

Quick Prototyping of Web Applications^{*}

Lukasz Olek, Bartosz Michalik, Jerzy Nawrocki, Mirosław Ochodek

Poznań University of Technology, Institute of Computing Science,
ul. Piotrowo 3A, 60-965 Poznań, Poland

{Lukasz.Olek, Bartosz.Michalik, Jerzy.Nawrocki,
Miroslaw.Ochodek}@cs.put.poznan.pl

Abstract. Web applications are getting more and more complex. Because of this, effective communication with prospective end user is essential. In the paper the concept of quick prototyping of web applications is presented and early experience with implementation of the concept is discussed. This approach assumes that requirements are written in form of use cases supplemented with screen designs. Quick prototype is generated automatically from these artifacts.

1 Introduction

Web applications are getting more and more complex. They serve as a basis for e-governance, enterprise content management etc. (it is predicted that in 2008 the ECM market will amount to \$3.9 billion [8]). Designing such complex and innovative applications require good communication with end-users at requirements elicitation stage. It is well-known that end-user feedback is very important, so experts suggest to use prototyping [21,20,4] to enhance this feedback. However, in constantly changing business environment, a prototype becomes additional artifact that requires maintenance and involves additional costs.

In *Quick Prototyping* we generate a *mockup* from use cases and screen designs. Such choice seems to be reasonable, since most projects use scenarios or use cases to write requirements (>50%) and user interfaces designs to visualize future system (>65%) [15]. The prototype is "quick" as it can be automatically generated from the requirements specification. It is also "cheap" as after changing requirements you can easily regenerate the *mockup*. *Mockup* is a simple web application that presents use cases together with screen designs attached to each step. An end user can animate the use cases to understand the application under development and can perform a review (the prototype collects feedback from a number of end-users and presents reports).

Since we generate a *Mockup* automatically, the requirements must be written using a semi-formalized model, understandable by a computer. This rises two main questions that are answered in this paper:

^{*} This research has been financially supported by the Ministry of Scientific Research and Information Technology grant N516 001 31/0269.

- Is the formalization usable for analyst, so is this model flexible enough to describe real web applications?
- People don't like formal models, because they seem to be difficult to understand. Thus, is generated *mockup* easy to understand?

In the paper various methods of prototyping are discussed (in Section 2) and the concept of quick prototyping of web applications is presented (Section 3). The concept was implemented as a part of the *UC Workbench* [13]. We describe our early experience with the *mockup* (Section 4) and present the mechanism of end-user-feedback measurement and analysis (Section 5).

2 Other Prototyping Methods

There are many well known and commonly used approaches to prototyping. In *Paper Prototyping* [17,20] the vision of future software is presented to an user using paper sketches. An analyst acts as a computer: draws and presents different sketches according to user's actions. User tries to execute some tasks with this "system". The analyst quickly sketches low-fidelity screens on paper during the session with the user, thus this method does not require extensive preparations. However, presentation costs are quite high: the analyst needs to invest his time to examine different functions with many users. The presentation of the prototype cannot be done remotely. Maintainability costs are also high: when requirements change, you need to repeat the process.

Storyboarding [10] approach is somewhat similar to Paper Prototyping. Here analyst does not prepare a single screen sketch, but the whole sequence of screens for the particular task. Storyboards are presented to users, and feedback is received. Storyboarding is less interactive than Paper Prototyping, because user can only watch it, but it is much easier to present such a prototype remotely (e.g. you can scan storyboard and send it using e-mail). The biggest problem arises with maintainability of storyboards. It is difficult to update the prototype - usually you need to throw away earlier version, and prepare a new one.

In prototyping using *RAD approach*, programmer uses Rapid Application Development (RAD) tools to produce a Graphical User Interface (GUI) layer using programming languages. It requires quite substantial effort to prepare such prototype, because RAD tools are optimized for the software development not prototyping. This approach is not very agile because it is not easy to introduce changes. RAD prototyping belongs to high-fidelity methods. It is difficult to say, whether high- or low-fidelity is better [18,23,22], but some studies [23] show that users tend to focus too much on graphical aspects of system on high-fidelity prototypes. One big advantage of this approach is that it is possible to use the prototype as a basis for GUI of future application.

