

UNIVERSITY OF
Southampton

Vocabularies and Applications

COMP6215 Semantic Web Technologies

Dr Nicholas Gibbins - nmg@ecs.soton.ac.uk

Overview

- Social Networking
- Bibliographic Metadata
- Cultural Heritage
- Knowledge Organisation
- Provenance
- Web Annotation
- Public Sector Information
- Geographic Information

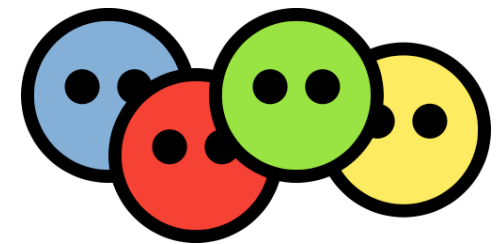


Social Networking

<http://www.flickr.com/photos/bigblackbox/4269397346/>

Friend of a Friend (FOAF)

- Semantic Web realisation of “six degrees of separation”
- Fully distributed system
 - Each user writes an RDF file describing themselves and who they know (using the FOAF schema)
 - Files are harvested and aggregated
 - Data is presented in a variety of interfaces
- Avoids siloing of existing social networking systems
- Data from many providers



Multiple Interfaces

The image displays two browser windows illustrating FOAF (Friend of a Friend) web views.

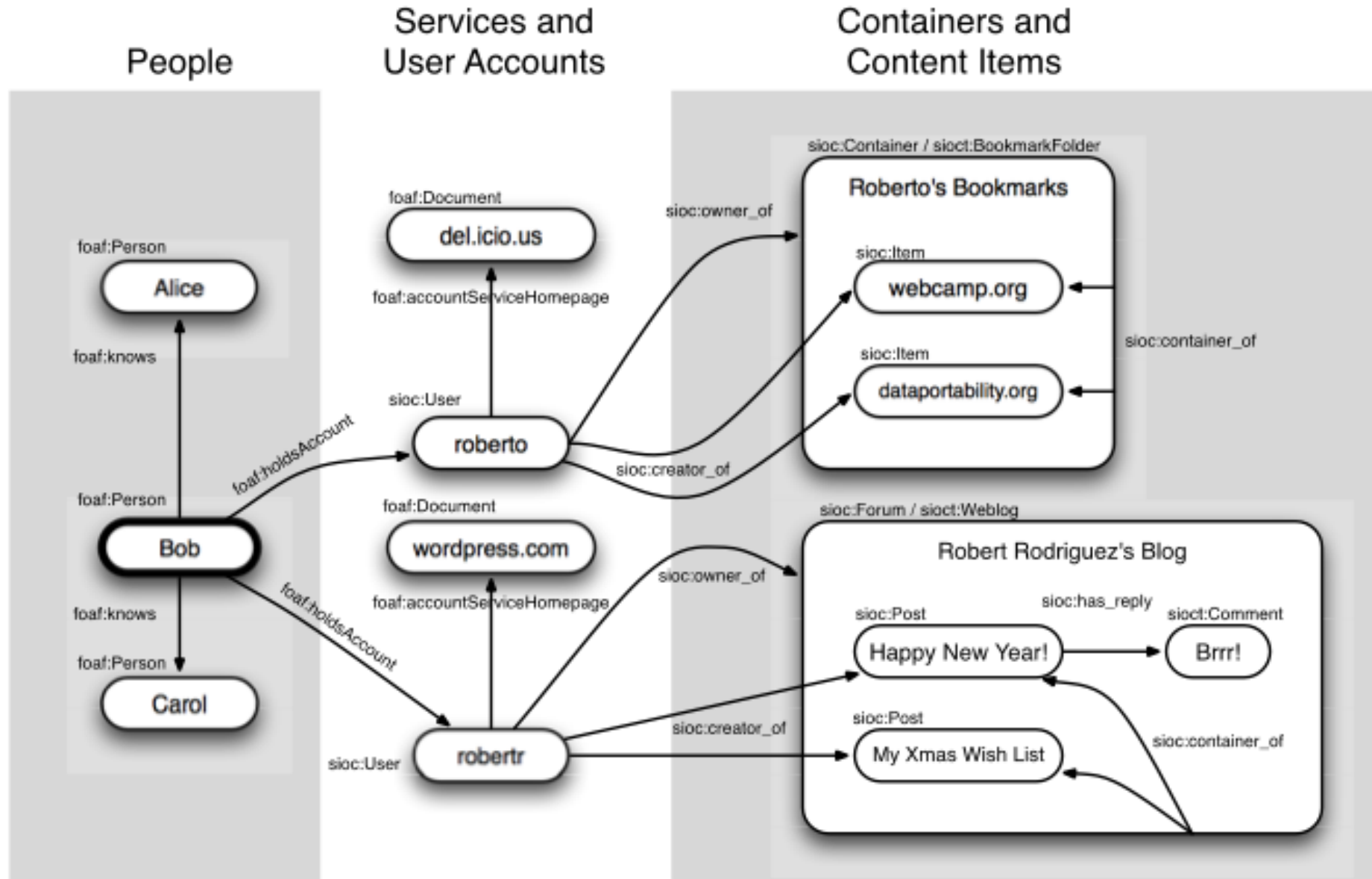
Left Window (Mozilla): Shows a text-based FOAF profile for Nick Gibbins. The profile includes:

- FOAF: Web View** (with a FOAF icon)
- Nick Gibbins | Nick Gibbin | Nicholas Gibbins**
- Show depictions**
- Nickname:** nmg
- Homepage:** <http://purl.org/net/ngibbins/>
- workplaceHomepage:**
 - <http://www.soton.ac.uk>
 - <http://www.iam.ecs.soton.ac.uk>
 - <http://www.ecs.soton.ac.uk>
- A current project this person works on:**
 - [_VZaVRjew4126](#)
 - [_VrHOMgya2612](#)
- Knows:**
 - Brian Kelly
 - Dan Brickley
 - Ian Hickson
 - Isobel Stark
 - Libby Miller | libby | Libby Miller
 - Max Froumentin

At the bottom, it states: "This site is running on Redfoot 1.3.0 / RDFLib 1.3.0 and is powered by RDF and W3C." Logos for RDF and W3C are visible.

