## Southampton

School of Electronics and Computer Science

# RKBExplorer.com:

Anatomy of a Semantic Web Application

**DERI** 

Hugh Glaser & Ian Millard 16th December 2008



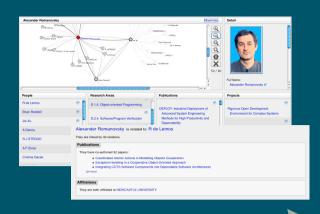


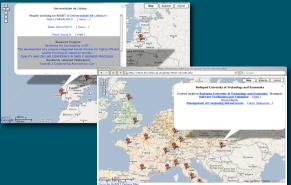
## Context

- CSAKTiveSpace
  - AKT Project
  - First Semantic WebChallenge winner2003



- ReSIST EU Network of Excellence in Resilient Systems
  - Knowledge-enabled infrastructure
  - Jan 2006 Dec 2008

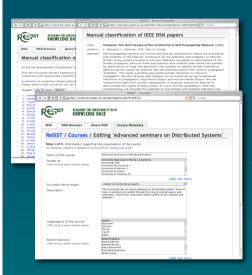


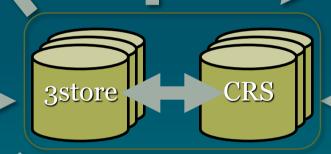


Southampton

School of Electronics and Computer Science







**Offline Conversion** & Versioning





#### acm.rkbexplorer.com

#### SPARQL Query Interface

This interface permits queries to be

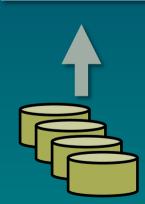
Result format: HTML Table | \$ Query:

PREFIX rdf: <a href="http://www.w3.org/1999/02/22">http://www.w3.org/1999/02/22</a>
PREFIX rdfs: <a href="http://www.w3.org/2000/01/rd">http://www.w3.org/2000/01/rd</a>
Presults in repository and PREFIX owl: <a href="http://www.w3.org/2002/07/ov">http://www.w3.org/2002/07/ov</a> PREFIX akt: <a href="http://www.aktors.org/ontology">http://www.aktors.org/ontology</a> PREFIX akts: <a href="http://www.aktors.org/ontology">http://www.aktors.org/ontology</a>

SELECT \* WHERE { ?s ?p "string" }

Subject	Property Object/Value	Source
Yannis Kalfoglou	sAtthd- Yannis Katloglou (Explore)	acm- periodicals.sdf >>>
Yannis Kalloglou	akt.hil- name Yannis Kalfoglou (Explore)	acm- proceedings.rdf >>>
Yannis Kalfoglou	sitches - NASAVW/U Selviere Research Lab. 100 Univ. Drive, Fairmort, WV, USA Dept. of Elec. and affiliation. Comp. Eng., 2355 Main Mell, Vencouver VST1Z4 BC, Canada, tim Remendes.com.	acm- periodicals.xdf >>
Yannis Kalloglou	akt.has- affiliation. http://scm.4besplorer.com/id/og-tatio855550cee703b7ct3bb7255697f	acm- proceedings.rdf >>>
Yannia Kalfoglou	atchas. Adv. Knowl. Tech., Clert. for Intell. Sys., and their Appa., Sich. of Informatice, The Univ. of atfastion. Exchange, UK manufactive about and Excels Superior de Tecnologies direformatics) affastion. Committees, Ut-W. Int. de Catalanya, Spain	acm- periodicals.nd >>>
Yannis Kallaglou	xf/type akt/Affiliated-Percon	acm-

**Ontology Mapping Service** 





## Communication

- Ontologies
  - General Scientific Endeavour
  - Domain-specific
  - Support (geospatial, etc)
- Open Local Knowledge HTTP
  - Resolvable URIs
  - SPARQL
- Uses Remote Knowledge
  - Resolves URIs with caching



## Components 1

- Semantic Web infrastructure throughout
- Triplestore for each source
  - Putting the Web in Semantic Web
  - Stores RDF (Subject, Predicate, Object)
  - We use 3store
- Linked Data
  - 303 and content negotiation architecture with caching



## Components 2

- Co-Reference Subsystem
  - CRS more later
- Community of Practice Analysis
  - Why do you think that?
- Ontology Mapping
  - Dealing with other Ontologies
- NLP for text classification
- Caching everywhere



## Components 3

- Application Middleware
  - URI Equivalence Closure
  - RDF Graph Closure
- Semantic Sitemap
  - Facilitate Search Engines



