GQR Model to Support Requirements Elicitation for Content Rich Systems

Leah Goldin
Afeka Tel-Aviv Academic College of Engineering, Tel Aviv, Israel

Keywords: Requirements Elicitation, Goal-Oriented Analysis, GQR Model, Content Rich Systems, GQM (Goal-Question-Metric), Biomedical, Requirements Management, CMMI.

Abstract: Requirements elicitation for content rich systems is a huge challenge. Even using known techniques, does not avoid diving into content details and get lost in hyperspace. We propose the GQR (Goal-Question-Result) model to support the Requirements Elicitation process in elevating the elicitation discussions from specific data elements to contextual structures of goal-question-result related content requirements, while scoping the discussion to specific stakeholders. This paper characterizes the GQR model, shows its advantages relative to the previous methods and illustrate its application by means of two case studies, one biomedical and another related to CMMI requirements processes.

1 INTRODUCTION

The purpose of requirements elicitation is initial extraction of the requirements. What methods are best used for discovering and gathering requirements, and how can we encourage customers to better express their needs, are still a great challenge to both Requirements Engineering (RE) practitioners and researchers (Christel and Kang, 1992).

Requirements elicitation for content rich systems is a further challenge. Even using known techniques, does not avoid diving into content details and get lost in hyperspace. While stakeholders' identification (Robertson and Robertson, 1999) is nowadays practical and achievable, when inquiring the requirements stakeholders to state their need and requirements for developing a content rich system, it is often impossible for them to envision their requirements for the new system.

We propose the GQR (Goal-Question-Result) model to support the higher demands of rich content systems' Requirements Elicitation process. The latter deals with contextual structures of goal-question-result related content requirements, while scoping the discussion to specific stakeholders. This paper characterizes the GQR model through its core concepts, and illustrates its application by means of two case studies, one biomedical and another referring to CMMI requirements processes (CMMI, 2005).

1.1 Large Scale Information Systems Requirements

Large scale Information Systems entail loads of data at different levels of details and abstraction. Sometimes the content is generated via different interests not always thinking about who will use it eventually. Thus requirements elicitation for such content rich system is a significant challenge.

1.2 Requirements Elicitation

The work on Issues in Requirements Elicitation (Neetu and Pillai, 2013) and (Christel and Kang, 1992) reported by the Software Engineering Institute (SEI), surveys the problems identified in implementing requirements elicitation processes, and suggest a requirements elicitation framework to cope with these inherent problems. Elicitation problems are classified to classes of scope, understanding and volatility. Problems of scope arise from ill-defined boundary of the system along with unnecessary incorporation of design information into requirements. Problems of scope result from lack of:

1. *Understanding the Organization* – in which the system under development will be placed, i.e.,

Goldin, L..
GQR Model to Support Requirements Elicitation for Content Rich Systems.
In *Proceedings of the 6th International Workshop on Software Knowledge (SKY 2015)*, pages 46-52
ISBN: 978-989-758-162-5

submitters of input to the target system and users of the target's system output, and

2. *Understanding the System's Mission* – within the organization, i.e., ways in which the target system will adapt/change the organization's means of doing business.

1.3 Goal-Oriented Requirements Methods

Goal Oriented analysis methods like KAOS, Tropos, i*, GBRAM (Bertrand et al., 1998), (Christel and Kang, 1992), (Anton, 1996), (Lapouchnian, 2005) and (Kavakli and Loucopoulos, 2003), include concepts like: actions; entities; agents; goals and constraints. These methods are formal and require from the analyst high abstraction skills at different levels of formal modeling. The current goal oriented analysis method expect from the requirement engineer to be creative and actually invent concepts that are not mentioned in the request for proposal (RFP) or other customers request for product (Zdravkovic et al., 2013).

1.4 Related Work

Current requirements elicitation methods in information systems (Neetu and Pillai, 2013) and (Christel and Kang, 1992) include:

- interviews as basic activity for discussing requirements with customers,
- using forms and questionnaires,
  - analyzing requirements via Business Process Modeling (BPM), Data Flow Diagram (DFD), and Entity Relation Diagrams (ERD) (Biedermann and Grierson, 1995) and (Chen, 1966) etc.

1.5 Paper Organization

The remaining of the paper is organized as follows. In section 2 we characterize the new GQR model. In section 3 we give the biomedical case study as an illustration of the application of the GQR model. In section 4 we provide the CMMI requirements' process areas case study as another illustration of the GQR model. In section 5 we evaluate the GQR model with respect to Requirements Elicitation. The paper is concluded with a Discussion in section 6.

