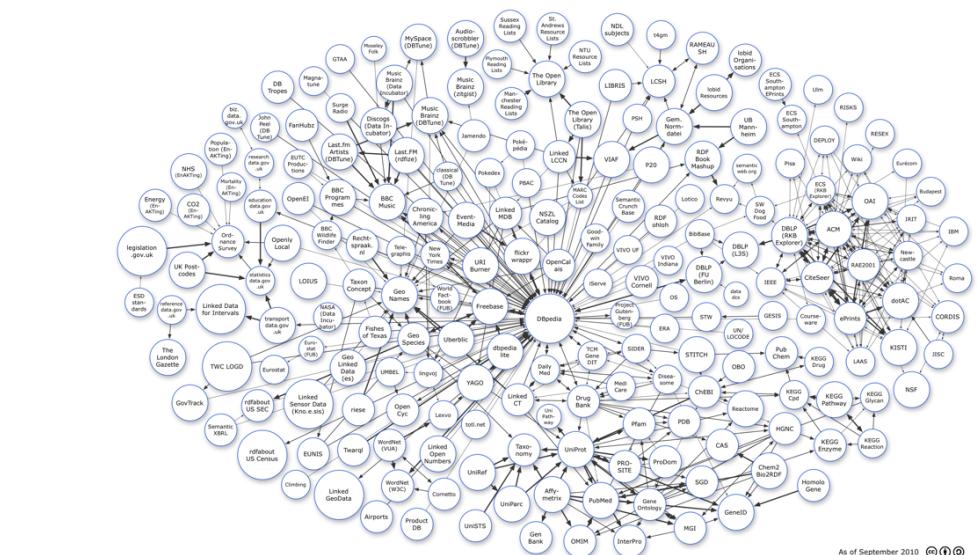




Applying Semantic Web technologies and Linked Data principles to facilitate the analysis of OU research data



Work funded by the Research School (Open University)
John Domingue, Mathieu d'Aquin and Vanessa Lopez

radar.open.ac.uk



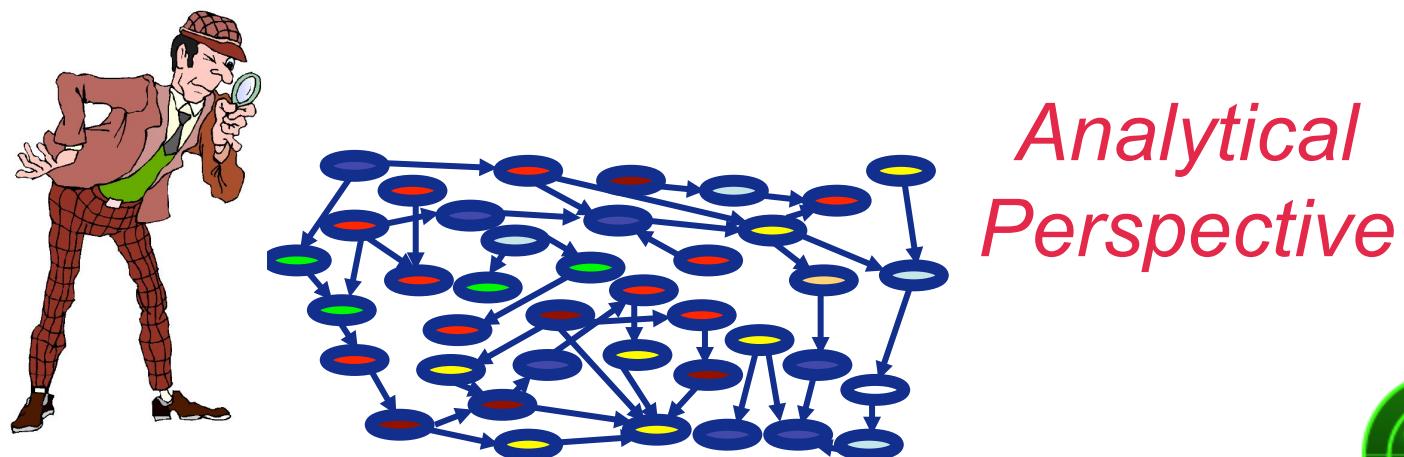
What is it?
What does it do?
What will it do?
Why should I care?





