



Prototype Knowledge Base: an on-line information service in dependability and security

Hugh Glaser

Electronics & Computer Science University of Southampton

Budapest, 22nd. March 2007







With



- Ian Millard
- Afraz Jaffri
- Benedicto Rodriguez
- ReSIST Partners
 - esp. Brian Randell





Background: Semantic Web Challenge 2003 Winner

- CS AKTive Space
 - Gather data
 - UK People, projects, publications
 - Research funding
 - Top Universities
 - Geographical presentation
- AKT Project (www.aktors.org)







- Scientific Intelligence
 - Who is doing what where?
 - What impact are they having?
- Integrating resources
 - CORDIS, Institutional DBs and web sites, ePrints, NSF, CiteSeer, RISKS list, ISO LoCodes...
- Information: distributed and heterogeneous
 - Not under own control
 - Not in a common format
 - Not where you expect it
- Presenting to users & agents

000	AKT: CS Aktive Space - Mozilla		0
AKT CS Aktive S	·		
About this page (** research area to general and the search area probability are authors discrete mathematical analysis and area to a search and a s	Radia: Map: 200 miles	Researcher No Shappin RC Travelens L Morres L Morres L A Carr	
Detail: NR Shadbolt			
Email nrs@ecs.soton.ac.uk Tel +442380597882 Fax +442380592865 Research interests Fluid Dynamics Aerodynamics Design and Testing T	onics and Computer Science, University of Southampton		
Biological Sciences D			E of





ReSIST - Start Again

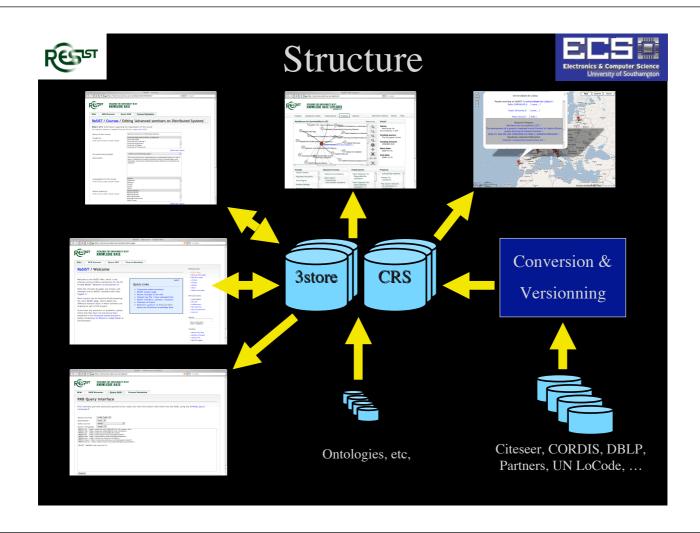
- A ReSIST Knowledge Base The *RKB*
- Project Infrastructure support
- Europe (no longer UK-centric), the World
- Up to date
- Extra subject targets (resilience)
- Browser & platform independent
- Engineer for maintenance
 - Empower partners and other contributors
 - Empower other application builders





ReSIST - and deliver

- D10 2007-01-01T00:00:00A
- In fact it is just a URI to a service:
 - http://resist.ecs.soton.ac.uk/sparql/
- Or the raw content can be browsed
 - <u>http://resist.ecs.soton.ac.uk/browse/</u>
- But there is a brand new faceted browser
 - <u>http://resist.ecs.soton.ac.uk/explorer/</u>
- The RKB is embedded in the infrastructure
- The prototype is already being used





Sources



- Publications
 - Partners
 - Citeseer
 - DBLP
 - ACM
 - DSN & FTCSSeries
- Documents
 - RISKS Digest

