

# Summary of the 11<sup>th</sup> Workshop on Requirements Engineering: Foundation for Software Quality (REFSQ'05)

Erik Kamsties, Vincenzo Gervasi, Pete Sawyer

Carneq GmbH, University of Pisa, University of Lancaster  
erik.kamsties@carneq.com, gervasi@di.unipi.it, p.sawyer@lancaster.ac.uk

**Abstract.** The eleventh edition of the International Workshop on Requirements Engineering: Foundation for Software Quality (REFSQ'05) series took place in connection with the International Conference on Advanced Information Systems Engineering in the historic city of Porto, Portugal on the 13th and 14th of June, 2005. The workshop was organised by Erik Kamsties, Vincenzo Gervasi, and Pete Sawyer with Eric Dubois, Andreas Opdahl and Klaus Pohl serving in the REFSQ Advisory Board. This summary gives an overview of the presentations and a summary of the fruitful discussions that took place at REFSQ'05.

## 1. Introduction

Clearly identified more than a quarter of a century ago, the importance of establishing high-quality *requirements* as a prerequisite for designing and building high-quality systems (software or otherwise) has never been contested. In today's interconnected societies, more and more of our everyday life is entrusted to software-intensive systems; the need for high-quality requirements is becoming correspondingly more pervasive. *Requirements engineering* (RE) is the discipline that studies the process of eliciting, negotiating, documenting, verifying, and validating requirements, applying techniques from computer science, psycho-social sciences, economics, and engineering.

The REFSQ workshop series was established in 1994 to foster research in requirements engineering, and over the years has published, to date, 175 papers on the subject – many of them presenting ground-breaking and seminal work – as well as a number of special issues of several journals. The format of the workshop, focusing on interactive, detailed discussions among participants over presented papers, has been highly successful in promoting active involvement and encouraging the emergence of innovative ideas.

REFSQ'05 is the 11<sup>th</sup> edition of the workshop. The call for papers invited submissions on any aspect of RE and its relation to other fields, including:

- Understanding and improving RE processes, RE method engineering
- RE for special needs (web services, mobile appliances, open source development, ubiquitous computing, etc.)

- Social issues in RE
- Innovation-driven RE
- RE economics

Moreover, case studies, experience reports and industrial problem statements were particularly encouraged.

In response to the call for papers, 36 papers were submitted, of which 18 were accepted for discussion at the workshop: 14 as full papers and 4 as short position papers. This represented a healthy interest in REFSQ and a significant increase in number of submissions over recent years. Each submitted paper was reviewed by 3-4 program committee members; most of the reviews were – as per long REFSQ tradition – rather detailed and included many suggestions for improvement. More suggestions and opportunities for improvement of the papers before their final publication came from the plenary discussions at the end of each session, to which two thirds of the time budget of the workshop were reserved.

Twenty-eight people from eleven different countries attended REFSQ. The usual strengths of REFSQ were on show; there was a healthy mix of people ranging from grizzled veterans such as Dan Berry, Sjaak Brinkkemper and Neil Maiden, to students in the initial stages of their doctoral research. Many among the participants were either currently or recently employed in industry or had a foot in both academia and industry. As usual at REFSQ, the workshop benefited greatly from their rooted-ness in real problems. The carefully structured discussion sessions delivered a lively workshop in which everyone participated and from which paper authors gained real value.

## 2. Workshop Structure

We organised the workshop in 4 sessions, which were devoted to *Understanding and Improving the RE Process* (chaired by Ana Moreno, with discussion facilitated by Barbara Paech); *RE in Different Domains* (chaired by Andrea Herrmann, with discussion facilitated by Thomas Alspaugh); *Requirements and Quality* (chaired by Daniel Berry, with discussion facilitated by Pete Sawyer); and *Changes, Dependencies, Composition* (chaired by Neil Maiden, with discussion facilitated by Mike Poppleton).

To ensure the effectiveness of the workshop, each full paper presentation was limited to 15 minutes and followed by 20 minutes of discussion. Furthermore, each paper discussion was initiated by three discussants — usually other paper presenters from the same session. At the end of each session, the major topics raised by the talks or the related discussions were elaborated after introductions by the session discussion facilitators. Presentations of short papers were restricted to 10 minutes with 15 minutes set aside for discussions initiated by two discussants.

Another long-standing REFSQ tradition is the request to presenters to start their talk with a diagram showing the context of their work with respect to a given set of concepts. This set comprised *people* (or stakeholders) which participate in the software development process and *products* (or documents) established by the process. The arrows indicated the relations between the concepts addressed by the paper. While arrows between people indicate some kind of social interaction, an arrow be-

tween documents describes technical and logical relationships. Arrows between people and documents usually indicate production/usage relationships. Figure 1 shows an example of this “first slide”.

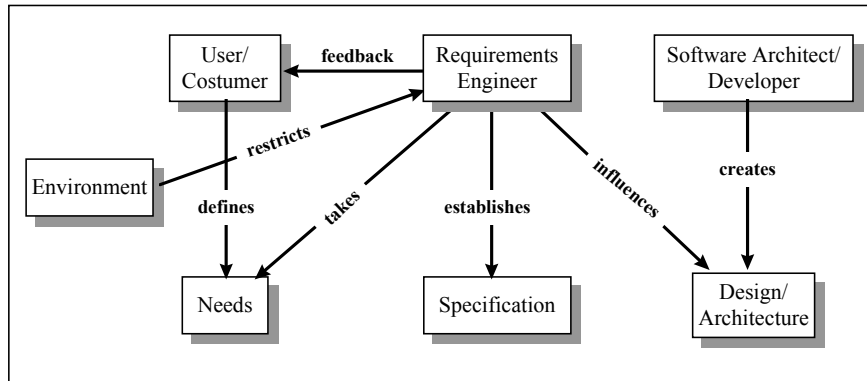


Figure 1: Framework for making the context of each talk explicit.

