

Linked Data

Making things more accessible

Hugh Glaser

Data-Intensive Research: how should we improve our ability to use data?
e-Science Institute, Edinburgh
17th March 2010

Seme4
18 Soho Square
London W1 3QL
T : + 44 20 7060 1590

© 2010 Seme4



Seme4.com

Overview

- **Technology -> Use**
- **Linked Data tutorial**
- **Doing the linking**
- **Using the linking**

Motivation

- **Publishing data becomes an imperative**
 - Freedom of Information
 - Public confidence
 - Reputation
- **Linking Data becomes an imperative**
 - Research benefits from the use of multiple datasets
 - Research benefits from datasets from different researchers
 - Connecting publications with the subject data
- **Data Fusion**
 - Breaking silos
 - Socio-political problem
 - data.gov.uk

How to Join?

- **Hypertext Style/ Cool URIs don't change.**
 - <http://www.w3.org/Provider/Style/URI>
- **Cool URIs for the Semantic Web**
 - <http://www.w3.org/TR/cooluris/>
- **Linked Data - Design Issues**
 - <http://www.w3.org/DesignIssues/LinkedData.html>
- **How to publish Linked Data on the Web**
 - <http://www4.wiwi.fu-berlin.de/bizer/pub/LinkedDataTutorial/>
- **Linked Data**
 - <http://linkeddata.org/>

1. Use URIs as names for things

- **Duh!**
- **Everything**
 - If you don't name something you can't talk about it
 - Things of course
 - Workflows
 - Ideas
 - ...
- **Cool URIs**
 - Think of the consumer/customer
 - https://secure.ecs.soton.ac.uk/gizmos/person_by_username.php?username=hg
 - <https://secure.ecs.soton.ac.uk/person/username/hg>
 - Library of Congress - *Chronicling America*: Historic American Newspapers
 - <http://chroniclingamerica.loc.gov/>
 - Cool URIs – 71233 page views per week to 447992 - factor of 6.3
- **Ambiguity**
 - URIs help to avoid it, especially if you...

2. Use HTTP URIs so that people can look up those names

- HTTP URI names come with an excellent look up mechanism
- And ownership, etc.
- Make sure they are Cool

3. When someone looks up a URI, provide useful information, using the standards (RDF, SPARQL)

- **Otherwise I don't know what you mean**
- **Deliver some human readable data**
 - html
- **Deliver some machine processable data**
 - RDF
 - JSON
 - CSV
 - text
- **Query is good**
 - SPARQL

RDF - Triples

- “Ian Millard works for the University of Southampton”

person-04860	name	Ian Millard
person-04860	works for	Southampton University
Southampton University	label	The University of Southampton

RDF - Triples

- “Ian Millard works for the University of Southampton”

soton-rkb:person-04860 foaf:name “Ian Millard” .

soton-rkb:person-04860 akt:works_for dbpedia:Southampton_University .

dbpedia:Southampton_University rdfs:label “The University of Southampton” .

- This is actually RDF!
 - Given the namespaces
- I could put this in an RDF store now

SPARQL

PREFIX soton-rkb: <http://southampton.rkbexplorer.com/id/>

PREFIX akt: <http://www.aktors.org/ontology/portal#>

PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

SELECT * WHERE

{ soton-rkb:person-04860 akt:works_for ?org . ?org rdfs:label ?name . }

org	name
http://southampton.rkbexplorer.com/id/UoS	University of Southampton
http://southampton.rkbexplorer.com/id/org-ecs	School of Electronics and Computer Science, University of Southampton

4. Include links to other URIs. so that they can discover more things

- **“Foreign” URIs**
 - `dbpedia:Southampton_University`
- **Equivalence**
 - `owl:sameAs`
 - `skos:exactMatch`
 - ...

Where to keep the equivalence?

