

Linking Data through Dryad and the HIVE project


International Conference on Dublin Core and Metadata Application
SEOUL, KOREA
Tuesday, October 13, 2009



School of Information and Library Science, Univ. of North Carolina at Chapel Hill

- Jane Greenberg, Professor
- Lina Huang, MSIS Student, Research Assistant
- Robert M. Losee, Professor
- Jose R. Pérez-Agüera, Clinical Assistant Professor
- Hollie White, Metadata Research Center Doctoral Fellow

NESCent (National Evolutionary Synthesis Center)

- Hilmar Lapp
Assistant Director for Informatics
- Ryan Scherle
Data Repository Architect
- Todd Vision, Associate Director of Informatics 

Overview

1. Introduction
2. Scientific data and Dryad
3. Dryad and HIVE: Linked data applications
 - Automatic metadata propagation
 - Dryad's Dublin Core application profile
 - HIVE—Helping Interdisciplinary Vocabulary Engineering
4. Conclusions



Scientific data and Dryad



Dryad's Goals

- One-stop deposition and shopping for data objects supporting published research...
- Support the acquisition, preservation, resource discovery, and reuse of heterogeneous digital datasets

Dryad Team

- National Evolutionary Synthesis Center (NESCent)
- Metadata Research Center at the School of Information and Library Science, University of North Carolina at Chapel Hill (UNC/CH)
- North Carolina State University
- University of New Mexico
- Yale University
- Major societies and journals in the field of evolutionary biology



Dryad and HIVE: Linked data applications



1. Automatic propagation of article citation metadata
2. Metadata generation via Dryad's Dublin Core application profile
3. Vocabulary Integration via HIVE



Citation metadata from a published research article in *Molecular Biology and Evolution*.



Search Dryad

[Advanced Search](#)

[Dryad >](#)
[Main >](#)
[Publications >](#)

Please use this identifier to cite or link to this item: <http://hdl.handle.net/10255/dryad.162>

Title: Compensatory Evolution in RNA Secondary Structures Increases Substitution
Authors: Knies, Jennifer L.
Dang, Kristen K.
Vision, Todd J.
Hoffman, Noah G.
Swanstrom, Ronald
Burch, Christina L.
Issue Date: 2008
Publisher: Oxford University Press
Citation: Compensatory Evolution in RNA Secondary Structures Increases Substitution Rate Variation among Sites. 2008. Knies, Jennifer L., et. al. *Molecular Biology and Evolution*. 25(8):1-10. doi:10.1093/molbev/msn130
Series/Report no.: *Molecular Biology and Evolution* 25(8):1-10
Description: There is growing evidence that interactions between biological molecules (e.g., RNA-RNA, protein-protein, RNA-protein) place limits on the rate and trajectory of molecular evolution. Here, by extending Kimura's model of compensatory evolution at interacting sites, we show that the ratio of transition to transversion substitutions (j) at interacting sites should be equal to the square of the ratio at independent sites. Because transition mutations generally occur at a higher rate than transversions, the model

[Home](#)

Browse

- [Communities & Collections](#)
- [Titles](#)
- [Authors](#)
- [Subjects](#)
- [By Date](#)

Sign on to:

- [Receive email updates](#)
- [My Dryad](#)
authorized users
- [Edit Profile](#)
- [Help](#)
- [About Dryad](#)



Data object metadata for item underlying published research in previous slide



[Dryad](#) >
[Main](#) >
[Data](#) >

Please use this identifier to cite or link to this item: <http://hdl.handle.net/10255/dryad.169>

Title: 16S alignment and tree
Authors: Knies, Jennifer L.
Dang, Kristen K.
Vision, Todd J.
Hoffman, Noah G.
Swanstrom, Ronald
Burch, Christina I

Issue Date: 30-Jul-2008

URI: <http://hdl.handle.net/10255/dryad.169>

Described By Publication: <http://dx.doi.org/10.1093/molbev/msn130>

Appears in Collections: [Data](#)

Files in This Item:

File	Description	Size	Format	
16SandTREE.nex		151.74 kB	Nexus	View/Open



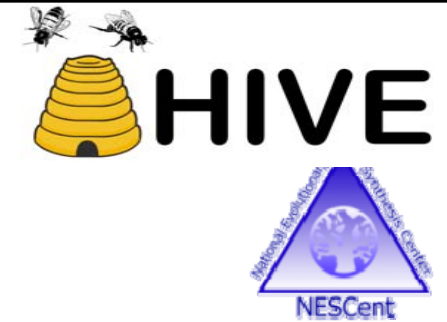
<DRYAD application profile> and the Singapore Framework