Additionally, each presenter and each discussant was asked to summarise his or her own views on the talk by answering the following questions:

- Which quality features are addressed by the paper?
- What is the main novelty or contribution of the paper?
- How will this novelty or contribution improve RE practice or RE research?
- What are the main problems with the novelty/contribution and/or with the paper?
- Can the proposed approach be expected to scale to real-life problems?

The workshop was closed by a general discussion, including an evaluation of the workshop itself by the participants.

### 3. Sessions Summaries

#### Session 1: Understanding and Improving the RE Process

Erik Kamsties opened REFSQ’05 by welcoming the attendees. He summarised the purpose of the workshop and the logistics of how it was organised. Following Erik’s brief introduction, Ana Moreno, chair of the first session on “Understanding and Improving the RE Process”, invited **Beatrice Alenljung**<sup>1</sup> to present the opening paper: “**Factors that Affect Requirements Engineers in their Decision Situations: A Case Study**”.

<sup>1</sup> We use **bold face** to denote the name of the author presenting the paper at the workshop.

The general theme of the paper was RE decision-making. It reported a case study of Ericsson Microwave Systems who develop military radar applications. The study revealed several factors that impact on RE practice within the company. These included a mixture of factors that are well-recognised factors by the RE community, such as weaknesses in analysts' training in RE, missing domain knowledge and the problems of workload pressure. The study also revealed factors that are less widely recognised. These included the low status accorded to requirements work and the high cognitive load imposed on analysts caused, at least in part, by weaknesses in RE tools.

The discussants applauded the qualitative approach to assessing RE quality and the insights the study provided into the vulnerabilities of RE within an organisation. They felt that such assessments could be combined with quantitative quality indicators to gain a more complete view of the health of an organisation's RE processes. The study did, of course, concern a single company and it would be interesting to repeat the study across a range of companies in different domains.

In the open discussion that followed, parallels were drawn between RE and testing; another critical activity that has traditionally been accorded low social status. Several people felt that the community needed to raise the profile of RE. However, others in the audience were more up-beat, pointed to some high-profile success stories in increasing the recognition of RE within companies. What was needed to make this more commonplace, it was suggested, was more evangelising by the RE community.

**Tony Gorschek** presented the second paper; "**Assessing the Quality of Requirements Process Changes**". Tony observed that the quality of RE is measured either as part of the process itself against a model such as the CMMI, or measured against the quality of the outputs; the documents, traces, etc. These provide a project-centric view of quality. Tony argued that focusing on projects risked failure to understand the big picture in terms of, for example, the impact on other projects within the company by one project being late or over-budget. As a consequence, RE quality assessment needed instead to be performed on several interconnected levels; the requirements level, project level, product level, company level, and societal level. This is a hard problem, of course, since although we know how to measure the quality of requirements and the impact of RE on projects, measuring company-wide impact was much harder and wider societal impacts harder still. Nevertheless, Tony's point was that the role of the requirements engineer needed to be much wider than that traditionally assumed in order to assure the effectiveness of companies' RE processes.

The discussants agreed that a wider approach to quality assessment was needed but noted that there were several difficulties to be overcome if it was to become practicable. In particular, we don't understand how to handle all the interdependent factors and there is a large time lag (as much as ten years) in some projects between critical RE activities and product delivery. Tony made the point that SMEs generally do product-driven development with much shorter project timescales, so his proposals were feasible for this large sector of the industry.

In the open discussion, tracking defects as they emerged during development was suggested as a means to cope with long-duration projects. Tony countered that this represents a subset of the overall problem, since it addressed problems related to the selected requirements. It couldn't cope with selection of the wrong requirements. The discussion concluded with a break for coffee and with people still discussing how Tony's welcomed ideas might be realised.

Following coffee the session resumed with the paper “**A Communication Protocol for Requirements Engineering Processes**” presented by **Bhavani Palyagar**. This work is based on the hypothesis that that communication problems within a development team negatively influence both RE process capability and compliance with the process. If the capability of an RE process represents its effectiveness and efficiency, then the performance of the process is the product of capability and the effort needed to comply with the process. Bhavani reported empirical evidence from a study of two companies that communication increases the compliance effort. A good example is how marketing estimates based on inadequate consultation with the development experts can lock a company into a project. To help deal with this, the authors have developed a high-level representation of a communication protocol for RE that uses bandwidth as a metaphor for how well communication is facilitated within a process. Bhavani reported that one of the companies used in the case study plans to test pilot the recommendations that have emerged from applying the communication model to their company.

The paper’s discussants welcomed how the author had shown the consequences of poor communication within RE processes. However, they questioned the communication model’s grounding. They wondered why the model borrowed ideas from the data communications area and whether it would cope with highly distributed development teams. It was noted that some of the problems described in the case study were also related to information management. The difference between a “protocol” and a “best practice” was also questioned.

The open discussion session centred on verifying improvements introduced as a result of applying the model, and establishing the cost and benefits of the improvements. This was acknowledged to be a hard problem because few companies keep metrics allowing this to be measured. It was also noted that the communication protocol focused on rights of access to information while, by extending the data communications metaphor, it might also consider (for example) expected or acceptable error rates (as a metaphor for robustness).

The final paper of the session was “**A Goal-Based Round-Trip Method for System Development**” presented by **Gemma Grau**. The starting point for the work was the observation that information systems development usually involves the partial automation of an existing business process. In this respect, it has strong parallels to process engineering. Given this similarity, the authors’ approach to information systems development is to adapt *i\** to the modelling of an existing business process based on the outputs of a RESCUE analysis process. This idea results in a round-trip through five phases: domain information gathering (using RESCUE), *i\** specification of the as-is system, systematic search for process alternatives, evaluation of the modelled alternatives, and prescriptive specification of the system-to-be. The advantage is that the five steps provide a well-defined method for system development, yet all use existing RE techniques. The limitations are that the approach is only feasible with tool support and anyway it is limited to reengineering contexts.

