

# Towards a Dashboard Environment for Repeating and Reusing Research

Ian Taylor

Ian Taylor, Cardiff University, UK  
and University of Notre Dame, USA

Jarek Nabrzyski

University of Notre Dame, USA

## ABSTRACT

Over the past 18 months, we have been working on a dashboard concept that enables researchers a means of interacting with existing research. This work was motivated by the National Data Service (NDS [1]), which is an emerging vision of how scientists and researchers across all disciplines can find, reuse, and publish data. NDS intends to provide an international federation of data providers, data aggregators, community-specific federations, publishers, and cyberinfrastructure providers by linking data archiving and sharing efforts together with a common set of tools. This abstract provides a status of the two existing proof-of-concept pilot dashboard implementations and how we plan to evolve this work.

At a high level, the researcher dashboard aims to provide an intuitive Web-based interface to expose fully *interactive research containers* that support the lifecycle of scholarly communication. Research containers enable executable and repeatable research by supporting methods, source code, and data within dynamically created Docker containers. To date, there have been two versions of the dashboard, which are described below.

The V1.0 dashboard integrated with the NDS Labs *Epiphyte* API, which interfaced with Docker containers and data management systems. This dashboard used Yii [2] and an SQL engine that was modelled using the OAI-ORE Web aggregation standard[3] to allow interoperability with other repositories. The system has been demonstrated at NDS meetings (Washington DC and Austin), and at SC14 at the NCSA booth. It is also checked into the NDS open source repository[4]. The main issue of this style of approach is the synchronization of research data and metadata with external repositories, since generally there are no automated means of achieving this. This led us to a more lightweight integration approach.

The Dashboard V2 interfaced to the Open Science Framework (OSF), which is an environment that supports open materials, data, tools to connect projects and initiatives and easy online publishing of results. Using this system, a researcher can create a project on the OSF, connect data management tools to it (e.g. Google Drive, Dropbox, Box, Dataverse, and so on) and then use the dashboard (implemented as an add on) to execute methods on data, stored on OSF, in a container using the Boatload API[5]. Boatload is an API for automating deployment and operations of Docker containers on clusters. It provides a thin layer on top of Fleet and CoreOS for further operations, such as getting container “diffs”, downloading files, streaming stdout, and so on.

Looking forward, we are planning on adopting a more lightweight federated approach; that is, we will provide similar interfaces to research containers but instead adopt a thin client based approach to multiple authentication and data infrastructures. To this end, we are now working on an Ember[6] toolkit of reusable components that bind to the OSF API. This implementation will verify the proof of concept that it is possible to create a common lightweight interface to support multiple distributed computing infrastructures. On the roadmap is to provide a similar binding to Globus[7] and DataOne[8].

## REFERENCES

- [1] The national Data Service Website: <http://nationaldataservice.org>
- [2] The Yii Framework: <http://www.yiiframework.com/>
- [3] Open Archives Initiative Object Reuse and Exchange: <https://www.openarchives.org/ore/>
- [4] Dashboard Repository: <https://bitbucket.org/nds-org/nds-dashboard>
- [5] The Boatload API: <https://bitbucket.org/keyz182/boatload>
- [6] EmberJS: <http://emberjs.com/>
- [7] Globus API: <https://docs.globus.org/api/transfer/>
- [8] DataOne API: <http://jenkins-1.dataone.org/documentation/unstable/API-Documentation-development/apis/index.html>