RHODES UNIVERSITY

INNOVATIVE USERS GROUP
SOUTHERN AFRICA

Our data – our responsibility:

The semantic web, deep linking and repositories

14th IUG-SA Annual Conference 2016
23 November 2016
Vanderbijlpark

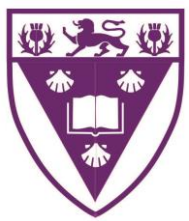
Wynand van der Walt
Head Librarian: Technical Services
Rhodes University

Content

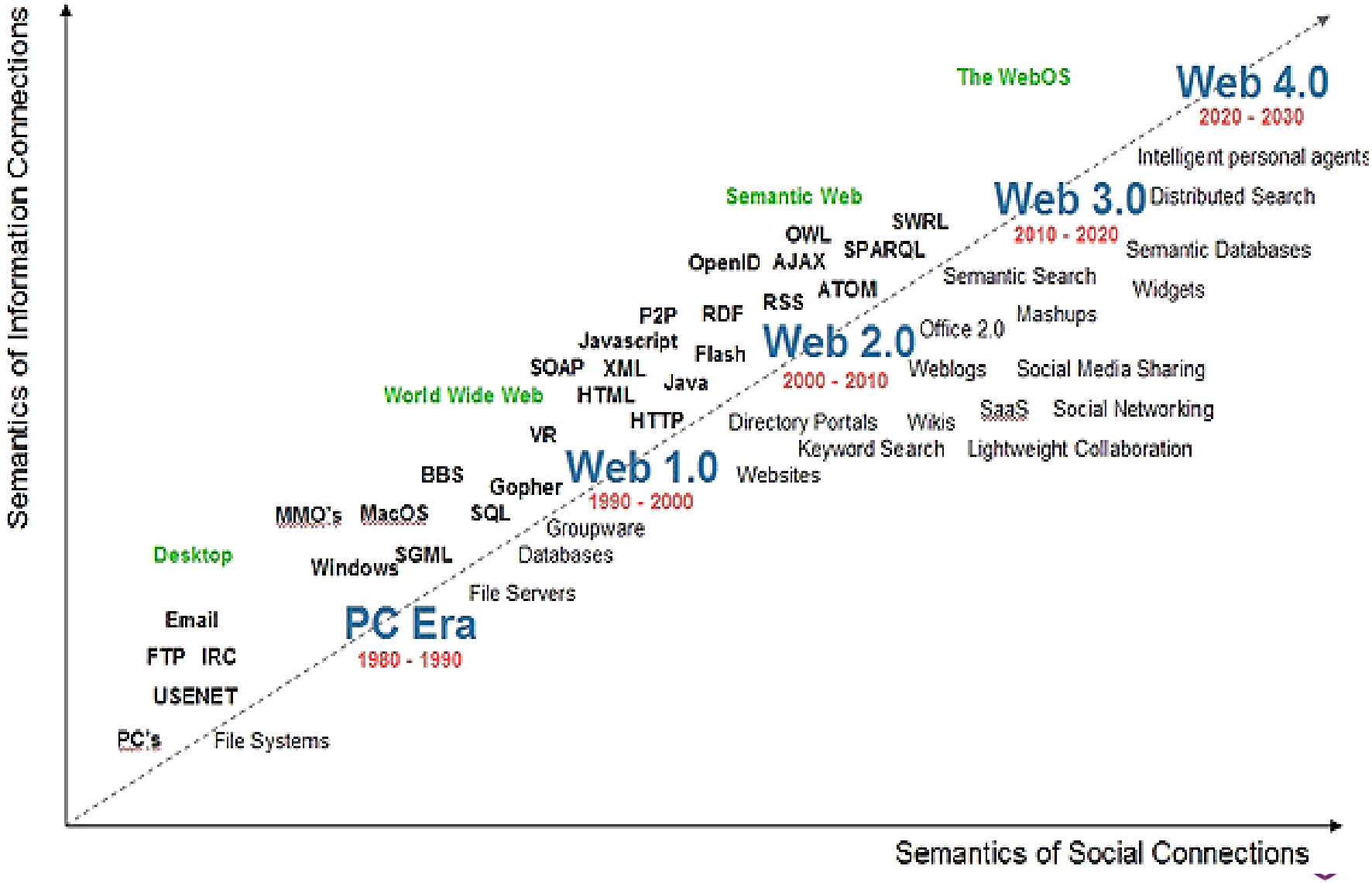
- The semantic web
- Deep linking
- Rethinking repositories and/or rethinking our role/s

Context

- Understand web developments
- Reflect on what we do within the context of web developments
- Explore and embrace new cataloguing forms
- Support interconnectedness of our data
- Develop our ability to optimally express data

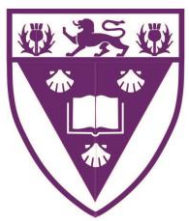


Web developments



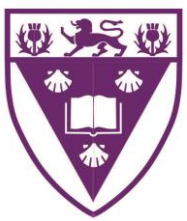
The semantic web - background

- Current web:
  - Fine for humans but not machines processing
  - Lacks semantic abilities
    - Semantic - Indicates meaning
    - Human language
    - Page structure and layout
    - Graphics and multimedia



