VecNet Digital Library & Sharing Vector Disease Simulations

A Presentation for Open Mathematical and Physical Sciences Data Preservation Workshop 1: Gauging the Impact of Requirements for Public Access to Data
Arlington, VA Nov 19, 2015

Natalie K. Meyers, E- Research & VecNet Digital Librarian
University of Notre Dame

JAMES COOK UNIVERSITY AUSTRALIA

UNIVERSITY OF NOTRE DAME

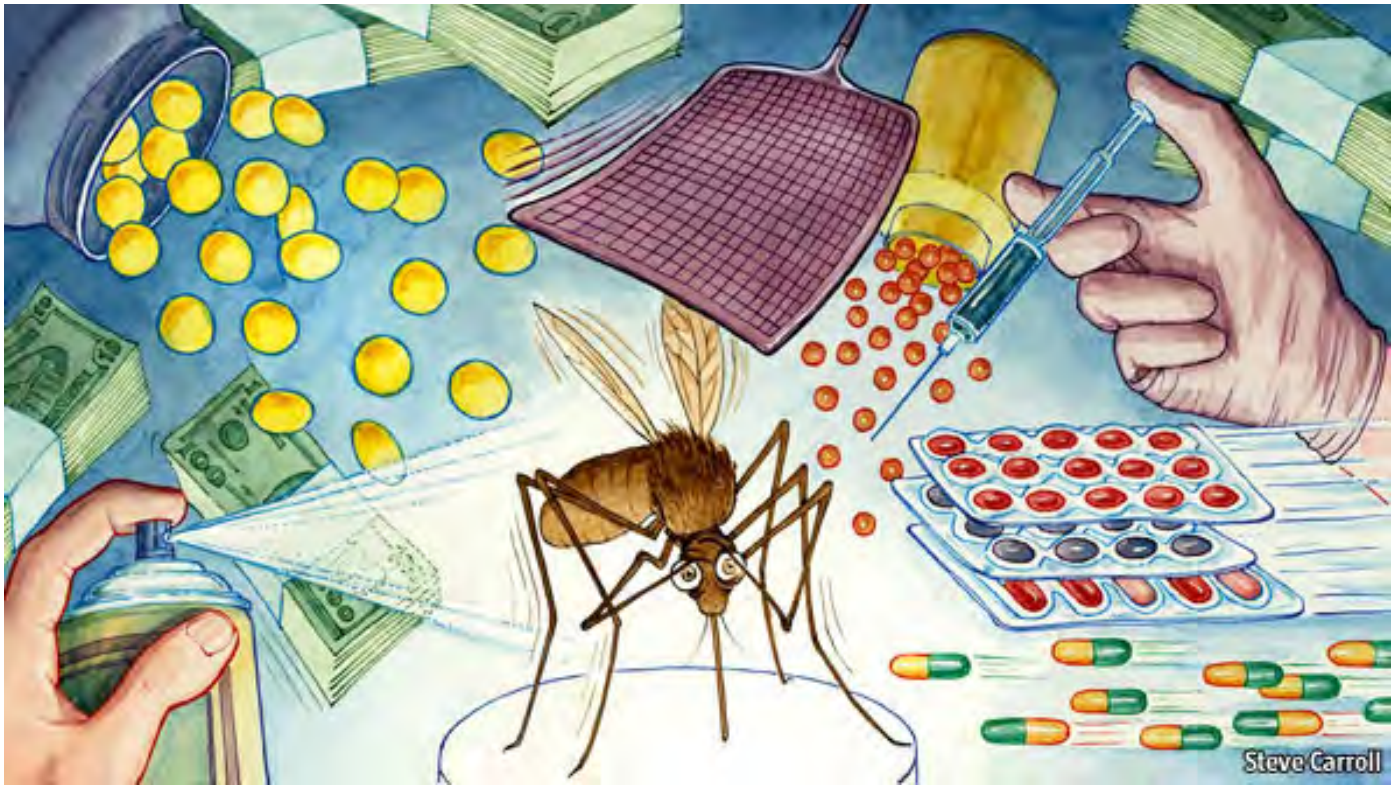
UNIVERSITY OF OXFORD

globalgood Ideas by INTELLECTUAL VENTURES

Swiss TPH
Swiss Tropical and Public Health Institute
Schweizerisches Tropen- und Public Health-Institut
Institut Tropical et de Santé Publique Suisse

R. Farlow Consulting, LLC

Amidst Heterogeneous Data
How can information be shared with and between modelers?