Right Window (Internet Explorer): Shows a graphical network view of Nick Gibbins' FOAF profile. The central node is NMG (Nick Gibbins), represented by a blue circle with a FOAF icon. He is connected to several other nodes, each represented by a smaller circle with a name and a FOAF icon: IAN, NICK, BERT, LIDDY, SRIAN, LIBBY, DANBRI, ISOBEL, JIBBER JIM, CHARLES, MAX, and NADIA. A small inset window shows a photo of Nick Gibbins. A tooltip for NMG provides the following information:

- NMG**
- NICK GIBBINS**
- nmg@ecs.soton.ac.uk
- nick@totl.net
- nick@gark.net
- Knows:** 6
- Codepictions:** 4
- Known by:** 1
- [homepage](#)



568

GAL. DE TEST

eiusdem sunt generis,
 querint musculis: quid
 que sinum magnum ad
 ioriora fuissent?
 tumentum, quod omnia

Bibliographic Metadata

<http://www.flickr.com/photos/celesterc/1069893367/>

Dublin Core

- A standard set of 15 metadata elements for cross-domain information resource description
- Formally endorsed by ISO, NISO and CEN
- Representation language-neutral
 - Plain XML serialisations in addition to RDF/XML
 - Predates RDF (started in 1995)
- Possibly the most widely used Semantic Web vocabulary
 - Libraries (electronic and physical)
 - Museums
 - Web archives
 - Open Directory Project
 - UK Mirror Service
 - MusicBrainz
 - ...



Dublin Core Element Set

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

Bibliographic Ontology

- More detailed bibliographic ontology
- Describes types of publication
 - Conference paper
 - Journal article
 - ...
- Builds on Dublin Core



MusicBrainz

- An open replacement for the Gracenote CD database
- Community-maintained service
 - Submissions and moderations
- Vocabularies used:
 - Dublin Core
 - Custom ontology for metadata specifically relating to audio recordings



Publishing Standards for Industry Standard Metadata

- Set of metadata vocabularies for automating publishing production processes and content exchange
- Simplified profile of RDF/XML
 - Lower hurdle to adoption
- Commercial users include Time, Inc., Lexis-Nexis and McGraw-Hill
- Vocabularies used:
 - Dublin Core
 - Custom ontologies for
 - Controlled vocabularies
 - Relational information
 - Resource types
 - Rights management

BBC Programmes

- Simple ontology for describing programmes
 - Brands
 - Series (seasons)
 - Episodes
 - Broadcast events
 - Broadcast services
- Data describing all BBC programmes



