RWTH AACHEN UNIVERSITY

Building an Effective Rewarding System based on Reputation

Study Programme:

M.Sc. in Media Informatics

Master of Science Thesis
Fraunhofer Institut für Angewandte Informationstechnik FIT

by

Henri Agaraj

First Examiner:
Prof. Wolfgang Prinz, PhD

Second Examiner:
Prof. Volker Wulf, PhD

Supervisor:
Dipl.-Inf Christian Prause

Semester:
Winter Semester 2011/12

Submission Date:
26 March 2012

Fraunhofer FIT

Building an Effective Rewarding System based on Reputation

THESIS

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE

in

MEDIA INFORMATICS

by

Henri Agaraj
born in Tirana, Albania
286109

Media Informatics
Fakultät für Mathematik, Informatik und Naturwissenschaften
Rheinische Westfalische Technische Hochschule
Aachen
Germany

Layout based on a design by the Media Computing Group,
RWTH Aachen University

http://hci.rwth-aachen.de/karrer_
thesistemplate

I hereby declare that I have created this work completely on my own and used no other sources or tools than the ones listed, and that I have marked any citations accordingly.

Hiermit versichere ich, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie Zitate kenntlich gemacht habe.

Aachen, March 2012
Henri Agaraj

Contents

List of Figures

List of Tables

List of Abbreviations

FBM	Fogg Behavior Model for Persuasive Design
Captology	Computers as Persuasive Technologies
IP	Integrated Project
ebbits	Enabling the business-based Internet of Things and Services
Moknow	Mobile Knowledge
CDE	Collaborative Development Environment
SNRS	Social Network-Based Recommender Systems
SSO	Single Sign-on
CAPLE	Context and Attention in Personalized Learning Environments
CWE	Collaborative Working Environment
CSCW	Computer Supported Cooperative Work
MVC	Model-View-Controller
GWT	Google Web Toolkit
UIDL	User Interface Defintion Language
JSON	JavaScript Object Notation
FP7	Framework Programme 7
IoP	Internet of People
IoT	Internet of Things
IoS	Internet of Services

IoPTS	Internet of People, Things, and Services
SVN	Subversion
AJAX	Asynchronous JavaScript and XML
GWT	Google Web Toolkit
HTTP	Hypertext Transfer Protocol
UI	User Interface
JSON	JavaScript Object Notation
CSS	Cascading Style Sheets
JAR	Java Archive
API	Application Programming Interface
URI	Uniform Resource Identifier
POJO	Plain Old Java Object
JPA	Java Persistence API
JSP	JavaServer Pages
RIA	Rich Internet Applications
OSGi	Open Services Gateway initiative

Abstract

This thesis project presents a persuasive reputation-based social game named *ReputationForge*, aiming at increasing the participants' engagement in the team wikis of two different working environments. The result of an extensive conceptual investigation into a range of behavioral studies and existing reputation-based systems led to the final concept of ReputationForge. Aspects like *community/social, reputation, personal-enjoyment and self-fulfillment* are the main ingredients of the concept.
In more detail, ReputationForge promotes a soft competition between the participants, who are divided into newbies and advanced ones. Participants start, and further complete tasks, that are related with the every-day work in the team wiki, such as creating or improving articles, and rating them. Tasks are to be fulfilled individually and their successful completion awards participants karma points that increase their overall reputation. Clear feedback related to the progress of each task is given. Last but not least, ReputationForge introduces triggers, which help making the underlying wiki more visible and thus attracting users to keep coming back to it.
An evaluation in two case studies (ebbits and Moknow) was performed in combination with a control group field study. It showed an increased engagement of the experimental group in the wiki of both case studies, thus proving evidence that a social-game approach based on reputation can actually increase the participants' engagement in a corporate/team wiki. Still, the experimental group of ebbits registered a statistically significant increase in the majority of cases, while there was no statistically significant increase (except the number of rated articles) for the experimental group of Moknow. The result in the latter study can be attributed to a short evaluation time, a well established wiki, and the further involvement of the participants of the control group in the existing system Moknowpedia.

Acknowledgements

First, I would like to express my deep gratitude to my supervisor, Dipl.-Inf. Christian Prause, for his continued support and guidance during the writing of this thesis. His door was always open to discuss my work and his understanding, encouraging and personal guidance have provided a very good basis for the present thesis. I wish you all the best in the future and a successful finalization of your PhD-thesis.
I would also like to express my gratitude to my examiners, Prof. Wolfgang Prinz, PhD and Prof. Volker Wulf, PhD, for reviewing my thesis and for taking an important part in the process at an early stage. Your feedback was a valuable foundation of my work.
A sincere gratitude goes to Dr. Markus Eisenhauer, for his important tips during the evaluation process of this thesis.
Thank you to all the members of the MoKnow group at Fraunhofer FIT.ICON, ebbits, and the students who have worked in C5-124, for their friendly collaboration, and support throughout the design process of this thesis. I would particularly like to thank M.Sc. Jonathan Simon for his valuable advice and friendly help. Our extensive discussions around my work and his interesting suggestions have been very helpful for this study.
An important prerequisite to begin and successfully finish my studies was the backing from my parents. All the support and above all the love they have provided me over the years was the greatest gift anyone has ever given me. Thank you mom, for having always believed in me, and for making me smile when I most needed it.
Thank you Hergys, for everything. You are the most wonderful brother in the world. I wish you, Julia, and my adorable nephew Christian a wonderful life together.
I would like to express my thanks to Simona, Ramesh, Edlira, and Markus. You have encouraged me a lot, especially in hard times, when I felt lost. Thank you all so much!
Finally, a big "Thank you" goes to the DAAD, who made my study in Germany possible and gave me a great support in getting to know better this great country, its beautiful language, and its people.

Conventions

Throughout this thesis the following conventions are used:

Definitions of technical terms or short excursus are set off in coloured boxes.

EXCURSUS:
Excursus are detailed discussions of a particular point in a book, usually in an appendix, or digressions in a written text.

Definition:
Excursus

Source code and implementation symbols are written in typewriter-style text.

```
myClass
```

Citations and bibliography references are written in APA[1] style.

The whole thesis is written in American English and in passive voice.

[1] http://www.apastyle.org/

Chapter 1

Introduction

"If you are trying to change the way you run a company, one of the most visible things you have to change is the way you compensate, reward and recognize people."

—Paul Allaire, CEO XEROX CORP, Harvard Business Review, October 1992.

Introduction of wiki

Ward Cunningham, the developer of the first wiki software, WikiWikiWeb, described it originally as "the simplest online database that could possibly work." Since its introduction, wiki has become a popular Web 2.0 tool, which reshaped the process of collaborative knowledge creation.

End-user development

Furthermore, wikis are an example of the emerging paradigm of *End-user development*, that makes it possible to users with insufficient background in computer science to modify their own application to their needs, thus allowing them to efficiently apply information and communication technologies (Lieberman et al., 2006).

Corporate wikis

Based on the success of the public wikis, the process of adopting the wiki concept to working environments tries to address the long-lasting challenge of knowledge management. Poole & Grudin (2010) identify the following types of enterprise wikis: (a) single-contributor wikis, (b) group wikis, (c) and company-wide encyclopedias, and fur-

ther discuss the emerging challenges specific to these wikis. This thesis project focuses on the group/team wikis, which are used by a team as a means of communication, sharing information and coordination.

Collaborative working environment

The use of a team wiki, as an example of collaborative software creates a Collaborative Working Environment (CWE), whose core is the interplay of Computer Supported Cooperative Work (CSCW) and groupware (Martínez-Carreras et al., 2007). A lot of research has been done on this field and its future perspectives, including the design of a generic CWE, which allows several groupware systems to be interoperable (Martínez-Carreras et al., 2007).

Expertise Sharing

The intention of a team wiki is to be the primary location where the internal process knowledge is collected and also to serve as the center stage of managing and disseminating the collected knowledge focusing on the self-organized activities of the team members, described as *Expertise Sharing* by Ackerman et al. (2003).

1.1 Research Problem

Challenge of a team wiki

The greatest challenge of every team wiki is to encourage its long-term growth. Observing the participation rate in the team wiki of the EU-funded research project "Enabling the business-based Internet of Things and Services (ebbits)", there were only sporadic contributions from very few of its contributors.

The logical question that arises is: *Why did so few people participate?*

Distorted view about documentation

In comparison to the public wikis, where the majority of the users tend to be involved, the participants in a team wiki tend to be inactive and not contribute with great enthusiasm. This is confirmed by Selic (2009), according to which, developers have a distorted view about documentation, perceiving it as unpleasant, because it presents no value to them and it is not created for their own benefit. Contributing to the internal wiki is considered by the par-

ticipants as a part of their work that they would better prefer to avoid (although the team itself would benefit).

Lack of time and suitable participation incentives

Additionally, the lack of time among the participants reduces their willingness to contribute with informative documents even more. Last but not least, the goal of having a thriving wiki is very often not achieved, because there is no careful consideration regarding participation incentives.

An existing attempt to increase participation

Dencheva et al. (2011) introduced some extensions to Moknowpedia—an internal team wiki, which is used by the groups Mobile Knowledge (Moknow) and Context and Attention in Personalized Learning Environments (CAPLE) in Fraunhofer FIT to produce a large number of technical documentation. The goal of the introduced extensions was to explicitly recompense users for their efforts and time in contributing to Moknowpedia, to reward them for their participation, and to further motivate them to higher levels of participation.

Remarks

The results were satisfactory, but still the evaluation results pointed out the following remarks:

- *The rewarding mechanisms applied offer only a temporary motivation,*
- *Needed additional rewards, that are directly related with a participant's performance,*
- *Lack of structured goals and dynamic tips,*
- *Lack of notification emails about the own engagement in wiki.*
- *Lack of a list displaying all the new articles, and the articles not yet rated by the participants.*
- *A combination of punishment and rewarding mechanisms is missing.*

Need to investigate the motivational values of the target audience

Understanding the motivational values of the participants is a prerequisite to find out why participants were only temporarily motivated and how to achieve that they keep coming back to it.

Importance of reputation in a team wiki

Returning back the focus to the team wiki, the collaborators in such a wiki have pre-existing relationships, contributors are less anonymous, thus participants are concerned with holding their reputation, which is in a team wiki pervasive and inescapable (Poole & Grudin, 2010). Indeed, Wasko & Faraj (2005) show that people contribute more knowledge if the contributions lead to increasing reputation. The key to success is combining reputation with motivational values, which aims at maximizing the users' experience leading them towards higher level of participation (Farmer & Glass, 2010).

Creating thriving wikis

An extensive research into the state of the art of motivation theories (cf. Chapter 2—"Theoretical Foundations"), indicates a long-standing controversy between those studies who believe that desired behavior is redirected at best by intrinsic reinforcement and those who believe that the extrinsic reinforcement (rewards, etc) is the best way. Still, studies agree that purely external rewards, (i.e money) are considered to be motivators that last only for a short-period, and thus ineffective (French et al., 1973). The majority of the analyzed studies identify factors to consider while trying to create thriving successful public wikis. There is unfortunately very few research about the effect of such factors in a work environment, thus making the direct transfer of them to team wikis very difficult, because of the different nature that these wikis exhibit, including the participants' own motivation values, goals and prejudices.

Little research about motivational values in corporate wikis

Two studies on team wikis

Nevertheless, Prause et al. (2010) and Yetim et al. (2011) investigate a wide range incentive mechanisms to motivate users to contribute with invaluable information. Both studies show a preference towards the following motivational values in a working environment: *(a) Community/Social Effect, (b) Reputation, (c) Personal enjoyment, (d) Self-fulfillment and (e) Curiosity.*

1.2 Goal

System goal

Inspired by the findings of Dencheva (2010), and the above-mentioned immaterial incentives particularly and the current state of the art in general, this Master's thesis presents a web-based system based on reputation, aiming at:

increasing the engagement of the members of Moknow and ebbits in their respective team wikis.

Engagement metrics

The participants' engagement is measured through quantifiable participation metrics as follows: *(a) Number of author contributions to the team wiki—creating new wiki articles, or improving existing wiki articles (either any article authorship, or at least 30% article authorship), (b) Number of rated articles, and (c) Number of log-ins* (for more details see Section 5.2.1—"Experience Scores".

Difference between wikis

The group Moknow has a well-established team wiki with a satisfactory participation rate, while the group ebbits uses a relatively new wiki with a general low engagement[1].

Documenting in wiki is a benefit for the team

Contributing to documentation is perceived as an unpleasant activity (as discussed in Section 1.1—"Research Problem") by the majority of the users but "a necessary chore for the sake of higher good and for the benefit of the team in general" (Selic, 2009).

1.3 Strategy

Captology

Computers as Persuasive Technologies (Captology) is a new research area, which deals with the efforts of persuasive computing technologies to influence people's attitudes in order to bring some constructive changes in many of their domains (Fogg, 1998). A lot of definitions, approaches, and perspectives can be found related to this topic.

[1]For more details about these working environments refer to Section 4.1.1—"Target Users/User Environment"

A socio-technical approach

Additionally, Huysman & Wulf (2004) believe that deep, and enduring changes in the participants behavior can be fulfilled by the adaption of a socio-technical perspective, which introduces to the participants a "social capital". Additionally, a later study focuses on the design requirements of tools that foster the concept of social capital supported by information technology (Huysman & Wulf, 2006).

Project thesis approach

This thesis project uses a game-based approach as an example of persuasive technology. It presents the target audience a *social game* called *ReputationForge*, which promotes a soft competition between the members, who try to start, and to complete tasks that are related to the every-day work in the team wiki, such as creating or improving articles, and rating them.

Game objectives

The game is designed to the following objectives, which help achieve the primary goal of this thesis project:

- The game motivates the target groups to increase their engagement in their respective team wikis by introducing the following motivational aspects:
  - *Community/Curiosity:* The drive to know new things regarding the community.
  - *Reputation:* The drive to grow one's status in the community, and one's perception of his importance to the community.
  - *Fulfillment incentives:* The desire to complete a task, assigned by oneself, a friend or the application.
  - *Recognition incentives:* The desire for the praise of others.
- The game motivates users more than the system by Dencheva (2010).

Design process and system evaluation

This thesis project will use a user-centred design process to develop a first prototype of the social game. Furthermore, the system will be evaluated in two case studies (ebbits and Moknow) in a field test, with the help of a control group in each case.

By analyzing the evaluation results, it is aimed to show that a game-based approach leads to a behavior change (i.e increased user engagement) among users participating in team wikis.

Does the chosen approach increase engagement?

1.4 Research Questions

This thesis project pursues the goal described in Section 1.2—"Goal" and respectively addresses the following general research question:

Can a persuasive game effectively fulfill the goal of increasing the participants' engagement in corporate/team wikis?

However, giving an overall and valid answer to this question in large scale is impossible within the scope of this thesis. Contextual factors like the underlying motivation of each participant to contribute to the wiki, the type of rewards introduced, and the amount of free time available are strongly interrelated with each other, and can possibly enhance or disturb the relationship between motivation and desired behavior. Enduring changes in the participation rate can only be proved through a longitudinal study.

Interrelation between contextual factors

The general research question can be further refined into these questions:

Research questions

- *Does a social-game, with features as described in Section 1.3—"Strategy" increase the participant's engagement in team-wikis (cf. Chapter 6—"Outcome Evaluation")?*
- *What is the overall satisfaction and motivation of the target users (cf. Chapter 7—"Implementation Evaluation")?*
- *What are the most suitable aspects of the concept that lead towards a higher participation rate in team wikis (cf. Chapter 7—"Implementation Evaluation")?*

1.5 Outline

Theoretical foundations

The following Chapter 2—"Theoretical Foundations") outlines the range of research areas that mainly influenced this thesis project, namely *(1) persuasive technology, (2) motivation and rewarding systems, (3) reputation and social software, and (4) game mechanics*.

State of the art

Subsequently, related work is presented in Chapter 3—"Related work". After discussing several design patterns applicable to reputation-based systems, examples of reputation systems are given, and why they are similar or different to this thesis: *(a) ranking systems, (b) rating systems, (c) collaborative filtering systems, (d) peer-based reputation systems, and (e) social systems with game mechanics principles*.

Final concept of thesis project

The next Chapter 4—"Finding a suitable concept" discusses the user-oriented process of gathering requirements which led to the final concept of *ReputationForge*. Before implementing this concept, it was previously validated by the target audience.

Implementation of a first prototype

The implementation concepts, along with the baseline systems needed to develop a working prototype, are discussed in Chapter 5—"Approaching a concrete implementation". Class diagrams and the detailed architecture show the interplay between the different components of the system and the baseline systems. Last but not least, a detailed description of the data structures, and the user interface of the system is given.

Evaluating the prototype

The following two chapters (cf. 6—"Outcome Evaluation" and 7—"Implementation Evaluation") describe the design and methods of the evaluation of ReputationForge. Both quantitative and qualitative results are described, which help giving answers to the research questions of this thesis project.

Conclusions and Future work

Chapter 8—"Conclusions" describes the contributions of this thesis project, regarding to which extent the set objectives were met. Finally, an outlook on both future work and lessons learned is given.

Chapter 2

Theoretical Foundations

"May you have a strong foundation when the winds of changes shift."

—from Bob Dylan's "Forever Young"

Foundation areas

This chapter addresses the theoretical foundations for research in the following areas: (1) persuasive technology, (2) motivation and rewarding systems, (3) reputation and web 2.0, and (4) game mechanics. State of the art in these areas can be seen as providing a theoretical foundation for further research and development in motivation theory.

2.1 Persuasive Technology

Definitions related to persuasion

This section describes the notions of persuasion, captology and persuasive technology, which aim at motivating and persuading people by creating successful information systems.

Definition: *Persuasion*

PERSUASION:
"a non-coercive attempt to change attitudes or behaviors" (Fogg, 2003).

Definition: *Captology*

CAPTOLOGY:
"a design, research, and analysis of interactive computing products created for the purpose of changing people's attitude or behavior" (Fogg, 2003).

Definition: *Persuasive Technology*

PERSUASIVE TECHNOLOGY:
"the application of persuasion strategy by means of computing technology" (Fogg, 2003).

Fogg and Captology

According to Fogg (1998), computers can change what individuals think and do. He mentions Captology to include the domain of research, design, and applications of persuasive systems. He tries to reveal how different information systems can be used to change people's attitudes and behavior.

Fields of application

Captology has been applied in several fields, including health, safety, education, environmental conservation, personal relationships and self-improvement, occupational productivity, reputation community involvement/activism, consumerism, etc. Existing empirical results have shown that persuasive technology can change people's attitude and behaviors to some extent (Fogg & Nass, 1997; Lenert et al., 2003).

Persuasive technology with respect to this thesis

This thesis project is an example of persuasive technology, because its aim is to change the behavior and attitude of its main users. Its purpose belongs to the major application domains for persuasion defined by Fogg (2003).

2.1.1 Fogg Behavior Model for Persuasive Design (FBM)

Introduction to FBM

Fogg (2009) argues that human behavior is a product of the following factors, which will be described in more detail in the following sections:

- *Motivation/Consequence:* having the reason for doing a specific desired action, which is related to the things that happened to an individual when he or she takes actions or performs certain behaviors.
- *Ability:* having the ability, competency of doing a specific desired action.
- *Triggers:* which give people impulse in starting a behavior, and additionally:
- *Focus:* knowing exactly what is required to be done.
- *Feedback:* knowing exactly how well a specific desired action is being done.

FBM main principles

This psychological model, asserts that for a desired target behavior to happen, a person must have sufficient motivation, sufficient ability/competence and an effective trigger. These three factors must occur at the same instant, or else the behavior will not occur.

Exploring FBM

FBM provides designers and scientific researchers with a systematic way to think about the underlying factors behind human behavior change. Figure 2.1 illustrates the above-mentioned factors that influence what people do. As shown, FBM has two axes. The vertical axis is for motivation, while the horizontal axis is for ability. A person who has low motivation to perform a specific target behavior registers low on the vertical axis, while a person who has low competence to perform a target behavior registers low on the horizontal axis. Both axes define a plane. The arrow shown indicates that once an increased motivation and ability is registered (directing to the symbolic start), it is more likely that a target behavior will be registered.

Behavior change according to FBM

Again according to this theory, in general, a behavior change would succeed faster if the behavior is made simpler instead of putting huge effort on motivation (nevertheless a considerable effort need to be put). People often resist attempts to motivation, but they naturally love simplicity. Fogg (2009) writes that elements like time, money, physical effort, brain cycles, social deviance, and routine actions are crucial in achieving the desired simplicity.

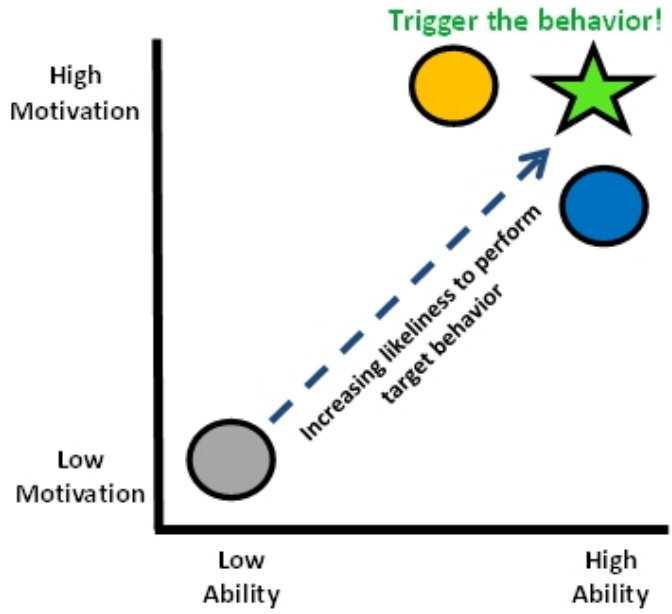


Figure 2.1: Fogg Beviour Model Graph. Source: Fogg (2009).

FBM implication

Briefly summarized, in order for behavior to occur, the motivation should be increased, and made simpler. But there is still a very important piece missing: Behavior must be triggered, with the help of triggers, as described below.

Triggers and Feedback

Importance of triggers

Without proper triggers, the desired target behavior will not occur even if high ability and motivation is registered. A trigger takes a lot of forms— alarm, mailing lists, text messages, notifications, etc. Successful implementation of triggers share all the following characteristics:

- *Making triggers visible.*
- *Associating triggers to the desired target behavior.*
- *Suitable timing of triggers.*

This thesis project relies on this behavior model, which encompasses behavior and persuasive design by combining the best practices from psychology, change management and design.

FBM with respect to this thesis

2.2 Motivation/Rewarding Systems

Rewards in a rewarding system serve as a means to motivate employees to perform in ways conforming to the organization's strategic goals. Their goal is furthermore to provide a systematic way to deliver positive consequences, which are to be understood as pleasurable, satisfying things that happen to an individual when he or she takes certain actions or performs certain behaviors (Wilson, 2003). According to Hoisl et al. (2007), implementing rewarding mechanisms in wiki systems generates benefits for its participants and serves the goal of achieving higher contribution rates.

Rewards and their relation to motivation

2.2.1 Motivation Theory

The motivation theory attempts to explain how and why individuals are able to achieve their goals (Petri, 1991).

MOTIVATION:
" the willingness of a person to exert high levels of effort to satisfy some individual need or want "(Kroth, 2007).

Definition: *Motivation*

The following describes different types of motivation theories.

Intrinsic vs. Extrinsic Theory

Wilson (2003) focuses on transforming compensation systems into truly effective reward systems. The book provides an outline of different extrinsic rewards that can serve

External Reinforcers

Category	Type
Verbal / Social	Specific Compliments
	Recognition
	Commendation Letters
	Award Dinners
	Celebration lunches/activities
	Take an Interest in Their Work
Work-Related	Promotion
	Special Development Programs
	Increased Decision Authority
	Increased Control Over Resources
	Access to Top Executives
	More Challenging Assignments
Tangible / Symbolic	Trophies
	Special Recognition Clubs
	Work related Tools
	Office Equipment
	Personal Items of Interest (e.g trips, time off, wine, etc.)
Monetary	Special Recognition Awards
	Individual Bonuses
	Group Incentives
	Pay Increases based on Merit
	Stock Related Rewards

Table 2.1: Four different categories of external reinforcers. Source: Wilson (2003).

as possible positive reinforcers (cf. Table 2.1). These reinforcers are categorized as *(1) verbal-social, (2) work-related, (3) tangible-symbolic, and (4) monetary*.

Rewards dimensions in a corporate wiki

Prause et al. (2010) determine several dimensions of rewards after discussing characteristics of beneficial competition. The dimensions identified can be shortly summarized:

- *Punishment vs. Reward*, this dimension determines whether best users are rewarded, or worse users are punished.
- *Virtual vs. Physical*, this dimension specifies whether the incentive is paid virtually or physically.
- *Anonymous vs. Personal*, this dimension specifies whether rewards are anonymized or are related to true identities.
- *Durable vs. Consumable*, this attribute determines whether awarded rewards to users cannot be lost again or endure just temporarily.
- *Individual vs. Group*, this dimension determines whether participants fight alone or collaborate together and therefore receiving the award together.
- *Show-off vs. Take-away*, this dimensions specifies whether the given reward is shown to other participants or is only presented to the person to whom the reward was given.
- *Intermittent vs. Continuous*, this dimension determines whether rewards are given continuously or only in predefined intervals of time.
- *Targeted vs. Open-ended*, this dimension determines whether a reward is given only after reaching a target goal or will grow and grow the more is achieved.
- *Public vs. Closed*, this dimension determines whether rewards are kept public or hidden to other participants.
- *Guarantee vs. Luck*, this dimension determines whether there is a guaranteed award of rewards after reaching a predefined goal, or a higher probability of being awarded a reward.
- *Competitive vs. Achievement*, this dimension specifies whether individuals compete against each other, or collaborate with each other to the achievement of a common goal.
- *Proclaimed vs. Discoverable*, this dimension determines whether the award is known before reaching a target, or only upon reaching the target.

Results of Prause et al. (2010)

The study performs a survey to find out an ideal reward system for software developers, in the context of large research projects. The results of the survey showed that a most-effective reward system is a system which allows *non-anonymous groups* compete in a *public contest* for *durable*, *physical*, *proclaimed rewards* through *specific targets* in a *predictable way*. Above all, *Curiosity and Learning*, *Self-Fulfillment*, and *Community* are considered as the participants main motives.

Prause et al. (2010) with respect to this thesis

The target users (cf. Section 4.1.1—"Target Users/User Environment") of this thesis project work in the same working environment analyzed in this study, thus its results are invaluable to understand user needs and wants.

Motivational values according to Yetim et al. (2011)

Yetim et al. (2011) determine common motivational values in two *working environments*, along with design features:

- *Reputation(-building)*: building reputable identities, point and status reward systems, etc.
- *Self-benefit*: feedback through rating of actions.
- *Self-development*: tagging awareness, allowing access to extra information.
- *Community*: explaining community benefits, informing benefits of contributions, motivating by goal-setting, social comparison, rewarding cooperative behavior.
- *Personal enjoyment*: integrating fun (game) as part of the participants' experience.

Rewards in open communities

Antikainen & Väätäjä (2010) try to find out what type of rewards can be used to motivate members, the reasons why people join, visit, participate and contribute in open innovation communities such as (a) open source communities, (b) problem solving communities, and (c) volunteer work in Wikipedia. This study summarizes the most important internal drives to participate in online communities as: *(a) altruism, (b) care for community, (c) attachment to the group, (c) enjoyement and fun, (d) peer recognition, (e) friendships, (f)*

relationships, (g) social support, (h) ideology, (i) interesting objectives, (j) knowledge exchange, (k) personal improvement, (l) reciprocity, (m) sense of efficacy, (n) and influencing.

Intrinsic drives in a volunteer environment

Clary et al. (1998) identify the following intrinsic categories as of main importance for behavioral change in a volunteer environment: *(a) values, (b) social importance, (c) understanding through learning experience, (d) career, (e) and enhancement*. Adding to these categories *fun and ideology*, Nov (2007) tries to find out which of these categories mostly lead people to freely share their time and knowledge with others. The result of the study shows *that fun, ideology, and understanding through experience* are perceived as *strong motivators to contribution*, while social, career, and enhancement are *not significantly strong motivators*.

Interaction between intrinsic and extrinsic motivation

Sicart (2008) defines the goal of achieving a balance between intrinsic an extrinsic motivation by means of rewarding mechanisms; a combination that Calder & Staw (1975) find in conformance with the self-perception theory, which envisions that intrinsic and extrinsic motivation interact with each other rather than combine cumulatively.

Pink's motivation theory

Another sustainer of the intrinsic theory Dan Pink, examines the puzzle of motivation during the TED Conference[1]. He starts with a fact that social scientists know but most managers do not: Traditional rewards are not always as effective as we think. According to him, tangible rewards by their very nature narrow our focus, concentrate the mind. Monetary incentives do not work or often do harm.