The comparison of selected methods is shown in Table 1.

Table 1. Comparison of prototyping methods.

Attribute	Paper Prototyping	Storyboarding	RAD Prototyping
Preparation costs	Low	Low	High
Maintainability costs	Moderate	Moderate	High
Understandability	Easy	Easy	Moderate
Remote presentation	Difficult	Easy	Easy

3 The Concept of Quick Prototyping Method

Requirement engineering experts [2,3] say that requirements should only present behavioural aspects of future system. These aspects shouldn't be mixed with user interface details, because it would disturb in seeing the overall vision of future system. On the other hand people are holistic beings, so the textual description is often not enough to present the vision. Because of that many experts advice to create prototypes (e.g. [21]) and present them to end-users to validate requirements.

Quick Prototyping tries to combine this two aspects (behavioural and graphical). The aspects are linked together, but kept separately (see Fig. 1). This allows analyst to focus only on system behaviour in the beginning, without being disturbed by user interface details.

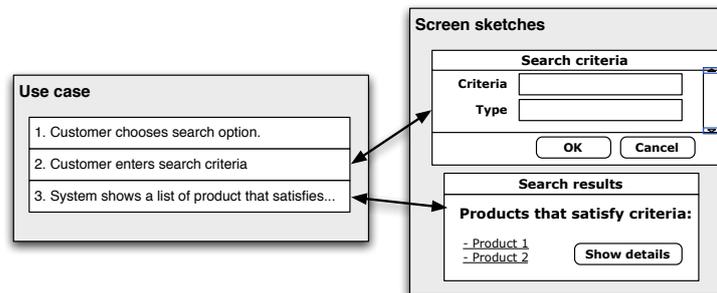


Fig. 1. A concept of keeping behavioural and user interface aspects separate. This orthogonality helps in introducing changes to the model, because they need to be introduced in one place only.

Behavioural aspects are written in form of structured use cases. Single use case [3,7,19] describes interaction between user and future system formed in sequence of steps. Use case has main scenario – steps that describe the most common way to reach user goal, and extensions that describe exceptional situations. A use case comprises a set of scenarios with the same goal. Each extension describes another scenario. Such use cases are semiformal. They are expressed in a natural language but the description has a structured form of a sequence

of steps and extensions. Each step can be connected with lo-fidelity sketches of application screens visible by user at this step.

Generated *Mockup* is similar to a storyboard (see Section 2) – it presents sequence of screen sketches according to some scenario. *Mockup* combines use cases (i.e. behavioural description) with attached screen designs on one view. A generated *mockup* is based on a web browser and it consists of two frames (see Fig. 2):

- the scenario window presents the currently animated use cases (it is the left frame in Fig. 2) and the current step is shown in bold;
- the screen window shows the screen design associated with the current step (it is the right frame in Fig. 2).

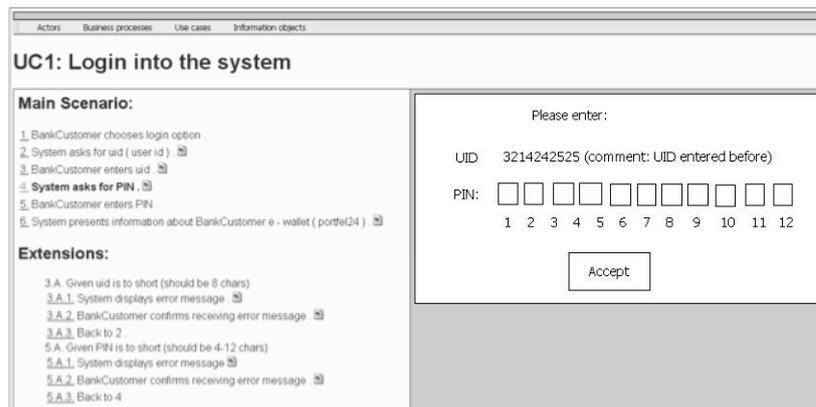


Fig. 2. An Example of *mockup* generated with *UC Workbench*. You can see a use case on the left – viewers use it to navigate through the mockup, and screen design on the right – a particular screen attached to selected step (in bold).