- **Equivalence is different to the substantive data**
 - Accuracy
 - Provenance
 - License
 - Trust
 - Maintenance
 - Process
- **Separate knowledge base**
- **Too many identifiers?**
 - The answer is not to create another one
- **A little middleware**
 - sameAs aware

Discovery v. Description

- **Equivalence**
 - <http://sameas.org/html?uri=http://dbpedia.org/resource/London>
 - <http://sameas.org/html?uri=http://webscience.org/person/23>
 - <http://sameas.org/html?uri=http://dbpedia.org/resource/14-Methoxymetopon>
 - <http://sameas.org/html?uri=http://purl.uniprot.org/core/Mitochondrion>
- <http://sameas.org/> is *discovery*
- For *description* you need your own
- Many advantages to distributed architecture
 - Scalable
 - Multiple
 - Rollback
 - Deprecation
 - Canonisation (URI reduction)

<sameAs>

interlinking the Web of Data

The Web of Data has many equivalent URIs.
This service helps you to find co-references
between different data sets.
Enter a known URI, or use Sindice to search first.

<sameAs>



Why not try searching for the string "[London](#)", or finding other equivalent identifiers for <http://dpbedia.org/resource/London> ?

Currently serving **30918791** URIs in **10796012** bundles!

[about](#) · [contact](#)



The Web of Data has many equivalent URIs.
This service helps you to find co-references
between different data sets.
Enter a known URI, or use Sindice to search first.

<sameAs>



Equivalent URIs for http://dbpedia.org/resource/London –

- <sameAs> {
1. <http://dbpedia.org/resource/London>
 2. <http://dbpedia.org/resource/Leondeon>
 3. <http://dbpedia.org/resource/Lundúnir>
 4. [http://dbpedia.org/resource/London_\(UK\)](http://dbpedia.org/resource/London_(UK))
 5. [http://dbpedia.org/resource/London_\(Britain\)](http://dbpedia.org/resource/London_(Britain))
 6. [http://dbpedia.org/resource/London_\(England\)](http://dbpedia.org/resource/London_(England))
 7. [http://dbpedia.org/resource/London_\(Great_Britain\)](http://dbpedia.org/resource/London_(Great_Britain))
 8. [http://dbpedia.org/resource/London_\(United_Kingdom\)](http://dbpedia.org/resource/London_(United_Kingdom))
 9. <http://www.okkam.org/ens/id1d7d3fb0-44f8-41a1-8866-89c1458a1c66>
 10. <http://www.okkam.org/entity/oke0ab0619-e28a-438d-b53b-cdd5c8157f30>
 11. <http://os.rkbexplorer.com/id/osr7000000000041428>
 12. <http://data.ordnancesurvey.co.uk/id/7000000000041428>
 13. <http://linkededgeodata.org/triplify/node/107775#id>
 14. <http://mpii.de/yago/resource/London>
 15. <http://rdf.freebase.com/ns/guid.9202a8c04000641f80000000000242b2>
 16. <http://statistics.data.gov.uk/id/eer/07>
 17. <http://statistics.data.gov.uk/id/eer?name=London>
 18. <http://sw.opencyc.org/concept/Mx4rvVjWPJwpEbGdrcN5Y29ycA>
 19. <http://sw.opencyc.org/2008/06/10/concept/en/CityOfLondonEngland>
 20. <http://sw.opencyc.org/2009/04/07/concept/en/CityOfLondonEngland>
 21. <http://sw.opencyc.org/2008/06/10/concept/Mx4rvVjWPJwpEbGdrcN5Y29ycA>
 22. <http://sw.opencyc.org/2009/04/07/concept/Mx4rvVjWPJwpEbGdrcN5Y29ycA>
 23. <http://sws.geonames.org/2643743/>
 24. <http://telegraphis.net/data/capitals/GB/London#London>
 25. <http://umbel.org/umbel/ne/wikipedia/London>
 26. <http://www.ontologyportal.org/SUMO.owl#LondonUnitedKingdom>

<sameAs>

interlinking the Web of Data

The Web of Data has many equivalent URIs.
This service helps you to find co-references
between different data sets.
Enter a known URI, or use Sindice to search first.