2 THE GQR MODEL FOR REQUIREMENTS ELICITATION

The Goal-Question-Result (GQR) was inspired and based upon the Goal Oriented requirements elicitation and Goal-Question-Metric (GQM) methods (Solingen and Berghout, 1999). The contextual structures of goal-question-result sets the scene for the requirements elicitation process with rational and operational context related to the expected content requirements, while scoping the discussion to specific stakeholders.

2.1 GQR Model

We have defined a high level model, shown in Figure 1, enabling simple but effective structuring of content investigations in terms of three primary concepts – Goals, Questions and Results (GQR). The model is grounded by goal oriented requirements engineering principles (Dardenne et al., 1993) and is inspired by the Goal-Question-Metric (GQM) method (Solingen and Berghout, 1999). The traditional GQM method used in software measurements enables a planned collection of data along with a systematic approach of analyzing the metrics in order to get an effective managerial feedback.

The Goal Question Result (GQR) method rationale is to:

1. define your research ***goal,***
2. state the research ***questions*** that are asked in order to check if the goal is met, and
3. identify the ***Results*** that will be used to answer these questions.

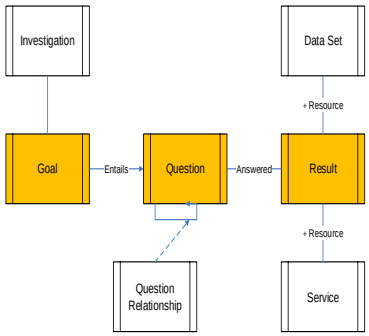

Figure 1: ***Core concepts in the GQR method***: Investigation represent an elicitation inquiry that conveys Goals, a Goal entails Questions that are answered by Results. Results resource may be a data set or a service that convey the content requirements.

A ***Goal*** represents the investigation's objective. Goals can be more or less specific like "the role of diet in cancer" or "investigate whether a disease responds to a drug". A goal entails one or more questions that must be answered to achieve its fulfilment. ***Questions*** are answered by way of data sets or services (which in turn may consume data sets) building up the actual ***Results*** that the analyst was looking for. The goals, questions and results are extracted from the customer request descriptions, and are structured via GQR. This GQR model helps scoping the content from which the requirements will be further elicited consistently via every GQR relation to Data-Set or Service of the content rich system.

Due to its simplicity (natural/cognitive presentation), people without a software modelling skills can easily connect to the model and act upon it. The GQR concepts appear in the customer requests or user cases, as oppose to KAOS or i* that require the invention of abstract concepts that are not described in customer requests.

2.2 Potential Stakeholders

The requirements stakeholders are the ones that have interest in the system, and can gain or lose something as a result of this project. This interest includes: functionality, revenue, status, compliance with rules, etc. (Robertson and Robertson, 1999).

The Questions in the GQR model are related to specific stakeholders which envision their operational usage of the system to be developed. Thus GQR is based on stakeholders identification, i.e., Researcher, Data Manager, Biotechnician for the PRM case study1 in Section 3, or the organization roles like Product Manager, System Engineer, Project Manager, Tester for the RMR case study 2 in Section 4.

2.3 Existing Resources

Once the results are defined in the GQR model, they are related to the resources of existing data from which the content is retrieved. Existing data must include documentation that stakeholders are using, knowledge repositories or services.

3 CASE STUDY 1: CANCER RESEARCH PLATFORM REFERENCE MODEL

The NCRI Informatics Initiative (NCRI, 2015) has set the goal to increase the impact of UK cancer research and improve prevention and treatment of cancer by effective use of informatics to manage and exploit the diverse types of information currently generated via an integrative platform (Begent et al., 2005). The main aim is to enable the creation of an open community where the different informatics tools and resources available in the UK and worldwide can interoperate with one another as a whole.

Data sets have been generated by different research groups around the world working across the cancer research spectrum, that is, from basic to clinical cancer research. Many of these data sets have evolved separately and so present an incoherent, fragmented landscape.

As an initial step in the platform development, the author participated in a project focusing on the requirements analysis and modeling of the Platform Reference Model (PRM) (Finkelstein et al., 2006). The analysis has been driven by a set of use cases acquired by interviewing practitioners working in the field. The main goals of the analysis have been to understand how the various initiatives operating in the cancer research field would effectively cooperate with one another, what the relative roles are, how they are actually used by practitioners. This required a higher level perspective in defining the use cases which are closer to user stories as defined in the Agile development (Cohn, 2004) than to standard use cases as described in (Cockburn, 2001).