Automatically generated *mockup* has one great advantage over other prototyping methods – it is easy to change. When it requires a change, there is no need to do this twice: once in prototype, and second time in requirements specification. You just change use case or screen sketch and regenerate *mockup*. Moreover, if one screen is used many times in use cases (e.g. log-in screen), you change it in one place only (this is the main maintenance problem with Storyboards).

Mockup generation mechanism does not need to know the meaning of particular steps. The steps can still be expressed using natural language and remain readable by humans. Only the structure of steps need to be formalized, because the generator needs to know the order of steps and screens designs connected with them. Use cases and screen designs are internally stored in XML format for generation purposes.

3.1 Quick Prototyping with *UC Workbench*

UC Workbench [1,14,13] is a tool helping analysts during elicitation of requirements. It helps with use-case authoring according to model presented on Fig. 1. It reads use cases and screen designs in a special XML format. It has an integrated use-case editor – text editor optimized for use cases, and screen sketch editor – graphical editor for quick screen sketching. *UC Workbench* has an integrated *Quick Prototype* generator, thus is able to generate *mockups* (see Fig. 2).

3.2 Specification for low-fidelity screen sketches

Screen sketches could be stored as bitmaps, but then semantical information would be lost. *UC Workbench* tries to formalise screen structure by introducing components that can be used for sketching application screens quickly. Currently we support components for designing web applications. Screens can contain three types of components:

- **Static** – user cannot act on them, but can read information: e.g. text or image
- **Dynamic** – a user can interact with them (read information and act): e.g. edit box, check box, combo box, radio buttons, list box or button. They represent HTML controls.
- **Grouping** – used to structure screen properly. Currently there is one component of this type: *Repeater*, used for grouping elements that can be repeated on the screen (see Fig. 3).

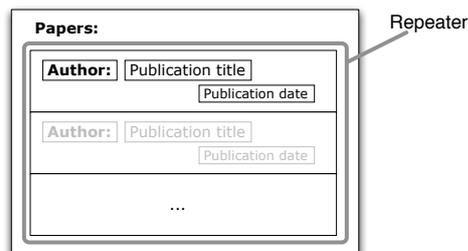


Fig. 3. Example of Repeater component. This component is used to model list and tables in web applications. It contains other elements that can be repeated on the screen.

4 Early Experience with Quick Prototyping

4.1 Flexibility of the Model

Usually an analyst uses general text editor and sheets of paper to author requirements, thus he/she has a full freedom. *Quick Prototyping* generator needs semi-formalized requirements (see Sec. 3), so the question arises: is this model flexible enough to be used for real systems? To answer this question following case study was conducted.

Three web systems were chosen to check if modelling existing systems in details is possible. The focus was put on one use case from each system:

- transfer cash in a *banking application*: www.bzwbk.pl
- bid an auction in an *auction system*: www.allegro.pl
- buy an item from an *on-line shop*: www.amazon.com

Use cases were gathered by browsing through each described system. Then screen designs were sketched using low-fidelity approach. Afterwards three *mockups* were generated.

We did not find any obstacles to model selected functionality in details. There was a straightforward mapping between screens from real application to our *mockup*'s screens. However, we have found one interesting aspect to consider: what is the best way to specify a tree of components (e.g. tree of categories of books in e-shops)? Further research is needed to find the best metaphor for such situation, because it is not easy to specify them in formal, but understandable way.

Low-fidelity static screens approach has some limitations. It is hard to show complicated interaction (eg. specify how the system will react to a drag and drop operation). Analyst can overcome this problem using a comment layer on screens to describe how the system will react. This problem would also need a further research.

4.2 Is Mockup Easy to Understand?

We have decided to verify thesis that *mockup* is an easy to use and self-explanatory prototype, which provides comprehensive understanding of system business logic and user interface (even with a limited analyst support).

Another valuable property of *Quick Prototyping* might be receiving feedback concerning usability of application, at the early stage of the design process. A simple experiment was conducted in order to compare system usability evaluation based on the *mockup* and pure use-case description.