<sameAs>



Equivalent URIs for <http://webscience.org/person/23> –

- <sameAs> {
1. <http://southampton.rkbexplorer.com/id/person-1ec5be2368d15446d69e61...d7ff327d1e>
 2. <http://southampton.rkbexplorer.com/id/person-2a8f4c2138ec4ce326c712...fd42c1a2fb>
 3. <http://southampton.rkbexplorer.com/id/person-2a8f4c2138ec4ce326c712...34d38c7878>
- ...
Show 699 more
703. <http://webscience.org/person/23>
- rdf+xml · n3 · json · text

Currently serving **30918791** URIs in **10796012** bundles!

[about](#) · [contact](#)



<sameAs>

interlinking the Web of Data

The Web of Data has many equivalent URIs.
This service helps you to find co-references
between different data sets.
Enter a known URI, or use Sindice to search first.

<sameAs> 🔍

 🔍

Equivalent URIs for <http://dbpedia.org/resource/14-Methoxymetopon> –

<sameAs> {
1. <http://dbpedia.org/resource/14-Methoxymetopon>
2. <http://bio2rdf.org/dbpedia:14-Methoxymetopon>
3. <http://rdf.freebase.com/ns/guid.9202a8c04000641f800000000812d3e8>
rdf+xml · n3 · json · text

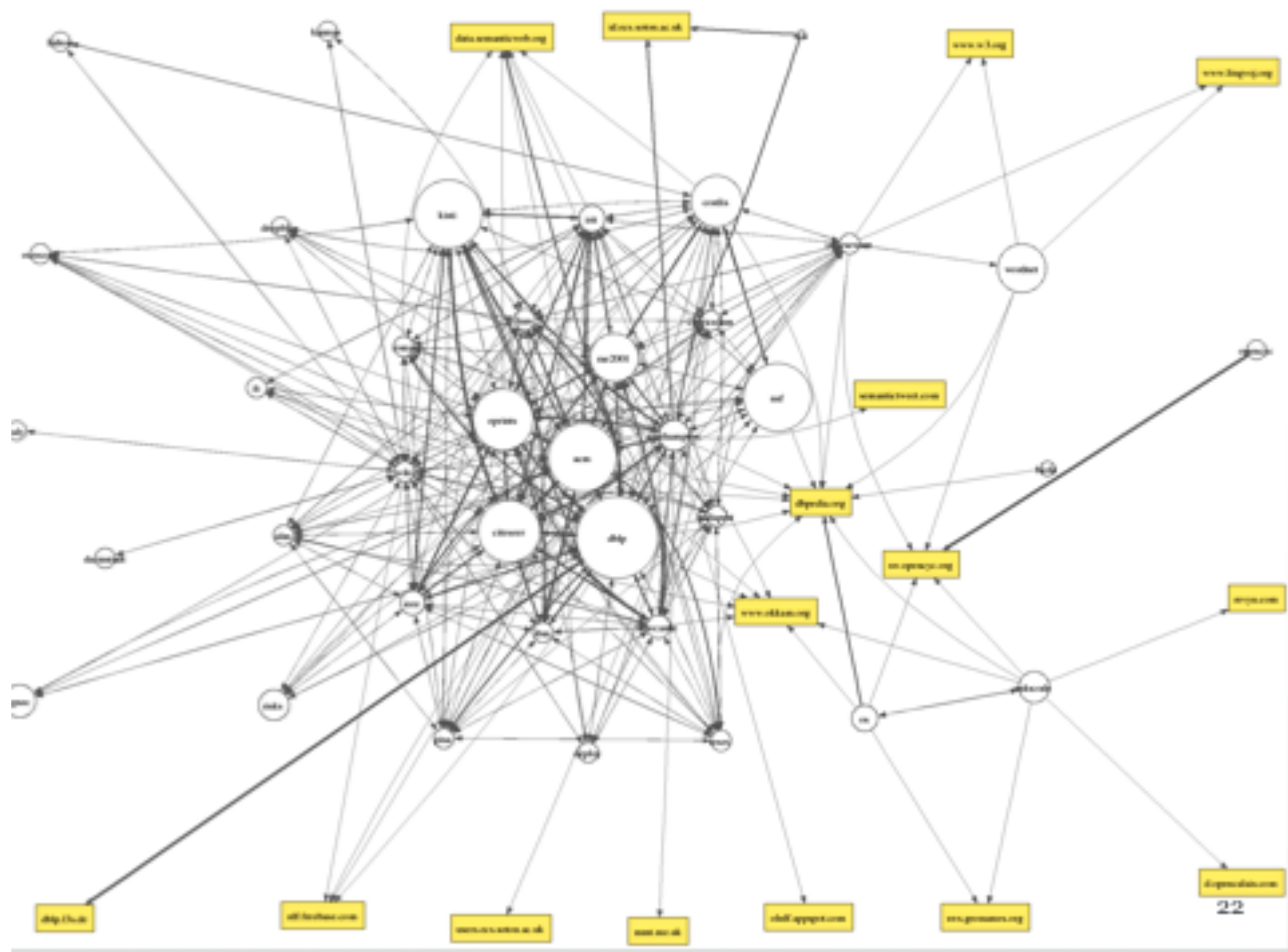
Currently serving **30918791** URIs in **10796012** bundles!

