

canarie



White Paper:

Introduction to CANARIE's Research Software Program

Author: Scott Henwood

Director, Research Software

scott.henwood@canarie.ca

613.943.5434 | Follow: @shensicle

Introduction

As is the case with most countries, Canada does not have a clearly identifiable research software development community. The community certainly exists, but members tend to interact only within specific research disciplines and the population is fairly transient due to the project-based approach to research software development used today. As a result, opportunities for collaboration and software re-use are missed and we see the same functionality being developed over and over again by the different software teams. This is particularly true in the case of non-research facing support software including user authentication components, visualization tools and digital infrastructure management.

Ultimately more time and money is spent on redundant software development, leaving less of both for research-facing software and for the actual research itself. To help alleviate this situation, CANARIE has established our Research Software Program.

For readers unfamiliar with CANARIE, we are a non-profit with major funding from the Government of Canada and a mandate to deliver digital infrastructure and drive its adoption by Canada's research, education and innovation communities. CANARIE's roots are in advanced networking, and CANARIE continues to evolve the national ultra-high-speed backbone network that enables data-intensive, leading-edge research and big science across Canada and around the world. Through its Research Software Program, CANARIE also leads the development of software that enables researchers to more quickly and easily access research data, tools, and peers.

Platforms and Services: Some Terminology

A research platform is a program or collection of software that supports the bulk of what you need to do in your research workflow. Typically, a platform's capabilities would include data acquisition and management, processing, storage and visualization. Platforms are almost always discipline-specific and frequently have a collaborative component.

Research software services, on the other hand, are software components that do one thing in the research workflow. Conceptually, a platform can be seen as an aggregation of different services. There could be a user authentication service, a cloud batch processing service, a virtual file system service and a visualization service, all combined together in a single research platform.

Indeed, re-usable software services are the real key to accelerating research. Suppose you have a research platform designed for a specific discipline, say astronomy. Other research groups involved in astronomy may be able to use this platform directly, resulting in a huge saving in software development cost and time. This, in itself, is a worthwhile outcome. But now suppose that your platform was made up of easily extractable software services that could be used in platforms developed by others. A component for visualizing 2D multi-spectral data could be used in platforms designed for research disciplines other than astronomy. A platform infrastructure component like group management may be of use to many platforms, regardless of their area of research.

I get the theory, but what is a Research Software Service?

For the purposes of our research software program, CANARIE has defined three types of research software services.

Managed Services are traditional web services. Service owners deploy and maintain the service on their own compute resources and users access the service over the network using a well-defined API, typically a web services interface.

Self-Deployed Services are also traditional web services. Unlike Managed Services however, Self-Deployed Services are deployed by the service users. A Self-Deployed Service is released to potential users in a format (source code, binaries, installers, VM images) that allows these users to run an instance of the service on their own hardware or cloud resources. In situations where the service depends on a piece of hardware, a sensor network or physical storage, a Managed Service is the way to go. Self-Deployed Services are more appropriate when there is no need to be co-located with expensive equipment and the service owners do not have the compute or storage resources required to support other users. CANARIE recommends that owners of Self-Deployed Services deploy a demo/evaluation instance to allow potential users to try out the service before committing to it.

Integrated Services are software services that don't fit into the web service model. Examples of this type of service would include Javascript code and VM images - software that would typically be included directly in your platform.

Where can I go to find *Platforms and Services*?

As part of our Research Software program, CANARIE has established a research software portal. Our vision is that this portal will become a major resource for finding and learning about software to aid in scientific research.

You can visit the portal at <https://science.canarie.ca>

Currently the portal serves as a registry for software services and platforms used in scientific research. We will continue to add support for different types of research resources depending on feedback from the community. This software is applicable to many areas of scientific discovery and covers all aspects of the research workflow.

The portal is not, however, merely a list of platforms and services. Contributions are made by working researchers and research software developers, people who understand the needs of other research teams. CANARIE adds an amount of governance to the contributions by asking for specific documentation in order to provide a complete, consistent user experience. We also monitor contributed services and platforms for availability/reliability and this information is made available to potential users.

Can I incorporate the Research Software Services available on the portal into my own Research Platform?

Yes! All services registered with the CANARIE research software portal are free to use for research purposes and many are open source. If you decide to use one or more of the services registered at the portal, please let us know so we can better track service adoption and value. In the near future, we'll be releasing a small software widget that platform developers can include in one of their UI pages. This widget will tell the portal which services are being used by each platform and, will automatically notify the platform owners when an update is available for one of the services they use.

