

Estimativa de Esforço em Projetos de DDS

Carlos Peixoto, Tiago Bortoluzzi, Thais Ebling, Rafael Prikladnicki, Jorge Luis Nicolas Audy, Mauricio Cristal, John Wheeler

Faculdade de Informática (FACIN)
Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS)
90.619-900 – Porto Alegre – RS – Brasil

carlos.peixoto.001@acad.pucrs.br, tiago.bortoluzzi@acad.pucrs.br,
thais.ebling@pucrs.br, rafaelp@pucrs.br, audy@pucrs.br,
mauricio.cristal@gmail.com, john@walkerwheeler.com

Abstract. *This paper presents the results of a systematic literature review on software effort estimation applied to distributed software development projects. Because the results did not explicitly explore effort estimation within distributed development, the paper also presents the results of interviews conducted with software project managers in order to understand the particularities of the effort estimation process and techniques in the context of distributed projects.*

Resumo. *Este artigo apresenta os resultados de uma revisão sistemática da literatura sobre estimativas de esforço em projetos de desenvolvimento distribuído de software. Como os estudos encontrados não exploram explicitamente estimativa de esforço em ambientes distribuídos, o artigo também apresenta os resultados de entrevistas conduzidas com gerentes de projetos de software buscando entender as particularidades do processo e das técnicas de estimativas de esforço no contexto de projetos distribuídos.*

1. Introdução

O desenvolvimento de um *software* requer que seja executada uma série de tarefas. Atualmente, existe uma grande preocupação com o gerenciamento de projetos de software em especial no que diz respeito a controle de custos e esforço necessários para o desenvolvimento do produto. Quando se planeja um projeto de um software, estimar o esforço necessário para o desenvolvimento é uma atividade crítica de gerenciamento, uma vez que estimativas são crucias para correta alocação de recursos e podem determinar a viabilidade ou não de um projeto de software, bem como perda de competitividade para a organização (Agarwal et al, 2001).

Embora a importância de estimativas acuradas seja conhecida, estimar o esforço e o custo associado a desenvolvimento de software continua a ser o elo fraco no gerenciamento de projetos de software. As organizações necessitam de acuradas estimativas de esforço nos estágios iniciais de um projeto. Entretanto, quanto mais cedo são feitas essas estimativas, maior é a probabilidade de que elas sejam baseadas em informações incompletas sobre o escopo do projeto, o que aumenta a probabilidade da imprecisão das estimativas. Nomeia-se como processo de estimativa o conjunto de técnicas e procedimentos que uma organização usa para estimar um projeto.

Em um contexto de Desenvolvimento Distribuído de Software (DDS) a estimativa de esforço também é crítica. Mas ainda não se tem conhecimento de quais abordagens e técnicas de estimativas poderiam ser utilizadas no sentido de aumentar a precisão dos projetos de DDS. Ou ainda, como as propostas existentes poderiam ser adaptadas para melhorar a estimativa de esforço neste contexto. As atuais abordagem (por exemplo, estimativa por analogia, modelos paramétricos) e técnicas (por exemplo, COCOMO, pontos de função, pontos de caso de uso), levam em consideração fatores importantes tais como experiência e senioridade da equipe, informações sobre o ambiente e sobre o produto a ser desenvolvido. Mas não se tem conhecimento de muitos estudos que exploram estas abordagens em um contexto de DDS (Almeida et al, 2009).

Por este motivo, este artigo possui dois objetivos. Inicialmente, são apresentados os resultados de uma revisão sistemática da literatura de estimativa de esforço em projetos de desenvolvimento de software, onde se buscou identificar estudos que exploram estimativa de esforço em projetos de DDS. Posteriormente, devido a falta de evidências de estudos existentes nas bases consultadas, uma pesquisa exploratória foi conduzida, seguindo-se o método de estudo de caso conforme Yin (2003), com o objetivo de entender alguns dos aspectos que torna a estimativa de esforço em projetos de DDS diferente. O estudo de caso foi conduzido através de entrevistas semi-estruturadas com colaboradores de um centro de desenvolvimento de uma empresa multinacional localizada no Sul do Brasil. Os resultados são apresentados e os próximos passos são discutidos.

