

Session 1: [Advancing Data Integration approaches of the structured data web](#)

Attendees: please add your details to the table below

Sylvain Grellet, Abdelfettah Feliachi, BRGM, France - ['Linked data' the glue within interoperable information systems](#)

Adam Leadbetter, Rob Thomas, Marine Institute, Ireland - [Using RDF Data Cubes for data visualization: an Irish pilot study for publishing environmental data to the semantic web](#)

Question: How does CubiQL compare to GraphQL-LD?

I wasn't aware of GraphQL-LD. From a quick search it looks like parallel development but I wouldn't be in a place to compare at the moment. Thanks for highlighting, will investigate.

Follow up: the example links on the [GitHub repo](#) return 502 Bad Gateway. Is there another location the examples are hosted or a plan to get this service back online anytime soon?

<https://github.com/Swirrl/cubiql> link worked for me just now

None of the **example links** are working, e.g.,

<http://graphql-qb.publishmydata.com/index.html?query=%7Bcubiql%7B%0A%20%20datasets%20%7B%0A%20%20%20%20uri%0A%20%20%20%20%20title%0A%20%20%20%20description%0A%20%20%7D%20%0A%7D%7D>

Sorry, misunderstood. I'll email the partners and inform them.

Boyan Brodaric, Eric Boisvert, Geological Survey of Canada, Canada; David Blodgett, USGS, USA - [Toward a Linked Water Data Infrastructure for North America](#)

Question: Do you have a specific pattern for URIs for vocabulary terms?

EB:What do you mean by “Vocabulary”, terms as string/words, or concepts/ontology ? Ontologies - similar how you have uris for other 4 types of uri patterns.

Boyan would be the best person to explain at length, but the URL patterns are really conveniences and don't have function in the architecture (they are opaque).

BB: we at NRCan are exploring a specific url structure for each url type. To date we have been using this structure primarily for non-ontology urls. They could be applied to urls for ontology contents, but we haven't tried that, given that there are often different groups governing the ontologies and the data network. Also, as Eric says above, the url structure is a convenience for understanding more about the related resource.

Dalia Varanka USGS, USA - [The Map as Knowledge Base: Integrating Linked Open Topographic Data from The National Map of the U.S. Geological Survey](#)

Alistair Ritchie, Landcare Research, New Zealand - [ELFIE at Landcare Research, New Zealand](#)

Irina Bastrakova, Geoscience Australia, [Australia - Location Index Project \(Loc-I\) – integration of data on people, business & the environment](#)

Jonathan Yu, CSIRO, Australia - [Building Loc-I: Spinning up a spatial knowledge graph should be easy... right?](#)

William Francis, Geoscience Australia, Australia - [Geoscience Australia's structured data on the web: PID governance, resolution, profiles and relationships](#)

Ilya Zaslavsky, Lucila Ohno-Machado, Peter Rose, Hua Xu, Kirk Roberts, Joe Hamman, George Alter, David Valentine, SDSC, USA - [Towards a knowledge-based query system for semantically-enhanced biomedical and geoscience data](#)

Attendees

Name	Organisation	email	Session 1	Session 2
Irina Bastrakova	Geoscience Australia	irina.bastrakova@ga.gov.au	Y	Y
Eric Boisvert	Geological Survey of Canada	eric.boisvert2@canada.ca	Y (remote)	
Alistair Ritchie	Landcare Research NZL	ritchiea@landcareresearch.co.nz	Y	Y
David Blodgett	USGS	dblodgett@usgs.gov	Y	Y
Jonathan Yu	CSIRO Australia	jonathan.yu@csiro.au	Y	Y
Rob Thomas	Irish Marine Institute	rob.thomas@marine.ie	Y (remote)	
Sylvain Grellet	BRGM (French Geological Survey)	s.grellet@brgm.fr	Y (remote)	Y (remote)
Boyan Brodaric	NRCan	boyan.brodaric@canada.ca	Y (remote)	
Joan Damerow	Lawrence Berkeley National Lab	JoanDamerow@lbl.gov	Y	
Alex Kerney	Gulf Of Maine Research Institute	akerney@gmri.org	Y	
Martin Seul	CUAHSI	mseul@cuahsi.org	Y	
Jacqueline Le Moigne	NASA ESTO	Jacqueline.LeMoigne@nasa.gov	Y	
Ethan Davis	UCAR Unidata	edavis@ucar.edu	Y	

Dalia Varanka	U.S. Geological Survey	dvaranka@usgs.gov	Y (remote)	
Nick Jarboe	UCSD	njarboe@ucsd.edu	Y	Y
Steve Aulenbach	USGS	saulenbach@usgs.gov	Y (remote)	Y (remote)
Mike Bobak	NCSA	mbobak@illinois.edu	Y (remote)	Y (remote)
Mike Flaxman	Geodesign Technologies	mflaxman@geodesigntech.com	Y	Y
David Harris	California Natural Resources Agency	David.Harris@Resources.ca.gov	Y (remote)	
Andrew Macleod	Federation University Australia (CeRDI)	a.macleod@federation.edu.au	Y (remote)	
Will Francis	Geoscience Australia	william.francis@ga.gov.au	Y	Y
Nancy Wiegand	U of Wisconsin-Madison	wiegand@cs.wisc.edu	Y (remote)	
Chuck Heazel	OGC	cheazel@heazelttech.com	Y	
Peter Baumann	Jacobs University	baumann@rasdaman.com		

Session 2: [Structured data web and coverages integration working session](#)

Attendees: please add your details to the table above

Agenda:

Introduction

Pitches

Breakouts

Report Back

Introduction:

This session is a time to get together and move the ball forward on a shared problem.