## **User Interaction**

- Semantic MediaWiki
- Custom form interfaces
- Google Maps
- Raw Knowledge Browser

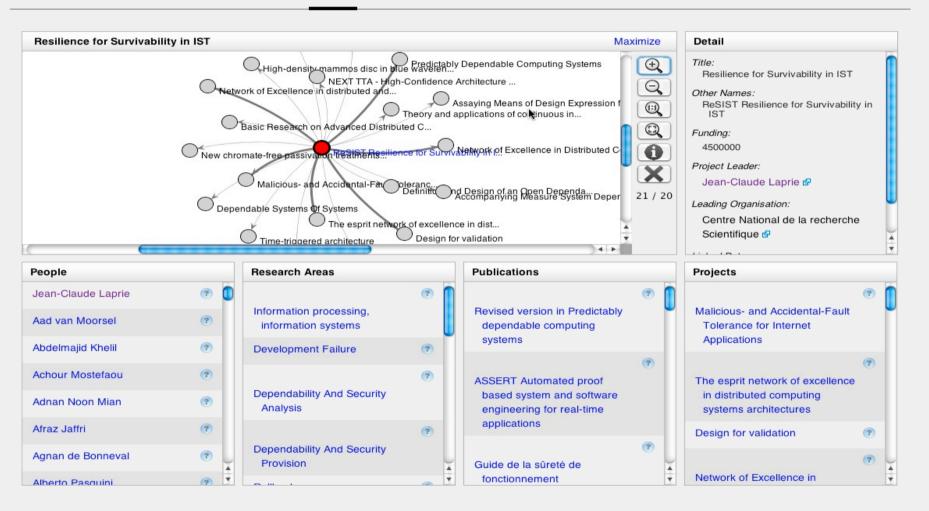
- RKBExplorer
- Why do you think that? information