The paper’s discussants particularly liked the method’s explicit exploration of alternative solutions and would have liked more information on this step. There was some concern about the usability of *i\** and the resultant complexity of the method. This general theme was taken up in the open discussion where it was suggested that

the method needed some means to reduce or modularize complexity; it was noted that it is possible to collapse or expand  $i^*$  models. Still on the theme of  $i^*$ 's complexity, participants worried that even experienced analysts would inevitably produce different models of the same problem. Gemma countered with the belief that the as-is model is prescriptive because of the nature of applying the RESCUE method. The intentional models will have less commonality, though. Her co-author Neil Maiden defended this perceived weakness with the observation that it is possible to factor out much of the variability by providing adequate guidelines.

Following conclusion of the last paper's discussion, Barbara Paech opened the session discussion by summarising the main themes that had emerged from the presentations as:

- Value and scope of RE
- Measuring quality of RE
- Contexts
- Modelling
- Tool support

This nicely kick-started discussion as people resumed the theme of RE's value. A positive spin was put on this problem by characterising it as an opportunity for us to prove its economic benefits. One way to do this is by stealth: to focus on the things that are clearly acceptable to industry but which serve the wider purpose of RE; things like doing cost estimation, writing test cases, etc. The issue of tool support also provoked discussion. It was noted that Rational has now integrated its toolset with Eclipse. This presents the RE community with a real opportunity to integrate research results with industry-strength tools. This is important because it makes it far less costly for companies to evaluate and adopt new tools. Barbara concluded the session by noting that we need to work better as a community and build on each others' work if we are to increase the impact of RE research on industrial practice.

## Session 2: RE in Different Domains

Following lunch, the second session was chaired by Andrea Herrmann with the focus on "RE in Different Domains". The session was opened by **Andre Rifaut** with the paper "**Towards a Risk-based Security Requirements Engineering Framework**". The presentation started by focusing on the definition of *Operational risk*; that is the risk of losses deriving from system malfunctions or external events. Risk-based requirements engineering studies requirements as precautions that have to be taken to avert the risk of losing or damaging valuable assets (both physical and immaterial). For example, if proper authority is considered an asset, a security requirement is a way of preventing damage to this asset (e.g., as in an intrusion attack). A Risk is modelled as a combination of Threat, Vulnerability and Impact; all these factors have to be considered in RE: Threat and Vulnerability to know what to avoid, Impact to properly prioritize requirements.

The authors proposed to integrate known techniques (e.g.,  $i^*$ ) with risk analysis during both early and late RE, and use the results to drive architectural engineering. Four steps are cyclically executed in risk analysis:

1. asset identification

2. security goal determination
3. security requirements elicitation
4. countermeasures selection

The reported research proposes to develop tool support for these activities. An example system for administrative support to physicians was used to illustrate the technique. From an i\* model of the system, business assets are identified, and the model is then refined through goal/task decomposition. During the process, security attributes (e.g., system security, privacy, etc.) attached to the goals are propagated through a decomposition structure. A tool could then be capable of generating architecture proposals from such models: for example, it could specify which tasks should be conducted behind a firewall and which are not as critical, or which functions should be restricted by a user login, which ones need privileged access, etc.

The proposed framework builds on results from various fields: security, RE, RE and AE co-engineering. It provides benefits in that early engineering of security requirements can avoid costly and pervasive changes in the architecture of the designed system later on during development.

The discussants generally liked the original point of view taken by the authors; even if some of the problems addressed by the paper are rather old, this different perspective offers an opportunity to have a “new look” at these problems. Scalability seems to be a problem for non-toy projects, though, especially due to the lack of tool support. It was also pointed out that the approach does not consider the problem of business-IT alignment. The presenter, however, maintained that on the contrary the proposed approach materialized the alignment in the selection of assets and risks, many of which are business assets and business risks. Actually, many large businesses (e.g., banks) are more accustomed to think in terms of risks than of functions and even goals. There were reports from the audience about experiences with known risk management methods from the business literature, and about their applicability to the RE context. As even getting less-than-optimal ROI is considered a risk in such an environment, efficiency of investment in security requirements is automatically taken into account (e.g., cost of averting the risk of breakage vs. future losses deriving from a possible breakage). It was also discussed that mitigation can take different forms: one could decide to take an insurance against a given risk rather than trying to prevent it by adding more security requirements. The need for a classification or taxonomy of assets, risks and mitigation measures surfaced during the discussion; the authors declared that they are working on that.

The second paper, “**How Service-Centric Systems Change the Requirements Process**” was presented by **Neil Maiden**. The subject of the paper is the emerging world of service-centric systems, e.g. web services. Such systems integrate independent web and software services over a network through a set of well-defined interfaces, and in turn offer a well-defined interface to potential users. In such context, the roles of end-users and developers blend into each other: there is no concept of a “user interface” which is final and cannot be used to build further services. This is a huge departure from traditional RE research: neither RUP nor most RE papers recognize this situation.

The EU Integrated Project *SeCSE* (<http://secse.eng.it>) was launched to address this situation. The project has four areas:

1. *service engineering*

2. *service discovery*
3. *systems engineering*
4. *service delivery*

The authors identified the following changes to a traditional RE process induced by the service-based context:

1. the process is iterative and informed
  - a. derivation of queries to a discovery service from user requirements
  - b. design process using retrieved services
2. use of divergence strategies to invent requirements
  - a. analogical matching for creative thinking
  - b. constraint removal to challenge system scope
  - c. random matching to provoke new ideas
3. use of convergence strategies to refine requirements
  - a. decomposing requirements
  - b. structuring requirements for service monitoring

The authors have developed a web-based tool for doing the querying and composing resulting services. The tool uses Wordnet to expand the query and match the resulting set of keywords to an infrastructure-provided set of service descriptions. Service registration and discovery is performed based on UDDI (with SeCSE-specific extensions), and results are presented to the user for selection and composition in the web-based interface. The project is still in its initial stages (being 9 months into a 4-year project), but the authors expect significant breakthroughs to be obtained in the future, improving practice and placing new problems into the research agenda.