RADAR: what and how

- Support researchers in analysing the research outputs and work of OU research staff to support, for example, the OU submission to the **REF**
- Based on an underlying data model (ontology) of the academic domain
- Pilot tool: visualization of general statistics and data analytics for individual researchers and groups





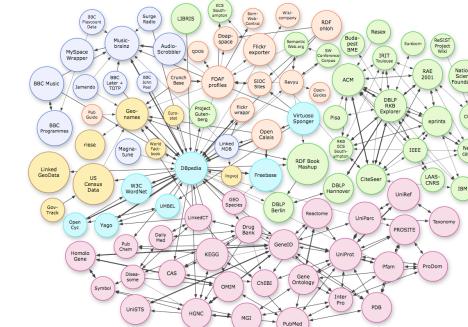
Yes.. but why?, what are the benefits of using Linked Data at the OU?



Enrico Motta
Professor of

Professor of Knowledge Technologies

```
<foaf:name>Enrico Motta</foaf:name>
  <foaf:firstName>Enrico</foaf:firstName>
  <foaf:surname>Motta</foaf:surname>
  <foaf:phone rdf:resource="tel:+44-(0)
1908-653506"/>
```



4





Yes.. but why?, what are the benefits of using Linked Data at the OU?

- Implements Linked Data principles:
 - unique resource identifier – URI
 - SPARQL end-point (OWLIM Sesame repository)
- Data enrichment at speed – not feasible with standard technologies and limited resources
- Data integration from different datasets
- Visualize and discover relationships and connections between entities (e.g., researchers) in large data

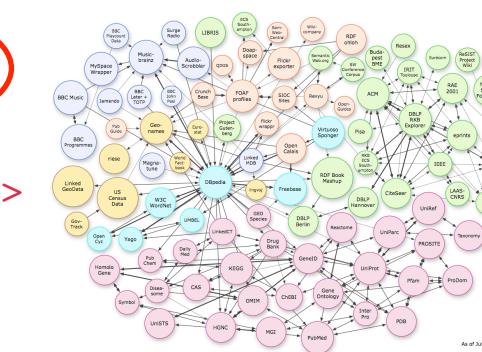


Enrico Motta

Professor of Knowledge Technologies



```
<foaf:name>Enrico Motta</foaf:name>
<foaf:firstName>Enrico</foaf:firstName>
<foaf:surname>Motta</foaf:surname>
<foaf:phone rdf:resource="tel:+44-(0)
1908-653506"/>
```





Dataset integration

- Internal
 - ORO (Open Research Online): OU repository of research publications, <http://oro.open.ac.uk>
 - Lucero (KMi data): <http://lucero-project.info>, <http://data.open.ac.uk>
 - Research audit data from different OU departments
- Re-using public schemas:
 - Foaf (people)
 - Bibo (bibliographic resources)
 - Small additions in <http://radar.open.ac.uk/ontology> namespace
- External
 - Structured semantic sources: DBLP, DBpedia, etc.
 - Unstructured sources: Google scholar (citations), news, impact reports.



Pilot tool (KMi)

- Staff data: projects and publications for KMi members
- Summary of classes, subclasses, properties and members
- Flexible approach: automatically classify individual members depending on their properties
- Graphs of distribution of property values: compare properties of entities belonging to the same group or across groups (e.g., professors, postdocs, etc.)



Demo



Person <http://xmlns.com/foaf/0.1/Person>

More specific classes ▾

Person (40 members)

- PostDoctoralResearcher
- ResearcherNonDoctor

PostDoctoralResearcher ResearcherNonDoctor

Properties ▾

- identifier [40 values]
- averageFundingFromProjects [11 values]
- hasUnitOfAssesment [2 values]
- isExternalSupervisor [2 values]
- numberOfProjects [7 values]
- numberOfPublications [24 values]
- numberOfStudentsSupervised [9 values]
- totalFundingFromProjects [11 values]
- label [75 values]
- hasMembership [40 values] [21 classes]