Purely extrinsic rewards do not work

Hoisl (2007) deeply studies the social rewarding in wiki systems aiming at motivating the community. Going to the conclusion that clean extrinsic motivations like money *cannot be used as a motivating factor*, they studied other factors like *(a) status, (b) (c) power, (d) acceptance, and (f) glory*. They tried to explain different social rewarding mechanisms which aim to meet the needs of users.

[1]For more information, see the website of the TED Conference.`http://blog.ted.com/2009/07/24/twitter_snapsho_54/`

Expectancy Theory

Expectancy theory as a motivation theory

In the study of organizational behavior, expectancy theory is a motivation theory first proposed by Vroom (1964). He argues that people will decide to behave or act in a certain way that will lead to their most valued rewards and outcomes. Rewards must be attractive to employees in order to attract, satisfy, and get the most out of the employees (Wabba & House, 1974).

Theory fundamentals

At the core of the theory are two important components:

- *The performance-reward relationship:* the degree to which the person believes that performing at a predefined level will lead to a specific desired outcome (Robbins, 2004).
- *The rewards-personal growth relationship:* the degree to which the rewards offered by the organization satisfy accordingly the changing needs and goals of the individuals (Robbins, 2004).

Reinforcement Theory

Reinforcement theory as a motivation theory

This theory takes a behaviorist approach — a combination of rewards and/or punishments is used to reinforce desired behavior or extinguish unwanted behavior. This suggests that a reward system is principal in creating the desired behavior (Corr, 2008).

Goal-Setting Theory and Tasks

Goal-Setting theory as a motivation theory

The goal-setting theory first developed by Edwin Locke takes another approach: the cognitive one.

Definition: *Goal*

GOAL:
"what a person tries to attain, accomplish, or achieve" (Locke & Latham, 2002).

Theory fundamentals

Goals tell an employee what needs to be done and how much effort will need to be expended. This theory suggests that it is important to establish goals, tasks as a means of internal stimulus and provide feedback to employees conforming to the organization goals (Locke & Latham, 2002). The individuals have a drive to reach a clearly defined end state which often is a reward in itself. The goals itself conform to:

- *Proximity*: A goal can be short-term or long-term. Combination of both types are recommended.
- *Difficulty*: Neither too hard nor too easy. Finding the correct balance between challenging but achievable is crucial in setting a goal that is motivating to an individual.
- *Specificity*: The description of the goal should be clear and it should be known to the user the amount of effort he needs to put in to successfully accomplish that goal. Furthermore, there should be a substantial probability that the user will succeed.

SMART criterias

Additionally, according to Wilson (2003), every effective set goal in a rewarding system should inherit the following features which make up the acronym *"SMART"*:

- *Specific:* There should be a clear connection between the results desired and the actions performed to achieve those results.
- *Meaningful:* Participants need to feel valued for their own accomplishments.
- *Achievable:* Participants need to feel sure that they can achieve the desired outcome.
- *Realistic:* The goal should be challenging yet achievable within the timeframe.
- *Timely:* Participants need to be rewarded "as timely as necessary" to stimulate them to achieve desired outcomes.

Concept of "task"

Similar to this theory, Valle et al. (2006) defines the concept of a "task" that can be a mission, an assignment, obligation, or an atomic process which is impossible to subdivide further. According to this paper, tasks are related to deadlines, production, decisions, projects, implementation, finalization of activities, reports, etc.

Goal-Setting theory with respect to thi thesis

The core of this theory is of major interest for this thesis project. Applying tasks, and goals to be achieved by the participants of the social game, introduced by this thesis project might indeed be the right choice.

Maslow's Theory of Needs

Maslow's theory of needs as a motivation theory

Maslow (1987) is one of the most widely discussed theories of motivation and can be summarized as follows (cf. Figure 2.2):

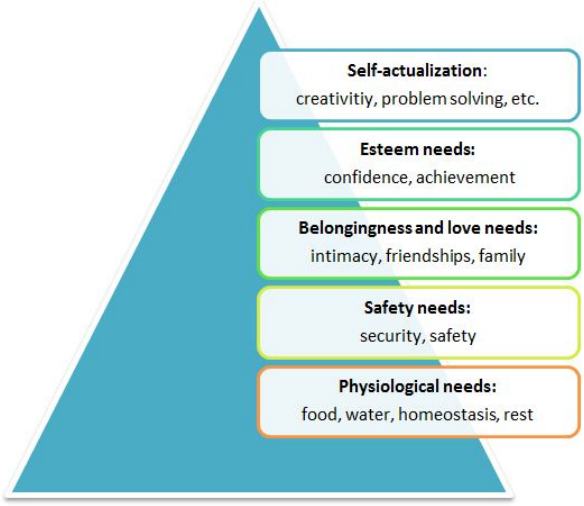


Figure 2.2: Theory of Needs. Source: Maslow (1987).

- *Only unsatisfied needs influence behaviors. Satisfied needs do not.*
- *Arrange needs in the order of their importance from the basic to complex.*

- *The person advances to the next level of needs only after the lower level of needs is at least minimally satisfied.*

- *The further the progress up the hierarchy, the more satisfaction a person will show.*

Theory fundamentals

This theory emphasizes the importance of having a hierarchy of needs, which helps people in the process of self-actualization—a process of continuous growing aiming at achieving individual potential.

Maslow's theory of needs with respect to this thesis

Applying the core of this theory, to the social game of this thesis project, namely categorizing tasks in a "hierarchical" order might be an interesting solution.

Equity Theory of Motivation

Fundamentals of the theory

This theory suggests that individuals compare their job inputs and outcomes with those of other individuals and in case of inequities they behave in such a way to eliminate them (Carrell & Dittrich, 1978).

Equity theory with respect to this thesis

The equity theory of motivation is used to describe the relationship between the employees' perception of how fairly is he being treated and how hard he is motivated to work. This is strongly related with the "ultimatum game experiment", in which players have to divide a given sum of money between them (Thaler, 1988). This is important from a sociological perspective, because it illustrates the human unwillingness to accept injustice and social inequality.

2.3 Reputation and Social Software

Notions of reputation and social software

This section describes the notions of social software and reputation, which aim at creating environments where social interactivity based on reciprocal judgements is promoted.

Definition: *Social Software*

SOCIAL SOFTWARE:
defines web applications that facilitate interactive interoperability, user-centered design, information sharing, etc. It is not just a technical progress of the Web, but its main purpose is to transform already passive users into producers of user-generated content (Koch & Prinz, 2008).

Technical aspect of social software

Additionally, Koch & Prinz (2008) discuss technical aspects, that contribute to the success of social software. Among the analyzed technologies are Ajax, RSS/Atom and Microformats.

Definition: *Reputation*

REPUTATION:
"the general opinion (judgement) (more technically, a social evaluation) of (and by) the public (or a group or a person) toward an entity (person, a group of people, or an organization or brand or object)—as distinct and different from the background (others)—concerning the likelihood of the entity to behave in a certain way in the future under certain circumstances. It is a ubiquitous, spontaneous and highly efficient mechanism of social control (Nadeau, 2007)."

Reputation role

A reputation system keeps track of the contributions that the participants make and keeps a running history of it all (Resnick et al., 2000). It additionally tracks desirable behaviors and recognizes those publicly (Farmer & Glass, 2010).

2.3.1 Competitive Spectrum

The most important part of every reputation system is the degree of *competitiveness* that should be cultivated within the community (cf. Figure 2.3).

The degree of competitiveness of a community

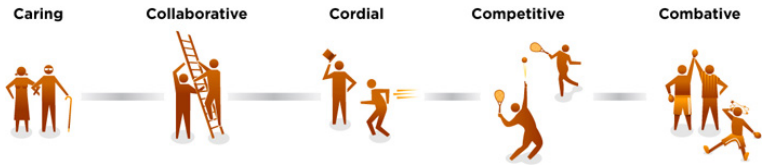


Figure 2.3: Competitiveness spectrum in a reputation-based system. Source: Crumlish & Malone (2009).

Goals	
Caring	Participant's main motivation is helping other participants.
Collaborative	Participants collaborate together to achieve main shared goals.
Cordial	Participants tend to achieve goals, that are not in conflict with other participant's goals.
Competitive	Participants must compete against each other to achieve shared goals.
Combative	Participants must compete against each other to achieve opposed goals.

Table 2.2: Competitiveness degree goals. Source: Crumlish & Malone (2009).

Table 2.2 describes the competitiveness spectrum, based on which decisions are taken about which reputation pattern to apply as described in Section 3.1—"Reputation Design Patterns". The competitiveness degree here is to be understood as the combination of: *(a) personal goals of the community members, (b) actions that community members engage in, and (c) person-to-person comparison* (Crumlish & Malone, 2009).

2.4 Game Mechanics

Several definitions of "Game Mechanics"

Game researchers and designers have provided several definitions of game mechanics that have been used in different contexts and research areas, from analysis (Järvinen, 2008) to game design (Hunicke et al., 2004).

Definition: *Game Mechanics*

GAME MECHANICS:
"something that connects players' actions with the purpose of the game and its main challenges " (Sicart, 2008).

Nevertheless, the definition is not always clear and the following definitions do also exist:

- Lundgren & Björk (2003) define game mechanics as "any part of the rule system of a game that covers one, and only one, possible kind of interaction that takes place during the game, be it general or specific [...] mechanics are regarded as a way to summarize game rules".
- Rouse (2000) defines game mechanics as "the guts of a design document", since they describe "what the players are able to do in the game-world, how they do it, and how that leads to a compelling game experience".
- Jarke et al. (2009) claim that successful game design/mechanics is characterized by "a good game *flow* of continuous engagement and decision-making with immediate feedback and continuous challenges, i.e through multiple game levels".

2.5 Summary

A deep investigation into motivational values

This chapter traversed much landscape about the research fields that mainly affect this thesis. A detailed analysis helps to develop a suitable concept of an effective rewarding system, by being aware of the behaviors that need to be encouraged and discouraged by this thesis project.

Insight into the next chapter

The next chapter discusses a possible categorization of reputation systems and design patterns to be used while designing such systems. The related work, together with the theoretical aspects of this chapter inform the requirements of the concept of this thesis project.

Chapter 3

Related work

"Research is formalized curiosity. It is poking and prying with a purpose."

—Zora Neale Hurston, Dust Tracks on a Road

The main goal of this master thesis is building an effective rewarding system based on features like reputation, game mechanics and social relationships which motivates users to contribute to wiki documentation and keep them coming back to the system. It is not only necessary to have a look at the current state of art of reputation systems but also to the different design patterns applied nowadays in building social websites, aiming at creating an environment that rewards users and encourages them to move toward higher levels of participation (Jøsang et al., 2007).

Importance of state of the art for this thesis

In the following a number of design patterns are discussed that are important for reputation systems. Afterwards a possible classification of reputation systems is discussed. A framework of comparison is given and why such systems are similar or different to this thesis.

Design patterns applicable to reputation systems

3.1 Reputation Design Patterns

Crumlish & Malone (2009) discuss a collection of *social user experience interface* design patterns that are the result of several years of experience in designing social software. The coming sections shortly summarizes some of these patterns that are relevant for this thesis.

3.1.1 Identity

Identity and reputation

Participant identity is the core ingredient of a social system based on reputation. The existence of such an identity makes it possible to build relationships between participants and the own reputation inside the community.

Profile

Profile as a means of displaying reputation

This design pattern is particularly helpful when it one place is needed to show a specific user's contribution, and reputation within the community. This pattern allows users to learn more about other users through their profiles.

Profile with respect to this thesis

Figure 3.1 shows the personal profile page in Mendeley[1]. Other websites that use this design pattern are: Linked in[2] , MySpace[3], Facebook[4], Foursquare[5], etc. This thesis project is based upon reputation, and this design pattern is suitable to display the user's reputation in the social game.

Avatars

Avatars and identity

This design pattern is used when the users prefer visualising their on-line identity. By doing so, they associate

[1] www.mendeley.com
[2] www.linkedin.com
[3] www.myspace.com
[4] www.facebook.com
[5] www.foursquare.com

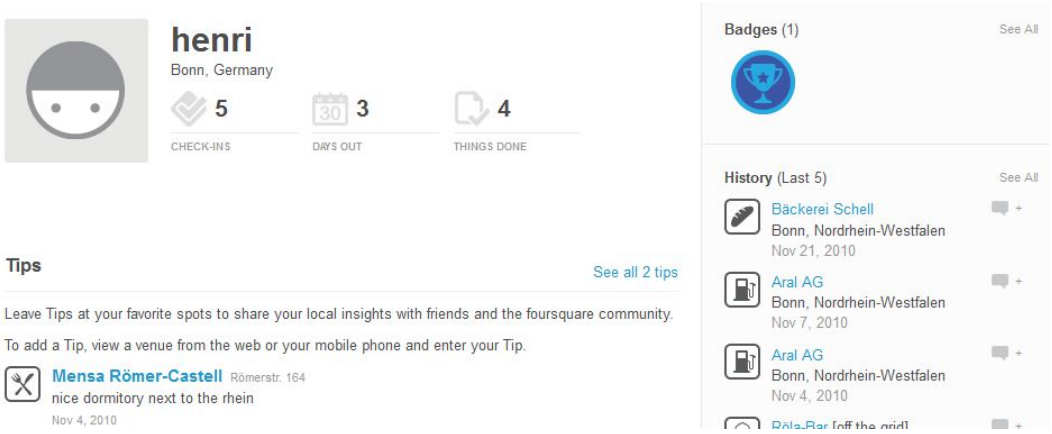


Figure 3.1: Foursquare displays a personal profile for each user. Source: Retrieved 13 Feb. 2012 from: `www.foursquare.com`

their reputation with their person. Users are free to upload any possible image that they believe represents themselves most. In case no image was uploaded, there should be a predefined default avatar.

Figure 3.2 shows the avatar in Gforge. Gforge allows users to upload a personal image, or having default avatars. This design pattern is strongly related to the *Profile Pattern* (cf. 3.1.1—"Profile"), and it suits the needs of a system based on reputation such as in this thesis project.

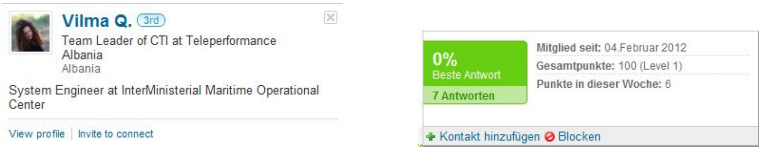
Avatars with respect to this thesis

Figure 3.2: Gforge assigns to each user a default avatar. Source: Retrieved 13 Feb. 2012 from: `http://forge.fit.fraunhofer.de`

Identity Cards or Contact Cards

Identity cards and identity

This design pattern is used in cases when the user requires further information about another participant's identity, without interrupting his or her current task. Figure 3.3 shows how Yahoo Answers and Linked in display contact cards of their users.



(a) Contact card in Linkedin.

(b) Contact card in Yahoo Answers.

Figure 3.3: Contact card examples. Source: Retrieved 13 Feb. 2012 from: `http://linkedin.com`, and `http://de.answers.yahoo.com/`

Personal Dashboard

Recent activities of the community

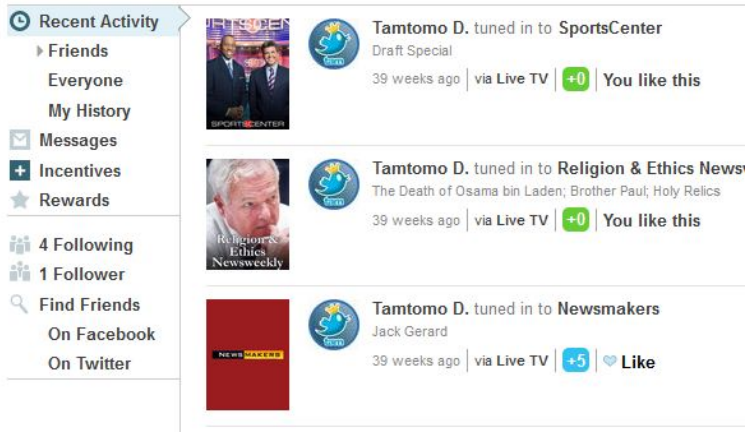
This design pattern allows having a dashboard, which revolves around recent activities of all kinds related to the system. As shown in Figure 3.4, the dashboard redirects the user to the profile of the other members and the recent activities they performed recently.

Recent activities and dynamic awareness

Being aware of user's activities plays a vital role in online systems. Wang et al. (2007) discuss the emergence of many groupware systems that provide users with information about recent activities in different manners. This paper implements the concept of a dynamic awareness system which allows for a balanced notification intensity in cases of huge information overload.

Personal dashboard with respect to this thesis

This thesis project proposes the idea of a social game. In a social environment, recent activities play a vital role for further involvement of its participants. Therefore, this design pattern proves to be suitable for this thesis project.



(a) Recent activities in Screentribe. Source: Retrieved 13 Feb. 2012 from `www.screentribe.com`

(b) Recent activities in Stackoverflow. Source: Retrieved 13 Feb. 2012 from: `www.stackoverflow.com`

Figure 3.4: Personal Dashboard examples.

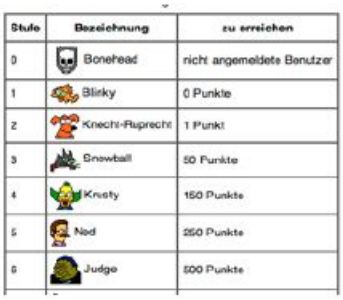
3.1.2 Levels

Levels and reputation

Levels of accomplishment form a basis upon which the own reputation of participants is not only explicitly measured but also shared between individuals in the community. Levels come as named or numbered ones. Figure 3.5 shows examples how levels are used.

Levels with respect to this thesis

Distinguishing different types of users within the social game through levels, is suitable for this thesis project, whose goal is to motivate users and influence their behavior.



Stufe	Bezeichnung	zu erreichen
0	Bonehead	nicht angemeldete Benutzer
1	Blinky	0 Punkte
2	Knecht-Ruprecht	1 Punkt
3	Snowball	50 Punkte
4	Krusty	150 Punkte
5	Ned	250 Punkte
6	Judge	500 Punkte

(a) Numbered levels in Moknow-Pedia.

(b) Named levels in Yahoo answers.

Figure 3.5: Examples of levels. Source: Retrieved 13 Feb. 2012 from `http://answers.yahoo.com`.

Named Levels/Labels

This design pattern is used in cases when the community has different type of members at a specific moment in time and these members need to be distinguished along the competitive spectrum as described in Section 2.3.1—"Competitive Spectrum"; when desirable behaviors have been identified and they want to be promoted; when the individual growth of the user inside the community is tracked; when the community is softly competitive.

Numbered Levels

This design pattern is similar to the *Named Levels* pattern in that it enables users to track their individual progress in the system. On the other hand, it promotes a higher competitive spirit in the community, and it is used when a high number of levels is desired to gauge how far participants progressed within the community.

3.1.3 Awards

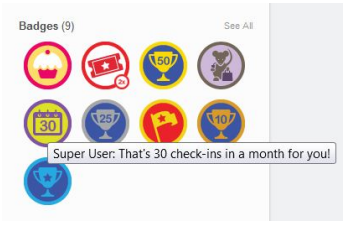
The main goal of awards is to promote positive behavior in the system.

Award as promoters of positive behavior

Collectible Achievements

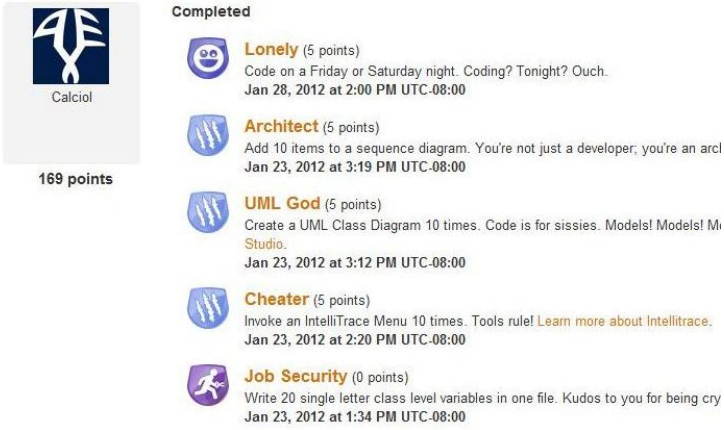
This design pattern allows a user to earn or win awards after reaching some predefined goals. These awards are than displayed either in his personal profile or to other participants, as shown in Figure 3.6. Its goal is to encourage participants in the community to try out all aspects of the system that are being promoted.

Encouraging participation through collectible achievements



(a) Badges in Foursquare.

(b) Badges in Stackoverflow.



(c) Badges in Microsoft Visual Studio Achievements.

Figure 3.6: Examples of Collectible Achievements- "Badges". Source: Retrieved 13 Feb. 2012 from: `www.foursquare.com`, `www.stackoverflow.com`, and `http://channel9.msdn.com/achievements/visualstudio`

Collectible achievements with respect to this thesis

Collectible achievements are strongly related with the concept of *goals* as described in Section 2.2.1—"Goal-Setting Theory and Tasks". This thesis project uses the ideas behind the goal-setting theory, thus this design pattern is applicable.

3.1.4 Peer to Peer Awards

Peer to peer awards and their collaborative touch

This design pattern fosters a more collaborative and cooperative peer-to-peer relationship, by allowing the community participant to give and receive awards reciprocally.

Peer to peer awards with respect to this thesis

Allowing participants to award or rate other participants in this thesis project would be unsuitable because it leads to unfair or subjective ratings.

3.1.5 Rankings

Rankings, as the name says, enable a fast and clear comparison of the participant's performance within the community.

Points

Points as a means to display the reputation

This design pattern determines a competitive community, where its participants involve in competitive activities such as games, and player vs. player competitions. Through this pattern, participants can compare their progress, and accomplishments with other competitors. The system counts up activities that the users engage in and keeps a running sum of the achievements (cf. Figure 3.7) (Farmer & Glass, 2010).

Points with respect to this thesis

This thesis project proposes the idea of a social game. Therefore, this design pattern suits the purpose of every simple game: associating the participant's progress in the game with points.

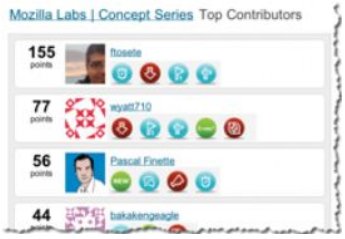
Figure 3.7: ScreenTribe associates "tune-in" of users with points which they can redeem with rewards. Source: Retrieved 13 Feb. 2012 from: `www.screentribe.com`

Leaderboards

Leaderboards promoting a competitive spirit

This design pattern allows users to know who are the best performers in a category of the system, or overall. It can be combined with the *Points design pattern* (as seen in Figure 3.8), in that both of the design patterns promote a competitive spirit within the community and thus are likely to be perceived as an encouragement of competition between users.

(a) Leaderboard in Moknowpedia.



(b) Leaderboard in Stackoverflow.

(c) Leaderboard in Yahoo Answers.

Figure 3.8: Examples of Leaderboards. Source: Retrieved 13 Feb. 2012 from: `www.screentribe`, and `www.stackoverflow.com`

Pitfalls about leaderboards

Based on Farmer & Glass (2010), leaderboards are considered to be harmful if not used carefully. Leaderboards excel in recognizing the performance of the best users, but users being at the bottom of the ranks get demotivated. Additionally, whatever is measured, there is the danger that it is taken way too serious by the participants, and that they do not engage in the activities promoted by the site.

Leaderboards with respect to this thesis

The progress of each participant in the game can be compared through such leaderboards. Therefore, this design pattern is suitable for this thesis project. Nevertheless, as leaderboards are considered to be harmful, a different version of this pattern makes sense.

Top X Leaderboards

Top X Leaderboards—another type of leaderboard

This design pattern is similar to the *Leaderboard Pattern* in that it displays top performers of the community. On the other hand, the ranks of the worst users are not displayed, which removes their sense of shame. Last but not least, users being near to the top X places are motivated to increase their ranks.

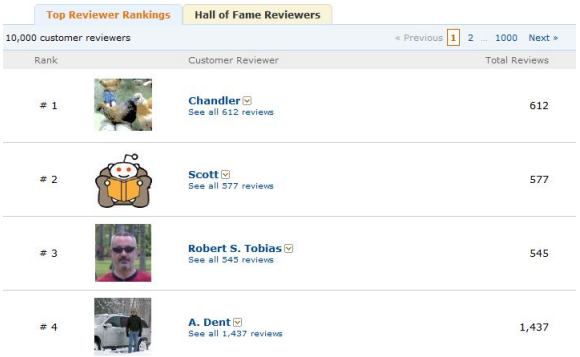


Figure 3.9: Amazon's list of top costumer reviewers. Source: Retrieved 13 Feb. 2012 from: `www.amazon.com`.

After having discussed a set of design patterns used in social interfaces, the next section concentrates on the different types of reputations systems, based upon a comparison framework.

3.2 Types of Reputation Systems

There is an ever-growing literature around web reputation systems which gives a strong sign that this is an important technology. Nevertheless, the systems being proposed and those already proposed unfortunately lack coherence (Jøsang et al., 2007).

Defining a comparison framework

The following describes a framework of comparison for different reputation systems, aiming at highlighting areas, aspects, and design patterns that are important for the purpose of this thesis project:

- *Target users.*
- *Amount of effort needed to achieve a high personal reputation.*
- *Actual benefit of having a high reputation.*
- *Applied design patterns, as described in Section 3.1—"Reputation Design Patterns".*

Categorization principles

Additionally, according to Jensen et al. (2002), reputation systems can be grouped based on:

- *the nature of information given about the object of interest, and*
- *the way how the rating is generated.*

Type of Reputation Systems

Applying this method of categorization, Jensen et al. (2002) groups reputation systems in:

1. Ranking Systems (see Section 3.2.1—"Ranking Systems"),
2. Rating Systems (see Section 3.2.2—"Rating Systems"),

3. Collaborative Filtering Systems (see Section 3.2.3—"Collaborative Filtering Systems"), and

4. Peer-based Reputation Systems (see Section 3.2.4—"Peer-based Reputation Systems").

Further Reputation Systems

According to Farmer & Glass (2010), the following subjects are also related to reputation:

1. Recommender Systems (Resnick & Varian, 1997),

2. Trust Systems (Jøsang et al., 2007), and

3. Search Relevance (Langville & Meyer, 2006).

However, the focus of this thesis is not in using reputation systems to facilitate trust between users in a commercial context, but to establish a status within the community.

3.2.1 Ranking Systems

Notion of Ranking systems

Ranking systems list items, and/or users, according to different criteria and indicators. A classic example of a ranking system are leaderboards/achievements.

Ranking systems with respect to this thesis

Such systems suit the needs of this thesis project, in that ranking systems are suitable for goal-oriented activities and the results displayed through leaderboards are easy to interpret.

Karma Model

Notion of Karma model

Such a model is used to track and create incentives for user behavior. The following forms of karma models can be distinguished (Farmer & Glass, 2010):

- *Participation Karma* measures the amount of user participation.
- *Quality Karma* measures the quality of contributions.
- *Robust Karma* combines the participation and quality karma scores into one score, representing so the overall value of a user's contributions.

Karma model with respect to this thesis

The Robust Karma Model is similar to the system proposed in this thesis in that the overall reputation score of a user of this thesis project would be a combination of different partial karma scores.

Examples of Ranking systems are:

- *Moknowpedia*: A collaborative review system for corporate wiki, which explicitly recompenses users for their efforts and time in contributing to it, rewards them for their participation, and further motivates them to higher levels of participation.
- *Ebay*[6]: An online auction and shopping website in which people and businesses buy and sell a broad variety of goods and services worldwide.
- *Amazon*[7]: A US-based multinational electronic commerce company.

3.2.2 Rating Systems

Notion of Rating systems

When an application gives users the possibility to give an explicit opinion about something, it typically employs a ratings model. Ratings can be handled in a number of ways, such as bars, stars, 10-point scale. These ratings are collected from all users and rolled up as a community average score, meaning that the ratings are global and all users will see the same score for the same object of interest (Farmer & Glass, 2010).

[6] http://www.ebay.com/
[7] http://www.amazon.com/

Ratings-Review Model

In this model, a user gives the target a series of ratings and provides one or more free-text opinions. Each individual part of that review feeds into a community reputation average score (Farmer & Glass, 2010).

The following gives some examples of rating systems:

- *Moknowpedia*,
- *StackOverFlow*,
- *Ebay*,
- *Amazon*, and
- *Digg*[8]: A social news website. Its cornerstone function consists of letting people vote stories up or down, called digging and burying, respectively.

Rating systems with respect to this thesis

Direct ratings of users by users are allowed in some of the above-mentioned systems. This thesis project is about wiki articles, thus a direct rating of users in this context is inapt. On the other hand, rating wiki articles contributes in increasing their quality, thus relevant for this thesis project.

3.2.3 Collaborative Filtering Systems

Notion of Collaborative Filtering systems

Collaborative filtering systems differ from rating systems in that reviewers are matched to the users based on the similarity of past ratings. These systems are more sophisticated than rating systems because they reflect more personally relevant information, are more difficult to understand and expensive to be implemented (Jensen et al., 2002).