VecNet Digital Library

dl.vecnet.org Vector-borne Disease Network

VecNet digital library supports mathematical modeling of malaria transmission & eradication.

It is a repository for curating & sharing information about simulations used to model malaria transmission & the impact of interventions

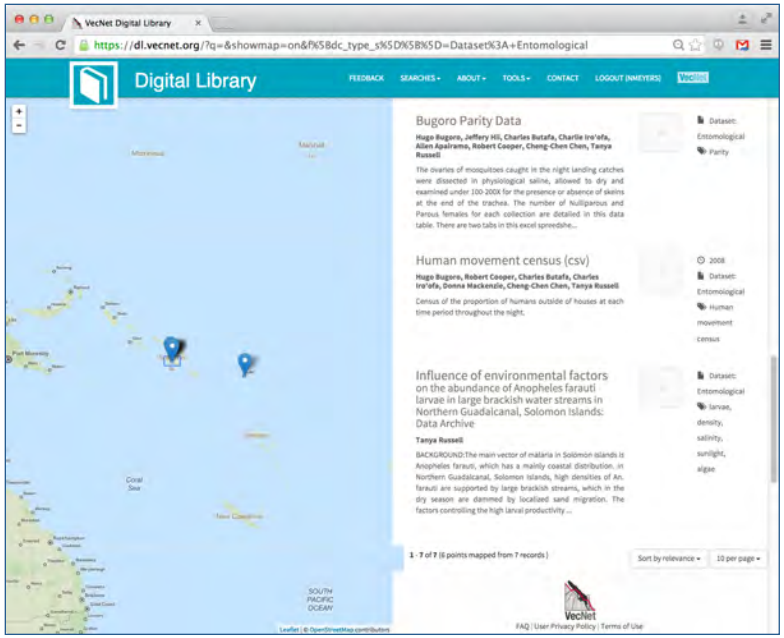
Contains: field, lab, survey, climate, demographic, and simulation data, input file code snippets, input file sets for models, simulations, tagged bibliographic citations, articles, maps, reports and more on entomology, epidemiology, demography, climatology, and interventions

VecNet

VecNet Digital Library

dl.vecnet.org Vector-borne Disease Network

The repository fulfills B&MGF mandate for global access, provides constituents from software developers to researchers with an api & sharing platform that has embargo and privacy features during active project phases, serves data sharing and needs during publication lifecycle a as well as preservation needs.



VecNet Digital Library

dl.vecnet.org Vector-borne Disease Network

Leverages strong existing open source communities and standards based choices

Built on Hydra Project technology stack:
Fedora Commons repository software
Ruby on Rails
SOLR
Blacklight & GeoBlacklight discovery interface
Dublin Core Metadata Initiative
Extended w/Darwin Core, NCBI Taxonomy, NLM's MeSH, GeoNames

In the spirit of the B&MGF global access mandate, the computer code used to create the VecNet tools (including the Digital Library) can be accessed at the https://github.com/vecnet/.

