable and up-to-date information about their gateway users.

TECHNICAL DESCRIPTION OF INTEGRATION PROCESS

The technical process of integrating ORCID into XUP was very streamlined. ORCID provides extensive developer documentation, integration guides, and API specifications on their website. Organizations interested in using ORCID must first become a member.

Figure 2: The ORCID OAuth sign-in window

After this step, developers are provided with keys and asked to create accounts in the ORCID “sandbox” instance, which is an instance specifically set up to develop against. Developers must also provide specific redirect URIs to ORCID. The implementation details are straightforward for the developer. They must place an OAuth link, provided by ORCID, on the desired page. This link contains the client id and the desired scope (described next), and opens a window for the user to sign in as shown in Figure 2.

Once the user signs in, they either approve or deny XSEDE’s request to access their information. What information XSEDE is asking to access, determined specifically by the scope provided, is displayed clearly in Figure 3.

Figure 3: The ORCID OAuth approve or deny screen

Once the user decides to approve, the redirect URI is given a code as a parameter. The developer takes this code and posts to ORCID to request an authorization token, which will respond with the user’s ORCID profile information. The developer is expected to store this information, such as the user’s ORCID iD and their authorization token, in their database. Once this process is complete, the portal can query the ORCID API on the user’s behalf using this authorization token. Currently, the XSEDE User Portal will display the user’s ORCID iD on their profile page and the dashboard once they have logged in. As shown in Figure 4, the ORCID iD also serves as a URL that links directly to the user’s ORCID profile. This provides an easy to find and access mechanism for viewing and updating the user’s ORCID iD information.

Figure 4: ORCID iD displayed under user picture on the XSEDE user dashboard.

There are a variety of scopes available to developers that affect what information they’re allowed to access on the user’s behalf. For example, the *read-limited* scope will only allow developers to fetch a user’s information, but also requesting the */activities/update* scope will allow developers to post updates to user information. Of course, users always have the option to revoke access, which requires the portal to make an API call to ORCID, as well as purge the user data from the database. It is important to note that each time the portal wants to change their scope, they will need to request new tokens specifically for that scope. This should not necessarily be a complication for other integration projects, but should certainly be a consideration.

CONSIDERATIONS FOR SMALLER GATEWAYS: IS ORCID RIGHT FOR YOU?

Integrating ORCID into XSEDE provided not only organizational challenges such as developing a relationship with ORCID and having XSEDE become a provider, but also developer challenges with API integration and ensuring permissions and accessibility are up to date. ORCID has extensive documentation on their site detailing how to access, store, and update data via their API. Though integration itself is relatively simple, there is a somewhat lengthy process, including providing a live demo, required to get production credentials, which must be factored into development time. This overhead may impact the decision for smaller gateways to invest resources into integration. It is also necessary to mention that becoming a member is not free, and this cost may be a barrier to projects on small budgets, especially those that are not backed by larger organizations like XSEDE.

ORCID: CURRENT USAGE

In mid-March 2019 XSEDE promoted the use of ORCID integration via the XSEDE User Portal. In less than a month time, over 400 users have linked their ORCID ID with their XSEDE User Profile. The community feedback and enthusiasm for the service has exceeded expectations. As ORCID integration capabilities are extended, we expect to see addition increase in adoption. In order to adequately obtain user feedback, the XUP team will launch a micro-survey - a short three question survey to gauge interest and capabilities in expanding ORCID integration.

FUTURE WORK

By integrating the initial components of ORCID into the XSEDE User Portal, we are able to demonstrate the viability and usefulness of further integration. The initial work done with ORCID and XSEDE gives a template of how this can be used with other science gateway projects.

Upcoming work will include promoting and publicizing ORCID integration in XSEDE to ensure adoption and expanding the integration points. Future work will focus on adding additional integration into areas including - personal profile information, project allocations, and publications. Currently, all users can do in the user portal is link their ORCID iD, and doing so does not sync any information with the portal itself. Our more immediate goals include adding the ability for users to sync their publications from ORCID to XUP, and vice versa. Later goals would involve adding more profile functionality, including career information. By adding these additional features we can ensure a user's profile, publications, and project information can exist in a central location in ORCID and live beyond their work in XSEDE.

ACKNOWLEDGMENTS

This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1548562.

Special thanks to the XSEDE Resource Allocation Service (RAS) Team for helping to get this work started.

REFERENCES

- [1] Gesing, S., Wilkins-Diehr, N., Dahan, M., Lawrence, K., Zentner, M., Pierce, M., Hayden, L.B., and Marru, S. Science Gateways: The Long Road to the Birth of an Institute. Proc. of HICSS-50 (50th Hawaii International Conference on System Sciences), 4-7 January 2017, Hilton Waikoloa, HI, USA, <http://hdl.handle.net/10125/41919>
- [2] Haak, Laurel L., Martin Fenner, Laura Paglione, Ed Pentz, and Howard Ratner. "ORCID: a system to uniquely identify researchers." *Learned Publishing* 25, no. 4 (2012): 259-264.
- [3] OAuth. 2019. "An open protocol to allow secure authorization in a simple and standard method from web, mobile and desktop applications." Retrieved from <https://oauth.net/>.
- [4] Towns, John, Timothy Cockerill, Maytal Dahan, Ian Foster, Kelly Gaither, Andrew Grimshaw, Victor Hazlewood et al. "XSEDE: accelerating scientific discovery." *Computing in Science & Engineering* 16, no. 5 (2014): 62-74.
- [5] David L. Hart, Amy Schuele, Ester Soriano, Maytal Dahan, and Matthew Hanlon. 2014. XTRAS: Allocations software as a service in XSEDE. In Proceedings of the 2014 Annual Conference on Extreme Science and Engineering Discovery Environment (XSEDE '14). ACM, New York, NY, USA, Article 65, 8 pages. DOI: <https://doi.org/10.1145/2616498.2616562>