The discussants highlighted that the endeavour is addressing an old problem – namely, reusability and COTS selection. However, the importance of reuse cannot be underestimated, and a fresh look at the various problems in this area is welcome. It was also suggested that the question of which business model is appropriate for selling and advertising web services should be addressed in the approach. Also, while ontologies might look like the best solution, structured natural language seems to be the most probable candidate for providing service descriptions. In any case, commercial issues should be considered (e.g., advertising, pollution of requirements due to competitive pressure, availability of commercial indexing services, etc.).

Weaknesses of UDDI and various ontologies emerged as issues during the discussion. Wordnet and natural language seem to provide a better match, while simpler to use and to implement than many ontology languages. The issue of description pollution was acknowledged to be important and – as yet – not addressed in the project. Suggestions were made that a provenance-based policy could be used to determine the trustworthiness of a description. Also, information brokers could be established and serve the need. Relationships between the presented project and COTS research were also discussed at length, and similarities and differences between the two contexts were highlighted.

The position paper by **Dan Berry**, Betty Cheng and Ji Zhang titled “**The Four levels of Requirements Engineering for and in Dynamic Adaptive Systems**” analyzed the peculiarities of RE in Dynamic Adaptive Systems (DASs), i.e. systems whose function can change depending on “programs” (that could also be considered as a very rich configuration). In such systems, four levels of RE are recognizable, revolving around the functions of the base system, of the adaptation, of the adapted sys-



tem, and of the adaptation mechanism in itself. Each level, in increasing order of *metaness*, makes decisions about the lower levels. The level structure does not indicate a temporal ordering – actually, co-evolution occurs quite often. The authors’ position is that knowing at which level each requirement has to be placed actually simplifies development, and helps in assigning responsibilities. It also improves our understanding of RE in general.

The discussants found the contribution of the paper very interesting, and its explicit treatment of knowledge that is often left implicit valuable. The lack of operational guidance can be seen as a shortcoming, but since the paper tries to provide an answer to a knowledge problem, this is not a major shortcoming; hence, the paper was considered a promising start for a new line of research. The extent of adaptability addressed by the approach was also discussed by the audience. In particular, the example used in the paper (an adaptive mail system) only adapted to the needs of its potentially impaired users. Other forms of adaptability were exemplified.

The following presentation of the paper “**Selling Web Services as Capabilities**”, given by **Anju Jha**, returned to the theme of service-oriented development, with a particular emphasis on web services. Countering the argument of Neil Maiden and colleagues, the authors judged that NL descriptions of web services are *not* enough to allow effective querying and retrieval of services. Instead, they concentrated on matching “value propositions” to “web service value propositions”, using a well-known internet bookshop web service as an example. Web service value propositions are declarations of what a web service offers, expressed as a progression of problem frames in Jackson’s tradition. Services needed by associates to accomplish *their* value proposition are matched to those offered by Amazon.com. In this sense, the matching provides to both sides of the deal an answer to the question *why are we using/offering this service?*, in that the selection of web service is tied to business objectives of the service provider *and* customer.

Discussants frankly criticized the paper as ignoring much research on the *why* question, and as misrepresenting several of the cited papers. Also, the application of problem frames to describe web services was seen as questionable. In particular, the problem frames notation was considered as more cumbersome for the particular purposes than other competing notations, e.g. *i\**; moreover, the lack of analytic capability was considered a serious weakness for a paper that proposes to solve concrete problems. Finally, while the attempt to connect business value and service specifications was considered as interesting and worthwhile, the particular approach chosen seemed very difficult to sell to the business side of the deal. On the other hand, since problem frames come with a good notion of abstraction and refinement, the approach appears to scale convincingly, and very abstract problem frames could be discussed with the business side with a little extra effort.

The presenter acknowledged that the work is at an initial stage, and many issues still need to be explored. Notational issues in the end boil down to personal preferences – Anju found problem frames much easier and more intuitive than *i\**. It was observed, though, that people in the business domain are accustomed to decomposition trees (e.g., organizational structure diagrams), as in *i\**, and of course to natural language descriptions, but not to problem frames notation, so beyond personal preferences there is an advantage in reusing notations that are analogous to well-known ones.

The last presentation of the day was given by **Colin Snook**, on the paper entitled “**The Engineering of Generic Requirements for Failure Management**”. The research reported studied the formal modelling and analysis of failure management requirements of a product line of engine controllers expressed in B, as part of the *Rodin* project. Failure management subsystems are those parts of reactive systems that handle the failure of sensors or actuators, and take appropriate initiatives like activating a backup device, ignoring data from the faulty component, or degrading the service offered to stay in a safe range.

Such requirements are addressed by performing a domain analysis (building a domain ontology, identifying relationships between the elements, and writing a generic domain specification in rigorous natural language), followed by a generalization/specialization step. The taxonomy obtained for the failure management domain identified six types of functional requirements; these were then modelled in a first-cut UML class diagram. This model was then revised by translating it to U2B, a UML profile tailored for carrying B annotations. From the U2B model, a B specification was automatically generated and imported into the B toolset. Here, violations of invariants were checked, and a validated generic model of the domain was obtained. The generic model can finally be instantiated down to specific cases (i.e., to a specific set of failure management requirements), and validated in the same way.

This research, too, is in its early stages, yet the authors believe that the coupling of formal techniques and failure management requirements will yield important benefits, and that the proposed method provides a friendly introduction to formal methods in this specific domain.

Discussants raised questions about how “friendly” a tool that requires knowledge of UML, B *and* sensors and actuator hardware and failure modes can be. This can impede adaptation and adoption in real cases, which is the final test for scalability. Also, the use of model checking techniques seemed to constitute a limit to scalability. Overall, the presented approach appeared reasonable, although restricted to a well-delimited class of problems and to especially trained users. The presenter did not consider the “language barrier” of B to be of particular importance, as after instantiation, much work on the model is done using UML (although annotated), not B. During the general discussion, the issue of finding a meeting point between formal methods and requirements engineering was discussed at some length.