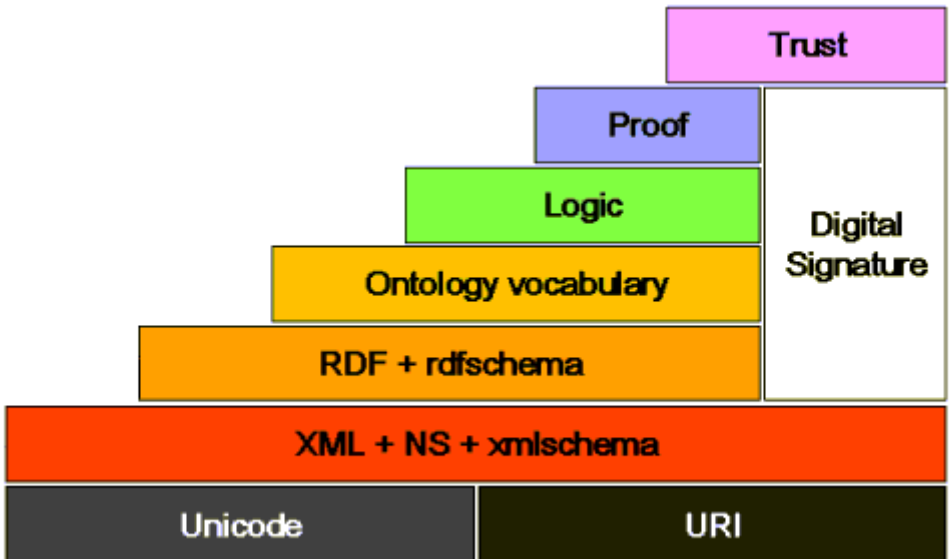
The semantic web - background

- “The past was document sharing, the future is data sharing”- Tim Berners-Lee
- Relating to meaning in language or logic
- Vision:
  - To provide information in a **machine-interpretable** format
  - In order to enable intelligent agents to act on our behalf through interpretation of the relationships between objects



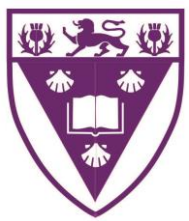
What is the semantic web

- Extension of the web
  - Common framework achieved through standards (notably W3C)
  - These standards promote common data formats and exchange protocols