[about](#) · [contact](#)



Data Fusion

- **Break the silos**
- **At low risk**
- **And low cost**
- **With little or no process/workflow change**



Using Linked Data and sameAs: RKBExplorer.com

- **Unified view of disparate resources**
- **Open Web**
- **Distributed query**
- **Distributed sameAs**
- **Services delivering RDF, JSON, csv, ...**
 - <http://www.rkbexplorer.com/service/>
- **Different interfaces possible**

RKBExplorer » People » David De Roure

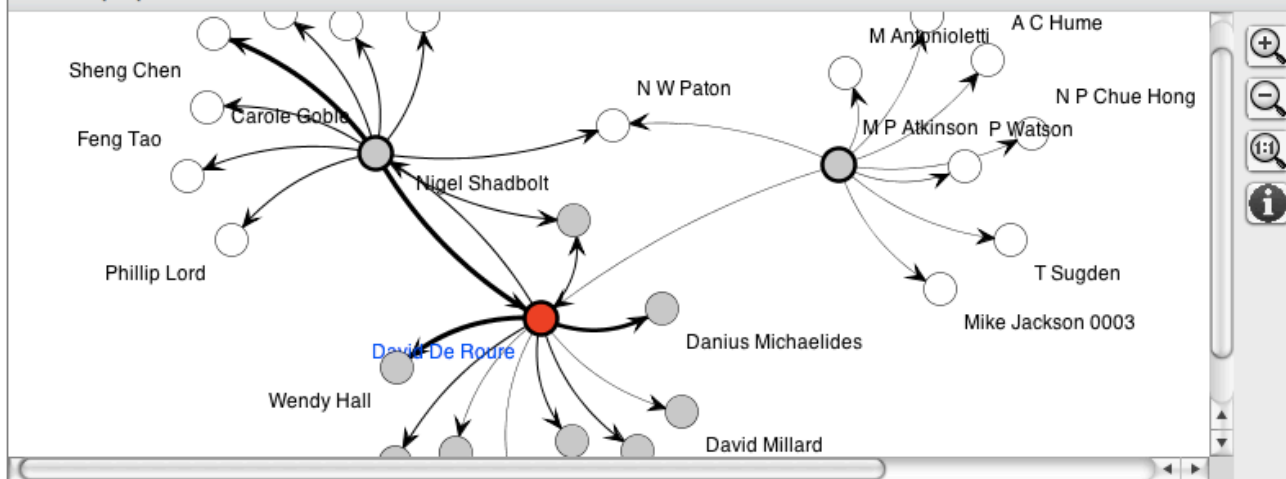
Malcolm Atkinson

Search

→ about → system requirements
→ help → data sources
→ contact → acknowledgements



Related people



Details

Full Name:
[David De Roure](#)