3.1 The PRM Stakeholders

The PRM stakeholders identified by the NCRI Unit are people involved directly or indirectly in the research itself, i.e., scientific and clinical researchers, clinicians and patients. Scientific researcher include: Biologists, Chemists, Physicists, Statisticians Bioinformatics experts, and Computer Scientists, using the Platform to support their experiments or studies. Clinical Researchers include: Radiologists, Pathologists, Oncologists, using the PRM to support their clinical trials.

3.2 The PRM Resources

The context where the NCRI platform will operate is made up of a multitude of projects, resources and initiatives that aim to support cancer research in different ways. The field of cancer research is highly dynamic reflecting the emergence of new technologies, developing infrastructure and scientific discovery.

The systems that interface with the PRM can be other systems such as informatics repositories. The biomedical Terminologies and Ontologies include all the standard vocabularies needed to access the various informatics repositories, i.e., NCI (NCI, 2015) Thesaurus. Biomedical Informatics Resources include the existing informatics systems that contain raw data from different research areas, i.e., EBI databases such as, ArrayExpress (EBI, 2015) for micro-array data. Biomedical Publications include approved and validated research results published in medical journals, i.e., PubMed website (NCBI, 2015), PLOS (the Public Library of Science) (PLOS, 2015), PharmaKGB (an integrated resource about how variation in human genes leads to variation in our response to drugs) (PharmaKGB, 2015) etc.

3.3 The GQR Model for the PRM

The ever-increasing amounts of cancer research data are collected and recorded in non-standardized ways and are not in a suitable form for sharing, re-use and integration. Thus, opportunities to gain new knowledge are lost, results are not translated for clinical use and experiments are repeated wastefully. Planning cancer research via the GQR method enables a planned collection of bioinformatics data along with a systematic approach of analyzing the results answers to the research questions, in order to get an effective research result (wrt research goal).

The GQR analysis of the "Genetic Variation in Response to chemotherapy" case study in figure 2 includes the identification of research goal and questions, along with the required Results mapped to the existing Bioinformatics repositories.

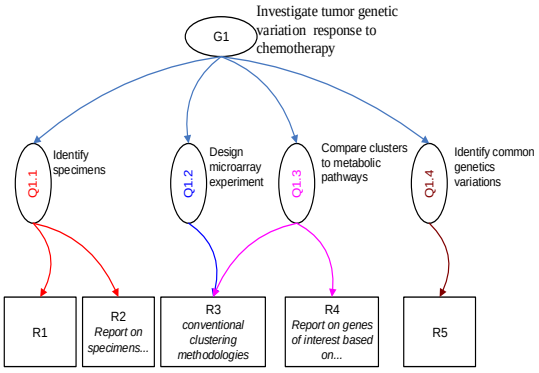

Figure 2: ***GQR analysis of the PRM***. For the research Goal of investigating the genetic response to chemotherapy, Question Q1.1 rose by the Pathologist regarding specimens' identification while Question Q1.2 rose by the Radiologist regarding microarray experiments (data and protocol). Each Question relays to a special repository of relevant content Results.

This case study analysis obtained 1 research goal, 4 research questions and 5 Results as described in Figure 2:

G1→ {Q1.1, Q1.2, Q1.3, Q1.4}
Q1→ {R1, R2} | Q2 →{R3} | Q3→ {R3, R4} | Q4 → {R5}

As part of the GQR analysis of the "Genetic Variation in Response to chemotherapy" case study, we have accompanied each dataset or Result required to answer a research questions with reference to existing bioinformatics repositories. For example, European Bioinformatics Institute (EBI) /Arrayexpress (EBI, 2015) required for microarray analysis of R3, see table 1.

Table 1: GQR Reference to Bioinformatics Repositories.

Result	Data asset (data set / service)	Bioinformatics repositories
R3	Analyze [and store] microarray data using conventional clustering methodologies	R3 ← Journals(JCO,Nature) ← clustering tools → EBI/ArrayExpress → Pub-Med → GEO → caArray
R4	Report on genes of interest based on expression in samples, and presence in metabolic pathways of interest, which are known to be chemotherapy targets	R4 ← REACTOME ← KEGG ← Journals ← Pub-Med

4 CASE STUDY 2: REQUIREMENTS MANAGEMENT REPOSITORY

As part of the CMMI (CMMI, 2015) initiative, more and more organizations decide to manage their requirements via a requirements management (RM) tool, thus establishing an RM Repository (RMR). Most of the RM tools are based on Databases and provide features of traceability required by the CMMI, and it is a major technology transfer to implement such a tool.