The general session discussion, facilitated by Thomas Alspaugh, opened on the observations that the various papers in the first two sessions addressed different types of RE:

- RE done by requirements engineers (traditional view)
- RE done by other developers and managers (business view)
- RE done by end users (service-oriented view)
- RE done by systems themselves (autonomic view)

So, the problem of defining RE as the common factor between these diverse activities is raised. Different views were expressed about whether RE in an autonomic context is really RE done by the same system that is the subject of RE, i.e. if there is some degree of self-consciousness involved in hypothesizing such a scenario. A simple description of RE as “deciding what the system should do” does not seem adequate in all these contexts. In the business view, there is a kind of percolation of strategic objectives, where strategies and goals are naturally refined as they go down the

organizational hierarchy, so it cannot be assumed that full requirements are already present.

### **Session 3: Requirements and Quality**

The first session of the second day of the workshop was on “Requirements Quality” and chaired by Dan Berry. The session opened with a paper by Hang Pang, **Neil Maiden**, Konstantinos Zachos and Cornelius Ncube. The paper, “**Do Rich Media Scenario Support Requirements Discovery?**”, analyzed the impact of rich media (e.g., pictures, videos and audio captures) on the effectiveness of scenario-based requirements elicitation sessions.

The authors moved from the observation that recognition works better than recall for humans, i.e., identifying valuable pieces in information in a list is easier than figuring out the information from scratch. In the ART-SCENE approach, text-based cues were generated from scenarios, helping participants in an elicitation meeting to identify potential alternative scenarios. In the present work, the technique has been extended by taking into account rich media, and a lab experiment (with students) was conducted to assess the increase in effectiveness, if any. The application used in the experiment was an information display application for bus stops in London. Results were quite encouraging, with an 8-fold increase in the number of alternative unique scenarios obtained by the rich media group vs. the text-only group. While the study is small-scale and not conclusive, it appears that rich media does favourably impact creativity and recollection. Based on these results, the technique was extended (by using video in addition to photographs) and applied to a project involving a major London hospital. The stakeholders participating in the project expressed satisfaction with the technique, and the number of scenarios elicited was actually higher than those recorded in the past with text-only scenarios. However, there were numerous technical problems, and the cost of collecting and processing all video fragments was found to be substantial. Differently from the lab experiment, only weak evidence was found that the addition rich media to scenarios actually improved the elicitation of new scenarios. In post-experiment interviews, domain experts declared that they were of course very familiar with the scenes shown by ART-SCENE, hence the addition of rich media did not contribute significantly to their “mental visualization” of the scenario. As a result, rich media seems to provide significant improvements for inexperienced analysts<sup>2</sup>, but not to domain specialists, hence the technique might be not cost-effective in real applications, which poses a limit to scalability.

The discussants appreciated the conclusions reached by the authors, and upheld those in their comments. The rigour of the various experiments was pointed out as an especially positive feature of the paper, which sets an example for empirical research in RE.

The general discussion that followed suggested that more experiments might be needed to better assess the relative value of different media (pictures vs. audio vs. video). Also, Neil Maiden communicated that experiments using eye-tracking equip-

---

<sup>2</sup> Note, however, that most Londoners do have to catch buses in their everyday life, so the students were inexperienced at analysis, but familiar with the application.

ment were being designed while the workshop was being held, and suggested that artist-drawn animation might be a more cost-effective alternative to video capture. Links between this usage of video footage and ethnographic studies collecting the same kind of video material were discussed. Also, the effect that the observer has on the behaviour of observed subjects entered the discussion, and it was pointed out that new, less invasive technologies (e.g., third generation mobile phones with video capabilities) might help in this respect. Another interesting question that was raised was about the quality and usefulness of the additional scenarios that were obtained through rich media; no special assessment of this feature was conducted, and this might be the subject of further research.

The second paper, by **Asa Dahlstedt**, was entitled “**Guidelines Regarding Requirements Engineering Practices in order to Facilitate System Testing**”. The aim of the paper was to gain an overall view of the needs of system testers on requirements. A literary review and an interview study (based on open-ended interviews with five experienced system testers) were conducted, resulting in a number of guidelines for requirements engineers that aim at facilitating system testing. These included:

- Involve the testers in the requirements analysis work
- Use test design to check the quality of the requirements
- Get the users involved in both RE and testing
- Use scenarios to make the functionality of the software more specific and concrete
- When formulating requirements, make sure that the correct outcome of a certain action is clearly and easily identifiable
- The more precise, complete and measurable the requirements are, the easier it is to identify test cases
- Carefully decide which additional information about the requirements needs to be collected and stored
- Plan carefully for requirements change
- If software is developed during several releases, use bug reports from testing to identify missing and erroneous requirements.

While this advice is certainly not big news, it is comforting to find that these are actually the things emphasized and often missing in practice. The matching of well-known good practice from RE with the needs of system testers indicate that the main problem is in transferring these good practices to RE practitioners and in knowing *how* the advice can be implemented in practice – a critical subject not covered by the paper.

The discussants agreed with these points (especially regarding the weak novelty), but felt the advice provided was rather general and lacked operational guidance. Also, methodological problems were raised: lack of triangulation, small sample size and lack of assessment of how expert the “expert” interviewees actually were, all cast a shadow on the reliability of the results obtained. Another issue that was raised was the lack of distinction between advice for functional and non-functional requirements, while it appears that these should be treated differently in terms of how testable they are and how to test them. The discussants also pointed out that such guidelines should in some way become part of a certification process for professional requirements engineers.

The general discussion that followed focused on how to improve the work. Access to the best experts available in a field was discussed, and the author was encouraged to contact recognized authorities. Also, considering commonalities in different sub-fields (e.g., testing with use cases vs. testing with formal models) was indicated as a potentially interesting venue for further research. The applicability of widely general advice in situated processes was the subject of some disagreement between the participants, and the RE'05 workshop on situated processes was identified as a venue where such issues could be discussed in more depth.