VecNet

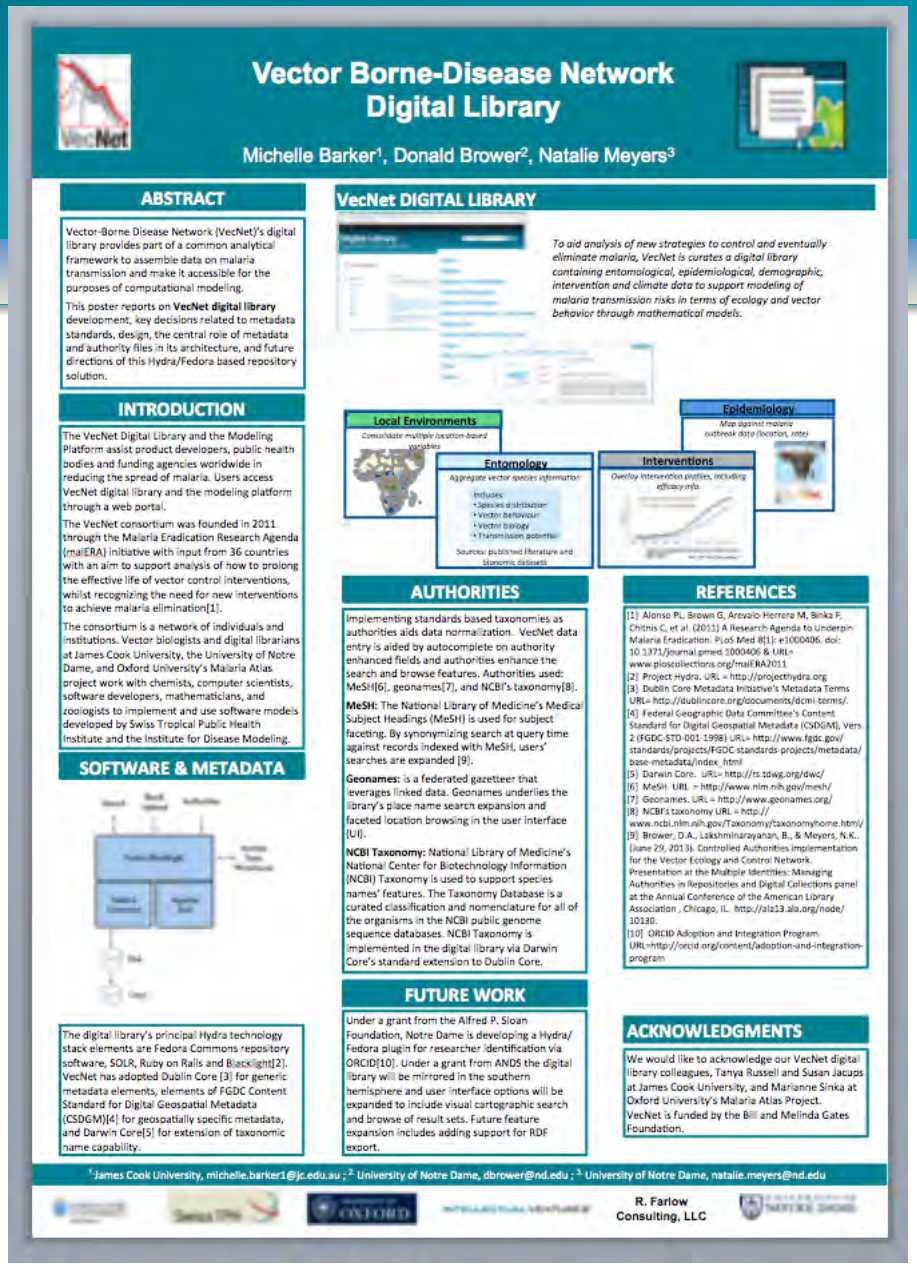
VecNet Digital Library Development Phases

Our digital library software stack and features were first developed and presented for beta feedback in 2013:

Brower D, Lakshminarayanan B, Meyers N. ***Multiple Identities: Managing Authorities in Repositories and Digital Collections*** presented at American Library Association Annual Conference, Chicago, IL 2013.

and then again at last year's ACM/IEEE JCDL conference :

Barker M, Brower D, and Meyers N. ***Vector-Borne Disease Network Digital Library*** presented at Digital Libraries 2014 IEEE(978-1-4799-5569-5) London, UK, Sept 9, 2014 .



Vector Borne-Disease Network Digital Library

Michelle Barker[1], Donald Brower[2], Natalie Meyers[3]

ABSTRACT

Vector-Borne Disease Network (VecNet)'s digital library provides part of a common analytical framework to assemble data on malaria transmission and make it accessible for the purposes of computational modeling.

This poster reports on **VecNet digital library** development, key decisions related to metadata standards, design, the central role of metadata and authority files in its architecture, and future directions of this Hydra/Fedora based repository solution.

INTRODUCTION

The VecNet Digital Library and the Modeling Platform assist product developers, public health bodies and funding agencies worldwide in reducing the spread of malaria. Users access VecNet digital library and the modeling platform through a web portal.