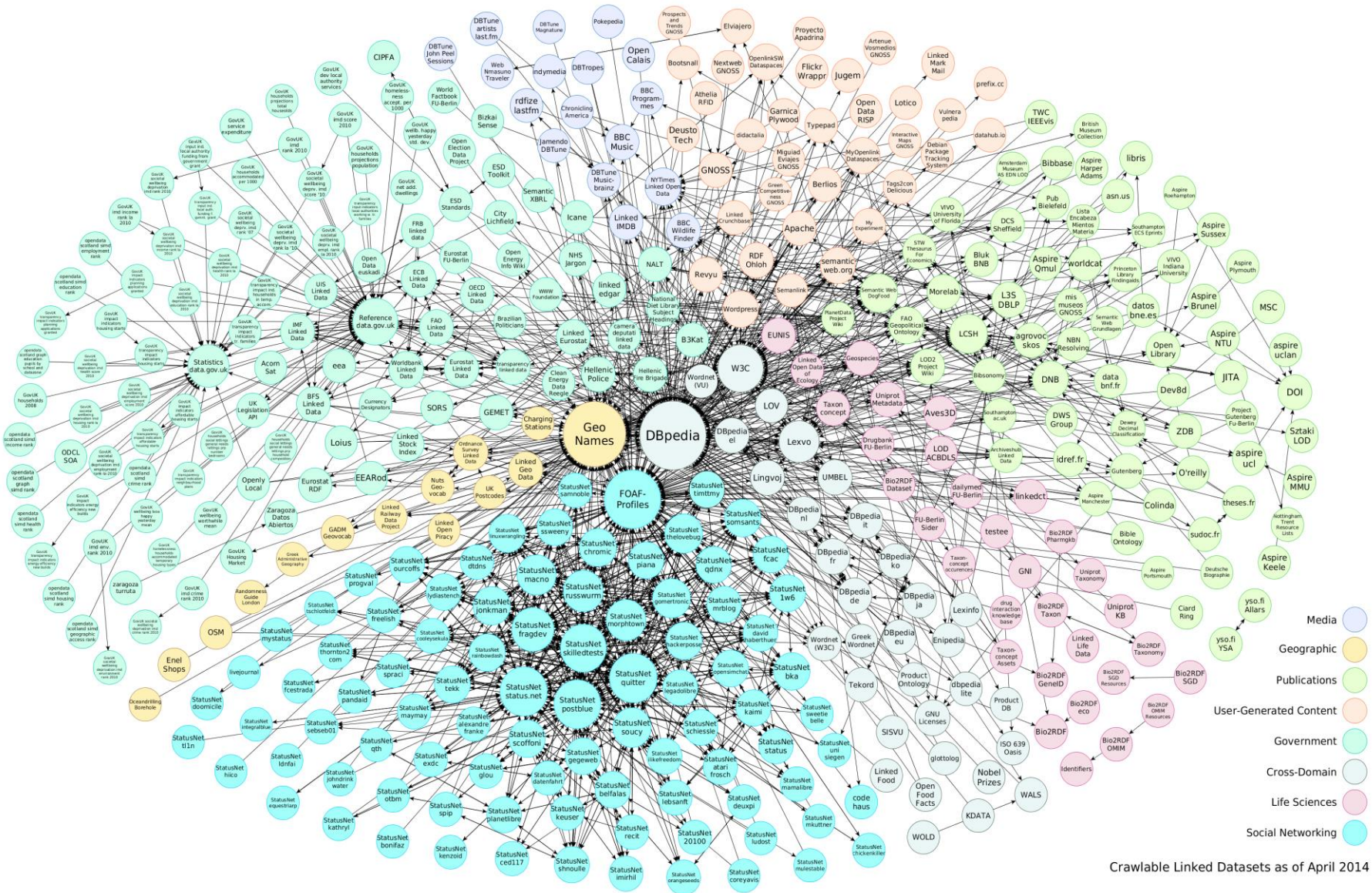


What is the semantic web

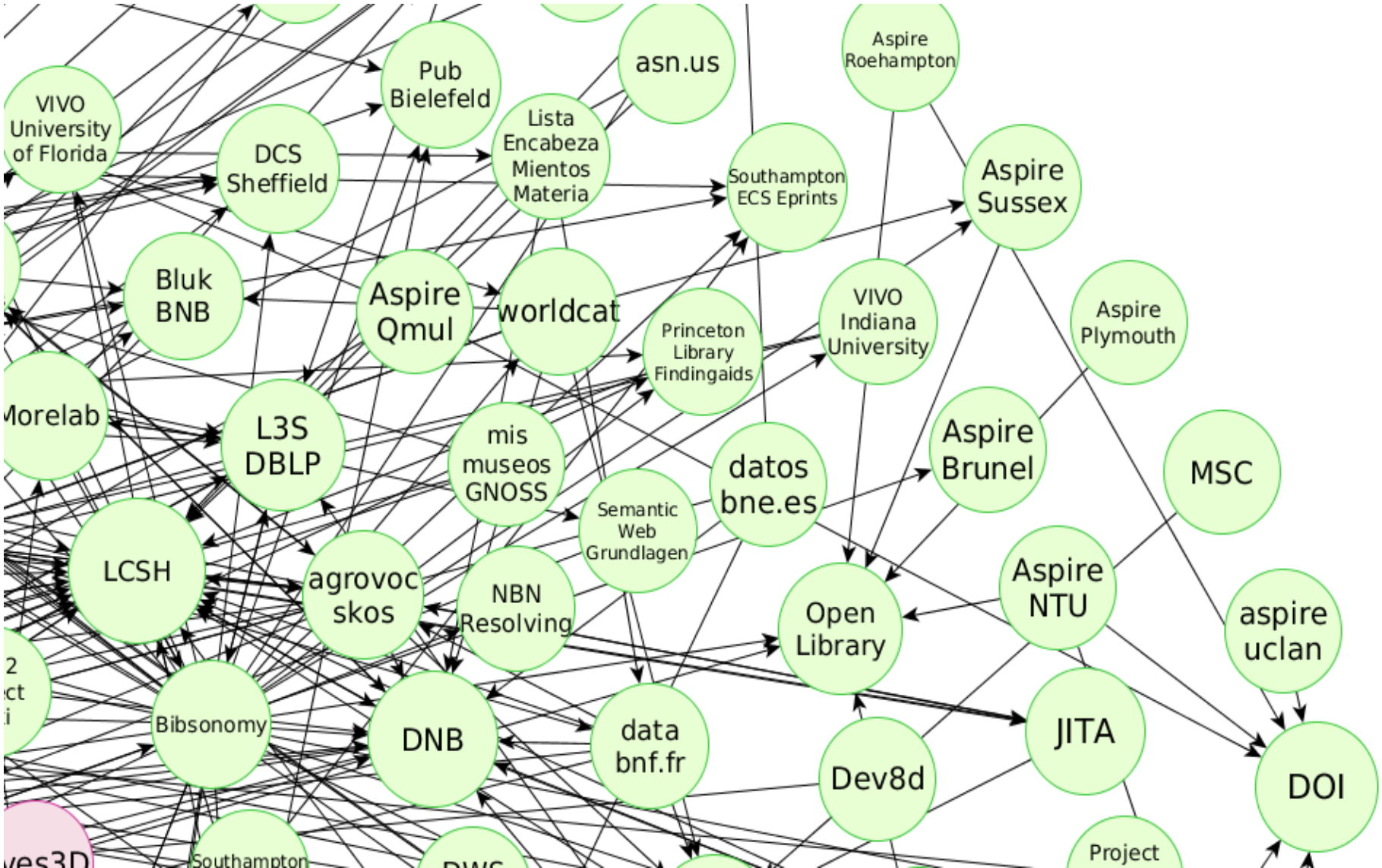
- Fundamental frameworks:
  - Resources Description Framework (RDF)
    - Expresses data models in terms of relations between objects
  - Web Ontology Language (OWL)
    - Describe taxonomies and classification networks
    - E.g. Dublin Core, Resource Description and Access (RDA)
  - Extensible Markup Language (XML)
    - Syntax for content structure
  - SPARQL
    - Protocol and query language
  - RIF
    - Rule Interchange Format for expressing web rules



What does it look like?