Experiment process description. The main idea of the first experiment was to verify whether potential end users are able to write down all operations required to attain established business goal with the use of *mockup* instead of the real system. We have decided to choose an electronic shop system under the

assumption that e-commerce domain is well known to the participants. The experiment was conducted in 2006 at the Poznan University of Technology. The participants were 4th year students working on their master degrees in Software Engineering (SE). There were 15 students in the examined group. The experiment procedure consisted of the following steps:

1. Introductory presentation, covering a *mockup* concept (15 min).
2. Survey: verification of students past experiences regarding the e-commerce domain (especially electronic shops) and requirements engineering with use cases.
3. Main task: achieving business goal from the user perspective, using the online shop *mockup*.
4. Survey: participants personal assessment of the *mockup* prototyping concept and usability of the presented e-commerce system.

The second experiment structure was similar with two exceptions: participants were 5th year Master Course in Software Engineering (SE) students (23 people) and *mockup* prototype was replaced by the use cases based functional requirements specification. Although the task remained, the purpose of the experiment was different. The main goal was to compare participants opinions about the presented system usability.

Main task description. Participants main task was to execute some operations on a *mockup*. We assumed that proper system for this experiment purpose, should conform two requirements:

- In the real projects customer have a extensive knowledge about the system business domain. Thus, it was very important that participants were also familiar with the problem domain.
- Proposed system should be more complex and less usable than application participants might have encountered in the past. If not, participants could solve the task by heart, basing on their previous experiences.

We decided to create a *mockup* of electronic shop (known domain), but we have introduced several usability problems, mostly concerning breaking common conventions and complicated user interface flow. To verify complexity of the system in procedural aspects of usability, we decided to use a GOMS [5] model (NGOMSL implementation). It enabled comparison of predicted execution times for buying single product scenario in the proposed *e-Shop* and *Amazon.com*. According to the GOMS simulation results, time required to complete the same task using two systems was:

- *e-Shop* - 184 seconds,
- *Amazon.com* - 121 seconds.

It seemed that buying a single product required 52% more time to complete in case of the *e-Shop* system. Thus, we assumed that proposed system fulfilled complexity requirement.

Participants were asked to use the *e-Shop* system *mockup* (based on 7 use cases) in order to write down a sequence of operations one should perform on the real system, to buy two types of products: apples and books. Each operation entry was logged using the following format (for example see figure 4):

- use case unique id,
- step number,
- action describing single user activity put down using following format: <Type> '<Name>'[:<Value>], where:
 - <Type> is a type of an object, e.g. Button, Link, Edit etc.,
 - '<Name>' is a name of an object, e.g. 'Login', 'Surname' etc.,
 - <Value> (optional) is value used to fill a field.

UC_ID	STEP	Action
UC2	1	Button 'Search'
UC2	3	Edit 'Name' : Lobo
UC2	3	Combo 'Category' : Fruits
UC2	3	Button 'Search'

Fig. 4. Example of participant operations log

Although students were working individually, they could request analyst support (1 analyst was assigned to 5 students). Each analyst intervention was reported and described.

Participants were explicitly provided with all data necessary to achieve the goal (e.g. product name, category, user personal data etc.).

Is *mockup* self-explanatory - experiment results. Before participants were asked to proceed to the main task, they were surveyed in order to examine their familiarity with the e-commerce domain and use cases. Questions included in questionnaire and results of the survey are presented in the table 2. It seems that participants were browsing through the content of electronic shop before and nearly all of them have bought a product at least once. Participants were also familiar with the concept of creating functional requirements specification with the use cases.

Before analysing participants operations logs, assumption was made that business goal was achieved if there were no major errors concerning business logic in the log (e.g. omitting the whole use case). We have also defined two types of errors which might occur:

- *Flow error* - minor defect in business logic (e.g. omitting in execution use case extension, when it was triggered),
- *Data error* - wrong input data format (form validation errors) or inaccurate data entered.

Table 2. Summary of the domain knowledge verification survey. Possible answers where: never, once, few times (less than 10), many times (10 or more).