[8] http://www.digg.com/

Examples

Examples of Collaborative Filtering Systems are: (1) *Amazon*, and *Netflix*[9]. Users of Amazon receive book recommendations, based on the users' buying behavior or their explicit product ratings. Netflix predicts user ratings for films, based on previous ratings.

Collaborative Filtering Systems with respect to this thesis

The main goal of this thesis project is not meant to assist users in taking every-day decisions based upon past ratings from other users. The proposed system is not a system of trust, but a system of pure personal reputation, where respect and status are its main ingredients.

3.2.4 Peer-based Reputation Systems

Notion of Peer-based Reputation Systems

Such reputation systems, known also as Social Network-Based Recommender Systems (SNRS) give participants personalized recommendations based on social friends influence (He & Chu, 2010).

Peer-based systems with respect to this thesis

This thesis project differs from peer-based systems, because their main goal is to assist users in entering in transactions based on peer recommendations.

3.2.5 Social/Reputation based Systems with inherited "Game Mechanics" Principles

Reputation systems with integrated game mechanics

This section gives an outline of modern reputation-based systems, which integrate *game mechanics* as *part of the users' experience* in those systems.

- *Foursquare*[10]: A location-based social networking site that allows users to "check-in" at places by selecting from a list of venues the application locates nearby. Through Foursquare, people can meet with their friends in the vicinity, and keep in contact with friends from all over the word.

[9] http://www.netflix.com/
[10] http://foursquare.com/

- *StackOverFlow*[11]: A collaboratively edited question and answer site for programmers.
- *Screentribe*[12]: A website which allows users to see what their buddies watch. Upon sharing a video, participants earn points, which can later be redeemed for real life prizes.
- *Gowalla*[13]: A website which allows users to discover their most loved places while sharing the places that mean to them.

Common design patterns

All of these systems use design patterns such as: (a) Points, (b) Leaderboards, (c) Badges (see Section 3.1—"Reputation Design Patterns"), etc. to get their users to continue to perform desired behaviors. Their overall number of users keep increasing drastically, which infers the growing success of such systems.

3.2.6 Comparison Framework

Table 3.1 compares the above-mentioned systems to each other based on the comparison framework described in Section 3.2—"Types of Reputation Systems".

Implications from the comparison framework

An analysis of such a framework reveals the emerging paradigm of integrating a social game as part of the user experience. It also points out the most important design patterns applied in such systems, aiming at increasing the participant's engagement. Finally, all but one (Moknowpedia) system belong to public online communities.

[11] http://stackoverflow.com
[12] http://www.screentribe.com/
[13] http://www.gowalla.com

Systems	Target Users	Amout of effort to be put	Benefit	Applied design patterns
Moknowpedia	Researchers (Team wiki)	Users contribute to the internal wiki to increase their reputation scores.	The best participant of the week is awarded the prize "Hero of the Week". Appearing in the top of the leaderboard, increases the status within the internal wiki.	Weekly achievements, Leaderboards, and Points.
Amazon/Ebay	People, Businesses	Buyers can rate a seller, and sellers can also rate buyers. The average rating defines the degree of trustworthiness within both the sites.	A higher reputation, attracts more users to a specific seller That increases of course the chance of making more money.	Points, Profile, Leaderboards, Avatars, etc.
StackOverflow	Programmers	Participants answer questions, that vary in difficulty to increase their reputation within the site.	Intrinsic drive to appear in the top of the lists; companies hire programmers based on their reputation within the site.	Badges, Leaderboards, Profile, Recent Activities, etc.
Digg	Online Community	The site lets people vote stories up or down. The more votes a story has, the more points gets the users that created that story.	Intrinsic drive of the users to appear among the users with the best stories ever.	Points, Leaderboards, Profile, Avatars, Recent Activities, etc.
Foursquare	Online community	Users need to "check-in" venues. The more venues they check-in the more points they get.	The more check-ins, the more recommendations the user gets. Users get bonuses from checked-in restaurants.	Profile, Leaderboards, Personal dashboard, Badges, etc.
Screentribe	Online community	Users can rate something, comment on it, like it or even tune in. The more comments they submit, the more credits they get.	Every tune in earns the users points that they can redeem for real life things.	Points, Profile, Leaderboards, Contact cards, Personal dashboard, Badges.
Gowalla	Online Community	User need to "check in" at different spots in their local vicinity.	Users receive virtual "items" from check-ins.	Points, Profile, Leaderboards, Recent Activities, Badges.

Table 3.1: Comparison framework of reputation-based systems

3.2.7 Summary

Emerging paradigm of a social game

This chapter summarized the state of the art regarding reputation-based systems as well as design patterns that are applied to such systems. An interesting finding of the performed research is the emerging paradigm of integrating a social game based on competition, recognition, status, and awards as part of the user experience.

Little research on corporate wikis

Another result is that except Dencheva et al. (2011), no example of a reputation-based social-game deployed in the context of a working environment could be found. Thus, evaluating whether such an approach proves to be successful in a corporate wiki is an important contribution of this thesis project.

Insight into the next chapter

Analyzing carefully the results of this research evaluation, along with the theoretical foundations described in Chapter 2—"Theoretical Foundations" contributed to the definition of the final concept of *ReputationForge*, a social game that was created for the purpose of this thesis project. The overall process of formulating a final concept is described in detail in the following chapter.

Chapter 4

Finding a suitable concept

"An invasion of armies can be resisted, but not an idea whose time has come."

—Victor Hugo, Histoire d'un Crime, 1852

A user-oriented design

This chapter describes the gathering of requirements through an elicitation process. Requirements elicitation is a part of the user-centered design process—a design based on: *(a) a clear definition of the target audience, (b) empirical measurement of the feedbacks given from the target users, and (c) iterative design* (Gould & Lewis, 1983). Afterwards, the gathered requirements are analyzed and specified. The overall concept of this thesis is described in detail in Section 4.2—"System concept". It results from the elicited requirements, the problems of the existing system Moknowpedia, the researched theoretical foundations in chapter 2—"Theoretical Foundations", and finally from the design patterns and reputation systems studied in Chapter 3—"Related work".

4.1 Gathering Requirements

In order to ensure a successful outcome, every project has to satisfy the needs and wants of the target users. The fol-

lowing outlines the user-centered design techniques used in this thesis project to elicit users needs.

4.1.1 Target Users/User Environment

This section describes the characteristics of the target users and their working environments.

ebbits

ebbits platform

ebbits[1] is a European Union Integrated Project (IP) within the Framework Programme 7 (FP7) in the area of Internet of Things and Enterprise environment. The ebbits platform transforms the Internet of People (IoP), the Internet of Things (IoT), and the Internet of Services (IoS) into the Internet of People, Things, and Services (IoPTS). The way how it achieves this is by providing a semantic resolution of the Internet of Things transforming so every device into a semantic-based web service.

ebbits features

ebbits is characterized by: *(a) the cooperation and collaboration of many partners (a minimum of 3 partners coming from 3 different countries) with different backgrounds, and (b) a high degree of distribution of the partners.*

ebbits wiki

Members of ebbits project do rarely use their internal wiki. The contribution and participation rate is not satisfactory.

Moknow

Moknowgroup

The Moknow group is a self-organizing team at the Fraunhofer FIT research institute with a flat hierarchy.

Moknow features

There is an overall flexibility in achieving predefined goals. Staff members are mostly researchers, one third working students and two executive managers (Dencheva et al., 2011).

[1] http://www.ebbits-project.eu/news.php

The Moknow group uses its internal wiki Moknowpedia extensively, as a means of everyday information exchange between its members.

Moknow wiki

4.1.2 Requirement Elicitation

Requirement Elicitation encompasses all activities involved in discovering the requirements of a system/describing its functionality, and understanding the application domain. The goal of this phase is a model representing the requirements of the system seen from the user's perspective, which describes what the system should do but not how to do it. It is not just a simple process about gathering requirements, but a highly complex one because different users have very often conflicting changing requirements, and an unclear picture of them (Bruegge & Dutoit, 1999).

Requirement elicitation and its importance

Bruegge & Dutoit (1999) define the following techniques as ways to gather requirements during the "Requirement Elicitation Phase": *(a) Interviews, (b) Observation, (c) Brainstorming, (d) Scenarios, (e) Questionnaires, (f) Evaluating existing systems*.

Techniques of eliciting requirements

Referring to these techniques in the early stages of software development ensures a lower probability of design mistakes in the later stages.

Remarks about Moknowpedia

The results of Dencheva (2010) were satisfactory. Nevertheless the following remarks were pointed out:

Desired improvements for Moknowpedia

- *The rewarding mechanisms applied offer only a temporary motivation,*
- *Needed additional rewards, that are directly related with a participant's performance,*
- *Lack of structured goals and dynamic tips,*

- *Lack of notification emails about the own engagement in wiki.*
- *Lack of a list displaying all the new articles, and the articles not yet rated by the participants.*
- *A combination of punishment and rewarding mechanisms is missing.*

Moknowpedia role in this thesis project

As this thesis project continues the idea of Moknowpedia, the results of this study provide valuable information about the extent to which the current system satisfies user needs and helps identifying potential problems to be avoided, and additional features to be introduced in this thesis project.

Brainstorming Session and Results

Participants' personal data

Eight people, all of them members of Moknowpedia volunteered to be part of a workshop for this thesis project. Ages of the users varied between 25 and 35 years. All of them were male, computer scientists. One participant had a research background in rewarding systems. All of them participate in EU projects.

Brainstorming setting

The session began with a short introduction into the topics of motivation and reputation. Additionally, a specific problem/opportunity statement was drawn up to inform the participants what this thesis project tries to achieve. The statement did not suggest a specific solution to not hinder the idea generation process, being careful so not to bias the participants. The discussion lasted almost two hours and a lot of new ideas were gathered. The focus of the brainstorming session was on identifying a list of extrinsic and intrinsic rewarding mechanisms to be fostered by a system based on reputation and finally to identify a list of system design patterns and requirements. Figure 4.1 shows the main results of the brainstorming session, namely: (a) the basic requirements the rewarding system should fulfill, (b) a list of intrinsic and extrinsic rewards, and (c) the possible problems related to the introduction of rewards.

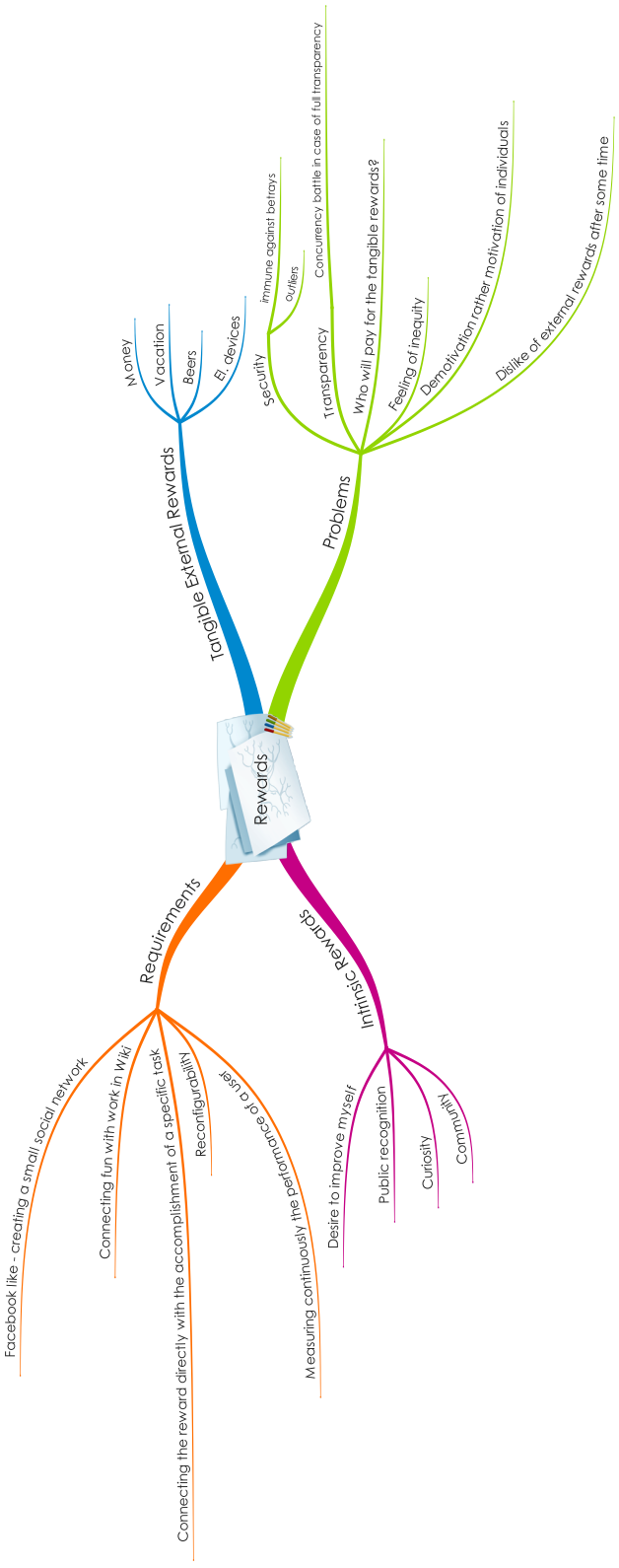


Figure 4.1: Mindmap representing the results of the brainstorming session.

Implication from the brainstorming session

The results of this workshop confirm the idea given in Chapter 2—"Theoretical Foundations" that cash and in general pure tangible external rewards are not the leading driver for employees, thus leading to discouragement if applied for a long time (Schawel & Billing, 2009). Once an individual reaches a level of comfortable living, satisfaction is not likely to get better as a result of such external rewards (Robbins, 2004).

4.1.3 Requirements Definitions

Requirements sources

Based on the outcomes of the brainstorming session, and the theories and the design patterns studied in Chapter 2—"Theoretical Foundations" and Chapter 3—"Related work", an extensive and comprehensive list of requirements was created. However, only the most important requirements have been selected, that respect a degree of completeness to verify the ideas proposed in this master thesis.

Requirements

Requirements for an ergonomic system

While developing a system, design guidelines and standards from the usability engineering field should be followed. They give developers an invaluable information on ergonomic issues. The rewarding system, proposed by this thesis project, should as a precondition be an ergonomic one, which provides a user-friendly interface that allows users to effectively and efficiently work with it. The standards ISO 9241 Part 110 and Part 11[2], define guidelines regarding usability and dialog principles such as efficiency, self-descriptiveness, learnability, task suitability, etc.

[2] http://www.handbuch-usability.de/iso-9241.html

Based on these standards, the following requirements are defined:

- *R1: The system should be easy to use.*
- *R2: The system should be easy to learn.*
- *R3: The way the system is presented is clear and understandable.*

The entertaining side of the system

The workshop results pointed out the necessity of relating fun with the everyday work in the Wiki. Fun is known to be a factor that motivates participants of online communities toward higher participation. Additionally, Prause et al. (2010) and Yetim et al. (2011) point out that *Community* is a major motivating factor (see more in Chapter 2—"Theoretical Foundations". This led to the following requirement:

R4: The participation and activity in Wiki is associated with a social game, as a way of social interaction, promoting fun within the participants.

Public contest/ partial transparency of reputation scores

Based on the workshop results and the results of the studies by Prause et al. (2010) and Yetim et al. (2011), the following requirements related to the concept of reputation were gathered:

- *R5: The social game promotes a public contest based on reputation. Its competitive degree is competitive as shown in Section 2.3.1—"Competitive Spectrum".*
- *R6: There should not be a full transparency of reputation scores.*

The "task" concept/Task requirements

The results of the first workshop and the goal-setting theory as described in Section 2.2.1—"Goal-Setting Theory and Tasks", and the studies by Prause et al. (2010) and Yetim et al. (2011) led to the following requirements:

- *R7: Small tasks are given to the player.*
- *R8: The task difficulty should be neither too easy nor too difficult.*
- *R9: Clear feedback about the progress of each task is given to the player.*
- *R10: The rules on which the game is based are transparent and clear.*
- *R11: Every day actions are made visible.*

Categorization of tasks/Equity between participants

Additionally, based on the *Maslow Theory of Needs* (see Section 2.2.1—"Maslow's Theory of Needs"), and the *Equity Theory of Motivation* (see Section 2.2.1—"Equity Theory of Motivation"):

- *R12: The tasks should belong to different categories.*
- *R13: Users should not feel disadvantaged, or excluded by the game.*

Rewards and triggers

Based on the *Fogg Behavior Model for Persuasive Design* (see Section 2.1.1—"Fogg Behavior Model for Persuasive Design (FBM)"), the *STUFF template for goals* and the study by Prause et al. (2010) (see Section 2.2.1—"Intrinsic vs. Extrinsic Theory"), the following requirements were gathered:

- *R14: The system should increase the motivation by introducing a variety of different proclaimed, and durable rewards.*
- *R15: The rewards should be attractive, and acceptable to the participants.*
- *R16: The system should be based on triggers, which start a change in behavior of the participants.*

Technical requirements

Furthermore, the following technical/architectural requirements are defined:

- *R17: It should be easy to add a new task to the already existing game.*
- *R18: Rules define when a task is completed or not.*

Final concept based on the elicited requirements

The elicited requirements help to define the scope of the solution of this thesis project. In the following the conceptual model is described, which is formulated based upon the identified requirements *R1-R18*.

4.2 System concept

The following sections describe in detail the conceptual model of this thesis project. Afterwards a scenario of use is described which helps clarify how users are supposed to carry out tasks in specified contexts. Finally, a paper and software prototype were developed based on this concept.

4.2.1 Concept

Integration with Gforge and Moknowpedia

The system name is ReputationForge. It is a web-based reputation system. Its service is available to users who are members of the Gforge ebbits project and Moknowpedia.

Main aspects of the social game

This concept addresses the goal of motivating its users to actively contribute to the wiki by providing users with a **social game** (*R4*) based on reputation, which embeds the following categories of incentives (*R14*):

1. *Community/Curiosity:* The drive to know new things regarding the community.
2. *Reputation:* The drive to grow one's status in the community, and one's perception of his importance to the community.

3. *Fulfillment incentives:* The desire to complete a task, assigned to oneself, a friend or the application.

4. *Recognition incentives:* The desire for the praise of others.

Introduction of tasks

The game is broken down into several tasks (*R7*) that need to be successfully completed by its users. A task will mainly require the user to contribute with informative wiki articles, rating such articles, increase their metric scores and ranks related to the scores (*R12*) .

Game objectives

What the game wants to achieve is that the participants go through the tasks available and complete them. The following tries to explain how the game tries to achieve this:

1. The design motivates the following types of users (*R13,R15*):

   (a) **Newbies,** which know little to nothing about the new system and the tasks they need to go through. They need to complete a specific number of tasks to be considered as advanced. As long as they have this status, they get special benefits from the system.

   (b) **Advanced users,** which already know about the game. As such, the system helps them advance through the game.

2. Based on the fact that users are quite busy, the design gives the user the full choice when to start competing for a task. It is totally the user's decision when to decide to accept a task. Still, the design **pushes the user to accept a new task by**:

   (a) giving him a **clear description and exact effort he needs to put in** (*R9,R10*). This is based upon the idea that busy people want to know exactly how much work they need to invest in to achieve specific goals.

   (b) making clear to the user that in case he accepts the task and successfully completes it, he would

increase his personal reputation score by getting karma points (*R15-R16*, see also 3.1.5—"Rankings"). Karma points are to be understood as virtual rewarding units that are given to the user, upon successful interaction with the system, in our case: successful completion of a task. The reputation of the user would then be defined by the overall sum of these karma points and further define the status and rank (a quantitative number to define the best users in a competition (*R5*) of the user.

(c) proposing **balanced tasks**, (*R8*) neither too easy nor to difficult.

(d) making clear to the user that only if he accepts the task and successfully completes it, the **remaining tasks will be unblocked** (*R14, R15*). This stimulates the curiosity of the users which ask themselves what comes next to the game.

3. The design **helps the user to successfully complete an already accepted task by**:

   (a) giving the user a clear **feedback** of how much effort is left.

   (b) reminding the user about the **karma points** (*R16*) he would win upon successful completion.

   (c) giving the user **information about other users** (*R16*) which are currently competing or already competed for that task. By doing this, not only a sense of competitiveness is leveraged but also users are more likely to complete the task if they recognize via the system that others are performing the same task along with them.

4. Through successful completion of tasks, the user increases his reputation, which defines his status and power. A higher reputation means a higher status for the user which increases his satisfaction and motivation to use the game (see Chapter 2—"Theoretical Foundations". The design proposes a personal profile (see 3.1.1—"Profile"). Besides the karma points, it includes the following:

(a) **Recent activities (*R11*) and personal achievements of the user** (when a user starts, completes or fails a task, becomes an advanced user) (see 3.1.1—"Personal Dashboard" and 3.1.3—"Awards"). This lies behind the idea that the user will be more motivated to perform a task if she can observe others performing the behavior and see the outcomes of their behavior.

(b) **Ranking position of the user**, which is only visible to the owner of the profile (*R6*). This is done to not contradict the idea of having a top 5 users leaderboard (as explained below), which displays only the ranks of the 5 best users (see 3.1.5—"Rankings").

5. By introducing **ranking lists/leaderboards**, the design gives a competitive touch (*R5*) to the game. The reason is that users will have a greater motivation to progress through the game if they can compare their performance with the performance of the others. Based on the two different types of users, the introduction of: a) *a top 5 users leadboard* and b) *a leaderboard of all newbies* is proposed. This addresses the following:

   (a) All the users would be driven by the desire to appear in the top 5 list.

   (b) Users that are at the bottom of the list would not feel ashamed about their actual score (still having the inner desire to progress) which would not be the case of a ranking list displaying all the users participating in the game.

   (c) Competition between newbies is increased. To avoid a possible demotivation, the leaderboard of all newbies is not visible to advanced users (*R6*).

6. ReputationForge increases its presence and thus the probability to start a change in its participants behavior by introducing the following triggers:

   - Changing the main page of ebbits web site in Gforge (see Section 5.4.3—"Creating a Gforge Plugin").
   - Regularly sending emails to users, regarding their activity in Wiki.

7. As a last consideration, based on the results of Section 4.1.2—"Remarks about Moknowpedia" the design introduces a punishment task, which stops the normal progress of the game, in case users fail to maintain some minimal requirements (for more detail see Section 5.4.5—"Punishment Task"). Only upon successful completion of the punishment task, the user will be able to continue playing the game as normally.

4.2.2 A Scenario describing the Concept of ReputationForge

Wednesday morning, Bill is drinking a coffee in his office in Denmark, while he receives an email. He was chosen to take part in ReputationForge, a social game that shall motivate him to participate more actively in the internal wiki of the project he is a member of. "There are a bunch of tasks, that are waiting to be completed from you. Their level of difficulty vary and they belong to different categories. Upon completing the tasks, you will be awarded some points, which will increase your reputation inside the wiki". The idea of embedding pleasure with the work in wiki attracts him.

Bill does not need to do anything in order to install the game. He just needs a browser. ReputationForge is displayed as an additional tab in the main page of his project. He clicks it, and starts navigating in the new system. ReputationForge is self-explanatory, and after some time he decides to start the first task, which is colored green. The remaining tasks are colored gray. He cannot open them. This very task wants him to "Rate two articles". He has no time restriction in doing that. Additionally, to help him complete this task, ReputationForge displays a list of the articles that he did not yet rate, and additionally forwards him to those articles. After a careful reading, he rates the required two articles with the help of the embedded rating form. After a while, he receives an email notification: "Congratulations. You just completed a task. It is now time for a new one. Remember, completing tasks increases your reputation within the wiki". Additionally, he notices that this activity was made public to the other members and that he belongs to the best users. Unfortunately, he does not have time to start a new task because he has an important meeting.

The next days are very busy for him, and he forgets about ReputationForge. ReputationForge sends him an email-notification "You did not complete any tasks recently", aiming at attracting him back to the system. Indeed, Bill logs in again and starts the second task.

Jana, a member of the same project, while working in her office in Germany, did not yet try ReputationForge. ReputationForge sends her an email-notification: "You belong to the worst users. Try to start a task and increase your reputation within wiki". She logs in and starts her own task. The game shows her also the users that completed or are completing the same task as her. She finds this thriving. After completing her first task, she can choose between the remaining tasks. She does not know which task to start. Having a look to the recent activities, she sees that Tim successfully completed Task "Contribute with two articles". She decides to start that task also. Her reputation increased and soon she will appear among the best users in ReputationForge.

4.2.3 Prototyping Techniques

Role of fidelity prototypes

Every user-oriented system uses prototyping techniques that enables developers to explore the application's functionality and interaction model with the target users. Hakim & Spitzer (2000) evaluate different approaches to prototyping based on the degree of fidelity which they define as follows:

Definition: *Fidelity*

FIDELITY:
the degree of closeness to the "depth, breadth and finish of the intended product".

In the following two widely used prototyping techniques during the *Requirement Elicitation Phase* are discussed:

Low Fidelity Prototyping

The low fidelity approach

Example of low fidelity prototypes are rough paper or pencil sketches or User Interface (UI) dialogs, which are not that detailed allowing designers and target users to focus on important high-level UI design. The disadvantage with this technique is that the dialog sequence is hard to be conveyed.

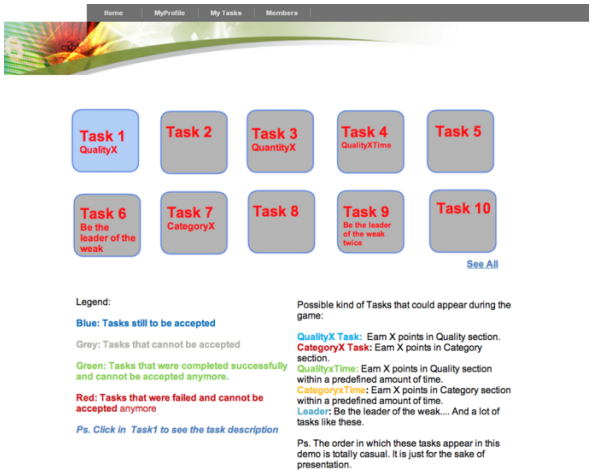
Medium Fidelity Prototyping

The medium fidelity approach

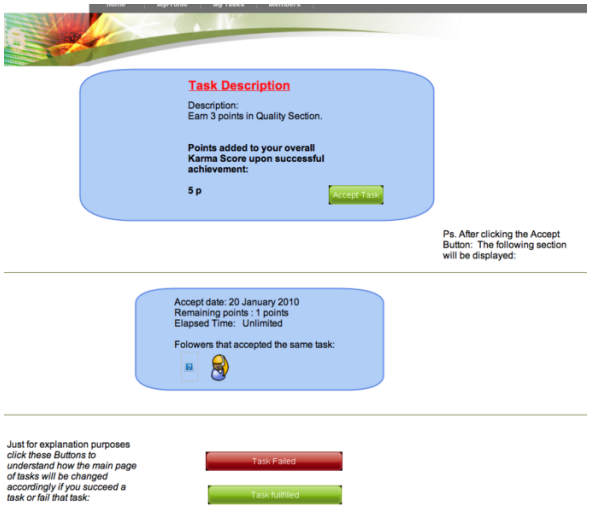
This approach takes designers very near to a true representation of the user interface of the future system. A medium fidelity prototype is more detailed, precise and interactive. It mocks some aspects of the final UI. A disadvantage of this technique is that target users focus on design details and may overlook larger problems.

Software prototype of this thesis project

A medium fidelity prototype[3] was developed and presented to the target users. The most important milestones of the conceptual model and user interface were explained. Figure 4.2a, and 4.2b show some screenshots of the software prototype.



(a) Software prototype—Overview of tasks



(b) Software prototype– Description and feedback about a task

Figure 4.2: Screenshots of the software prototype

[3] http://www.youtube.com/watch?v=UZHVlX6T4Is

4.3 Concept Validation

Having formulated a detailed concept for this thesis project, the next step was to validate it. Eight participants, where three of them were students, were again briefed on the goals of this thesis project. The scenario described in section 4.2.2—"A Scenario describing the Concept of ReputationForge" was presented to them. Its aim was to make the concept of this thesis project clear to them. Additionally, a software prototype (refer to Section 4.2.3—"Prototyping Techniques") was demonstrated to them.

Procedure

It is important to mention that this section aims at finding out the preference towards ReputationForge concept before starting its implementation. The detailed empirical evaluation of ReputationForge, along with its concept and motivational values is described in Chapter 6—"Outcome Evaluation", and Chapter 7—"Implementation Evaluation".

Real evaluation of ReputationForge is carried out later

4.3.1 User Questionnaire

After this, each participant was given a web based questionnaire (cf. Appendix A—"User Evaluation Questionnaire") that asked him to give his overall reactions to the concept. The majority of the questionnaire items were based on a five-point Likert scale. Each of the items had five potential choices — "strongly agree"(1), "agree"(2), "neutral"(3), "disagree"(4), "strongly disagree"(5). The collected data was than analyzed through descriptive statistics[4].

5-point Likert scale items

Response options

General Questions

Participants were given questions that aimed at identifying their profile, and their degree of familiarity with different reputation systems in the market.