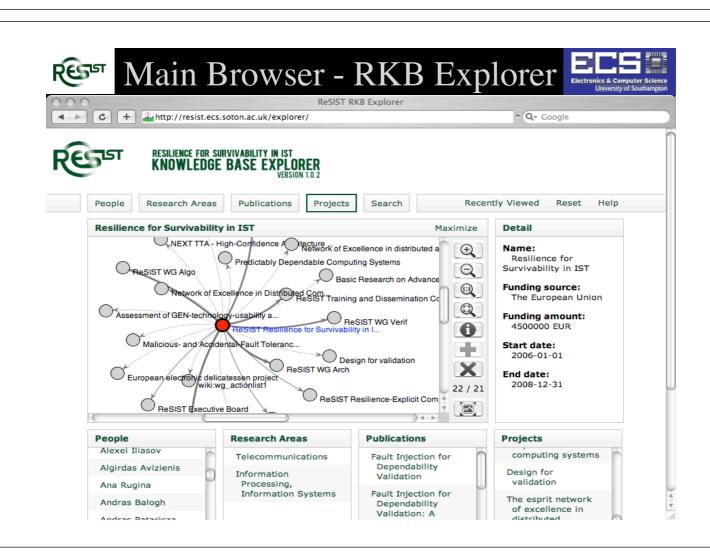
- Projects
 - CORDIS
 - -NSF
- People
 - Partners
- Support
 - UN LoCode

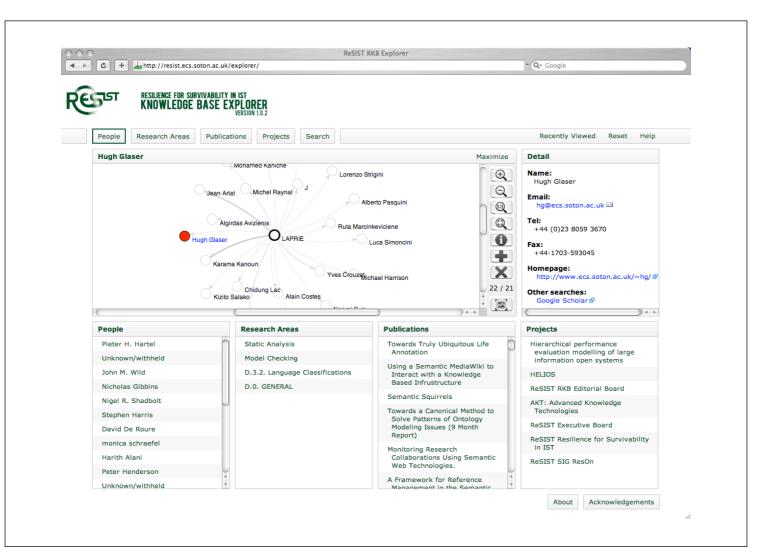


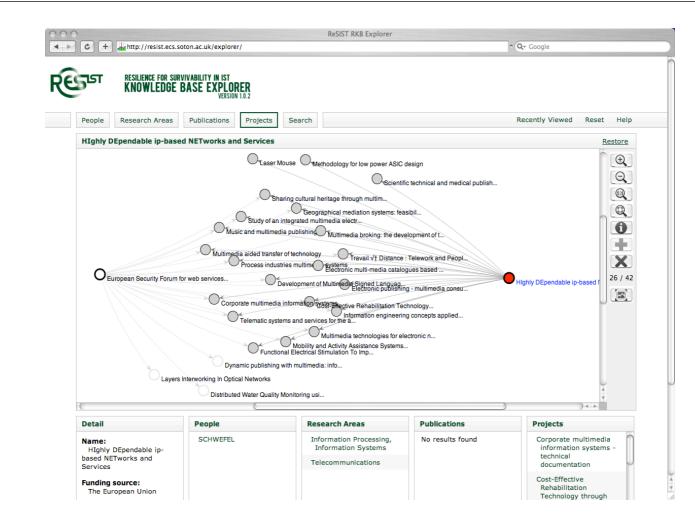


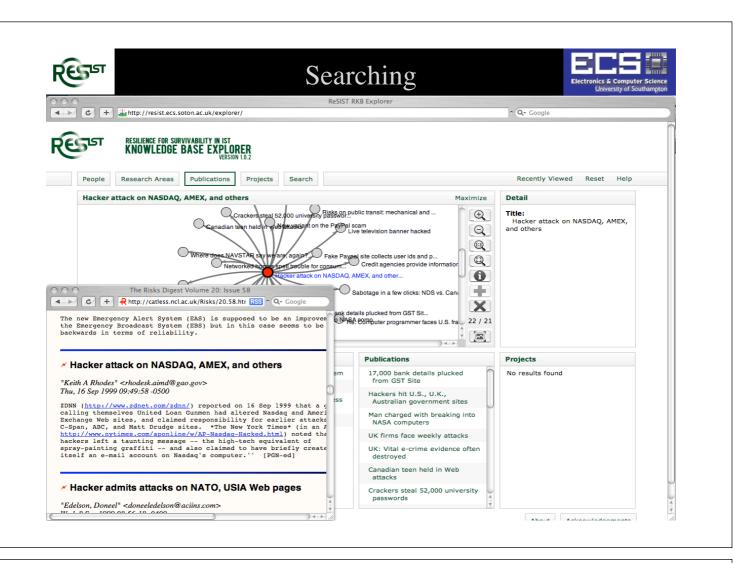
Ontologies etc.

