ResIST's Plans

for a

Resilience Knowledge Base

Brian Randell

The Aim

- To provide convenient access to extensive up to date *integrated* information from across Europe about people, projects, institutions, events, publications, etc., of relevance to dependability and security research

- To obtain this information regularly and automatically from multiple heterogeneous sources, and to enable it to be queried and analyzed, from various viewpoints, *without* requiring uniformity of terminology

The Starting Point

- RKB will not be yet-another conventional web-site providing a manually-collected and classified collection of links to relevant information

- Rather it will be based on the very sophisticated CS AKtiveSpace System - a pilot knowledge base covering (UK) computer science as a whole

- CS AKTiveSpace was developed by the University of Southampton, as part of the AKT (Advanced Knowledge Techniques) IRC - a five-university six-year interdisciplinary research project, similar to DIRC, the Dependability IRC
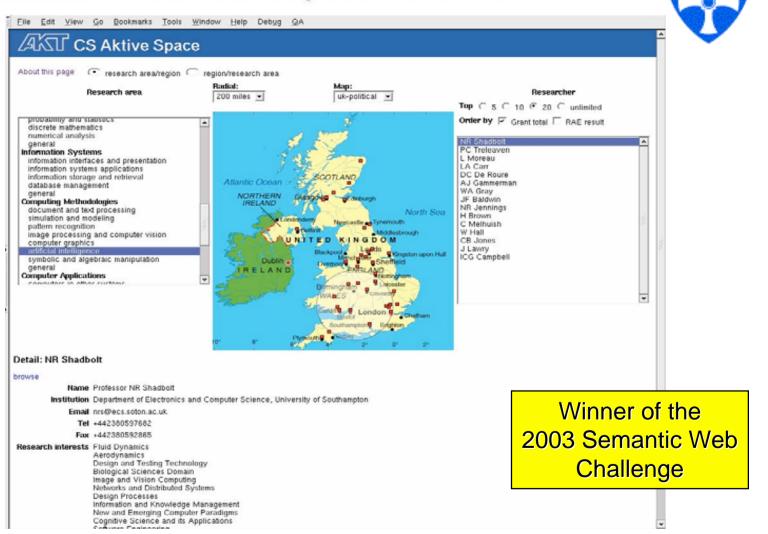
CS AKTiveSpace

- CS AKTiveSpace is the Computer Science version of a family of AKTiveSpace applications that "mediate between heterogeneous data sources, and present the results to users in digested and informative ways".

- It contains, *and inter-relates,* automatically-collected and updated information (some 7M RDF database triples) about UK research, in particular CS research, e.g.

  - The EPSRC research grant database
  - The HEFCE Research Assessment Exercise database
  - Information "scraped" from CS department web-sites

- It makes inferences from this inter-related information, e.g.

  - Identifying individuals forming a "community of practice"
  - Identifying major current research themes, based on funding and publication counts

CS AktiveSpace - one view

Winner of the
2003 Semantic Web
Challenge

An Example Use of CS AKTiveSpace - 1

- The user has begun by selecting a research area of interest (AI).

- The map shows, in red, those places that are considered from the sources to have activity in that area.

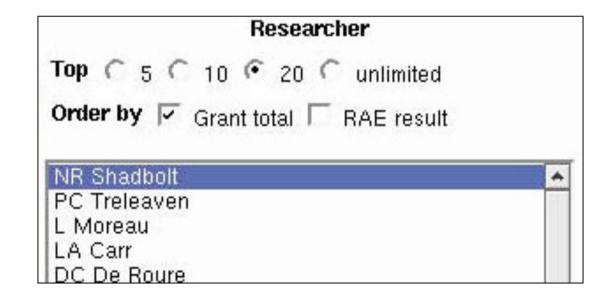- The chosen geographic area of interest has been indicated, by a circle

An Example Use of CS AKTiveSpace - 2

- This choice caused the 20 highest-funded researchers in AI within the area to be listed.

- One can then choose which researcher's details are to be displayed

Researcher

Top ○ 5 ○ 10 ● 20 ○ unlimited

Order by ☑ Grant total ☐ RAE result

NR Shadbolt
PC Treleaven
L Moreau
LA Carr
DC De Roure

An Example Use of CS AKTiveSpace - 3

- Choosing the first of these has caused Nigel Shadbolt's details to be displayed

- Scrolling down would show funded projects, publications, etc.

- This information comes from various entirely-separate sources

Detail: NR Shadbolt

browse

Name Professor NR Shadbolt

Institution Department of Electronics and Computer

Email nrs@ecs.soton.ac.uk

Tel +442380597682

Fax +442380592865

Research interests Fluid Dynamics
Aerodynamics
Design and Testing Technology
Biological Sciences Domain
Image and Vision Computing

Integration

- Information from each separate source is associated with an appropriate set of so-called "ontologies" - which are in effect dictionaries defining words and relations between words in terms of other words.

- Existing general ontologies will be used for such domains as "people", "publications" and "projects".

- "Basic Concepts and Taxonomy of Dependable and Secure Computing" (IEEE TDSC, Jan 2004) will be the starting point for a resilience ontology

- The equivalent of bi-lingual dictionaries will be used to indicate relations between words from different domains, e.g.:

    - is equivalent to
    - is a subset of

Integration

- Information from each separate source is associated with an appropriate set of so-called "ontologies" - which are in effect dictionaries defining words and relations between words in terms of other words.

- Existing general ontologies will be used for such domains as "people", "publications" and "projects".

- "Basic Concepts and Taxonomy of Dependable and Secure Computing" (IEEE TDSC, Jan 2004) will be the starting point for a resilience ontology

- The equivalent of bi-lingual dictionaries will be used to indicate relations between words from different domains, e.g.:
  - is equivalent to
  - is a subset of

- And perhaps :-)
  - is a poorly defined version of
  - is this year's Eurospeak for
  - is what everyone except Microsoft calls

Planned Contents of the RKB

- Information from ReSIST members' web-sites

- The entire CiteSEER CS publications database

- Contractual and publicity information from CORDIS on EU IST projects (we hope)

- Similar information from Member States' research funding agencies - EPSRC, etc.

- Contributed information

- etc., etc.

- Indeed anything else that is appropriate and relatively easily obtained and incorporated

Obtaining Information for the RKB

- Via public Web Service interfaces, i.e. data in XML or RDF format
- Through the provision of copies of (large) databases
- Via cooperative web-masters who provide us with means, such as SQL, of extracting all appropriate information from their database
- Via (possibly hidden) pages that contain simply-formatted exports of database tables
- "Scraped" from actual HTML pages - feasible, esp. from pages that were generated from a database, but preferably avoided
- Though the initial concentration is on ResIST members and EU research funding sources, the hope is to expand to cover other sources, throughout Europe - and beyond