[4]For detailed information about descriptive statistics see: `http://www.le.ac.uk/bl/gat/virtualfc/Stats/descrip.html`

Participants' "profile"

Three out of eight participants had a HiWi position in FIT. Two of them were females. Seven out eight users strongly like challenges in their everyday work. All the participants are motivated by new ideas in their everyday work. Six out of eight participants do not only finish their individual tasks, but also contribute to the team. Five out of eight participants need a clear visible path to complete assigned tasks. Five out of eight participants like to complete tasks that are neither too easy nor too difficult.

Results

Based on these results, there is an obvious tendency of users toward balanced tasks, with a clear description of what needs to be done. This corresponds to the conceptual model of this thesis project.

Degree of familiarity with Reputation-based Systems

Seven out of eight participants were familiar with Amazon, and Ebay. Five out of eight participants were familiar with Stackoverflow. None of the participants were familiar with systems like Foursquare, Digg, Screentribe, which introduce the concept of a social game as part of their users' experience.

Analyzing the Likert-Scale Items

Rated statements by the participants

Table 4.1 shows the detailed descriptive statistics of the following statements:

Statement 1: Systems like Moknowpedia would be used by the manager to control the performance of participants.

Statement 2: I am overall satisfied with Moknowpedia.

Statement 3: It is easy to get a first understanding of the new system ReputationForge.

Statement 4: Regular emails sent by ReputationForge would increase its presence.

Statement 5: Advertisements would increase the presence of ReputationForge.

Statement 6: The rewards/incentives promoted by the so-

cial game in ReputationForge overall motivate me.

Statement 7: I feel excluded by this concept.

Statement 8: Overall I am satisfied with the concept.

Statement	n	Mean[a]	StDev	Median	Range
St. 1	8	3.875	0.641	4	2
St. 2	8	2.5	0.535	2.5	1
St. 3	8	1.625	0.916	1	2
St. 4	8	2.625	1.408	2.5	4
St. 5	8	2.75	1.165	3	4
St. 6	8	2.125	0.641	2	2
St. 7	8	3.875	0.991	4	3
St. 8	8	2.25	0.886	2	2

[a]Response values range from 1 (strongly agree) to 5 (strongly disagree).

Table 4.1: Results of the concept validation

Figure 4.3 displays the overall descriptive statistics shown in Table 4.1 through boxplots.

Displaying statistic results through boxplots

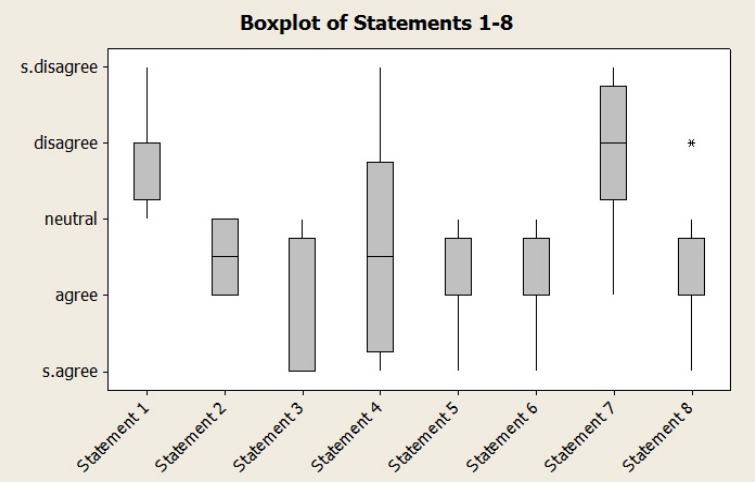


Figure 4.3: Boxplots describing the results of the concept validation

Controlling performance

75% of the participants think that the manager would not use Moknowpedia as a means of controlling their performance. Only one participant was afraid of such a side effect. Based on this outcome, it is assumed that ReputationForge which continues the idea of MoknowPedia, is perceived as a system which promotes and does not control the participants' performance.

MoknowPedia Motivation

50% of the participants were motivated by Moknowpedia , meanwhile the other half had a neutral feeling on this point.

Understandability of the new system

75% of participants found it easy to understand the concept of ReputationForge. The remaining 25% of the participants are undecided. Based on this outcome, it is assumed that the participants did not overlook major milestones of ReputationForge conceptual model.

The aim of statement 6 and 8 was to gain an overall feeling about how motivated the participants from ReputationForge feel and how much they like its concept. The focus was not in each specific reward promoted by the ReputationForge concept, but in finding a tendency of the participants. For a complete detailed evaluation of the concept of ReputationForge, refer to Chapter 6—"Outcome Evaluation", and Chapter 7—"Implementation Evaluation".

Overall feeling towards ReputationForge concept

75% of the participants like the concept of ReputationForge. 12.5% of the participants is undecided, while one participant does not like the concept (shown as an outlier in Figure 4.3). His main concern is the need of emails, to increase the system presence. According to him, *"if a tool is useful, it is used without the need of further triggers. If a tool is not used, this is because it is not useful enough"*.

Overall motivation

75% of the participants believe ReputationForge would motivate them in general. The remaining 25% had a neutral opininion.

Preference of participants

75 % of the participants feel motivated by the concept of ReputationForge compared to 50% of the participants that are motivated by the old system Moknowpedia. Thus, the participants prefer ReputationForge to Moknowpedia.

Does the concept exclude participants?

75% of the participants do not feel excluded by the concept. 12.5% of the participants are undecided. One participant felt excluded. That was the same participant that disliked the concept.

Effectiveness of triggers

Statements 4 and 5 discuss the effectiveness of emails and advertisements at increasing ReputationForge presence. 50% of the participants agree about the effectiveness of email notifications. 25% of the participants are undecided, whereas the rest rather disagree. Based on this outcome, it is hard to say whether such a trigger would prove to be effective. Still, since the median of all answers is 2.5, it is decided to apply it in the upcoming prototype. Regarding advertisements, 75% of the participants believe that they would increase the presence of the new system. 25% of the participants are undecided. Thus, it is decided to apply advertisements in the upcoming prototype.

Concept improvements

While asked for any kind of concept improvements, the participants do think the original concept is better and would like to try a first prototype based on that concept.

4.3.2 Summary

Summary

This chapter described the main steps followed to elicit requirements, build a suitable concept based on the requirements, and furthermore validate it.

Insight into the next chapter

The next chapter describes the implementation of this concept, through a high fidelity prototype.

Chapter 5

Approaching a concrete implementation

"If you can dream it, you can create it."

—WALT DISNEY, "The Man Behind the Mask"

Technical tools needed to implement a first prototype

This chapter focuses on designing a prototype, which suits the concept defined in Chapter 4—"Finding a suitable concept". To fully realize this concept, different technical tools are needed, which will be presented in the following sections.

5.1 Gforge

Gforge platform

Gforge (cf. Figure 5.1) is a collaboration tool for managing software projects[1]. It assists in the management of the entire development life cycle using a Collaborative Development Environment (CDE)[2] where software development project's stakeholders collaborate with each other, no matter what timezone or region they are in, to discuss, document, and produce project deliverables.

[1] http://gforge.org/gf/

[2] http://www.alphaworks.ibm.com/contentnr/cdeintro

Figure 5.1: GForge AS is an extensible platform with an intuitive interface that ties together a huge toolset controlled by a centralized permission system and maintained automatically by the system. Source: Retrieved 13 Feb. 2012 from: `gforge.com/gf`

Gforge technical features

Gforge Software is PHP-based, has a modular architecture and is open source. Gforge AS can be installed on any cleanly-installed Fedora Core 4-7, Red Hat Enterprise Linux 4-5, Debian Etch, Ubuntu 6, or CentOS 4-5. SuSE 10 and Solaris 10 can also be used, although it requires more effort to configure all the required packages.

5.1.1 Gforge Projects

Role of a project in Gforge

A single Gforge installation hosts several projects, which users that have a Gforge account can join. A project in such a collaborative environment can be thought of as an or-

ganizational entity combining people and activities. Most things in Gforge are related to a specific project. Once joined to the system, members can browse around the project.

Each project has a Summary page, highlighting:

- *Recent News*, which displays the news about the project.
- *Latest Activities*, which displays what users did recently in the project.
- *Activity*, which gives a visual graph of the project's progress.
- *Developer Info*, which contains a list of all the members of the project.
- *Trove Categorization*, which displays details related to the project.

5.1.2 Gforge Modules/Plugins

Gforge has a modular architecture

Modules or plugins in Gforge define the modular architecture of Gforge. They extend the basic functionalities of Gforge. Through the help of these additional modules Gforge adapts itself to the user or project needs, making it highly flexible.

Gforge offers several useful built-in modules/plugins such as: *(1) issue tracker mechanisms, (2) task manager, (3) embedded wiki, (4) Subversion (SVN) repositories, (5) and some modules like News and File Releases.*

Gforge Wiki

Collaborative feeling through wiki

One important feature of Gforge that makes it highly collaborative is the embedding of a built-in Wiki. Figure 5.2 and Figure 5.3 show relevant parts of the PostgreSQL

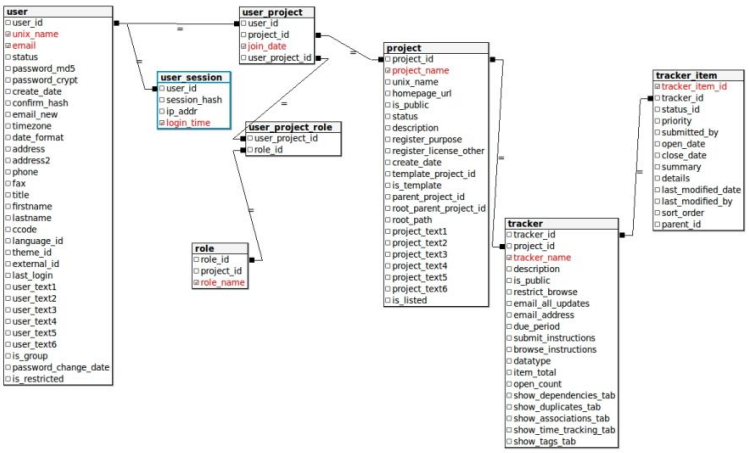


Figure 5.2: Part of the PostgreSQL database schema of Gforge that is relevant for this thesis project.

database schema of Gforge and relevant parts of the relational schema of the Gforge wiki plugin, respectively.

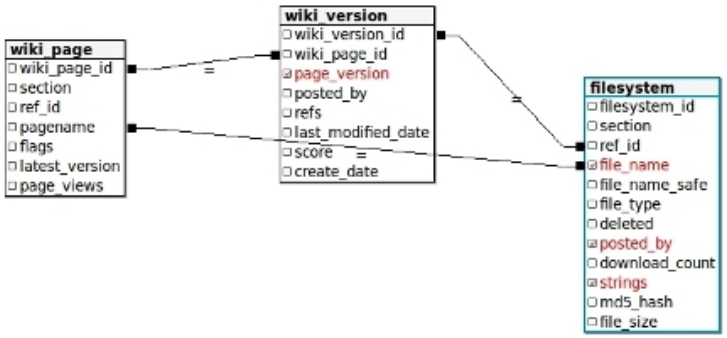


Figure 5.3: Part of the wiki relational schema that is relevant for this thesis project.

Analyzing these schemes is important for this thesis project, because this data need to be read from the PostgreSQL Gforge database for the successful implementation of a first prototype.

5.1.3 Gforge Plugin Architecture

As described above, Gforge has a modular plugin architecture which makes it very easy for developers to add new plugins. There are three types of plugins:

Plugin types

- *Site plugins*, which are available as a single instance.
- *Project plugins*, which are available in multiple instances, one per project. The administrator of a Gforge project can activate or deactivate the plugin through his admin page.
- *User Plugins*, which are also available in multiple instances, but one per user. If needed, each user can activate or deactivate the plugin.

Plugin Development

GForge Group (2011) describes the whole plugin development in Gforge. The following is a brief summary of it.

The Gforge convention for each plugin directory is as following:

Gforge plugin directory

- `bin`: containing non cronjobs executable scripts,
- `conf`: containing configuration files,
- `cronjobs`: containing cronjobs executable scripts,
- `languages`: containing internationalization files,
- `lib`: containing all the plugin classes and functions,
- `litb/html`: containing the main plugin template class,
- `wwwlib/lib`: containing the web interface scripts, and
- `wwwlib/admin`: containing the web scripts for the admin UI.

GFPlugin interface

Every new plugin needs to implement some basic methods that are defined in the `GFPlugin` interface located in `/lib/plugin/GFPlugin.php`:

- `getType()`: which returns one of the plugin types as described above.
- `getName()`: which returns the plugin name.
- `getListeners()`: which returns an array of predefined actions that the plugin will handle. Gforge AS notifies the plugin, whenever a new action is registered, and the plugin itself handles it.
- `handleAction(listenername,params)`: handles every registered action.

This interface extends the important `GFProjectPlugin` interface, located under `lib/plugin/GFProjectPlugin.php`, and adds some methods that are commonly used in project plugins.

Web interface scripts

Scripts that are responsible for the web interface are stored in the directory: `plugins/pluginname/wwlib/lib`. A specific file `handler.php` is first called, which is a dispatcher and maps the URL *action* parameter to a given script. The action itself is handled by the file `index.php`.

Plugin installation steps

To install a new plugin, the following steps are necessary:

- Store the plugin subdirectory inside the Gforge Plugin directory.
- Update the plugin table of Gforge Database by adding an entry into it.
- Update the configuration and language cache files.

  After installation of a project plugin, Gforge appends it as a tab in the main project table modules as seen in Figure 5.4.

Figure 5.4: Gforge appends any new project plugin as a tab in its table module

5.2 CollabReview

CollabReview is a reputation system for collaboratively written text. It was originally implemented, with the main goal of increasing the quality of source code during the software development process.

5.2.1 Experience Scores

Collabreview computes "Experience Scores"

CollabReview allows users to give anonymous reviews of source code, and/or Wiki articles and additionally computes for each user a list of experience scores/points, defined as "Erfahrungspunkte" in Dencheva (2010). Some of the most important computed scores by this tool are:

- *quality points,*
- *contribution points,*
- *and review points.*

Description of the experience scores

The first two scores are formed by the combination of values such as the article weight, the authorship and the average rating of an article. These scores represent the average article quality users contributed to, and the sum of contributions of each user respectively. They are proportionately

distributed between users. *The third score* represents the number of articles rated by a user.

In doing so, this tool triggers a very useful quality concurrence between users. More information about this tool, can be found in Dencheva et al. (2011).

5.3 Vaadin Framework

The Vaadin platform

Vaadin is a framework used for the development of web-based Rich Internet Applications. It is open source and has a server-side architecture. Asynchronous JavaScript and XML (AJAX) technology is used on the browser-side. Vaadin is built on top of the client-side and can be extended with the help of Google Web Toolkit (GWT). The following summarizes the basic principles of Vaadin along with its modular architecture, as described in Grönross (2011).

5.3.1 Vaadin Architecture

Brief description of the Vaadin architecture

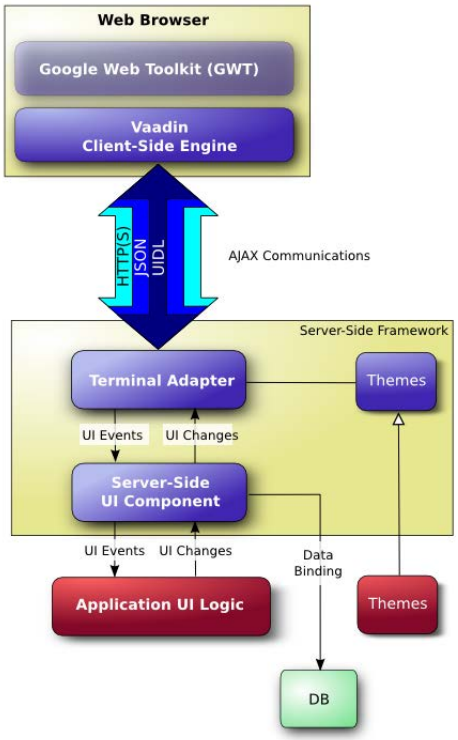
A Vaadin-based application runs as a servlet in a Java application server, serving Hypertext Transfer Protocol (HTTP) requests. The terminal adapter functioning in the background, receives client requests from web browsers and renders changes to the UI components made by the application logic in the web browser. The terminal adapter renders the changes to the UI components made by the application logic in the web browser.

Application class

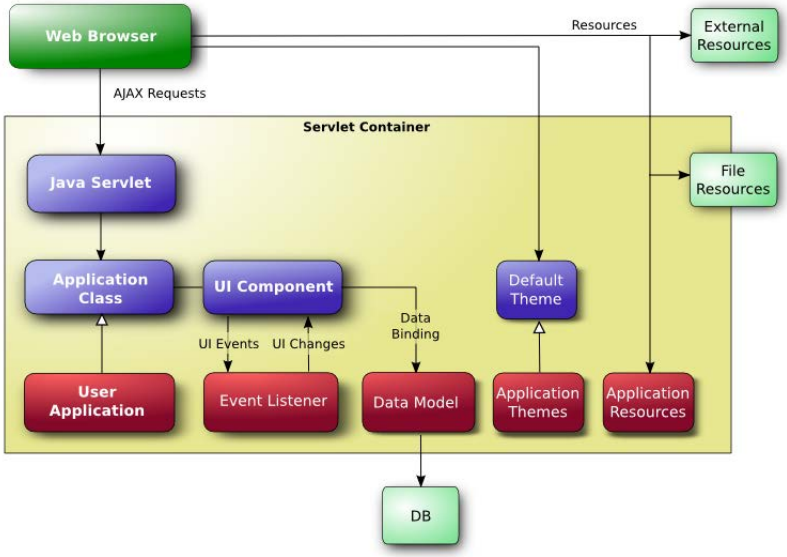
The top level of a user application consists of an application class that inherits `com.vaadin.Application`. Its main function is creating UI components, receiving events related to the UI components, and finally making necessary changes to the components.

Figure 5.5a and 5.5b describe the major parts of the architecture and their function as follows:

- *UI Components:* Vaadin offers developers a full set of



(a) Vaadin Architecture—Interaction between different components. Source: Grönross (2011).



(b) Vaadin Architecture—A second depiction. Source: Grönross (2011).

Figure 5.5: Vaadin Architecture

UI components. It also allows developers to define custom components. Most components are bound to a data source.

- *Client-Side Engine:* The Client-Side Engine of Vaadin is responsible for rendering UI changes in the web browser using GWT. This is done through the help of the server-side Terminal Adapter using the User Interface Defintion Language (UIDL)[3], a JavaScript Object Notation (JSON) based language.

- *Terminal Adapter:* It is responsible for rendering the UI Components, which communicate their changes to it. These changes, known as events are communicated to the terminal adapter through the webserver as asynchronous AJAX requests. The terminal adapter finally renders the UI changes in the user's browser.

- *Themes:* Vaadin applies a Model-View-Controller (MVC) design pattern, by separating between presentation and logic. The UI logic is handled as pure Java code, the presentation is defined in Cascading Style Sheets (CSS) based themes.

- *Events:* While users of a Vaadin application interact with its UI components, events are created, which are first processed on the client-side and than passed to the application.

- *Data Model:* Based on the inherited MVC design pattern of Vaadin, it provides to developers a data model for interfacing data presented in UI components. The flexibility that the data model gives consists in the fact that UI components can update the application data directly, without the need for any control code.

[3] `http://www.uidl.net/`

5.3.2 Extending Vaadin Functionality through Add-Ons

Vaadin has a modular architecture, extendable through add-ons

The Vaadin architecture is modular which gives developers the ability to add specific functionalities to it, through so called Add-ons. Both commercial and free components that can be used in open source projects exist under various licenses.

A rich collection of add-ons can be found in Vaadin directory at `http://vaadin.com/directory/`. To install an add-on the following steps are necessary:

- Download the add-on package from the details page.
- Copy the Java Archive (JAR) file inside the downloaded add-on package to the web project under the `WEB-INF/lib` directory.
- Compile the included widget set, if needed (the client-side implementation).

Navigator7 Add-On

Navigator7—an add-on that helps the purpose of this project

Vaadin lacks the concept of a navigation system which is crucial for successful web oriented applications and also for the successful realization of this thesis concept. The add-on Navigator7[4] helps achieving this gap of Vaadin. Through this add-on Vaadin window management is completely revised by supporting multi-window (known as pages) applications, tabbed navigation and web-style applications in an easy and logical way.

Navigator7 Application Programming Interface (API) include:

[4] `https://vaadin.com/directory#addon/navigator7`

- the concept of application level window that includes a page,
- clickable links between pages which make the navigation between pages easy,
- templating around pages (as header/footer/...),
- support for fixed and fluid designs,
- a rich analysis of Uniform Resource Identifier (URI), with entry points for the Plain Old Java Object (POJO) persistency code (i.e Java Persistence API (JPA), Hibernate),
- parameter injection (in page object's fields), and
- a general interceptor (filter) mechanism.

The main classes[5] of this add-on are as follows:

AppLevelWindow

The `AppLevelWindow` introduces in Vaadin an application level mechanism. It extends the Vaadin `Window` class, and plays the role of the original `Application.mainWindow`.

NavigableAppLevelWindow

`NavigableAppLevelWindow` is the descendant of `AppLevelWindow` and it is responsible for defining the page template, and finally displaying it. It keeps a reference to the `Navigator` and the current page.

MyWebApplication

`MyWebApplication` specifies a list of all page classes that need to be registered.

Navigator

The `Navigator Class` is responsible of handling URI changes that are fired by an `UriFragmentUtility` part of Vaadin API. Through the help of the class `NavigatorConfig`, the page name, i.e *Profile* (`ProfilePage.class`) is mapped into the URI *Profile/id=101*.

UriAnalyzer

Through `UriAnalyzer`, the application becomes bookmarkable. It is responsible for the dynamical mapping

[5] Retrieved 15 February 2012 from: `http://code.google.com/p/navigator7/wiki/ArchitectureOverview`

of application objects and creating of the URIs. To realize a URI Profile/id=123 based on the application objects `ProfilePage.class` and int 123, the following is necessary:

- `UriAnalyzer` extracts the page name.
- `ParamUriAnalyzer` manipulates parameters as strings.
- `EntityUriAnalyzer`: translates the parameter value into an object.
- `MyUriAnalyzer` defines how to find entity.

NavigableApplication

`NavigableApplication` implements the `ThreadLocal` pattern to get an instance of the main application and main window from anywhere.

5.4 Implementation of the Prototype

The following outlines the basic steps towards the final prototype of ReputationForge, namely: *(a) logging-in and logging-out, (b) embedding the review form in Gforge, (c) creating a Gforge plugin, (d) extending Collabreview through Vaadin, and (e) email notifications and punishment tasks*.

5.4.1 Logging-in and Logging- out

Single Sign-On and Single Sign-Out

As already described, there will be three different systems that will be involved in the successful realization of the first prototype: Gforge, CollabReview and a Vaadin-based application. To give the users the feeling that they are using a single system instead of three different systems, it is very important to connect them through a Single Sign-on (SSO) mechanism. The Open Group defines SSO as a mechanism of a single action of user authentication and authorization that permits the users to access all the other software systems (in this case Vaadin and CollabReview), where they

have granted access without the need to login at each of them (Clercq, 2002). In a similar way, when the user signs out from Gforge, a single sign-off mechanism terminates access to the other systems.

Connection between Gforge and CollabReview

Not logged-in users in Gforge do not have access to project-specific data. In order to use Gforge Wiki, to commit source code, and add new issue trackers, users must log-in to Gforge. The same is also valid for CollabReview. Additionally, it is important that information exchange should occur between the two systems.

Different approaches to SSO

There are some well-known practical SSO models as described in Dencheva (2010), where advantages and disadvantages of each systems are described. The approach chosen by the author here is a central instance server that makes the authentication. A similar approach will be used in this thesis project.

The logical sequence of events is described as follows:

1. The user logs in to Gforge.
2. The credentials are checked from Gforge internal logic whether they exist in the Gforge database.
3. If the credentials are correct, the user gains access to Gforge; otherwise he is prompted to login again.
4. Immediately after the successful login in Gforge, the hashed credentials of the users are sent to CollabReview, through a log-in JavaServer Pages (JSP) file.
5. A request to the CollabReview database is sent, to check whether the user exists or not.
6. In case the user does not exist, CollabReview logic creates a new user with the given credentials.
7. The user logs in and CollabReview assigns him a session id. At the same time, this session id is forwarded to the Vaadin-based application, which is a CollabReview extension. At this time, the rating form, and the Vaadin-based application *ReputationForge* are available for the user.

5.4.2 Embedding the Review Form in Gforge

Review Form Requirements

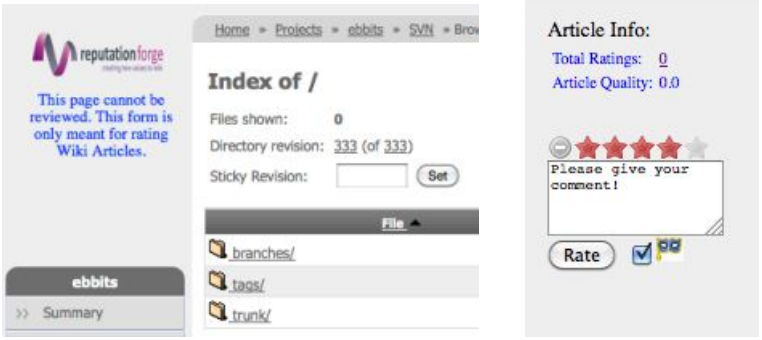
Dencheva (2010) developed a tool that allows users rating the quality of Wiki articles. Such a form, needs to be integrated in Gforge also. The form should be highly visible to increase the probability that users will rate articles. Additionally, the form should visible only to logged-in users and in specific projects in Gforge. To satisfy such requirements, the following implementation steps are required:

Changes to the Main Gforge theme

`GFTheme.php` defines the template of Gforge. It implicitly defines the look and feel of the *Summary Page*, which is the most important page of each project in Gforge. Therefore, it makes sense to embed the form in this main page. To realize that, `GFTheme.php` was changed by including an iframe that refers to the form, which itself was implemented as a jsp file in CollabReview.

Only wiki articles can be rated

The form is meant only for reviewing Wiki Articles (cf. Figure 5.6a). Internal code checks whether the user is looking Wiki articles. If that is the case, the name of the Wiki article is forwarded to the JSP form. CollabReview checks in the database whether there is an artifact related to that Wiki article. If yes, the user is then able to rate an article (cf. Figure 5.6b).



(a) Only wiki articles can be rated

(b) Rating form for a wiki article

Figure 5.6: Embedding the rating form in Gforge

5.4.3 Creating a Gforge Plugin

The basic steps of creating a Gforge plugin have been already described in Section 5.1.3—"Gforge Plugin Architecture". The name of the plugin is ReputationForge.

index

The file `index.php` is the main file of the ReputationForge plugin. As ReputationForge is a Java (Vaadin) based application, the `index.php` contains just some initialization code and refers to the Java application through an i-frame.

handler

The file `handler.php` checks the different actions under which the `index.php` is called. Based on a specific action the handler calls dynamic versions of the `index.php` file. This is crucial to give the users a feeling of a single system.

In addition to the basic plugin structure, the ReputationForge plugin consists of the following files:

Single Sign-on and Single Sign-out

`SingleSignon.php` file is included in `GfTheme.php` and logs the user automatically in/out to/from CollabReview and Vaadin. It forwards the hashed credentials of the users to CollabReview (by embedding an invisible image), through the file `gforgelogin.jsp` situated under the `gforgeWidgets` CollabReview directory.

Allowed Members

The file `AllowedMembers.php`, situated under `/lib/html` plugin directory is needed for the sake of the evaluation process (cf. Section 6.3.1—"Defining Comparable Groups in ebbits"). It lists all the users of ebbits experimental group, which are entitled to use ReputatationForge and the rating form.

Changes to the "Summary Page"

In order to have a visible system, the `index.php` the *Summary Page* of ebbits project, was changed by including (cf.Figure 5.7):

1. The four best users in ReputationForge.
2. Information whether a user started a task or not.
3. A table showing the most recent activities in Reputationforge.