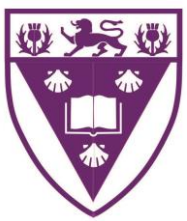
What does it look like?



Deep Linking

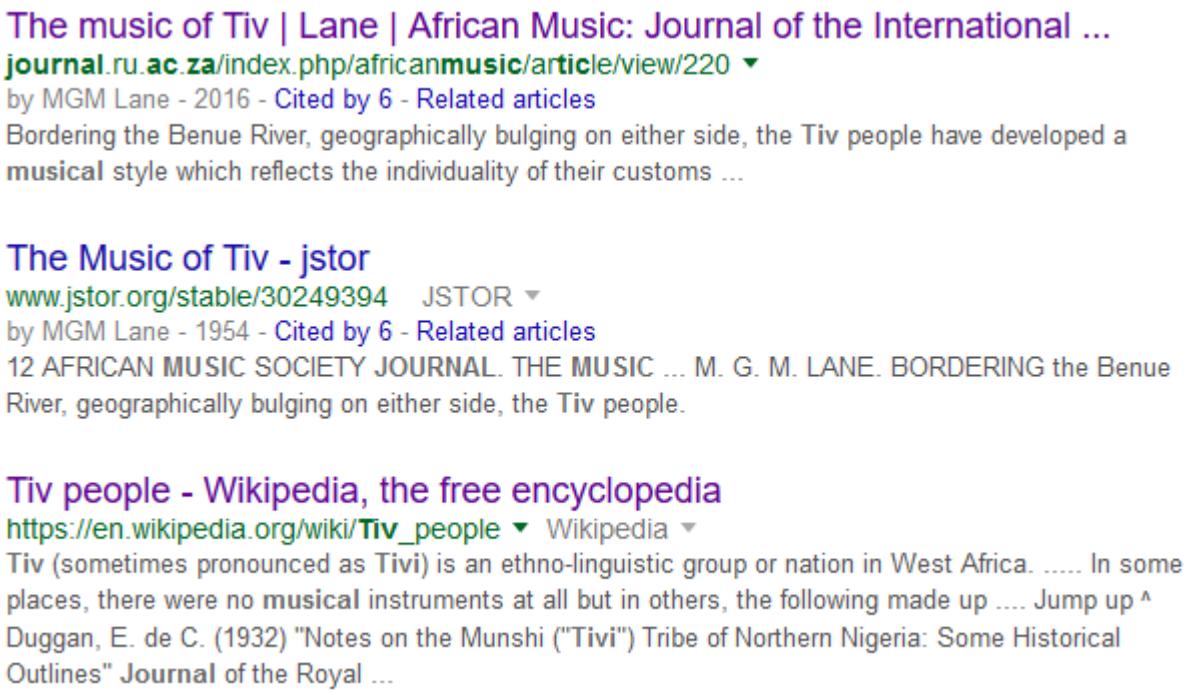
- The use of a hyperlink that links to a specific, generally searchable or indexed, piece of web content on a website
  - (e.g., "http://example.com/path/page"),

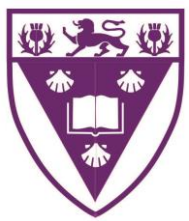
    rather than the website's home page
  - (e.g., "http://example.com/").



Objects & Relationships

- Example:
  - Tiv music





Objects & Relationships

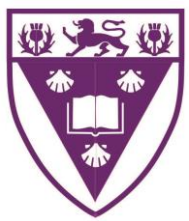
- Example established certain relationships through identifiers:
  - Title to context
  - Author (ORCid) to context
  - ISSN to context
  - **DOI** to context
  - Publisher to context
  - Etc.