The position paper “**Quality Misuse**” by **Andrea Herrmann** and Barbara Paech, reported on initial work in the application of Misuse Cases to quality attributes other than security/safety. As in the paper by Meyer et al., the view here is that *assets* are to be protected from *threats*, and in so doing *countermeasures* are defined. In contrast with standard literature, each asset is paired here with a quality attribute, and misusers do not necessarily have a misuse goal – e.g., misuse can happen unintentionally as a consequence of other decisions. A meta-model for misuse cases paired with qualities was defined, including quality deficiencies caused by a threat and vulnerabilities deriving from quality deficiencies. A small case study concerning the integrity of a medical record management system was presented, hierarchically refining threats and countermeasures in order to derive requirements for a system of sufficient quality. Future work includes the selection of further case studies, the integration of risks (as probabilities for a misuse case to occur), and conflict identification and negotiation (i.e., when the countermeasure for a threat negatively affects a different desired quality attribute).

The discussants liked the idea of extending misuse cases outside of the domain of security requirements to arbitrary qualities. The contribution may improve the identification and analysis of non-functional requirements using well-established concepts from use cases and (to a less extent) misuse cases. It was also pointed out that even big use-case research efforts (e.g. CREWS) never come to a satisfactory treatment of non-functional requirements, so the idea is certainly worth pursuing. Also, it was suggested that an investigation of recurring patterns in misuse cases might provide valuable insights on the nature of the corresponding qualities. The suggestion was accepted by the author as a subject for future research. The role of the satisfaction and satisficement relationships popularized by other methods in this context was also discussed, and tracing links between misuse cases applied to quality attributes and goal models was recommended.

The paper “**Scenario-driven Specification-based Testing against Goals and Requirements**”, by **Thomas Alspaugh**, Debra Richardson, Thomas Standish and Hadar Ziv, discussed ways to improve the effectiveness of testing by focusing it on the actual stakeholders’ needs, possibly giving less coverage to other features that are not in their immediate interest. Also, this would allow more accurate coverage metrics, giving measures of the *needs* tested rather than the *features* tested, and possibly support the perfecting of requirements by correcting errors uncovered through testing. Moving from the current high cost of testing, the authors proposed to leverage RE work by annotating tests with stakeholders’ goals, and use comparison of goal and event traces to help distinguish implementation and domain modelling errors. Creating test cases from requirements scenarios was proposed as a way to realize

these goals, using ScenarioML, an XML encoding of scenarios by Thomas Alspaugh, as a vehicle for research in this direction.

Discussants pointed out that the approach has potential, and introduces a clear operational way to link in a tight relationship, stakeholders' needs and test cases. In a sense, however, such a tight relationship can be a disadvantage, in that it reduces the degree of *independent* validation of the requirements. Also, scale problems might make the approach unpractical without good tool support, which is not currently available. Testing on an industrial case study should be a priority for the prosecution of this research, as is some evidence of its cost-effectiveness.

The general discussion started with a debate on testing and requirements coverage. The basic way to obtain good coverage at reasonable expense is to collect specific behaviours into equivalence classes, and then testing one representative for each class. The way these equivalence classes map to requirements was analyzed, distinguishing between testing systems developed in a purely feature-composition style (for which a mapping can probably be found) and those developed in more traditional ways (for which a mapping probably does not exist). Compositionality problems arose again in a discussion about highly concurrent systems, where the number of possible scenarios (i.e., linear traces) can be enormous, and difficult to handle as proposed by the paper. This in turn led to the issue of prioritization of goals and tests, which is needed to cope with rich goal hierarchies and a limited budget for testing, but which is not yet handled by the approach. It was suggested that any established prioritization technique could be applied to goals, depending on how the stakeholders value them, and then propagated to test cases.

The general session discussion, facilitated by Pete Sawyer, identified two general themes in the papers presented, namely quality of requirements and quality of RE processes. Most papers elaborated on scenarios and use cases, linking them to coverage of needs (as in Neil Maiden's paper) and of tests (as in Thomas Alspaugh's paper) – and thus, to the general quality attribute of *completeness*. Emerging issues were empirical issues, importance of context (i.e., situated processes), and more reflection on the scope of RE as a discipline.

An ongoing switching of focus from *consistency* to *completeness* was seen as a positive sign of maturation in the discipline. Still, the "good enough" question was not addressed – i.e., taking budget limitations into proper account to decide when to stop requirements engineering as well as testing. The subject has been studied extensively in testing, so reusing or at least considering those techniques for requirements work was pointed out as an important area of research. The relative importance that industry assigns to completeness vs. consistency was also discussed: participants reported that at times industry people *do not want* too many requirements, even when the budget capability would be there, in order to keep the set of requirements manageable (the same was reported as happening in testing, too). So, stop criteria are not necessarily linked to resource constraints. Still, the availability of tools changes the threshold for such discussions. A better understanding of stop criteria would also help in justifying the cost of upfront RE, by placing a cap on the time and financial costs of the activity.

More considerations on the state of the practice in industry followed, focusing in particular on the link between strategic organization objectives and the constraints that are imposed on the RE process as a consequence. All kinds of feedback loops

exist inside organizations between business decisions and technical decisions, and these should be studied carefully in each distinct situation. The role of knowledge, natural language and semi-formal structures built around natural language (e.g. scenarios), was also mentioned.

Following the general session discussion, a special presentation was given by **Ana Moreno** on the paper “**Requirements Engineering Research: A Microcosm of International Economic and Political Trends**” (Eustace Asanghanwa et al.). The authors have studied the statistical relationships between RE research output and a number of social, economical and political factors of different countries, as a contribution towards establishing a statistical truth on the impact that RE practices have on the general economy.