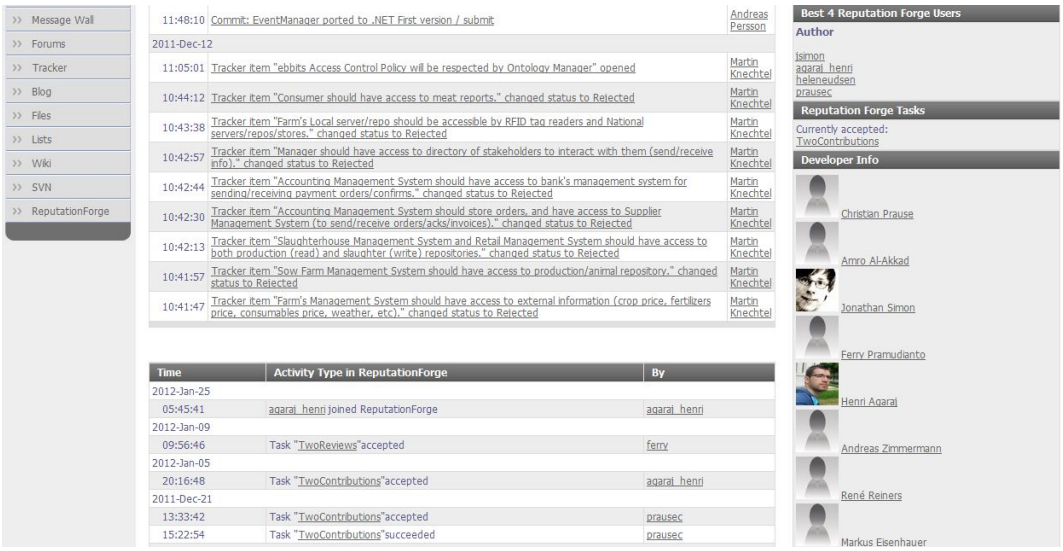


Figure 5.7: Changes performed to the Summary Page of ebbits project

The first two changes are put in a separate file called `ReputationForgeStatsRender.php`, included by `index.php`. The last change is realized through `ReputationForgeSummary.php`, which is also included in `index.php`.

5.4.4 Prototype as a CollabReview Extension using Vaadin Technology

This section describes the extensions to CollabReview that are necessary for the realization of a first successful prototype. An `AuthorManager` class manages the user data and makes it possible to use the same access data in both of the systems. At the very first start of the application, `AuthorManager` imports from the Gforge database all the wiki articles belonging to a specific Gforge project, and saves them to the CollabReview database as *artifacts*. At a second run, the data is updated and only the newest articles are imported.

Author Manager

Integrating ReputationForge with CollabReview

The most important part of this prototype is the Vaadin-based application, which is integrated in CollabReview. For this to happen, the `web.xml` file needs to be updated with a suitable Vaadin servlet mapping.

The Vaadin-based application has several tabs, that are realized as *Navigator7 pages* as described in Section 5.3.2—"Navigator7 Add-On". The connection of pages with the rest of the system architecture can be seen in Figure 5.8. The tabs are as following:

1. the *"Start Page"* gives users information about what other users are doing and the rankings of users based on the status they have (either beginner or advanced one, as explained in the concept session),
2. the *"My Profile Page"* gives personal information about the user and his progress in ReputationForge,
3. the *"Tasks"* page gives users an overview of all the tasks available, that they can complete,
4. the *"Registered Members"* page gives users an overview of all the members of the ebbits project,
5. the *"Administration"* page allows the administrator of the project to add new tasks to the social game, and
6. the *"Wiki Articles"* pages gives the users an overview of all the wiki pages in ebbits project.

Start Page

The Start Page (cf. Figure 5.9a) is the web page that automatically loads when the user presses the tab *ReputationForge*. This page contains the following information:

Recent Activity

- A Recent Activity section that lists users' activity as shown in Figure 5.9c, such as joining ReputationForge, starting, completing, and failing tasks, and becoming advanced users. It also contains links to other

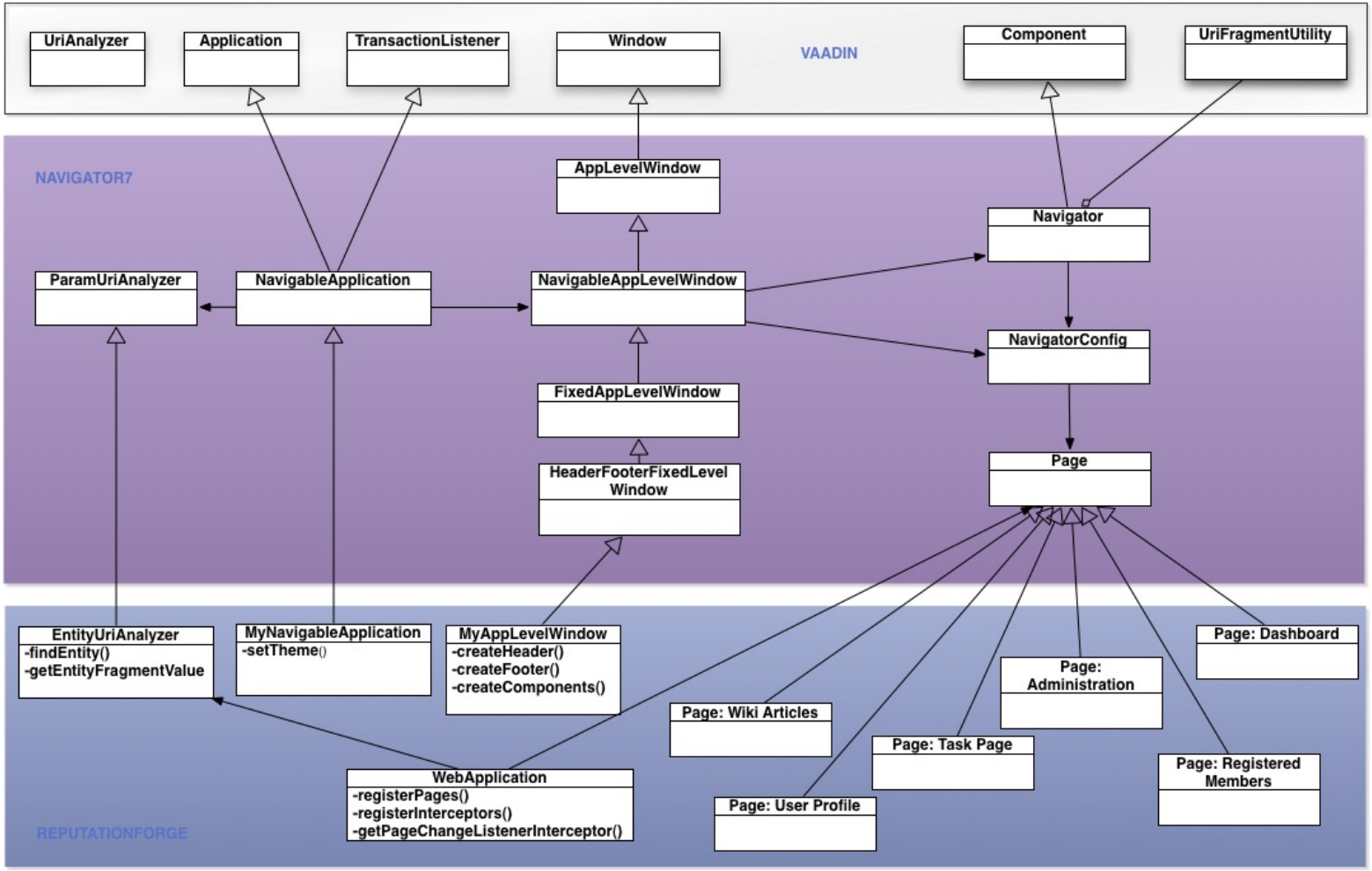
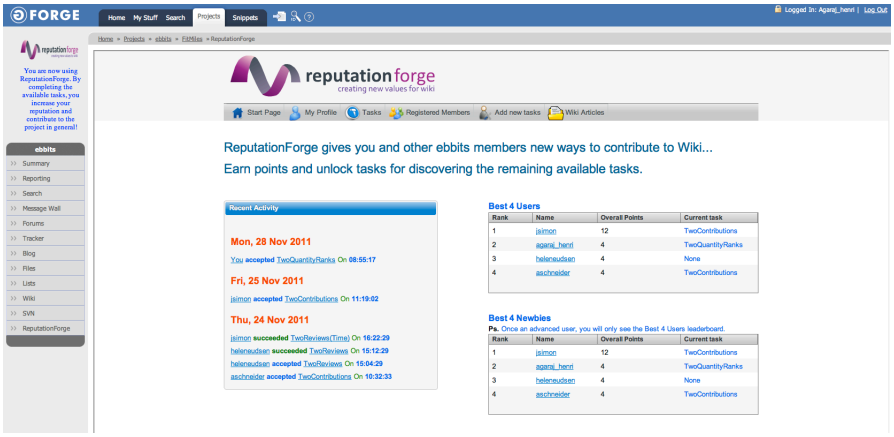


Figure 5.8: System Architecture. The relation between Vaadin, CollabReview, Navigator7



(a) ReputationForge—StartPage

Best 4 Users

Rank	Name	Overall Points	Current task
1	agaraj_henri	4	None
2	jsimon	4	TwoReviews(Time)
3	aschneider	4	TwoContributions
4	martin.knechtel	4	TwoContributions

Best 4 Newbies

Ps. Once an advanced user, you will only see the Best 4 Users leaderboard.

Rank	Name	Overall Points	Current task
1	agaraj_henri	4	None
2	jsimon	4	TwoReviews(Time)
3	aschneider	4	TwoContributions
4	martin.knechtel	4	TwoContributions

(b) ReputationForge—Leaderboards of the participants.

(c) ReputationForge—Recent activities of all participants related to the social game.

Figure 5.9: Start Page of ReputationForge—Main parts

parts of the website, like personal profile of other users, or the task page the user started, completed, or failed.

Leaderboards

- Leaderboards based on the status the users have: a) newbie or b) advanced. Both leaderboards (cf. Figure 5.9b) are a list showing a fixed number of top competitors, ranked by score from highest to lowest. Once the user becomes advanced, the newbie leaderboard disappears.

My Profile

The best way how to display the reputation of each user (as described in Chapter 3—"Related work"), is by associating each user with a personal profile, which shows the most important information related to the game as below:

User image

- Each user has a predefined image, based on his status in ReputationForge. If the user decides to upload a personal image in Gforge, then ReputationForge displays that uploaded image (cf. Figure 5.10).

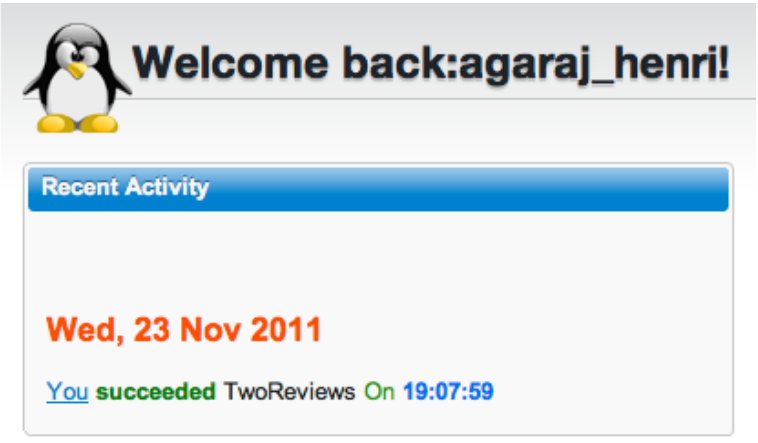


Figure 5.10: Personal profile of a participant in ReputationForge

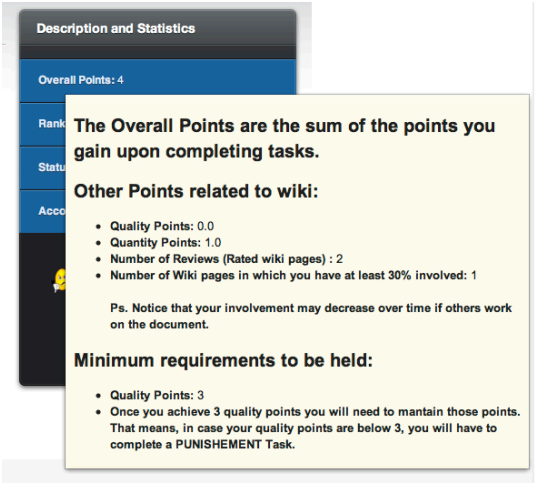


Figure 5.11: Reputation scores in ReputationForge

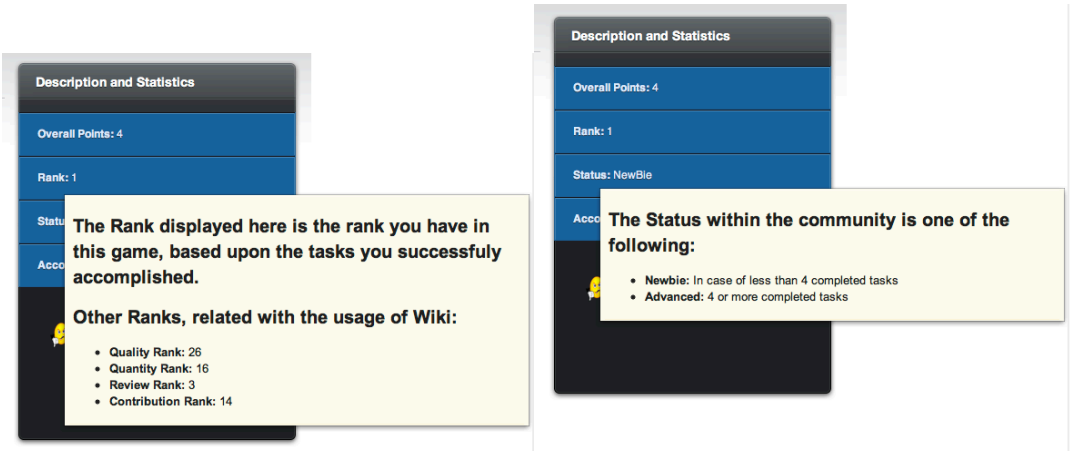
Recent Activity

- The *Recent Activity* section displays a list of the activities only related to the logged-in user, such as joining ReputationForge, starting, completing, and failing tasks, and becoming advanced users, since the last five days (cf. Figure 5.10).

Description and Statistics

- The *Description and Statistics* section gives personal information about the user and his progress in ReputationForge as below:

  1. *Overall Points*: the sum of points a user gains upon completing tasks (cf. Figure 5.11). The overall points are not visible to other users. Besides these points, this section displays scores that are directly related to the usage of the ebbits wiki and computed by CollabReview as described in Section 5.2.1—"Experience Scores". Additionally to these scores, ReputationForge displays *contribution points*, which represent the number of wiki articles of a user, with an authorship of at least 30%.
  2. *Rank*: The overall points define the user's rank in the game and so build his reputation (cf. Figure 5.12a). The rank is not visible to other users. Besides this rank, this section displays the ranking scores directly related to the usage of the ebbits wiki and computed by CollabReview.



(a) Rank(s) of a user in ReputationForge

(b) ReputationForge categorizes users in newbies and advanced ones.

Figure 5.12: Rank and status of a user in ReputationForge

3. *Status*: A user has a newbie or advanced status (cf. Figure 5.12b). Once he completes four tasks, he gains the advanced status.
4. *Task currently started*: This section aims at notifying the user about the task he is currently completing. It links the user to that task page.
5. *Accomplished tasks*: This section display the most recent completed tasks of a user. Achieved tasks are not visible to other users.

Tasks Section

Tasks Overview

The *Tasks Overview* section is the integral part of ReputationForge and represents the social game itself. It displays an overview of all the tasks that are to be completed. As initial point, the first task is available, meanwhile the remaining tasks are disabled. Users can accept only one task at a time. Only after completing a task, other tasks are unlocked. By hovering the mouse over each task icon, a short information is displayed for each task as shown in Figure 5.13).

Immediately after starting, completing or failing a task, the

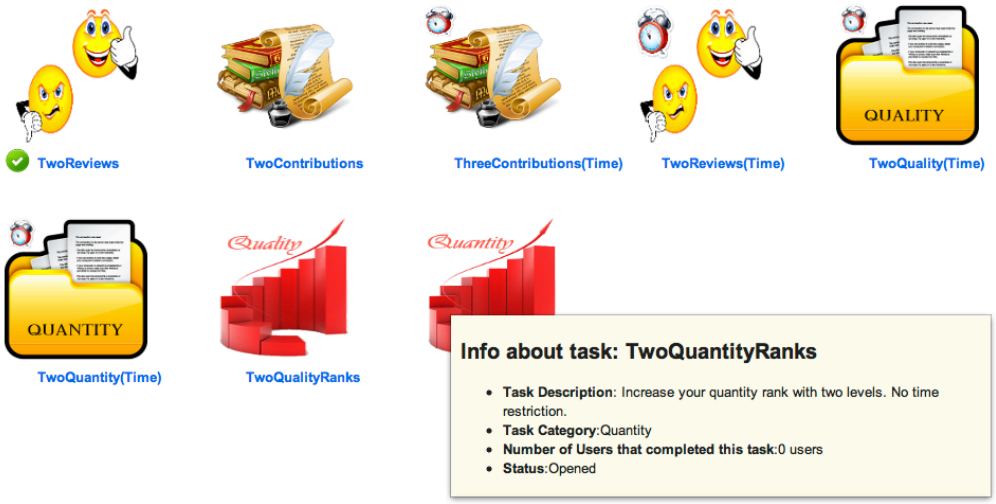


Figure 5.13: Tasks of ReputationForge

Notifying the change of the task status

icon of the respective task is changed accordingly as seen in Figure 5.13, where the first task is completed and signed with a green check icon. In this way, the user is notified about the change of the task status.

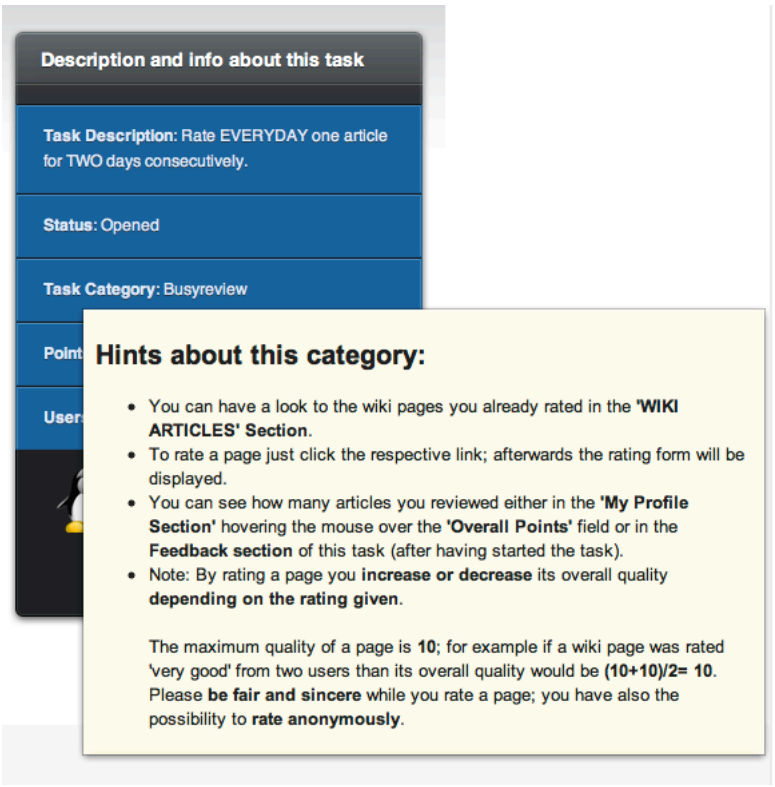
Task Page

Each task page contains the following information:

Info about this task

- The section *Info about this task* gives detailed information about the task. The information displayed consists of:

  1. *Task Description*, which describes what the task is about.
  2. *Task Status*, which shows the actual status of the task.
  3. *Task Category*, which shows the category to which the task belongs (cf. Figure 5.14a). The categories are based on the CollabReview scores and are: *(a) Quality Category, (b) Quantity Category, (c) Review Category, (d) Rank Category.*



(a) Detailed information about the category of a task

(b) Clear feedback for each task

Figure 5.14: Information and feedback about a task

4. *Points won if completed*, which shows the amount of points won upon completing a task. If the user is a newbie, ReputationForge awards twice the amount of points.
5. *Users competing for the task*, which shows all the users that started the task and did not yet successfully complete it.

Feedback Section

Upon starting a task, ReputationForge displays a detailed *Feedback section* (cf. Figure 5.14b), which shows the user the remained effort needed to be put in. While having a look into a task started by another user, ReputationForge shows the user whether he completed that task also. If not completed, it forwards the user to the task page (cf. Figure 5.15).

Figure 5.15: ReputationForge incents users to start already completed tasks by other users

Registered Members

List of registered members

This section gives an overview of the registered members (cf. Figure 5.16). By hovering the mouse over each user icon, ReputationForge displays the name and status of the user.

Figure 5.16: List of all registered members. Additionally, a contact card for each member is shown

Wiki Articles

List of Wiki articles

This section gives an overview of all the wiki articles added in the ebbits project. Its aim is to help users accomplish the tasks easily and allows the user to have a seamless navigation between Gforge and ReputationForge.

Info about each wiki article

Wiki articles are sorted out by latest modification date, so that users get to know immediately which wiki article was changed recently. For each wiki article, the following is displayed: *(a) Link to the ebbits wiki article, (b) last modified date, (c) article quality, (d) contribution of at least 30 percent, (e) and whether the article is rated or not by the user.*

Searching for wiki articles

The users can also search for a specific wiki article. The search is case insensitive. Upon search completion, feedback is given about the list of articles found.

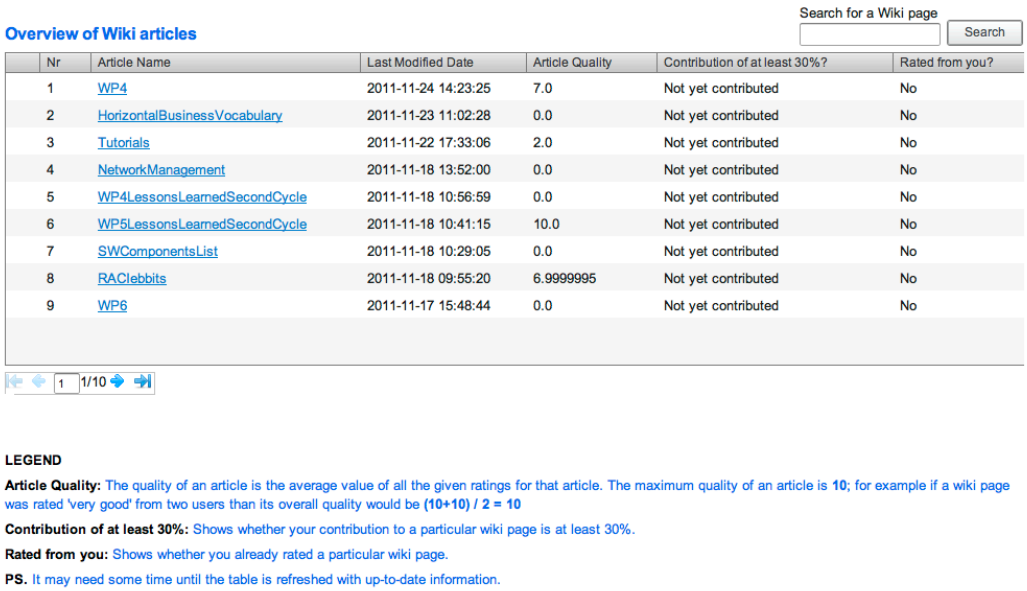


Figure 5.17: List of all wiki articles sorted out by latest modification date.

Administration Page

This section allows the administrator of the ebbits project (or a person with a specific Gforge role) to easily add new tasks to the system (cf. Figure 5.18). The template of each class is defined by the underlying POJO of that task (see Section 5.4.4—"Class Diagram/Model of ReputationForge". The administrator can additionally specify an image for a task.

Type of tasks

Depending on the use case, the administrator can choose to add:

- *General Tasks*, which are high level tasks, related to the Wiki in general. The following example tries to makes this definition clearer: "Try to keep the kitchen clean please". The task given to the users here does not require from them to clean specific parts of the kitchen. It just says the users: keep the kitchen clean. It is up to the users, how they want to clean the kitchen. Example: *Add two quantity points*.

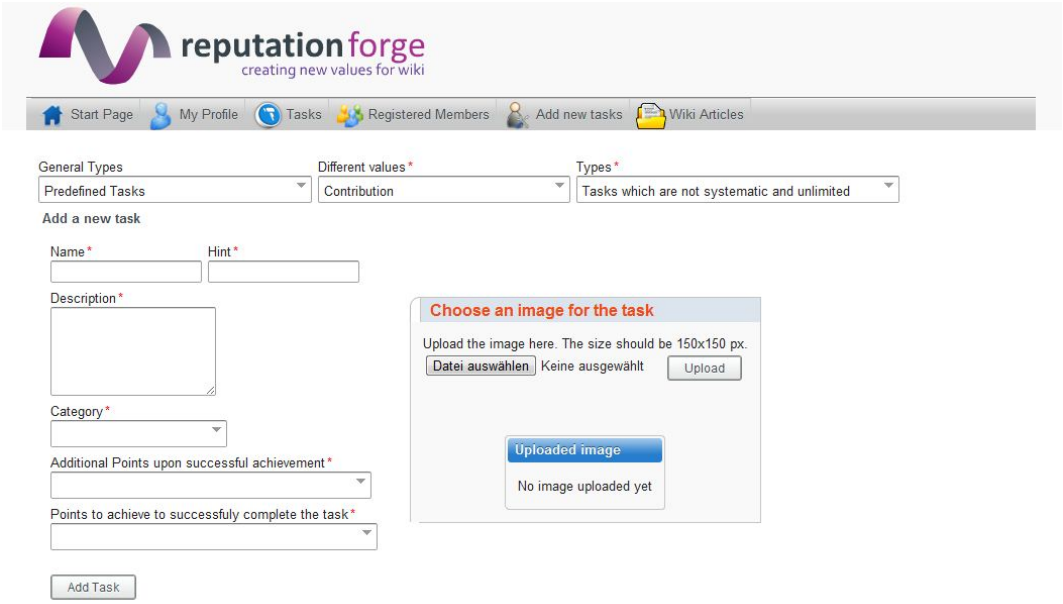


Figure 5.18: The administrator can easily add different type of tasks

- *Specific Tasks*, which are low level tasks, related to very specific parts of wiki. Referring back to the above example, if the task now would be: "Please clean the fridge in the kitchen", now every user knows exactly what to do. *Make two contributions of at least 30%*.

Additionally, each general or specific task can be a:

1. *Non Systematic and Not Limited Task*: Such a task does not put a time restriction to the user. The user is free to take the time he needs in completing the task. The user is also not required to work systematically towards the achievement of the task. Example: *Rate two wiki articles*.

2. *Non Systematic and Limited Task*: Such a task puts a time restriction on the users. The user should complete the task within a deadline. If the due date is not held, the task is considered to be failed. Again, the user is not required to work systematically towards the achievement of the task, he should just complete

the task within the required frame of time. Example: *Rate two wiki articles within two days*.

3. *Systematic and Limited Task*: Such a task puts a time restriction on the users, similar as in the previous described task. Still, the difference consists in the fact that this task requires the user to complete systematically (a defined time-frame) a predefined amount of points in order to complete the task while the previous task does not put such a requirement. An example would be: *Rate two articles every day for three days consecutively*.

Upon adding the task, that task is visible for all the users in the ReputationForge in the *Task Section*.

Class Diagram/Model of ReputationForge

The design strategy that allows adding tasks in an easy way, is the abstract parent class as described in Gamma et al. (1995). As shown in Figure 5.19, the abstract class *Task* contains:

- Implementation of methods, which represent the invariable part of the class functionality.
  - `accept()`, called when a task is started.
  - `succeed()`, called when a task is completed.
  - `fail()`, called when a task is failed.
- Abstract methods, which represent the variable parts of the class behavior.

Concrete subclasses

The three concrete subclasses defined (cf. Figure 5.19) correspond to the types of tasks described in Section 5.4.4—"Administration Page".