#### www.rkbexplorer.com/explorer/

people research areas publications projects search

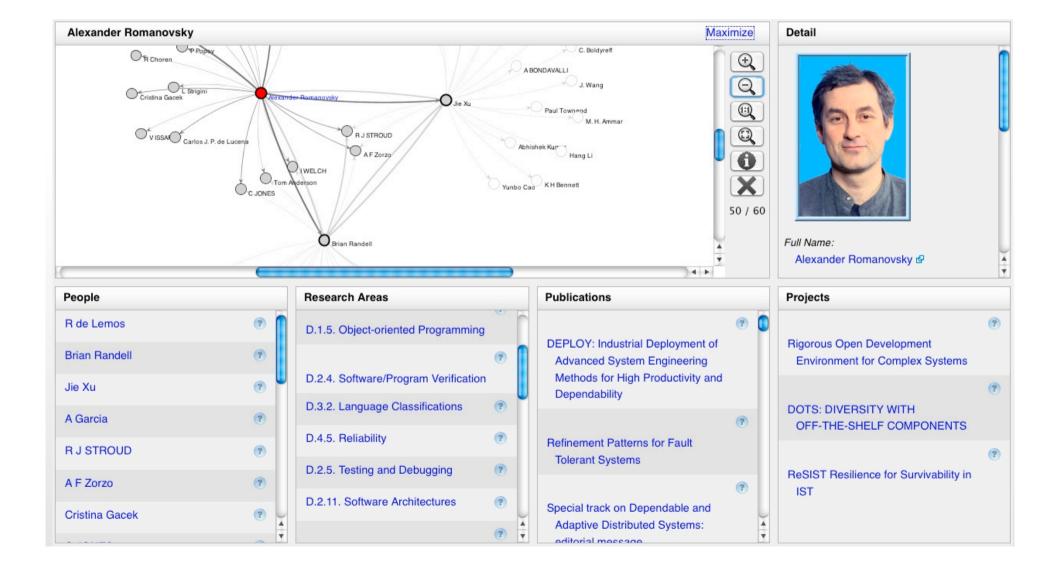
recently viewed reset help

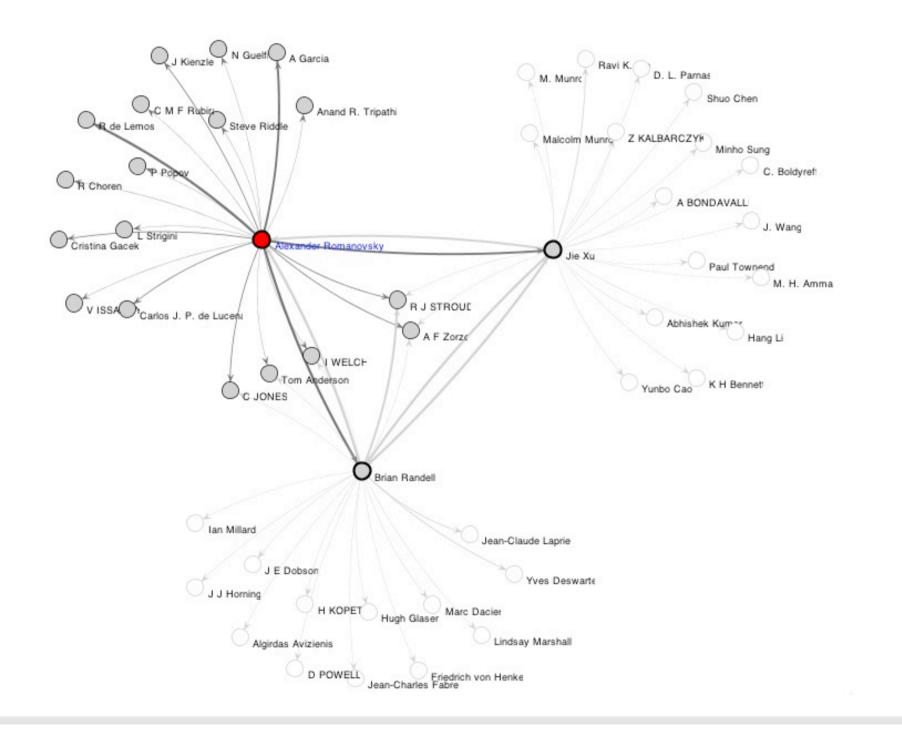


about I news I system requirements I acknowledgements I contact



## Focusing on a Person







## Why do you think that?

#### Alexander Romanovsky is related to R de Lemos

They are linked by 33 relations.

#### **Publications**

They have co-authored 32 papers:

- Coordinated Atomic Actions in Modelling Objects Cooperation
- Exception Handling in a Cooperative Object-Oriented Approach
- Integrating COTS Software Components into Dependable Software Architectures

(29 more)

#### **Affiliations**

They are both affiliated to NEWCASTLE UNIVERSITY.

This is a page that gives a simple demonstration showing papers which have been deemed related through textual analysis by IAI Saarbrucken. Up to the top 20 are listed for each paper, when they meet a simple thresholding:

The 1980 paper Exception Handling and Software-Fault Tolerance [browse]

is very strongly related to

- [browse] 2003 "Automatic detection and masking of non-atomic exception handling" [PDF]
- [browse] 1989 "Formal Verification of Programs with Exceptions"
- [browse] 1983 "Programming Reliable and Robust Software in ADA"

is strongly related to

- [browse] 1998 "Improving software robustness with dependability cases" [PDF]
- [browse] 1999 "Wrapping windows NT software for robustness" [PDF]
- [browse] 1981 "Exception Handling and Error Recovery Techniques in Modular Systems An Application to the Isaure System"
- [browse] 2003 "Deadlock resolution via exceptions for dependable Java applications" [PDF]
- [browse] 2002 "Robust software no more excuses" [PDF]

is related to

- [browse] 1995 "Fault tolerance in concurrent object-oriented software through coordinated error recovery" [PDF]
- [browse] 2004 "Implementing simple replication protocols using CORBA portable interceptors and Java serialization" [PDF]
- [browse] 1984 "Fault Tolerance Using Communicating Sequential Processes"
- [browse] 2001 "Middleware support for voting and data fusion" [PDF]



## RESILIENCE FOR SURVIVABILITY IN IST KNOWLEDGE BASE

Wiki

**RKB Browser** 

Query RKB

Course Metadata

#### **ReSIST / Welcome**

Welcome to the ReSIST Wiki, which is the internal communication mechanism for the EU funded ReSIST "Network of Excellence" ®.

Note that virtually all pages are private, and viewable only to ReSIST members who have logged in.

Most content can be found by firstly browsing the main ReSIST page, which details the different research areas in which activities are ongoing as part of the project.

If you have any questions or problems, please check that they have not previously been answered in the frequently asked questions, before contacting lan Millard or Hugh Glaser at Southampton.

