Exposing HUB Objects for Aggregation Using OAI-ORE and Linked Data

Michael Witt, mwitt@purdue.edu
Aswathy Sivaram, asivaram@purdue.edu

Distributed Data Curation Center (D2C2)
Purdue University Libraries

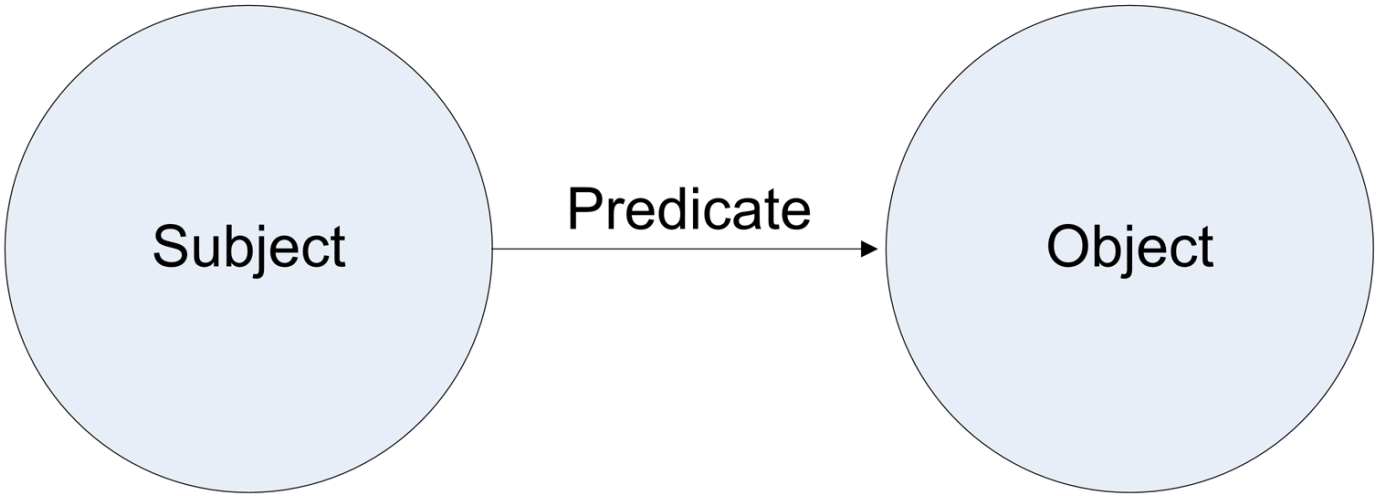
HUBbub 2011, IUPUI, Indianapolis, Indiana: April 6, 2011