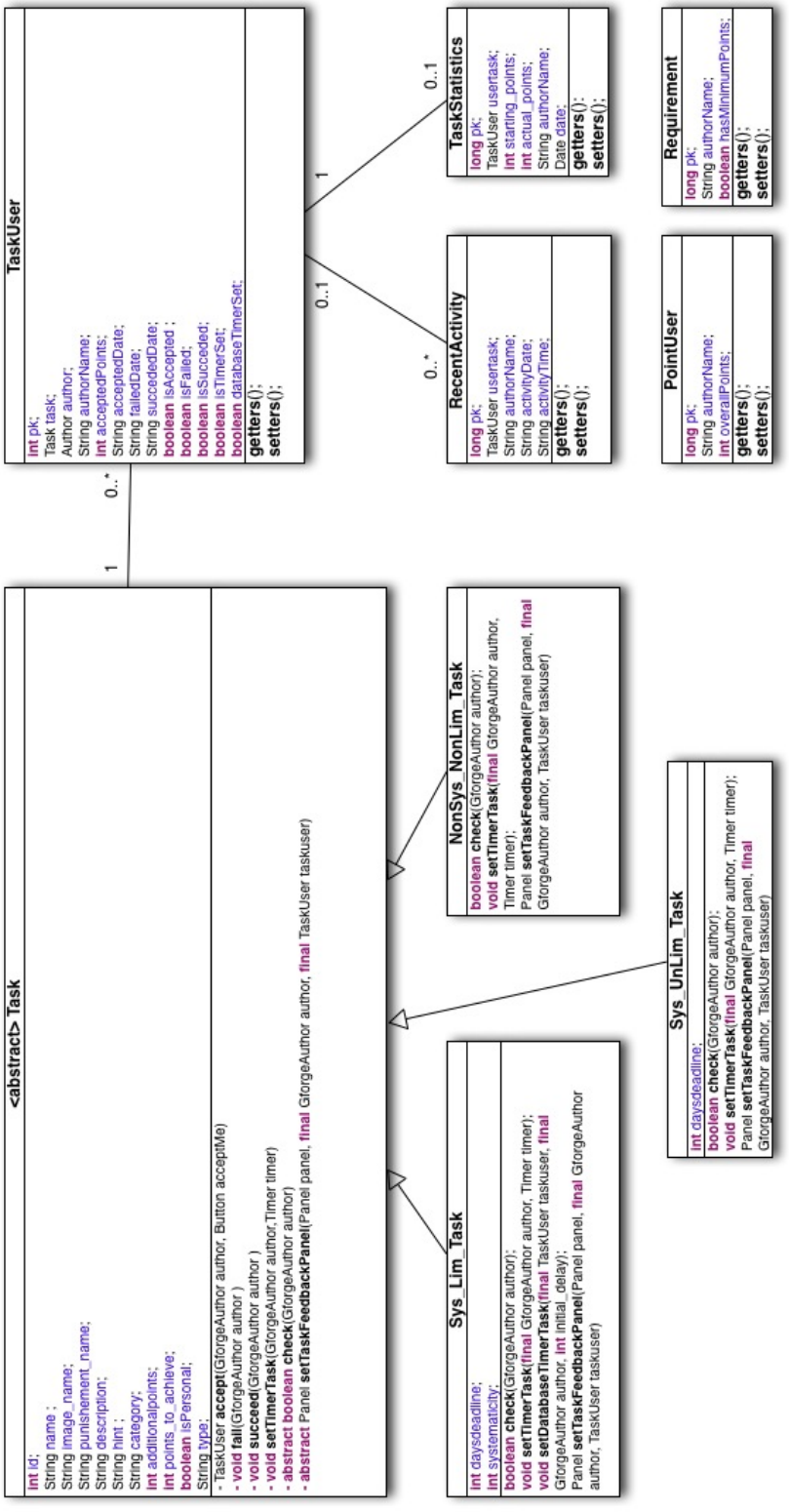


Figure 5.19: Class diagram of ReputationForge

5.4.5 Email Notifications and Punishment Tasks

Triggering users

Based on Section 2.1.1—"Fogg Behavior Model for Persuasive Design (FBM)", ReputationForge was made visible by introducing the following:

- Changing the "Summary Page" of ebbits as described in Section 5.4.3—"Creating a Gforge Plugin".
- Sending regularly email notifications to the users. `GforgeNotification` class is responsible for sending every week the following systematic emails: *(a) If the users belongs to the four worst users they get notified, and (b) If the users do not complete any task for a week, they get notified.* Additionally, a single email is sent upon task completion.

Punishment Task

Normal progress of the game stops, unless a punishment task is completed

As described in Section 4—"Finding a suitable concept", a first prototype should introduce a punishment task. Indeed, ReputationForge demands the completion of a punishment task, in case participants fail to maintain a minimum of three quality points. The normal progress of the game stops, and the participant needs to complete a punishment task, which is added automatically by the system. Such a task aims at bringing the participant back to the level of three quality points.

State Pattern

Task lifecycle

A task has its own life cycle as indicated in Figure 5.20: `open` represents the time when a task is created and the user did not start any other tasks; `completed or failed` when a task is accomplished and archived; `started` represents the time when the user started a task but did not yet complete it.

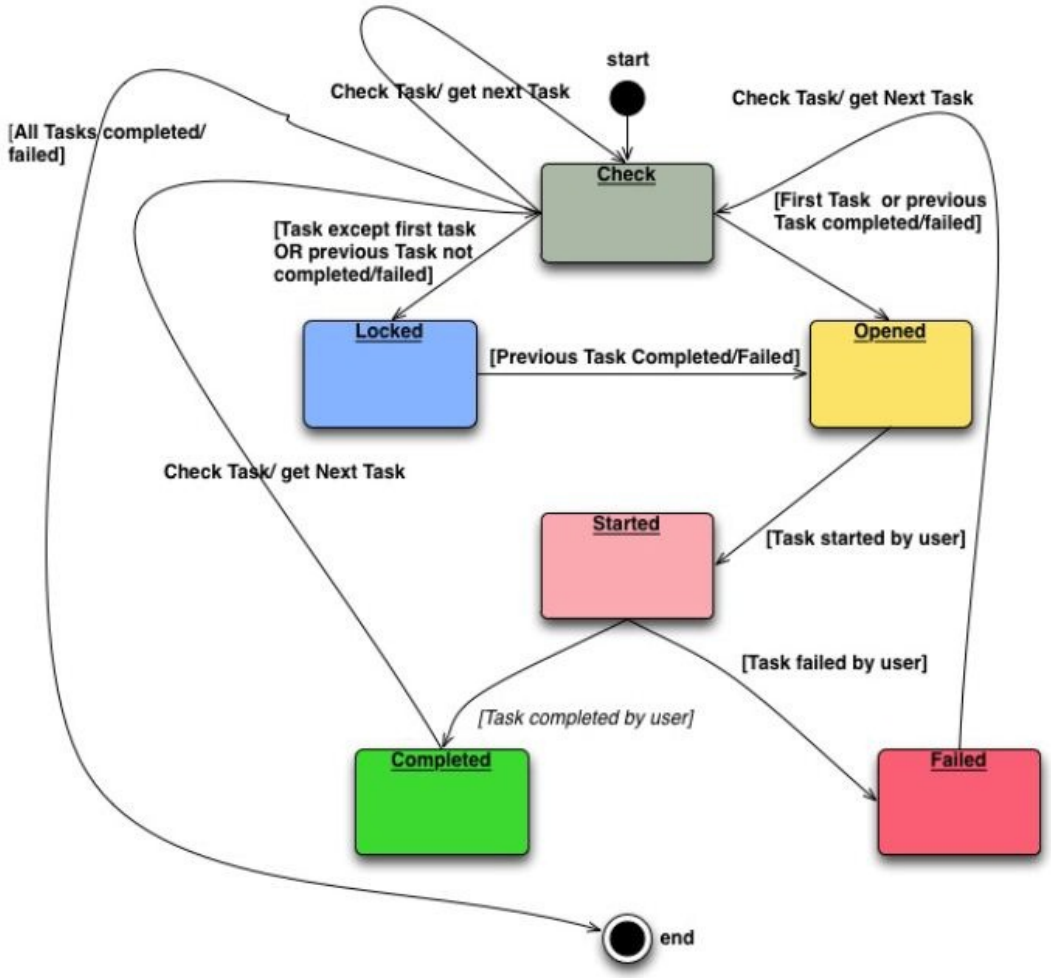


Figure 5.20: Lifecycle of a task in ReputationForge

As a starting point, the first task is opened and the remaining tasks are locked. Users have to complete this first task, to be able to progress through the other available tasks.

5.5 Summary

Outlook of this chapter

This chapter described the actual implementation of the prototype based on the final concept defined in Chapter 4—"Finding a suitable concept". The technical tools used, along with implementation details were extensively dis-

cussed. Next, the system architecture, along with its realized components was explained. Screenshots of the realized prototype *ReputationForge* tried to give a clearer picture of the system.

Insight into the next chapter

The following two chapters present the outcome and implementation evaluation performed on both the ReputionForge prototype and its concept. Based on the results of such evaluation methodologies, an extensive analysis was performed.

Chapter 6

Outcome Evaluation

"The true worth of a researcher lies in pursuing what he did not seek in his experiment as well as what he sought "

—Claude Bernard, French physiologist

The present chapter and chapter 7—"Implementation Evaluation" constitute the evaluation of this thesis project, aiming at finding answers to the research questions formulated in 1.4. The research questions were:

1. *Does a social-game, with features as described in Section 1.3—"Strategy" increase the participant's engagement in team-wikis?*
2. *What is the overall satisfaction and motivation of the target users?*
3. *What are the most suitable aspects of the concept that lead towards a higher participation rate in team wikis?*

Evaluation methods of this thesis project

The reader is first given an insight into the used evaluation methods: *(a) process/implementation evaluation, (b) outcome evaluation*. The main focus of this chapter is the outcome-based evaluation of ReputationForge, which tries to find out whether the introduction of ReputationForge made a difference, thus giving an answer to the first research question.

Furthermore, this chapter starts the implementation evaluation of ReputationForge, which is continued by Chapter 7—"Implementation Evaluation", aiming at giving detailed answers to the remaining research questions.

6.1 Evaluation Methods

Evaluation background

Each good evaluation should build on existing evaluation knowledge and resources. Such a knowledge is of particular interest for the evaluation of this thesis project.

Questions answered by evaluation methods

Innovation Network (2011) gives insight into two major forms of evaluation, namely *(a) implementation evaluation, and (b) outcome evaluation*, that seek to answer questions like:

- What was done?
- How well was it done?
- What changes occurred because of the introduction of the system?

The following briefly summarizes the main considerations of Innovation Network (2011) regarding implementation, and outcome evaluation.

6.1.1 Process / Implementation Evaluation

The core of the method

This method addresses what was done and how well was it done, but it does not address what changes occurred because of the program. This method gives an answer to the following questions:

- *Participation:* What and how much was accomplished by the target users? Why?
- *Quality:* Which of the features of the program were perceived as valuable by the intended audience?

- *Satisfaction:* Were the target users overall satisfied with the introduced system?

Giving an answer to such questions, helps at the same time answering the research questions of this thesis project.

6.1.2 Outcome Evaluation

The core of the method

This evaluation method tries to assess the outcomes and sought changes in behavior, and awareness that result from the system. This method gives an answer to the question: *What differences/outcomes occurred?*

6.2 Establishing and Evaluating Causal Relationships

As described in the previous section, the outcome evaluation determines outcomes, which are changes that are expected to see as a result of the introduced program/system. Outcomes are strongly related to the concepts of a "cause-effect" relationship and validity. The following describes these concepts in detail.

6.2.1 Validity

The concept of validity is crucial in quantitative studies.

Definition: *Validity*

VALIDITY:
"the best available approximation to the truth or falsity of propositions" (Shadish et al., 2001)

Internal vs. external validity

Shadish et al. (2001) distinguish between internal and external validity. They claim that the ideal case for a good study would be that it proofs both type of validities. According to them, internal validity is an essential one. On the other

side, answering that the same result of a given study can be observed in other situations will always remain incomplete because every critic could find a counter-example that would weaken the external validity of a study.

The concept of both internal, and external validity are described as follows:

Internal Validity

Definition: *Internal validity*

INTERNAL VALIDITY:
An experiment/study is considered to be internally valid, if it can show a cause-effect relationship between the independent (manipulated condition) and dependent variable (the behavior measured). This relationship is only valid, if and only if the conditions other that the manipulated one are ruled out as potential causes for the sought behavior change (Trochim, 2006).

Threats that disturb the internal validity

In order to achieve a strong internal validity, Shadish et al. (2001) considers threats that must be ruled out. Some of these threats are:

- *History threats*: Some historical events that occurred during the evaluation produce changes in outcomes.
- *Maturation*: the processes within subjects which act as a function of the passage of time.
- *Testing*: if participants take a pretest, they would be ready in a way that they would not have been without the pretest.
- *Instrumentation*: the changes in the data collection methods which may produce changes in outcomes.
- *Statistical regression*: False outcome caused by the selection of participants on the basis of extreme scores or characteristics.
- *Selection of subjects*: the biases which may result in selection of non randomized assignment of participants in groups.

External Validity

EXTERNAL VALIDITY:
"asks the question of generalizability: To what populations, settings, treatment variables and measurement variables can this effect be generalized? " (Shadish et al., 2001)

Definition: *External Validity*

6.3 Experimental Setting

The following describes the experimental procedure followed during the course of investigation, along with the dependent measures that were collected.

6.3.1 Defining Probabilistic Equivalent Groups

Trochim (2006) analyzes the most important experimental designs while assessing causal hypothesis, which are based on: (a) the manipulated condition, (b) the independent variable, and (c) the random assignment. All the designs fall than in two main categories: *(a) single group, and (b) multiple group designs*.

Experimental Designs

After having analyzed the advantages and disadvantages of each design (ie. one shot case study, pre- test-single group, static group comparison, salomon groups, and pretest-post-test control group), the best solution would be to introduce a control group that helps assessing whether the introduction of ReputationForge increases the participant's engagement in the internal wikis. In such an experimental design, the experimental group receives the program, the control group does not. Additionally, the groups should be comparable to each other, so that the only difference is the program. The existence of a comparable control group is the most effective way to rule out single-group threats to internal validity as seen in Section 6.2.1—"Internal Validity".

Experimental Design for this thesis project

Why a control group?

Defining Comparable Groups in ebbits

ebbits partners

There are in total 38 participants in ebbits, which come from nine different project partners.

ebbits workpackages

The participants of ebbits work in so called *workpackages*, which are responsible for specific parts of the project.

Strategy

The following strategy (cf. Table 6.1) was chosen to create two comparable groups:

- Each workpackage was represented approximately 50% in each group.
- Each partner was represented approximately 50% in each group.
- A list of combined measures (contributions and logins) was created. The participants at the top of the list represented the participants with the highest scores. The participants were than randomly assigned to both the groups. A sample-two-test showed no statistical difference between those groups ($H_0 =$ The means of the groups are different: $\rho = 0.979 > 0.05$). As such, the groups can be thought of as probabilistically equivalent.

Measure	Distribution	
	Exp.Group	Con.Group
Workpackage (%)	48	52
Partner (%)	49	51
Contribution[a]	174	111
Contribution[b]	342	263
Logins[c]	91	86

[a]Sum of contributions to wiki from 5.10.2011 to 15.11.2011 (comparable with the period of evaluation).
[b]Sum of contributions to wiki since the start of the project.
[c]Number of logins from 5.10.2011 to 15.11.2011.

Table 6.1: Defining comparable groups in ebbits

Social differences between participants

Demographic differences between participants like education, organization tenure, race, sex, and social position were not taken into consideration.

Defining Comparable Groups in Moknow

Categorization in Moknow

A similar procedure was chosen for Moknow. Here the members are categorized in: *(a) students, (b) junior/senior researchers of the Moknow department, and (c) junior/senior researchers of the CAPLE department*.

Strategy

For each of the above categories, half of the participants were assigned to the control group, half of the participants to the experimental group (cf. Table 6.2). It is important to mention that four participants from Moknow group already participated in ebbits, thus they were automatically assigned to the control group of Moknow. The rest of the Moknow participants were assigned to the experimental group. This decision was taken to give all of them the opportunity to test ReputationForge. Such a decision has a drawback, because the remaining Moknow participants happened to be the most active users in Moknow, leading to a difference between the control and experimental group (Average rank for the experimental group of 10.9 vs. an average rank for the control group of 14.5). Still, two participants of the experimental group (with ranks 2 and 5, respectively) informed that they could not participate (lack of time, vacation leave), so they were indeed considered being part of the experimental group. The average rank after this reassignment procedure was 12.3 for the experimental group and 12.9 for the control group (cf. Table 6.2).

Measure	Distribution	
	Exp.Group	Con.Group
Students (%)	50	50
Researchers Moknow (%)	60	40
Researchers CAPLE (%)	50	50
Rank Average	12.3	12.9

Table 6.2: Defining comparable groups in Moknow

6.3.2 Experimental Procedure

Field study evaluation

It is expected that a real change in the participants behavior would occur when users use the system in their own normal work environment rather than in a laboratory study. For this reason, a field study was conducted, where the participants used ReputationForge in their normal work context. Furthermore, such a research study is associated with greater external validity.

Overall Observation

Observation in ebbits wiki

A detailed analysis of ebbits participants with relation to their participation/engagement in the wiki in different intervals of time proved evidence that only a few participants contribute regularly to the ebbits wiki. For the rest of the participants, once some users are very active, and once other users become more active, but overall their participation is very low.

Observation in Moknowpedia

On the other hand, Moknowpedia (the internal wiki of Moknow) is an established wiki. The majority of the participants participate in it regularly.

Task Preparation

Before starting the evaluation, the following tasks were added to ReputationForge for both groups ebbits and Moknow.

- *Not periodic and not limited tasks*:
  1. Rate two articles.
  2. Make two contributions of at least 30 percent.
  3. Increase the quality rank with two levels.
  4. Increase the quantity rank with two levels.
- *Not periodic but limited tasks*:

1. Make three contributions of at least 30 percent in wiki within 3 days.
2. Gain two quantity points within four days.
3. Gain two quality points within four days.

- *Periodic and limited tasks*:

1. Rate everyday one article for two days consecutively.

Instructions given to the Participants

Introducing ReputationForge to the participants

Just before the evaluation started, the Head of the Ubiquitous Computing Department in FIT sent an official email to the participants, informing them about the introduction of ReputationForge. A short description of ReputationForge and its goal was given. Instructions how to access ReputationForge were given. Participants were assured that the collected data would be treated confidentially and evaluated anonymously. It was also made clear to the participants why they were assigned to the experimental group.

Participants and Evaluation Period

Participants of ebbits wiki

Nineteen participants representing the "ebbits experimental group" tested ReputationForge for 5 weeks.

Moknowpedia participants

Thirteen participants representing the "Moknow experimental group" tested ReputationForge for 10 days.

The characteristics of the target users and their working environments have been already described in Section 4.1.1—"Target Users/User Environment".

Participants' activities

During the evaluation period, participants received systematic email notifications from ReputationForge. While accepting and trying to complete tasks of the social game, the participants: *(a) reviewed articles, (b) added new articles to wiki, and (c) improved existing articles in the wiki.*

To facilitate their work, ReputationForge displayed them: *(a) a list of articles that were not reviewed by them, and (b) a list of articles in which they did not have a sufficient contribution*.

Overall Observation during the Evaluation Period

Observation in the ebbits wiki

While comparing the five weeks before the introduction of ReputationForge to the five weeks of the evaluation period, both groups (experimental and control group) in ebbits contributed less to ebbits wiki during the evaluation period. This is an indication that ReputationForge was evaluated in a period of no considerable progress in the EU research project ebbits.

Observation in Moknowpedia

A similar situation was seen in Moknowpedia.

Such a situation can also be attributed to the Christmas period (especially the evaluation in Moknowpedia).

6.3.3 Dependent Measures

The dependent measures were based on database data and data collected from the final on-line questionnaire.

Objective measures

From the saved data the following was extracted: *(a) the number of log-ins, (b) the number of started/completed/failed tasks, and the (c) wiki experience scores, such as the quality score, contribution and review score*, as described in Section 5.2.1—"Experience Scores".

Subjective measures

A post-test questionnaire asked participants to rate on a five-point Likert scale, various aspects of ReputationForge

along with its motivational values. The evaluation and results of such subjective measures are given in detail in Chapter 7—"Implementation Evaluation".

6.4 Hypotheses

Formulating hypotheses based on the thesis goals

The hypotheses below are related to the main goals of this thesis project: *(a) to increase the target users' engagement in wiki, and (b) to compare ReputationForge and the existing team-wiki Moknowpedia with relation to the overall satisfaction and motivation*.

The hypotheses are categorized as: *(a) Causal Hypotheses, and (b) Non-Causal Hypotheses*, as following:

- Causal Hypotheses:
  - *Participation Hypotheses*:
    - H_0: There is no statistically significant difference between the objective measures (cf. Section 6.3.3) collected in the experimental and control group of ebbits.
    - H_a: There is a statistically significant difference between the objective measures collected in the experimental and control group of ebbits.
    - H_0: There is no statistically significant difference between the objective measures collected in the experimental and control group of Moknow.
    - H_a: There is a statistically significant difference between the objective measures collected in the experimental and control group of Moknow.
- Non-Causal Hypotheses:
  - *Usability Hypothesis:* Participants of Moknowpedia and ebbits wiki are satisfied with the usability of ReputationForge.

- *Motivation Hypothesis:* The social game promoted by ReputationForge and its incentives motivate the participants.
- *Comparison Hypothesis:* ReputationForge motivates participants more than the already existing system Moknowpedia.

6.5 Objective Measures Results

This section outlines the results achieved by the participants of both experimental groups and the impact of ReputationForge on their engagement in the respective wikis.

6.5.1 Social Game Usage/Tasks completed

The following tries to give an answer to the question "What and how much was accomplished by the users? Why¿' as presented in Section 6.1.1—"Process / Implementation Evaluation". The other questions described in that section are analyzed in Chapter 7—"Implementation Evaluation".

Results in ebbits wiki

Tasks completed in ebbits wiki

At the end of the evaluation period, eight out of nineteen participants of the ebbits experimental group used ReputationForge. Thirteen tasks were completed by them. Six out of the remaining eleven participants never logged in, neither in ReputationForge nor in Gforge. The rest logged in from time to time, but they did not start any task of ReputationForge social game. Some of them mentioned the *lack of time* as the major reason. The following describes the results more in detail:

- Task *"Rate two articles"* was completed by six users.
- Task *"Make two contributions of at least 30%"* was completed by two users. Three users started this task but did not complete it within the evaluation period.
- Task *"Rate everyday one article for two days consecutively"* was completed by two users.
- Task *"Gain two quantity points within four days"* was started by one user and failed by him.
- Task *" Gain two quality ranks"* was completed by one user.
- Task *"Gain two quantity ranks"* was completed by one user.

Results in Moknowpedia

Tasks completed in Moknowpedia

At the end of the evaluation period, six out of eleven participants of the Moknow experimental group used ReputationForge. Nine tasks were completed by them. The rest never logged in to ReputationForge. Some of them mentioned the *lack of time* as the major reason. The following describes the results more in detail:

- Task *"Rate two articles"* was completed by six users.
- Task *"Make two contributions of at least 30%"* was completed by two users. One user started this task but did not complete it within the evaluation period.
- Task *"Make three contributions of at least 30%"* was started but failed by one user.
- Task *"Rate everyday one article for two days consecutively"* was completed by one user. One user failed this task.

Comparison of Results

ebbits vs. Moknow

The average participant in ebbits completed 8.5% of the tasks available, whereas the average participant in Moknow completed 10.25% of the tasks available.

Greater interest of Moknow

The evaluation period of Moknow lasted ten days in comparison to a five week evaluation of ebbits, thus leading to the conclusion that Moknowpedia participants showed a greater interest towards ReputationForge than ebbits participants. Such a fact is also confirmed by two participants of ebbits wiki, who believe that ReputationForge is more suitable for more active wikis than the ebbits one.

6.5.2 Participation Hypothesis

This section gives an answer to the validity of the causal hypothesis of this thesis project. The participant's engagement of the experimental and control group in both team wikis was compared using a two-sample t-test[1] (cf. Figure 6.1). The following measures were considered:

- Wiki experience scores (cf. Section 5.2.1—"Experience Scores"), namely:
  - Contributions (adding new articles to wiki, and/or improving existing wiki articles) of any authorship.
  - Contributions of at least 30% authorship.
  - Quality—an average value from 0 to 10. It is calculated based on the authorship for a specific wiki article and the average rating the article has.
  - Reviews—the number of articles rated by the participants.
- Logins, the number of times the participants logged in.

[1] A quick analysis with the Anderson-Darling distribution test showed normal distribution of the collected data, as such a distribution is a prerequisite for the two-sample t-test

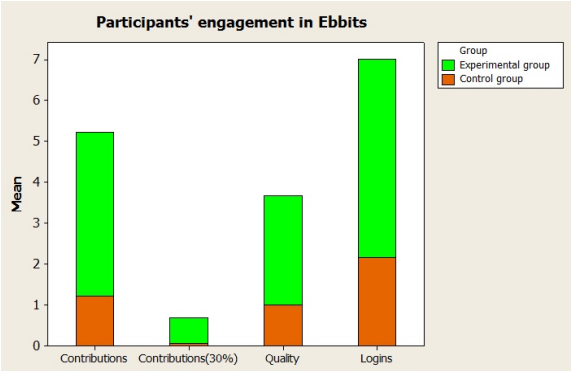
Results in ebbits wiki

Contributions of any authorship

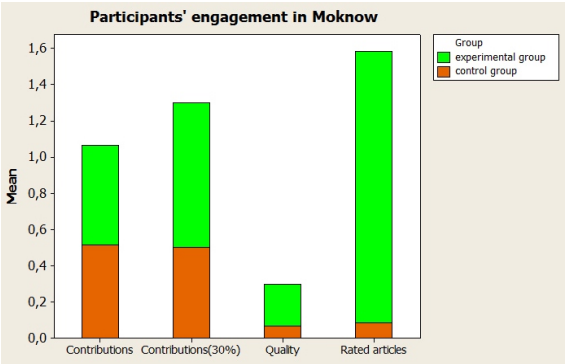
For contributions of any authorship, the contribution rate of the experimental group ($M = 4.05, SD = 5.48$)[2] is 3.45 times bigger than that of the control group ($M = 1.22, SD = 2.44$). Indeed, a sample two-test shows a significant difference ($t(25) = 2.05, \rho = 0.05 <= 0.05$).

Contributions of at least 30%

For contributions of at least 30% authorship, the contribution rate of the experimental group ($M = 0.63, SD = 1.34$) is 12 times bigger than that of the control group ($M = 0.056, SD = 0.236$). A sample two-test shows a borderline significant difference ($t(19) = 1.84, \rho = 0.081 > 0.05$).



(a) Participants' engagement in ebbits wiki



(b) Participants's engangement in Moknowpedia

Figure 6.1: Participant's engagement in both ebbits wiki and Moknowpedia

[2] M = "Mean", SD = "Standard Deviation"

Quality change

The average quality of articles of the experimental group ($M = 2.67, SD = 2.67$) increased by 2.43 times more than that of the control group ($M = 1.01, SD = 1.85$). A sample two-test shows a significant difference ($t(32) = 2.21, \rho = 0.034 < 0.05$).

Reviews

A comparison of the total number of articles rated by both groups in ebbits was not possible, because the rating form was only available for the experimental group.

Number of logins

Participants of the experimental group ($M = 4.84, SD = 6.79$) logged in 2.34 times more than the participants of the control group ($M = 2.17, SD = 2.26$). However, the difference was not statistically significant ($t(35) = 1.59, \rho = 0.121 > 0.05$).

Results in Moknowpedia

Contributions of any authorship

For contributions of any authorship, no noticeable difference between the groups was found. The experimental group ($M = 0.55, SD = 0.743$) contributed 1.057 times more than the control group ($M = 0.52, SD = 1.03$). Indeed, a sample two-test showed no significant difference ($t(20) = 0.09, \rho = 0.933 > 0.05$).

Contributions of at least 30%

For contributions of at least 30% authorship, the contribution rate of the experimental group ($M = 0.62, SD = 1.04$) is 1.34 times greater than that of the control group ($M = 0.5, SD = 0.798$). Still, a sample two-test shows no significant difference ($t(23) = 0.31, \rho = 0.76 > 0.05$).

Quality change

The experimental group had an increase in its average quality of articles ($M = 0.177, SD = 0.676$) by 3.13 times more than that of the control group ($M = 0.067, SD = 0.089$). Still, a sample two-test shows no statistical difference ($t(12) = 0.58, \rho = 0.571 > 0.05$).

Rated articles

The experimental group ($M = 1.15, SD = 1.41$) rated 15 times more articles[3] than the control group ($M =$

[3]One can argue that the quality was increased, due to participants rating their own articles more positively. However, a fast look on the database values proved evidence that participants did not rate their own

$0.083, SD = 0.289$). Indeed, a sample two-test shows statistical difference ($t(13) = 2.69, \rho = 0.019 < 0.05$).

This analysis was not possible for ebbits, because of lack of data.

Number of logins

Hypotheses Satisfaction

Based on these results, the following can be concluded:

- The null hypothesis in ebbits wiki was rejected with high confidence in the majority of cases, except for the number of logins (still the ebbits group scored better). These results suggest that the introduction of ReputationForge increased significantly the participants' engagement in the ebbits wiki.

  Result in ebbits wiki

- The null hypothesis in Moknowpedia was not rejected in the majority of cases, except for the number of rated articles. However, the experimental group of Moknow showed in any case a greater engagement than the control group. Such a result can be attributed to various factors like *(a) the short and unsuitable evaluation time (just before Christmas Eve), (b) the well established internal wiki, and (c) the further involvement of the participants of the control group in the existing system Moknowpedia*.

  Result in Moknowpedia

6.5.3 Summary

The main focus of this chapter was to find out empirically whether the introduction of ReputatationForge made a difference in the participant's engagement. In order to achieve a higher internal as well external validity of the evaluation results, a control group field study evaluation was performed in both ebbits wiki and Moknowpedia. Before actually evaluating the impact of ReputationForge

articles

Greater interest of Moknowpedians towards ReputationForge

in both working environments, an analysis of "what and how much was accomplished by the participants" was performed. It pointed out that the Moknowpedia participants showed more interest towards ReputationForge than the ebbits participants. Such a result can be attributed to the fact that the internal wiki of Moknow is used massively for an every-day information exchange within the group, while the internal wiki of ebbits is a relatively new one, and the participants' engagement depends on the actual progress of the project. If there is no progress, there is no need to document in wiki.