The VecNet consortium was founded in 2011 through the Malaria Eradication Research Agenda (malERA) initiative with input from 36 countries with an aim to support analysis of how to prolong the effective life of vector control interventions, whilst recognizing the need for new interventions to achieve malaria elimination[1].

The consortium is a network of individuals and institutions. Vector biologists and digital librarians at James Cook University, the University of Notre Dame, and Oxford University's Malaria Atlas project work with chemists, computer scientists, software developers, mathematicians, and zoologists to implement and use software models developed by Swiss Tropical Public Health Institute and the Institute for Disease Modeling.

SOFTWARE & METADATA

The digital library's principal Hydra technology stack elements are Fedora Commons repository software, SOLR, Ruby on Rails and Blacklight[2]. VecNet has adopted Dublin Core [3] for generic metadata elements, elements of FGDC Content Standard for Digital Geospatial Metadata (CSDGM)[4] for geospatially specific metadata, and Darwin Core[5] for extension of taxonomic name capability.

VecNet DIGITAL LIBRARY

To aid analysis of new strategies to control and eventually eliminate malaria, VecNet is curates a digital library containing entomological, epidemiological, demographic, intervention and climate data to support modeling of malaria transmission risks in terms of ecology and vector behavior through mathematical models.

Local Environments

Entomology

Epidemiology

Interventions

AUTHORITIES

Implementing standards based taxonomies as authorities aids data normalization. VecNet data entry is aided by autocomplete on authority enhanced fields and authorities enhance the search and browse features. Authorities used: MeSH[6], geonames[7], and NCBI's taxonomy[8].

MeSH: The National Library of Medicine's Medical Subject Headings (MeSH) is used for subject faceting. By synonymizing search at query time against records indexed with MeSH, users' searches are expanded [9].

Geonames: is a federated gazetteer that leverages linked data. Geonames underlies the library's place name search expansion and faceted location browsing in the user interface (UI).

NCBI Taxonomy: National Library of Medicine's National Center for Biotechnology Information (NCBI) Taxonomy is used to support species names' features. The Taxonomy Database is a curated classification and nomenclature for all of the organisms in the NCBI public genome sequence databases. NCBI Taxonomy is implemented in the digital library via Darwin Core's standard extension to Dublin Core.

FUTURE WORK

Under a grant from the Alfred P. Sloan Foundation, Notre Dame is developing a Hydra/Fedora plugin for researcher identification via ORCID[10]. Under a grant from ANDS the digital library will be mirrored in the southern hemisphere and user interface options will be expanded to include visual cartographic search and browse of result sets. Future feature expansion includes adding support for RDF export.

REFERENCES

[1] Alonso PL, Brown G, Arevalo-Herrera M, Binka F, Chitnis C, et al. (2011) A Research Agenda to Underpin Malaria Eradication. PLoS Med 8(1): e1000406. doi: 10.1371/journal.pmed.1000406 & URL= www.plosollections.org/malERA2011

[2] Project Hydra. URL = http://projecthydra.org

[3] Dublin Core Metadata Initiative's Metadata Terms URL= http://dublincore.org/documents/dcmi-terms/.

[4] Federal Geographic Data Committee's Content Standard for Digital Geospatial Metadata (CSDGM), Vers 2 (FGDC-STD-001-1998) URL= http://www.fgdc.gov/standards/projects/FGDC-standards-projects/metadata/base-metadata/index_html

[5] Darwin Core. URL= http://rs.tdwg.org/dwc/

[6] MeSH URL = http://www.nlm.nih.gov/mesh/

[7] Geonames. URL = http://www.geonames.org/

[8] NCBI's taxonomy URL = http://www.ncbi.nlm.nih.gov/Taxonomy/taxonomyhome.html/

[9] Brower, D.A., Lakshminarayanan, B., & Meyers, N.K.. (June 29, 2013). Controlled Authorities Implementation for the Vector Ecology and Control Network. Presentation at the Multiple Identities: Managing Authorities in Repositories and Digital Collections panel at the Annual Conference of the American Library Association , Chicago, IL. http://ala13.ala.org/node/10130.