#### **Quick Links**

- · Frequently asked questions
- · ReSIST project page
- · Recent changes to the wiki
- · Upload new file / View uploaded files
- ReSIST members / photos / locations
- · Calendar of Events
- Browse ☑, query ☑, or find out more about the Resilience Knowledge Base

#### **Editing tools**

- » View Page
  » Discuss this page
- » Edit this page
- » History
- » Protect
  » Delete
- » Move

[edit]

» Watch this page

#### Personal tools

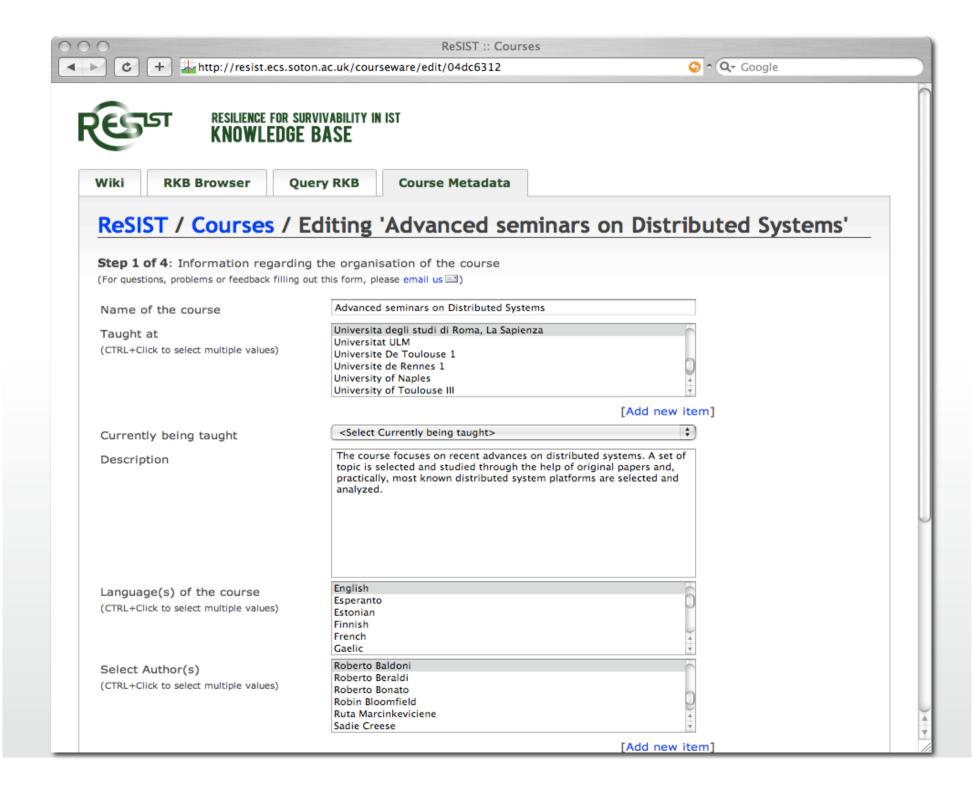
- » hugh glaser
- » My talk
- » Preferences
- » My watchlist
- » My contributions
- » Log out

#### Search



#### Toolbox

- » What links here
- » Related changes
- » Upload file
- » Special pages



#### **ReSIST / Resilience-Explicit Computing Mechanisms**

Name of the resilience mechanism (A title to identify your mechanism) N-Version Programming/1/1

Southampton
School of Electronics
and Computer Science

#### Submitted by

(The person(s) identified here shall be the point of contact for any queries relating to data entered into this form about this mechanism)

#### Zoe Andrews

#### Author of mechanism

(Click on the "add new item" link to search for, and add, authors of this mechanism. These people should have a good understanding of the mechanism and may be the same as those identified in the previous question)

#### Algirdas Avizienis

#### Associated projects

(Click on the "add new item" link to search for, and add, projects that are associated with this mechanism. Possible associations include projects that: funded research on the mechanism; address similar aims; or use similar techniques) <None>

#### Mechanism Objectives

(Summary of the purpose of your mechanism in a sentence or two) To utilise design diversity and voting in order to tolerate software faults

#### Detailed Description

(Either enter a detailed description of the mechanism here, should be detailed enough for the reader to be able to re-create the mechanism, or reference a paper with such text in below) The information here applies to the specific variant of the mechanism NVP/1/1, described in "Definition and Analysis of Hardware- and Software-Fault Tolerant Architectures". The specific variant considered, NVP/1/1, has three diverse implementations of a software module.

For a more general overview of the mechanism please see "The N-Version Approach to Fault-Tolerant Software".