Question	Never	Once	Few times	Many times
1. How many times have you been browsing through a content of the electronic shop?	0	0	1	14
2. How many times have you been buying products in the electronic shops?	2	1	8	4
3. How many times have you been preparing use cases (for how many systems)?	0	1	13	1
4. How many times have you been reading requirements specification written with use cases?	0	0	12	3

Another important aspect was the analysts involvement in helping participants attaining business goal. We divided analyst interventions into two categories:

- *R1* - small problems or questions concerning screens, fields validation etc.,
- *R2* - more serious problems concerning understanding business logic (flow of control in use cases),

To our surprise, all participants achieved business goal. Results summary is presented in table 3. The average analyst interventions per participant was 1.6, including both R1 and R2 issues. Moreover, approximately 30% of the questions concerned business logic (R2). Average number of errors made by a single

Table 3. Is *mockup* self-explanatory experiment results summary

Participant	Time[min]	\sum Flow errors	\sum Data errors	\sum Errors (F+D)	R1	R2	\sum (R1+R2)
Avg	53.93	0.40	0.80	1.20	1.13	0.47	1.60
Std dev	6.64	0.63	0.86	1.26	0.99	0.64	1.40

participant was 0.40 for flow problems and 0.80 for data defects.

It seems that *mockup* is easy to understand. It provides information necessary to understand the developed system business logic (average total errors number per single use case was 0.17). It also does not require much analyst support (average analyst support requests per use case was 1 / 4.38).

Mockup helps to unveil usability problems. In real projects customer representatives are supposed to review requirements and find defects: e.g. business logic defects or usability defects. The question is how to present requirements in order to get the best feedback? Is it easier for reviewers to observe defects if they have application screens presented together with text of requirements?

After completing experiment's main task, participants were asked to express their opinion about the *mockup* concept and usability of the *e-Shop* system. On

question regarded usability of the *e-Shop* system, 93% of the responders stated that presented *e-Shop* system, has serious design problems concerning usability. Thus, another experiment was conducted: participants were given only use cases without screen sketches.

Another group of students (5th year SE students) was asked to complete the same task but using use cases for the shop. Students familiarity with the e-commerce and use cases was also verified. Participants operations were logged on higher level (use case steps rather than GUI controls).

Although, most of the *e-Shop* system usability problems concerned operational aspects, which should be easy to identify with use cases, the second group was not so unanimous in their evaluations. Nearly the same percentage of participants decided that system has (43%) and does not have (39%) usability problems.

The comparison of both groups usability evaluation is presented in the figure 5. It seems that enhancing use cases scenarios with screen sketches gives a better understanding of system possible usability pitfalls at the early stage of the design process.

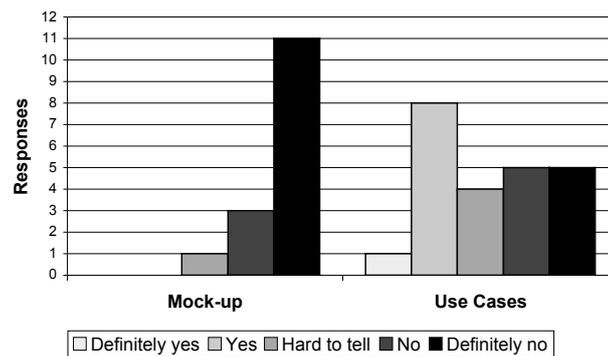


Fig. 5. The *e-Shop* system usability evaluation by *mockup* and Use Cases group. The question asked was "Do you think, that presented *e-Shop* system is properly designed in the matter of usability?". Possible answers where: definitely yes, yes, hard to tell, no, definitely no.

Experiments conclusions The *mockup* prototype seems to fulfil it's main role. Firstly, it presents designed system in easy to understand way (all participants achieved business goal). End users working with such a prototype do not require much analyst support (1 analyst support requested per 4.38 use case was observed). Thus, we believe that such a prototype is self-explanatory enough to be accessed remotely by the end user with only limited help of analyst.

Secondly, presenting system being developed as a *mockup* can also help to identify both operational and GUI usability problems. Common sense says that this should be true for prototyping in general, but the main advantage of *mockup* is that it is very quick and cheap to produce and maintain.