Can we foster some cross-project consistency?

Can we set the stage to solve some problems consistently?

Opportunity to revisit on the IT&I Webinars in the next several months.

Outcomes should be tangible and aim to (for example):

- 1) Establish mutually beneficial collaborations.
- 2) Create issues or other documentation of shared goals.
- 3) Pose questions to be answered in future engagements.
- 4) Plan to propose follow on collaborations.

Potential Topics:

- 1) Testing linked data systems for interoperability and/or consistent behavior.
- 2) Information resource content variability (alternate views, conneg by profile, etc.)
- 3) Links to "in-band" linked (meta)data and "out-of-band" linked (meta) data.
- 4) Incorporating spatial coverages into linked data / knowledge graphs.

Breakout Ideas:

- Linked data system federation (convenience api, linked data, ?QL)
- Content negotiation and in-band and out of band links
- Incorporating spatial coverages into linked data / knowledge graphs.
- How to incorporate other database systems in graph queries / what is role of 'spatial' knowledge graph in relation to arbitrary data indexes, data services, other database backends?

Breakout Notes:

Linked Data System Federation

- Initial Discussion:
- Linked data system federation (convenience api, linked data, ?QL)
- Should everyone export files in common format?
- Common API that everyone exposes?
- VOID dataset description - directory of
- SKOS to link / disambiguate / map IDs
- Trying to walk a JSON-LD graph is slow, 21M triple graphs
- Issues with standing up APIs and keeping them up. Too much to ask of providers.
- Recommended approaches
 - To get started: One-time static dumps of repos
 - Eventually: Continuous harvesting
 - Allow push/publication model as well
 - Who gets to update/publish?
 - Different levels of checks for different submissions
 - Institutional submissions, tightly controlled
 - Voluntary submission, not so much
 - What formats to publish in?
 - JSON-LD, RDF XML, Turtle
 - What API for consuming this data
 - SPARQL (but...be careful)
 - And: what's *your* favorite graph query language?
 - (GREMLIN says Doug)
 - How to provide DOS protections for graph queries?
 - Does registered user access only help?
 - What to do for the rest then?
 - <https://data.geolink.org> - demo of federated graph query from EarthCube
 - SPARQL endpoint
 - Identifiers: strongly encourage (require?) common identifiers:
 - Orcids (not so good for historical folks--ISNI)
 - [ROR](#) (but how to handle non-research orgs?)
- N.B.: Determine beforehand whether non-public data will be collected/published!

Incorporating Spatial Coverages in Knowledge Graphs

- Do we actually have this problem? Why can't we just do the operation on the fly?

- Because it doesn't scale well.
- Seems like the DGGs as an intermediate technology solves some of the scalability problem.
- Are there ways to update a graph rapidly?
- What about a static dataset?
 - If static, you can have a geospatial workflow that generates and inserts data.
- Level 1: Just point to the coverage
- Level 1.5: Work with a static file in an "out of band" process.
- Level 2: Reproducible GIS workflow that modifies the graph + provenance
- Level 3: Automated / real time updates
 - Requires a high performance graph editing worker.
- Alternate approach to embed a service call to get some coverage data
 - Would require some resolver that could follow the links and use them.
 - OGC-API could serve an RDF form of a coverage query.
 - Is more or less in CIS 1.1, see <http://schemas.opengis.net/cis/1.1/rdf/> (feel free to contact Peter Baumann, baumann@rasdaman.com, for questions on CIS)
 - Could we have neighbor cells available in the OGC-API
- .
- How does this kind of data show up in a landing page or visualization?
 - Level 1 is required if you want to show the coverage as a map.
 - Level 2 is categorization that might tell people which coverages would be interesting to look at.
 - Level 3: might be saved metrics over time that give you baseline stats about coverages for later reference.
 - Alternate approach is a bit limited unless it's a real time dashboard or data preview.
- Could also model a coverage (or disaggregated) as linked data?
 - Could there be sub-coverages that get put into linked data directly?
 - Quad tree type disaggregation would work well.
 - A lot in common with cloud optimized geotiff
 - Are there convenience APIs that would help here? OGC-API?
- Next Steps?
 - Need to explore more on tessellations as an intermediate index.
 - Will carry forward some of these ideas at the EDR SWG
 - Will represent some of these ideas to the OGC-API Coverages SWG
 - Will mention these ideas to the UFOKN

Role of 'spatial' knowledge graphs

- Will spatial data analysis and transformation tools grow to adopt/support RDF as an underlying data structure for spatial information or will RDF continue to be a 'view' of existing (legacy) spatial data in GI systems?
- The BIM (building information modeling) world has set an example for fully adopting RDF as the core data structure and 'single-source-of-truth', making all inter-division software interoperable with the same (self-consistent) model.
- Even the mere conceptualization of geographic 'entites' has far-reaching implications, for example, geometry boundary objects can tend to over-inform policy and ecological analysis.
- Perspective-oriented characteristics > Topology > Geometry.
- Existing data models for spatial objects does not equate to the core concepts of spatial *information*.
-

Content negotiation and in-band and out of band links

Use blank nodes with link properties for rdf elements that are URI for out of band content. Link properties need to include identifier for profile

Identify in band links with sdo @id, out of band links with sdo:URL

W3C content negotiation by profile draft recommendation can provide some ideas for describing/profiles

In-band URI-- known to return rdf compatible with containing document vocabulary

Out-of-band-- URI dereferences to non-rdf representation, or can use content negotiation to get various representation profiles

connegbyP next steps (E2SIP activity)

-implementations for profiles in SELFIE

-GA pyLDAPAPI implementations

-practical profile definitions: define profiles with SHACL shapes <https://www.w3.org/TR/shacl/>