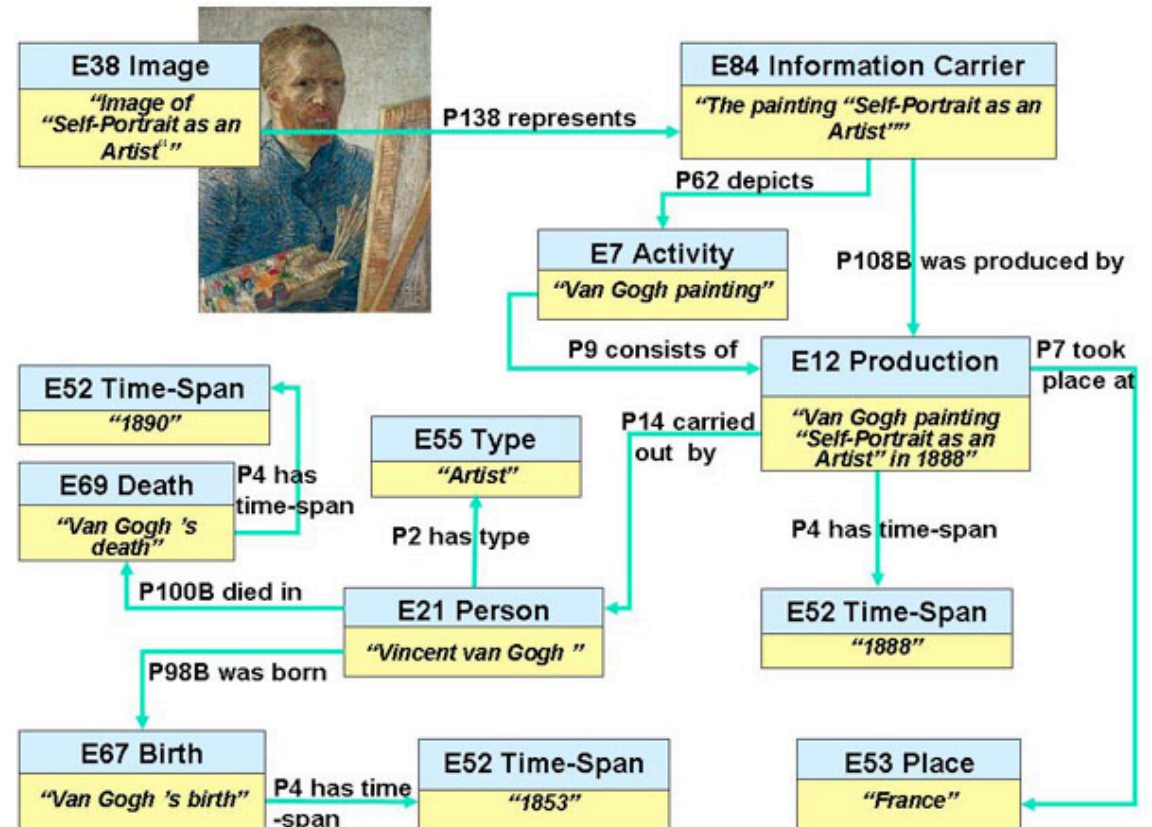


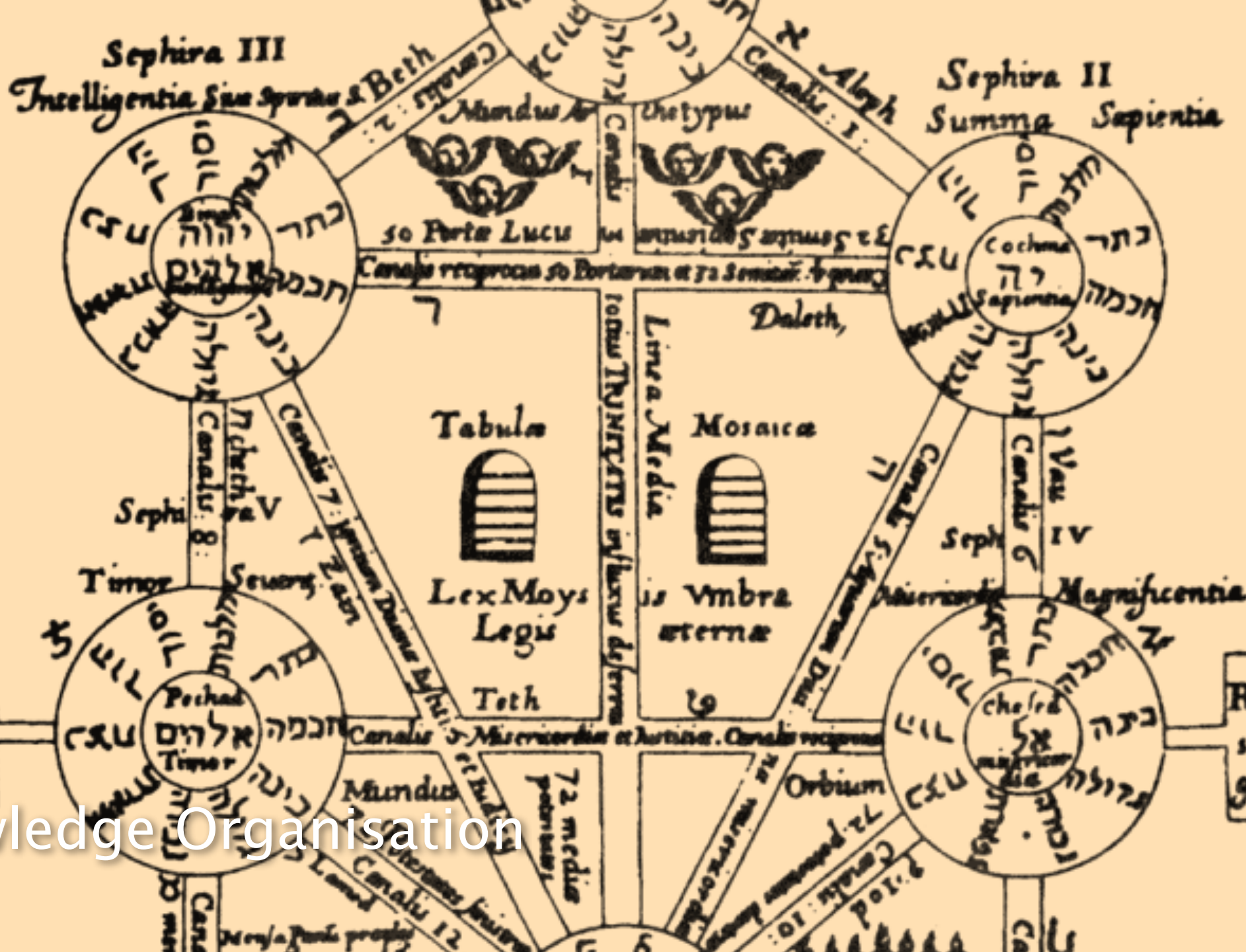
Cultural Heritage

<http://www.flickr.com/photos/juanillooo/476627356/>

CIDOC CRM

- Ontology for heterogeneous cultural heritage information
- Represents more than just bibliographic information
- Event-based model (compare with Dublin Core's attribute-based approach)





Knowledge Organisation

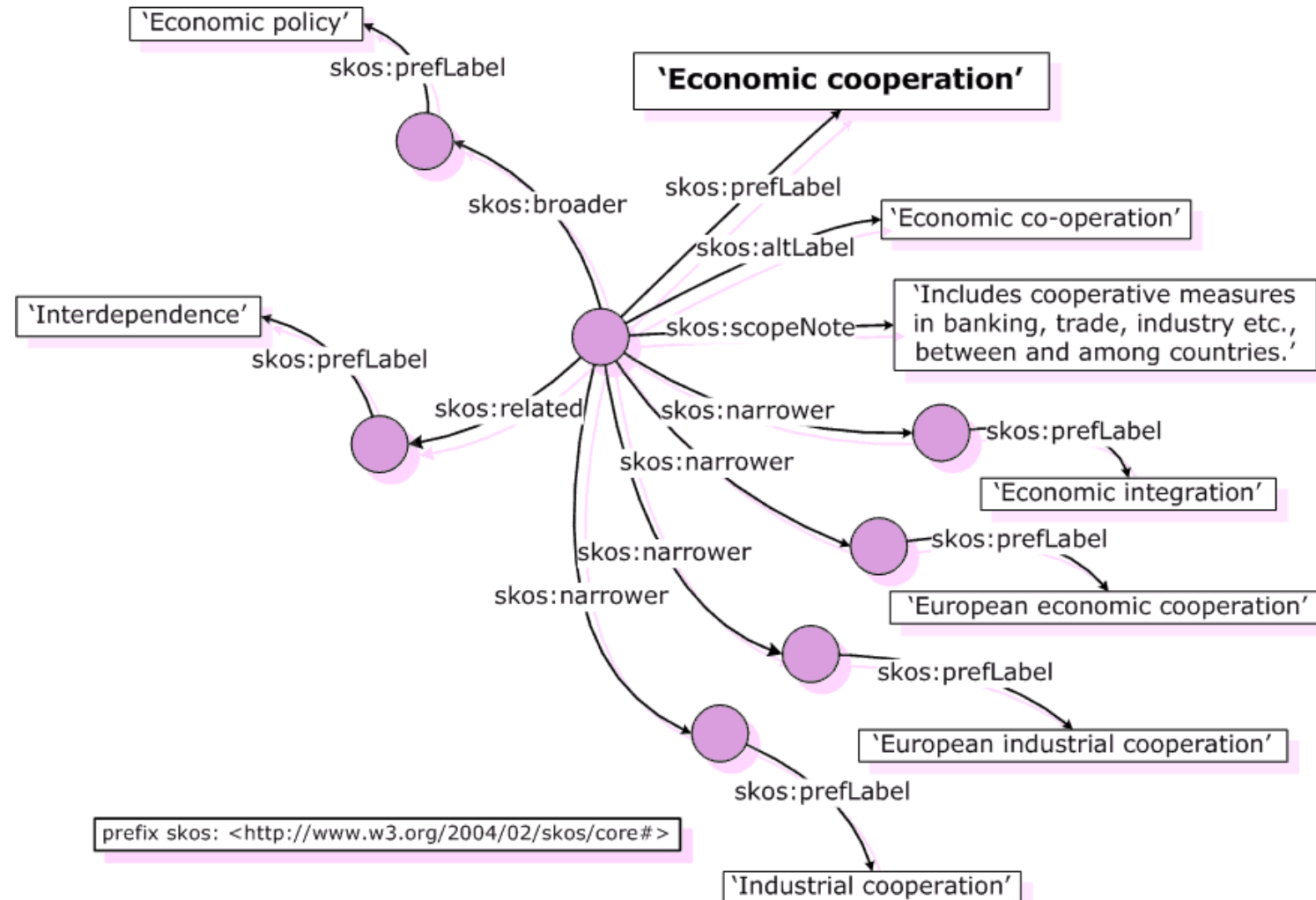
35
Principes a se
veritate origi-
nem ducentes