Na próxima seção são apresentados os conceitos relacionados com estimativa de esforço em projetos de desenvolvimento de software. A Seção 3 apresenta os resultados da revisão sistemática da literatura, e a Seção 4 apresenta os resultados do estudo de caso. Por fim, na Seção 5 discutem-se os resultados e os próximos passos desta pesquisa.

2. Estimativa de Esforço em Projetos de Software

De acordo com Agarwal et al (2001), o processo de estimar continua sendo uma das tarefas mais complexas e que apresenta a maior quantidade de falhas no gerenciamento de projetos. Os autores definem o processo de estimativa como o conjunto de técnicas e procedimentos que uma organização utiliza para chegar a uma estimativa. Um aspecto importante a ressaltar é que este é um processo contínuo, ou seja, o resultado de uma execução do processo atualiza os dados históricos para uma nova execução do processo.

Agarwal et al (2001) caracterizam o esforço necessário para desenvolver um projeto como um dos principais fatores que contribui para custo de desenvolvimento do mesmo. Portanto, baseado nessa afirmação, conclui-se que existe uma forte relação entre as estimativas de esforço e estimativas de custo e que muitas vezes essas duas estimativas são utilizadas na literatura como sinônimas. Segundo Sommerville (2007), não há uma maneira simples de se estimar o esforço necessário para desenvolver um software, entretanto as organizações necessitam estimar custo e esforço de software. Para realizar essas estimativas pode-se utilizar uma ou mais das técnicas de estimativas existentes, organizadas em algumas abordagens, que Sommerville (2007) agrupa segundo suas características (Tabela 1).

Tabela 1 – Abordagens de Estimativa (adaptada de Somerville (2007))

Técnica	Descrição
Modelagem Algorítmica	Um modelo é desenvolvido usando informações históricas de custos que relacionam alguma métrica de software. É feita uma estimativa dessa métrica e o modelo prevê o esforço necessário.
Julgamento de Especialista	Vários especialistas nas técnicas de desenvolvimento de software e no domínio da aplicação são consultados. Cada um deles estima custo e o esforço do projeto. Essas estimativas são comparadas e discutidas. O processo de estimativa prossegue iterativamente até que uma estimativa em comum seja atingida.
Estimativa por analogia	Essa técnica é aplicável quando outros projetos do mesmo domínio de aplicação foram concluídos. O custo e o esforço necessários para o novo projeto são estimados com base nos projetos análogos a ele concluídos.
Lei de Parkinson	A Lei de Parkinson estabelece que o trabalho se expande para preencher o tempo disponível. O custo e o esforço são determinados pelos recursos disponíveis e não pela avaliação subjetiva.
Atribuição de preço para ganhar	O custo do software é estimado para ser o que o cliente tiver disponível para gastar com no projeto. O esforço estimado depende do orçamento do cliente e não da funcionalidade do software.

3. Revisão Sistemática da Literatura sobre Estimativa de Esforço em DDS

Nenhum processo de pesquisa científica deve acontecer sem a existência de um processo de revisão da literatura, ou seja, a identificação dos estudos existentes na área deve ser parte do processo de pesquisa. No contexto de pesquisas em Engenharia de Software (ES), essas revisões são normalmente realizadas sem um processo rigoroso ou de uma forma “*ad hoc*”, o que dificulta sua replicabilidade (Biolchini et al, 2005).

Recentemente, Kitchenham (2004) introduziu o conceito de revisão sistemática da literatura na área de ES. Esse método vem sendo utilizado em outras áreas, como a Medicina, ao longo do tempo, e se consolidou como uma forma para sumarizar evidências sobre uma área de interesse bem como para identificar possíveis lacunas na pesquisa nessa área. Seu sucesso pode ser observado pelo crescente número de revisões sistemáticas que têm sido conduzidas na área de ES.

Como a literatura na área de estimativa de esforço é ampla, uma possível identificação dos estudos existentes através de uma pesquisa não sistematizada torna-se complexa devido a grande quantidade de informações existentes na área. Além disso, pode ser igualmente difícil identificar quais estudos têm realmente investigado estimativa de esforço em um contexto de DDS. Nesse sentido optou-se por identificar estudos primários relevantes para o tema em questão, através da execução de um estudo secundário, ou seja, a revisão sistemática da literatura da área.

O processo de revisão sistemática adotado nesse trabalho seguiu as recomendações apresentadas por Biolchini et al (2005), que sugere um processo de revisão sistemática em cinco fases: formulação da questão, seleção das fontes, seleção dos estudos, extração da informação e síntese dos resultados.