In order to benefit from the RMR, the RM process should be defined and tailored accordingly into the RM tool. An organization that wishes to use an RMR needs to change its culture of work, and it is not simple to elicit from people their RMR requirements and needs for a new technology that they have not experienced yet. To that end the author has been using the GQR method in order to elicit the RMR requirements and needs, while focusing on project management goals that can be answered by RMR reports.

4.1 The RMR Stakeholders

The RMR stakeholders are the usual project

stakeholders, i.e., Project Manager, Product Manager, System Engineer, Test Manager, Software Engineer, etc. In this RMR case study, Pre-sales Engineer, Proposal Manager, and Product Managers played a major role in using the RMR for answering RFPs (Request for Proposal).

4.2 The RMR Resources

The RMR resources are the usual project documentation including: PRD (Product Requirements Document) defining the product features as from the R&D group to develop , System and Subsystem Specifications (SSS), Software Requirements Specifications (SRS) (DOD-Mil-std 498, 2000), Request for Proposal (RFP) in cases that the customers define their needs specifically, etc. Other project documentation items like Software Test Plan (STP), User Manuals, Product Release Notes, etc., can also be included in the RMR depending on the project goal as will be described in the following.

In this RMR case study, the project is based upon requirements information found in written documents such as product boilerplates, product requirements definition (PRD) and product bulletins.

4.3 The GQR Model for the RMR

This case study presents the ``RFP traceability`` project in a large telecom company. The project goal was to facilitate knowledge retrieval assisting in RFP internal compliance and proposal reply, see Figure 3. Internal compliance relates to the real product knowledge needed for building the offering by Pre-sales people to the customers.

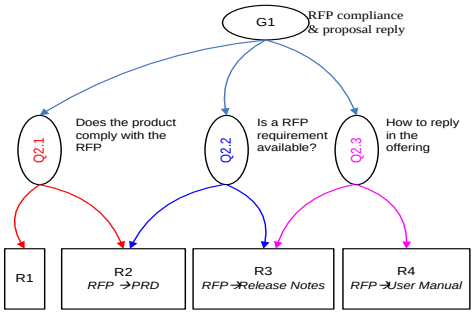

Figure 3: ***GQR model of the RMR***. For the Goal of investigating RFP compliance, Question Q2.1 raised by the Sales Engineer regards whether the product requirements definition (PRD) complies with RFP, while Question Q2.2 raised by the R&D VP refers to product features availabity in existing product releases. Each Question relates to a specifc product documentation of relevant content Results.

The questions that are asked in an RFP traceability project include:

Q2.1. Does the existing product *comply* with a given RFP requirement?

Q2.2. Is a given RFP requirement *available*, and in what release of product?

Q2.3. How to *reply* in the offering to a RFP requirement? (What references are used for that?)

This case study analysis resulted in 1 goal, 3 questions and 4 Results as described in Figure 3:

G1→ {Q1, Q2, Q3 }

Q1→ {R1, R2} | Q2 →{R2, R3} | Q3→ {R3, R4}

As part of the GQR analysis of the "RFP traceability" case study in Figure 3, we added to each Result required to answer a research question the respective references to existing requirements resources. For example, both RFP and Release notes are required for R3, see table 2.

Table 2: GQR reference to requirements repositories.

Result	Data asset (data set / service)	Requirements repositories
R2	Trace RFP requirements to product existing features	R2 ←RFP ←PRD ←RFP2PRD
R3	Trace RFP requirements to actual release notes	R3 ← RFP ←Release notes ←RFP2Release

5 EVALUATION OF THE GQR MODEL AS SUPPORT TO RE

In the biomedical case study, the GQR investigation model can be used as a mechanism for both organizing and checking the biomedical informatics needed to support a cancer research. The GQR model is in the first place useful for designing a well-defined research via goal/question modelling in order to achieve effective research analysis of outcomes. Once the research is managed via GQR, it can be used as guidance for checking the required bioinformatics resources. Due to the specialized nature of the various cancer research sub-fields, often only the experts interviewed for defining each use case could provide us with the required information for structuring the use cases and identifying the content sources for each result.

In the CMMI requirements repository (RMR), the GQR model encouraged the stakeholders to express their ultimate questions that usually are asked at most urgency and require lots of meetings

and discussions, and put a lot of pressure on people. The GQR was found to be very useful in structuring the content in the RMR, while specifying the required content to fill in in order to provide the required result.