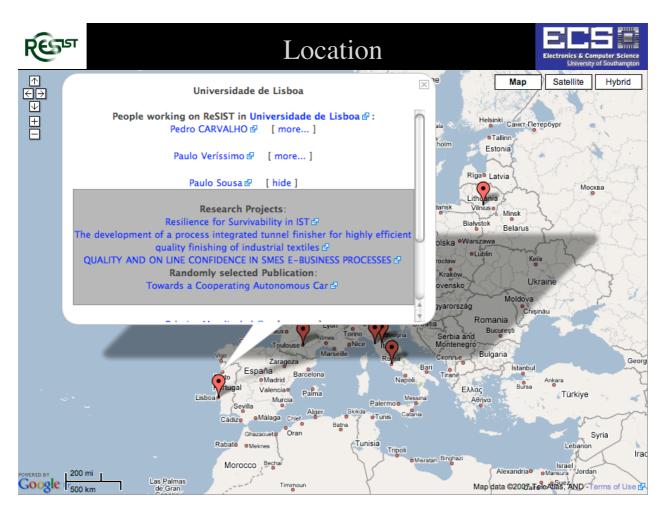
- AKT Ontology
 - Scientific Research Activity
 - Dates
 - Location
 - **–** ...
- ALRL Paper
- Courseware (extension of LOM)
- RISKS Codes
- ACM Classification

