Interoperability, Linked Data and Reasoning - Lecture 3 of the GIServices module

Anthony Beck

Abstract

Figure 1:

This document has been written in CommonMark: an unambiguous implementation of Markdown for scholarly writing.

Contents

1 Preamble **7**

1.1 Licence . . . 7
1.2 Using Ipython for presentations . . . 8
1.3 The environment . . . 9
1.3.1 Set up watermark . . . 9
1.4 Running dynamic presentations . . . 16
1.5 Pre load some useful libraries . . . 17
1.6 About me . . . 18
1.7 About this presentation . . . 18

2 Contribution to GIScience learning outcomes **19**

3 A potted history of mapping **19**

3.1 In the beginning was the geoword . . . 19

4 And then there was data ……… **22**

5 Why is this an issue? **26**

6 INSPIRE **27**

7 INSPIRE principles **28**

8 Interoperability **30**

9 Technical interoperability - levelling the field **31**

10 Syntactic Heterogeneity **32**

11 Semantic Heterogeneity **33**

12 Schematic Heterogeneity **34**

13 The role of the OGC (a geospatial standards body) **35**

14 The role of the OGC (a geospatial standards body) **36**

15 OGC Technologies **37**

16 The main OGC standards **38**

17 Other OGC standards **39**

18 Interoperability in action 40

19 What did technical interoperability facilitate 41

20 The world was a happy place....... 43

21 Then .... along came open data 44

22 The Open landscape integrates formal and informal data 45

23 Background - originally a grass roots (community) movement.. 47

24 Open in government 48

24.1 Valuing Open Data 49

25 Open in academia 50

25.1 The next generation open data in academia 51

26 The underlying rationale of Open Data is: 52

27 Free and Open Source Software in in Geo 53

28 So...... we have access to lots of data and software 54

29 Where are these new data products? 54

30 Non-technical interoperability issues? 55

31 Non-technical interoperability 56

31.1 Policy Interoperability 58

31.2 Social (or human) Interoperability 59

31.3 Legal Interoperability 60

31.3.1 Licence Interoperability 61

32 Example of applying the semantic web to licence interoperability 62

32.1 What is a licence? 63

32.2 Creative Commons Zero 67

32.2.1 Constraints on CC0 68

32.2.2 CC license combinations 68

33 Why are licenses important? 69

33.1 Which is important when we mash up data 70

34 A rudimentry logic example **73**
34.1 Here's something I prepared earlier 74

35 A more robust logic **75**

36 OGC and Licence interoperability **76**

37 Linked data and the Semantic Web **77**
37.1 The web of Documents 77
37.2 The web of Linked Data 77

38 Linked Data **78**
38.1 Why publish Linked Data 78

39 Linked Data Basics **79**
39.1 Four rules for Linked Data from Tim Berners Lee 79

40 The Resource Description Framework (RDF) data model **80**
40.1 Data expressed as RDF 81
40.2 Data expressed as RDF Linked Data 82
40.3 RDF notation 83
40.4 One step beyond.... Linked Open Data 84
40.4.1 Is your Linked Open Data 5 star 84

41 The Supporting Semantic Web Stack **85**
41.1 It's about re-use 86
41.1.1 Vocabularies 86
41.2 It's about re-use 87
41.2.1 Ontology 87
41.2.2 Ontology 88
41.2.3 Ontology 88
41.2.4 Ontology Example 89
41.3 SPARQL the SQL of the semantic web 90
41.4 GeoSPARQL the SQL of the spatial semantic web 91
41.5 GeoSPARQL employs spatial calculus 92

42 Querying Linked Data in the wild **93**
42.1 The Ordnance Survey 93
42.2 Geonames 100

43 Geo Vocabularies **103**

44 Conclusions **104**

45 Geo reasoning example (if time) **105**

46 Questions **107**

References **109**

1.6 About me

Figure 4: It's all about me - details about Anthony Beck

- Honorary Research Fellow, University of Nottingham: orcid
- Director, Geolytics Limited - A spatial data analytics consultancy

1.7 About this presentation

- Available on GitHub - https://github.com/AntArch/Presentations_Github/
- Fully referenced PDF

2 Contribution to GIScience learning outcomes

This presentation contributes to the following learning outcomes for this course.

1. Knowledge and Understanding:
   - Appreciate the importance of standards for Geographic Information and the role of the Open Geospatial Consortium.
   - Understand the term ‘interoperability’.
   - Appreciate the different models for database design.
   - Understand the basis of Linked Data.
   - Find UK government open data and understand some of the complexities in the use of this data.
   - Appreciate the data issues involved in managing large distributed databases, Location-Based Services and the emergence of real-time data gathering through the ‘Sensor-Web’.
   - Understand the different models for creating international Spatial Data Infrastructures.
2. Intellectual Skills:
   - Evaluate the role of standards and professional bodies in GIS.
   - Articulate the meaning and importance of interoperability, semantics and ontologies.
   - Assess the technical and organisational issues which come into play when attempting to design large distributed geographic databases aimed at supporting ‘real-world’ problems.

3 A potted history of mapping

3.1 In the beginning was the geoword

and the word was ***cartography***

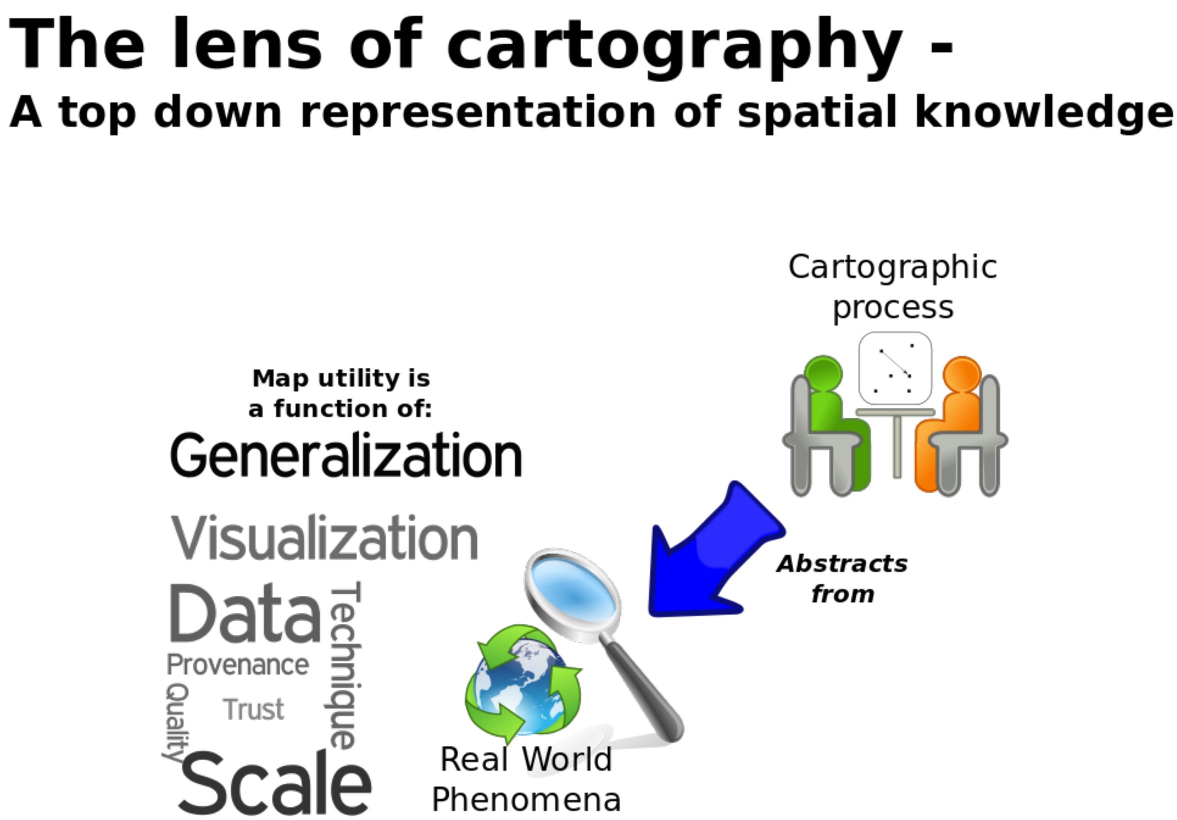

Figure 5: The lens of cartography A. Beck (2015h)

Figure 6: A static map (public domain) encapsulating spatial knowledge in a portable manner

- Cartography was king.
- Static representations of spatial knowledge with the cartographer deciding what to represent.
  - Hence, maps are domain specific knowledge repositories for spatial data

4 And then there was data ....

Figure 7: Data A. Beck (2015j)

Figure 8: But the data was siloed (restricted use)

Restrictive data

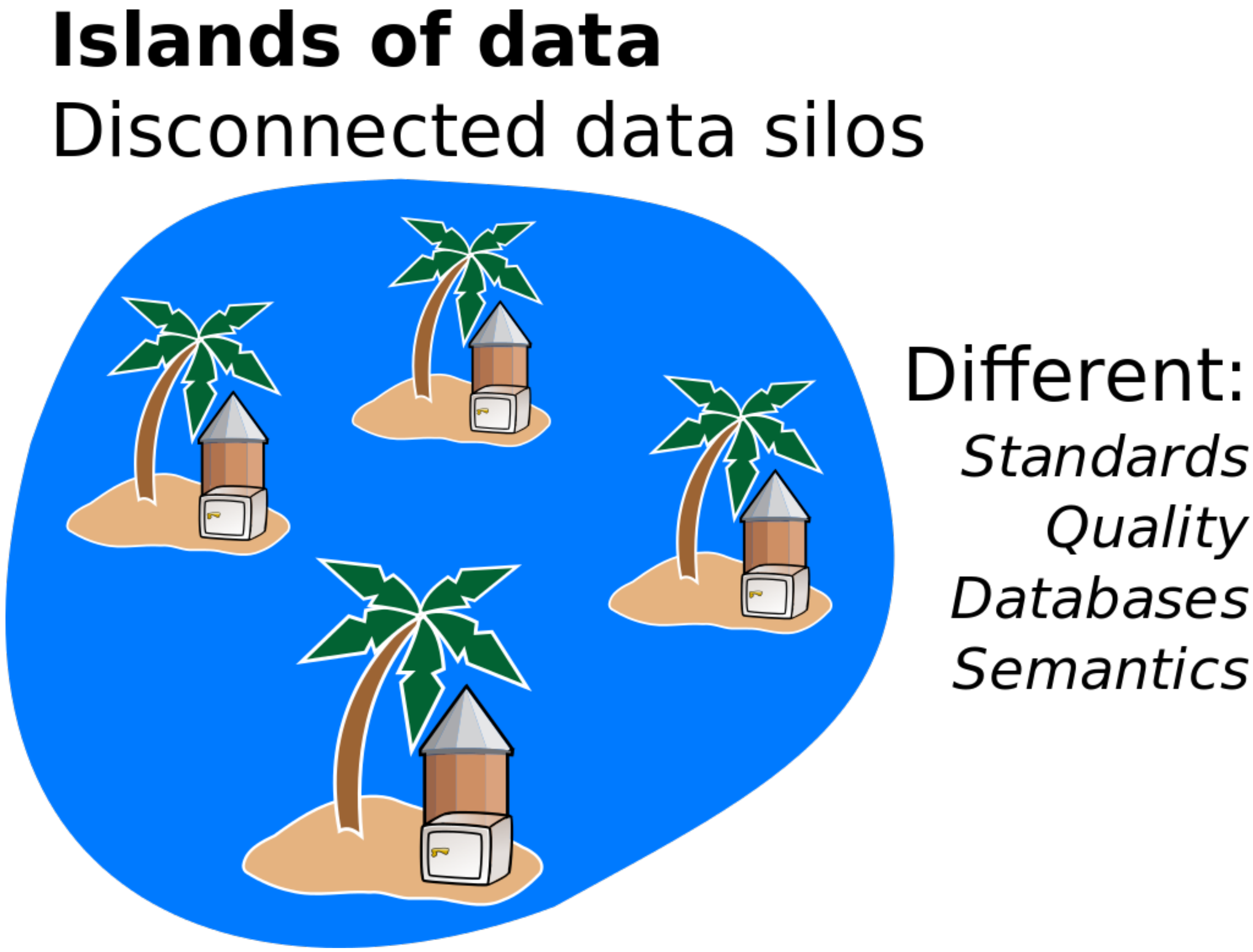

Figure 9: The implications of a landscape of silo-ed data A. Beck (2015e)

Disconnected data with different:

- Standards
- Quality
- Databases
- Semantics

5 Why is this an issue?

Over to you……

- Decision Making
  - certainty
  - uncertainty
- Co-ordination
- Policy formation
- Efficiencies
- Best Practice

6 INSPIRE

```
from IPython.display import YouTubeVideo
YouTubeVideo('xew6qI-6wNk')
```

```
    <iframe
        width="400"
        height="300"
        src="https://www.youtube.com/embed/xew6qI-6wNk"
        frameborder="0"
        allowfullscreen
    ></iframe>
```

7 INSPIRE principles

- Data should be collected only once and kept where it can be maintained most effectively
- It should be possible to combine seamless spatial information from different sources across Europe and share it with many users and applications
- It should be possible for information collected at one level/scale to be shared with all levels/scales; detailed for thorough investigations, general for strategic purposes
- Geoinformation needed for good governance at all levels should be readily and transparently available
- Easy to find what geoinformation is available, how it can be used to meet a particular need, and under which conditions it can be acquired and used

Figure 10: Concerted efforts to de-silo data and make data interoperable (restricted use)

Making data interoperable and open

8 Interoperability

> is a property of a product or system, whose interfaces are completely understood, to work with other products or systems, present or future, without any restricted access or implementation.

Wikipedia (2016)

The Defense domain are a bit more explicit……

> As defined by DoD policy, interoperability is the ability of systems, units, or forces to provide data, information, material, and services to, and accept the same from, other systems, units, or forces; and to use the data, information, material, and services so exchanged to enable them to operate effectively together. IT and NSS interoperability includes both the technical exchange of information and the end-to-end operational effectiveness of that exchanged information as required for mission accomplishment. Interoperability is more than just information exchange; it includes systems, processes, procedures, organizations, and missions over the life cycle and must be balanced with information assurance.

Watson (2010)

9 Technical interoperability - levelling the field

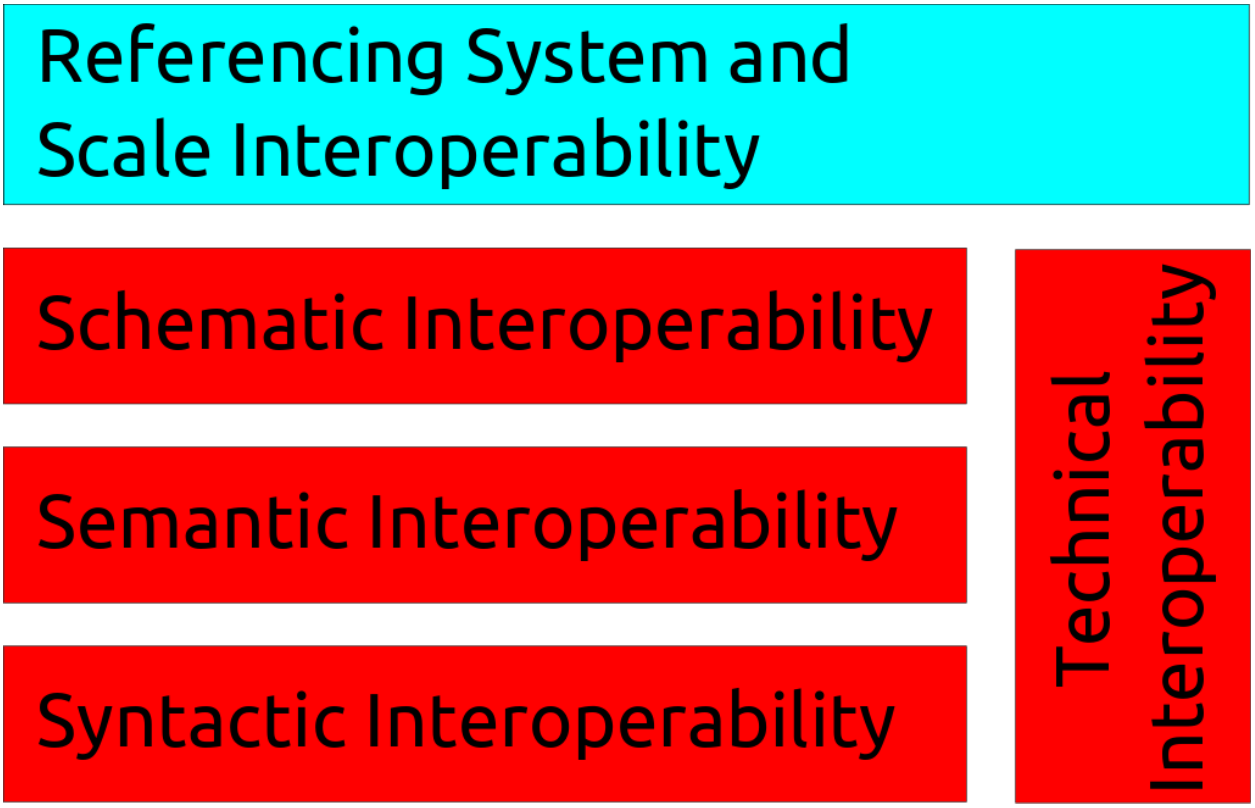

Figure 11: Interoperable integration of spatial data - the technological issues A. Beck (2015d)

10 Syntactic Heterogeneity

> the difference in data format. The same **logical model** can be represented in a range of different **physical models** (for example ESRI shape file or Geography Mark-up Language (GML)).