Works For:
[University of Southampton](#)

Telephone:
+44 (0)23 8059 2418

Fax:
+442380592865

Email Address:
dder@ecs.soton.ac.uk

Web Address:
<http://www.ecs.soton.ac.uk/people/c>

People

[Wendy Hall](#)[Danus Michaelides](#)[Les Carr](#)[Mark Weal](#)[Carole Goble](#)[Don Cruickshank](#)[Nigel Shadbolt](#)[GJ Hill](#)

Organisations

[University of Southampton](#)[Grid and Pervasive Computing Group](#)[Intelligence, Agents and Multimedia](#)[Intelligence, Agents, Multimedia Group](#)

Publications

[A Dynamic Size Distributed Program Image Cache for Wireless Sensor Networks](#)[A networked registration scheme to support open science](#)[Benchmarking Workflow Discovery: A Case Study From Bioinformatics](#)[Building a Distributed Infrastructure for Scalable Triple Stores](#)[Designing authoring tools for the](#)

Projects

[Meeting Memory Technology Informing Collaboration](#)[MyExperiment](#)[myExperiment Repository Enhancement](#)

RKBExplorer » People » David De Roure

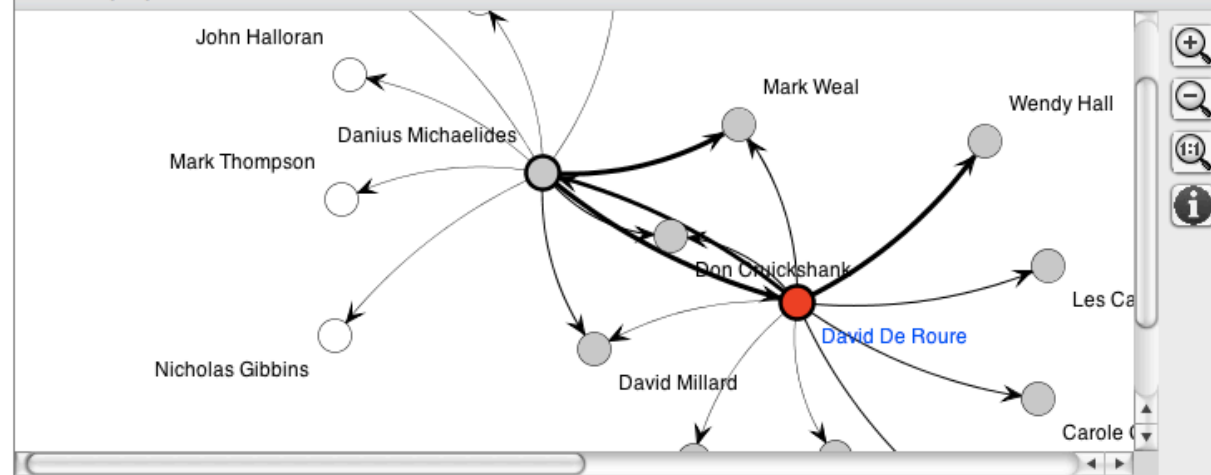
Search...

Search

- about
- help
- contact
- system requirements
- data sources
- acknowledgements



Related people



Details

Full Name:
[David De Roure](#)

Works For:
[University of Southampton](#)

Telephone:
+44 (0)23 8059 2418

Fax:
+442380592865

Email Address:
dder@ecs.soton.ac.uk

Web Address:
<http://www.ecs.soton.ac.uk/people/dder/>

People

[Wendy Hall](#) ?

[Danus Michaelides](#) ?

[Les Carr](#) ?

[Mark Weal](#) ?

[Carole Goble](#) ?

[Don Cruickshank](#) ?

[Nigel Shadbolt](#) ?

[GJ Hill](#) ?

Organisations

[University of Southampton](#) ?

[Grid and Pervasive Computing Group](#) ?

[Intelligence, Agents and Multimedia](#) ?

[Intelligence, Agents, Multimedia Group](#) ?