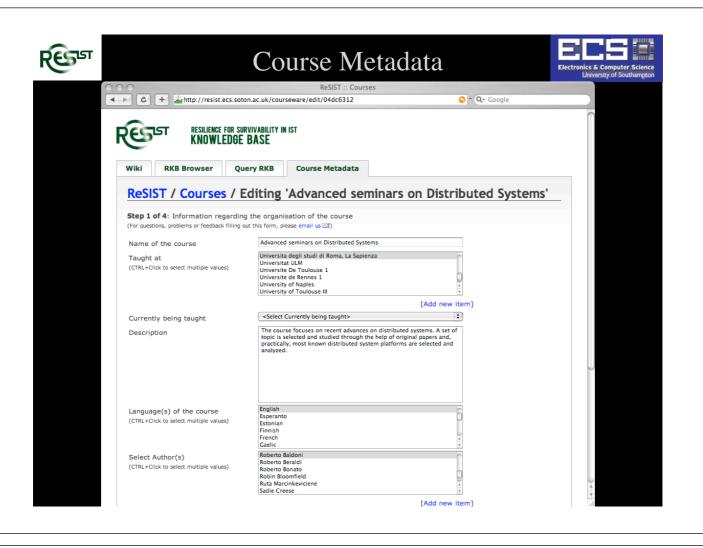


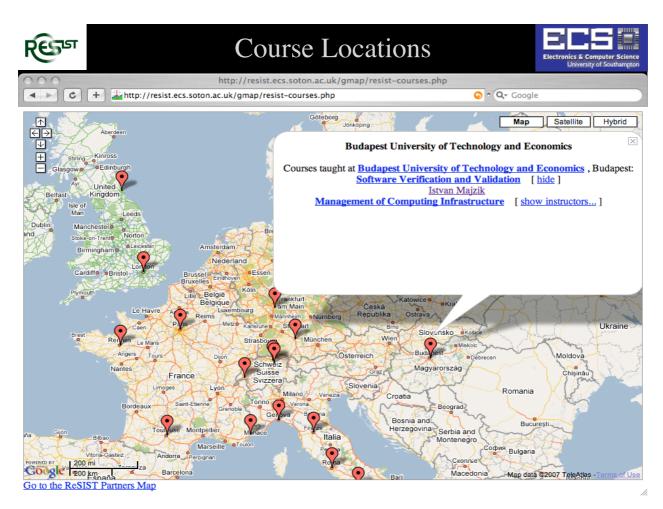


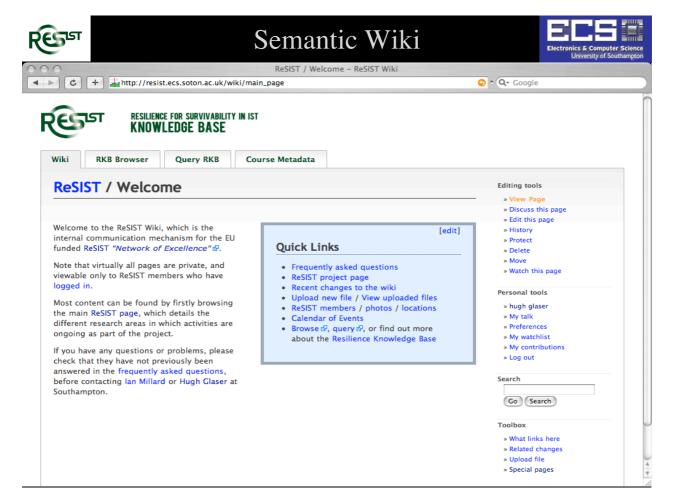














Please select the topics from within the hierarchy below that best match your research interests within the ReSIST NoE.

It is best to "drill down" as far as possible, and to select the most specific topics. Selecting higher level topics will indicate that your are interested in all of the sub-topics, which are selected for you.

Note however that this is not strictly a tree, as some topics appear in multiple places within the hierarchy. In these cases the "other" instances are automatically selected when you tick a topic area which exists in more than one category.

As is usual within the wiki, clicking a blue link should take you to a page describing the subject of that link.

Happy clicking:)

akt:Research Area

Dependability And Security, Trustworthiness

Two somehat overlapping concepts, with dependability being an integrating concept that encompasses the attributes: availability, reliability, safety:, integrity and maintainability, while security encompasses comfidentiality as well as integrity and availability.

Dependability, High Confidence, Survivability

The original definition of dependability is: the ability to deliver service that can justifiably be trusted. The alternate definition, that provides the criterion for deciding if the service is dependable, is: the dependability of a system is the ability to avoid service failures that are more frequent and more severe than is acceptable.

□ Dependence

The dependence of system A on system B represents the extent to which system A's dependability is (or would be) affected by that of System B.

Accepted dependence - where the dependence of a user on a given system represents the extent to which the user's dependability is (or would be) affected by that of the system. (The acceptance of this state of affairs by the user may be willing or unwilling, and

Attribute Of Dependable Systems

The dependability properties that are expected from a system, and in terms of which a system's dependability can be assessed with respect to the threats and the means to oppose these threats.



Classifying





RESILIENCE FOR SURVIVA KNOWLEDGE BAS

RKB Browser

Query I

Manual classification of

To aid the development of automatic cl

conference with appropriate research a

A selection of randomly-chosen publica Please select a title you think you can

Suggest 100 \$ titles. Refresh

- Hotspots: The Root Causes of Non
- Dataflow anomaly detection, 2006
- Tracking Probabilistic Correlation The final nail in WEP's coffin, 2006
- SubVirt: implementing malware wi
- Cost-Effective Configuration of Con
- Dynamic Verification of Memory Co A Component-Level Path Compositi
- A Dependable System Architecture
- 10. Lucky Read/Write Access to Robust

- Cobra: fine-grained malware analy Privacy and contextual integrity: fr 12.
- 13. BlueGene/L Failure Analysis and Pro
- Performance Assurance via Softwa
- 15 A General Framework for Scalabilit Deterring voluntary trace disclosure
- 16.
- A large-scale study of failures in hi
- Using Attack Injection to Discover
- Fast Abstracts, 2006
- cting and Analyz