Tim Berners-Lee on the next Web

TED2009, Filmed Feb 2009; Posted Mar 2009

11:05 | 16:20 Share Rate

Building blocks: RDF triples

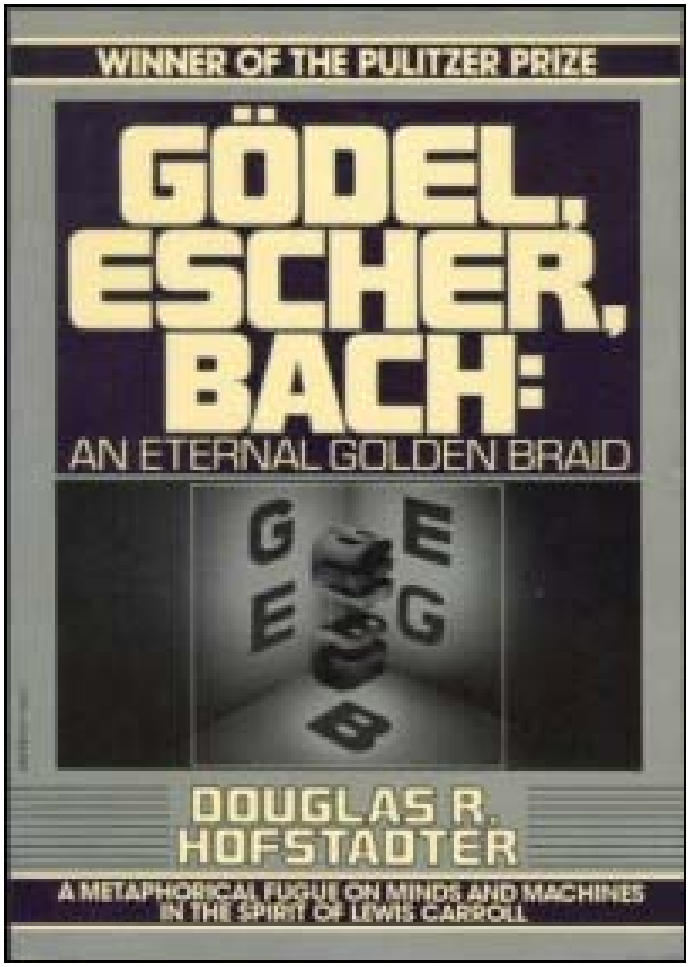


A vocabulary: Dublin Core

- Contributor
- Coverage
- Creator
- Date
- Description
- Format
- Identifier
- Language
- Publisher
- Relation
- Rights
- Source
- Subject
- Title
- Type

Dublin Core Element Set, http://dublincore.org/documents/dces/

One object we may care to describe: a book



GÖDEL,
ESCHER,
BACH:

an Eternal Golden Braid

Library of Congress Cataloging in Publication Data

Hofstadter, Douglas R. (1945-)
Gödel, Escher, Bach: an Eternal Golden Braid

Bibliography: p. 746
Includes index.