#### Detailed Description Publication (If applicable (see above), click on the "add

Definition and Analysis of Hardware- and Software-Fault-Tolerant Architectures

#### Editing "N-Version Programming/1/1"

Step 5 of 7: Resilience metadata - how the mechanism helps a system's resilience (For questions, problems or feedback filling out this form, please email us )

#### Failure Modes

(Select the ways in which your mechanism can fall to function as intended. To help you to decide what the appropriate failure modes are you could treat your mechanism as a black box and think about the kinds of failures you expect to observe from it. The terms in this list are taken from the ReSIST ontology on security and dependability.) (CTRL+Click to select multiple values)

Consistent Failures	
Content And Timing Failure	
Content Failure	
Early Timing Failure	
Erratic Failure	
False Alarm	

#### Accidental Fault **Budget Failure** Catastrophic Error Catastrophic Failure Commission Fault Complete Development Failure

#### Threats Addressed

(Select the threats to resilience that your mechanism aims to address, le the faults it aims to remove, the errors it aims to compensate for and the failures it aims to prevent. The terms in this list are taken from the ReSIST ontology on security and dependability.)

(CTRL+Click to select multiple values)

#### Resilience Metadata

In this question you are asked to think about the effect your mechanism has on the resilience of a system. If you were to compare your mechanism to a different mechanism addressing a similar aim, what data would you use to choose which was fit for a specific purpose? This question allows you to define such metrics and associate a value with them for your mechanism. New resilience metadata metrics and values can be added to this list by clicking on the "add new item" link. Existing metadata instances can be deleted or edited by clicking the cross or the pencil next to them respectively. Note that when you edit some metadata a new version is saved as well as the old one, which can then be deleted.)

Time-dependent probability (P(t)) of undetected failure POFOD (Undetected) \* application software's execution rate \* t Probability

Southampton Southampton

**School of Electronics** 

and Computer Science

Time-dependent probability (P(t)) of failure POFOD \* application software's execution rate \* t Probability

Time-dependent probability (P(t)) of detected failure



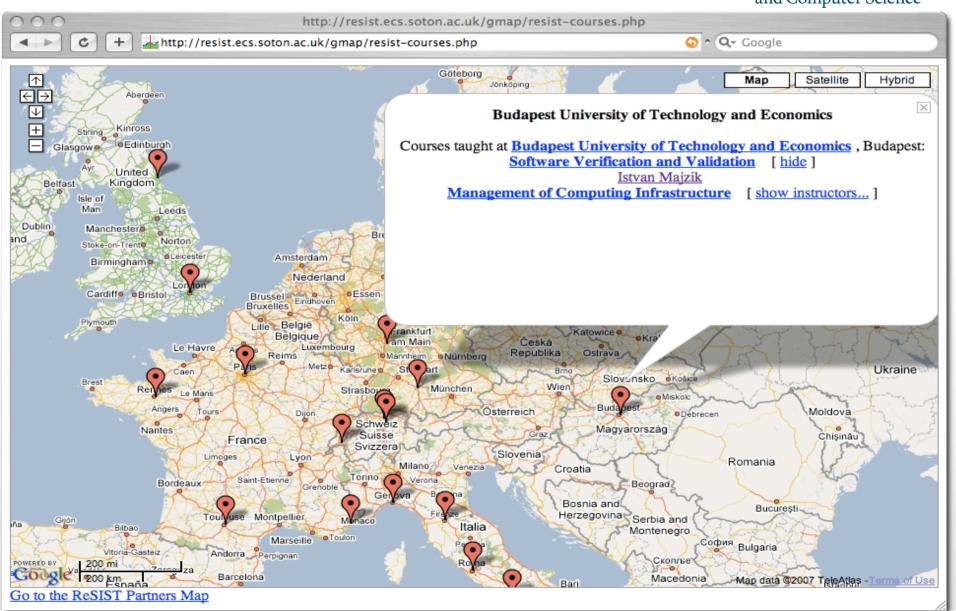




## Where is it Taught?



School of Electronics and Computer Science





## Knowledge Sources

- Partners
- Publications
- Funding Agencies
- Project Wiki
- Courseware
- Resilient-Explicit Computing