Authorities as example

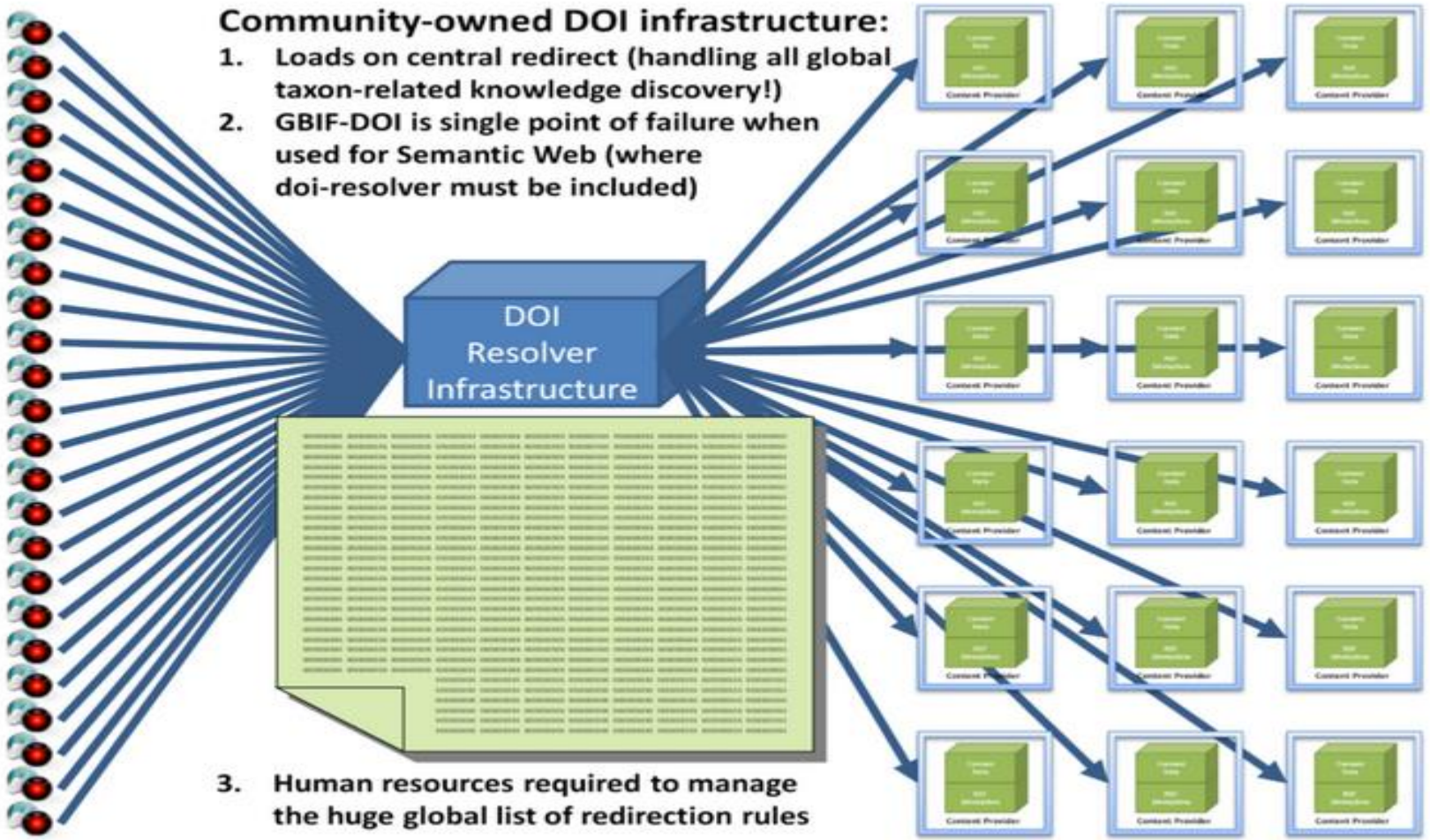
- JLB Smith example:
  - Different “authority” records
- Virtual International Authority File (VIAF)
  - “super” authority record with contributions from
    - German National Library
    - Library of Congress
    - OCLC
    - Other national libraries
    - Others

Identifiers

- Identifiers in the Library and Information Provisioning Sector
  - Barcodes
  - Bibliographic number
  - DDC (Dewey Decimal Classification)
  - ISBN (International Standard Book Number)
  - ISSN (International Standard Serial Number)
  - URL (Uniform Resource Locator)
  - URN (Uniform Resource Names)
  - DOI (Digital Object Identifier)
  - Handle
  - ORCid (Open Researcher and Contributor ID)
  - ResearcherID
  - Patron Codes
  - Etc.



Identifiers



Digital Object Identifiers

- Zootaxa megajournal

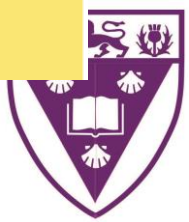
CrossRef reports the year should be "2015" not "2014" in reference "Corbella, Møller, 2014".

Nye V, Copley J, Plouviez S. A new species of *Rimicaris* (Crustacea: Decapoda: Caridea: Alvinocarididae) from hydrothermal vent fields on the Mid-Cayman Spreading Centre, Caribbean. J Mar Biol Assoc U K. 2012;92(05):1057–1072. doi:10.1017/S0025315411002001.