Publications

[A Dynamic Size Distributed Program Image Cache for Wireless Sensor Networks](#) ?

[A networked registration scheme to support open science](#) ?

[Benchmarking Workflow Discovery: A Case Study From Bioinformatics](#) ?

[Building a Distributed Infrastructure for Scalable Triple Stores](#) ?

[Designing authoring tools for](#) ?

Research Areas

[H.5.4. Hypertext/Hypermedia](#) ?

[Computer Science](#) ?

[H.5.1. Multimedia Information Systems](#) ?

[Unknown/Withheld](#) ?

[Unknown/Withheld](#) ?

[Unknown/Withheld](#) ?

[Unknown/Withheld](#) ?

RKBExplorer » People » Chris Rusbridge

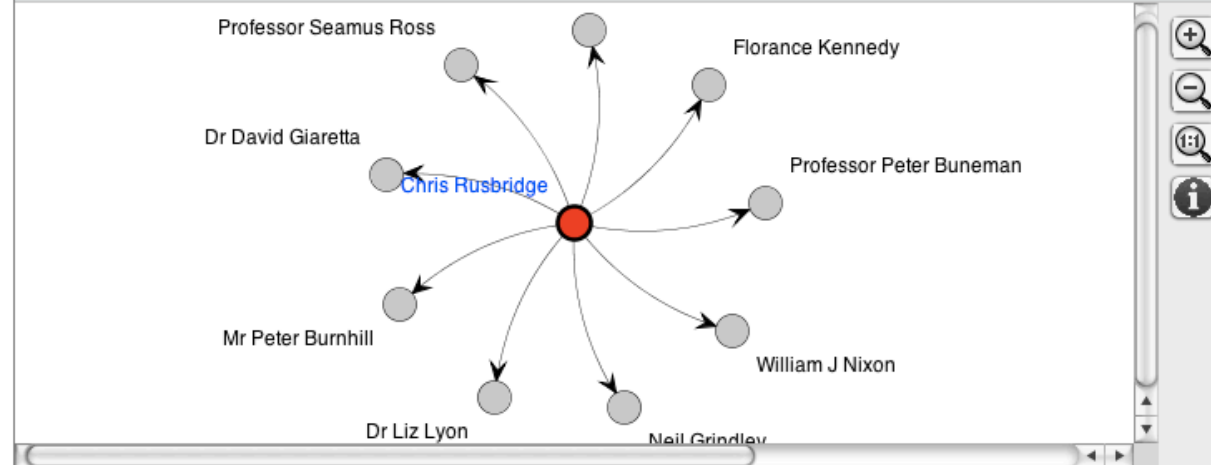
Search...

Search

- about
- help
- contact
- system requirements
- data sources
- acknowledgements



Related people



Details

Full Name:

[Chris Rusbridge](#)

Full Name:

[Chris Rusbridge](#)

Affiliated to:

Digital Curation Centre,
University of Edinburgh

Email Address:

munnari!max.sait.oz.au!cccar@u

Email Address:

C.Rusbridge@ed.ac.uk

How Linked Data:

People

[Neil Grindley](#) ?

[Dr David Giaretta](#) ?

[Dr Liz Lyon](#) ?

[Florance Kennedy](#) ?

[Mr Peter Burnhill](#) ?

[Professor Malcolm Atkinson](#) ?

[Professor Peter Buneman](#) ?

Organisations

[Digital Curation Centre, University of Edinburgh](#) ?

Publications

[Human automata and inhuman automata](#) ?

[Australian Tertiary Education Networking](#) ?

[Digital Curation Centre - Final Progress Report \(Phase 1\)](#) ?

[JCEI requirements for JCAS](#) ?

[Managing local institutional assets: OAI on Campus](#) ?

[Parallel searching and the](#) ?

Projects

[Digital Curation Centre](#) ?

[NSF economic sustainability taskforce](#) ?