5 User-feedback Measurement and Analysis

5.1 Tracing Review Preparation Time

Having requirements presented in digital form (*mockup* is a web application) gives another big advantage: it allows to track time that each person spent on reviewing it (of course they should be aware, that their activity is being tracked). It motivates customers and end users (which are usually very busy) to review *mockup* more thoroughly. It also allows analyst to check which parts of *mockup* were paid enough attention, and which parts should be reviewed once again.

5.2 Number of Defects Estimation

Common understanding of system requirements is crucial for the project success. However it is often hard to fully involve customer representative into process of requirements elicitation and elaboration (13% of project has failed because of lack of client involvement and 12% because of wrong or incomplete requirements [11]). Our proposal is to use asynchronous reviews to develop system requirements with client involvement in continuous way. We have chosen *mockup* as a subject for review.

Requirements reviews can save time (decision made in this project phase have great impact on a final product) and reduce project costs (cost of fixing errors in requirements increases exponentially from phase to phase) [11]. However review itself costs time and money. Distributed (in the context of place and time) access to specification being reviewed, can reduce this cost. There is also a question of desired level of quality. How do we know how many defects do requirements have? Is this level accepted, or not? An old method called *capture-recapture* used to count number of animals in some bounded environment [16] can help to count number of defects in requirements specification.

As it is shown in the figure 6 asynchronous review process consists of three phases: *Review*, *Data Cleaning* and *Risk Analysis*.

First phase, *Review* is based on n-fold review[24] and runs on *mockup* server. Each reviewer is equipped with special checklist (generated from good practises for use cases [2,3,6,9]) to let him focus on some basic problems. This list should be combined with the business domain questions to focus reviewers on business logic defects. In this process reviewers do not communicate and do not see defects marked by others. Defect is identified as a violation of a checklist item attached to particular step (or whole use case). This method allows automatic grouping of defects for capture-recapture method.

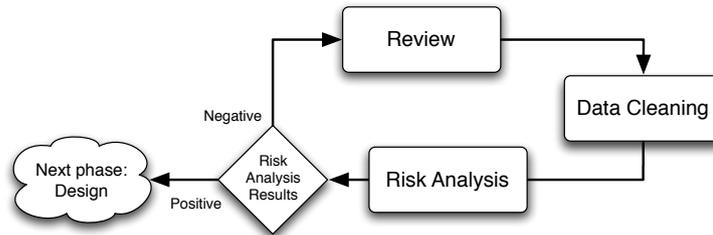


Fig. 6. Asynchronous review process cycle. It begins with distributed *Review* by a few of inspectors, then an analyst gathers defects and cleans the data (*Data Cleaning*), finally *Risk Analysis* is performed.

Second phase, *Data Cleaning* is preparing data (defects set) for the capture-recapture analysis. To improve the capture-recapture results accuracy we must correct wrong defects and remove duplicates.

In third phase, *Risk Analysis* a number of defects is estimated using *capture-recapture*. These calculations can help to decide whether the quality of requirements is satisfactory. If the quality is good, the process finishes. If not, analyst has to correct requirements and the process starts from the beginning. There are four *capture-recapture* models that are used [12] in software reviews. The simplest model is called *two-sample model*. First sample is catch, marked and returned to population. Then second sample is catch. The population size is calculated from simple proportion 1:

$$\frac{m}{n} = \frac{M}{N} \Rightarrow N = \frac{M * n}{m} \quad (1)$$

where:

m – the number of marked elements in second sample, n – second sample size, M – first sample size, N – population size,

There must be four conditions fulfilled to make this computations reliable: population must be closed, marking should be persistent, all marks needs to be correctly recorded, each animal should have constant and equal probability of capture. All these conditions can be easily translated into software reviews domain [12]:

1. Specification doesn't change during the review.
2. Inspectors don't reveal their defects to others.
3. Inspectors must ensure that found defects are accurately documented.
4. All inspectors are provided with identical information, and the inspectors should have similar knowledge and experience.

6 Conclusions

In the paper we have presented the concept of quick prototyping and its implementation within the *UC Workbench*. The pilot version of the tool has been used by a number of local companies and the first feedback is very positive. Moreover, the experiments we have conducted suggest that the *mockup* can support communication with end users in an effective way.