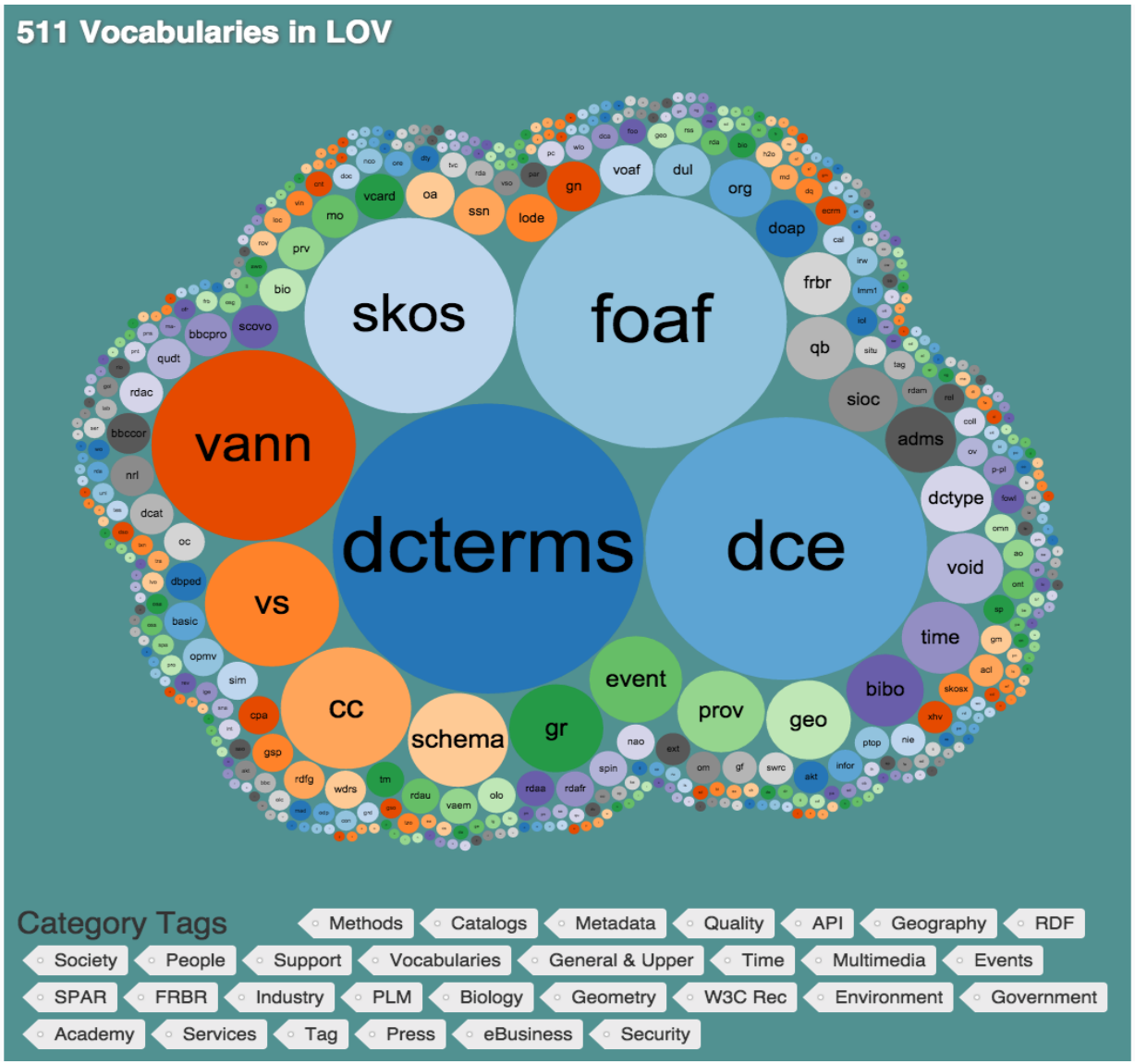
Nye VE, Copley JT, Plouviez S, Van Dover CL. A new species of *Lebbeus* (Crustacea: Decapoda: Caridea: Hippolytidae) from the Von Damm Vent Field, Caribbean Sea. J Mar Biol Assoc U K. 2013;93(03):741–751. doi:10.1017/S0025315412000884.

Nye VE, Copley JT, Linse K, Plouviez S. *Iheyaspira bathycodon* new species (Vetigastropoda: Trochoidea: Turbinidae: Skeneinae) from the Von Damm Vent Field, Mid-Cayman Spreading Centre, Caribbean. J Mar Biol Assoc U K. 2013;94(04):1017–1024. doi:10.1017/S0025315412000823.

CrossRef reports the volume should be "93" not "94" in reference "Nye, Copley, Linse, Plouviez, 2013".