1. Philosophy. 2. Metamathematics. 3. Artificial Intelligence. 4. Bach, Johann Sebastian, 1685-1750. 5. Escher, Maurits Cornelis, 1898-1971. 6. Gödel, Kurt, 1906-1978. I. Title.
QA9.8.H63 510'.1 78-19943
ISBN: 0-465-02685-0

Slimy Woodland Salamander at Great Smoky Mountains National Park

Posted 25 Mar 2010 in Images & Videos

About Citations Reviews

Contributor(s) Emily Gore

Abstract Video clip documenting the slimy woodland salamander at the Great Smokies.

Additional Authors Edward Pivorun; Kristen Paysinger; Clemson University

Cite this work Researchers should cite this work as follows:

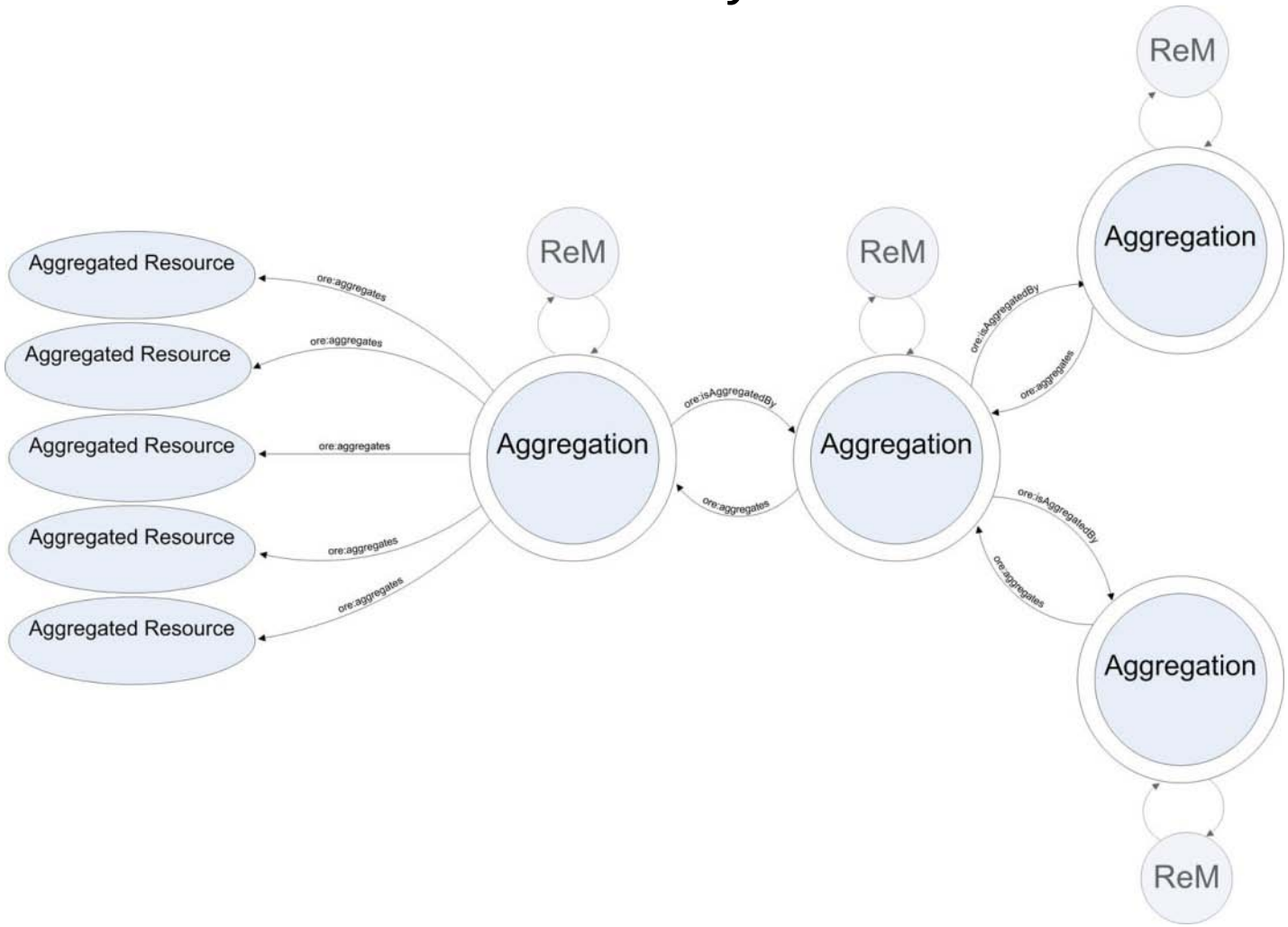
Great Smoky Mountains National Park Interactive CD Series: Salamanders. Produced by Clemson University.