Significant improvement in the ebbits wiki

Partial significant improvement in Moknowpedia

The outcome evaluation demonstrated a significant increase in the engagement of the ebbits experimental group (ie. contributions of any authorship, contributions of at least 30% authorship, average quality of articles, and the number of rated articles) . A statistically significant effect was only partially obtained in Moknowpedia (ie. number of rated articles), albeit better participation scores than the control group were registered. The result of the latter can be attributed to factors like the relatively well-established wiki, the short and unsuitable evaluation period, and the further involvement of the participants of the control group in the existing system Moknowpedia. A longitudinal evaluation in a future would probably show a significant increase in the Moknow participants' engagement, based on the high interest shown within such a short evaluation time, as discussed above.

The participants' engangement increased

Summarizing, the outcome evaluation of ReputationForge proved evidence that a social-game approach based on reputation can actually increase the participant's engagement in a corporate/team wiki.

Insight into the next chapter

The next chapter continues the evaluation of ReputationForge, namely its implementation evaluation.

Chapter 7

Implementation Evaluation

"Evaluate what you want—because what gets measured, gets produced"

—James Belasco

Questions answered

This chapter continues the implementation evaluation of ReputationForge that was begun in the previous chapter by focusing on the following research questions:

1. *What is the overall satisfaction and motivation of the target users?*
2. *What are the most suitable aspects of the concept that lead towards a higher participation rate in team wikis?*

Overall respond rate in both cases

It mainly consists of the quantitative and qualitative results jointly collected from a post-test questionnaire. 30 participants from the two working environments (Nineteen from ebbits and eleven from Moknow) were supposed to fill the questionnaire. The overall respond rate was: *42%* for ebbits and *72%* for Moknow. However, if considering only participants who used ReputationForge as well completed the survey, the overall respond rate looks like: *87.5%* for ebbits and *100%* for Moknow.

Survey statements

The survey presented statements, which were rated on a five-point Likert scale— "strongly agree"(1), "agree" (2), "neutral" (3), "disagree" (4), "strongly disagree" (5). The statements were than analyzed through descriptive statistics. Participants assessed topics such as: *(a) the usability of ReputationForge, (b) the overall satisfaction and motivation, (c) the tasks' difficulty, and (d) the motivational aspects of ReputationForge concept*. Additionally, the questionnaire gave the participants the possibility to add open-ended comments.

Both the quantitative and qualitative research methods help to give an answer to the research questions and collect evidence whether to support or deny the hypotheses set in Section 6.4—"Hypotheses".

7.1 Usability

This section describes the qualitative and quantitative procedure for analysing the *usability* of ReputationForge. The existence of an ergonomic system is a prerequisite for a consequent behavior change in the target audience.

7.1.1 Quantitative Results

The following examines the agreement of the participants' ratings with the statements regarding "Usability".

Statements

Statement 1: The way the system is presented is clear and understandable.

Statement 2: I am able to use the software without help of coworkers.

Statement 3: The software provides the necessary help information about conceptual aspects of the program.

Statement 4: The system was fast enough.

Statement 5: The software did always what I was expecting.

Statement 6 I would recommend this software to colleagues.

Statement	Group	n	Mean[a]	StDev	Median	Range
St.1	ebbits	7	1.714	0.756	2	2
	Moknow	5	2	0.707	2	2
St.2	ebbits	7	2.429	0.535	2	1
	Moknow	5	2	0.707	2	2
St.3	ebbits	7	2.429	1.134	2	3
	Moknow	5	2	0	2	0
St.4	ebbits	7	1.857	1.069	2	3
	Moknow	5	2.4	0.894	2	2
St.5	ebbits	7	2.143	0.69	2	2
	Moknow	5	2.6	0.894	2	2
St.6	ebbits	7	1.857	0.69	2	2
	Moknow	5	2.2	0.837	2	2

[a]Response values range from 1 (strongly agree) to 5 (strongly disagree). Lower mean, median ratings indicate a more positive rating.

Table 7.1: Descriptive statistics—"Usability"

Comparison

Referring to Table 7.1, a descriptive comparison of the rated statements highlights the differences in ebbits and Moknowpedia orientation towards "Usability".

Compared to Moknowpedia participants, the mean ratings of the majority of the statements for the ebbits participants are slightly smaller, thus ebbits participants rated slightly more positive. Still, the median ratings are equal in both cases.

The highest rated statement from both of the groups is the statement 1, namely: "The way the system is presented is clear and understandable".

Moknowpedia participants *are able to use the software with-*

out help of the coworkers and they believe *enough help is given about the conceptual aspects of the program*, while ebbits participants slightly agree.

With regard to individual items, the greatest discrepancies between ebbits and Moknow involved statements 4 and 5, namely: "The system was fast enough", and "The software did always what I was expecting". These statements were rated with a mean of 1.857 and 2.143 by ebbits participants compared with 2.4 and 2.6 by Moknowpedia participants.

Both groups would recommend this software to their colleagues.

The mean ratings for each statement in every group is less than 2.6 (Median 2), which shows an overall satisfaction of both groups regarding the usability principles of ReputationForge.

7.1.2 Qualitative Results

This section shows the main remarks given by the participants of both groups related to "Usability.

Remarks in ebbits

Remarks regarding "Task Description"

Trajan[1] claims: "One drawback I remember is that I was only able to read a quest description by holding the mouse pointer over a field. I would expect this to be displayed on an ordinary Web page".

Remarks regarding "Browser Problems"

Eda complained that ReputationForge was not functioning properly in Internet Explorer. Vaadin has compatibility issues with Internet Explorer. Additionally, Trajan claimed that during the evaluation there was a version of the ReputationForge online which did not work with his browser (Mozilla Firefox). After updating his browser it worked again.

[1] Names changed for anonymization

Remarks in Moknowpedia

Remarks regarding "Task Description"

Jim and Tom claimed that the description of some tasks was sometimes a bit confusing or not detailed enough, thus the text could be improved there. Additionally, Tom claimed that there was some kind of formula explaining the contributions and that he does not get motivated from mathematical formulas in a task description.

Remarks regarding "Single sign-on"

Jim and Jerry claimed that the single sign-on failed to work in some cases and that is why they got unexpected crashes. They said: "I just tried to open ReputationForge which did not work. I got the 404 Status error although I logged in to Moknowpedia before".

7.1.3 Conclusion

Positive feeling

Based on the quantitative results explained above, there is a positive feeling regarding the usability of ReputationForge. Thus, the "Usability Hypothesis" along with the defined usability requirements in Section 4.1.3—"Requirements" is satisfied.

7.2 ReputationForge Tasks

This section describes the qualitative and quantitative results of the statements related to the "ReputationForge tasks".

7.2.1 Quantitative Results

This section examines the agreement of the users' ratings with the statements related to the "ReputationForge tasks".

Statement 1: The tasks are diverse.

Statement 2: The tasks are interesting.

Statement 3: The feedback of each task is clear enough and I knew the exact effort to put in to complete the task.

Additionally, participants were asked about the perceived difficulty of the tasks.

Statement	Group	n	Mean[a]	StDev	Median	Range
St.1	ebbits	7	2.286	0.488	2	1
	Moknow	5	3	1	3	2
St.2	ebbits	7	2.143	0.378	2	1
	Moknow	5	3	0.707	3	2
St.3	ebbits	7	2.571	0.976	2	2
	Moknow	5	2	0	2	0

[a]Response values range from 1 (strongly agree) to 5 (strongly disagree). Lower mean, median ratings indicate a more positive rating.

Table 7.2: Descriptive statistics—"ReputationForge Tasks"

Comparison

Referring to Table 7.2, a descriptive comparison of the rated statements highlights the differences in ebbits and Moknowpedia orientation toward "Tasks in ReputationForge".

Figure 7.1 shows clearly that both groups believe the introduced tasks are neither too easy nor too difficult.

Compared to Moknowpedia participants, the mean and median ratings of ebbits participants for the first two dimensions (statements) were less than 2.5 (positive feeling), while Moknowpedia participants share an undecided opinion thereof.

With regard to the last item, Moknowpedia participants believe that the feedback of each task is clear enough, while ebbits participants slightly agree.

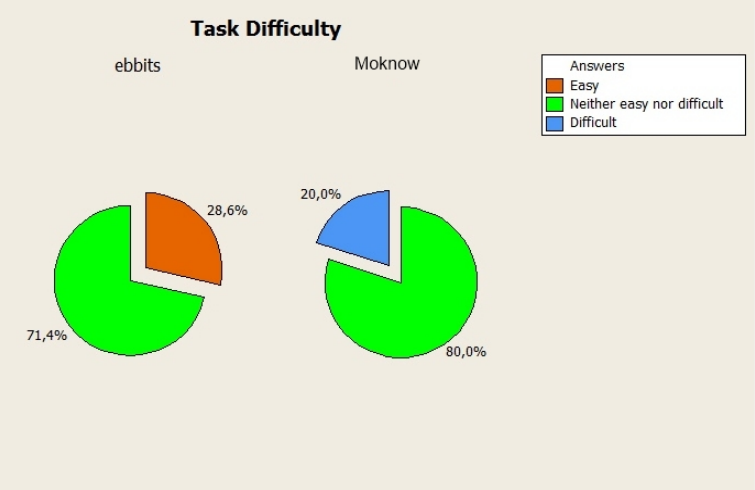


Figure 7.1: Task difficulty in ReputationForge

7.2.2 Qualitative Results

This section shows the main remarks given by the participants of both groups related to "Tasks in ReputationForge".

Remarks in ebbits

While asked what could have been done differently with tasks, *Tom claimed*: *"As I am busy, I find the tasks difficult. Some easier tasks would have been good.* The participant unfortunately did not mention what kind of tasks he would have liked.

Tim claimed: *"Better prepare tasks, their difficulty varied a lot"*.

Jina commented the difficulty of the task titled "Two reviews in two days" as too restrictive and not sufficiently explained that the user is not allowed to do the two reviews on the same day. The explanation was indeed given, but the participant was confused by the title of the task.

7.2.3 Conclusion

Based on the above-mentioned results, introduced tasks were perceived as neither easy too nor too difficult. However, groups had different opinions regarding the variety of tasks and their motivating effect. Thus, an upcoming prototype could introduce additional tasks to attract the participants even more.

7.3 Popularity of Motivation Mechanisms

Table 7.3 shows the overall ranking (means) of each single aspect of ReputationForge concept for each team wiki: Moknowpedia and ebbits wiki.

Similar ratings

According to a performed Man-Whitney U test, the medians are statistically similar indicating that both groups rated in a similar way ($W = 269, \rho = 0.8648 > 0.05$).

Generalisation of the outcomes

Thus, a generalisation of the tested motivational aspects is possible to some extent, albeit a perfect external validity can never be achieved, because of the existence of confounding factors, that could disturb the validity of the results.

Conclusions based on the outcomes

Based on the rankings shown in Table 7.3, the following qualified statements can be made about the common preference of both groups Moknow and ebbits towards the conceptual aspects of ReputationForge.

7.3.1 Aspects that motivate both groups

31.5% of the aspects of ReputationForge concept motivate/partly motivate the participiants of both Moknowpedia and ebbits wiki. Those are:

- Displaying recent activities.
- Receiving an email-notification upon task completion.

Keywords[a]	ebbits[b]	Moknow	Difference
Displaying recent activities	2,14	2,2	0,06
Email *You completed a task*	2,14	2,2	0,06
Differentiating between newbies and advanced users	2,43	2,8	0,37
Additional info about own task status while viewing others' tasks	2,57	2	0,57
Choosing between remaining tasks	2,57	2,4	0,17
Displaying users who started the same task	2,57	2	0,57
Displaying leaderboards in Gforge	2,70		
Starting manually tasks	2,71	2,8	0,09
Associating with every user a personal profile	2,71	2,2	0,51
Email *You did not complete any task*	2,86	3,8	0,94
List of unrated articles	2,86	3	0,14
Leaderboards of both newbies and advanced users	2,86	3	0,14
Embedding Recent Activities in Gforge	3,00		
List of articles, a user contributed to	3,00	3	0,00
Displaying the started task in Gforge	3,14		
Partial transparency of reputation scores	3,29	3	0,29
Email *You belong to the worst users*	3,57	3,6	0,03
Unlocking of new tasks only if previous tasks are completed	3,71	2,8	0,91
Punishement task	4,00	4,2	0,20
Average			0,32

[a] The keywords reflect the original statements of the questionnaire in Appendix A—"User Evaluation Questionnaire".

[b] Response values range from 1 (strongly agree) to 5 (strongly disagree). Lower mean ratings indicate a more positive rating. Lower differences indicate similar ratings.

Table 7.3: Ranking of the major aspects of ReputationForge concept

- Additional info about own task status while viewing others's tasks.
- Choosing between remaining tasks.
- Displaying users who started the same task.

7.3.2 Aspects with a neutral effect on motivation on both groups

31.5% of the aspects of ReputationForge concept had a neutral effect on the participants of both Moknowpedia and ebbits. Those are:

- Manually starting tasks. Two interviewees claimed as follows:
  - *Jina*, a participant of ebbits claimed that she failed to see the need to start the task by herself.
  - *Tom*, a participant of Moknow claimed: "I started the Rating-In-Two-Days task on a Friday, and of course I failed it, because I didn't rate on Saturday. I propose having tasks, which are automatically checked for completeness, and not started by the participants".
- List of unrated articles.
- Leaderboards of both newbies and advanced users.
- List of articles, a user contributed to.
- Partial transparency of reputation scores.

Additionally, ebbits participants are undecided whether the introduced changes in the "Summary Page" of the Gforge project motivate them to progress in ReputationForge, and thus participate more actively in the team wiki.

Discussion

Longitudinal study for the aspects with a neutral effect

Based on such an outcome, it would be a good idea to test ReputationForge in a longer evaluation period, aiming at finding out whether these aspects with a neutral effect on the target audience would then be accepted or rejected. Additionally, the existence of a neutral response in the survey

can discourage the in-depth thinking of participants, because they would tend to choose the easy way out by selecting "a neutral opinion". A survey without that choice, would make participants think harder.

One of the requirements of this thesis project, was the development of a prototype based on partially transparent reputation scores. However, both groups were rather undecided whether a partial transparency would motivate them to progress through the game, and thus participate more actively in wiki. Additionally, showing a list of unrated articles, and articles with no contribution was marked as desired by Dencheva (2010). Still, both groups had a neutral opinion on the motivating effect of such lists.

Participants' needs change

These results confirm the fact that requirements change during the development process of a software, and that several iterations are needed for a successful final system.

7.3.3 Aspects that motivate none of the groups

12.5% of the aspects of ReputationForge concept did not motivate the participants of both Moknowpedia and ebbits. Those are:

- Systematically receiving a notification email "You belong to the worst users".
- Introduction of a punishment task.

7.3.4 Biggest discrepancies between groups

For the the remaining 25% aspects (Statements 3, 9, 10, and 18), groups shared different opinions.

ebbits participants do not agree with the idea of "unlocking new tasks only if previous tasks are completed", while Moknowpedia participants are slightly undecided.

Moknowpedia participants are undecided whether splitting users in newbies and advanced is effective, while ebbits participants slightly agree of its effectiveness. On the other hand, ebbits participants are slightly undecided about the effectiveness of a personal profile, while Moknowpedia participants are confident about its effectiveness. However, median ratings were the same for both groups (Median = 2).

These discrepancies can be attributed to the local differences that exist between the groups. Nevertheless, the only logical connection that could be found was the discrepancy in answers regarding statement 10: "Associating a profile with each user. Noting that Moknow has a relatively old wiki, with established reputations of each participant, the concept of a profile, which displays the participant's reputation suits better this group as ebbits. It could not be found any logical connection with the remaining discrepancies described above.

7.3.5 Conclusions

Based on the above considerations, the conceptual aspects can be categorized in the following motivational *categories*:

Social/Curiosity

Participants prefer: *(a) displaying recent activities of other users, (b) displaying users who started the same task as them, and (c) displaying additional info about own task status while viewing others' tasks*.

Notification Category

Participants prefer receiving a single email upon task completion, rather than systematic emails from time to time.

Reputation

Participants partially agree about the effectiveness of "differentiating between newbies and advanced users", and "associating with every user a personal profile".

Self-Determination

Participants prefer choosing between remaining tasks, and letting the system check tasks automatically for completeness without the need to start them manually.

Rewarding Category

Participants strongly believe that enduring changes are achieved by the existence of only positive reinforcers. Punishment mechanisms should be excluded.

7.4 Overall Feeling

This section describes the procedure for analyzing the "Overall feeling towards ReputationForge".

7.4.1 Quantitative Results

The following examines the agreement of the participants' ratings related to the "Overall feeling towards ReputationForge".

Statement 1: I feel disadvantaged by the game.

Statement 2: The game motivates me.

Statement 3: I like the concept of social game implemented by ReputationForge.

Statement	Group	n	Mean[a]	StDev	Median	Range
St.1	ebbits	7	4,143	0,69	4	2
	Moknow	5	4,4	0,548	4	1
St.2	ebbits	7	2,429	0,535	2	1
	Moknow	5	2,6	0,894	2	2
St.3	ebbits	7	2,286	0,756	2	2
	Moknow	5	2,4	0,894	2	2

[a]Response values range from 1 (strongly agree) to 5 (strongly disagree). Lower means and medians indicate a higher level of agreement with the specific statements.

Table 7.4: Descriptive statistics— "Overall Feeling towards ReputationForge"

Comparison

Referring to Table 7.4, a descriptive comparison of the rated statements highlights the orientation differences in ebbits and Moknowpedia participants towards the overall feeling of ReputationForge. Figure 7.2 gives the reader a better visualization of the descriptive statistics displayed in Table 7.4.

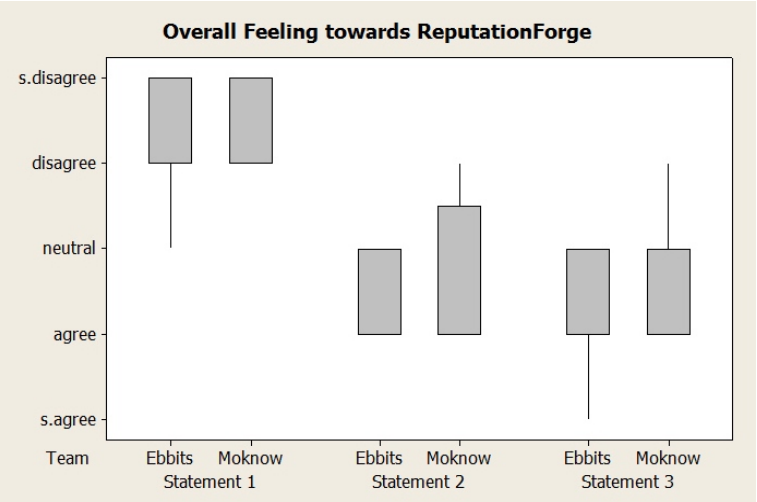


Figure 7.2: "Overall Feeling" towards ReputationForge— Boxplots

No disadvantage

Both groups strongly believe that they do not get disadvantaged by ReputationForge.

General positive feeeling

With regard to the "Overall motivation", and "Concept approval", there are no noticeable discrepancies between the groups. Based on the mean ratings, ebbits participants rated slightly more positive than Moknow participants. However, the median scores show no difference between the groups (Median=2). As a whole, the social game concept was positively received from both the groups. Similarly, both groups feel motivated from ReputatationForge (Median = 2), though median ratings show a rather slight motivation.

Preference of Participants

Participants prefer ReputationForge

While explicitly asked, which system motivates them more: ReputationForge or the existing system (only relevant for those users who already used the existing system), 62.5% of the eight asked participants claimed to be more motivated from ReputationForge (cf. Figure 7.3). 12.5% were undecided while the rest preferred the existing system more.

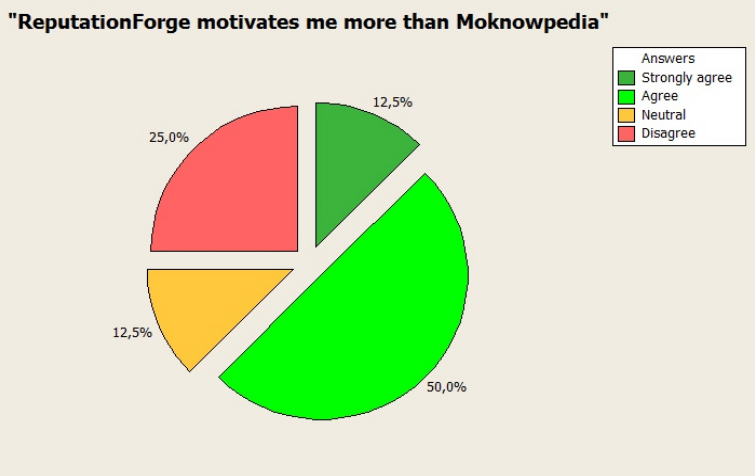


Figure 7.3: ReputationForge vs Moknowpedia—Participants' preference

7.4.2 Qualitative Results

This section shows the main remarks given by the participants of both groups related to "Overall Feeling of ReputationForge".

Remarks in ebbits

While asked what could have been done differently to increase the motivation effects of ReputationForge, *Tom claimed* the following: "Motivating people to contribute to wiki is a highly complex problem. When I have something interesting to put in the wiki I do it. If not, I don't write something regardless if there is a task to do so. During the last two weeks I did not have much to write in the Wiki so it was hard to select a task after having finished the rating-tasks."

Jerry, as the single participant (displayed as an outlier in Figure 7.2), who gave overall negative ratings for ReputationForge claimed: *"In my opinion, the underlying CollabReview system is useless. I totally disagree that CollabReview is needed. Therefore, I also do not think that ReputationForge is needed because its only sense is to support the use of CollabReview"*.

Remarks in Moknowpedia

Trajan claimed that ReputationForge motivated him to rate articles. He would not have rated articles without ReputationForge. Overall he finds the concept of embedding a social game in wiki very promising and interesting.

George claimed that he did not actively use the ReputationForge system, but used it passively to see the activity level of himself compared to others, which indirectly influenced him to increase his participation in the wiki. A last user

Jina claimed that she is not experienced with social games, and is skeptical whether the introduction of a game would increase the participation and motivation to use wiki. In some circumstances, it can have a totally contrary effect.

7.4.3 Conclusion

Participants were generally motivated and agreed that ReputationForge concept was suitable. Furthermore, participants who used the old system Moknowpedia prefer using ReputationForge from now on. Thus, the *Motivation Hypothesis*, and the *Comparison Hypothesis* (cf. Section 6.4) are satisfied.

7.5 Summary

Outlook of this chapter

This chapter continued the implementation evaluation of ReputationForge, and completed therewith the whole evaluation process of this thesis project. The post-test questionnaire allowed to get insight into the usability of ReputationForge, the overall satisfaction and motivation in general, and the motivational aspects of ReputationForge concept in particular. It also helped to collect evidence whether to support or deny the hypotheses set in Section 6.4—"Hypotheses". Furthermore, the performed qualitative analysis identified remarks of single participants, that need to be taken into consideration in the design process of upcoming prototypes.

Insight into the next chapter

The following chapter concludes this thesis by describing the main contributions of this project, namely the design, implementation, and evaluation of *ReputationForge*, a social-game based on reputation, aiming at increasing the participants' engagement in the wiki of ebbits and Moknowpedia. It furthermore discusses to which extent the objectives were met, and the implications of the findings. Last but not least, it gives future research suggestions, that could be considered in the upcoming iterations of ReputationForge.

Chapter 8

Conclusions

"The future is not a result of choices among alternative paths offered by the present, but a place that is created–created first in the mind and will, created next in activity. The future is not some place we are going to, but one we are creating. The paths are not to be found, but made, and the activity of making them, changes both the maker and the destination"

—John Schaar

A user-oriented design of a social-game based on reputation

This thesis project presented the entire user-centred design process adopted to develop *ReputationForge*, a social persuasive reputation-based game to increase the participants' engagement in the team wikis of the groups: (a) Moknow and ebbits. The *entertaining side* of persuasive tools, aiming at changing the end users' behavior is an emerging novel paradigm nowadays.

Initial spark

The initial spark for this project came from the findings of Dencheva et al. (2011). The results of the introduced extensions to the team wiki of the groups Moknow and CAPLE, aiming at increasing its participation rate were satisfactory, albeit only a temporary motivation was observed, and additional incentives were desired.

Deep conceptual investigation

An important contribution of this thesis project was the extensive conceptual investigation aiming at identifying motivational values that lead the target audience towards higher levels of participation. Yet, the motivational values derived from the conceptual investigation were of hypothetical nature, and most of them were applicable for public wikis.

Final concept of ReputationForge

An extensive analysis of the state of the art of the areas that most influenced this thesis project, and the involvement of the target audience throughout the design process led to the definition of the final concept of this thesis. Aspects like *community/social, reputation, goal-setting, personal-enjoyment and self-fulfillment* are the main ingredients of the concept of ReputationForge, which promotes a public soft competition between the participants, who are divided into two levels: *(a) newbies, and (b) advanced*. Participants are awarded respectively reputation/karma points upon completion of individual tasks. Tasks are directly related with the work in the team wiki, such as creating and/or improving documentation articles, and rating them. Clear feedback related to the progress of each task is given. Finally, ReputationForge introduces triggers such as email notifications, changes to the main page of the research project, and a list of unrated articles, aiming at making the underlying wiki more visible and thus attracting users to keep coming back to it.

Technical Contributions

A further contribution of this thesis project was the implementation of a first prototype of *ReputationForge*, based on the final concept. Before starting the implementation of *ReputationForge*, the final concept was first validated and accepted by the target audience. A prerequisite for designing a successful prototype was its interaction with: (a) Collabreview, and (b) Gforge in case of the ebbits evaluation and MediaWiki in case of the Moknowpedia evaluation. ReputationForge itself was developed as CollabReview extension using Vaadin—a framework used for the development of Web-based Rich Internet Applications (RIA). To give the users the feeling that they were using a single system instead of three different systems, a SSO mechanism was created. Additionally, the goal of integrating ReputationForge within Gforge was achieved by creating a Gforge project plugin and linking it to the ebbits project. Further-

more, a rating form was embedded in Gforge. A further technical contribution was the design of an intuitive UI, which incorporated all of the motivational values of the final concept. The realization of a modular architecture allowed to add new programmatic tasks easily, thus a principal technical contribution of this thesis project.

Last but not least, a major contribution of this thesis project was to evaluate ReputationForge, along with the hypothetical motivational values of its concept in the two different working environments—Moknow, having a well-established team wiki (Moknowpedia) with a satisfactory participation rate—and ebbits having a relatively new wiki with low participation rate. A case study evaluation was performed in combination with a control group field study. It was opted for a field study, mainly because a real change in the participants behavior is expected to occur when users use the system in their own normal work environment rather than in a laboratory study. Furthermore, two case studies were chosen due to their indicative dimension, which aim is as Denscombe (2007) states "to illuminate the general by looking at the particular", and thus leading to a more plausible external validity of the results. Additionally, the introduction of a control group contributed to a higher degree of the internal validity of this thesis project.

Evaluation of ReputationForge through a control group field study in both ebbits wiki and Moknowpedia

8.1 Main Findings

The research questions presented in Section 1.4—"Research Questions" set the following objectives for this thesis project:

Main research questions

- *Design a social-game that increases the participants' engagement in the team wikis.*
- *Find out the overall satisfaction and motivation of the target users regarding the new concept of ReputationForge.*
- *Find out the most suitable aspects that lead towards a higher participation rate in team wikis.*

The following describe how these objectives were met.

ReputationForge and its Ergonomicity

Positive ratings related to the ReputationForge usability

Most evaluation participants, in both groups, without receiving any training, found ReputationForge easy to use, easy to learn and be understood. Thus, the goal of having an ergonomic system as a prerequisite for a consequent behavior change in the target audience was satisfied.

Participant's Engagement in both Case Studies

Greater interest from the Moknowpedia participants

The evaluation of the social game usage in both cases pointed out more interest towards ReputationForge in the Moknowpedia participants than in the ebbits participants. Such a result was due to differences between the team wikis of both groups, namely Moknowpedia is used massively for daily information exchange, whereas the engagement in the internal wiki of ebbits depends on the actual progress of the project. If there is no progress, there is no need to document in wiki.

Improvement of engagement in both cases

Furthermore, the outcome evaluation showed an increased engagement of the experimental group in the wiki of both case studies, thus proving evidence that a social-game approach based on reputation can actually increase the participantsí engagement in a corporate/team wiki.

Overall significant improvement in ebbits; a significant increased number of rated articles in Moknowpedia participants

Still, participants of the ebbits experimental group had statistically significant higher mean engagement metrics, i.e contributions of any authorship ($\rho = 0.05$), contributions of at least 30% authorship ($\rho = 0.081$), and the average quality of articles ($\rho = 0.034$), whereas participants of the Moknowpedia experimental group showed a statistically significant increase only in relation to the number of rated articles ($\rho = 0.019$). Such a result of the latter was attributed to factors like the relatively well-established internal wiki, the short and/or unsuitable evaluation period, and the further involvement of the participants of the control group in the existing system Moknowpedia.