This mismatch between underlying data models implies that the same information could be represented differently in different organisations.

The most profound difference is in the storage paradigm:

- relational,
- object orientated or
- hybrids.

Beck et al. (2008), Bishr (1998)

11 Semantic Heterogeneity

> Semantic heterogeneity refers to differences in naming conventions and conceptual groupings.

This can be subdivided into **naming** and **cognitive** heterogeneities.

- Naming (synonym) mismatch arises when semantically identical data items are named differently.
- Cognitive (homonym) mismatch arises when semantically different data items are named identically.
  - Cognitive semantics can be subtle, reflecting the domain of discourse.

Beck et al. (2008), Bishr (1998)

12 Schematic Heterogeneity

> refers to the differences in data model between organisations modelling the same concepts.

This reflects each organisation's abstracted view of their business and physical assets. Hence, different hierarchical and classification concepts are adopted by each organisation to refer to identical or similar real world objects.

Beck et al. (2008), Bishr (1998)

13 The role of the OGC (a geospatial standards body)

- To serve as a global forum for the development, promotion and harmonization of *open and freely available* **geospatial standards**
- To achieve the full societal, economic and scientific benefits of integrating electronic location resources into commercial and institutional processes worldwide.

Figure 12: The OGC Logo

14 The role of the OGC (a geospatial standards body)

OGC's Open Standards are:

- Freely and publicly available
- Non discriminatory
- No license fees
- Vendor neutral
- Data neutral
- Adopted in a formal, member based consensus process

OGC's Open Standards are submitted to other industry and National Standards Development Organisations in the vertical area and to global organisations like ISO for standard branding.

15 OGC Technologies

- The OGC publish standards that have been agreed by OGC members
- Current standards can be found at: http://www.opengeospatial.org/standards
- These are implementation standards
  - written for a more technical audience and detail the interface structure between software components
- Predicated on abstract specifications
  - the conceptual foundation for most OGC standards development activities
  - http://www.opengeospatial.org/specs/?page=abstract

16 The main OGC standards

- WMS – Web Map Service
  - Provides rendered images of maps
  - Current version: 1.3
- WFS – Web Feature Service
  - Provides vector data on demand
  - Current version: 2.0
- WCS – Web Coverage Service
  - Provides raster data (e.g. satellite data) on demand
  - Current version: 2.0
- GML – The Geography Markup Language
  - Used as an interoperable standard for transmitting geographic data (2D, 3D, topology, etc.)
  - Versions 2.1.x and 3.2.1 are most relevant

17 Other OGC standards

```python
from IPython.display import IFrame
IFrame('http://www.opengeospatial.org/standards', width=1000, height=700)
```

```
    <iframe
        width="1000"
        height="700"
        src="http://www.opengeospatial.org/standards"
        frameborder="0"
        allowfullscreen
    ></iframe>
```

18 Interoperability in action

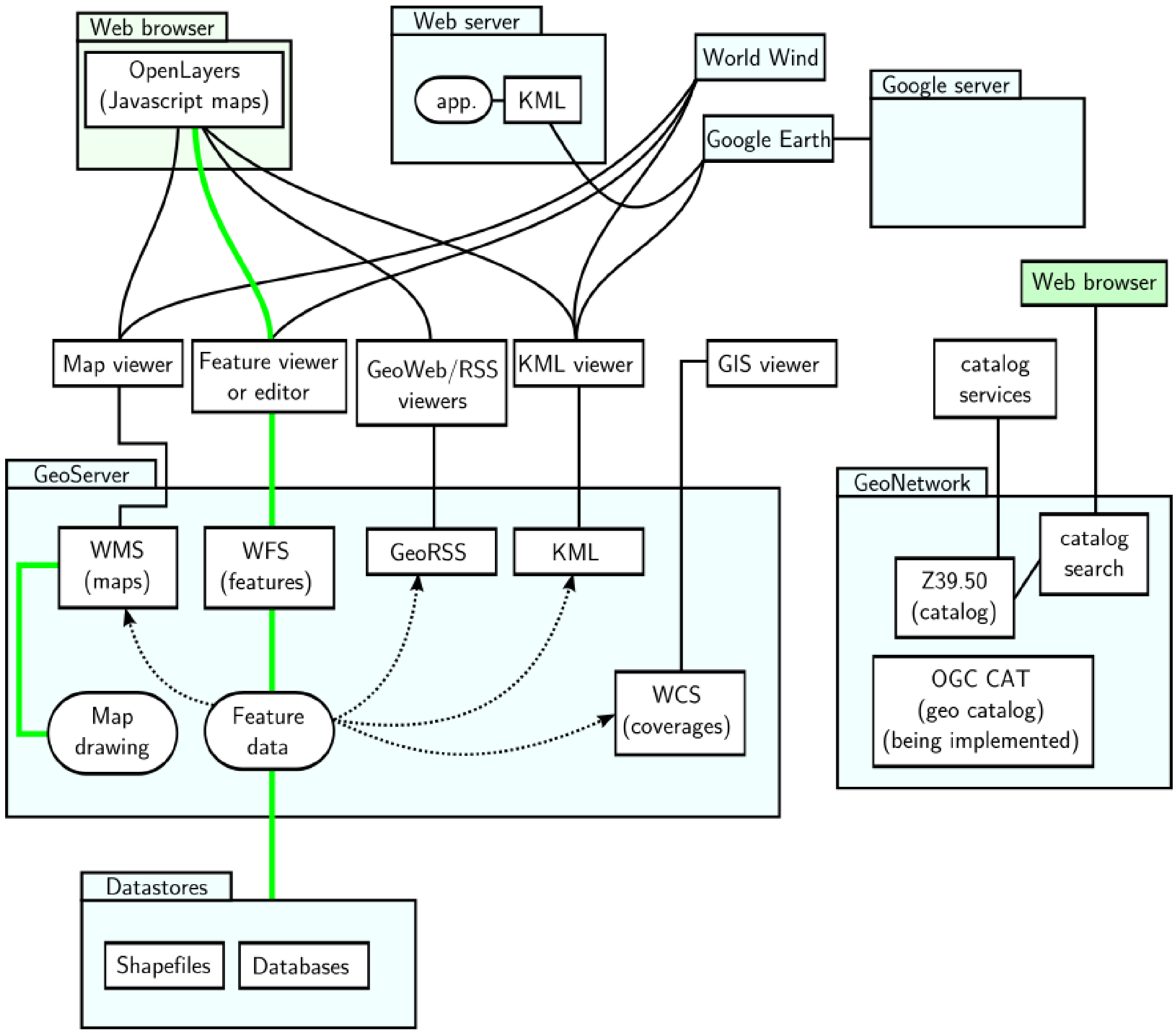

Figure 13: The use of OGC standards in an interoperable multi-data delivery environment SEWilco (2013)

19 What did technical interoperability facilitate

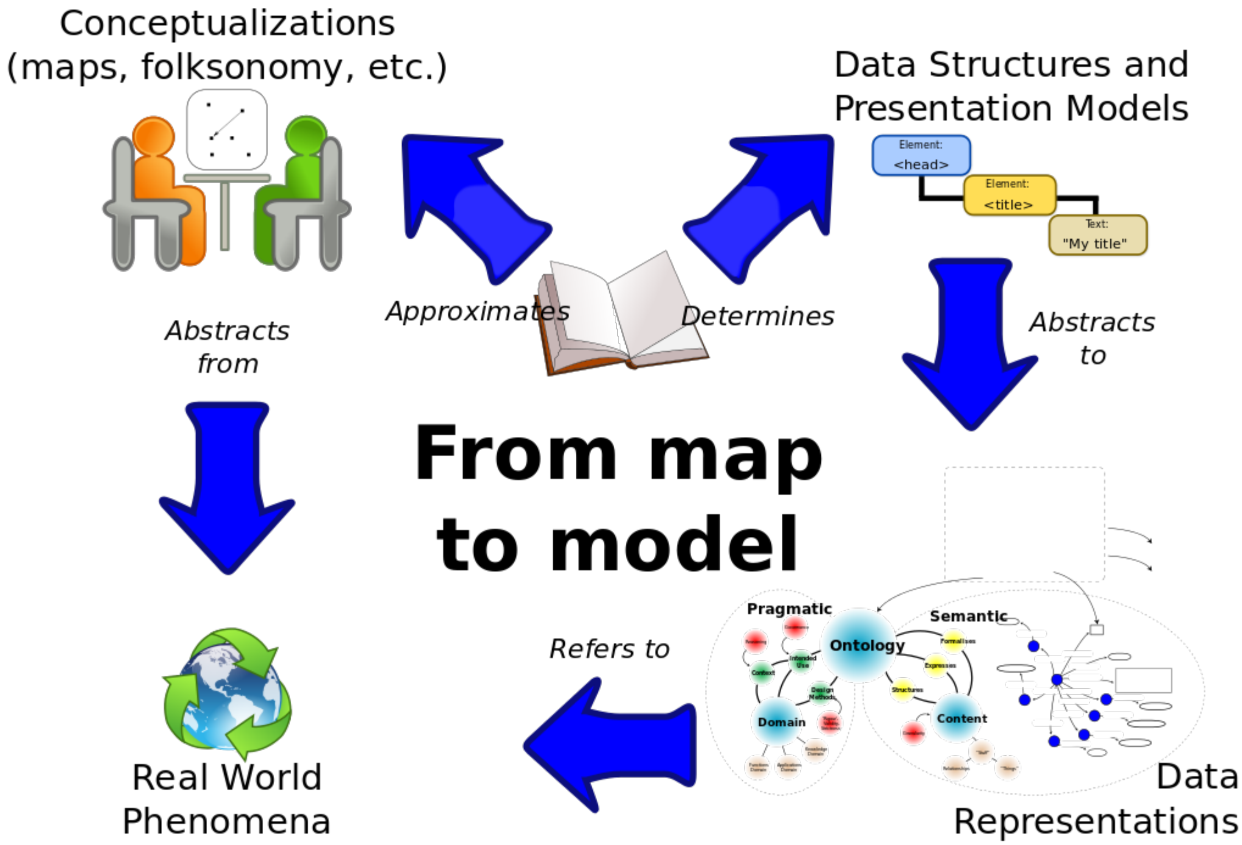

Figure 14: From Map to Model The changing paradigm of map creation from cartography to data driven visualization A. Beck (2015c)

From Map to Model The changing paradigm of map creation from cartography to data driven visualization

Figure 15: A new working paradigm (public domain)

20 The world was a happy place.......

Our data was interoperable!

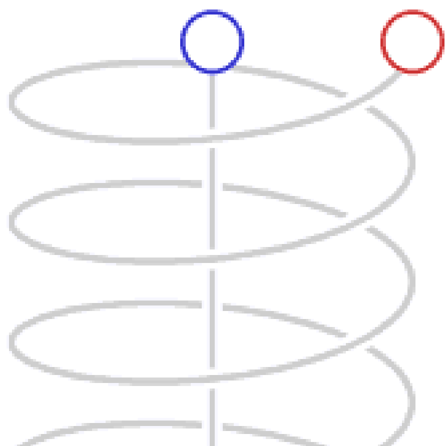

Figure 16: But time moved on cleonis (2006)

21 Then .... along came open data

Figure 17: open data word cloud of Anthony Beck

22 The Open landscape integrates formal and informal data

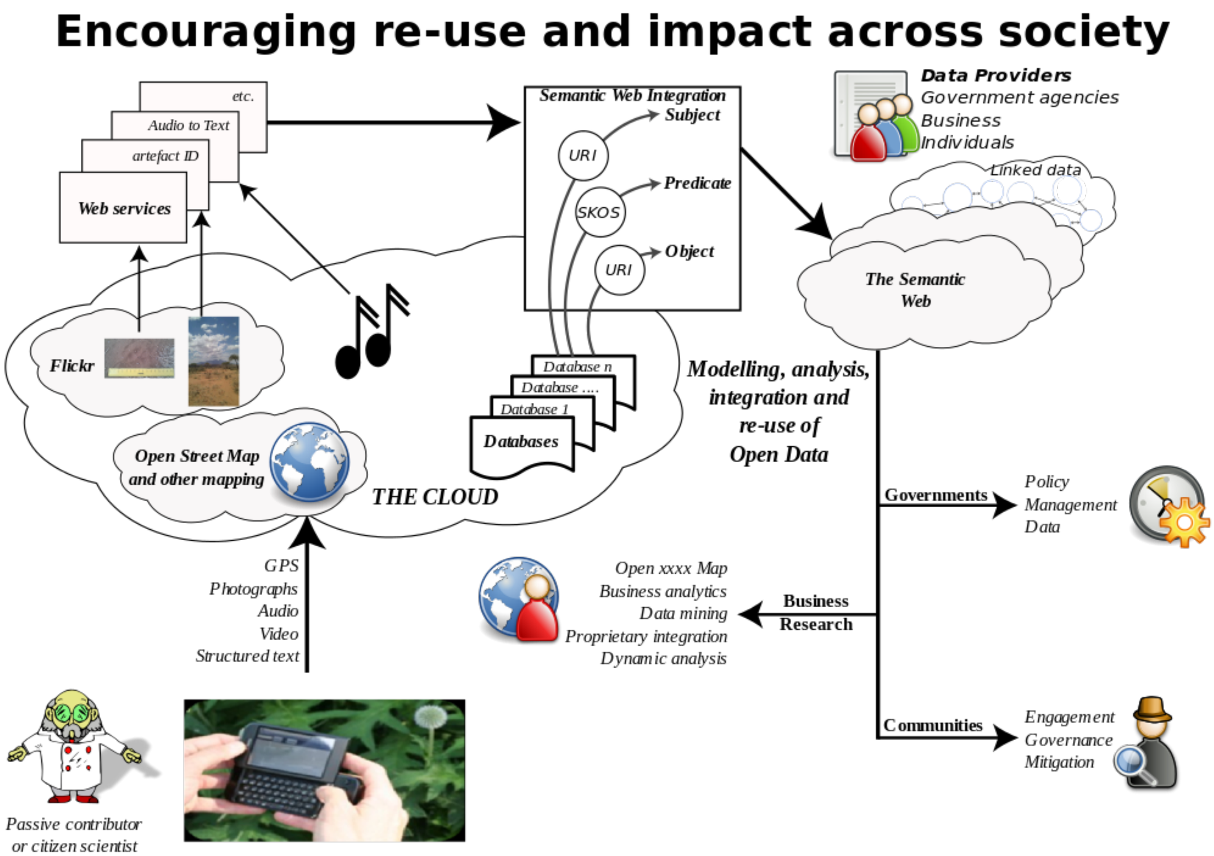

Figure 18: Local To Global integration of data to create multiple generic products A. Beck (2015f)