35
Principes a mi-
sericordia ori-
ginem ducentes

Simple Knowledge Organisation System

- Standard vocabulary to support thesauri, classification schemes, subject heading lists, taxonomies, and other types of controlled vocabulary
- Semantic relationships: broader, narrower, related terms
- Labels, preferred labels, alternative labels
- Notes, documentation, scope notes
- Mappings between vocabularies

SKOS Example



181 The Fox Hunt oil 38x68

1893

OEDDE met

4

N

Rec. W. Homer. June 15, 1893.

K836114

Sold B3105 Nov. 21, 1893

Penn. Acad of the Fine Arts.

met. Mus. Feb-Mar 1911. Cent P.A. of the F.A.

Provenance 85

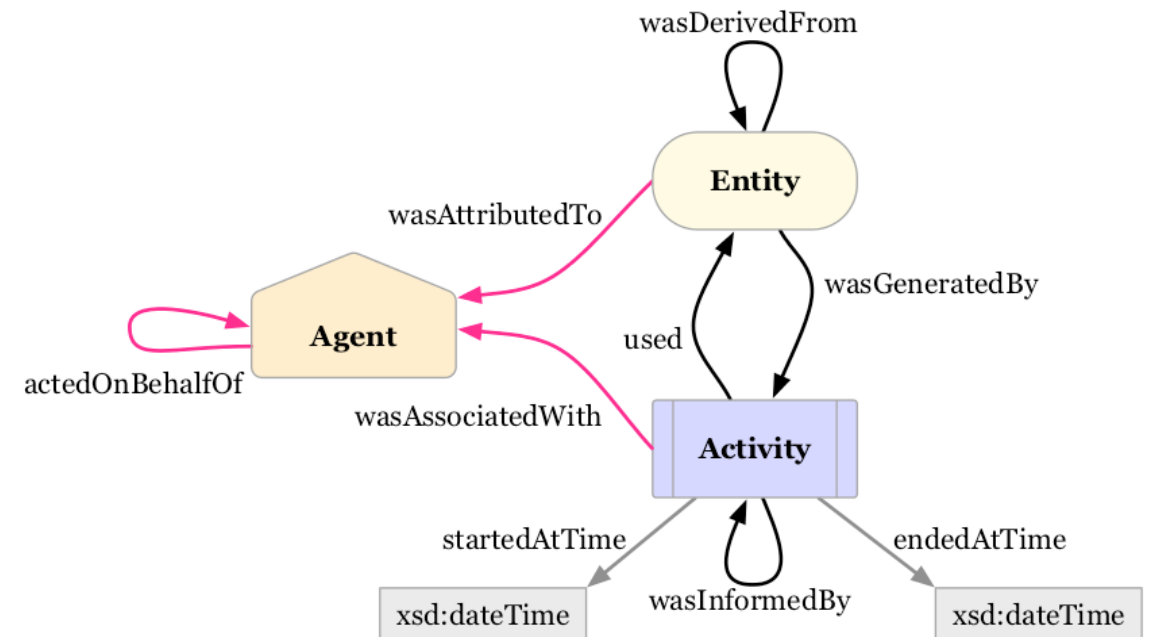
W.H.D. p 168

Provenance

- The *provenance* of a digital object represents its origin
- Provenance facilitates:
 - understanding how data was collected so it can be meaningfully used
 - determining ownership and rights over an object
 - making judgements about information to determine whether to trust it
 - verifying that the process and steps used to obtain a result complies with given requirements
 - reproducing how something was generated.

PROV-O

- A specification and ontology for expressing provenance records
- Records contain descriptions of the entities and activities involved in producing and delivering or otherwise influencing a given object
- Became a W3C Recommendation in April 2013



has a gift of seek on site, uncausally
dreamydeary. Heed! Heed! It may half been a missfired brick, as
some say, or it mought have been due to a collupsus of his back
promises, as others looked at it. (There extand by now one thou-
sand and one stories, all told, of the same). But so sore did abe
ite ivvy's holired abbles, (what with the wallhall's horrors of rolls-
rights, carhacks, stonengens, kisstvanes, tramtrees, fargobawlers,
autokinotons, hippohobbilies, streetfleets, tournintaxes, mega-
phoggs, circuses and wardsmoats and basilikerks and aeropagods
and the hoyse and the jollybrool and the peeler in the coat and
the mecklenburk bitch bite at his ear and the merlinburrow bur-
rocks and his fore old porecourts, the bore the more, and his

nam
outside
if own
destruction
sking
wings
over 75k
it.
any
pub
any
reshops
isell

refer here
unsure
falling up

Indiana
moving
may 21
the set
to
America
compared
the oscar
over 500
muse
wanderer