Members ▾

Members ▾

- Adrian Stuart Poulton (Person Agent Person SpatialThing Agent)
- Alec Goodyear (Person Agent Person SpatialThing Agent)
- Alexander Mikroyannidis (Person Agent Person SpatialThing Agent JuniorResea
- Alexandra Lilavati Pereira Okada (Person Agent Person SpatialThing Agent J
- Alison Bettley (Person Agent Person SpatialThing Agent)
- Alistair Guillaume Willis (Person Agent Person SpatialThing Agent)
- Allan Clive Jones (Person Agent Person SpatialThing Agent)
- Allan Third (Person Agent Person SpatialThing Agent)
- Alvaro Eduardo Faría (Person Agent Person SpatialThing Agent)
- Amir Shirzadi (Person Agent Person SpatialThing Agent)

< >

<http://radar.open.ac.uk/LDI/classview.php>





Individual researcher profile

Member ▲

Enrico Motta ▲

<http://data.open.ac.uk/person/c9aa7f2e582d191ed728ad414c5ea711>

Person Agent Person Spatial Thing Agent SeniorResearcher PostDoctoralResearcher Professor

- Number of publications
 - per year
 - by type: conference / journal / workshops
- Number of PhD students
- Total funding
 - from projects
 - projects per year
- Frequent co-authors and co-authored publications





Demo

Member Info ▲

Enrico Motta ◀ **Enrico Motta's Group Info** ▼

<http://data.open.ac.uk/person/c9aa7f2e582d191ed728ad414c5ea711>

Person Agent Person Spatial Thing Agent SeniorResearcher PostDoctoralResearcher Professor

• Number of publications [2008-11]: 47
• Total funding from projects: 4641992
• Number of PhD students: 16

2002: [1 publications]
2003: [6 publications]
2004: [13 publications]
2005: [3 publications]
2006: [5 publications]
2007: [4 publications]
2008: [17 publications]
2009: [20 publications]
2010: [11 publications]
2011: [1 publications]
2012: [3 projects]

2000: [1 projects]
2001: [1 projects]
2002: [1 projects]
2003: [1 projects]
2004: [3 projects]
2005: [3 projects]
2006: [8 projects]
2007: [8 projects]
2008: [5 projects]
2009: [7 projects]
2010: [7 projects]
2011: [3 projects]
2012: [3 projects]

Funding per Project (total= 4611278)

Project	Funding
OPEN KNOWLEDGE	391474.0
EPSRC DTA -UNIV.OF SOUTHAMPTON	48423.0
RESEARCH COLLABORATION IN AKT	1321728.0
NEON	1035507.0
SMARTPRODUCTS	470939.0
NEON (MANAGEMENT)	289197.0
STUDENTSHIP	30714.0
XMEDIA	537854.0
KNOWLEDGE WEB	259920.0
NEON (COMMON FUND)	225522.0

Number of Publications and Projects per Year

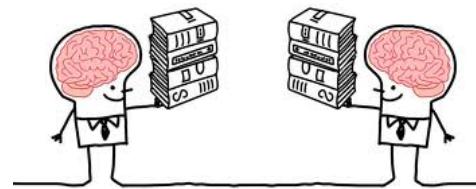
Year	Projects	Publications
2000	1	25
2001	1	0
2002	1	1
2003	2	6
2004	4	13
2005	3	4
2006	8	5
2007	8	4
2008	5	17
2009	7	20
2010	7	11
2011	3	1
2012	3	0





Research Group profile - groups

- Dynamic creation of groups based on relationships
 - Co-authors
 - Unit of assessment: computing, education, etc.
 - Topics (tag clouds): “show me everybody interested in leadership”, etc.
 - Career stage: Professors, Research Associates, PhD.
 - Manual: impact of adding or removing somebody
- Selecting and Filtering by year (e.g., 2008-2011)