Wide range, don't just look where you expect to find



## Some Underlying Sources

#### acm.rkbexplorer.com

budapest.rkbexplorer.com citeseer.rkbexplorer.com

cordis.rkbexplorer.com courseware.rkbexplorer.com

darmstadt.rkbexplorer.com

#### dblp.rkbexplorer.com

deep blue.rk bexplorer.com

deploy.rkbexplorer.com

epsrc.rkbexplorer.com

eurecom.rkbexplorer.com

ft.rkbexplorer.com

ibm.rkbexplorer.com

## ieee.rkbexplorer.com

irit.rkbexplorer.com

italy.rkbexplorer.com

kaunas.rkbexplorer.com

#### kisti.rkbexplorer.com

laas.rkbexplorer.com

lisbon.rkbexplorer.com

newcastle.rkbexplorer.com

#### nsf.rkbexplorer.com

pisa.rkbexplorer.com

rae2001.rkbexplorer.com

#### resex.rkbexplorer.com

roma.rkbexplorer.com

southampton.rkbexplorer.com

ulm.rkbexplorer.com

unlocode.rkbexplorer.com

wiki.rkbexplorer.com

Range from a few 100 to more than 10,000,000 "facts"



## For example

- Statistics for repository kisti.rkbexplorer.com
  - Last data assertion 2008-09-18 17:16:41
  - Number of triples 12815162
  - Number of symbols 3239105
  - Size of RDF dataset 671M



## Co-Reference

- Co-Reference is a Big Problem
  - Identifying multiple URIs for one resource
  - Rejecting incorrectly conflated resources
  - Publishing
  - Using
- Coldstart
  - A serious problem
  - Nothing is linked to anything

## Co-Reference Closure



## School of Electronics and Computer Science

#### Complete Co-Reference Information

This service computes the equivalence class within the known URIs for a specified URI, by consulting all relevent CRS knowledge bases.

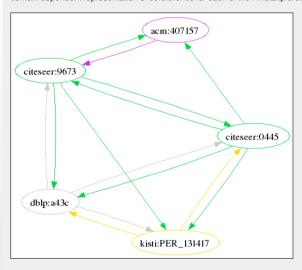
<a href="http://kisti.rkbexplorer.com/id/PER\_00000000000000131417">http://kisti.rkbexplorer.com/id/PER\_000000000000000131417</a>

#### Go

#### Equivalent URIs...

- 1. (Canon) http://acm.rkbexplorer.com/id/person-407157
- 2. http://citeseer.rkbexplorer.com/id/resource-CSP179673
- 3. http://citeseer.rkbexplorer.com/id/resource-CSP180445
- http://dblp.rkbexplorer.com/id/people-1ec5a600299222dd6374695ef5214f05-90d423eb148125a6e5c573dc5a15a43c
- 5. http://kisti.rkbexplorer.com/id/PER\_0000000000000131417

The following diagram shows the interconnectivity between the CRS knowledge bases which maintain the context-dependent representation of coreference for each of the RKBExplorer domains.



#### Seungwoo Lee

Showing information queried from all repositories ...

Subject	Property	Object/Value	Source
Seungwoo Lee	akt:full-name	Seungwoo Lee [Explore]	acm-periodicals.rdf >>
Seungwoo Lee	akt:full-name	Seungwoo Lee [Explore]	acm-proceedings.rdf >:
Seungwoo Lee	akt:full-name	Seungwoo Lee [Explore]	dblp-publications- 2001.rdf >>
Seungwoo Lee	akt:has-affiliation	Electrical and Computer Engineering Division, Pohang University of Science & Technology (POSTECH), Pohang, South Korea.gblee@postech.ac.kr	acm-periodicals.rdf >>
Seungwoo Lee	akt:has-affiliation	POSTECH, Pohang, Korea	acm-proceedings.rdf >:
Seungwoo Lee	kisti:engNameOfPerson	Seungwoo Lee [Explore]	datatypeproperties.ttl >
Seungwoo Lee	rdf:type	akt:Affiliated-Person	acm-periodicals.rdf >>
Seungwoo Lee	rdf:type	akt:Affiliated-Person	acm-proceedings.rdf >
Seungwoo Lee	rdf:type	Generic Agent	acm-periodicals.rdf >>
Seungwoo Lee	rdf:type	Generic Agent	acm-proceedings.rdf >:
Seungwoo Lee	rdf:type	Generic Agent	dblp-publications- 2001.rdf >>
Seungwoo Lee	rdf:type	akt:Person	acm-periodicals.rdf >>
Seungwoo Lee	rdf:type	akt:Person	acm-proceedings.rdf >
Seungwoo Lee	rdf:type	akt:Person	dblp-publications- 2001.rdf >>
Seungwoo Lee	rdf:type	PER_char(20)"^^	datatypeproperties.ttl >
Seungwoo Lee	rdf:type	PER_char(20)"^^	objectproperties.ttl >>>
Seungwoo Lee	rdf:type	PER_char(20)"^^	resources.ttl >>>
Subject	Property	Object	Source
Automatic acquisition of named entity tagged corpus from world wide web	akt:has-author	Seungwoo Lee	acm-proceedings.rdf >
A Corpus-Based Learning Method of Compound Noun Indexing Rules for Korean	akt:has-author	Seungwoo Lee	acm-periodicals.rdf >>
SiteQ: Engineering High Performance QA System Using Lexico-Semantic Pattern Matching and Shallow NLP.	akt:has-author	Seungwoo Lee	dblp-publications- 2001.rdf >>
A Corpus-Based Learning Method of Compound Noun			dblp-publications-