3.1. Definição das Questões de Pesquisa

No contexto de estimativa de esforço, o foco da pesquisa foi identificar as abordagens de estimativa de esforço em desenvolvimento de software e como estas abordagens são adaptadas para ambientes de DDS. Também se buscou nesse trabalho identificar boas práticas na utilização das abordagens de estimativas existentes bem como identificar

propostas de possíveis melhorias nessas abordagens. As seguintes questões foram formuladas em relação ao tema em estudo:

Q1: Quais são as abordagens de estimativas existentes e quais suas características?

Q2: Quais são as boas práticas de utilização dessas abordagens e quais as lições aprendidas foram obtidas com o uso dessas práticas?

Q3: Quais são as propostas de melhorias para as abordagens de estimativa existentes?

Q4: Quais destas abordagens têm sido utilizadas em ambientes de DDS e como elas têm sido adaptadas?

Baseado nessas perguntas identificou-se a questão foco da pesquisa como sendo:

“Quais são as abordagens de estimativa existentes na literatura e como elas tem sido aplicados no desenvolvimento de software e, em particular, no Desenvolvimento Distribuído de Software?”

Com objetivo de tornar a revisão o mais ampla possível foram considerados artigos que apresentem propostas de abordagens de estimativa bem como artigos que apresentem as abordagens existentes. Nesse contexto, buscou-se identificar pontos fortes e fracos da pesquisa existente na área e possíveis lacunas na mesma.

3.2. Execução da Revisão Sistemática

Essa revisão sistemática utilizou os seguintes mecanismos de busca como fonte para a pesquisa: IEEE Xplore, Compendex, Springer Link, Science Direct. Os critérios para essa decisão foram os seguintes: a base de dados deveria poder ser consultada através do convênio existente entre PUCRS e CAPES; a base de dados deveria permitir a utilização de mecanismos de busca que possibilitem a utilização de palavras-chave e expressões booleanas; a base de dados deveria permitir acesso ao artigo completo; a base de dados deveria possuir artigos na língua inglesa, uma vez que esse é o idioma em que a maioria dos artigos publicados na área é escrita;

O estudo de Kitchenhan et al (2007) serviu como fonte de referência para escolha das principais palavras-chave utilizadas nesta revisão. A Tabela 2 apresenta as palavras-chave selecionadas divididas nas categorias população, intervenção e resultados, seguindo as definições das propostas de revisão sistemática de Kitchenhan (2004) e Biolchini et al (2005).

Tabela 2: Palavras chave utilizadas na revisão sistemática

Categoria	Palavras-chave
População	“software”
Intervenção	“effort estimation”, “effort estimating”, “effort prediction”
Resultados	“model”, “method”, “methodology”, “process”, “approach”, “technique”, “experience”, “lessons learned”, “assessment”

Pela combinação das palavras-chave foi definida a seguinte expressão de busca:

“software” AND (“effort estimation” OR “effort estimating” OR “effort prediction”) AND (“model” OR “method” OR “methodology” OR “process” OR “approach” OR “technique” OR “experience” OR “lessons learned” OR “assessment”)

Pode-se perceber que não existem palavras-chave associadas com DDS. O motivo é que, após alguns testes, optou-se por considerar neste estudo apenas palavras relacionadas com a área mais ampla, estimativa de esforço em projetos de desenvolvimento de software. A razão desta escolha deveu-se principalmente ao fato de que se entendeu que projetos de DDS apareceriam como um subconjunto da pesquisa maior. Como as questões principais estavam relacionadas com estimativa de esforço como um todo, quando eram incluídas palavras-chave associadas com DDS não eram retornados artigos, ou os artigos acabavam não tendo relação com as questões definidas.

A seleção dos estudos foi realizada em duas etapas, na primeira etapa foram selecionados 85 artigos pela leitura do título e resumo, representando 8% do total de artigos selecionados pela expressão de busca; na segunda etapa foram selecionados 20 artigos pela leitura do texto integral, tendo como critério de seleção a identificação de estudos que abordassem conceitos de abordagens de estimativa ou apresentassem resultados de experimentos utilizando abordagens de estimativas ou relatassem o uso de abordagens de estimativa na indústria, incluindo artigos que explorassem aspectos do DDS. Esse processo ocorreu no período entre Abril e Maio de 2009 e foi realizado por dois pesquisadores, utilizando a opinião de um terceiro em caso de discrepância nas avaliações.