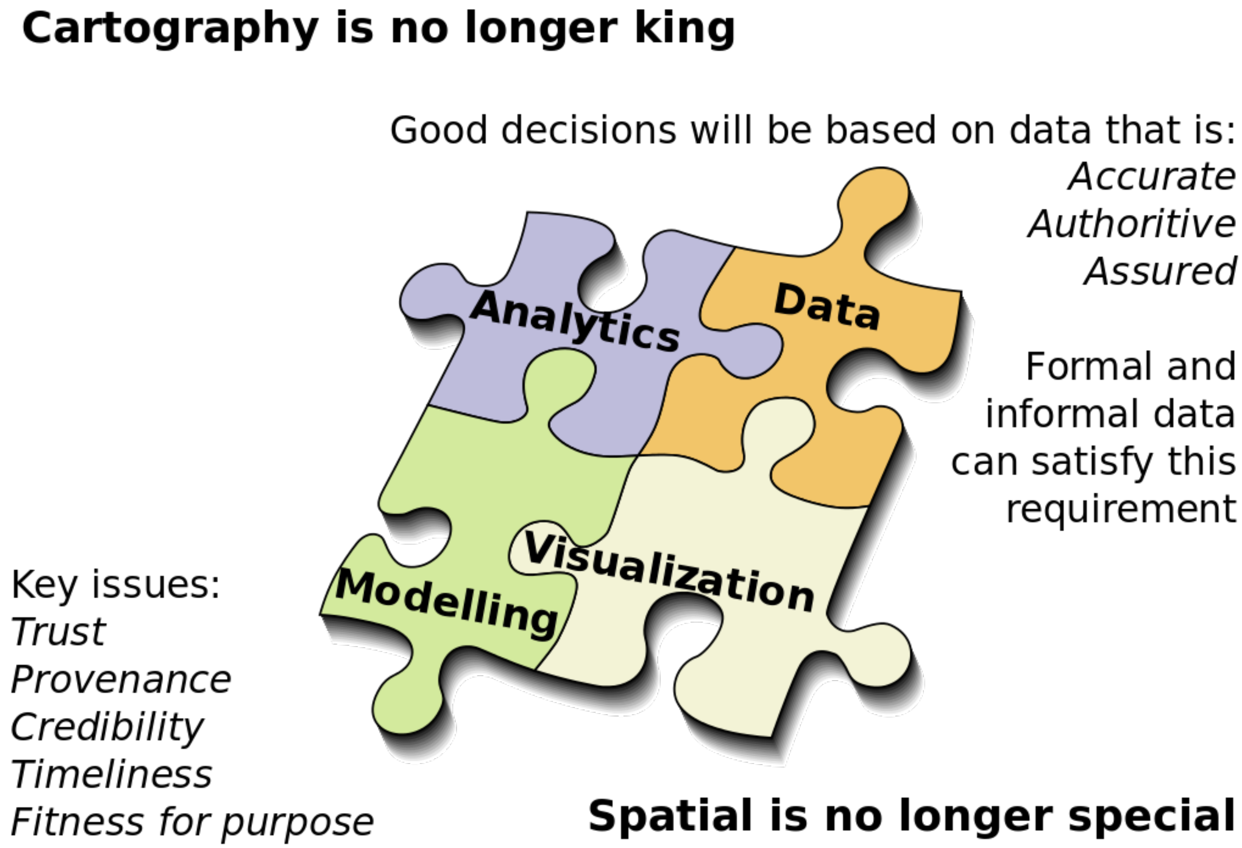

Figure 19: Cartography is no longer key. Spatial mapping is now about the the formal and informal data stack. Elements such as provenance, credibility are much more important for use and re-use of this data. A. Beck (2015b)

23 Background - originally a grass roots (community) movement..

Figure 20: The Open Knowledge Foundation (AlisonW (2014)), Open Street Map (Rillke (2014)), Wikipedia (Foundation (2010)) and OSGeo

Open access to knowledge gained significant momementum with the increased uptake of the World Wide Web. This is particularly seen in initiatives like Wikipedia (established in 2001) and Open Knowledge (was the Open Knowledge Foundation: established in 2004). Within the Geo community Open Street Map (also established in 2004) and the Open Source Geospatial Foundation (OSGeo - established in 2006) are key initiatives that promote accessible data and software resources respectively.

Critical to this is that these were **grass roots** (community) movements that have proven to be highly disruptive to incumbent data providers, practices and policies.

24 Open in government

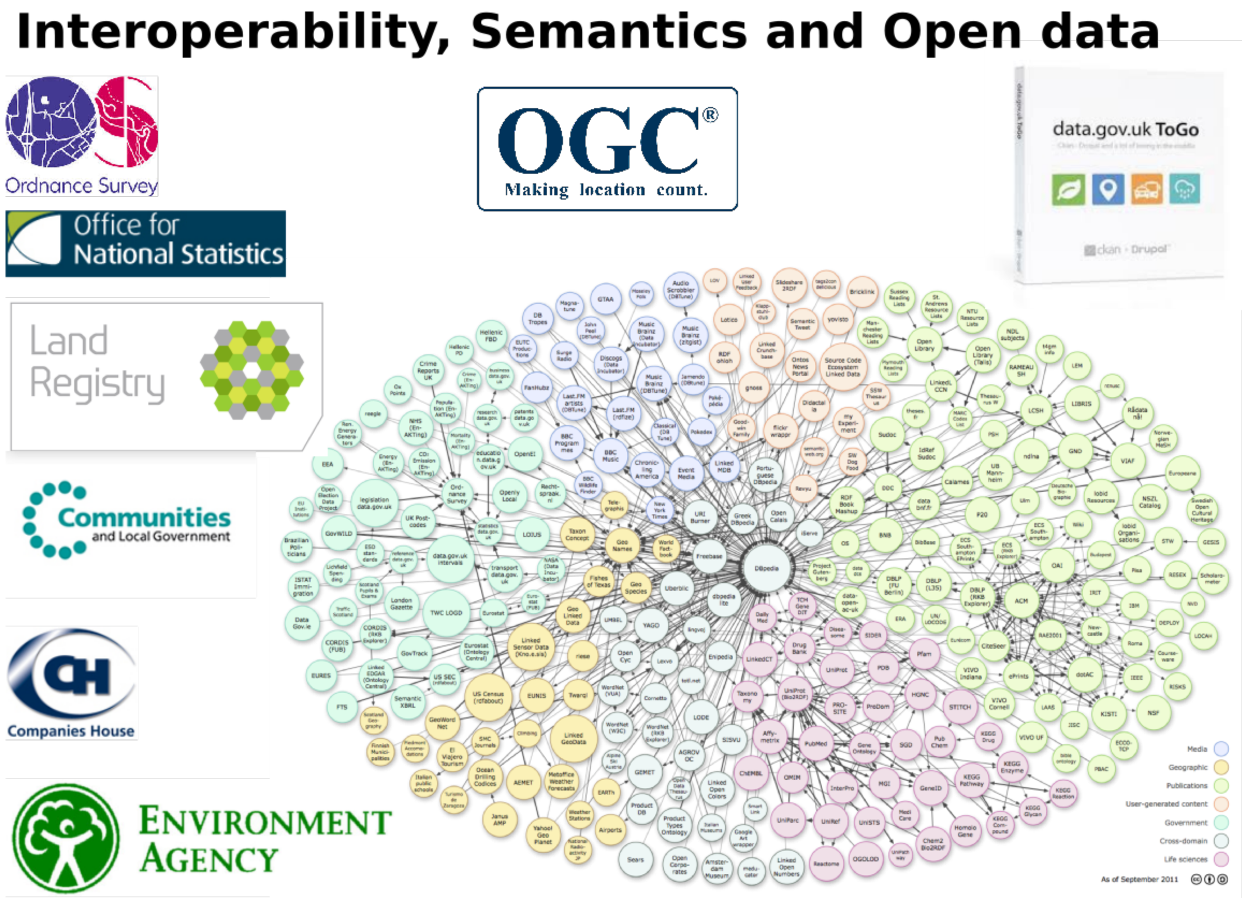

Figure 21:

The impact of these grass roots movements is seen in Open Data (dot) gov. Pioneered by leaders such as Tim Berners Lee and Nigel Shadbolt

The Shakespeare review (2013) indicate that the amount of government Open Data, at least in the UK, is only going to grow. Open data has the potential to trigger a revolution in how governments think about providing services to citizens and how they measure their success: this produces societal impact. This will require an understanding of citizen needs, behaviours, and mental models, and how to use data to improve services.

24.1 Valuing Open Data

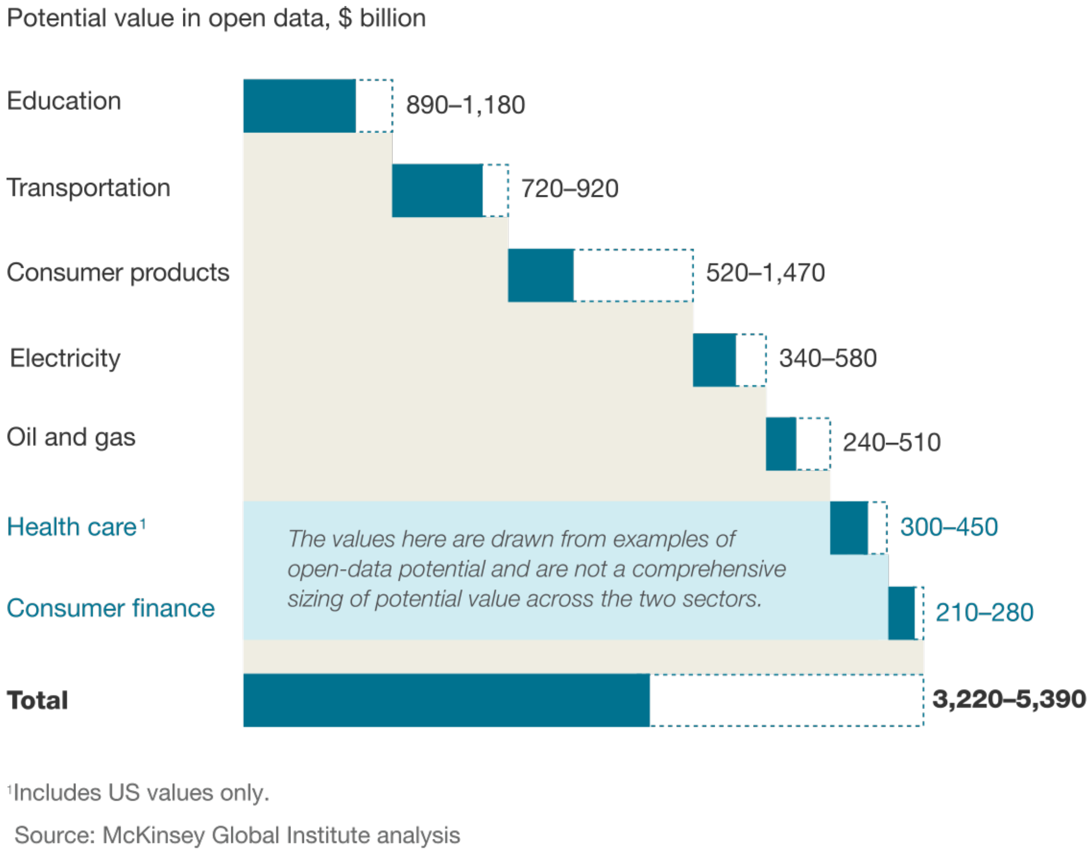

Figure 22: McKinsey report valuing *open data* (McKinsey 2013)

A McKinsey Global Institute report examines the economic impact of Open Data (McKinsey 2013) and estimates that globally open data could be worth a minimum of $3 trillion annually.

25 Open in academia

> Open inquiry is at the heart of the scientific enterprise..... Science's powerful capacity for self-correction comes from this openness to scrutiny and challenge.

Science as an open enterprise (The Royal Society 2012, p.7).

> Science is based on building on, reusing and openly criticising the published body of scientific knowledge.
>
> For science to effectively function, and for society to reap the full benefits from scientific endeavours, it is crucial that science data be made open.

The Panton Principles (Murray-Rust et al. (2010)) which underpin **Open Science**.

The Royal Society's report Science as an open enterprise (2012) identifies how 21st century communication technologies are changing the ways in which scientists conduct, and society engages with, science. The report recognises that 'open' enquiry is pivotal for the success of science, both in research and in society.

The Panton Principles pre-cursed this call with a clarion call to the academic community to open their data and start to conduct **open science**.

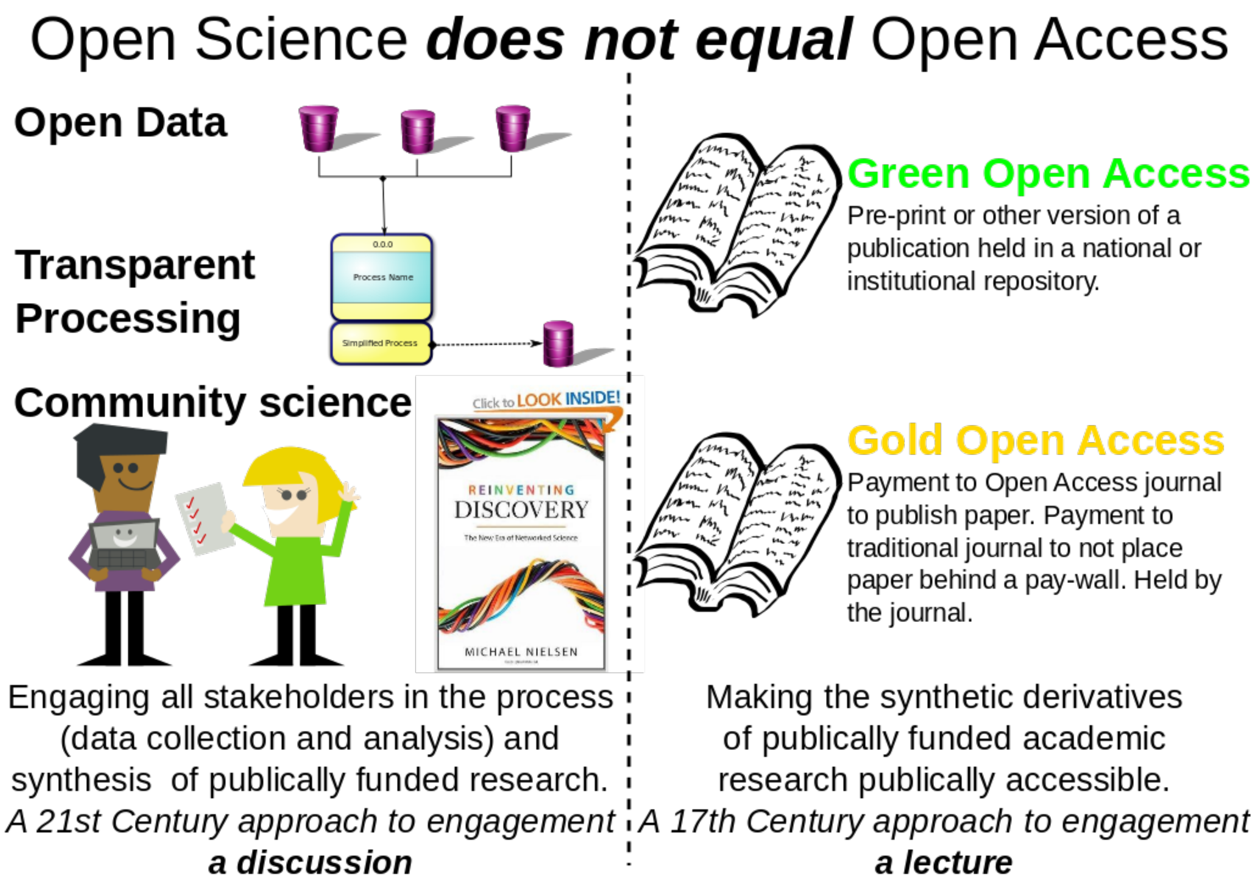

Figure 23: Beck (2013)

This goes beyond open access to publications (Open Access), to include access to data and other research outputs (Open Data), and the process by which data is turned into knowledge (Open Science).