6 DISCUSSION AND FUTURE WORK

Content rich systems are difficult for requirements elicitation, mainly because of the huge amounts of content they process and the variety of potential users that cannot envision their requirements from the new system to be developed. The proposed GQR (Goal-Question-Result) model directs and guides the Requirements Elicitation process, in elevating the elicitation discussions from specific data elements to contextual structures of goal-question-result related content requirements, while scoping the discussion to specific stakeholders.

Problems of requirements scope result from lack of understanding of the organization in which the system under development will be placed. Especially problematic in content rich systems is understanding the users of the target's system output, what output is required for them, and how the target system will change the organization's means of doing business.

The basic structure of Goal-Question-Result acts as a mean of defining the requirements scope, from which the requirements will then be elicited. Goals and questions are related to the requirements stakeholders, thus encouraging them to simulate their operational concept for the system under development.

The results of the two case studies show that the GQR model reveals most important content requirements while helping stakeholders to articulate their rationale via goals. This in turn is applied iteratively between Goal-Question-Result rounds, providing more and more contextual knowledge about the system.

Future work will focus on traceability between stakeholders and resources, while integrating between data, stakeholders, and presentation required from resources.

ACKNOWLEDGEMENTS

The author is grateful to Iaakov Exman for a critical reading of early versions of the paper.

REFERENCES

Anton A., 1996. Goal-Based Requirements Analysis, *Proceedings of JeRE '96*, College of Computing Georgia Institute of Technology Atlanta, Georgia,USA.

Bertrand P. et al, 1998. GRAIL/KAOS: an environment for goal driven requirements engineering. ICSE'98 - 20th International Conference on Software Engineering, IEEE-ACM.

Biedermann J. and Grierson D., 1995. A Generic Model for Building Design, Springer Publisher.

Begent, R. et al, 2005. Challenges of Ultra Large Scale Integration of Biomedical Computing Systems, 18th IEEE International Symposium on Computer-Based Medical Systems, IEEE.

Chen P. P., 1966. The entity-relationship model: toward a unified view of data, ACM Transactions on Database Systems, ACM.

Christel M. and Kang K., 1992. Issues in Requirements Elicitation, (CMU/SEI-92-TR-012). Software Engineering Institute, Carnegie Mellon University.

Cockburn A., 2001. Writing Effective Use Cases. Addison Wesley.

Cohn M., 2004. User Stories Applied: For Agile Software Development. Addison-Wesley.

CMMI, 2015. Capability Maturity Model Integration (CMMI) Overview. Software Engineering Institute. http://cmmiinstitute.com/

Dardenne A. et al, 1993. *Goal-Directed Requirements Acquisition*. Science of Computer Programming.

DOD-Mil-std 498, 2015. IEEE standards for Software Development Process, IEEE, https://en.wikipedia.org/wiki/MIL-STD-498

EBI, 2015. European Bioinformatics Institute, http://www.ebi.ac.uk

Finkelstein et al, 2006. Developing an Integrative Platform for cancer Research: a Requirements Engineering Perspective, Fifth e-science All Hands Meeting (AHM2006), e-science.

Kavakli E. and Loucopoulos P., 2003. Goal Driven Requirements Engineering: Evaluation of Current Methods, Department of Cultural Technology and Communication, University of the Aegean. Copyright EMMSAD'03.

Lapouchnian A., 2005. Goal-Oriented Requirements Engineering: An Overview of the Current Research, Department of Computer Science, University Of Toronto.

NCBI, 2015. The National Center for Biotechnology Information, http://www.ncbi.nlm.nih.gov

NCI, 2015, National Cancer Institute – Center for Biomedical Informatics & Information Technology, http://cbiit.nci.nih.gov/

NCRI, 2015. Cancer Informatics Initiative – www.cancerinformatics.org.uk

Neetu K. and Pillai A., 2013. A Survey on Global Requirements Elicitation Issues and Proposed Research Framework, ICSESS 2013, Proceedings of

2013 IEEE 4th International Conference on Software Engineering and Service Science, IEEE.

PharmaKGB, 2015. The Pharmacogenomics Knowledgebase, http://www.pharmgkb.org/

PLOS, 2015. The Public Library of Science (PLOS), http://www.plos.org/

Robertson S. and Robertson J., 1999. Mastering the Requirements Process, Addison-Wesley, Cambridge, MA, USA.

Solingen A. and Berghout E., 1999. The Goal/Question/Metric Method, McGraw-Hill Publishing Company, New York, NY, USA.

Zdravkovic et al., 2013. Using i* to Capture Consumer Preferences as Requirements for Software Product Lines, iStar 2013, *Proceedings of the 6th International i* Workshop.*