

Research Graph VIVO Cloud Pilot

To be archived



This page describes a project that was not completed. It will be archived.

Project Summary

Many VIVO implementers find collecting, mapping, and loading data into VIVO to be quite difficult. For example, data on publications, grants, and datasets produced by an institution's faculty can be difficult to find and disambiguate. Understanding the ontologies used to describe data in VIVO and mapping faculty data to those ontologies involves a steep learning curve. Also, transforming the data to a linked data format, such as VIVO RDF, has proven difficult for most implementers due to gaps in skills and knowledge. These barriers have prevented organizations from joining the VIVO community and adopting the technology that enables access, discovery, and analysis of scholarship data.

Research Graph is an integrated network of information about researchers, their publications, grants, and datasets, across global research infrastructures such as ORCID, DataCite, CERN, CrossRef, and funders such as National Institutes of Health (NIH). For example, when provided "seed data," such as a simple list of researchers, Research Graph will identify publications, grants, and/or datasets related to those researchers and represent the information in a graph. These are referred to as "first order" connections. Research Graph is also capable of identifying and linking collaborators of the people in the "first order" data and linking their publications, grants and datasets. These collaborator links are referred to as "second order" connections.

A recent collaboration between VIVO and Research Graph [1] developed a repeatable process for using seed data to build first and second order graphs, and to export, transform, and load those graphs in VIVO RDF format to a hosted VIVO instance. We believe 1) Repositories and Research Institutes, 2) Semantic Web Sites of government and research organisations, and 3) Current VIVO Sites that wish to enrich and augment their data can benefit from the collaboration between VIVO and Research Graph. The Cloud Pilot will have participants representing these three types of organizations.

As illustrated in the following figure, during this pilot we enrich the research repositories' data by (1) Transforming repository data to a graph database, (2) Augmenting the graph with the Research Graph data, (3) Making this graph available as a VIVO instance. In addition, we make the graph data available as GraphML or JSON for further analysis and visualisation.

[blocked URL](#)

The project will determine the value and potential of a long term collaboration between VIVO and Research Graph in the form of new services that could reduce barriers for organizations that want to find, disambiguate, transform, and map research data. To read a more detailed description of the project, please see our [Abbreviated Proposal](#). To ask a question or express interest in participating, please contact Erin Tripp, DuraSpace Business Development Manager at etripp@duraspace.org.

Footnotes:

1. Conlon, Michael, and Amir Aryani. "Creating an Open Linked Data Model for Research Graph Using VIVO Ontology," July 24, 2017. <https://doi.org/10.4225/03/58ca600d726bd>.