could also
be circle
mecklenburg
(plymouth?)
5

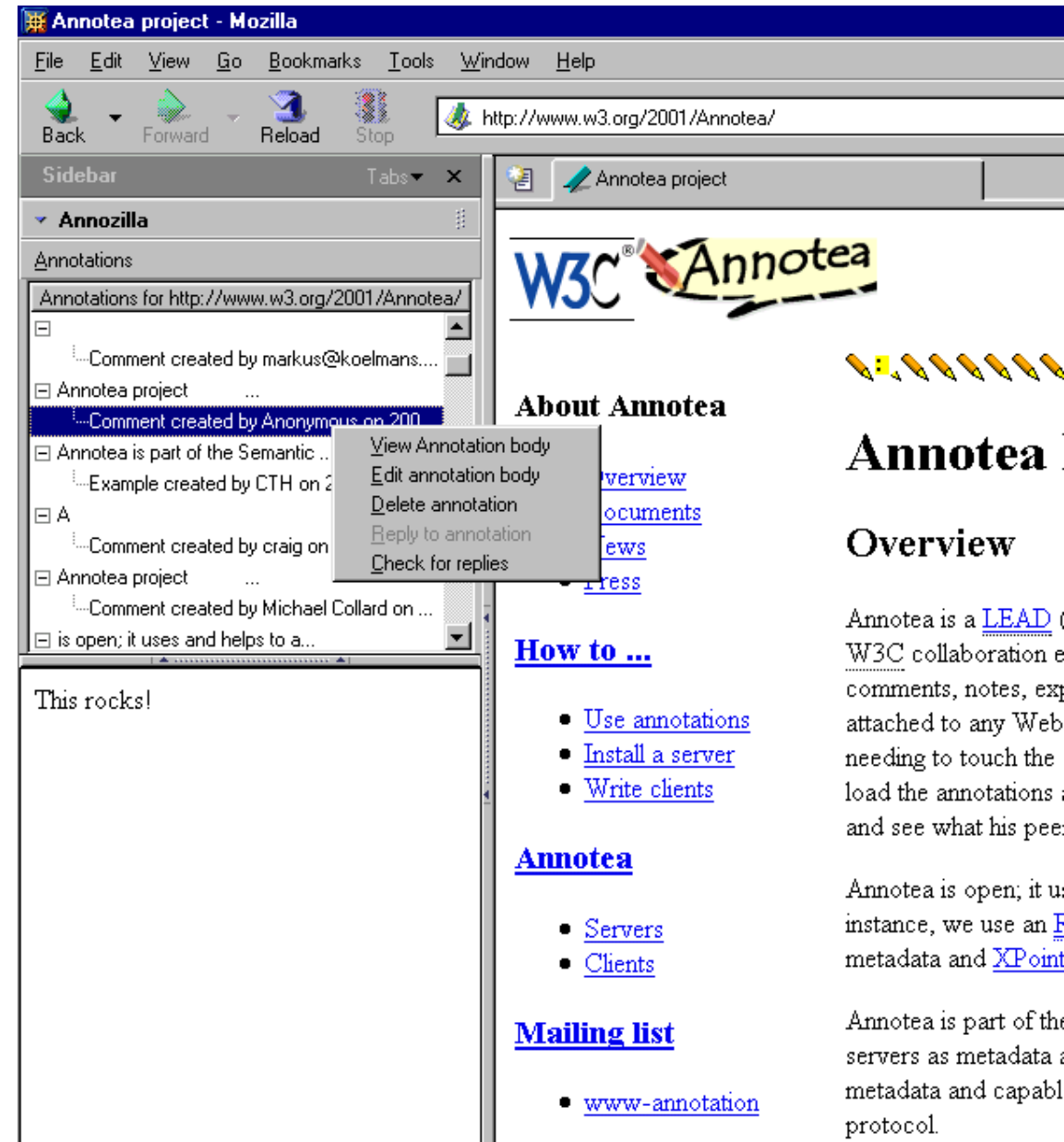
buried
cece
Anfanden
legends

and travel
through sea:
Othello's
"seraphy"