(2010), "Slimy Woodland Salamander at Great Smoky Mountains National Park," https://openparksgrid.org/resources/225.

BibTex EndNote

Tags Great Smoky Mountains National Park | Natural Resource Stewardship and Science | slimy salamander | woodland salamander

Another vocabulary: OAI-ORE



Project Status

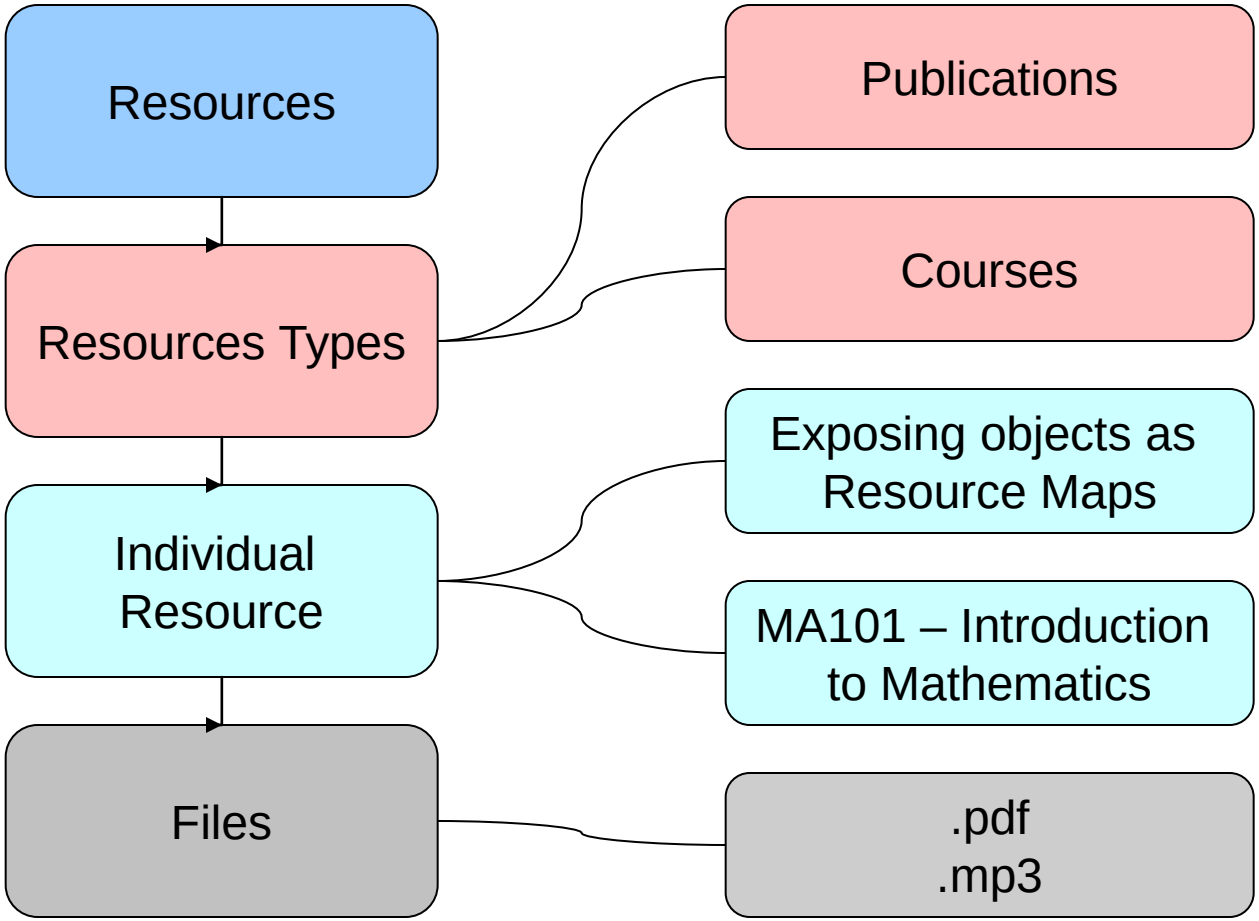
- Identify 'objects' in the HUB
- Create aggregations and expose these objects as ReMs
- Implementation stage
- Three-year project
    - Three months of work
- Work started in Jan 2011