Os artigos considerados relevantes para a revisão foram agrupados em três categorias: Conceitual, com estudos que apresentavam conceitos sobre abordagens de estimativas, proposta de abordagens de estimativas; Experimento, com artigos que apresentavam resultados de experimentos ou estudos de caso sobre a aplicação de abordagens de estimativas; e Experiência: com artigos que apresentavam relatos de experiência da aplicação de abordagens de estimativa na indústria.

Para a nossa surpresa, não foram encontrados, na revisão sistemática conduzida, estudos descrevendo experiências com abordagens de estimativa de esforço no contexto de DDS. Como não tínhamos conhecimento de trabalhos anteriores na área, não havia artigos utilizados como controle (artigos que fornecem indicações de que estudos relevantes podem estar sendo identificados). Sendo assim, as seguintes razões foram discutidas para estes resultados: pode ser que os estudos que relatam experiências com estimativa de esforço em DDS utilizem diferentes palavras-chave; pode ser que realmente não existam estudos específicos sobre estimativa de esforço em DDS; pode ser que o estudo de estimativa de esforço no contexto de DDS não seja tão relevante, ou não apresente diferenças significativas que mereçam um maior foco.

Optou-se por investigar a última razão, ou seja, coletar dados empíricos sobre quais são os possíveis impactos do DDS na estimativa de esforço em projetos de desenvolvimento de software. Estes dados foram coletados através de entrevistas conduzidas em uma empresa multinacional no Brasil, e são apresentados a seguir.

4. Um Estudo de Caso sobre Estimativa de Esforço em Projetos de DDS

Com o intuito de coletar informações acerca das abordagens de estimativas utilizadas e de como ocorre atualmente o processo de estimativas de esforço na prática, e em um contexto de DDS, optou-se pelo planejamento e execução de um estudo de caso, seguindo as orientações de Yin (2003).

Sendo assim entre Maio e Junho de 2009 foi desenvolvido protocolo de pesquisa contendo o plano para a execução do estudo, incluindo os roteiros para entrevistas semi-estruturadas. Logo após, foi feita a validação de face e conteúdo e pré-teste do instrumento de coleta de dados para então aplicar a versão final deste instrumento em uma série de entrevistas com doze funcionários de uma empresa multinacional com centros de desenvolvimento de software distribuídos ao redor do mundo. As entrevistas foram executadas no centro brasileiro e incluiu quatro desenvolvedores, dois líderes de desenvolvimento, dois gerentes de TI, um gerente de projetos, dois analistas de negócios de TI e um engenheiro de *performance*. As entrevistas ocorreram presencialmente, estando sempre presente dois pesquisadores e o entrevistado. Um dos pesquisadores foi responsável pela aplicação e gravação de áudio da entrevista e o segundo pesquisador foi responsável pela geração de manuscritos das mesmas.

Para analisar os dados coletados com as entrevistas optou-se pela utilização do método de pesquisa conhecido como *Grounded Theory* ou “Teoria fundamentada nos dados”. Este método de pesquisa compreende a geração sistemática de teorias a partir de dados (Creswell, 2009). Neste método o pesquisador investiga um determinado assunto sem ter uma teoria a ser testada, mas com o intuito de entender uma determinada situação. Os dados são geralmente analisados de forma incremental, de forma que a análise chega ao fim quando novos dados não agregam mais informações ao que já foi identificado.

Neste sentido, para a análise dos dados das entrevistas utilizou-se um subconjunto das fases deste método sugerido por Strauss e Corbin (1998), especificamente as fases de codificação aberta (*open coding*) e codificação axial (*axial coding*). A codificação aberta envolve a identificação, análise e categorização dos dados considerados relevantes, e a codificação axial examina as relações entre estas categorias. Primeiramente foram identificados códigos (conceitos e fenômenos de interesse) nos textos das entrevistas. Estes códigos foram agrupados em conceitos similares formando as categorias de informações, e por fim foram analisadas as relações entre estas categorias, obtendo assim algumas conclusões das entrevistas. A figura a seguir apresenta um exemplo de codificação feita nas entrevistas:

Nesta figura é possível observar a codificação das informações relacionadas a técnica de estimativas *Delphi*. Podem-se visualizar os pontos positivos (+), pontos negativos (-) e sugestões para esta técnica, segundo os entrevistados. Pode-se perceber também que os códigos são apresentados seguidos de dois números que representam respectivamente o seu grau de fundamentação (*groundedness*) e o de densidade teórica (*density*). O grau de fundamentação mostra o número de citações com as quais o código está associado, já o grau de densidade teórica mostra o número de relacionamentos do código com outros códigos.