Manual classification of IEEE DSN papers

Hotspots: The Root Causes of Non-Uniformity in Self-Propagating Malware (2006)

eSIST :: Manual Classifie

Authors: F. Jahanian, F. Jahanian, Z.M. Mao, E. Cooke

Self-propagating malware like worms and bots can dramatically impact the availability and reliability of the Internet. Techniques for the detection and mitigation of Internet threats using content prevalence and scan detectors are based on assumptions of how

threats propagate. Some of these assumptions have recently been called into question by observations of huge discrepancies in the quantity of specific threats detected at different points around the Internet. We call these deviations from uniform propagation 'hotspots". This paper quantifies and explains these influences on malware propagation. We then propose that hotspots can be explained by two fundamental influences on propagation: algorithmic factors and environmental factors. We use measurement data from sensors deployed at 11 locations around the Internet to demonstrate the impact of these factors on worm and bot propagation. With this understanding, we simulate the outbreak of new threats with hotspots and show how algorithmic and environmental factors reduce the visibility of distributed detectors

resulting in the inability to identify new threats.

Please select:

Keywords: None

akt:Research Area

Popendability And Security, Trustworthiness

encompasses the attributes: availability, reliability, safety:, integrity and maintainability, while security encompasses comfidentiality as well as integrity and availability.

Dependability, High Confidence, Survivability

The original definition of dependability is: the ability to deliver service that can justifiably be trusted. The alternate definition, that provides the criterion for deciding if the service is dependable, is: the dependability of a system is the ability to avoid service failures that are more frequent and more severe than is acceptable

Dependence

The dependence of system A on system B represents the extent to which system A's dependability is (or would be) affected by that of System B.

■ Trust

Accepted dependence - where the dependence of a user on a given system



Browsing Raw Data



Alternative representations

» RDF export

O O Source of http://resist.ecs.soton.ac.uk/resolve/?resource=http%3A%.



RKB Browser

Query RKB

Course Metadata

🖒 🕇 🚣 http://resist.ecs.soton.ac.uk/browse/?resource=http%3A%2F%2Fcatless.ncl.ac.uk%2Fperson%2360d6abae 📀 🤈 💽 Google

RKB Browser:: John Rushby

Identifiers...

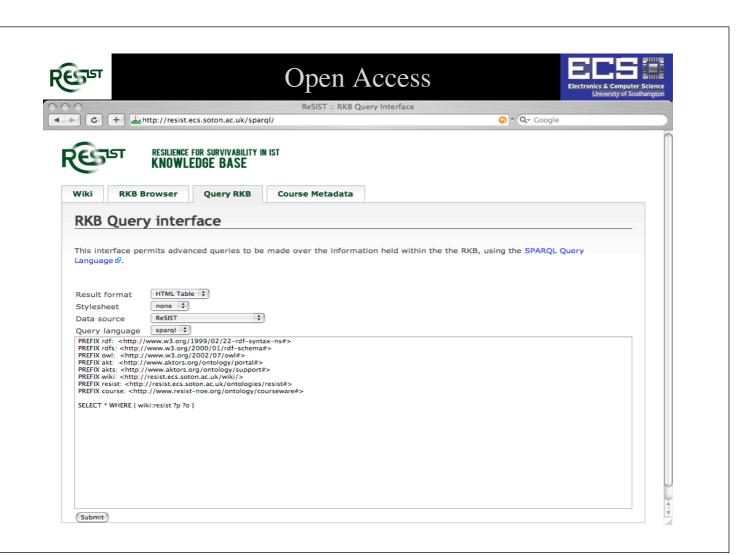
http://catless.ncl.ac.uk/person#60d6abae http://citeseer.ecs.soton.ac.uk/#CSP272905

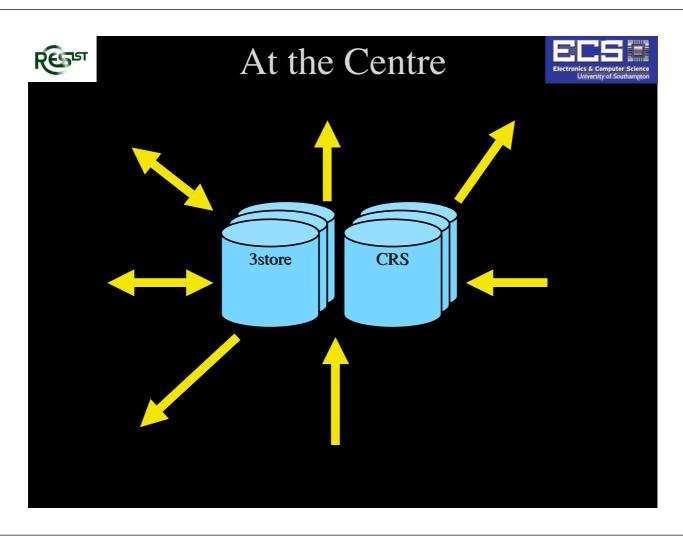
http://citeseer.ecs.soton.ac.uk/#CSP272906