Research Group Profile - charts

Group Info ▲

[Enrico Motta's Group Info](#) ◀ [Enrico Motta](#) ▼

<http://data.open.ac.uk/person/c9aa7f2e582d191ed728ad414c5ea711>

Enrico Motta Marta Sabou Mathieu d'Aquin John Bruno Gerard Domingue Victoria Uren Vanessa Lopez Andriy Nikolov Sofia Angeletou Miriam Fernandez Martin Dzbor Yuangui Lei Maria Vargas-Vera Liliana Cabral Fouad Zablith Stefan Markus Rueger Ainhoa Llorente Ning Li A. De Roeck Dnyanesh Rajpathak Carlo Allocca Zdenek Zdrahal Stefanie Nowak Silvio Peroni Miklos Nagy J. Zhu Chwhynny Overbeeke

- Number of publications, projects and PhD students of each researcher
- Total funding per researcher
- Number of publications:
 - per year
 - per type: journal, conference and workshops
- Most popular venues for conference publications





Demo

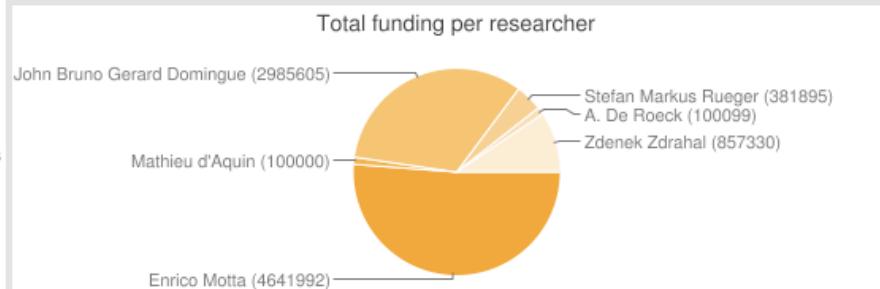
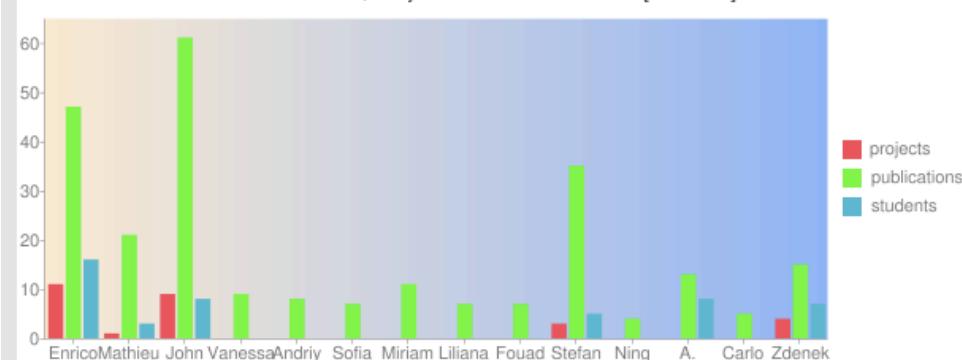
Group Info ▲

Enrico Motta's Group Info ◀ Enrico Motta ▼

<http://data.open.ac.uk/person/c9aa7f2e582d191ed728ad414c5ea711>

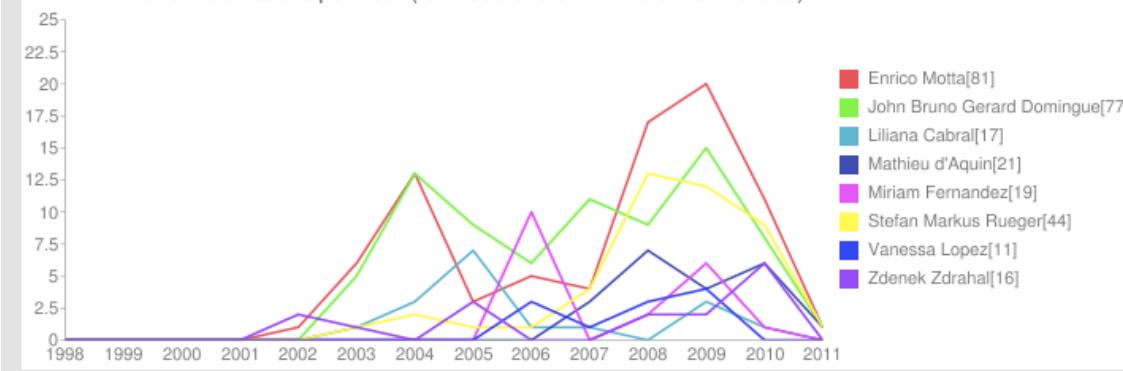
Enrico Motta Marta Sabou Mathieu d'Aquin John Bruno Gerard Domingue Victoria Uren Vanessa Lopez Andriy Nikolov Sofia Angeletou Miriam Fernandez Martin Dzbor Yuangui Lei Maria Vargas-Vera Liliana Cabral Fouad Zablith Stefan Markus Rueger Ainhoa Llorente Ning Li A. De Roeck Dnyanesh Rajpathak Carlo Allocca Zdenek Zdrahal Stefanie Nowak Silvio Peroni Miklos Nagy J. Zhu Chwhynny Overbeeke