25.1 The next generation open data in academia

Figure 24:

Zenodo is a **DATA REPOSITORY** which offers:

- accreditation
- different licences
- different exposure (private (closed), public (open) and embargoed (timestamped))
- DOIs
- is free at the point of use

- is likely to be around for a long time
  - supported by Horizon 2020 and delivered by CERN

26 The underlying rationale of Open Data is:

- unfettered access to large amounts of 'raw' data
  - enables patterns of re-use and knowledge creation that were previously impossible.
  - improves transparency and efficiency
  - encourages innovative service delivery
- introduces a range of data-mining and visualisation challenges,
  - which require multi-disciplinary collaboration across domains
  - catalyst to research and industry
- supports the generation of new products, services and markets
- the prize for succeeding is improved knowledge-led policy and practice that transforms
  - communities,
  - practitioners,
  - science and
  - society

27 Free and Open Source Software in in Geo

```python
from IPython.display import IFrame
IFrame('http://www.osgeo.org/', width=1000, height=700)
```

```
    <iframe
        width="1000"
        height="700"
        src="http://www.osgeo.org/"
        frameborder="0"
        allowfullscreen
    ></iframe>
```

28 So…… we have access to lots of data and software

- Formal and Informal
- Open and Proprietary

29 Where are these new data products?

Data, data everywhere - but where are the new derivatives and services?

30 Non-technical interoperability issues?

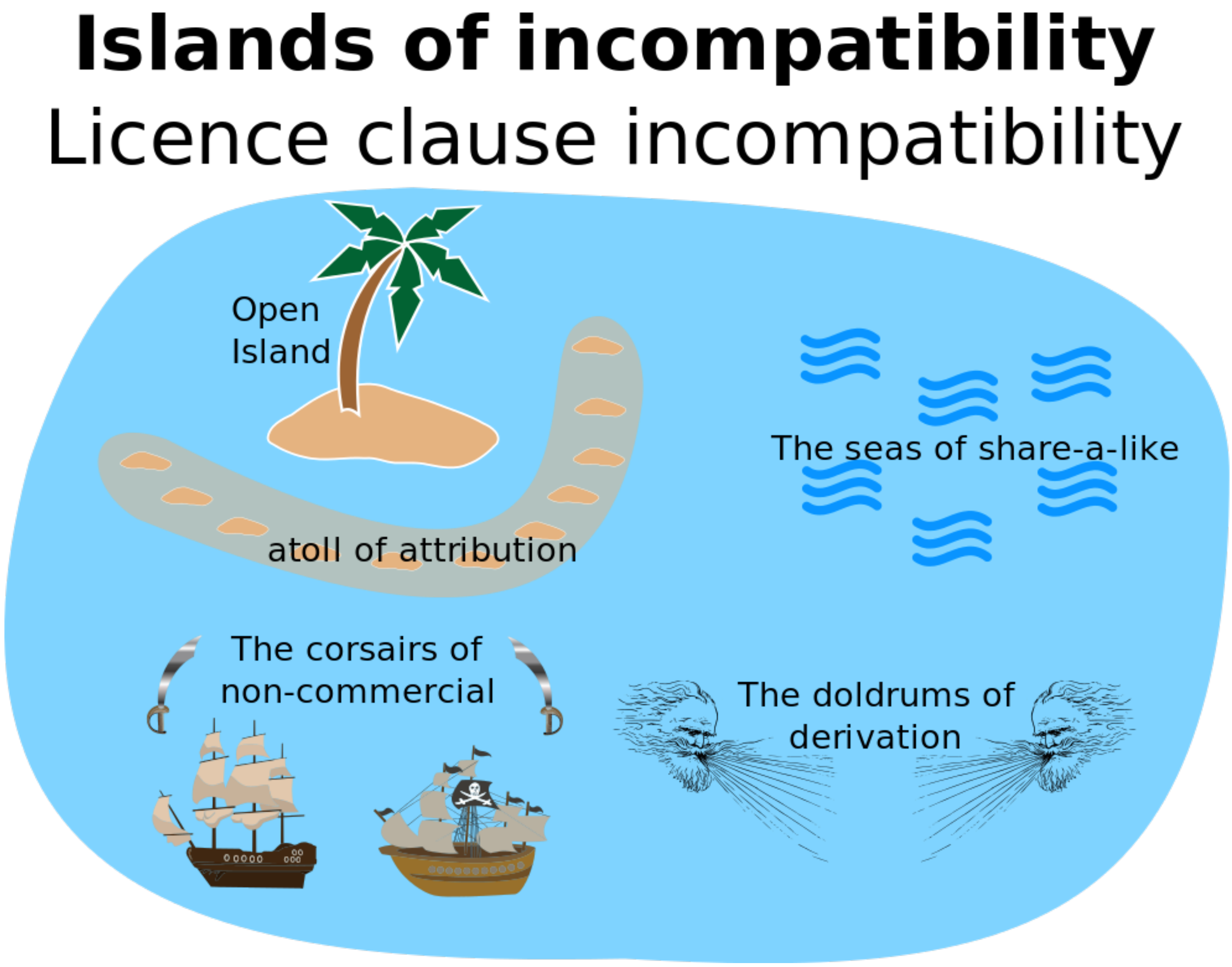

Figure 25: Islands of incompatibility (Beck 2016)

31 Non-technical interoperability

Issues surrounding non-technical interoperability include:

- Policy interoperabilty
- Licence interoperability
- Legal interoperability
- Social interoperability

We will focus on licence interoperability

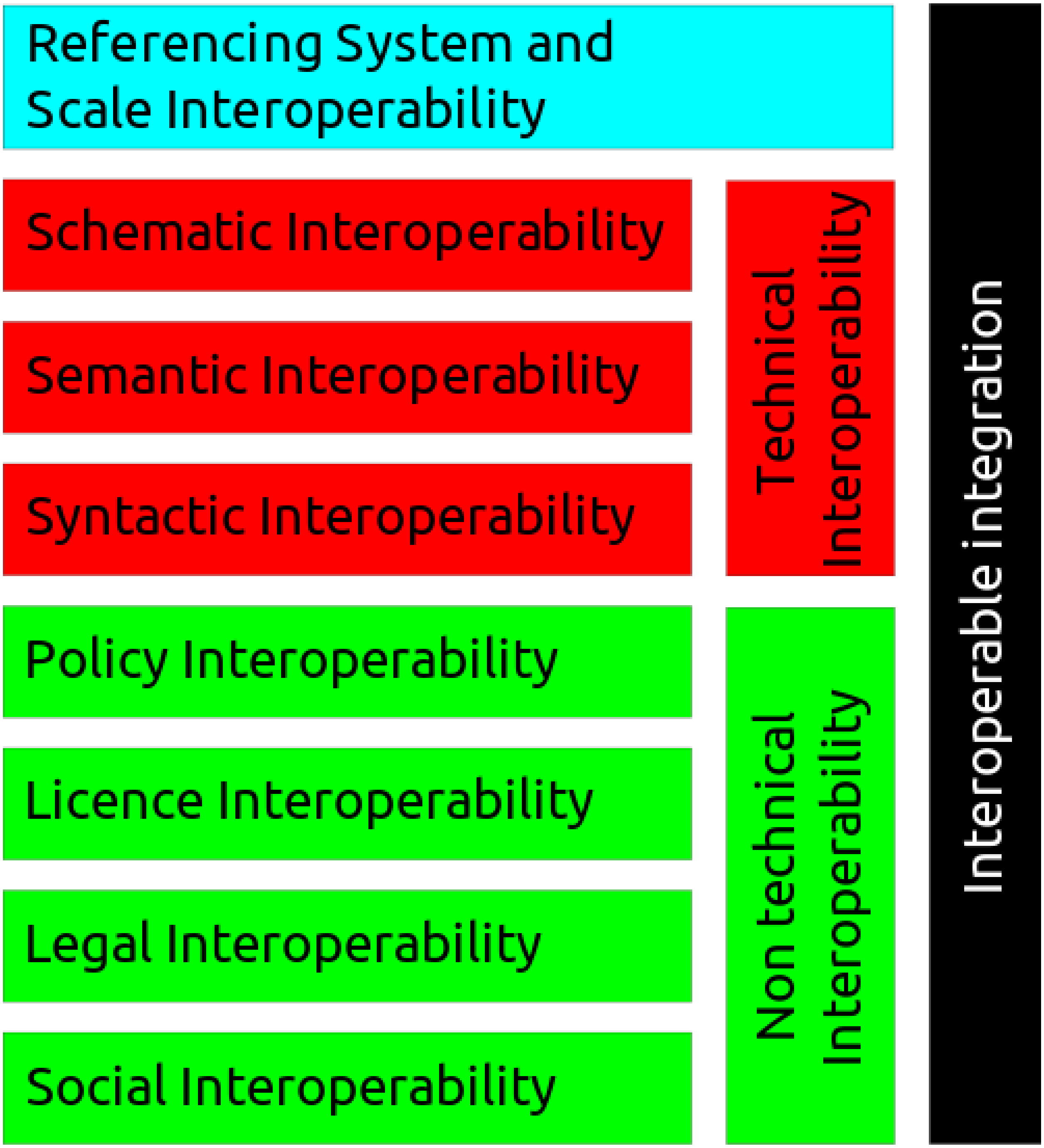

Figure 26: The full stack that enables interoperable integration of spatial data A. Beck (2015g)

31.1 Policy Interoperability

The relationship between:

- Individuals
- Organisations
- Countries

Policy determines what, who and how different content can be accessed.

In addition to other elements the policy statements determine:

- Authentication
- Authorization
- Audit

See Innocenti et al. (2011) for more details

31.2 Social (or human) Interoperability

Social interoperability is concerned about the environment and business and human processes.

- Tools are used by people
- The social dimension of operational use is underestimated (it's difficult)
- People form complex inclusive and exclusive networks
  - These operate at many scales

US Department of Defence researchers have advocated the development of Policy, Standards, and Operational Procedures for:

- forming human networks
- human to human communications
- organization to organization communications
- human system integration
- information sharing across disparate domains:
  - DoD-Coalition-Interagency-intercommunity

31.3 Legal Interoperability

> Legal interoperability addresses the process of making legal rules cooperate across jurisdictions, on different subsidiary levels within a single state or between two or more states.

(Weber (2014), p. 6)

The Research Data Alliance state that legal interoperability occurs among multiple datasets when:

- use conditions are clearly and readily determinable for each of the datasets,
- the legal use conditions imposed on each dataset allow creation and use of combined or derivative products, and
- users may legally access and use each dataset without seeking authorization from data rights holders on a case-by-case basis, assuming that the accumulated conditions of use for each and all of the datasets are met.

Legal interoperability also implies that the search for or tracking of licenses or other legal instruments and their compatibility with other legal conditions will occur in online environments.

31.3.1 Licence Interoperability

A specific form of legal interoperability

32 Example of applying the semantic web to licence interoperability

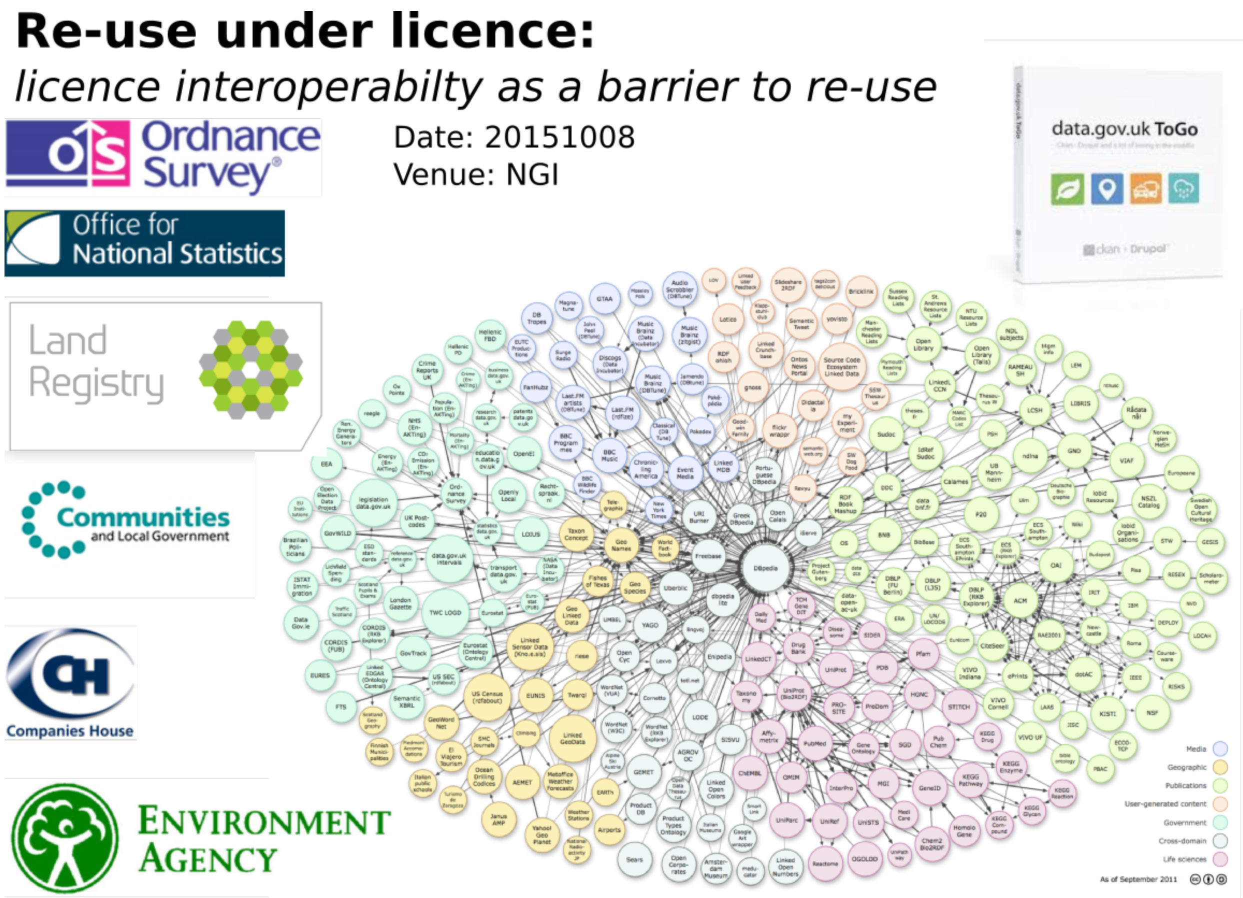

Figure 27: The modern data landscape (restricted)

There is a multitude of formal and informal data.

32.1 What is a licence?

Wikipedia state:

> A license may be granted by a party (“licensor”) to another party (“licensee”) as an element of an agreement between those parties.
>
> A shorthand definition of a license is “an authorization (by the licensor) to use the licensed material (by the licensee).”

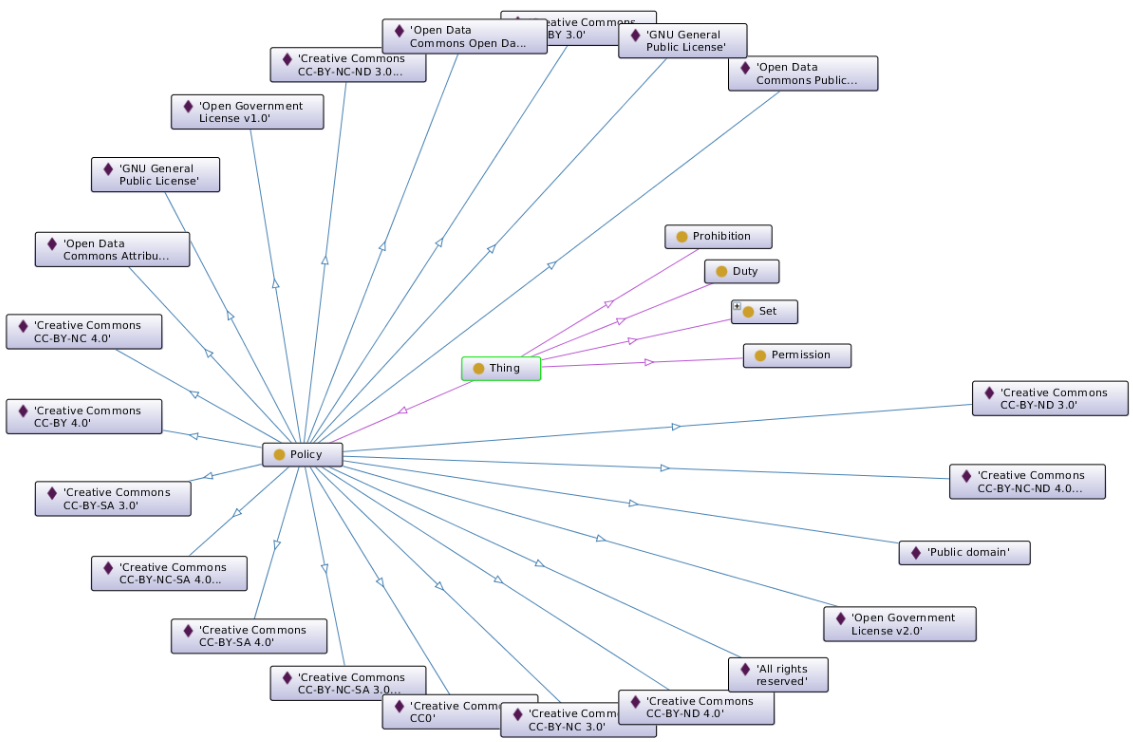

Figure 28: Some licences (Anon n.d.)