[10] ORCID Adoption and Integration Program. URL=http://orcid.org/content/adoption-and-integration-program

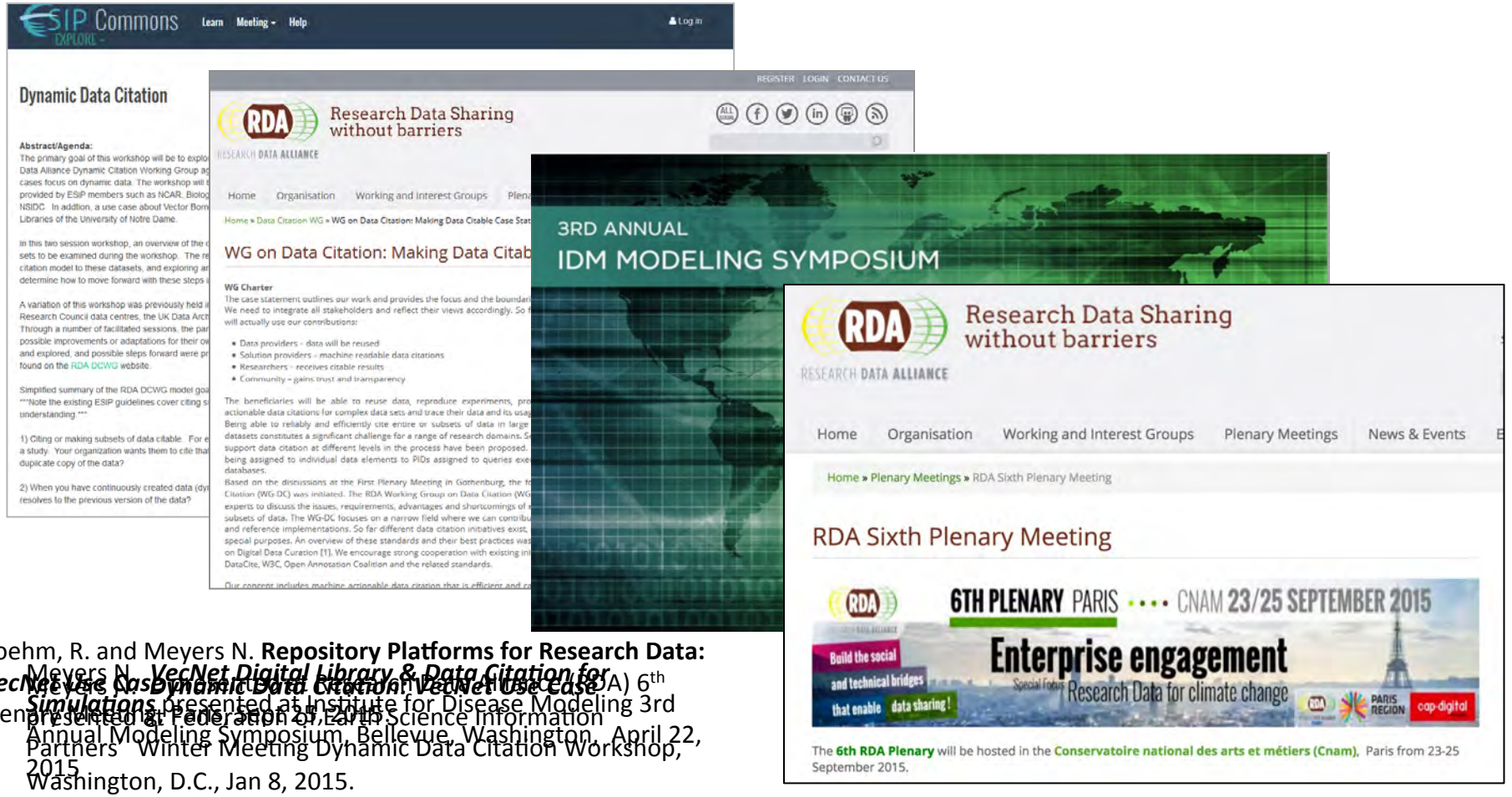
ACKNOWLEDGMENTS

We would like to acknowledge our VecNet digital library colleagues, Tanya Russell and Susan Jacups at James Cook University, and Marianne Sinka at Oxford University's Malaria Atlas Project. VecNet is funded by the Bill and Melinda Gates Foundation.