Number of Publications, Projects and PhD students [2008-11]



Publications

Number of Publications per Year (for researchers with more than 10 total)



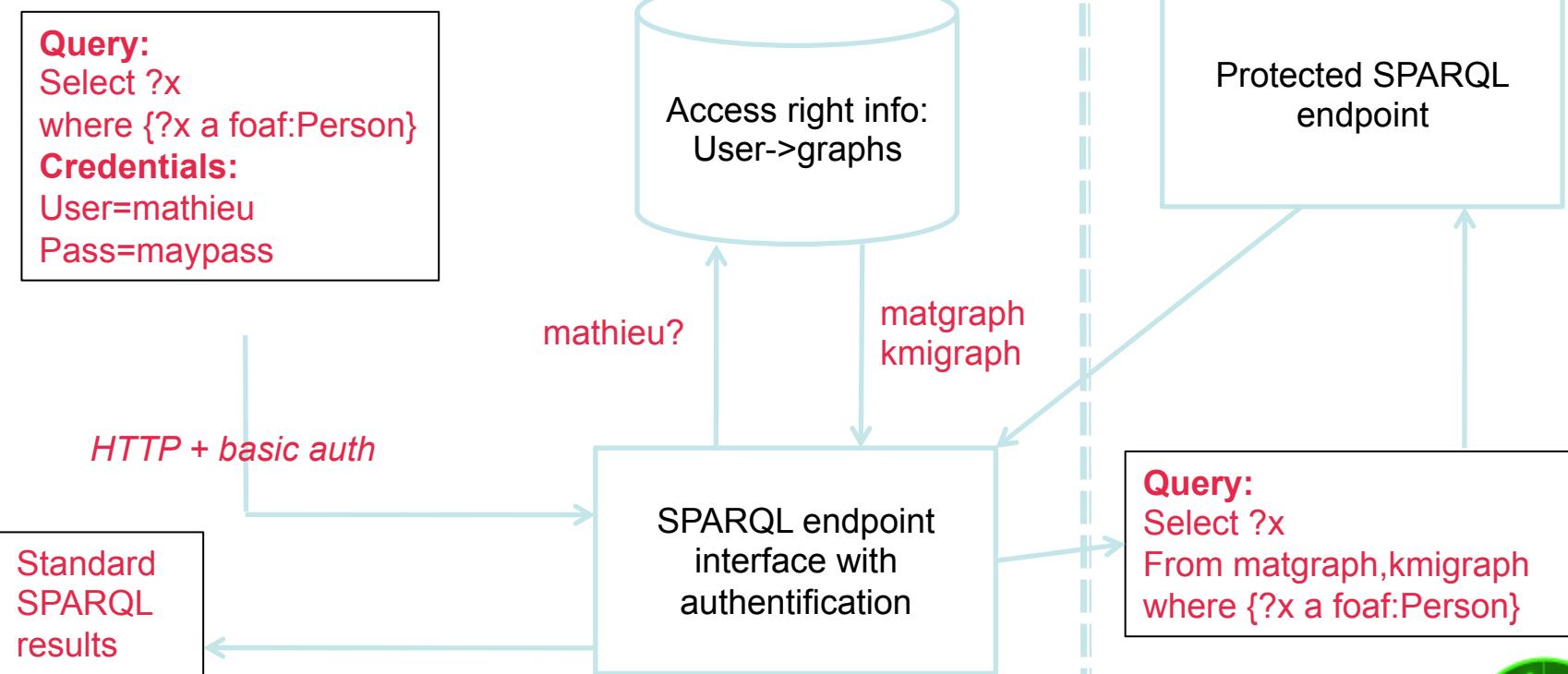
- A. De Roeck[5 publications: 0 journals, 3 conferences, 2 workshops]
- Andriy Nikolov[9 publications: 0 journals, 6 conferences, 3 workshops]
- Carlo Allocca[3 publications: 1 journals, 2 conferences, 0 workshops]
- Enrico Motta[81 publications: 4 journals, 52 conferences, 25 workshops]
- Fouad Zablith[9 publications: 2 journals, 4 conferences, 3 workshops]
- John Bruno Gerard Domingue[77 publications: 2 journals, 41 conferences, 34 workshops]
- Liliana Cabral[17 publications: 0 journals, 7 conferences, 10 workshops]
- Mathieu d'Aquin[21 publications: 2 journals, 11 conferences, 8 workshops]