## CRS – Consistent Reference Service

- A service to manage and publish co-referent information
- Identify co-referent pairs using a set of tools
- Assert into the CRS
- Query the CRS

$$- URI_{i} -> \{ URI_{1}, ..., URI_{i}, ..., URI_{n} \}$$

Recommend a Canon



## CRS continued

- CRS Policies are defined by context
  - Often one per Triplestore
  - Can be many per Triplestore for different purposes
  - May not be associated with a particular Triplestore
- Maintenance
  - Provenance
  - Rollback
- Can be used to infer owl:sameAs



## Dealing With Non-SPARQL KBs

- The RKBExplorer application uses SPARQL to query the KBs
  - But needs to access data from KBs that only offer resolvable URIs
- So resolve such a URI
- Cache the RDF with associated resolved RDF locally
- Query the local cache



## Dealing With Different Ontologies

- The RKBExplorer application uses a particular ontology
  - Some KBs will use different ontologies
  - Eg kisti.rkbexplorer.com
- One solution
  - Represent the ontology relationship in RDF (as far as possible)
  - Resolve the URI through the mapping service to get RDF in the required ontology



## Supporting resilience

- People, Publication, Projects, Research Areas
- Resilience-related topics
- Resilience-Explicit Computing
- Educational Resources

- In the future
  - Automating discovery of issues and solutions
    - -Design time
    - -Run time

# Southampton Finding mechanisms that are appropriate for Hardware and Aerospace

SELECT DISTINCT ?mechanismURI ?mechanismName ?metadataName ?metadataValue WHERE {
?mechanismURI rdf:type resex:Resilience-Mechanism .
?mechanismURI resex:applies-to-technology akt:Hardware-Platform .
?mechanismURI resex:has-application-domain acm:J.2.0 .
?mechanismURI rdfs:label ?mechanismName .
}

Result	Binding	Value
1	?mechanismURI	http://resex.rkbexplorer.com/id/resilience-mechanism-267972cd
	?mechanismName	N-Self-Checking Programming/1/1
2	?mechanismURI	http://resex.rkbexplorer.com/id/resilience-mechanism-e679bd05
	?mechanismName	N-Version Programming/1/1
3	?mechanismURI	http://resex.rkbexplorer.com/id/resilience-mechanism-7425f52f
	?mechanismName	Recovery Blocks/1/1

## Inspecting metadata, number of wahren school of Electronics

SELECT DISTINCT?mechanismURI?mechanismName?metadataName?metadataValueWHERE {

?mechanismURI rdf:type resex:Resilience-Mechanism.

?mechanismURI resex:applies-to-technology akt:Hardware-Platform .

?mechanismURI resex:has-application-domain acm:J.2.0.

?mechanismURI rdfs:label ?mechanismName.

?mechanismURI resex:has-resilience-metadata?metadata.

?metadata resex:metadata-type id:resilience-metadata-type-231c8583

?metadata resex:metadata-type ?mt . ?mt rdfs:label ?metadataName .

?metadata resex:has-value ?metadataValue

Result	Binding	Value
1	?mechanismURI	http://resex.rkbexplorer.com/id/resilience-mechanism-7425f52f
	?mechanismName	Recovery Blocks/1/1
	?metadataName	Number of variants
	?metadataValue	2
2	?mechanismURI	http://resex.rkbexplorer.com/id/resilience-mechanism-e679bd05
	?mechanismName	N-Version Programming/1/1
	?metadataName	Number of variants
	?metadataValue	3
3	?mechanismURI	http://resex.rkbexplorer.com/id/resilience-mechanism-267972cd
	?mechanismName	N-Self-Checking Programming/1/1
	?metadataName	Number of variants
	?metadataValue	4

and Computer Science

# Inspecting metadata, average cost of implementing faul tolerant system vs- cost of implementing non fault tolerant system

SELECT DISTINCT?mechanismURI?mechanismName?metadataName?metadataValueWHERE {

?mechanismURI rdf:type resex:Resilience-Mechanism.