Implementation

HUB contains the content:

- Our focus is on HUB resources
- Has most of the data that makes up the hub in terms of content
- A HUB is a collection of Resources
    - ➔ Resource Types (Seminars, Courses, Tools, etc.)
        - ➔ Individual Resource (Seminar X, Simulation Tool Y…)
            - ➔ Files (.pdf, .pps, .mp3, etc.)

Resources Hierarchy



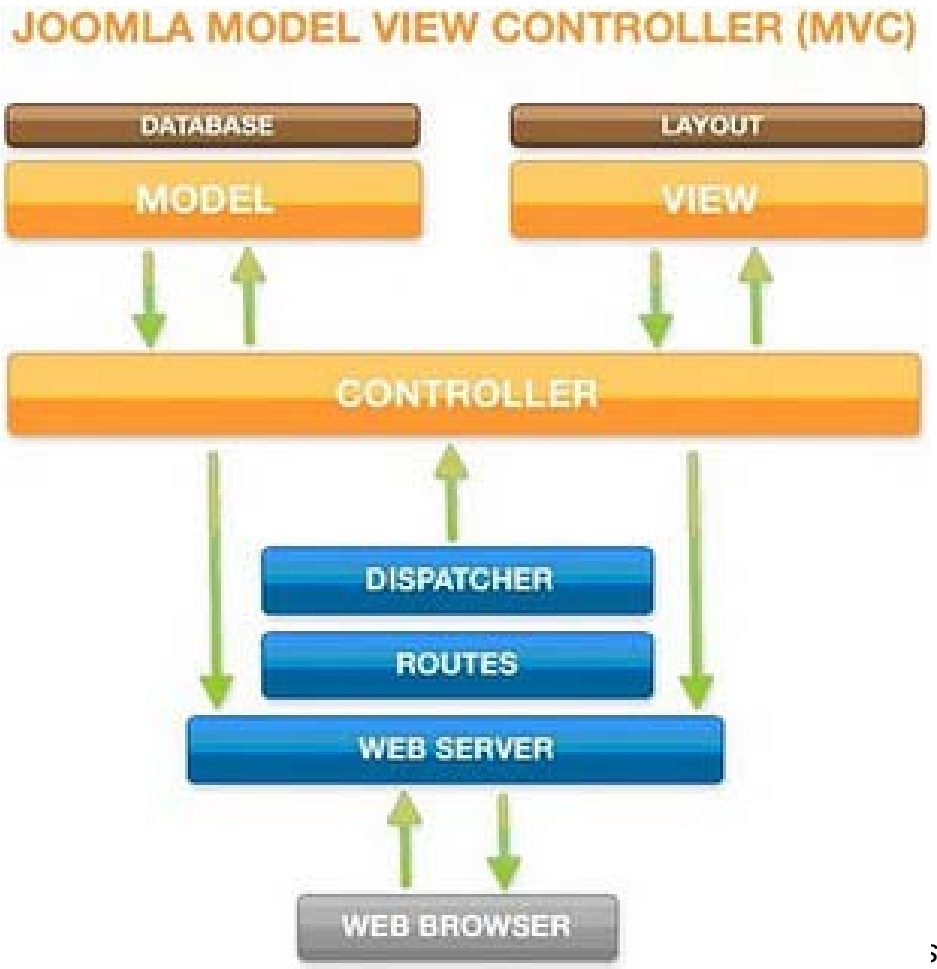
Resource Map

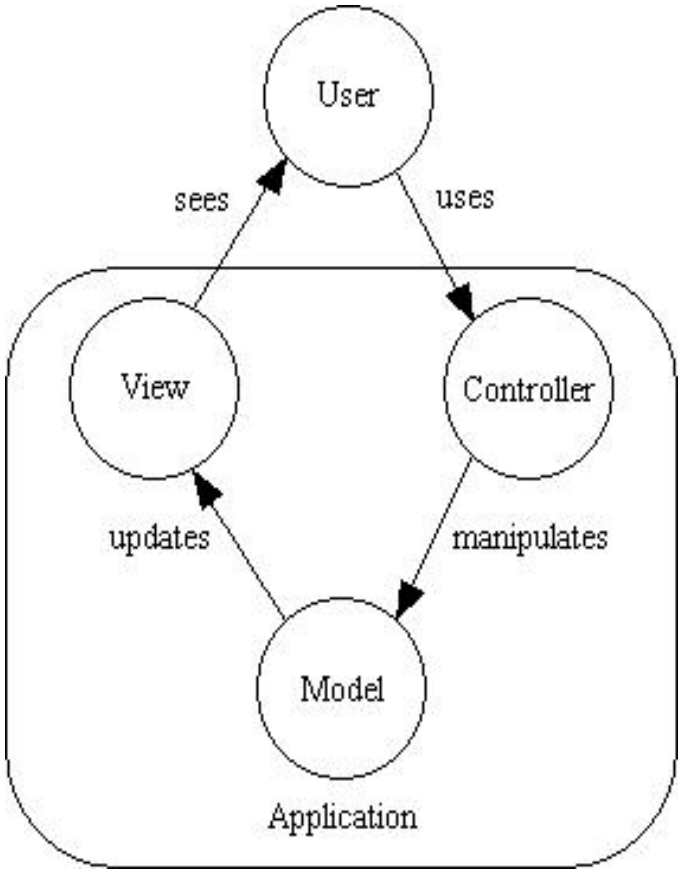
Embedded in the XHTML pages that constitute any resource using RDFa

Example:

- Resource type – Images & Videos
- Slimy Wood Salamander at Smokies
  - Has an author/creator
  - Has a description
  - Has a relation to About, Citations, Reviews
  - Has a video file
- Source Code