Web Annotation

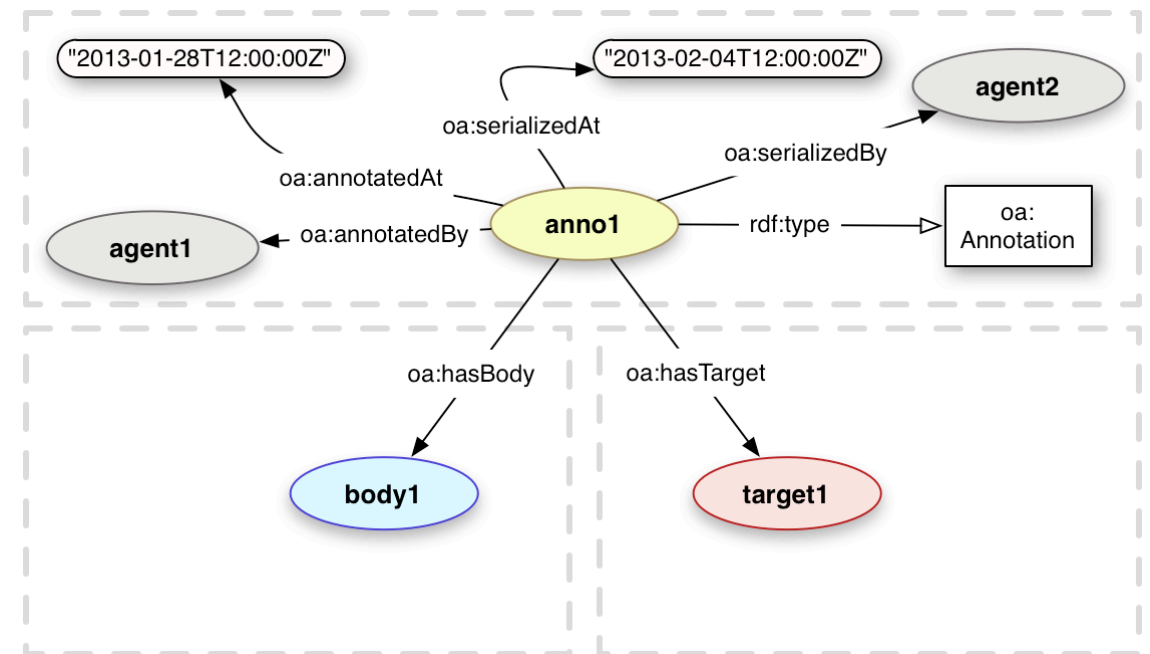
Annotea

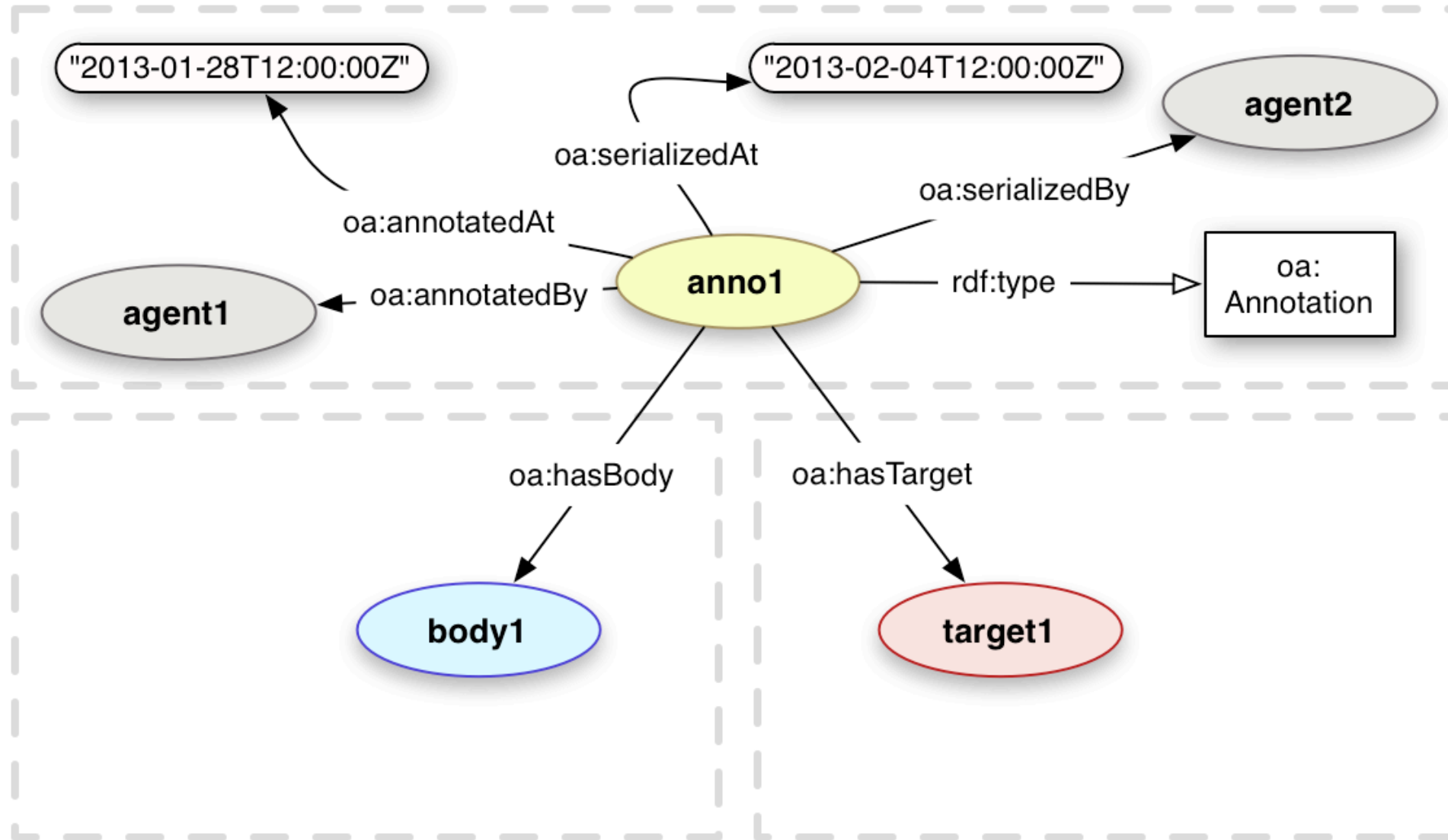
- Shared annotations for the Web
- Vocabularies used:
 - Dublin Core for bibliographic metadata about the annotation
 - Custom schema for relating annotation to annotated document
 - XPath and XPointer used to identify region of document to be annotated



Open Annotation Data Model

- W3C Proposed Recommendation (17 Jan 2017)
 - Ongoing work on annotation within W3C (Web Annotation Working Group)
- Vocabularies used:
 - FOAF
 - Dublin Core
 - Provenance
 - SKOS







Commerce

<https://www.flickr.com/photos/polycart/5786374464>

Good Relations

- Standardised vocabulary for product, price, and company data
- Represents:
 - Prices
 - Accepted payment methods
 - Opening times
 - Delivery methods and charges
 - Warranty promises and scope
- Integrates with:
 - UNSPSC vocabulary for offerings
 - SIC codes for company type





Public Sector Information

<http://www.flickr.com/photos/emsef/1398072554/>

AKTive PSI (2005)

- Use advanced knowledge management technology to improve the delivery of policy and public services across Government
- Build up a detailed picture of life in two London Boroughs, using as broad a collection of Public Sector Information as possible



Office of
PUBLIC SECTOR INFORMATION



Land Registry
Cymraeg



national **STATISTICS**

LONDON

CabinetOffice



Visualisations and Mashups

AKTive PSI - Mozilla Firefox

http://sasoki.aktivepsi.com/scraper/ch/?companyname=Abington+Autos

AKTive PSI

AKTive PSI
Companies House Web Scraper

Company Name:
Abington Autos

Get Company Information

Suggested Searches:
Abington Autos
Abington Autocentre

Company Information

The closest company that matched your search was called 'ABINGTON AUTO'S LIMITED'.