Given the lack of data about the state of RE practice in industry, the authors hypothesized that RE research output may be correlated with the state of practice, so they analyzed a database of 2.645 publications in international forums collected in a 16-year effort by Al Davis, and considered the country of origin of all authors as attributes of the publications. Various large-scale economic indicators (e.g., GDP) were considered for each of the countries with an active RE community. A wealth of statistical data was provided (e.g., distribution by continent), which readers can find listed in detail in the paper published in this volume. Limitations of the research and threats to validity were clearly stated, but the general outlook is believed to be essentially correct. A warning was also stated that statistical correlation and causality are different concepts, hence the results should not be taken to imply a direct effect of RE on economy (the reverse, though, seems very reasonable to assume). Most data was aggregated by continent, which may be not significant for very diverse continents (e.g., Asia, where China and Japan have very different economies), and the authors plan to study the data at a finer granularity as future work.

Discussants appreciated the rigour and extensiveness of the study, although the possible uses of such results were not clear. Also, the significance of some correlation was found to be debatable (e.g., RE has a strong correlation with agriculture decline in Europe!), and this cast a doubt shadow on other results – a doubt which, however, was first disclaimed by the authors themselves. The same conclusions were reinforced during the open discussion, and the similarity between correlations with RE and other type of scientific research were also proposed as a subject worth investigating to put the authors’ results into context. Still, the utility of the data so obtained was unclear. The research method, although rigorous, was criticised as unsuitable, as it compared macro-economic data with micro-phenomena like publishing a RE paper (and for several small countries, their entire RE research output could be credited to a single individual!) . Suggestions for further analysis (e.g., looking for a time lag between economic and research phenomena that might reveal some causality) were also provided, and accepted by the authors as part of future work.

#### **Session 4: Changes, Dependencies, Composition**

The fourth and final session began after lunch and was chaired by Neil Maiden. **Anju Jha** presented the first paper entitled “**Supporting the Consideration of Dependencies in Use Case Specification**”. The paper was about dependencies between events

in use cases, and about symbolic execution tools used as a means for their validation. Use case descriptions normally require statements of their pre- and post-conditions. This is a coarse level of granularity and the work reported focused on allowing the statement of finer-grained dependencies on individual events. State diagrams are the traditional and well-tried means to do this; however, they are hard to explain to most stakeholders and so inhibit their involvement in the derivation and validation of use cases. The paper's authors have instead developed a tool that allows intra-use-case dependencies to be expressed in terms of pre- and post conditions on events. A case study based on a company developing real-time software was presented and this produced evidence that the approach was useful in exposing some significant misunderstandings in use case models of their software.

Discussion of the paper was inhibited by the enforced absence of the main author (John Kanyaru), but discussants were generally supportive of any means to help highlight the dependencies between interconnected events that were intelligible to stakeholders. The tool-first approach was questioned, however. Since UML is extensible, a relatively straightforward extension to the underlying use case notation might have represented a useful first step.

**Andrew Stone** next presented a short paper entitled **“Finding Tacit Knowledge by Solving the Pre-Requirements Tracing Problem”**. This represents his initial ideas on how a sub-class of tacit (i.e., implicit or taken-for-granted) knowledge might be identified by examining the relationships between a specification and the sources from which the specification is presumed to be derived. The absence of any derivation relationship between a requirement and the source material indicates that the requirement potentially represents the concretisation of tacit knowledge, which should then be added to the list of source material. The work was scoped by a dependency on the availability of requirements sources in textual form – interview transcripts, domain documents, etc. This was because the basic enabler for testing Andrew's hypothesis is a range of language engineering tools; particularly Latent Semantic Analysis (LSA) as a means to detect transitive lexical relationships.

The following discussion raised two critical issues. The first was the general problem of achieving completeness. Even in the (rather unlikely) event of all the source material being available in textual form, the degree of recall and precision achievable would be relatively low. The second was a warning against the assumption that specifications are produced in a transformational process. Rather, requirements are invented in a creative way so direct relationships between requirements and source material may not exist even where tacit knowledge isn't involved. Language engineering experts in the audience also observed that the computational expense involved was likely to inhibit scalability. Nevertheless, the investigation of tacit knowledge, despite the probable limitations, was thought worth-while.

The next paper on **“Flexible Release Composition using Integer Linear Programming”** was presented by **Johan Versendaal** (co-authors J.M. van den Akker, S. Brinkkemper, G. Diepen). This paper reported work addressing the needs of a product manager planning the up-coming release of a software product. Given a large set of requirements, the product manager needs to scope them to select the optimal subset for implementation in the release. This poses many problems; there are typically many requirements to choose from, they are hard to prioritise and they are sourced from different stakeholders. Johan proposed Integer Linear Programming



(ILP) as a technique for managing the problem and presented a tool, validated on a set of real requirements for an ERP system that implements ILP. Given a set of requirements with known attributes (such as cost estimates) this allows a product manager to parameterise the ILP model with a number of important *managerial steering mechanisms* (such as deadline shifts, the time needed for a new team member to become productive, etc.) to discover the optimal set of requirements while concealing the complexities of the underlying model.

Johan acknowledged the need for further research on the model's ability to handle (for example) combinations of steering mechanisms, requirements interdependencies, etc. He observed that the real novelty of the work lay in its combination of existing techniques to provide utility for a real RE problem.

The paper's discussants applauded the work's contribution to fine-grained flexible release planning. They did, however, question the technique's vulnerability to the quality of the data. A key requirements attribute, for example, is revenue-earning potential yet the estimation of this is enormously hard. This particular problem is being examined by a concerted effort amongst software product development organisations in the Netherlands, so the authors are confident that it will not prove a show-stopper. In the open discussion session it was noted that the authors will face the problem of demonstrating that their technique is better than competing ones. However, it was suggested that this was solvable by separating validation of the model from that of the optimization process. Another observation was that requirements that have the potential to be high revenue-earners are often risky to commit to, and therefore that risk needed to be factored into the optimization algorithm.