Linked Open Vocabularies (LOV)



Latest insertion

sor - SORON: Social Relationships ONtology
2015-05-19

mexcore - MEX Core Vocabulary
2015-05-18

mexperf - MEX Performance Ontology
2015-05-18

dqc - The Data Quality Constraints Library
2015-05-17

qb4o - Vocabulary for publishing OLAP data cubes
2015-05-17

Latest Updates

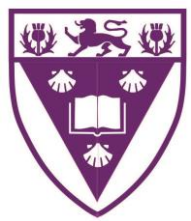
saws - Sharing Ancient Wisdoms Ontology
2015-06-02

vann - VANN: A vocabulary for annotating vocabulary descriptions
2015-06-02

rr - RDB to RDF Mapping Language Schema
2015-06-02

osspr - Spatial Relations Ontology
2015-06-02

bibo - The Bibliographic Ontology
2015-06-02



Linked Open Vocabularies (LOV) (the semantic glue)

arch - Archival collections ontology
http://purl.org/archival/vocab/arch
An RDF vocabulary for describing archival collections and the names associated with them @en

bf - BIBFRAME Vocabulary
http://bibframe.org/vocab
The BIBFRAME model and vocabulary consider resources that are cataloged as works with corresponding instances (physical and/or electronic). The metadata describing a work is an amalgamation of the some of the data that was formerly associated with the uniform title authority record combined with subject data that was associated with the bibliographic records. Thus each cataloging resource must have a work description, and, if the cataloging resource exists physically or electronically, one or more instance descriptions. @en

bibo - The Bibliographic Ontology
http://purl.org/ontology/bibo/
The Bibliographic Ontology Specification provides main concepts and properties for describing citations and bibliographic references (i.e. quotes, books, articles, etc) on the Semantic Web. @en

bibtex - BibTeX ontology
http://purl.org/net/nknouf/ns/bibtex
Transformation of bibTeX into an OWL ontology @en

blt - British Library Terms RDF schema
http://www.bl.uk/schemas/bibliographic/blterms

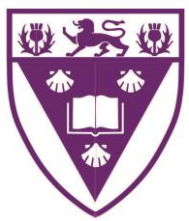
DBpedia


DBpedia
Towards a Public Data Infrastructure for a Large,
Multilingual, Semantic Knowledge Graph

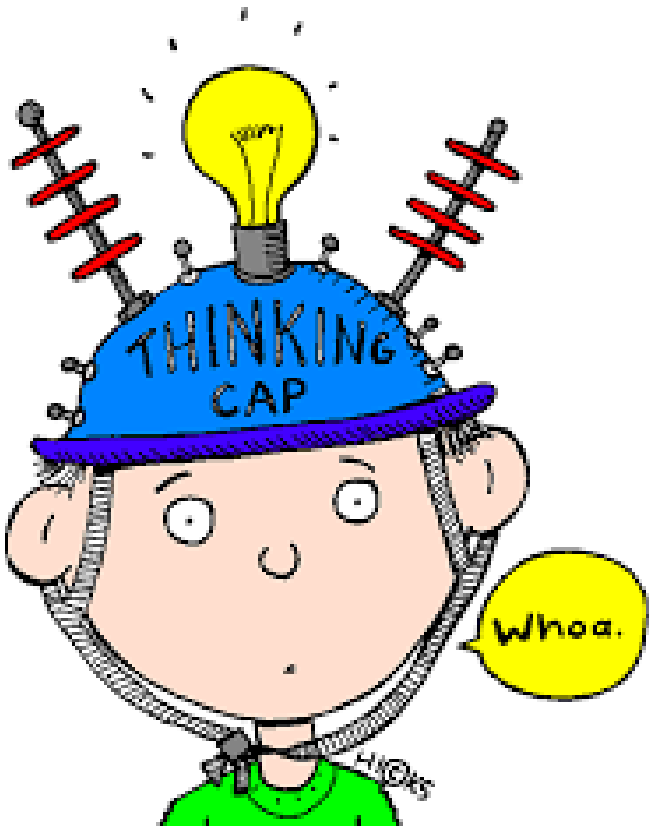
DBpedia

The English version of the DBpedia knowledge base currently describes 4.58 million things, out of which 4.22 million are classified in a consistent ontology (http://wiki.dbpedia.org/Ontology2014), including :

1,445,000 persons, 735,000 places (including 478,000 populated places), 411,000 creative works (including 123,000 music albums, 87,000 films and 19,000 video games), 241,000 organizations (including 58,000 companies and 49,000 educational institutions), 251,000 species and 6,000 diseases.



Where are we in this?

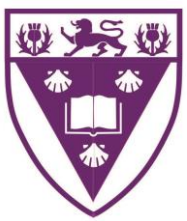


Point of Reflection

- Reflect:
  - Repositories expose & enhance accessibility to
    - Primary and secondary resources
    - Research output and
    - Research materials
- But
  - Data deluge
  - Linkrot
  - Crowdsourced information becoming more authoritative

Key role changes?

- Subject specialists
- Metadata specialists
- Repository managers
- Systems Librarians
- Authors
- Publisher
- Etc.



Key role changes?