ABINGTON AUTO'S LIMITED	MAYFIELDS INSOLVENCY PRACTITIONERS CHURCH STEPS WEST MIDLANDS B63 4AB
Company No. 03869315	
Status	Dissolved 29/12/2004
Date of Incorporation	01/11/1999
Country of Origin	United Kingdom
Company Type	Private Limited Company
Nature of Business (SIC(03))	5010 - Sale of motor vehicles
Accounting Reference Date	31/03
Last Accounts Made Up To	31/03/2001 (TOTAL EXEMPTION SMALL)
Next Accounts Due	31/01/2003
Last Return Made Up To	01/11/2001
Next Return Due	29/11/2002
Last Members List	01/11/2001
Insolvency History	

AKTive PSI on Google Maps - Mozilla Firefox

http://sasoki.ecs.soton.ac.uk/aktivepsi/!maps/

AKTive PSI

Land Registry Camden Sample Set

Show Markers

land and buildings on the south side of Platt Street at the Junction with Golding

Title Number: NGL185462
Owned by: CLARK KENT 1 Smith Street, London, E14 9NZ
Tenure: Absolute Freehold
Other Register Data:

Daisies Day Nursery

Risk Category
B

[Camden Food Premises](http://www.camden.gov.uk/food/PremiseOntology/#premises5083310)
<http://www.camden.gov.uk/food/PremiseOntology/#premises5083310>

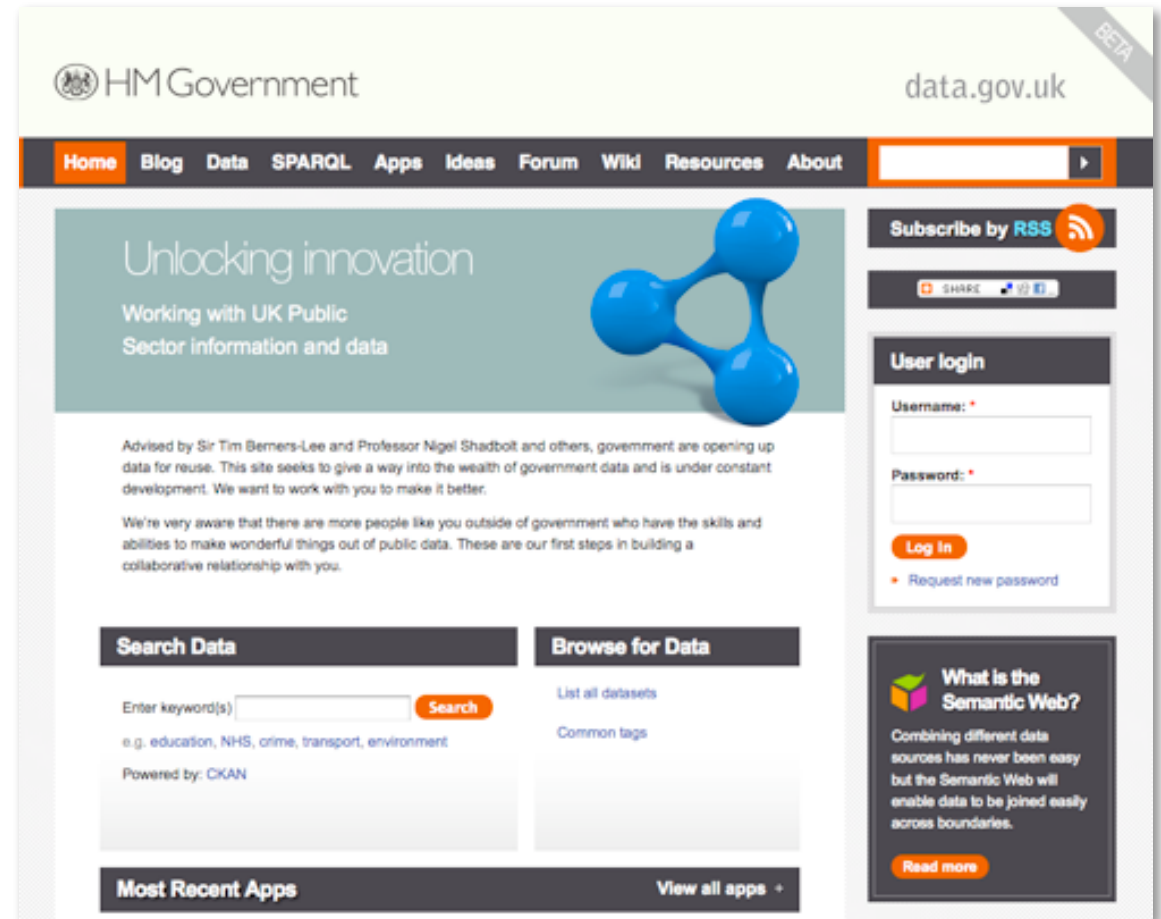
[PointX Points of Interest](http://www.ordnancesurvey.co.uk/PoiOntology/#business26453632)
<http://www.ordnancesurvey.co.uk/PoiOntology/#business26453632>

