# The ReSIST Resilience Knowledge Base

T. Anderson, Z.H. Andrews,
J.S. Fitzgerald, B. Randell
School of Computing Science
Newcastle University
Newcastle upon Tyne, NE1 7RU, UK
{Tom.Anderson, Z.H.Andrews, John.Fitzgerald,
Brian.Randell@ncl.ac.uk}@ncl.ac.uk

H. Glaser, I.C. Millard
School of Electronics and Computer Science
University of Southampton
Southampton, SO17 1BJ, UK
{hg, icm}@ecs.soton.ac.uk

## Abstract

*We describe a prototype knowledge base that uses semantic web technologies to provide a service for querying a large and expanding collection of public data about resilience, dependability and security. We report progress and identify opportunities to support resilience-explicit computing by developing metadata-based descriptions of resilience mechanisms that can be used to support design time and, potentially, run-time decision making.*

## 1. Introduction

Knowledge is the greatest asset that the scientific and engineering community has in facing the challenges of developing resilient systems, particularly those that are distributed, large-scale, heterogeneous and evolving dynamically. Our goal is to develop resources that help the community gain the most value from its knowledge assets. Semantic web technologies offer many benefits over and above the results of applying conventional web-based search engines, enabling rich annotation and the synthesis of data from disparate sources. For example, through the application of ontological mapping, researchers may access information which is relevant to their query terms, but which has originated from sources or been described in vocabularies with which they are not familiar.

The EU Network of Excellence on Resilience for Survivability in IST (ReSIST) aims to promote integration of researchers so that fundamental topics on resilient systems are addressed by a critical mass of co-operative work. Key to this is the provision of a comprehensive and accessible knowledge base to support this effort beyond the life of the network as a formal project. A prototype of such a resilience knowledge base (RKB) has been developed. We here provide an overview of the technologies employed, current status and the potential for exploitation of the RKB in supporting resilience-explicit approaches to system development and operation.

## 2. Resilience Knowledge Base

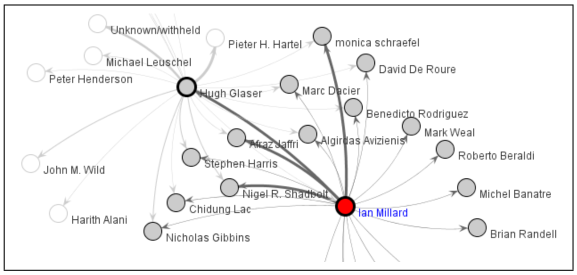
The ReSIST RKB is a collection of facts recorded as semantic RDF statements, describing real resources such as people, projects, publications, courses, and the relationships between them. It is implemented using 3store, an open source repository designed specifically to handle high volumes of RDF, and is an evolution of the 'CS AKTive Space' system that won the 2003 Semantic Web Challenge [4].

### 2.1. Content and Interfaces

The RKB offers a number of interfaces to facilitate data acquisition, querying and visualisation, for both human users and service oriented software processes. These include custom web-based interfaces integrated within a Semantic Wiki, enabling users to record their academic research interests and to upload data regarding training materials in an ontologically mediated fashion.

Data on people, publications and projects is provided by ReSIST participants, in addition to external database and web resources. Over 50 million RDF triples have been collected (as of early 2007) including all public data held by CORDIS, NSF (EU & US research projects), DBLP, Citeseer and the ACM (publications) and from the RISKS forum. Characteristic data on the corpus of papers published at DSN conferences and the (Oakland) symposia on Security & Privacy have recently been added. Resolving coreferences and name variants between different data sources (e.g. *H. Glaser* vs. *Hugh Glaser*) is an acute problem with publications and people, hence we are developing an automatic resolution scheme within the RKB.