RKBExplorer » People » Brian Randell

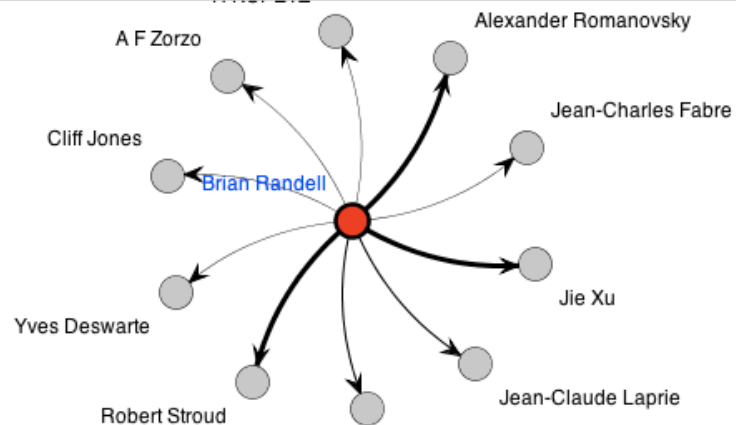
Search...

Search

- about
- help
- contact
- system requirements
- data sources
- acknowledgements



Related people



Details

0191 222 1923

Fax:

0191 222 8232

Email Address:

brian.randell@ncl.ac.uk

Web Address:

<http://www.cs.ncl.ac.uk/people/hid=4>

Description:

Brian Randell is a British computer scientist, specializing in research in software fault tolerance and dependability.

People

[Alexander Romanovsky](#)

[Jie Xu](#)

[Robert Stroud](#)

[Jean-Claude Laprie](#)

[David Powell](#)

[Cliff Jones](#)

[Yves Deswarte](#)

[A F Zorzo](#)

Organisations

[The University of Newcastle upon Tyne](#)

[Dependability](#)

[NEWCASTLE UNIVERSITY](#)

Publications

[The ReSIST Resilience Knowledge Base](#)

[Voting Technologies and Trust](#)

[Basic Concepts and Taxonomy of Dependable and Secure Computing](#)

[Dependable Pervasive Systems](#)

[Extending the Horizons of DSE \(GC6\)](#)

Projects

[CABERNET](#)

[DEVA](#)

[MAFTIA](#)

[PDCS 2](#)

[ReSIST Resilience for Survivability in IST](#)



www.rkbexplorer.com

Brian Randell is connected to [Andrea Bondavalli](#)

They are linked by 4 relations.

Publications

They have co-authored 3 papers:

- [Dependability Analysis of Iterative Fault Tolerant Software Considering Correlation](#)
- [Rational Design of Multiple-Redundant Systems: Adjudication and Fault Treatment](#)
- [Dynamic Adjustment of Dependability and Efficiency in Fault-Tolerant Software](#)

Projects

They are both members of the [ReSIST Resilience for Survivability in IST](#) project.



MyExperiment

Search...



Location of Project Members



Details

Names:

MyExperiment

Starts:

01/03/2007

Ends:

31/03/2009

Web Address:

<http://www.myexperiment.org/>

Description:

The myExperiment Virtual Research Environment enables you and your research

Project Members

[Carole Goble](#) ?

[Darius Michaelides](#) ?

[David De Roure](#) ?

[Frederique VanTill](#) ?

Participating Organisations

[University of Southampton](#) ?

Related JISC Projects

[Meeting Memory Technology Informing Collaboration](#) ?

[myExperiment Repository Enhancement](#) ?

[Collaborative Orthopaedic Research Environment](#) ?

[Internet-Based Behavioural Research](#) ?

[Virtual Research Integration Collaboration](#) ?

[Building Research and Innovation Networks](#) ?