**John Brier** presented REFSQ'05's final paper on "**Computing Change in Socio-Technical Systems with Problem Frames**". This paper proposed the use of problem frames for capturing patterns of change in socio-technical systems. This was novel since problem frames were conceived for systems where there is an explicit interface between the "machine" and its environment. Such a clean interface is mostly absent in socio-technical systems. John illustrated the proposed adaptation using a simple transaction completion scenario involving users, bank employees and an intermediary. He illustrated this with problem frames analysis diagrams adapted so that one of the key advantages was the explicit illustration of the impact of change to a business process in a single diagram.

The paper's discussants and the wider audience focused on scalability. It was noted that there are only five known problem frames patterns (with a sixth proposed at REFSQ'03). It was suggested that this fact alone will inhibit scalability for large, typically messy socio-technical systems. A counter argument was that a small number of patterns was actually a good thing in usability terms. There were some questions regarding the utility of modelling change in a single diagram and this led to further discussion on whether problem frames are truly compositional. One participant contrasted problem frames unfavourably with  $i^*$  which he felt was better suited to the task yet, like problem frames, also attracts criticism for the apparent complexity of the notation. The discussion concluded with a challenge to model an air traffic control system, successfully modelled in  $i^*$ , using the proposed problem frames-based technique: a kind of RE *High Noon*. Keep your heads down!

Mike Poppleton, the session discussion facilitator, set the scene for the final discussion by suggesting that the title of the session should have been "Requirements

knowledge, planning and change”. In the subsequent discussion, Mike’s Requirements Knowledge sub-theme was picked up. It was noted that there are many different types of tacit knowledge. One example was learning to tie your shoe laces. Once learned, it becomes tacit, yet it is enormously hard to describe in words how it is done to someone else. However, seeing a video of it *would* allow it to be made explicit. As a result of this, it was suggested it would be useful to look at a range of mechanisms for stimulating recall and identification of the tacit.

The workshop concluded with a lively discussion of the nature and value of research driven by existing industry priorities versus more speculative research. Ten out of the workshop’s eighteen papers were about capturing, analysing and modelling requirements. Doubt was expressed that this reflected the real priorities of the industry since in industry (or at least some sectors of the industry) these just weren’t considered problems. Testing against requirements, by contrast was posited as an example of a more valuable line of research. It was acknowledged that the community could do more about identifying real problems, but we also needed to be innovative rather than responding only to currently recognised problems. Modelling was offered as an example. Despite industry’s initial failure to recognise its utility, modelling has been proven to hold great general value for understanding. It was also noted that industry is heterogeneous and the problems they face are similarly heterogeneous. Limiting the foci for RE research would potentially impede our ability to address the wide range of problems faced now, and so far unanticipated problems in the longer term.

#### 4. Workshop Conclusion

As in its long tradition, REFSQ generated many discussions and good feedback. Three interesting conclusions can be drawn from this year’s papers. First, we observed a **comeback of quality features**, in particular product (requirements) quality (e.g., completeness, testability, security) and process quality, (e.g., communication protocol quality). Second, **new approaches to known problems** (e.g., eliciting tacit knowledge, release planning, risks and testing) were shown. Finally, we became aware of **fresh problems for RE research** (e.g., service-oriented systems, multiple level and meta-RE, impact of RE on business and society, wider role of RE).

A general wrap-up discussion was held at the end of the workshop, during which suggestions were sought for interesting themes to propose for REFSQ in 2006. The following themes were suggested by the audience:

- Estimation in RE
  - Artifacts
  - Processes
  - Maturity level models for organizations
- Best practices for specific industries and applications
  - Domain-specific RE
  - What is domain-dependent and what is domain-independent in RE?
  - Requirements in ERP: why are they failing?
  - Empirical validation of best practices

- Identification of worst practices?
- Technology transfer
- Ambient intelligence
- RE for autonomic systems
- RE and outsourcing
  - Is RE so critical that it cannot be outsourced?
  - Might outsourcing stimulate the accumulation of RE knowledge?
- Encoding and analysis of RE knowledge
  - Pooling of resources (see also above)
- Compositionality in RE
  - Problem frames, viewpoints, etc.
- The state of practice of RE in industry
- Multi-project RE
  - Independent development of (loosely coupled) subsystems
- ROI of RE

In time-honoured fashion, the co-chairs asked for feedback on the workshop and a number of useful perceptive and useful ideas were suggested. A proposal for the replacement of the ubiquitous first slide mandated for each presentation was presented; a source of much confusion. With this year's presenters' cooperation, the proposal will be prototyped in advance of REFSQ'06.

The 11<sup>th</sup> edition of the workshop was closed with a warm thank from the co-chairs to all participants for a very successful REFSQ'05; in turn, participants generously thanked the co-chairs for their work. Particularly enthusiastic thanks were given to Erik Kamsties for all his hard work and leadership in the last three years, since REFSQ'05 marks his retirements from the group of co-chairs.

## **Acknowledgements**

First of all, we want to thank Eric Dubois, Andreas Opdahl and Klaus Pohl who served on the REFSQ Advisory Board.

We would also like to thank the members of the program committee who acted as anonymous reviewers and provided valuable feedback to the authors: Thomas Alspaugh, Aybüke Aurum, Daniel Berry, Sjaak Brinkkemper, Pierre Jean Charrel, Karl Cox, Alan Davis, Eric Dubois, Peter Haumer, Patrick Heymans, Sara Jones, Giuseppe Lami, Soren Lauesen, Peri Loucopoulos, Kalle Lyytinen, Johan Natt och Dag, Andreas Opdahl, Barbara Paech, Anne Persson, Klaus Pohl, Norah Power, Naveen Prakash, Jolita Ralyté, Björn Regnell, Colette Rolland, Matti Rossi, Camille Salinesi, Kristian Sandahl, Andres Silva, Guttorm Sindre, Michel Wermelinger, Mario Winter, Claes Wohlin, Didar Zowghi.

---

*The proceedings from REFSQ'05, as well as those from previous editions of the workshop, can be ordered from REFSQ <refsq-proc@sse.uni-essen.de>.*

*Information on REFSQ'06 and past REFSQs is available at  
<http://www.refsq.org>*