

Curating and enabling discovery of New York State climate change science content using semantic web technologies

Huda J. Khan, Jon Corson-Rikert, Ingrid H. H. Zabel, Darcy Branchini, and Arthur DeGaetano

Cornell University, Ithaca, NY, U.S.A.

Abstract. The New York Climate Change Science Clearinghouse (NYCCSC) project arose in response to the need for a central portal for policy makers to access vetted data and information relevant to climate change adaptation and mitigation in New York State. This project involves multiple components including a publicly available website encompassing a software application that utilizes semantic web technologies to represent and categorize diverse types of data and information in a manner optimized for search. This application extends the VIVO open ontology and software suite (<http://vivoweb.org>), leveraging VIVO's ability to represent people, projects, and organizations and adding the capability for representing climate change science concepts and specific relationships between those concepts and data, maps, and documents.

1 NYCCSC Project: Background and Extension of VIVO

The impacts of climate change on natural and built infrastructure are wide-reaching and affect multiple sectors identified in the ClimAID report [1] such as water resources and transportation. Policy makers and researchers tasked with understanding these effects in their local regions and formulating adaptation and mitigation strategies must utilize information from multiple, diverse and scattered sources and, therefore, may draw different conclusions based on the information they encounter. The NYCCSC project emerged as a means to address these needs and aims to provide a web-based portal enabling access to data and information relevant to climate change adaptation and mitigation across New York State. Enabling a central source of regional information will support policy makers and researchers in assessing and addressing the effects of climate change. The curation process for determining what content is referenced in the clearinghouse is headed by a subject matter expert and will integrate feedback from sector experts. The NYCCSC site is scheduled to go into production by early 2016.

For this project, we extended the VIVO application and ontology [2, 3] to support ongoing curation and to represent the rich semantic relationships between the different types of content accessible through the clearinghouse. As part of this extension, we

VIVO is an open-source semantic application that supports representing academic research communities. VIVO extends the ontology instance editor and display application called Vitro [2, 3]. VIVO provides functionality for defining, displaying, and searching for researcher profiles and utilizes the VIVO-ISF ontology for describing information about researchers and their associated publications, grants, and organizations. More than 100 institutions and agencies across more than 30 countries are implementing VIVO or producing VIVO-compatible data. Extending VIVO allows us to leverage the linked data functionality in VIVO and to enable the information in the clearinghouse to be accessible through linked data to other linked data applications.

We constructed the NYCCSC ontology to represent climate change concepts and to support both curation and discoverability of content related to these concepts. These semantic relationships are central to both curation and search and discovery, forming the core of the curation process of the content that will be linked or made accessible through the clearinghouse site as well as informing the faceting and search in the Solr search index driving the clearinghouse site search. Figure 1 shows a subset of the NYCCSC ontology and a part of the related underlying VIVO ontology. All nodes in the image are within the NYCCSC ontology namespace except for those with a prefix preceding the class name.

2 Search and Discovery Example

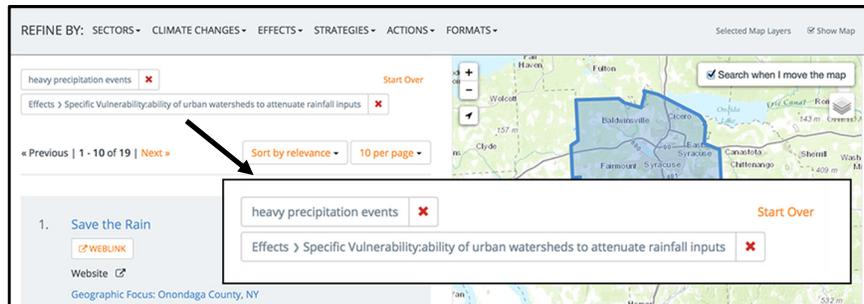


Fig. 2. Screenshot excerpt showing results for “heavy precipitation events” filtered by one of the Effects facet options and further refined by zooming into the map near the Syracuse region.

The NYCCSC VIVO extension added multiple queries to populate the Solr search index facets used on the front end. Figure 3, which represents individual resources using rectangular shapes and classes using ovals, shows the semantic relationships which populate the facet values leading to the “Save the Rain” website result in the clearinghouse search displayed in Figure 2. The “Effect addressed by” relationship is used to populate the Effect facet, enabling the selection of a particular Effect facet value in the interface to filter search results to those that utilize this semantic relationship.

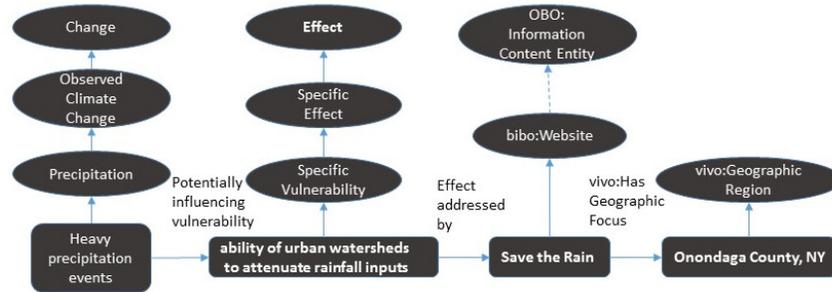


Fig. 3. Relationships supporting discovery of resources related to heavy precipitation events

3 Evaluation

After the late 2015 Beta release, we will continue to conduct usability testing to evaluate how well the site supports the exploration of the effects of climate change and strategies for addressing those effects. User feedback on faceted search will further inform ontology design as we refine what relationships should be represented to promote discoverability of content across various dimensions of interest for the user.

Acknowledgments. This project is funded by the NY State Energy Research and Development Authority and includes members Northeast States for Coordinated Air Use Management, State University of New York College of Environmental Science and Forestry, and the federal National Oceanic and Atmospheric Administration.

References.

1. Rosenzweig, C., Solecki, W., DeGaetano, A., O'Grady, M., Hassol, M., Grabhorn, P. (eds.): Responding to Climate Change in New York State: The ClimAID Integrated Assessment for Effective Climate Change Adaptation. Technical Report. New York State Energy Research and Development Authority (NYSERDA), Albany, New York. (2011)
2. Borner, K., Conlon, M., Corson-Rikert, J. VIVO: A Semantic Approach to Scholarly Networking and Discovery. Morgan & Claypool Publishers (2012)
3. Krafft, D. B., Cappadona, N. A., Caruso, B., Corson-Rikert, J., Devare, M., Lowe, B. J., & VIVO Collaboration. VIVO: Enabling National Networking of Scientists. In: Proceedings of the WebSci10: Extending the Frontiers of Society On-Line. (2010)
4. The Bibliographic Ontology (BIBO), <http://bibliontology.com/>
5. Data Catalog Vocabulary (DCAT), <http://www.w3.org/TR/vocab-dcat/>
6. The Event Ontology, <http://motools.sourceforge.net/event/event.html>
7. The Friend of a Friend (FOAF) project, <http://www.foaf-project.org/>
8. The Open Biological and Biomedical Ontologies, <http://www.obofoundry.org/>
9. PROV-O: The PROV Ontology, <http://www.w3.org/TR/prov-o/>
10. SKOS Simple Knowledge Organization System, <http://www.w3.org/2004/02/skos/>