Acknowledgements

First of all we would like to thank the students involved in the *UC Workbench* project. We would like to thank the IBM company for awarding Eclipse Innovation Grant to *UC Workbench* project. It allowed students focus on the development work. This research has been financially supported by the Ministry of Scientific Research and Information Technology grant N516 001 31/0269.

References

1. UC Workbench project homepage. <http://ucworkbench.cs.put.poznan.pl>.
2. Steve Adolph, Paul Bramble, Alistair Cockburn, and Andy Pols. *Patterns for Effective Use Cases*. Addison-Wesley, 2002.
3. Alistair Cockburn. *Writing Effective Use Cases*. Addison-Wesley, 2001.
4. Larry L. Constantine and Lucy A. D. Lockwood. *Software for Use: A Practical Guide to the Models and Methods of Usage-Centered Design*. ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, 1999.
5. Kieras D. A guide to GOMS model usability evaluation using NGOMSL. *The Handbook of Human-Computer Interaction*, 1996.
6. IEEE. IEEE Recommended Practice for Software Requirements Specifications (IEEE Std 830-1998), 1998.
7. Ivar Jacobson. Use Cases - Yesterday, Today, and Tomorrow. Technical report, Rational Software, 2002.
8. McNabb K. ECM Growth Outpaces The Overall Software Market. Technical report, Forrester Research, 2005.
9. Daryl Kulak and Eamonn Guiney. *Use Cases: Requirements in Context, Second Edition*. Addison-Wesley Professional, July 2003.
10. James A. Landay and Brad A. Myers. Sketching storyboards to illustrate interface behaviors. In *CHI '96: Conference companion on Human factors in computing systems*, pages 193–194, New York, NY, USA, 1996. ACM Press.
11. Dean Leffingwell and Don Widrig. *Managing Software Requirements: A Use Case Approach, Second Edition*. Addison-Wesley Professional, May 2003.
12. James Miller. Estimating the number of remaining defects after inspection. *Software Testing, Verification & Reliability*, 9(3):167–189, 1999.
13. Jerzy Nawrocki and Łukasz Olek. UC Workbench - A Tool for Writing Use Cases. In *6th International Conference on Extreme Programming and Agile Processes*, volume 3556 of *LNCIS*, pages 230–234. Springer Verlag, Jun 2005.
14. Jerzy Nawrocki and Łukasz Olek. Use-Cases Engineering with UC Workbench. In Krzysztof Zieliński and Tomasz Szmuc, editors, *Software Engineering: Evolution and Emerging Technologies*, volume 130, pages 319–329. IOS Press, oct 2005.

15. CJ Neill and PA Laplante. Requirements Engineering: The State of the Practice. *Software, IEEE*, 20(6):40–45, 2003.
16. D. L. Otis. Statistical inference from capture data on closed animal populations. *Wildlife Monographs*, 62(62), 1978. Wildlife Society.
17. Marc Rettig. Prototyping for tiny fingers. *Communication of the ACM*, 37(4):21–27, 1994.
18. Jim Rudd, Ken Stern, and Scott Isensee. Low vs. high-fidelity prototyping debate. *interactions*, 3(1):76–85, 1996.
19. G. Schneider and J. P. Winters. *Applying Use Cases: A Practical Guide*. Addison-Wesley, 1998.
20. Carolyn Snyder. *Paper Prototyping: The Fast and Easy Way to Define and Refine User Interfaces*. Morgan Kaufmann Publishers, 2003.
21. Yan Sommerville and Pete Sawyer. *Requirements Engineering. A Good Practice Guide*. Wiley and Sons, 1997.
22. Robert A. Virzi, Jeffrey L. Sokolov, and Demetrios Karis. Usability Problem Identification Using Both Low- and High-Fidelity Prototypes. In *Proceedings of the CHI Conference*. ACM Press, 1996.
23. Miriam Walker, Leila Takayama, and James A. Landay. High-Fidelity or Low-Fidelity, Paper or Computer? Choosing Attributes When Testing Web Applications. In *Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting*, pages 661–665, 2002.
24. Yuk Kuen Wong. *Modern Software Reviews: Techniques and Technologies*. IRM Press, 2006.