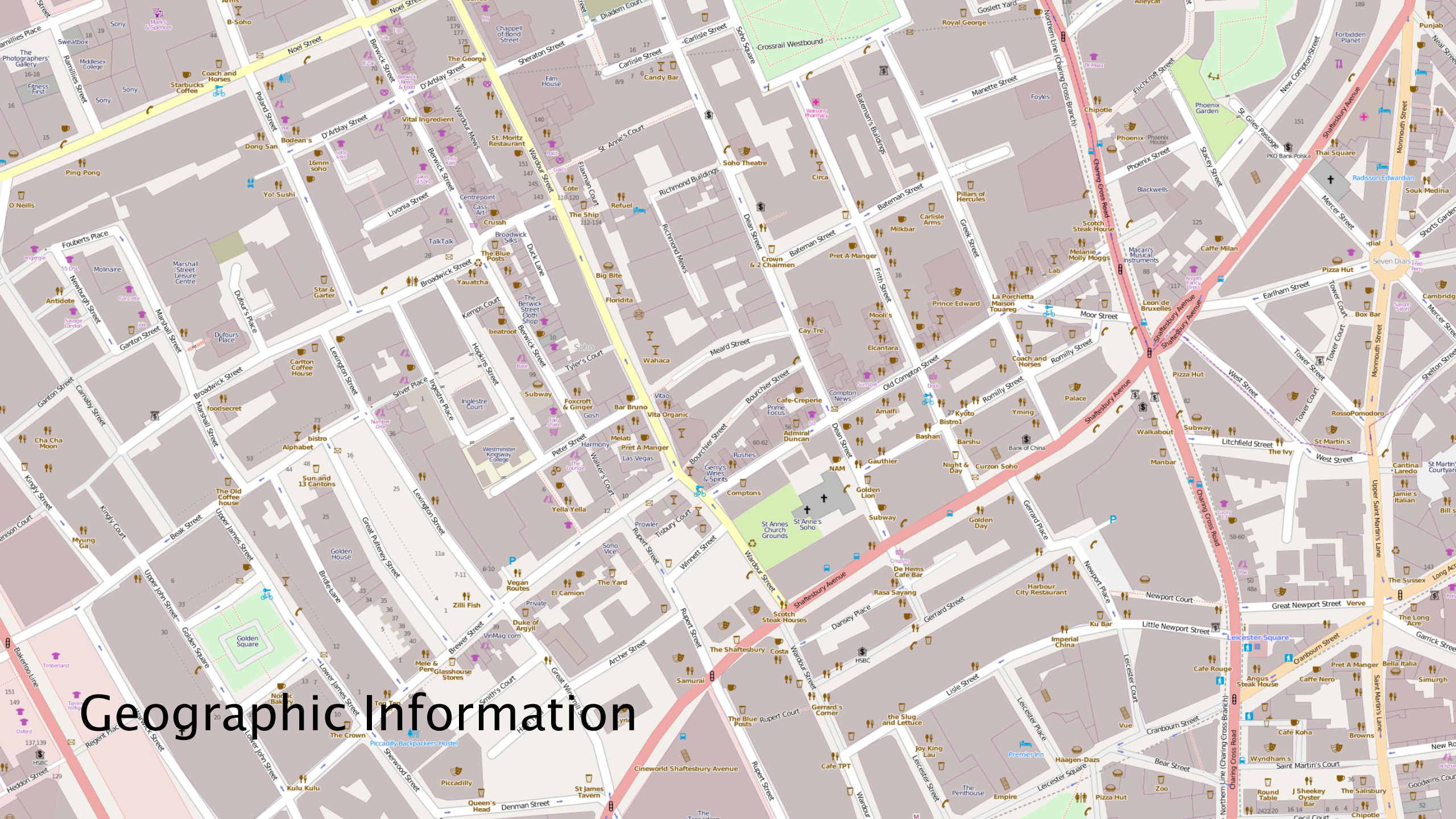
Integrated Public Sector Vocabulary

- Taxonomy for classifying public sector information resources
- Developed with the backing of the ODPM and the Cabinet Office e-Government Unit
- Available in a variety of formats, but RDF/XML is the definitive format
- Uses Dublin Core and SKOS

data.gov.uk (2010)

- In 2009, Tim Berners-Lee and Nigel Shadbolt appointed as advisors to HMG on opening up public data
- data.gov.uk website went public on 21 Jan 2010
 - 2890 datasets from central and local government
 - Uses SW technologies to represent data





Geographic Information

Ordnance Survey Linked Data

Opened up datasets as part of UK public sector data initiative:

- 1:50,000 Gazetteer
250,000 place names
- Code-Point
1,700,000 postcode units
- Boundary Line
Electoral and administrative boundaries

- Output formats:
 - JSON
 - RDF
- SPARQL query

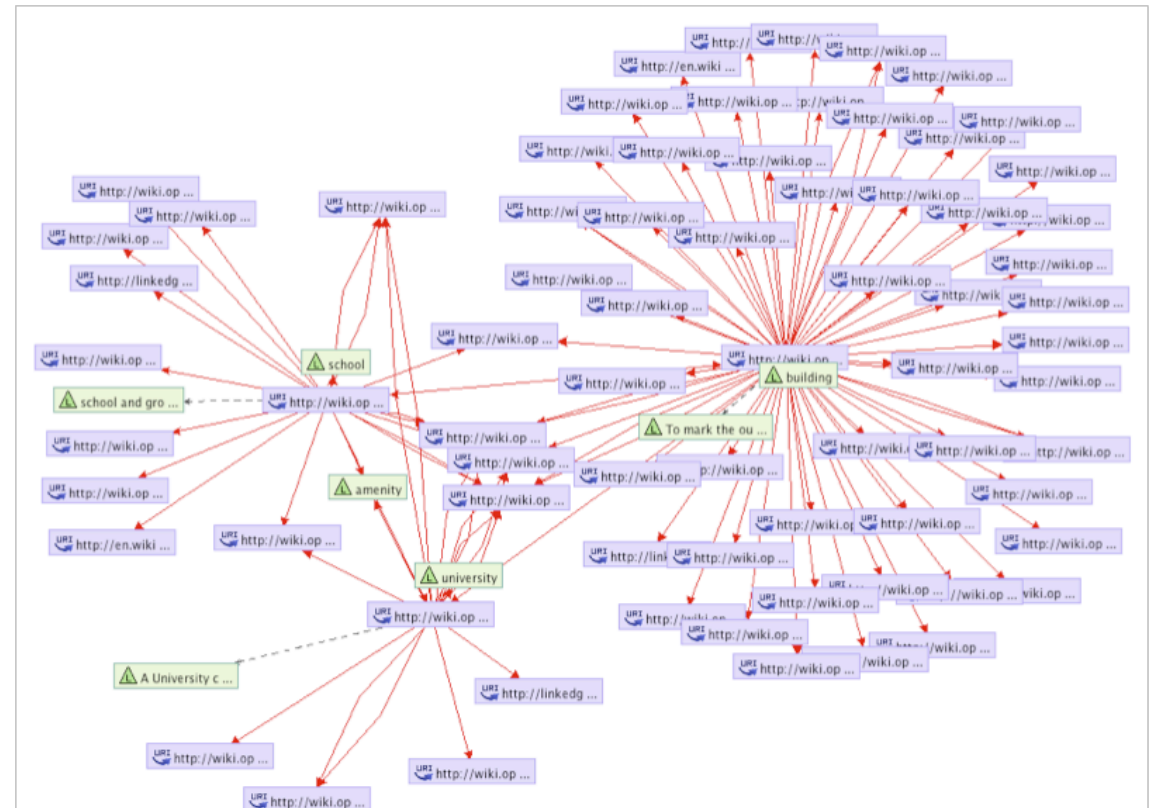
Open Streetmap Linked Data

OSM Semantic Network

- SKOS vocabulary built from map feature tags

LinkedGeoData.org

- Interlinks OSM data with other linked open data



Next Lecture: RDF