[1] James Cook University, michelle.barker1@jc.edu.au ; [2] University of Notre Dame, dbrower@nd.edu ; [3] University of Notre Dame, natalie.meyers@nd.edu

R. Farlow Consulting, LLC

Development Highlights: Dynamic Data Citation & Repositories for Research Data



Boehm, R. and Meyers N. **Repository Platforms for Research Data: *VecNet Use Case*** presented at Research Data Alliance (RDA) 6th Plenary Meeting in Paris, September 2015.

Meyers N. ***VecNet Digital Library & Data Citation for Simulations*** presented at Institute for Disease Modeling 3rd Annual Modeling Symposium, Bellevue, Washington, April 22, 2015

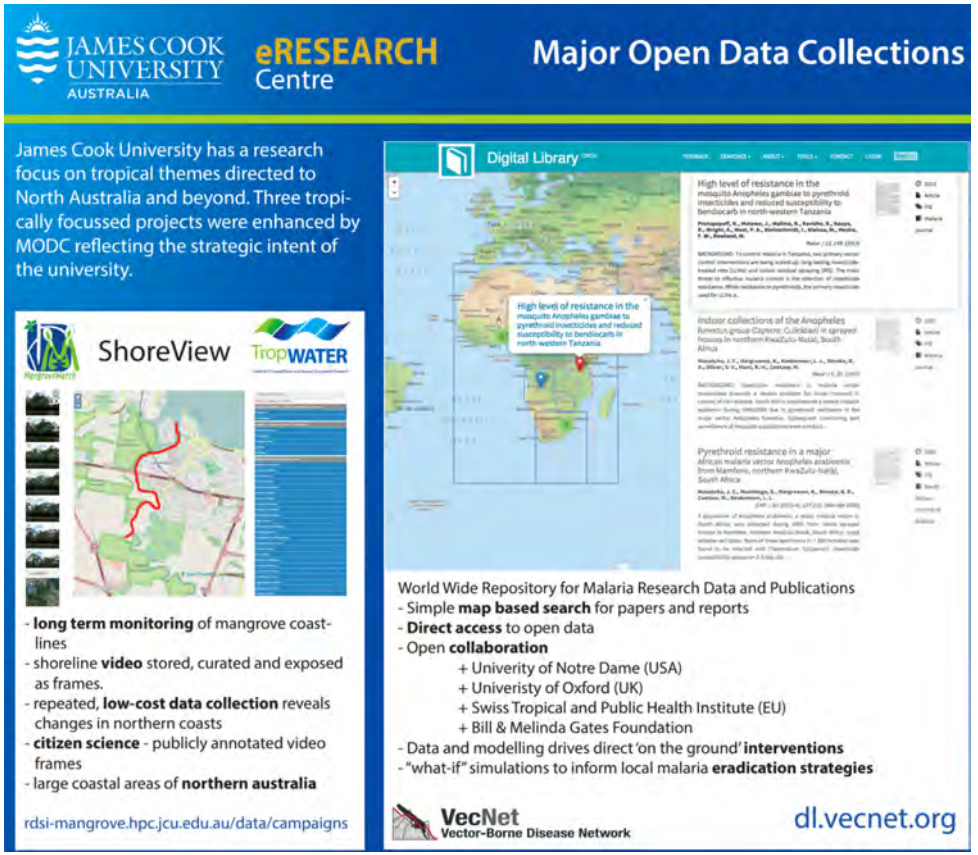
Meyers N. ***Dynamic Data Citation: VecNet Use Case*** presented at Federation of Earth Science Information Partners' Winter Meeting Dynamic Data Citation Workshop, Washington, D.C., Jan 8, 2015.