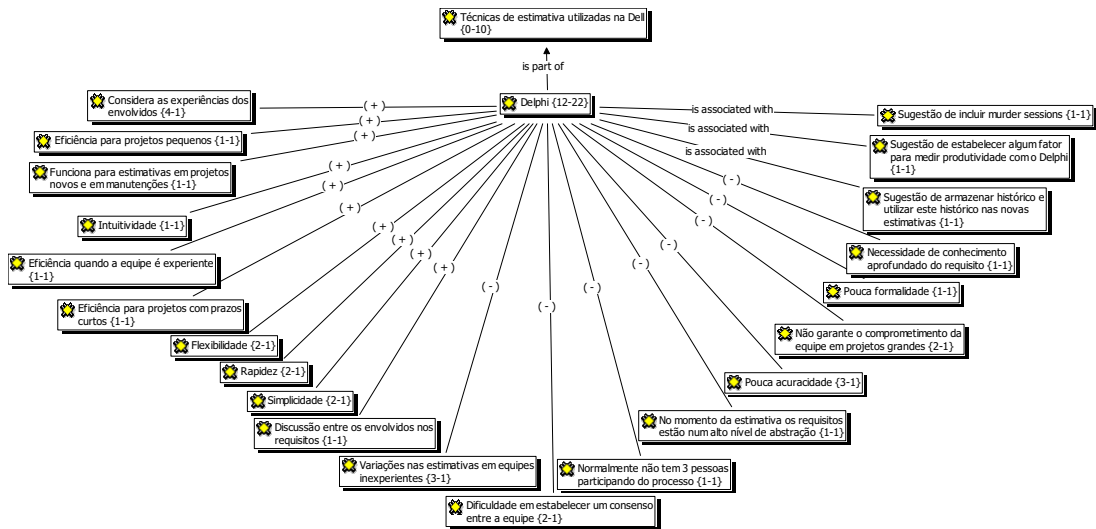


Figura 1. Codificação das entrevistas

4.1. Resultados do Estudo de Caso

A técnica de estimativa mais utilizada nos projetos sejam eles distribuídos ou não foi Delphi. O principal ponto positivo desta técnica é considerar as experiências dos envolvidos no processo de estimativas. Seus principais pontos negativos são as variações que podem ocorrer em estimativas de equipes inexperientes e a pouca precisão da técnica. Para a melhoria da técnica sugere-se definir um fator para medir produtividade, armazenar o histórico e utilizá-lo nas novas estimativas e incluir “murder sessions” (reunião para a discussão da relevância do requisito) em sua execução;

Um ponto negativo geral é que as estimativas ficam centradas na opinião somente de determinadas pessoas. Sendo assim, identificou-se que as técnicas eram utilizadas não por que eram as melhores para um determinado projeto, mas sim por que eram as técnicas de maior conhecimento da equipe. Uma sugestão neste sentido foi a de criar formas de identificar, baseado nas características dos projetos, nos níveis de distribuição das equipes, nos tipos de tecnologias e outros fatores, as melhores técnicas para determinadas situações. Além disso, armazenar o histórico também foi considerado relevante, bem como compartilhar conhecimento sobre as técnicas entre as equipes e medir a produtividade dos times, além de fornecer treinamentos sobre as técnicas existentes.

Em relação ao processo de estimativas, percebeu-se que existem processos distintos dentro da empresa e têm-se opiniões de que até nem exista um processo de estimativas. A maioria dos participantes primeiramente faz uma estimativa *high-level* (chamada *T-shirt*) baseando-se nos requisitos iniciais recebidos. Eles depois analisam esta estimativa com mais afinco e refazem a mesma se necessário. Os principais pontos positivos deste processo eficiência e a inexistência de burocracia, ou menos burocracia na execução do processo, além de ser simples, usar a experiência dos envolvidos e incluir mais pessoas no processo decisório se comparado com Delphi. Seus principais pontos negativos são a instabilidade do requisito, o uso da *T-Shirt* como estimativa final, o pouco detalhamento e tempo despendido na estimativa e o fato do desenvolvedor não participar do processo de estimativas. Para a melhoria do processo sugeriu-se refazer a

sugeriu-se refazer a *T-shirt*, conhecer e utilizar novas técnicas neste processo, analisar a estimativa sob outro ponto de vista, reduzir a quantidade de requisitos instáveis, gastar mais tempo no processo de estimativas, balancear o nível de conhecimento dos envolvidos e compartilhar conhecimentos de soluções de requisitos.