Each of these data objects can be licenced in a different way. This shows some of the licences described by the RDFLicence ontology

```python
### Export this notebook as markdown
commandLineSyntax = 'dot -Tpng FCA_ConceptAnalysis.dot > FCA_ConceptAnalysis.png'
commandLineSyntax = 'dot -Tsvg FCA_ConceptAnalysis.dot > FCA_ConceptAnalysis.svg'
print (commandLineSyntax)

os.system(commandLineSyntax)
```

Concepts (derived from Formal Concept Analysis) surrounding licences

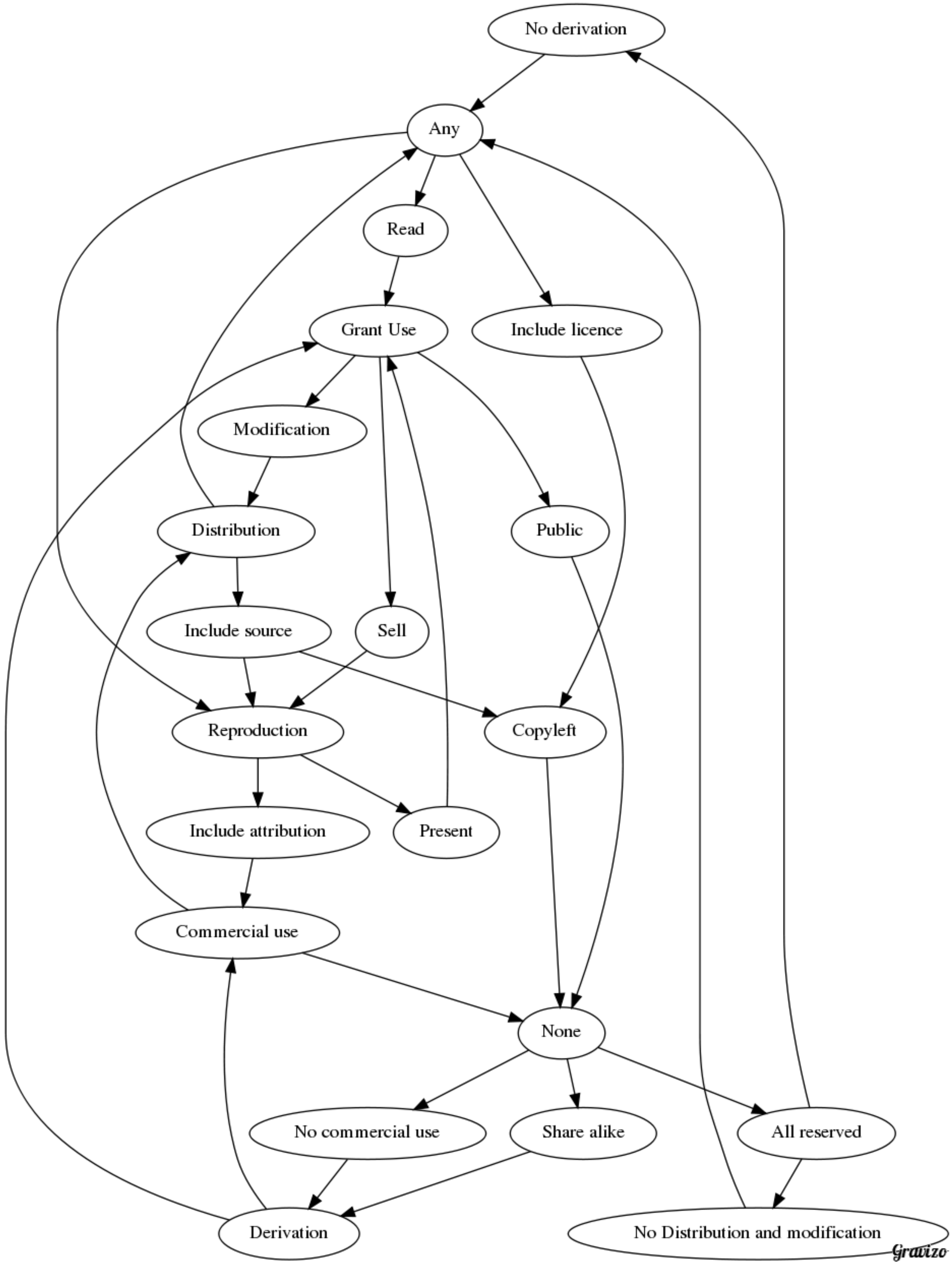

Figure 29: Concepts surrounding licences (derived from Formal Concept Analysis)

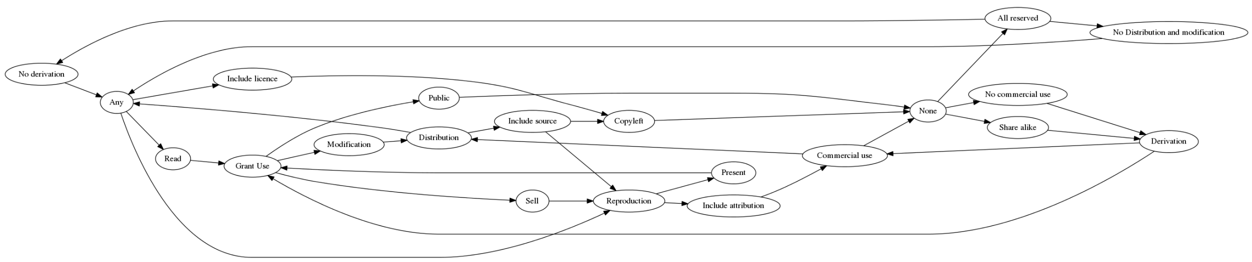

Figure 30: Concepts surrounding licences (derived from Formal Concept Analysis)

Two lead organisations have developed legal frameworks for content licensing:

- Creative Commons (CC) and
- Open Data Commons (ODC).

Until the release of CC version 4, published in November 2013, the CC licence did not cover data. Between them, CC and ODC licences can cover all forms of digital work.

- **There are many other licence types**
- Many are bespoke
  - Bespoke licences are difficult to manage
  - Many legacy datasets have bespoke licences

Figure 31: Creative Commons Gianni (2008)

I'll describe CC in more detail

32.2 Creative Commons Zero

Creative Commons Zero (CC0) is essentially public domain which allows:

- Reproduction
- Distribution
- Derivations

32.2.1 Constraints on CC0

The following clauses constrain CC0:

- Permissions
  - ND – No derivatives: the licensee can not derive new content from the resource.
- Requirements
  - BY – By attribution: the licensee must attribute the source.
  - SA – Share-alike: if the licensee adapts the resource, it must be released under the same licence.
- Prohibitions
  - NC – Non commercial: the licensee must not use the work commercially without prior approval.

32.2.2 CC license combinations

Table 1: Creative Commons license combinations

License	Reproduction	Distribution	Derivation	BY	SA	NC
CC0	X	X	X			
CC-BY-ND	X	X		X		
CC-BY-NC-ND	X	X		X		X
CC-BY	X	X	X	X		
CC-BY-SA	X	X	X	X	X	
CC-BY-NC	X	X	X	X		X
CC-BY-NC-SA	X	X	X	X	X	X

33 Why are licenses important?

- They tell you what you can and can't do with 'stuff'
- Very significant when multiple datasets are combined
  - It then becomes an issue of license compatibility

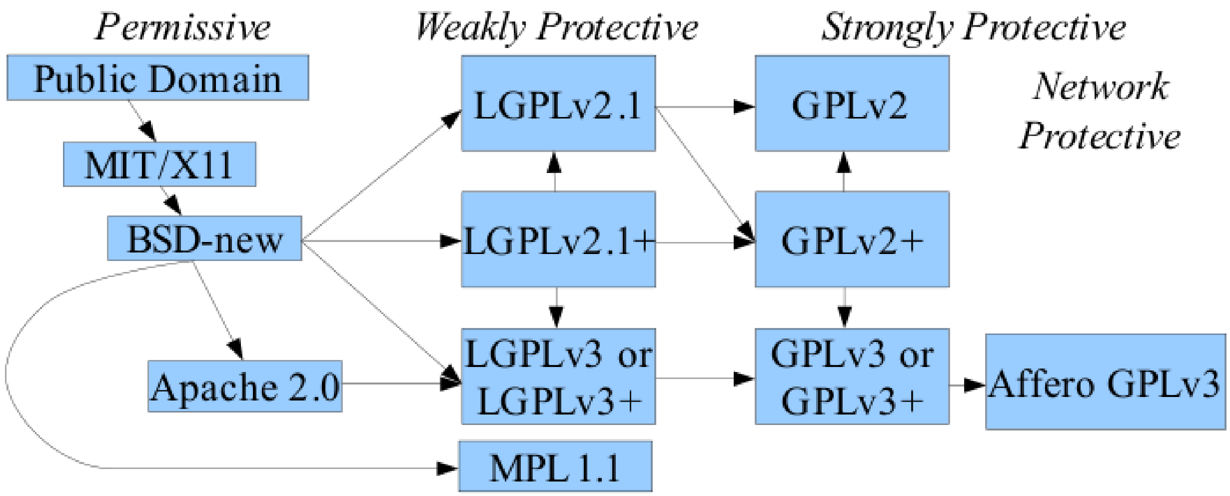

Figure 32: Compatibility of common open-source software licenses Wheeler (2007)

33.1 Which is important when we mash up data

Certain licences when combined:

- Are incompatible
  - Creating data islands
- Inhibit commercial exploitation (NC)
- Force the adoption of certain licences
  - If you want people to commercially exploit your stuff don't incorporate CC-BY-NC-SA data!
- Stops the derivation of *new works*

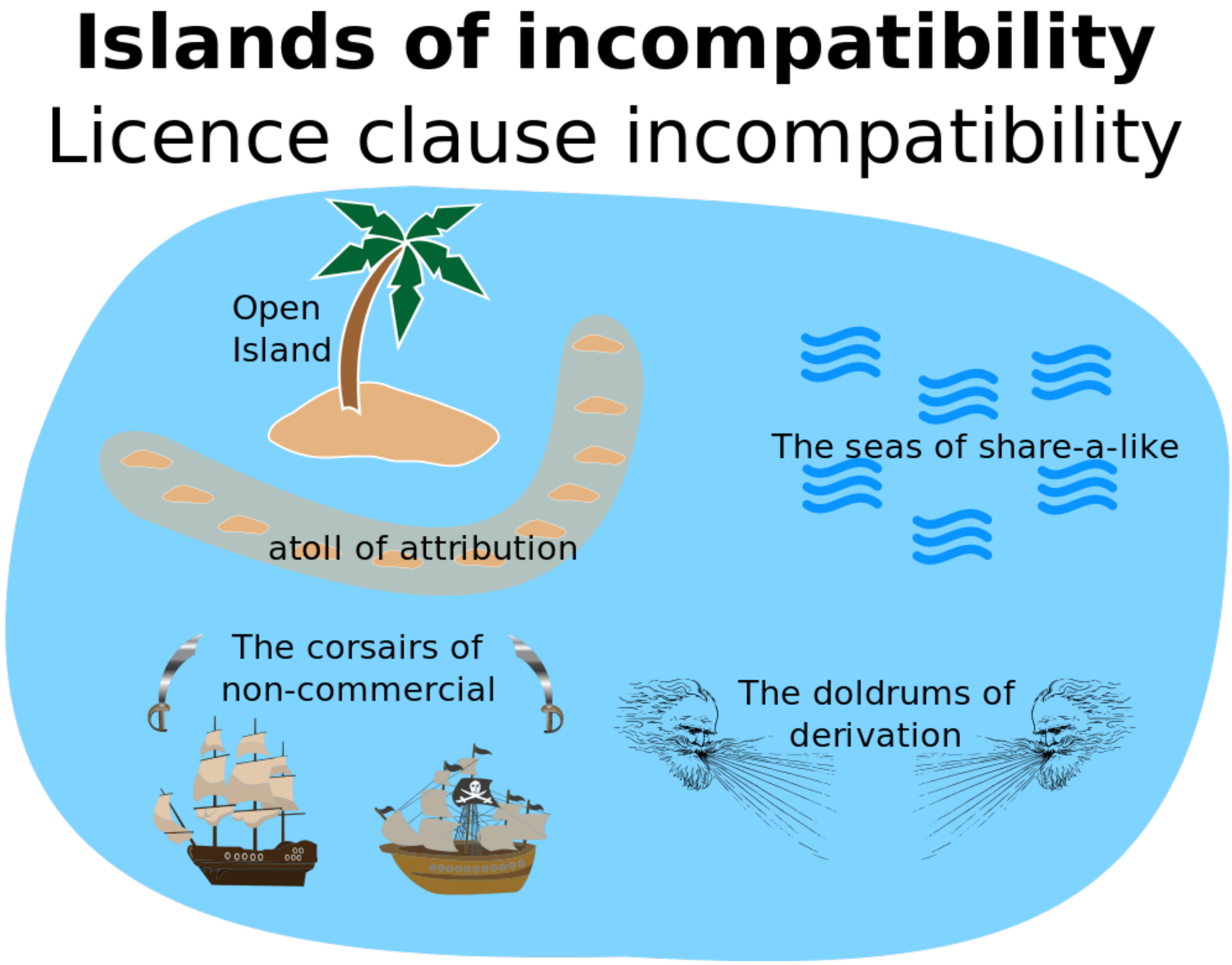

Figure 33: Islands of incompatibility (Beck 2016)

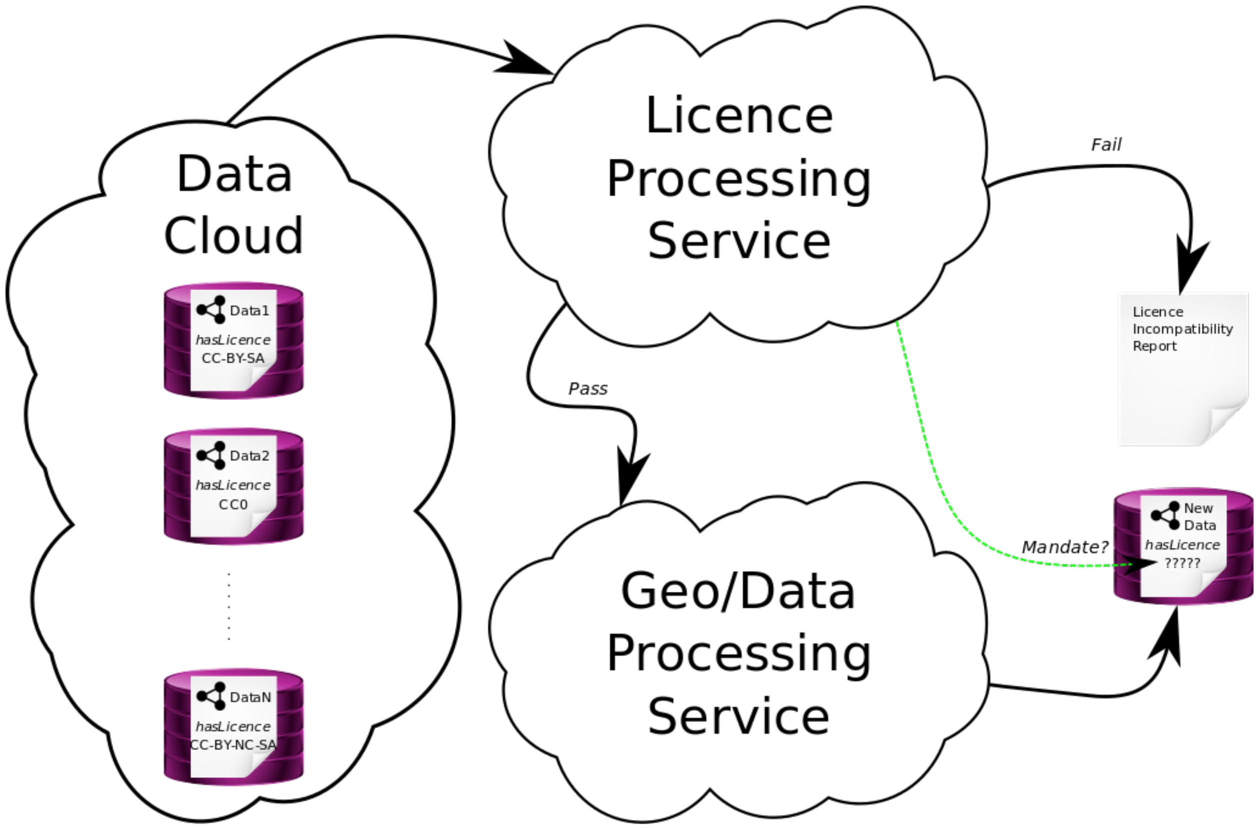

Figure 34: A conceptual licence processing workflow A. Beck (2015a)

A conceptual licence processing workflow. The licence processing service analyses the incoming licence metadata and determines if the data can be legally integrated and any resulting licence implications for the derived product.