- Functional requirements documented, and high level metadata functionalities (Dube et al, 2007, White, et al, 2008)
- Domain model (Carrier, 2007)
- Dryad description set profile, Usage Guidelines, and Encoding guidelines—all on the Dryad Development Wiki
(https://www.nescent.org/wg_dryad/Main_Page)

Greenberg, J., White, H.C., Carrier, S., and Scherle, R. (in press). A Metadata Best Practice for a Scientific Data Repository. *Journal of Library Metadata*. [24 manuscript pages.] [Special issue on metadata best practices] Journal homepage: <http://www.informaworld.com/smpp/title~content=t792306902~db=all>



Helping Interdisciplinary Vocabulary Engineering



- SILS Metadata Research Center and NESCent
- Vocabulary partners: Library of Congress, the US Geological Survey, and the Getty Research Institute.
 - *Library of Congress Subject Headings (LCSH)*
 - *National Biological Information Infrastructure Biocomplexity Thesaurus (NBII)*
 - *Thesaurus of Geographic Names (TGN)*



Helping Interdisciplinary Vocabulary Engineering

- HIVE implements the technological infrastructure to *store millions of concepts from different vocabularies* and make them available on the Web by a simple HTTP call
- Vocabularies can be imported in HIVE using SKOS/RDF format
- HIVE is divided in two different modules:
 - HIVE Core**
 - Automatic Metadata Extraction and Topic Detection (KEA++ and MAUI)
 - Concept Retrieval (Lucene and MG4J)
 - RDF storage and management (SESAME/Elmo)
 - HIVE Web**
 - Web user Interface (GWT)
 - Machine oriented interface (SOAP and REST)





Help with Interdisciplinary Vocabulary Engineering

[Home](#)

[Concept Browser](#)

[Indexing](#)

Current vocabularies sources: [X LCSH](#) [X NBII](#) [X MeSH](#) [+ Add](#)

Searching for Concept

[Search](#)

0-9 [A](#) [B](#) [C](#) [D](#) [E](#) [F](#) [G](#) [H](#) [I](#) [J](#) [K](#) [L](#) [M](#) [N](#) [O](#) [P](#) [Q](#) [R](#) [S](#) [T](#) [U](#) [V](#)
[W](#) [X](#) [Y](#) [Z](#) [À](#) [Á](#) [Â](#) [Ã](#) [Ä](#) [Å](#) [\[All\]](#)

- [LCSH](#)
- [NBII](#)**
- [MeSH](#)
- [aallonpituus kooste](#)
- [aapa mires](#)
- [aardvark](#)
- [abandoned houses](#)
- [abandonment](#)
- [abbreviations](#)
- [abdomen](#)
- [abductions](#)
- [ability](#)
- [ability to write](#)
- [abnormality](#)
- [abortion](#)
- [abrasion resistance](#)
- [abrasion tests](#)
- [absences](#)
- [absent-mindedness](#)
- [absorption](#)
- [absorption](#)
- [absorptive disorders](#)
- [abstract art](#)
- [abstraction of water](#)
- [abstraction of water](#)
- [abstract publications](#)

Retrieved Concepts: ✕

- NBII->Body
- LCSH->Body & Society
- NBII->Body fluids
- NBII->Body conditions
- LCSH->Body, Agnes, 1866-1952
- LCSH->Body & Mind
- NBII->Body cavity
- LCSH->Body and soul in literature
- NBII->Barr body
- NBII->Body
- LCSH->Body & Society
- NBII->Body fluids
- NBII->Body conditions
- LCSH->Body, Agnes, 1866-1952
- LCSH->Body & Mind

[view more...](#)

Current Concept: NBII->body

[View in SKOS](#)

Body	
Preferred Label	Body
Broader Term	No broader term
Narrower Term	Abdomen , Appendages , Attachment organs , Bladders , Body cavity more »
Related Term	Anatomy , Animal morphology , Biomechanics
Scope Notes	LSC Life Sciences
Alternative Label	Animal body Regions , Body regions
URI	Http:// www.nbii.com/body

Context viewer ✕

[+](#) [-](#)
View





National Science Foundation
WHERE DISCOVERIES BEGIN



Conclusion

- Continuing enriching our data descriptions
- Dryad and HIVE are real-world applications using Semantic Web technology

Links

- Dryad
 - <http://datadryad.org/>
- HIVE
 - <http://ils.unc.edu/mrc/hive/>
- Metadata Research Center <MRC>
 - <http://www.ils.unc.edu/mrc/>
- National Evolutionary Synthesis Center (NESCent)
 - <http://www.nescent.org/index.php>