De forma geral, os pontos negativos do processo de estimativas são principalmente a falta de padronização entre equipes e a falta da análise posterior da precisão das estimativas. As sugestões gerais para a melhoria do processo de estimativas são padronizá-lo, definir um time para controlar o uso das técnicas, incluir o Analista de Negócios na elicitação de requisitos, compartilhar experiências entre os times, discutir possíveis soluções de requisitos e definir papéis e fases do processo de estimativas.

5. Discussão e Considerações Finais

A partir da análise das conclusões das entrevistas pudemos perceber que a técnica mais utilizada na empresa é Delphi. Pelo fato desta técnica se basear nas experiências dos envolvidos no processo de estimativas, deve-se atentar ao fato de que os responsáveis pelas estimativas devem ter conhecimento das equipes que desenvolverão o requisito, para que suas estimativas sejam acuradas. Este tipo de conhecimento pode ser proveniente da definição de cenários e melhores casos para o uso de determinadas técnicas, do fornecimento de treinamentos sobre as técnicas, do uso de uma base histórica de estimativas, do compartilhamento de conhecimento entre as equipes, da análise da produtividade dos times, entre outros.

Analisando as entrevistas podemos perceber também a existência de processos de estimativas distintos dentro da empresa, existindo opiniões de participantes de que até nem exista um processo de estimativas. Este é um reflexo dos ambientes distribuídos, onde as equipes utilizam processos, ferramentas, técnicas diferentes por não terem contato umas com as outras ou por não existir uma padronização destes aspectos dentro da empresa, o que é justamente uma das sugestões dos participantes para o processo de estimativas da empresa.

Desta maneira, as conclusões obtidas com as entrevistas aplicadas nesta empresa indicam necessidades pertinentes aos ambientes distribuídos, tais como: padronização do processo e uso das técnicas de estimativas, base histórica de estimativas e produtividade de equipes, definição de cenários de uso de determinadas técnicas, compartilhamento de conhecimento e experiências entre equipes. Sendo assim, entende-se que estudos que explorem as diferenças e adaptações do processo e técnicas de estimativas de esforço em ambientes de DDS são necessários e relevantes.

Agradecimentos

Estudo financiado pela Dell Computadores do Brasil Ltda., com recursos da Lei Federal Brasileira nº 8.248/91.

Referências Bibliográficas

- Agarwal, R. et al. (2001). “Estimating Software Projects,” ACM SIGSOFT Software Engineering Notes, v. 26, n. 4, pp 60-67.
- Almeida, A. C. M., Souza, R., Aquino, G. Meira, S. (2009). “Effort Drivers Estimation for Brazilian Geographically Distributed Software Development,”. In: Proc of

- SEAFOOD (Software Engineering Approaches for Offshore and Outsourced Development), Zurich.
- Biolchini, F. et al. (2005). "Systematic review in software engineering," Technical report. Systems Engineering and Computer Science Department, Rio de Janeiro.
- Creswell, J. W. (2009). "Research Design: Qualitative, Quantitative, and Mixed Methods Approaches". Third Edition. University of Nebraska-Lincoln. SAGE Publications Ltd.
- Kitchenham, B., Mendes, E., Travassos, G. H. (2007). "Cross versus Within-Company Cost Estimation Studies: A Systematic Review," IEEE Transactions on Software Engineering, v. 33, n. 5.
- Kitchenham, B. (2004). "Procedures for Performing Systematic Reviews," Joint Technical Report SE0401 and NICTA technical report 0400011T.1. Software Engineering Group. Department of Computer Science. Keele University.
- Strauss, A. e Corbin, J., (1998). "Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory". 2 ed. London, SAGE Publications.
- Yin, R. K., "Case Study Research: Design and Methods", 3rd edition, Sage, 2003.