?mechanismURI resex:applies-to-technology akt:Hardware-Platform .

?mechanismURI resex:has-application-domain acm:J.2.0.

?mechanismURI rdfs:label ?mechanismName.

?mechanismURI resex:has-resilience-metadata ?metadata .

?metadata resex:metadata-type id:resilience-metadata-type-de1eddf9.

?metadata resex:metadata-type ?mt . ?mt rdfs:label ?metadataName .

?metadata resex:has-value ?metadataValue

•	
,	
_	

Result	Binding	Value
1	?mechanismURI	http://resex.rkbexplorer.com/id/resilience-mechanism-e679bd05
	?mechanismName	N-Version Programming/1/1
	?metadataName	Av CFT/CNFT
	?metadataValue	2.25
2	?mechanismURI	http://resex.rkbexplorer.com/id/resilience-mechanism-267972cd
	?mechanismName	N-Self-Checking Programming/1/1
	?metadataName	Av CFT/CNFT
	?metadataValue	3.01
3	?mechanismURI	http://resex.rkbexplorer.com/id/resilience-mechanism-7425f52f
8	?mechanismName	Recovery Blocks/1/1
	?metadataName	Av CFT/CNFT
	?metadataValue	1.75

and Computer Science

## Comparison of the operational overheads in Scientific of the operation of t fault has occurred

**School of Electronics** and Computer Science

SELECT DISTINCT ?mechanismURI ?mechanismName ?metadataName ?metadataValue WHERE {

?mechanismURI rdf:type resex:Resilience-Mechanism.

?mechanismURI resex:applies-to-technology akt:Hardware-Platform.

?mechanismURI resex:has-application-domain acm:J.2.0.

?mechanismURI rdfs:label ?mechanismName.

?mechanismURI resex:has-resilience-metadata?metadata.

?metadata resex:metadata-type id:resilience-metadata-type-3443934c.

?metadata resex:metadata-type ?mt . ?mt rdfs:label ?metadataName .

?metadata resex:has-value ?metadataValue

u		
>		
•		

Result	Binding	Value
1	?mechanismURI	http://resex.rkbexplorer.com/id/resilience-mechanism-7425f52f
	?mechanismName	Recovery Blocks/1/1
	?metadataName	Errors op time overheads
	?metadataValue	One variant and acceptance test execution
2	?mechanismURI	http://resex.rkbexplorer.com/id/resilience-mechanism-267972cd
	?mechanismName	N-Self-Checking Programming/1/1
	?metadataName	Errors op time overheads
	?metadataValue	Possible result switching
3	?mechanismURI	http://resex.rkbexplorer.com/id/resilience-mechanism-e679bd05
	?mechanismName	N-Version Programming/1/1
	?metadataName	Errors op time overheads
	?metadataValue	Usually negligible

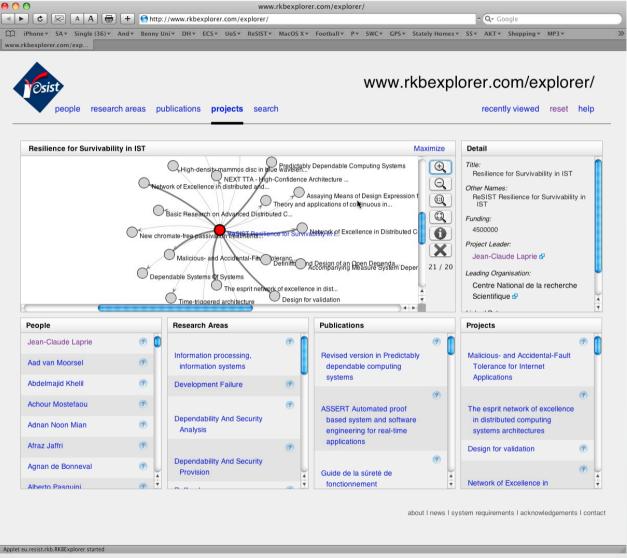


## Concluding Remarks

- Major Data Fusion using Semantic Web Technologies
- Many things can be cast in a Semantic Web framework
- Linked Data works pretty well
- RDF works pretty well
- A little Ontology goes a long way
- Co-Reference is the biggest problem
  - But is tractable

## RKBExplorer.com/explorer/ - Try it!





http://eprints.ecs.soton.ac.uk/17025