34 A rudimentry logic example

```
Data1 hasDerivedContentIn NewThing.

Data1 hasLicence a cc-by-sa.

What hasLicence a cc-by-sa? #reason here

If X hasDerivedContentIn Y and hasLicence Z then Y hasLicence Z. #reason here

Data2 hasDerivedContentIn NewThing.

Data2 hasLicence a cc-by-nc-sa.

What hasLicence a cc-by-nc-sa? #reason here

Nothing hasLicence a cc-by-nc-sa and hasLicence a cc-by-sa. #reason here
```

And processing this within the Protege reasoning environment

```python
from IPython.display import YouTubeVideo
YouTubeVideo('jUzGF401vLc')
```

```
    <iframe
        width="400"
        height="300"
        src="https://www.youtube.com/embed/jUzGF401vLc"
        frameborder="0"
        allowfullscreen
    ></iframe>
```

34.1 Here's something I prepared earlier

A live presentation (for those who weren't at the event)…..

```python
from IPython.display import YouTubeVideo
YouTubeVideo('tkRB5Rp1_W4')
```

```
    <iframe
        width="400"
        height="300"
        src="https://www.youtube.com/embed/tkRB5Rp1_W4"
        frameborder="0"
        allowfullscreen
    ></iframe>
```

35 A more robust logic

- Would need to decouple licence incompatibility from licence name into licence clause (see table below)
- Deal with all licence type
- Provide recommendations based on desired derivative licence type
- Link this through to the type of process in a workflow:
    - data derivation is, from a licence position, very different to contextual display

Table 2: Creative Commons license combinations

License	Reproduction	Distribution	Derivation	BY	SA	NC
CC0	X	X	X			
CC-BY-ND	X	X		X		
CC-BY-NC-ND	X	X		X		X
CC-BY	X	X	X	X		
CC-BY-SA	X	X	X	X	X	
CC-BY-NC	X	X	X	X		X
CC-BY-NC-SA	X	X	X	X	X	X
ODC-PDDL	X	X	X			
ODC-BY	X	X	X	X		
ODC-ODbL	X	X	X	X	X	
OGL 2.0	X	X	X	X		
OS OpenData	X	X	X	X	?	

36 OGC and Licence interoperability

- The geo business landscape is increasingly based on integrating heterogeneous data to develop new products
- Licence heterogeneity is a barrier to data integration and interoperability
- A licence calculus can help resolve and identify heterogenties leading to
  - legal compliance
  - confidence
- Use of standards and collaboration with organisations is crucial
  - Open Data Licensing ontology
  - The Open Data Institute
- Failure to do this could lead to breaches in data licenses
  - and we all know where that puts us……..

Figure 35: Breaching a data license can be serious (restricted = randomly!)

37 Linked data and the Semantic Web

37.1 The web of Documents

- a global filesystem
- Designed for human consumption
- Primary objects are documents
- Expresses links between documents (or sub-parts of)
- Degree of structure in objects is fairly low
- Semantics of content and links is implicit

37.2 The web of Linked Data

- a global database
- Designed for machines first, humans later
- Primary objects are things (or descriptions of things)
- Expresses links between things
- Degree of structure in (descriptions of) things is high
- Semantics of content and links explicit

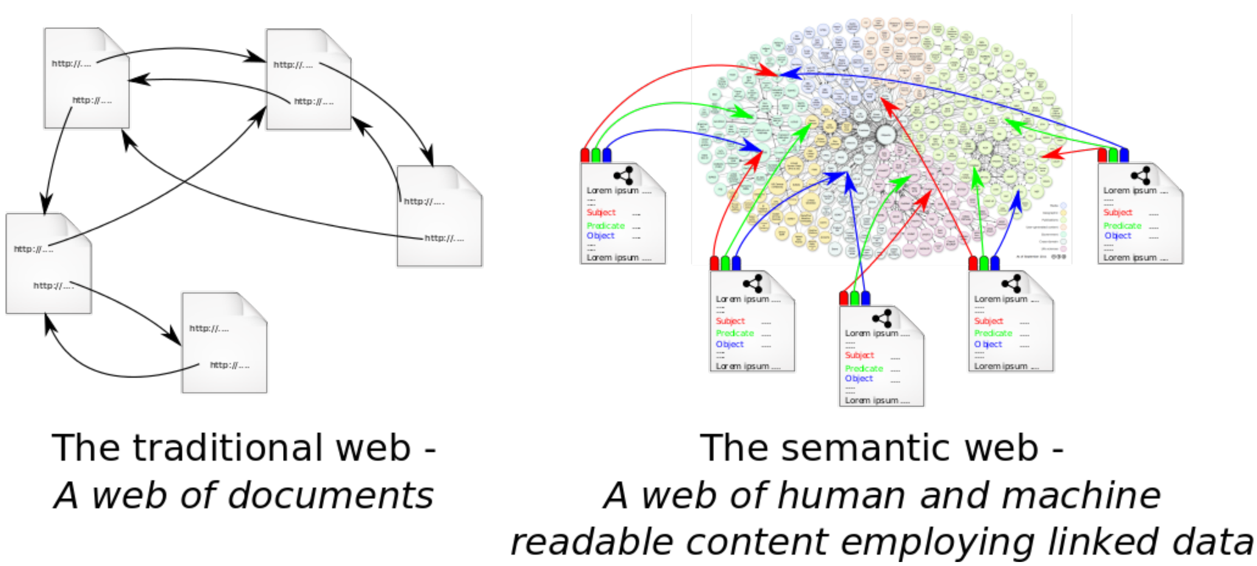

Figure 36: The Semantic Web Compared To The Traditional Web A. Beck (2015i)

38 Linked Data

a way of publishing data on the Web that:

- encourages reuse
- reduces redundancy
- maximises its (real and potential) inter-connectedness
- enables network effects to add value to data

38.1 Why publish Linked Data

- Ease of discovery
- Ease of consumption
    - standards-based data sharing
- Reduced redundancy
- Added value
    - build ecosystems around your data/content

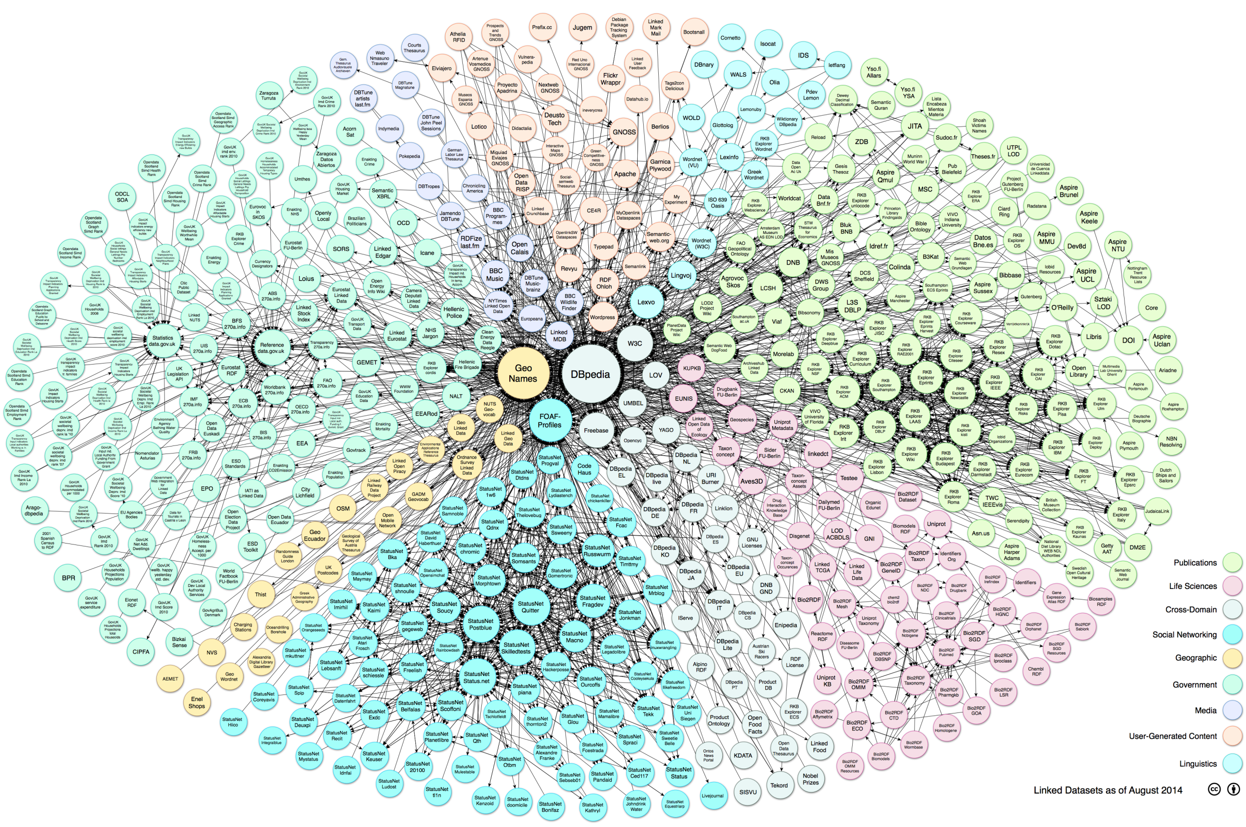

Figure 37: The Linked Open Data cloud in 2014 (Schmachtenberg et al. (2014))

39 Linked Data Basics

39.1 Four rules for Linked Data from Tim Berners Lee

1. Use URIs as names for things
2. Use HTTP URIs so that people can look up those names.
3. When someone looks up a URI, provide useful information, using the standards (RDF*, SPARQL)
4. Include links to other URIs, so that they can discover more things.

40 The Resource Description Framework (RDF) data model

RDF stores data as *triples* in the following manner:

Figure 38:

This is a graph model and edges (predicate).

40.1 Data expressed as RDF

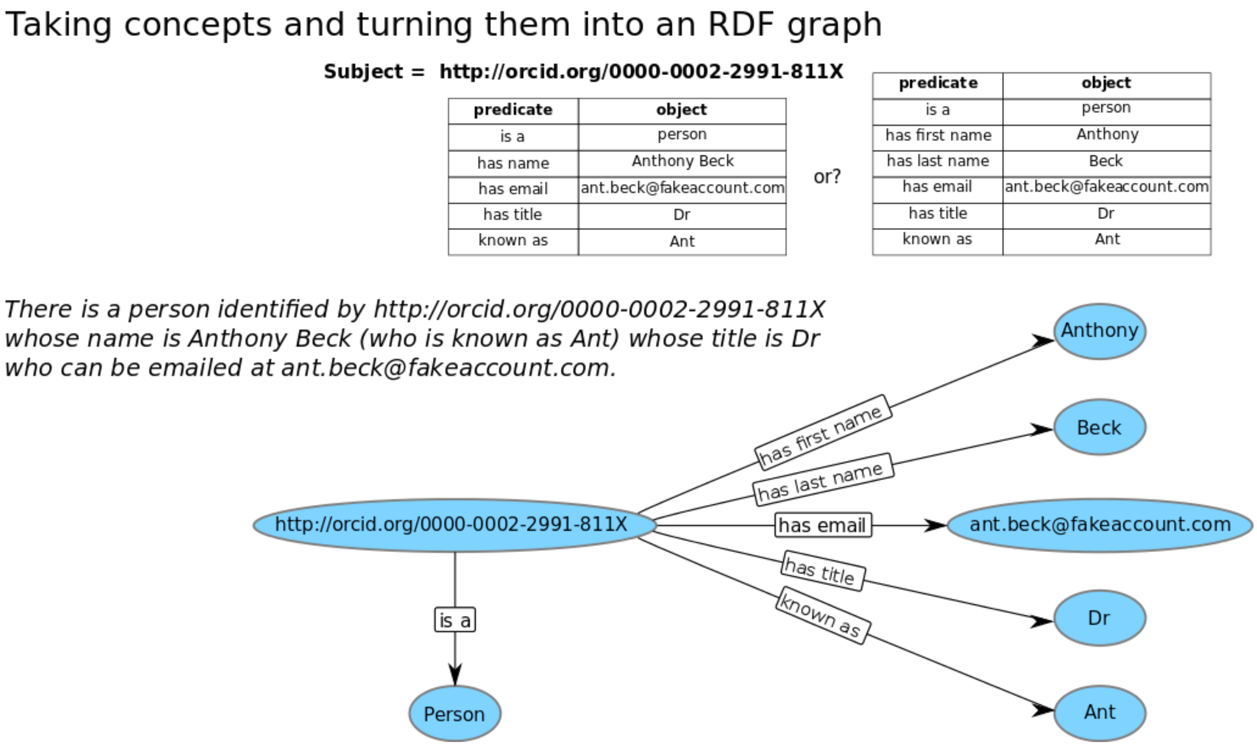

Figure 39:

40.2 Data expressed as RDF Linked Data

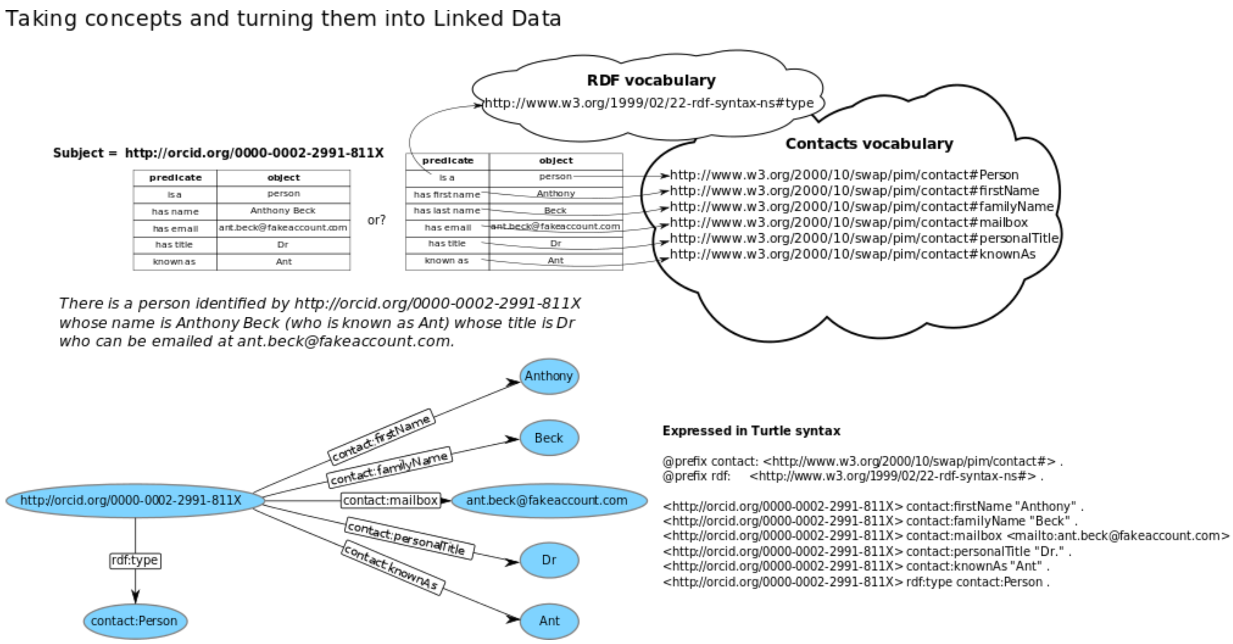

Figure 40:

40.3 RDF notation

RDF can be represented in different ways - each of which are interoperable. For example:

- RDF/XML,
- Notation-3 (N3),
- Turtle (.ttl),
- N-Triples,
- RDFa,
- RDF/JSON

Each represent *subject, predicate, object* triples in different ways

40.4 One step beyond.... Linked Open Data

40.4.1 Is your Linked Open Data 5 star

```
 Available on the web (whatever format) but with an open licence, to be Open Data
Available as machine-readable structured data (e.g. excel instead of image scan of a table)
as (2) plus non-proprietary format (e.g. CSV instead of excel)
  All the above plus, Use open standards from W3C (RDF and SPARQL) to identify things, so that people can p
   All the above, plus: Link your data to other people's data to provide context
```

Figure 41: Is your data 5 star

41 The Supporting Semantic Web Stack

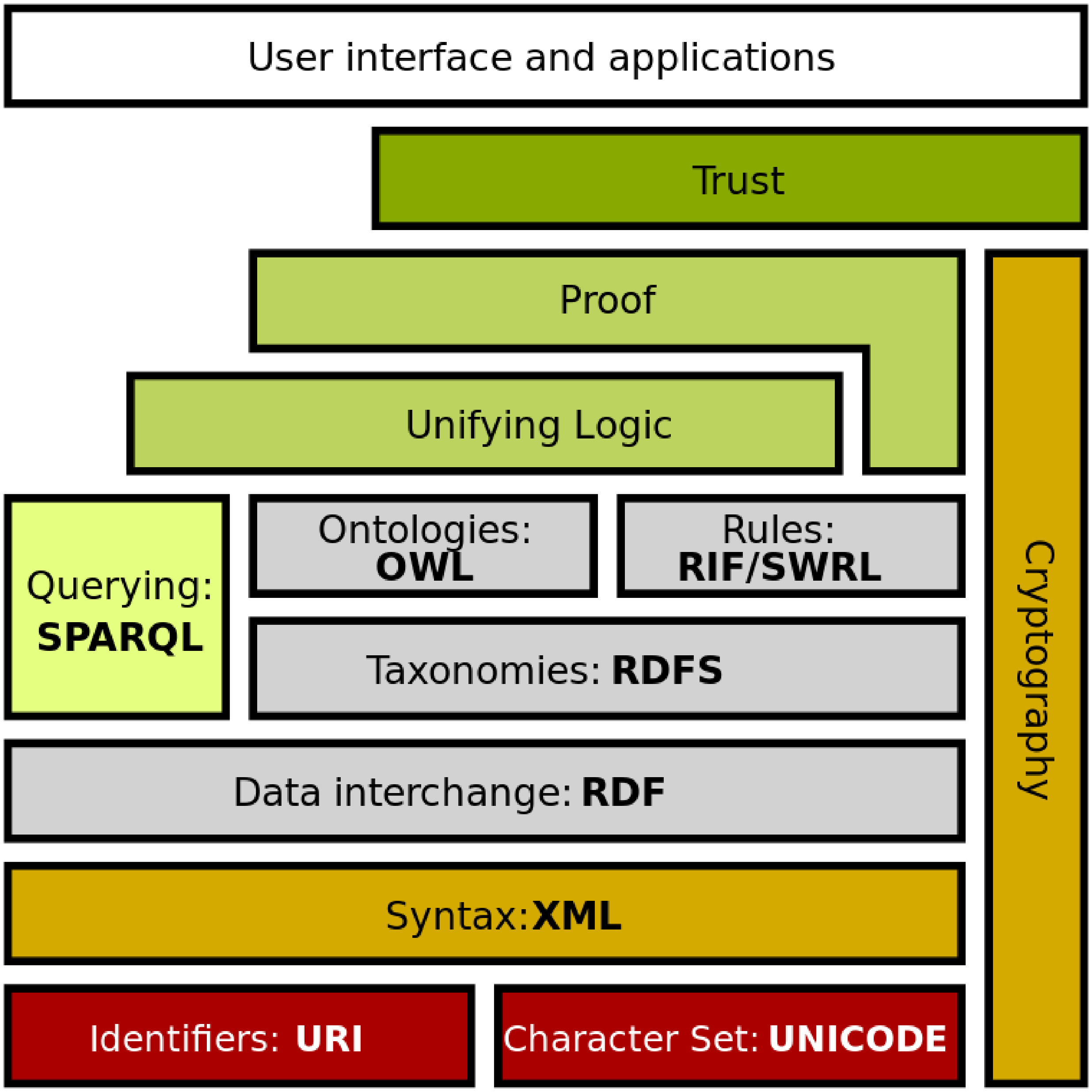

Figure 42: The semantic Web Stack

41.1 It's about re-use

41.1.1 Vocabularies

The glue that joins concepts together.

A concept shared is a link gained. By re-using concepts it makes it easier to understand what your data means and where and how it should be re-used.

```
from IPython.display import IFrame
IFrame('http://lov.okfn.org/dataset/lov/', width=1000, height=700)
```

```
    <iframe
        width="1000"
        height="700"
        src="http://lov.okfn.org/dataset/lov/"
        frameborder="0"
        allowfullscreen
    ></iframe>
```

41.2 It's about re-use

41.2.1 Ontology

An ontology is a shared formal explicit specialisation of a conceptualisation

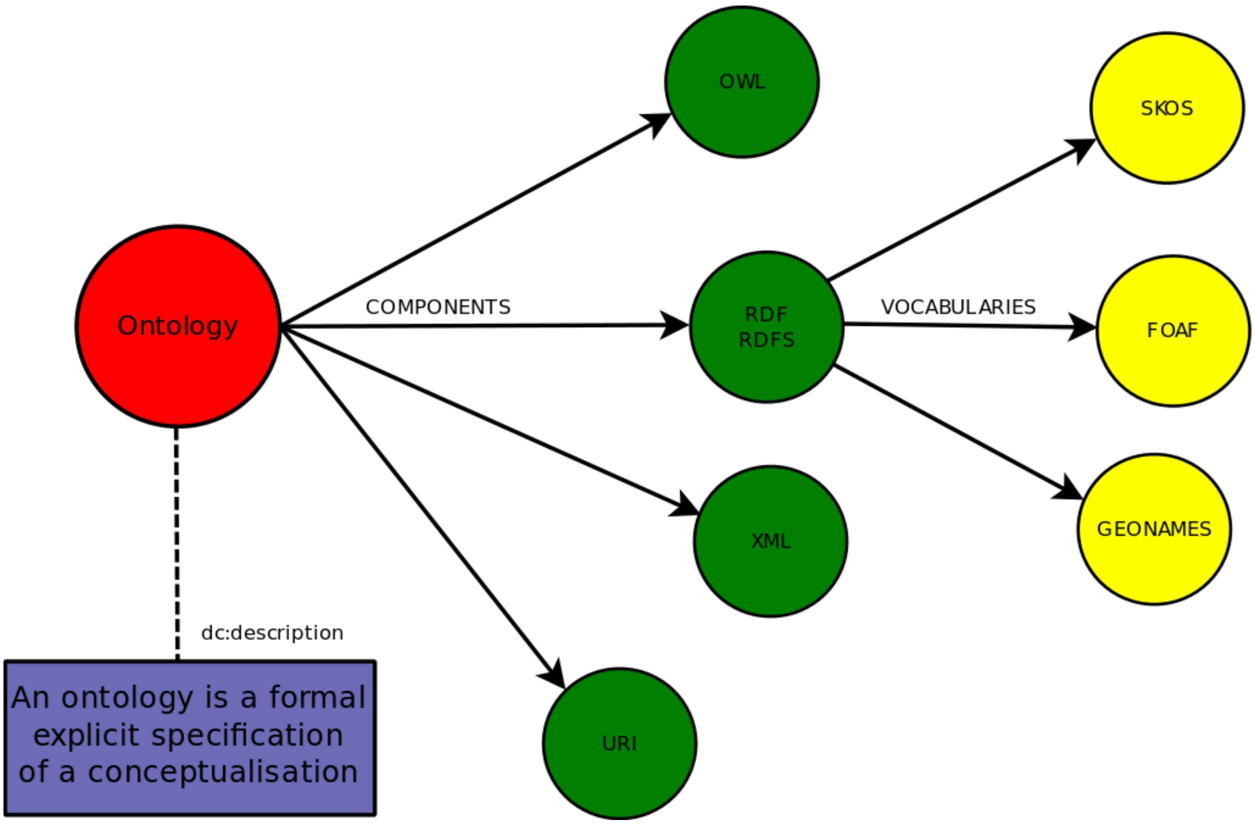

Figure 43: Technology used in Ontology definition(Beck (2012))

41.2.2 Ontology

- The term originated from a philosophy
  - which deals with the nature and organization of reality
- It tries to answer the questions:
  - What is being?
  - What are the features common to all beings?
  - How should things be classified?

41.2.3 Ontology

> An ontology is a shared formal explicit specialisation of a conceptualisation

After Agarwal -(Agarwal (2005)):

- *conceptualisation* is identifying relevant abstracted concepts of a phenomena suited to a specific domain
- *explicit* means that the concepts are explicitly defined
- *formal* refers to the fact that the ontology should be machine-readable
- *shared* refers to notion that on ontology captures consensual knowledge

41.2.4 Ontology Example

- A 'Carnivore' is a concept whose members are exactly those animals who eat only meat
- A 'Bear' is a concept whose members are a kind of 'Carnivore'
- A 'Cub' is a concept whose members are exactly those 'Bear' whose age is less than one year
- A Panda is a individual of a 'Bear'

We can use these concepts to infer new information from facts.

For example: from the fact 'Ching Ching' is a newborn Panda we know:

```
'Ching Ching' is a Panda.
'Ching Ching' is a newborn.
```

We can infer:

```
'Ching Ching' is a Bear.
'Ching Ching' is a Carnivore.  ????
'Ching Ching' eats only meat.  ????
```

If we had other logic that told us that 'newborn' is the same as saying less than one year then we can also infer

```
'Ching Ching' is a Cub.
```

In an ontology you can say *Anything about Anything*. Whilst carnivore is a generally useful concept about *bears* it is not *specifically* useful when considering *pandas*. The domain of application is clearly important.

41.3 SPARQL the SQL of the semantic web

Find me the capital of all countries in Africa:

```
PREFIX abc: <nul://sparql/exampleOntology#> .
SELECT ?capital ?country
WHERE {
  ?x abc:cityname ?capital ;
     abc:isCapitalOf ?y.
  ?y abc:countryname ?country ;
     abc:isInContinent abc:Africa.
}
```

There is a thing ('x') against which the following concepts exist:

- 'abc:cityname' (the name of a city: stored in the variable 'capital')
- 'abc:isCapitalOf' (the concept for which the city is capital: stored in the variable 'y')

The 'concept for which the city is capital' (stored in variable 'y') must also have the following concepts:

- 'abc:countryname' (the name of a country: stored in the variable 'country')
- 'abc:isInContinent' abc:Africa (isInContinent of the the individual Africa')

41.4 GeoSPARQL the SQL of the spatial semantic web

An OGC standard

```
SELECT ?f
WHERE { ?f my:hasPointGeometry ?fGeom .
?fGeom ogc:asWKT ?fWKT .
FILTER (ogcf:relate(?fWKT,
"<http://www.opengis.net/def/crs/OGC/1.3/CRS84>
Polygon ((-83.5 34.0, -83.5 34.3, -83.1 34.3,
-83.1 34.0, -83.5 34.0))"^^ogc:WKTLiteral,
ogc:within))
}
```

```python
from IPython.display import IFrame
IFrame('http://www.opengeospatial.org/projects/groups/geosparqlswg', width=1000, height=700)
```

```
    <iframe
        width="1000"
        height="700"
        src="http://www.opengeospatial.org/projects/groups/geosparqlswg"
        frameborder="0"
        allowfullscreen
    ></iframe>
```

```python
# Linked Data and Geo
```

41.5 GeoSPARQL employs spatial calculus

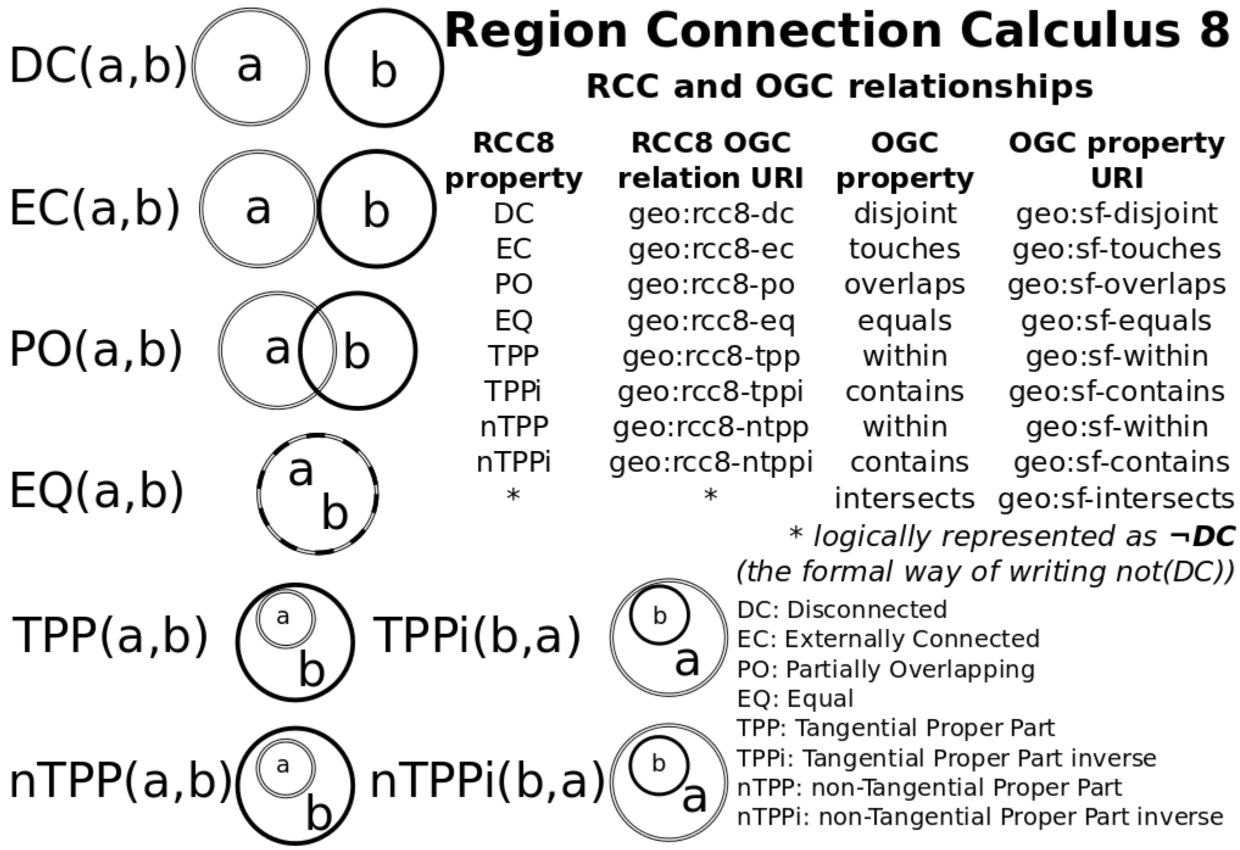

Figure 44:

42 Querying Linked Data in the wild

42.1 The Ordnance Survey

A URI for every place in the UK

```
from IPython.display import IFrame
IFrame('http://data.ordnancesurvey.co.uk/doc/50kGazetteer/177276', width=1000, height=700)
```

```
<iframe
    width="1000"
    height="700"
    src="http://data.ordnancesurvey.co.uk/doc/50kGazetteer/177276"
    frameborder="0"
    allowfullscreen
></iframe>
```

```python
from IPython.display import IFrame
IFrame('http://data.ordnancesurvey.co.uk/id/postcodeunit/NG72QL', width=1000, height=700)
```

```
<iframe
    width="1000"
    height="700"
    src="http://data.ordnancesurvey.co.uk/id/postcodeunit/NG72QL"
    frameborder="0"
    allowfullscreen
></iframe>
```

```python
from IPython.display import IFrame
IFrame('http://data.ordnancesurvey.co.uk/', width=1000, height=700)
```

```
<iframe
    width="1000"
    height="700"
    src="http://data.ordnancesurvey.co.uk/"
    frameborder="0"
    allowfullscreen
></iframe>
```

```python
from IPython.display import IFrame
IFrame('http://data.ordnancesurvey.co.uk/ontology/', width=1000, height=700)
```

```
<iframe
    width="1000"
    height="700"
    src="http://data.ordnancesurvey.co.uk/ontology/"
    frameborder="0"
    allowfullscreen
></iframe>
```

```
from IPython.display import IFrame
IFrame('http://data.ordnancesurvey.co.uk/datasets/code-point-open/explorer/sparql', width=1000, height=7
```

```
    <iframe
        width="1000"
        height="700"
        src="http://data.ordnancesurvey.co.uk/datasets/code-point-open/explorer/sparql"
        frameborder="0"
        allowfullscreen
    ></iframe>
```

```python
## Open Street Map
```

```python
from IPython.display import IFrame
IFrame('http://linkedgeodata.org/About', width=1000, height=700)
```

```
    <iframe
        width="1000"
        height="700"
        src="http://linkedgeodata.org/About"
        frameborder="0"
        allowfullscreen
    ></iframe>
```

```python
from IPython.display import IFrame
IFrame('http://browser.linkedgeodata.org/', width=1000, height=700)
```

```
    <iframe
        width="1000"
        height="700"
        src="http://browser.linkedgeodata.org/"
        frameborder="0"
        allowfullscreen
    ></iframe>
```

42.2 Geonames

```python
from IPython.display import IFrame
IFrame('http://www.geonames.org/ontology/documentation.html', width=1000, height=700)
```

```
    <iframe
        width="1000"
        height="700"
        src="http://www.geonames.org/ontology/documentation.html"
        frameborder="0"
        allowfullscreen
    ></iframe>
```

```python
from IPython.display import IFrame
IFrame('http://www.geonames.org/maps/google_52.94_358.8.html', width=1000, height=700)
```

```
    <iframe
        width="1000"
        height="700"
        src="http://www.geonames.org/maps/google_52.94_358.8.html"
        frameborder="0"
        allowfullscreen
    ></iframe>
```

```python
from IPython.display import IFrame
IFrame('http://lov.okfn.org/dataset/lov/vocabs/gn', width=1000, height=700)
```

```
    <iframe
        width="1000"
        height="700"
        src="http://lov.okfn.org/dataset/lov/vocabs/gn"
        frameborder="0"
        allowfullscreen
    ></iframe>
```

43 Geo Vocabularies

```
from IPython.display import IFrame
IFrame('http://lov.okfn.org/dataset/lov/vocabs/?q=geo+space+address+geonames+os+spatial', width=1000, he
```

```
    <iframe
        width="1000"
        height="700"
     src="http://lov.okfn.org/dataset/lov/vocabs/?q=geo+space+address+geonames+os+spatial"
        frameborder="0"
        allowfullscreen
    ></iframe>
```

44 Conclusions

- Technical interoperability is only one part of the problem
- Open data will become increasingly important as governments and other groups release resources under clear licences
  - Licences are a barrier to re-use
- Data shows its true value when combined with other data sources – linked data creates an opportunity
- Usability: common data model and reference of common URIs (for example, postcodes) allows for easy data aggregation and integration.
- Shift in focus from cartography and geometries to 'things' and the relationships between them.
- Spatial no longer special – part of the bigger information world….
- location is a very important information hub and provides a key underpinning reference framework which brings many datasets together and provides important context.

45 Geo reasoning example (if time)

Geo example:

```
Leeds is a city.

Yorkshire is a county.

Sheffield is a city.

Lancaster is a city.

Lancashire is a county.

Lancaster has a port.

What is Leeds?

Leeds isIn Yorkshire.

Sheffield isIn Yorkshire.

Lancaster isIn Lancashire.

What isIn Yorkshire?

If X isIn Y then Y contains X.

What contains Leeds?

Yorkshire borders Lancashire.

If X borders Y then Y borders X.

What borders Lancashire?

Yorkshire isIn UnitedKingdom.

Lancashire isIn UnitedKingdom.

#Transitivity

If X isIn Y and Y isIn Z then X isIn Z.

If X contains Y and Y contains Z then X contains Z
```

using proper isIn

```
Leeds is a city.

Yorkshire is a county.
```

```
Sheffield is a city.

Lancaster is a city.

Lancashire is a county.

Lancaster has a port.

What is Leeds?

Leeds is spatiallyWithin Yorkshire.

Sheffield is spatiallyWithin Yorkshire.

Lancaster is spatiallyWithin Lancashire.

What is spatiallyWithin Yorkshire?

If X is spatiallyWithin Y then Y spatiallyContains X.

What spatiallyContains Leeds?

Yorkshire borders Lancashire.

If X borders Y then Y borders X.

What borders Lancashire?

Yorkshire is spatiallyWithin UnitedKingdom.

Lancashire is spatiallyWithin UnitedKingdom.

#Transitivity

If X is spatiallyWithin Y and Y is spatiallyWithin Z then X is spatiallyWithin Z.

If X spatiallyContains Y and Y spatiallyContains Z then X spatiallyContains Z

What is spatiallyWithin UnitedKingdom?
```

Adding more……

```
Pudsey is spatiallyWithin Leeds.

Kirkstall is spatiallyWithin Leeds.

Meanwood is spatiallyWithin Leeds.

Roundhay is spatiallyWithin Leeds.

Scarcroft is spatiallyWithin Leeds.
```

and more

```
UnitedKingdom isPartOf Europe.

UnitedKingdom is a country.

If X isPartOf Y and X spatiallyContains Z then Z isPartOf Y.

What isPartOf Europe?
```

and `more`

If X spatiallyContains Y and X is a city then Y is a place and Y is a cityPart.

Every city is a place.

What is a place.

and `more`

UK isPartOf Europe.

UK is sameAs UnitedKingdom.

If X has a port then X borders Water.

What borders Water?

46 Questions

In terms of discussion I'm interested in how these issues affect you………

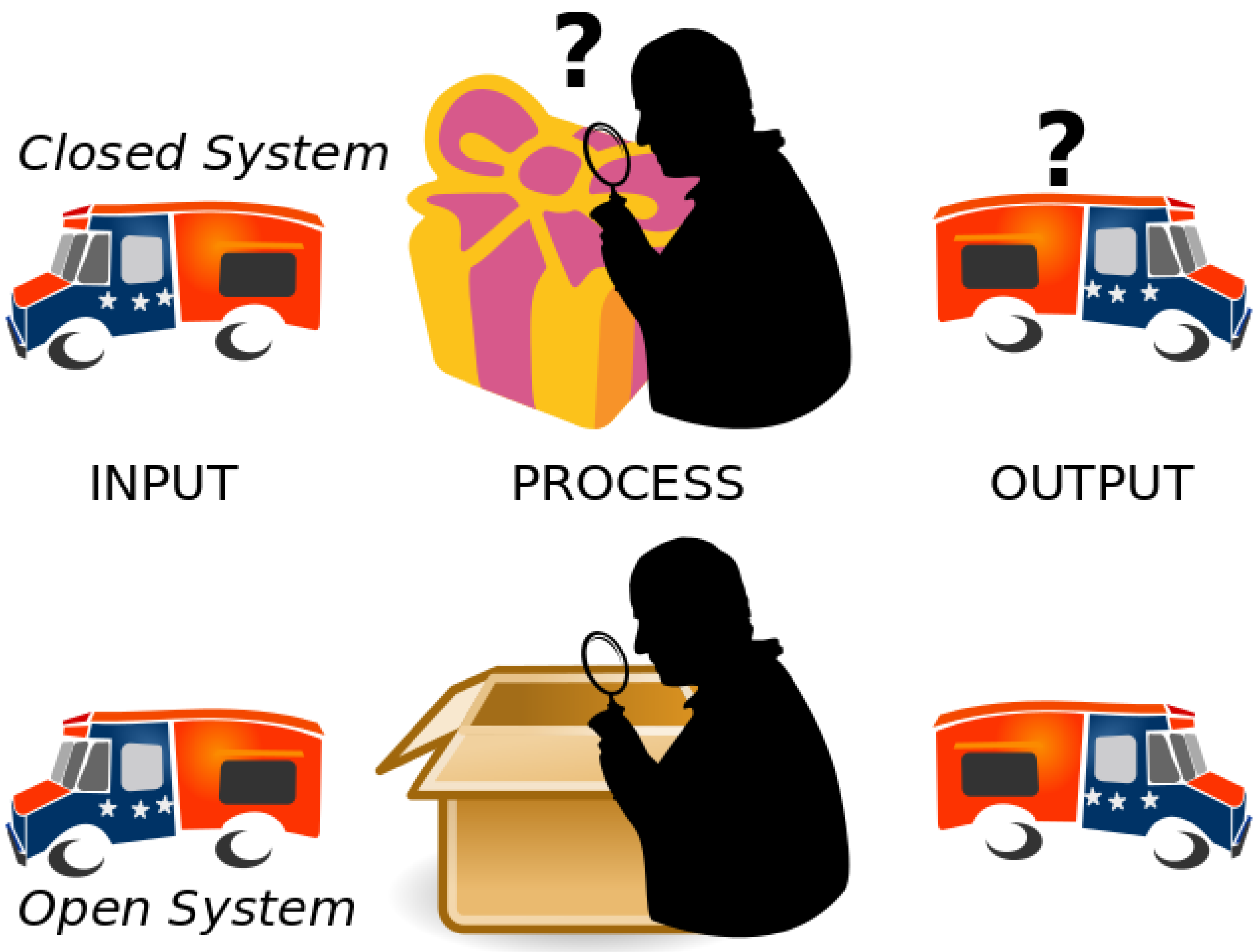

Figure 45: Processing transparency between open and closed systems

References

Agarwal, P., 2005. Ontological considerations in GIScience. *International Journal of Geographical Information Science*, 19(5), pp.501–536. Available at: http://dx.doi.org/10.1080/13658810500032321 [Accessed January 24, 2016].

AlisonW, 2014. Open knowledge logo. Available at: https://commons.wikimedia.org/wiki/File:Open_Knowledge_logo.png [Accessed December 18, 2014].

Anon, RDFLicense - the datahub. Available at: http://datahub.io/dataset/rdflicense [Accessed September 10TZ, 2015].

Beck, A., 2015a. A conceptual licence processing workflow. Available at: https://commons.wikimedia.org/wiki/File:ConceptualLicenceProcessingWorkflow.svg [Accessed September 10, 2015].

Beck, A., 2015b. Cartography is no longer key. Available at: https://commons.wikimedia.org/wiki/File:CartographyNoLongerKing.svg [Accessed September 7, 2015].

Beck, A., 2015c. From map to model the changing paradigm of map creation from cartography to data driven visualization. Available at: https://commons.wikimedia.org/wiki/File:FromMapToModel.svg [Accessed September 7, 2015].

Beck, A., 2016. Incompatibilities and licence clauses. Available at: https://commons.wikimedia.org/wiki/File:Incompatibilities_And_Licence_Clauses.svg [Accessed January 9, 2016].

Beck, A., 2015d. Interoperable integration of spatial data - the technological issues. Available at: https://commons.wikimedia.org/wiki/File:TechnicalInteroperableIntegrationOfSpatialData.svg [Accessed September 7, 2015].

Beck, A., 2015e. Islands of data. Available at: https://commons.wikimedia.org/wiki/File:Islands_Of_Data.svg [Accessed December 29, 2015].

Beck, A., 2015f. Local to global integration of data to create multiple generic products. Available at: https://commons.wikimedia.org/wiki/File:Local_To_Global_integration_of_data_to_create_multiple_generic_products.svg [Accessed September 7, 2015].

Beck, A., 2012. Ontology and technology. Available at: https://commons.wikimedia.org/wiki/File:SemWebOverview.svg [Accessed October 30, 2012].

Beck, A., 2013. Open science does not equal open access. Available at: https://commons.wikimedia.org/wiki/File:Open_Science_Does_Not_Equal_Open_Access.svg [Accessed February 23, 2015].

Beck, A., 2015g. The full stack that enables interoperable integration of spatial data. Available at: https://commons.wikimedia.org/wiki/File:InteroperableIntegrationOfSpatialData.svg [Accessed September 7, 2015].

Beck, A., 2015h. The lens of cartography. Available at: https://commons.wikimedia.org/wiki/File:TheLensOfCartography.svg [Accessed September 7, 2015].

Beck, A., 2015i. The semantic web compared to the traditional web. Available at: https://commons.wikimedia.org/wiki/File:The_Semantic_Web_Compared_To_The_Traditional_Web.svg [Accessed December 29, 2015].

Beck, A., 2015j. Types of data. Available at: https://commons.wikimedia.org/wiki/File:Data_types_-_en.svg [Accessed September 7, 2015].

Beck, A.R. et al., 2008. UK utility data integration: Overcoming schematic heterogeneity. In *Proceedings of GeoInformatics2008*. Gaizhuong, China.

Bishr, Y., 1998. Overcoming the Semantic and Other Barriers to GIS Interoperability. *International Journal of Geographical information Science*, 12(4), pp.299–314.

cleonis, 2006. Nonsymmetric velocity time dilation. Available at: https://commons.wikimedia.org/wiki/File:

Nonsymmetric_velocity_time_dilation.gif [Accessed January 28, 2006].

Foundation, W., 2010. Wikipedia logo v2 en. Available at: https://commons.wikimedia.org/wiki/File:Wikipedia-logo-v2-en.png [Accessed May 12, 2010].

Gianni, S., 2008. We love creative commons. Available at: https://www.flickr.com/photos/steren/2732488224/ [Accessed July 27, 2015].

Innocenti, P. et al., 2011. Towards a Holistic Approach to Policy Interoperability in Digital Libraries and Digital Repositories. *International Journal of Digital Curation*, 6(1), pp.111–124. Available at: http://www.ijdc.net/index.php/ijdc/article/view/167 [Accessed January 22, 2016].

McKinsey, 2013. Open data: Unlocking innovation and performance with liquid information. Available at: http://www.mckinsey.com/insights/business_technology/open_data_unlocking_innovation_and_performance_with_liquid_information [Accessed November 9, 2013].

Murray-Rust, P. et al., 2010. Panton Principles, Principles for open data in science. Available at: http://pantonprinciples.org/ [Accessed January 17, 2016].

Rillke, 2014. Open street map logo. Available at: https://commons.wikimedia.org/wiki/File:Openstreetmap_logo.svg [Accessed July 14, 2014].

Schmachtenberg, M. et al., 2014. The Linking Open Data cloud diagram 2014. Available at: http://lod-cloud.net/ [Accessed January 24, 2016].

SEWilco, 2013. GeoServer geoNetwork with web app. Available at: https://commons.wikimedia.org/wiki/File:GeoServer_GeoNetwork_with_web_app.svg [Accessed July 10, 2013].

Shakespeare, S., 2013. *Shakespeare review of public sector information*, Available at: https://www.gov.uk/government/publications/shakespeare-review-of-public-sector-information [Accessed July 2, 2013].

The Royal Society, 2012. *Science as an open enterprise*, The Royal Society. Available at: http://royalsociety.org/policy/projects/science-public-enterprise/report/ [Accessed September 24, 2012].

Watson, C., 2010. Joint Interoperability Certification: What the Program Manager should know. Available at: http://www.dau.mil/pubscats/atl%20docs/jan-feb/watson_jan-feb10.pdf [Accessed January 17, 2016].

Weber, R., 2014. Legal Interoperability as a Tool for Combatting Fragmentation. Available at: http://issuu.com/cigi/docs/gcig_paper_no4/1 [Accessed January 22, 2016].

Wheeler, D.A., 2007. Compatibility of common open-source software licenses. Available at: https://en.wikipedia.org/wiki/File:Floss-license-slide-image.png [Accessed July 27, 2015].

Wikipedia, 2016. Interoperability. *Wikipedia, the free encyclopedia.* Available at: https://en.wikipedia.org/w/index.php?title=Interoperability&oldid=698683288 [Accessed January 17, 2016].