Enhancing SciENcv through semantic research profile integration with the VIVO-ISF ontology

Marijane White, Matthew Brush, Shahim
Essaid, Robin Champieux, Adrienne Zell,
Melissa Haendel
Oregon Health and Science University
Portland, OR, USA

Colin Grove, Syeda Momina Tabish,
David Eichmann
School of Library and Information Science
University of Iowa
Iowa City, IA, USA

Keywords—SciENcv; VIVO-ISF; ontology; semantic integration

I. INTRODUCTION

SciENcv [1], the US Federal Science Experts Network Curriculum Vitae, is an online system for simplifying the creation of researcher biographical sketches or biosketches, which are required when applying for federal funding. SciENcv profiles are curated and controlled by researchers themselves - they own the data, they control what data are public, and they edit and maintain the information contained within the profiles, which includes expertise, employment history, educational background, and professional accomplishments. The system leverages data from myNCBI and eRA Commons, and includes links to ORCiD [2] researcher identifiers. The structure of SciENcv biosketches is defined by an XML Schema Definition and profiles can be downloaded in XML format. The system aims to eliminate the need for researchers to repeatedly enter biosketch information and reduce the administrative burden associated with federal grant submission and reporting requirements, as well as creating a repository of researcher profile data where researchers can describe their scientific contributions in their own language.

The VIVO Integrated Semantic Framework (VIVO-ISF) [3] is an ontology for representing people, works, and the relationships between them and was developed as a merger of the VIVO [4] and eagle-I [5] ontologies. VIVO-ISF is the designated standard for research profiling data in the Centers for Translational Science across the US, and is utilized by a variety of research profiling tools such as Harvard Profiles. However, many research profiling tools are on different versions and/or have not yet adopted VIVO-ISF as a standard representation. SciENcv is in this category.

II. ONGOING WORK

The SciENcv integration project aims to create interoperability between the SciENcv XML Schema and the VIVO-ISF ontology. This will facilitate generation and consumption of standard, compliant data, expand the role of SciENcv in exchanging data within the research landscape, and broaden the definition of researcher impact by facilitating the emergence of new standards relating to the contribution and attribution of researchers to software, datasets, and other scholarly products. Achieving this integration will support improved research profiling analytics across a much wider set of data sources and allow researchers to more effectively and specifically describe their contributions.

The project enables mapping between the SciENcv XML Schema and the OWL constructs in VIVO-ISF. The mapping is being validated via CTSAsearch [6], which ingests VIVO-ISF, ORCiD, PubMedCentral, Medline, and more, allowing for a fully integrated and curated set of institutional data. CTSAsearch is also adding community detection visualization features for search results.

Here we showcase the interpretation of the SciENcv schema, its alignment to the VIVO-ISF ontology, highlight some of the competency questions being used to evaluate exchanged profiles, and some example queries that can be used to access SciENcv profile data via a VIVO-ISF version 1.6 SPARQL endpoint.

III. REFERENCES

[1] National Center for Biotechnology Information, "SciENcv: Science Experts Network Curriculum Vitae," http://www.ncbi.nlm.nih.gov/sciencv/

[2] ORCiD, http://orcid.org/

[3] VIVO Integrated Semantic Framework Ontology, https://wiki.duraspace.org/display/VIVO/VIVO-ISF+Ontology

[4] VIVO, http://vivoweb.org/

[5] eagle-i ontology, https://search.eagle-i.net/model/

[6] CTSAsearch, http://research.icts.uiowa.edu/polyglot/