Can CANARIE help get the word out about my Software Service or Platform?

Absolutely! CANARIE has extensive reach into the Canadian Research and Education community and will be happy to promote your contributed research platform or service. We're also working on a new landing page for the research software portal that will showcase these platforms and services. Additionally, we'll soon be embarking on extensive marketing effort to promote the research software program to the wider research community.

As a contributor to the research software community, you may be eligible to send members of your team to one of CANARIE's research software developers' workshops to meet other research software developers, provide CANARIE with input to the future of the program, and explore new opportunities for collaboration.

I have a Research Software Service I'd like to contribute! What are the requirements for registering my Service with CANARIE's Research Software Portal?

The most important function of the portal is to provide useful, well-documented software services of the highest quality to accelerate research. For this reason, we have outlined some simple requirements for service contribution below.

Maturity

All research software services contributed to the registry should be under active use in a research platform, usually by the service owner, at the time of registration. The wide variety of services and their varied research disciplines make a formal qualification test by CANARIE impractical. Instead, we're providing a similar level of confidence by ensuring that all contributed software is being used in at least one active research platform. By the same token, all contributed services should be of production quality. If you would like to pre-announce a service that is still in either development or alpha/beta testing, please contact us to discuss having it included on our "Coming Soon" list.

Documentation

One of our main goals is to enable potential users to easily obtain the information they need in deciding whether or not to use a contributed service in their research platform. In order to maintain a consistent user experience on the portal, CANARIE asks that the following information be provided for all registered services:

- The name of the service
- A one or two sentence description of the service and its capabilities
- The name of the owner/creator institution
- The category (ie. type) of the service
- The research subject (from the [NSERC tables](#)) to which the service applies.
- Tags – descriptive strings about the service used to aid in site searching
- The service version
- Release notes
- Service user documentation
- Information on how to obtain support
- A link to the service source code (optional - for open source software only)
- A link where potential users can try out the service
- License statement
- Provenance statement

For an example of this documentation, please visit the details page for the CANARIE reference service at

<https://science.canarie.ca/researchmiddleware/researchresource/main.html?resourceID=16>

Licensing

We ask that all services registered with the portal be free to use by researchers. You may choose to offer additional capabilities on a "for-fee" basis, but the registered version should not be degraded in any way. CANARIE recommends that all software service contributors consider releasing their service's source code under an open source license, although this is not a requirement.

Monitoring

Currently, the portal supports only Managed Services. We will be completing support for Self-Deployed Services and Integrated Services over the summer of 2014. The portal will periodically poll the documentation links described above to ensure that they are available. For more comprehensive monitoring of Managed Services, consider implementing CANARIE's full service monitoring API as described in the document entitled Research Service Support for the CANARIE Registry and Monitoring System. This API, which is fully supported today, allows us to report on your service's availability, reachability and usage, as well as providing all of the information listed in the Documentation section above.

Maintenance

Research software services are really only of value to the community if they are available and supported for a period of time long enough to allow them to be effectively used in research platforms. Rather than attempting to enforce maintenance period requirements, CANARIE asks that you provide a development roadmap for your service that tells potential users about your development and maintenance plans. That doesn't necessarily mean that you will be expected to continue development of the service after it's been contributed - that's entirely up to you. We do, however, require that you provide instructions to users on obtaining support for your service, including a support email address.

Some users may depend on specific versions of your software service for experimental reproducibility. Please inform potential users of your policy regarding support for old versions in your documentation. This is particularly important for Managed Services.

We will be adding a user communication feature to the portal this year. This feature will allow service owners to communicate events of significance, including update notifications and service bulletins, to their users. As part of this feature, users will be able to "follow" a service in order to receive such notifications.

Can I register my Research Platform with the CANARIE Science Portal?

Yes. We accept both Managed and Self-Deployed platforms in the registry. The portal is about accelerating discovery and not about marketing, so if you contribute a Managed Platform, it must be freely accessible to all users with appropriate academic credentials. Contributing a platform to the portal is an excellent way to meet potential new collaborators. To discuss further, please contact us.

What should I do next?

Archived editions of our program newsletter can be found at <https://collaboration.canarie.ca/elgg/file/group/62/all#964> . To subscribe to this newsletter, simply send an email to support@science.canarie.ca indicating your wish to subscribe. If you would like more information about the program or are interested in contributing a research service or platform, please contact:

Scott Henwood
Director, Research Software
scott.henwood@canarie.ca
613.943.5434 | Follow @shensicle