Fitting into the Joomla Framework

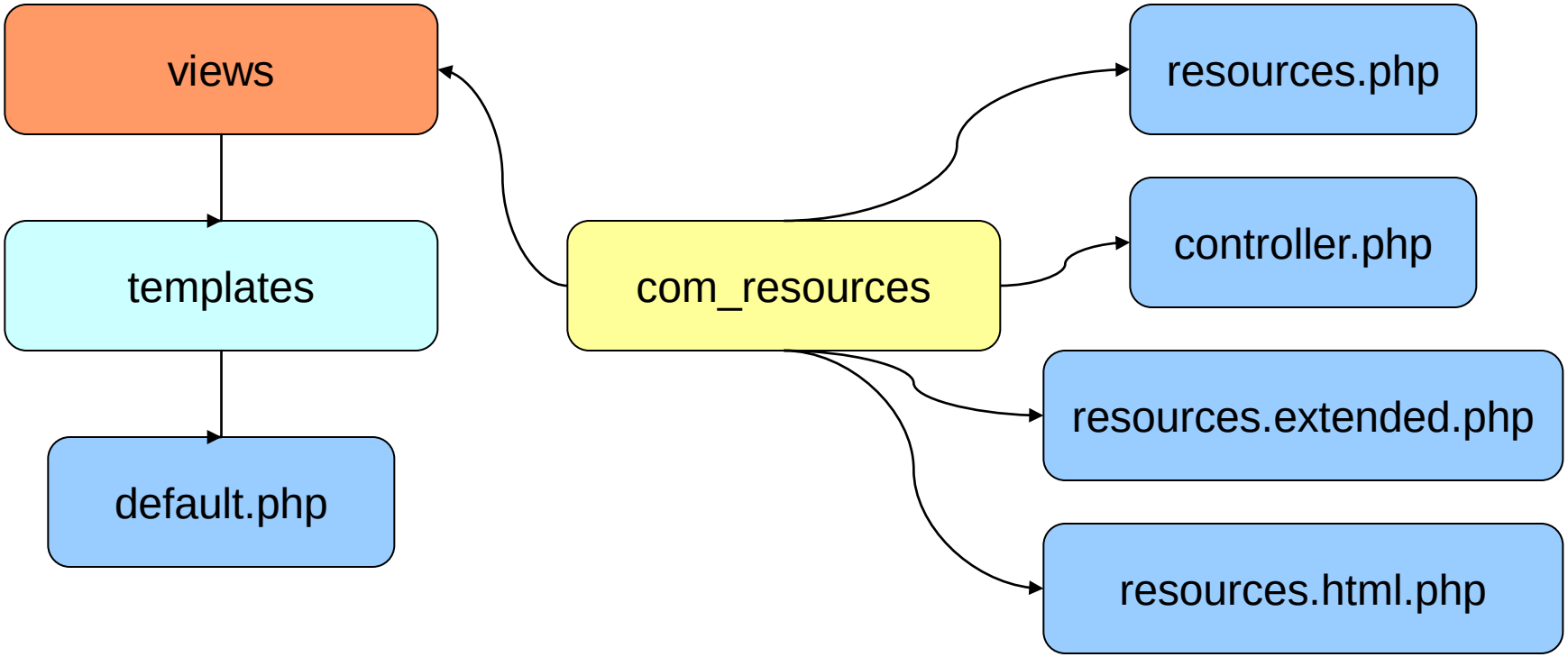


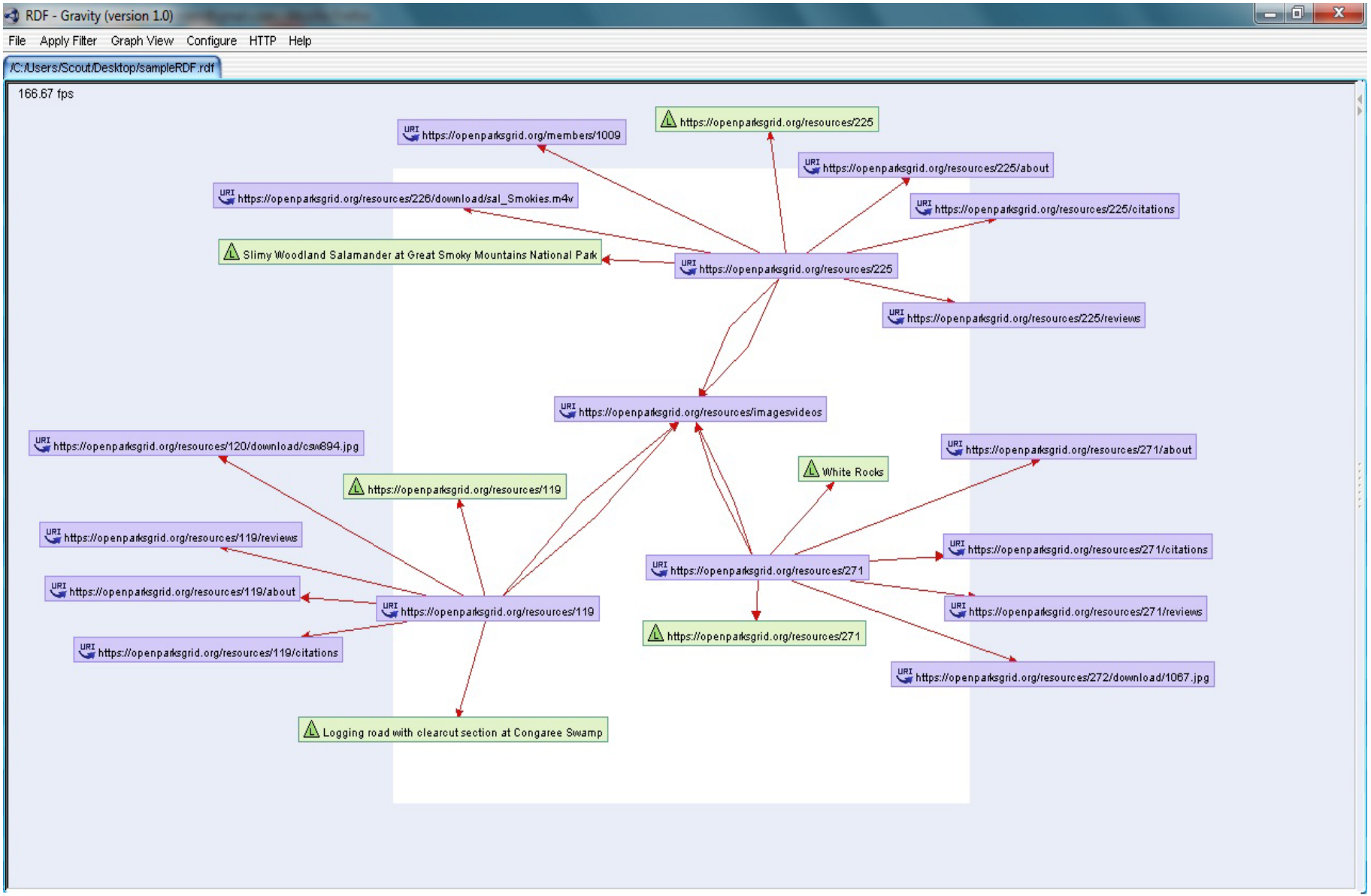


spot.com/2010/06/understanding-joomla-architecture-part1.html

http://www.vojtechovsky.net/joomla/component-helloworld-2-create-tutorial-guide-en.html

Resources Hierarchy




RDF - Gravity (version 1.0)
File
Apply Filter
Graph View
Configure
HTTP
Help
/C:/Users/Scout/Desktop/sampleRDF.rdf
166.67 fps
https://openparksgrid.org/members/1009
https://openparksgrid.org/resources/225
https://openparksgrid.org/resources/225/about
https://openparksgrid.org/resources/226/download/sal_Smokies.m4v
https://openparksgrid.org/resources/225/citations
Slimy Woodland Salamander at Great Smoky Mountains National Park
https://openparksgrid.org/resources/225
https://openparksgrid.org/resources/225/reviews
https://openparksgrid.org/resources/imagesvideos
https://openparksgrid.org/resources/120/download/csw894.jpg
https://openparksgrid.org/resources/271/about
White Rocks
https://openparksgrid.org/resources/119
https://openparksgrid.org/resources/119/reviews
https://openparksgrid.org/resources/271/citations
https://openparksgrid.org/resources/271
https://openparksgrid.org/resources/119/about
https://openparksgrid.org/resources/119
https://openparksgrid.org/resources/271/reviews
https://openparksgrid.org/resources/271
https://openparksgrid.org/resources/119/citations
https://openparksgrid.org/resources/272/download/1067.jpg
Logging road with clearcut section at Congaree Swamp

Why do all this?

For Tim Berners-Lee!

His Linked data principles in three rules:

- All kinds of **conceptual things, they have names now** that start with HTTP
- I get **important information back. I will get back some** data in a **standard format which is kind of useful data** that somebody might like to know about that thing, about that event.
- I get back that information it's not just got somebody's height and weight and when they were born, **it's got relationships.**

Why do all this?

Our goal: to enable a HUB to become a part of the Semantic Web

More specifically, to expose HUB content for as OAI-ORE aggregations and Linked Open Data

Two use cases, to be developed by us:

- To replicate the content, metadata, and basic semantic structure between a HUB and another repository environment (Fedora)
- Create a HUB tool that enables users to create and publish their own custom aggregations, e.g., a “My Collection” tool

Acknowledgements

Supported by IMLS National Leadership Grant, LG-06-07-0032-07.

Principal investigators: Emily Gore (Clemson, lead)
Michael Witt (Purdue) and Elizabeth Baldwin (Clemson)