http://citeseer.ecs.soton.ac.uk/#CSP272907

http://citeseer.ecs.soton.ac.uk/#CSP272908 http://citeseer.ecs.soton.ac.uk/#CSP272909

http://citeseer.ecs.soton.ac.uk/#P145810 http://citeseer.ecs.soton.ac.uk/#P570387 http://resist.ecs.soton.ac.uk/publications/person#f89fd02d http://resist.ecs.soton.ac.uk/wiki/User:john_rushby			<pre>drd:Description drd:Description rd:about="http://catless.ncl.ac.uk/Risks/13.77.html#subj2.1"> drd:Description rd:about="http://catless.ncl.ac.uk/person#8046aboe"/> drd:Description drd:Description rd:about="http://catless.ncl.ac.uk/Risks/13.77.html#subj3.1"> drd:Description rd:about="http://catless.ncl.ac.uk/Risks/13.77.html#subj3.1"> drd:Description rd:about="http://catless.ncl.ac.uk/Risks/13.77.html#subj3.1"> drd:Description rd:about="http://catless.ncl.ac.uk/Risks/13.77.html#subj3.1"> drd:Description</pre>			
Subject	Property		Object/Va			Ų
John Rushby	akt:family- name	Rushby		<pre>~rdf:Description rdf:about="http://catless.ncl.ac.uk/Risks/13.84.html#subj5.1" an8thas-author rdf:resource="http://catless.ncl.ac.uk/person#60d6abae"/> </pre>	>	¥
JOHN RUSHBY	akt:full- name	JOHN RUSHBY		acm- proceedings.rdf >>		
John Rushby	akt:full- name	John Rushby		acm- bookchapters.rdf >>		
John Rushby	akt:full- name	John Rushby		acm- periodicals.rdf >>		
John Rushby	akt:full- name	John Rushby		acm- proceedings.rdf >>		→









So what is RDF...?

- Resource Description Framework
- W3C recommendation
 - From Semantic Web research efforts
- Modelling language
 - Represents facts about resources
- Can model any abstract domain
 - Things do not have to be accessible on the web
 - But can be described in it



RDF: Basic components



• RDF graphs are formed by *triples*

subject predicate object

http://laas.fr/people#laprie laprie@laas.fr



Important Components 3store and CRS



- 3store
 - Open source semantic store
 - Scalable
 - ReSIST 50 million facts
 - (cf Wikipedia metadata)
- CRS Consistent Reference Service
 - Bridges between disparate sources



Openness



- Almost nothing shown was private
- Except
 - Wiki project discussion pages
 - But semantic relations go to RKB
 - Data entry
 - Controlled
 - Not moderated



Future for ReSIST & the RKB



- Improve on the Prototype
 - Sources
 - CRS
 - UI
- Resilient-Explicit Computing
 - Model expert knowledge
 - Model processes, components, mechanisms
- Support Engineer/Scientist
 - Move effectively between
 - System design
 - Knowledge Base
 - People
 - To choose cost, characteristics, etc
- Support Run-Time Deployment
 - Dynamic Reconfiguration



Future Resources



- Original proposal
 - Now primarily maintenance
- Victim of success?
 - Important infrastructure
 - Serious resources to be maintained
 - People want to provide data (costs)



Response



- ReSIST
 - Has increased future RKB resources
- Other Funding and Additionality
 - Lithuania & Saarbrücken
 - JISC
- Longer term
 - Self-funding SIGs, Clubs
 - Infrastructure EU, EPSRC, NSF
- Engineer for maintenance and Openess
- Open
 - Knowledge Sources
 - Knowledge Publishing



Some Review Highlights



- One year of work one RF funded
- ReSIST has done what it said it would do
 - And more
 - In particular, 1M -> 40M
 - Sophisticated UI
- Real tool for the network, from Day One
- Excellent Partner co-operation
 - Data
 - Evaluation
 - Ontology work
- Much Value in Expert Involvement