A prototype interface has been developed to allow users to navigate and explore this large information repository. The RKB Explorer shows derived connections between related resources. The figure below is a snapshot from a 'People' view, indicating the strengths of Ian Millard's connections to other people by the weight of the connecting lines. These strengths are calculated dynamically by the RKB, taking into account factors such as co-authorship of papers and co-membership of projects. There are also similar RKB views for papers, projects and research topics.




### 2.2. Resilience Ontology

Pre-existing ontologies have been used for general concepts such as projects and publications; an ontology for courseware was created. However, the primary ontology developed specifically for the RKB is on Dependability and Security (D&S) and is derived from the taxonomies of Avizienis et al. [1]. This is represented in OWL and incorporates 166 terms related to Dependability and Security, and 23 to Systems.

Much of the RKB content is derived from published articles, for which keywords are either absent, arbitrary or from existing (and unsuitable) classification schemes. Existing schemes are being mapped into the emerging D&S ontology. For publications without conformant tagging, an interface has been developed to allow domain experts to classify papers manually against the D&S ontology. A current task is to exploit linguistic based term-extraction tools (developed at Saarbrucken) to automate the classification of papers, using the manually classified papers as a training set.

## 3. Support for Resilience-Explicit Computing

*Resilience-Explicit (Res-Ex) Computing* refers to the explicit use of information (metadata) on the resilience characteristics of system components, infrastructure and environment to guide decision-making at either design time or in the running system [2, 3]. Support for such decision making requires access to resilience mechanisms (design patterns, tools, processes) described in terms of their anticipated effects on system resilience. The RKB has a natural role to play here. Machine-interpretation and manipulation of metadata would require that the semantics of various resilience-related concepts be defined within an ontology, to enable a search for information relevant to design (such as research or application reports) and even information to support run-time reconfiguration. For example, a search for a configuration to provide a specified level of service availability (described in terms of metadata) might form part of an off-line or on-line evolution step.

An aim for our current work is to support the description of a wide range of example resilience mechanisms within the RKB, conformant to the relevant ontologies. Competency questions that we expect the Res-Ex enhanced RKB to address for the benefit of researchers have been identified from the point of view of researchers wishing to develop policies and architectures for run-time resilience. We are developing an outline ontology and data acquisition page in the Semantic Wiki to support the addition of resilience mechanism descriptions that explicitly identify metadata. Ten candidate specific resilience mechanisms have been proposed by ReSIST partners, ranging from classical design patterns such as recovery blocks to validation tools such as model checkers and human factors techniques such as dynamic function allocation. We are thus developing new material to add value to the emerging RKB.

## 4. Conclusions and Future Work

The protoype RKB Explorer interface is accessible for use at http://resist.ecs.soton.ac.uk/explorer/

From a resilience-explicit perspective, we hope that proponents of new resilience mechanisms will see benefits in contributing standardised descriptions of them to the RKB, as an aid to research as well as a way of promoting their use.

**Acknowledgements**: We are grateful to our ReSIST (www.resist-noe.org) partners for their contributions, and to the EU for financial support under contract 026764.

## References


[1] A. Avizienis, J.-C. Laprie, B. Randell, and C. Landwehr. Basic concepts and taxonomy of dependable and secure computing. *IEEE Trans. on Dependable and Secure Computing*, 1(1):11–33, 2004.

[2] J. S. Fitzgerald, S. Parastatidis, A. Romanovsky, and P. Watson. Dependability-explicit computing in service-oriented architectures. In *Proc. Int. Conf. on Dependable Systems and Networks (Supp. Vol.)*, pages 34–35, 2004.

[3] M. Kaniche, J.-C. Laprie, and J. P. Blanquart. A frame-work for dependability engineering of critical computing systems. *Safety Science*, 40(9):731–752, 2002.

[4] N. Shadbolt, N. Gibbins, H. Glaser, S. Harris, and m.c. schraefel. CS AKTive Space, or how we learned to stop worrying and love the semantic web. *IEEE Intelligent Systems*, 19(3):41–47, 2004.