Publications



RKBExplorer » Networks (Communities of Practice)

Carl Lagoze

Found 1 result for "Carl Lagoze"

[Carl Lagoze](#)

Carl Lagoze

[Display Metadata](#)

[James Allan](#) [Massoud Amin](#) [Ross Anderson](#) [William Arms](#) [Jeroen Bekaert](#) [Donna Bergmark](#) [Les Carr](#) [Tim Cornwell](#) [Lorrie Cranor](#) [James R. Davis](#) [Pavel Dmitriev](#) [Naomi Dushay](#) [Doug Dyer](#) [Dea Eckstrom](#) [Adam Finkelstein](#) [J. French](#) [Stone](#)

RKBExplorer » Networks (Communities of Practice)

Les Carr

Les Carr

[Display Metadata](#)

[Christopher Bailey](#) [David Barron](#) [Sean Bechhofer](#) [Donna Bergmark](#) [Tim Brody](#) [Grainne Conole](#) [Hugh Davis](#) [David De Roure](#) [Samhaa El-Beltagy](#) [Quintin Gee](#) [Nicholas Gibbins](#) [Lester Gilbert](#) [Yves Gingras](#) [Hugh Glaser](#) [Carole Goble](#) [Simon Grange](#) [Christopher Gutteridge](#) [Wendy Hall](#) [Stevan Harnad](#) [Stephen Harris](#) [Jessie Hey](#) [G Hill](#) [Steve Hitchcock](#) [Neil Jefferies](#) [Zhuoan](#) [Timothy](#)

RKBExplorer » Networks (Communities of Practice)

Les Carr

Les Carr

Full Name:

[Les Carr](#)

Telephone:

+44 (0)23 8059 4479

Email Address:

lac@ecs.soton.ac.uk

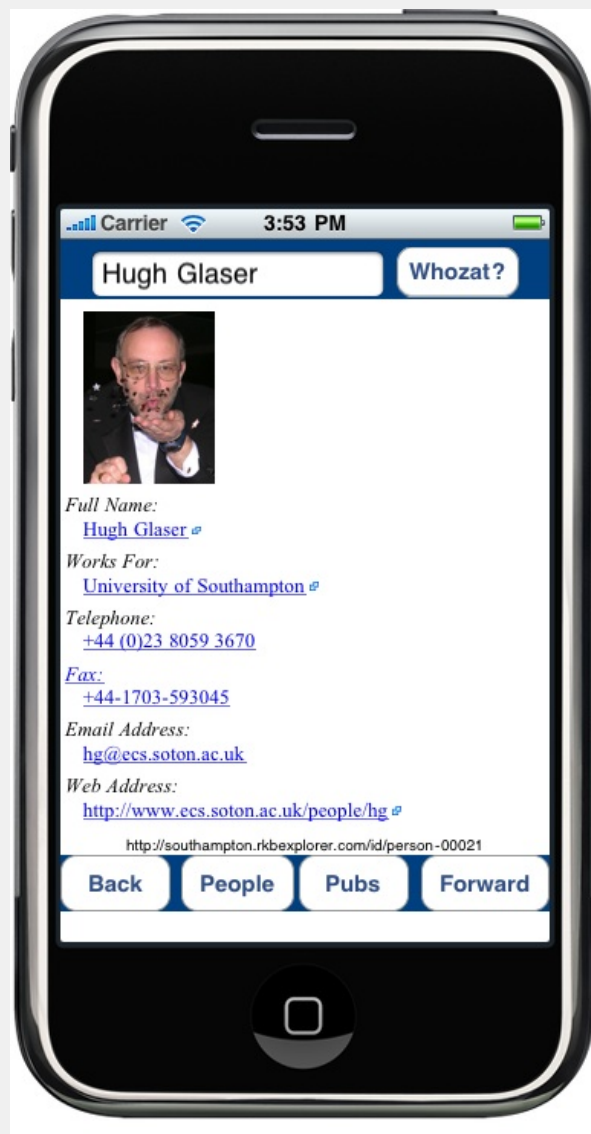
Web Address:

<http://www.ecs.soton.ac.uk/~lac/>

Raw Linked Data:

[Resolvable URI](#)

[All Known Co-reference Data](#)



Concluding Remarks

- **Linked Data is not hard**
 - URIs are easy
 - RDF is just another format
- **Coreference is important**
- **Ian Millard did the work**
- **<http://eprints.ecs.soton.ac.uk/18691/>**