Access control

- Fine-grained mechanism for access control based on dynamic creation of graphs
- Grant access to specific parts of data and specific groups of individuals





Directions – tagging and annotations

- Manual annotation helps to identify the researchers that will be submitted for the evaluation exercise:
 - Referenceable / Potential referenceable / Not likely
- Allow users to define their own group (e.g., by adding members)
- Rank quality of the research output: publications (e.g., 4 start publication) and venues
- Improve the visualization and quality of the data



Directions: enrich data with external sources

- Number of citations for publications
- Impact factor:
 - ISI Web of Knowledge for journal citation report
 - Core conference ranking (computing): A*, A, B or C

<http://admin-apps.webofknowledge.com/JCR/JCR>

Journal Title	ISSN	Total Cites	Impact Factor	5-Year Impact Factor
INT J INTELL SYST	0884-8173	1085	1.314	1.249

43623	International Joint Conference on Artificial Intelligence	IJCAI	A	0801	Artificial Intelligence and Image Proce
43624	International Joint Conference on Automated Reasoning	IJCAR	A	0801	Artificial Intelligence and Image Proce
43625	International Joint Conference on e-Business and Telecommunications	ICETE	C	0806	Information Systems
43626	International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management	IC3K	C	08	Information and Computing Sciences
43627	International Joint Conference on Natural Language Processing	IJCNLP	B	0801	Artificial Intelligence and Image Proce
43628	International Joint Conference on Qualitative and Quantitative Practical Reasoning	ESQARU	A	0801	Artificial Intelligence and Image Proce
43630	International KES Symposium on Agents and Multiagent systems - Technologies and Applications	KES AMSTA	C	0801	Artificial Intelligence and Image Proce
43632	International Machine Vision and Image Processing Conference	IMVIP	C	0801	Artificial Intelligence and Image Proce
45018	International Meeting on DNA Computing and Molecular Programming	DNA	B	0801	Artificial Intelligence and Image Proce
43635	International Natural Language Generation Conference	INLG	B	0801	Artificial Intelligence and Image Proce
43637	International Network Optimization Conference	INOC	C	0802	Computation Theory and Mathematic
43638	International Parallel Computing Workshop	PCW	C	0805	Distributed Computing
43643	International Python Conference	IPythC	C	0803	Computer Software
43646	International Semantic Web Conference	ISWC	B	0804	Data Format





Other aspects being explored

- Flexible methods to add data:
 - To allow research managers or individual researchers to link researchers/ projects to Web documents
 - Bookmarklet Web documents as evidence of impact
- More data from other departments will raise unexpected problems



Conclusions

- Lucero the public face of RADAR for Linked Open Institutional data
- Using Linked Data in research and education scenarios brings new possibilities over traditional data management models in terms of data integration and enrichment
- Reusing existent visualizations to analyze semantic federated data and discover connections: graphs, charts, tag clouds, search, QA, multimodal interaction, etc.
- Explore new ways to deal with data: Discover links between previously isolated data and external sources