ReputationForge Tasks

Task features

Participants perceived the introduced tasks as neither easy nor difficult. They found the tasks' feedback to be clear. Whereas, regarding the variety of tasks and their motivating effect, they had different opinions. Thus, an upcoming prototype could introduce additional tasks to attract the participants even more, as described in Section 8.2—"Outlook".

Main Aspects of ReputationForge Concept

Generalisability of the outcomes

An extensive analysis of the major aspects behind the ReputationForge concept showed overall similar ratings from both Moknow and ebbits. Such a result allows for replication of the aspects under study, and thus contribute to the external validity of this thesis project.

Major motivational values

Furthermore, the analysis pointed out the following motivational *categories* for both ebbits and Moknow:

- *Social/Curiosity*: Displaying recent activities of other users, displaying users who started the same task as them, and displaying additional info about own task status while viewing others' tasks were reported as motivating from both groups.
- *Reputation*: Participants partially agree about the effectiveness of "differentiating between newbies and advanced users", and "associating with every user a personal profile".
- *Triggers*: Receiving a single email upon task completion proved to be motivating, whereas receiving systematic emails from time to time were perceived as obtrusive.
- *Self-Determination*: Choosing between remaining tasks, and letting the system check the tasks automatically for completion without the need to start them manually were preferred by the participants.

- *No Punishment techniques*: The existence of only positive reinforcers is a strong motivational value for the participants.

Overall Feeling towards ReputationForge

General positive feeling

As a whole, the social game concept was positively received from both the groups. Similarly, both groups felt motivated from ReputatationForge. The majority of participants preferred using ReputationForge more than the existing system Moknowpedia (only relevant for those participants who already used the existing system).

Furthermore, the evaluation pointed out the main weaknesses of the concept, and technical issues of ReputationForge, which gives way to improvements and enhancement of the system as described in the following paragraphs.

8.2 Outlook

Participants do not prefer systematic e-mails

The idea behind the introduction of triggers was: *(a) to increase the visibility, and the presence of the underlying wiki, and (b) to increase the ability of the participants in performing specific tasks related to wiki*. As already described, systematic email notifications were perceived as obtrusive. Therefore, an upcoming prototype could allow the participants to self-configure the way how they are notified or even register future events (i.e tasks, or activities related to the wiki) they want to be notified about, which supports the concept of awareness of future events as described by Prinz et al. (2008).

No negative reinforcers

The evaluation proved evidence that an effective rewarding system must not combine rewards with punishment techniques to reinforce desired behavior or extinguish unwanted behavior. Participants clearly stated that a rewarding system should always positively reward them. Thus, an upcoming prototype should relinquish the introduction of punishments to influence the human behavior.

User needs change during the design process

Additionally, the concept evaluation proved evidence that user requirements change during the software design process. Aspects such as partial reputation transparency, and a list of unrated articles/no contribution were highly desired before the introduction of ReputationForge, but introduced only a neutral effect after the evaluation ended. Thus, having continuous iterations is the key to a final successful system.

Further investigate the differences between the groups

Discrepancies in the participants' preference towards particular motivational values were attributed to the local differences that exist between the groups. A further detailed investigation into the differences between groups, including the demographic differences should be part of a future research work.

Introduction of additional tasks

Aiming at keeping the participants interested in the social-game for a longer time, it would be meaningful to introduce a variety of new types of tasks related to wiki, or even source code or issue trackers. Introducing collaborative tasks to be completed in groups, and/or a public contest between groups would be a promising approach for the upcoming iterations.

A peer-based, recommender system

An upcoming prototype could apply the foundation principles of peer-based systems, allowing the participants to follow members, and join a group with similar interests and likings. Furthermore, the game could recommend the users at specific points in time to complete a particular task that is with high probability a task that is strongly related to his personal likings or interests.

Exchanging reputation scores with tangible rewards

An idea would be also the introduction of a bank metaphor, namely extending the overall score of a participant to be at the same time a credit score that could be exchanged with money or with different tangible rewards. Nevertheless, such an approach based on extrinsic motivations according to the studied literature would prove to be unsuccessful and cannot be used as a long-lasting motivating factor.

Technical issues with Vaadin

Of course, there were also some technical issues related to Vaadin—the web framework chosen to develop ReputationForge. The incompatibility with Internet Explorer, and a poor response performance from time to time indicate the need to further investigate on how to optimize ReputationForge, or even choosing other web frameworks.

Programming tasks as OSGi bundles

The introduction of a modular architecture, to easily add new tasks proved to be successful. Still, an even better solution for the upcoming prototypes would be programming new tasks as bundles in an Open Services Gateway initiative (OSGi) environment that can be at any time installed, started, stopped, updated and even un-installed without the need of the redeployment of the whole application.

Longitudinal study to prove evidence of enduring changes in behavior

As a last consideration, the reactions towards a social game approach based on reputation were positive, thus it is worth a further extensive investigation on the topic. However, a longitudinal nature of an upcoming study evaluated by a larger number of participants would allow to monitor an enduring preference and the long-term effects of such an approach on the participants engagement in a corporate/team wiki. Such a longitudinal study would further help find out whether the motivational aspects of ReputationForge that registered a neutral effect on the target audience would in a long term be accepted or rejected. This would be definitely a thriving challenge of future research work, that no one should be afraid of. As Steve Jobs said “Stay hungry, stay foolish”.

Appendix A

User Evaluation Questionnaire

The following contains the surveys that were used during the user-centred design process of ReputationForge. Both survey consisted of Likert scale questions, rated from 1 = Strongly disagree to 5 = Strongly agree. Further non Likert-based Itemes were asked also. Users were encouraged to add additional comments. The results of the first survey are discussed in Chapter 4—"Finding a suitable concept", while the results of the second survey are discussed in Chapter 7—"Implementation Evaluation".

A.1 ReputationForge Pre-Testing Survey

A big thank you to all the persons that will dedicate some precious time to answer the questions below.

A time of 15-20 minutes will be needed to completely fill this survey.

User Acknowledgement:
The aim of this questionnaire is to gather valuable data about my master thesis. Data gathered is stored anonymously and will not be used for any other purpose, and will not be given to any third party. The summarizing results of this questionnaire will be published in the final thesis document.

Preliminaries: Duration 2-4 minutes

Are you registered in Gforge?

○Yes ○No

If previously answered yes, how often do you use Gforge?

○	○	○	○	○
Never/Less than 1 month	Rarely/ Once in a moth	Sometimes/Once in a week	Often/Several times in a week	Every day

Did you use Moknowpedia?

○Yes ○No

If previously answered yes, how often do you use Moknowpedia?

○	○	○	○	○
Never/Less than 1 month	Rarely/ Once in a moth	Sometimes/Once in a week	Often/Several times in a week	Every day

Do you think that a system like Moknowpedia would be used as a means of controlling your performance and not actually promoting your performance?

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

In case you used Moknowpedia, were you overall satisfied?

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

How familiar are you with the following web reputation systems?

	Totally unfamiliar	Somewhat unfamiliar	Somewhat familiar	Totally familiar
Amazon	○	○	○	○
Ebay	○	○	○	○
Digg	○	○	○	○
Foursquare	○	○	○	○
Screentribe	○	○	○	○
Stackoverflow	○	○	○	○

Software Prototype

I will kindly ask you to listen before to this screencast video where I propose my system concept, before answering the upcoming questions:

http://www.youtube/com/watch?v=UZHVIX6T4Is

If you want to have a look to a written text or even try the software prototype yourself here the link:

http://moknowpedia.fit.fraunhofer.de/mediawiki/index.php5/Master_Thesis_Concept

Concept Validation/Duration: 10 min

It is easy to get a first understanding of the system.

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

Would the mechanisms below, already mentioned in the system concept description, motivate you?

Sending regular emails, where you get informed about the progress in the new system.

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

Links/Advertisements that redirect users to the new system.

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

Embedding the recent activity sections of ReputationForge and Ebbits wiki.

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

The proposed game concept relies on a predefined set of tasks. Upon task completion, the users can choose between the remaining tasks available which get unlocked. Tasks are organized in categories, of different level of difficulty. I am satisfied.

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

Would the social game motivate you during the project duration to participate actively in the internal wiki?

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

I have the feeling of being excluded by this concept.

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

If you feel excluded, please list the reason(s) why:

_____
_____
_____
_____
_____
_____
_____

I would feel comfortable while using the system.

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

I think this concept is conformant to internal corporate rules.

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

I am overall satisfied with this concept.

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

Please share any additional comments

_____
_____
_____
_____
_____
_____
_____

General Information/Duration: 5 min

What is your gender?

○Female ○Male

Are you a computer scientist or studied something related to computer science?

○Yes ○No

What is your profession? _____

I like challenges in my everyday work.

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

I get motivated by new ideas and proposals in my everyday work.

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

In your everyday work, do you just try to finish your assigned tasks, or try also to contribute to the team in general?

○	○	○
I only try to get my work done	Neutral	I try at the same time to help the team

In your everyday work, when tasks are assigned to you, what is more important for you?

○	○	○	○
A clear visible path to the completion of tasks.	Easy tasks	Challenging tasks	Neither easy nor challenging tasks

Thank you for taking the time to fill out this survey. Your input is greatly appreciated.

A.2 ReputationForge Post-Testing Survey

A big thank you to all the persons that will dedicate some precious time to answer the questions below. A time of 10 minutes will be needed to completely fill this survey.

User Acknowledgement:
The aim of this questionnaire is to gather valuable data about my master thesis. Data gathered is stored anonymously and will not be used for any other purpose, and will not be given to any third party. The summarizing results of this questionnaire will be published in the final thesis document.

Your Name:
This field is optional. Providing your name will allow me to make a deeper investigation of the effects of ReputationForge, and to therefore write a better evaluation. So, please provide your name.

ReputationForge- Participation during the evaluation period

○ Never used it ○ Used it

ReputationForge- General Usability questions

The way that system information is presented is clear and understandable.

○ Strongly agree ○ Agree ○ Neutral ○ Disagree ○ Strongly disagree

I was able to use the software right from the beginning by myself, without having to ask coworkers for help.

○ Strongly agree ○ Agree ○ Neutral ○ Disagree ○ Strongly disagree

The software provides the necessary help information about conceptual aspects of the program.

○ Strongly agree ○ Agree ○ Neutral ○ Disagree ○ Strongly disagree

The system was fast enough.

○ Strongly agree ○ Agree ○ Neutral ○ Disagree ○ Strongly disagree

How often did systems errors (i.e crashes) occur?

○	○	○	○	○
Never	1-5 times	5-10 times	10-20 times	More than 20 times

I would recommend this software to my colleagues.

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

Please share any additional comments.

_____
_____
_____
_____
_____
_____

ReputationForge- Questions about the game/tasks

A feature of ReputationForge is that it sends regularly emails. Please consider the following.

ReputationForge sends me every week an e-mail with the content "You belong to the worst users" in case I am among the worst users. This motivates me to progress through the game.

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

ReputationForge send me every week an e-mail "You did not complete any tasks the last week" in case I am inactive in ReputationForge. This motivates me to progress through the game.

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

Upon successful completion of tasks, ReputationForge send me a message with a link directing me to the task. Because of it, I use ReputationForge more often.

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

The tasks of the game are:

○	○	○	○	○
Too easy	Easy	Neither easy nor difficult	Difficult	Too difficult

ReputationForge's main idea is the introduction of tasks, that participants need to complete. Please consider the following:

The task variety is big enough.

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

The tasks are interesting.

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

The feedback of each task is clear and I knew the exact effort to put in to complete the task.

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

Users can start the next task, only if they complete the previous task. This increases my curiosity about the game.

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

Upon starting a task, users can see who are the users that started the same task. This motivates me to complete tasks.

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

In case a user has a look to the task of another user, ReputationForge shows to the user whether he also completed that task. This motivates me to complete the task.

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

Once I complete a task, I can choose between the remaining tasks. This give me the possibility to choose and this motivates me.

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

I can start a task whenever I want. This motivates me.

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

ReputationForge introduces a "punishment task". I agree with it.

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

Did you find the description and feedback of some tasks difficult to be understood? Please give the name of the task(s) and the reason why?

_____
_____
_____
_____
_____
_____

ReputationForge- Incentives

ReputationForge displays recent activites of other members. This feature motivates me.

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

ReputationForge motivates newbies by giving them more additional points as advanced users. As a newbie, I get motivated to progress through the game.

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

ReputationForge displays the reputation scores of the best 4 users and best 4 newbies. I struggle that my name appears in one of these lists.

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

ReputationForge does not show the reputation score (except the top x in the ranking lists) to other members. This motivates me to progress through the game.

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

ReputationForge associates a personal profile with every users. The existence of the profile gives me invaluable information about my achievements within wiki, and therefore motivates me.

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

ReputationForge displays a list of the articles, that you rated or not. I rate more articles because of it.

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

ReputationForge displays a list of the articles, that you contributed to. I contribute more to wiki because of it.

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

Here some questions about the frequency of usage in ReputationForge :

I used ReputationForge during the evaluation period:

○	○	○	○
Daily	Weekly	Monthly	Never

After the evaluation ends, I will user ReputationForge:

○	○	○	○
Daily	Weekly	Monthly	Never

ReputationForge- Overall satisfaction

I have the feeling the social game disadvantages me.

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

The social game overall motivates me.

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

I like the concept of the social game implemented by ReputationForge.

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

Overall, I like the system.

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

ReputationForge motivates me more than the original system, which displays a ranking list of all the members and gives a weekly "Hero of the week" prize.

○	○	○	○	○
Strongly agree	Agree	Neutral	Disagree	Strongly disagree

The game disadvantaged me. Please give us your comment.

_____
_____
_____
_____
_____

What could have been done differently to increase the motivating effects of the system ?

_____
_____
_____
_____

The game did not motivate me. Here are some reasons:

No time ○

No real trigger to use ReputationForge ○

I do not like the concept of ReputationForge ○

Other: _____

ReputationForge- Reasons why you did not user ReputationForge

No time ○

No real trigger to use ReputationForge ○

ReputationForge was not accessible to me ○

The evaluation time was too short ○

Other: _____

Thank you for taking the time to fill out our survey. Your input is greatly appreciated.

Bibliography

Ackerman, M. S., Pipek, V., & Wulf, V. (2003). *Sharing Expertise: Beyond Knowledge Management*. Cambridge, MA, USA: MIT-Press.

Antikainen, M. & Väätäjä, H. (2010). Rewarding in Open Innovation Communities - How to Motivate Members? *International Journal of Entrepreneurship and Innovation Management (IJEIM)*, *11*(4), 440–456.

Bruegge, B. & Dutoit, A. H. (1999). *Object-Oriented Software Engineering: Conquering Complex and Changing Systems*. Upper Saddle River, NJ, USA: Prentice Hall.

Calder, B. J. & Staw, B. M. (1975). Self-Perception of Intrinsic and Extrinsic Motivation. *Journal of Personality and Social Psychology (JPSP)*, *31*(4), 599–605.

Carrell, M. R. & Dittrich, J. E. (1978). Equity Theory: The Recent Literature, Methodological Considerations, and New Directions. *Academy of Management Review (AMR)*, *3*(2), 202–210.

Clary, E., Snyder, M., Ridge, R., Copeland, J., Stukas, A., Haugen, J., & Miene, P. (1998). Understanding and Assessing the Motivations of Volunteers: A Functional Approach. *Journal of Personality and Social Psychology (JPSP)*, *74*(6), 1516–1530.

Clercq, J. D. (2002). Single Sign-On Architectures. In *Proceedings of the International Conference on Infrastructure Security (InfraSec'02)*, (pp. 40–58)., Bristol, UK, October 1-3. Springer.

Corr, P. J. (2008). The Reinforcement Sensitivity Theory of Personality and Psychopathology. *International Journal of Psychophysiology (IJP)*, *69*(3), 151–152.

Crumlish, C. & Malone, E. (2009). *Designing Social Interfaces*. Beijing, China: O'Reilly.

Dencheva, S. (2010). Entwicklung und Evaluation einer soziotechnischen Komponente zur dynamischen Selbstmoderation in einem Firmeninformationssystem. Diploma thesis, University of Bonn, Germany. (English Translation: Development and Evaluation of a Socio-Technical Component Supporting Dynamic Self-Moderation in a Corporate Information System).

Dencheva, S., Prause, C. R., & Prinz, W. (2011). Dynamic Self-Moderation in a Corporate Wiki to Improve Participation and Contribution Quality. In Bodker, S., Bouvin, N. O., Wulf, V., Ciolfi, L., & Lutters, W. (Eds.), *Proceedings of the 12th European Conference on Computer Supported Cooperative Work (ECSCW '11)*, (pp. 1–20)., Aarhus, Denmark, 24-28 September. Springer.

Denscombe, M. (2007). *The Good Research Guide for Small-Scale Social Research Projects*. Milton Keynes, UK: Open University Press.

Farmer, R. & Glass, B. (2010). *Building Web Reputation Systems* (1st ed.). Sebastopol, CA, USA: O'Reilly Media, Inc.

Fogg, B. J. (1998). Persuasive Computers: Perspectives and Research Directions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'98)*, (pp. 225–232)., New York, NY, USA, April 18-23. ACM Press/Addison-Wesley Publishing Co.

Fogg, B. J. (2003). *Persuasive Technology: Using Computers to Change What We Think and Do*. San Francisco, CA, USA: Morgan Kaufmann Publishers.

Fogg, B. J. (2009). A Behavior Model for Persuasive Design. In *Proceedings of the 4th International Conference on Persuasive Technology (Persuasive '09)*, (pp. 40:1–40:7)., New York, NY, USA, April 26 - 29. ACM Press.

Fogg, B. J. & Nass, C. (1997). How Users Reciprocate to Computers: An Experiment that Demonstrates Behavior Change. In *CHI '97 Extended Abstracts on Human Factors in Computing Systems: Looking to the Future (CHI EA '97)*, (pp. 331–332)., Atlanta, Georgia, USA, March 22-27. ACM Press.

French, E. B., Metersky, M. L., Thaler, D. S., & Trexler, J. T. (1973). Herzberg's Two Factor Theory: Consistency Versus Method Dependency. *Personnel Psychology*, *26*(3), 369–375.

Gamma, E., Helm, R., Johnson, R., & Vlissides, J. (1995). *Design Patterns: Elements of Reusable Object-Oriented Software*. Boston, MA, USA: Addison-Wesley.

GForge Group, I. (2011). An Introduction to Plugin Development for Gforge. Retrieved 15 February 2012 from: `http://gforgegroup.wordpress.com/2011/01/19/an-introduction-to-plugin-development-for-gforge`.

Gould, J. D. & Lewis, C. (1983). Designing for Usability: Key principles and What Designers Think. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI' 83)*, volume 28, (pp. 50–53). ACM Press.

Grönross, M. (2011). *Book of Vaadin* (4th ed.). Turku, Finnland: Vadin Ltd. `https://vaadin.com/download/book-of-vaadin/current/pdf/book-of-vaadin.pdf`.

Hakim, J. & Spitzer, T. (2000). Effective Prototyping for Usability. In *Proceedings of the 18th Annual ACM International Conference on Computer Documentation: Technology & Teamwork (IPCC/SIGDOC '00)*, (pp. 47–54)., Cambridge, MA, USA, September 24-27. IEEE Educational Activities Department.

He, J. & Chu, W. W. (2010). A Social Network-Based Recommender System (SNRS). In *Data Mining for Social Network Data* (pp. 47–74).

Hoisl, B. (2007). Social Rewarding in Online Communities - A Focus on Wiki Systems. Master's thesis, Vienna University of Technology, Austria.

Hoisl, B., Aigner, W., & Miksch, S. (2007). Social Rewarding in Wiki Systems - Motivating the Community. In *Proceedings of the 2nd International Conference on Online Communities and Social Computing (OCSC'07)*, (pp. 362–371)., Beijing, China, July 22-27. Springer.

Hunicke, R., Leblanc, M., & Zubek, R. (2004). MDA: A Formal Approach to Game Design and Game Research. In *Proceedings of the Challenges in Games AI Workshop, Nineteenth National Conference of Artificial Intelligence (AAAI '04)*, (pp. 1–5)., San Jose, CA, USA, July 25-29. MIT Press.

Huysman, M. & Wulf, V. (2004). *Social Capital and Information Technology*. Cambridge, MA, USA: MIT Press.

Huysman, M. & Wulf, V. (2006). IT to Support Knowledge Sharing in Communities: Towards a Social Capital Analysis. *Journal on Information Technology (JIT)*, *21*(1), 40–51.

Innovation Network, I. (2011). Evaluation Plan Workbook. `http://www.innonet.org/client_docs/File/evaluation_plan_workbook.pdf`.

Jarke, M., Klann, M., & Prinz, W. (2009). Serious Gaming: The Impact of Pervasive Gaming in Business and Engineering. In C. M. Schlick (Ed.), *Methods and Tools of Industrial Engineering and Ergonomics* (pp. 281–292). Springer.

Järvinen, A. (2008). *Games without Frontiers: Theories and Methods for Game Studies and Design*. PhD thesis, University of Tampere, Finnland.

Jensen, C., Davis, J., & Farnham, S. (2002). Finding Others Online: Reputation Systems for Social Online Spaces. In *Proceedings of the SIGCHI Conference on Human factors in Computing Systems: Changing our World, Changing Ourselves (CHI'02)*, volume 4, (pp. 447–454)., Minneapolis, MN, USA, April 20-25. ACM Press.

Jøsang, A., Ismail, R., & Boyd, C. (2007). A Survey of Trust and Reputation Systems for Online Service Provision. *Decision Support Systems*, *43*(2), 618–644.

Koch, M. & Prinz, W. (2008). Technische Aspekte von Social Software. In A. Back, N. Gronau, & K. Tochtermann (Eds.), *Web 2.0 in der Unternehmenspraxis* (pp. 78–84). Oldenbourg. (English translation: Technical Aspects of Social Software).

Kroth, M. (2007). Maslow-Move Aside! A Heuristical Motivation Model for Leaders in Career and Technical Education. *Journal of Industrial Teacher Education (JITE)*, *44*(2), 5–36.

Langville, A. N. & Meyer, C. D. (2006). *Google's PageRank and Beyond: The Science of Search Engine Rankings*. Princeton, NJ, USA: Princeton University Press.

Lenert, L., Muñoz, R. F., Stoddard, J., Delucchi, K., Bansod, A., Skoczen, S., & Pérez-Stable, E. J. (2003). Design and Pilot Evaluation of an Internet Smoking Cessation Program. *Journal of the American Medical Informatics Association (JAMIA)*, *10*(1), 16–20.

Lieberman, H., Paternò, F., & Wulf, V. (2006). *End User Development*. Dordrecht, NL: Springer.

Locke, E. A. & Latham, G. P. (2002). Building a Practically Useful Theory of Goal Setting and Task Motivation: A 35-year odyssey. *The American Psychologist*, *57*(9), 705–717.

Lundgren, S. & Björk, S. (2003). Game Mechanics: Describing Computer-Augmented Games in Terms of Interaction. In *Proceeding of Technologies for Interactive Digital Storytelling and Entertainment (TIDSE'03)*, (pp. 45–46)., Darmstadt, Germany, June 24-26. Fraunhofer IRB Verlag.

Martínez-Carreras, M. A., Ruiz-Martínez, A., Gómez-Skarmeta, A. F., & Prinz, W. (2007). Designing a Generic Collaborative Working Environment. In *IEEE International Conference on WebServices (ICWS'07)*, (pp. 1080–1087)., Salt Lake City, UT, USA, July 9-13. IEEE Computer Society.

Maslow, A. H. (1987). *Motivation and Personality*. New York, NJ, USA: HarperCollins Publishers.

Nadeau, T. (2007). Reputation 2.0. Retrieved 17 February 2012 from: `http://www.slideshare.net/ted.nadeau/sxsw-2007-reputation-20`.

Nov, O. (2007). What Motivates Wikipedians? *Communications of the ACM*, *50*(11), 60–64.

Petri, H. L. (1991). *Motivation: Theory, Research and Application* (3d ed.). Belmont, CA, USA: Thomson Wadsworth.

Poole, E. S. & Grudin, J. (2010). A Taxonomy of Wiki Genres in Enterprise Settings. In *Proceedings of the 6th International Symposium on Wikis and Open Collaboration (WikiSym '10)*, (pp. 14:1–14:4)., Gdansk, Poland, July 7-9. ACM Press.

Prause, C., Reiners, R., Dencheva, S., & Zimmermann, A. (2010). Incentives for Maintaining High-Quality Source Code. In *Psychology of Programming Interest Group Work-in-Progress Workshop (PPIG '10)*, (pp.6̃)., Dundee, UK, January 7-8. Psychology of Programming Interest Group Work.

Prinz, W., Hinrichs, E., & Kireyev, I. (2008). Anticipative Awareness in a Groupware System. In *Proceedings of the 8th International Conference on the Design of Cooperative Systems (COOP'08)*, Carry-le-Rouet, Provence, France, May 20-23. Institut d'Etudes Politiques d'Aix-en-Provence.

Resnick, P., Kuwabara, K., Zeckhauser, R., & Friedman, E. (2000). Reputation Systems. *Communications of the ACM*, *43*(12), 45–48.

Resnick, P. & Varian, H. R. (1997). Recommender Systems. *Communications of the ACM*, *40*(3), 56–58.

Robbins, S. P. (2004). *Organizational Behaviour: Concepts, Controversies, Applications* (4th ed.). Upper Saddle River, NJ, USA: Prentice Hall.

Rouse, R. (2000). *Game Design: Theory and Practice* (2nd ed.). Plano, TX, USA: Wordware Publishing Inc.

Schawel, C. & Billing, F. (2009). *Top 100 Management Tools—Das wichtigste Buch eines Managers*. Wiesbaden, Germany: Gabler.

Selic, B. (2009). Agile Documentation, Anyone? *IEEE Software*, *26*(6), 11–12.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2001). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference* (2nd ed.). Boston, MA, USA: Houghton Mifflin.

Sicart, M. (2008). Defining Game Mechanics. *Game Studies*, *8*(2), `http://gamestudies.org/0802/articles/sicart`.

Thaler, R. H. (1988). Anomalies: The Ultimatum Game. *Journal of Economic Perspectives (JEP)*, *2*(4), 195–206.

Trochim, W. M. (2006). The Research Methods Knowledge Base. Retrieved 15 February 2012 from: `http://www.socialresearchmethods.net/kb/`.

Valle, C., Prinz, W., & Jarke, M. (2006). Using Tasks and Semistructured Messages to Support Decision Follow Up. In *International Conference on Collaborative Computing (CollaborateCom '06)*, (pp. 1–10)., Atlanta, GA, USA, November 17-20. IEEE Press.

Vroom, V. H. (1964). *Work and Motivation*. Chichester, UK: John Wiley Sons.

Wabba, M. A. & House, R. J. (1974). Expectancy Theory in Work and Motivation: Some Logical and Methodological Issues. *Human Relations*, *27*(2), 121–147.

Wang, Y., Gräther, W., & Prinz, W. (2007). Suitable Notification Intensity: The Dynamic Awareness System. In *Proceedings of the 2007 International ACM Conference on Supporting Group Work (GROUP '07)*, (pp. 99–106)., Sanibel Island, FL, USA, November 04-07. ACM Press.

Wasko, M. M. & Faraj, S. (2005). Why Should I Share? Examining Social Capital and Knowledge Contribution in Electronic Networks of Practice. *Management Information Systems Quarterly (MISQ)*, *29*(1), 35–57.

Wilson, T. B. (2003). *Innovative reward systems for the changing workplace*. New York, NY, USA: McGraw-Hill Professional.

Yetim, F., Wiedenhoefer, T., & Rohde, M. (2011). Designing for Motivation: Focusing on Motivational Values in Two Case Studies. In Anwitaman, D., Stuart, S., Baihua, Z., Shou-De, L., Aixin, S., & Ee-Peng, L. (Eds.), *Proceedings Social Informatics - Third International Conference (SociInfo'11)*, (pp. 255–268)., Singapore, October 6-8. Springer.

Typeset March 26, 2012

Curriculum Vitae

Henri Agaraj was born in Tirana, Albania on the 28th of May 1986. He graduated in 2004 at the secondary school ”Gjergj Kastrioti” in Durrës. He began his study of Computer Engineering in 2004 at the Polytechnic University of Tirana, Faculty of Information Technology.

After he received his B.Sc. degree in Computer Engineering in 2008, he was awarded a scholarship from DAAD and on October 2008 started the master programme “Media Informatics” at the RWTH University. He performed his thesis work at Fraunhofer FIT under the supervision of Christian Prause, graduate computer scientist. His thesis is titled ”Building an Effective Rewarding System based on Reputation”.

His research interests include: social computing, Web 2.0, pervasive games, software engineering, human-computer interaction, interaction design, ubiquitous computing, mobile computing, information systems, and data mining.