Cartographic Discovery Extension

GeoBlacklight discovery layer leverages standards based choices already in place in the VecNet digital library & future development benefits all parties

Builds on

- OpenGeoportal federated metadata sharing
- GeoRSS



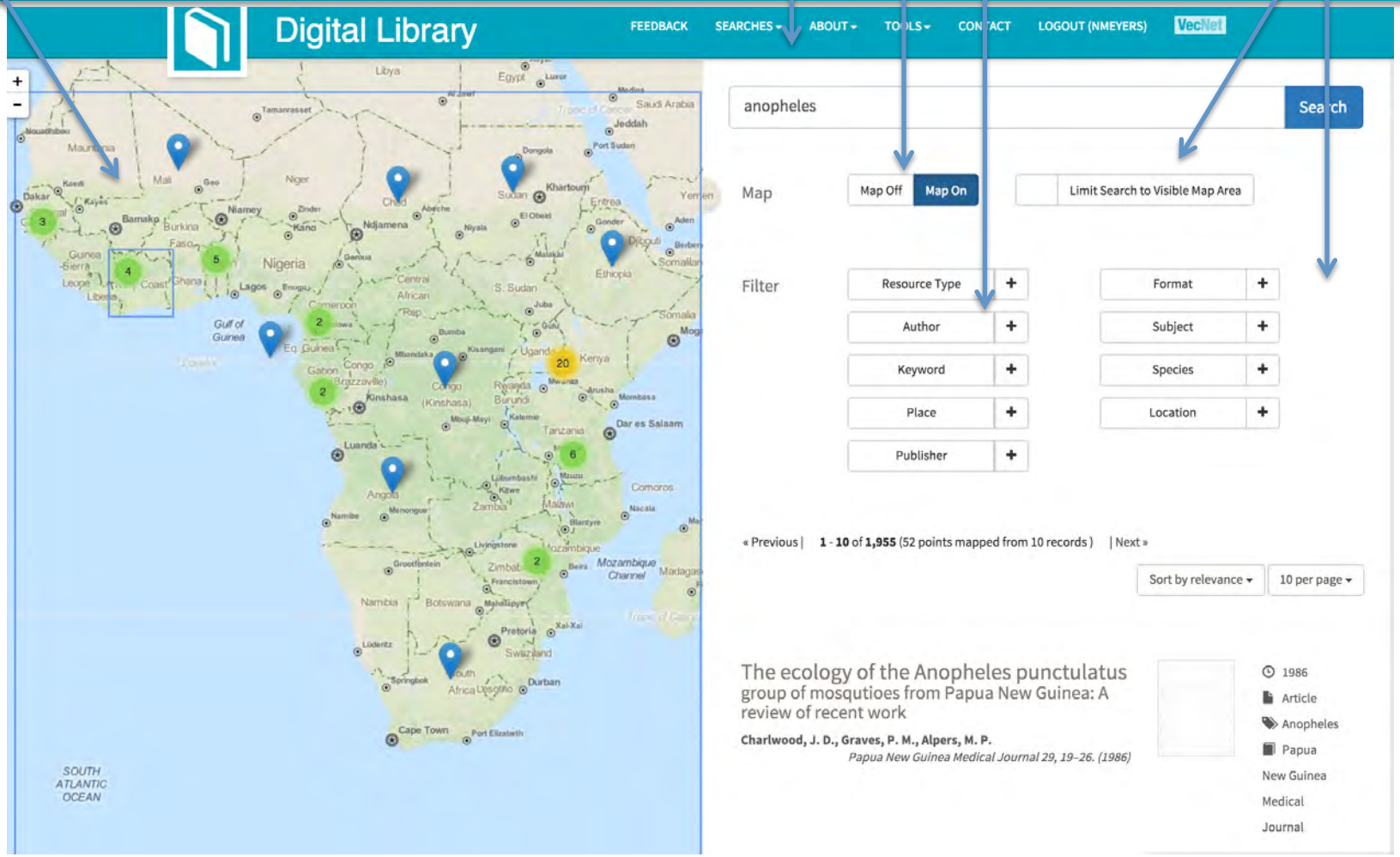
ands
AUSTRALIAN NATIONAL DATA SERVICE

Cartographic Discovery Extension

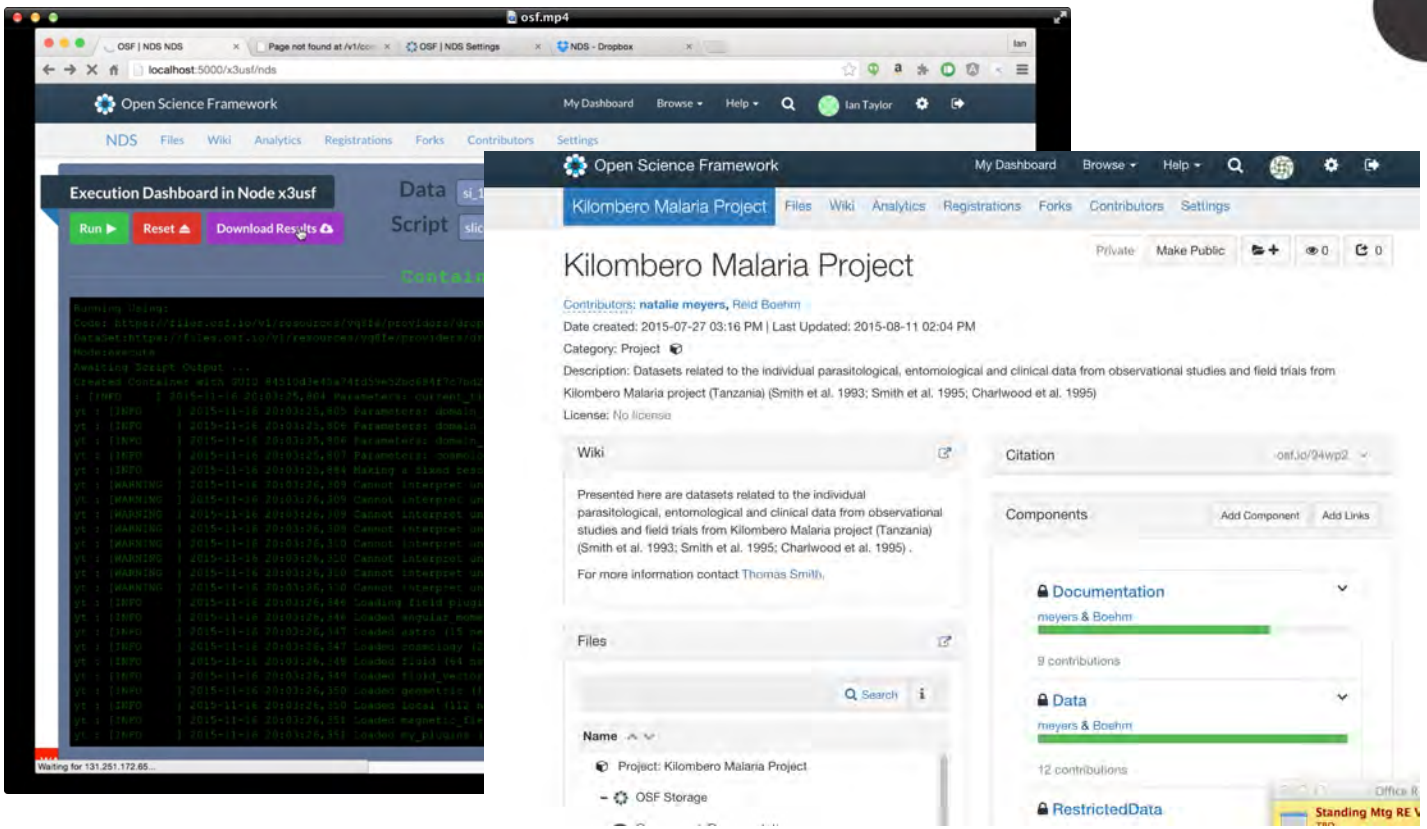
Big searchable map

Logical flow

Simple search filters



OSF Extensions & Pilots



Thanks for your interest in the VecNet Digital Library & data sharing for simulations

https://www.vecnet.org

Natalie K. Meyers
natalie.meyers@nd.edu

JAMES COOK UNIVERSITY AUSTRALIA

UNIVERSITY OF NOTRE DAME

UNIVERSITY OF OXFORD

globalgood Ideas by INTELLECTUAL VENTURES

Swiss TPH
Swiss Tropical and Public Health Institute
Schweizerisches Tropen- und Public Health-Institut
Institut Tropical et de Santé Publique Suisse

R. Farlow Consulting, LLC