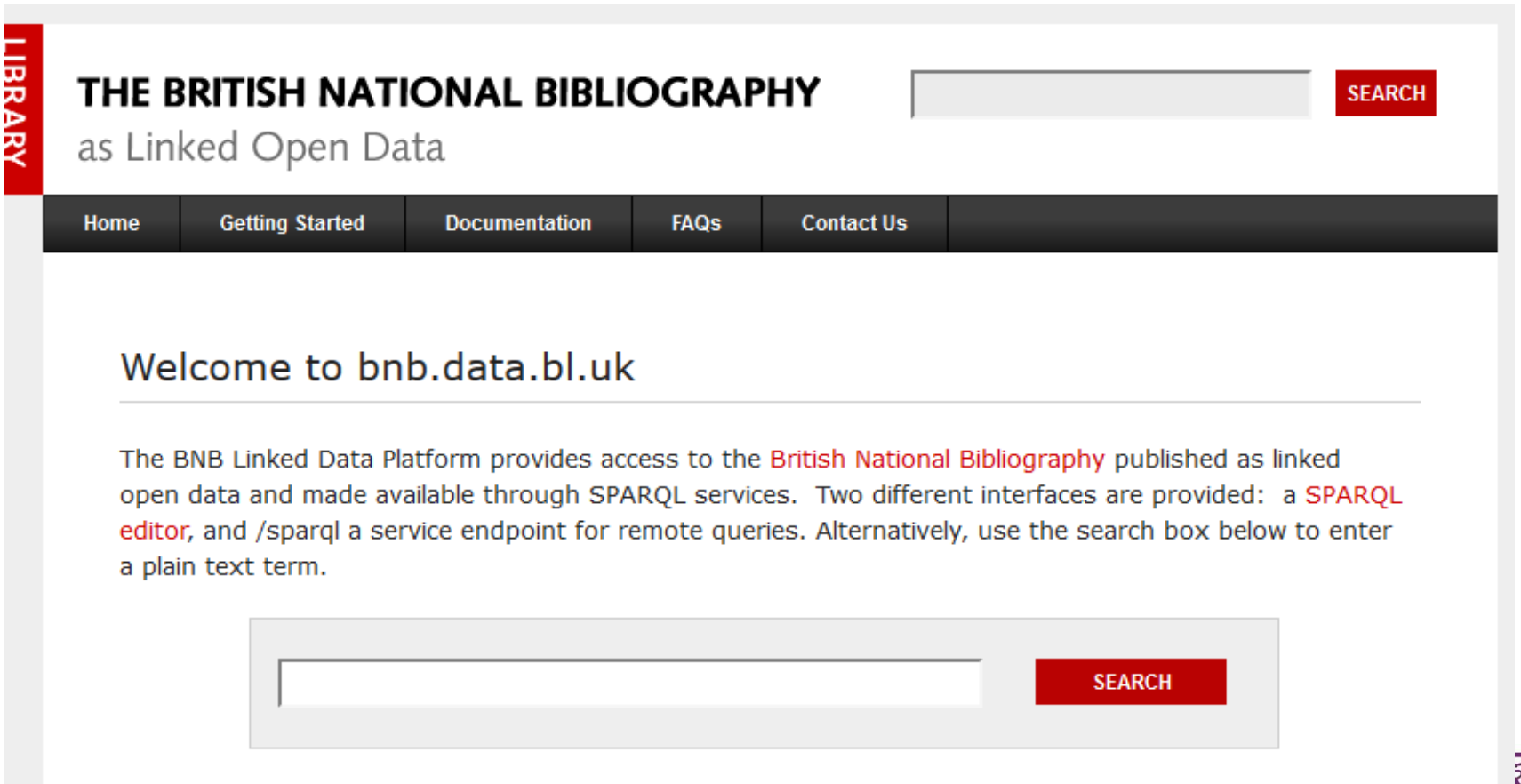
- Subject specialists
- Metadata specialists
- Repository managers
- Systems Librarians
- Authors
- Publisher
- Etc.

Knowledge Engineers

What is required?

- Mind shift from macro-bibliographic to micro-bibliographic and content expression
- Mind shift from micro-community to world-community
- Increased use and development of taxonomies and ontologies
- Increased understanding and use of underlying technologies
- Participate in and enable crowdsourcing

Enquire and learn



References

- Image: Semantics of Information Connections
  - Radar Networks & Nov Spivack, 2007. Available: https://blog.law.cornell.edu/voxpop/files/2010/02/radarnetworkstowardsawebos.jpg
- Image: Linked data cloud 2014
  - https://blogstats.wordpress.com/category/023-semantic-web/
- Image: Self-reflection
  - Factorlab.com. Self reflection – a simple way to improve performance. Available: https://factorlab.com/self-reflection-simple-way-improve-performance/
- Image: Semantic Web Tower
  - W3C. The Smeantic web made easy. Available: https://www.w3.org/RDF/Metalog/docs/sw-easy
- Deep linking. Wikipedia entry. Available: https://en.wikipedia.org/wiki/Deep_linking
- Dbpedia. Facts and Figures. Available: http://wiki.dbpedia.org/about/about-dbpedia/facts-figures
- Dority, Kim. LinkedIN post. For those interested in non-LIS taxonomy applications (i.e., jobs). Available: https://www.linkedin.com/groups/3126663/3126663-6138056180787933188
- Open Knowledge Foundation. Linked Open Vocabularies. Available: http://lov.okfn.org/dataset/lov/
- DOI or LOD or DOI and LOD. Available: http://wiki.pro-ibiosphere.eu/wiki/DOI_or_LOD_or_DOI_and_LOD
- W3C. Home Page. Available: https://www.w3.org/