MASTER THESIS

Double degree program
Technical University of Berlin and University of Twente
Master of Sc. in Innovation Management and Entrepreneurship
Master of Sc. in Business Administration

PUBLIC, LIMITED VERSION

DEVELOPING A PERFORMANCE MEASUREMENT METHOD IN ORDER TO EVALUATE THE BENEFITS AND COSTS OF WEB APPLICATIONS

FIRST SUPERVISOR UNIVERSITY OF TWENTE:
Dr. Efthymios Constantinides
Assistant Professor Marketing / E-media
School of Management and Governance
University of Twente

SECOND SUPERVISOR UNIVERSITY OF TWENTE:
Dr. Michel L. Ehrenhard
Assistant Professor Entrepreneurial Leadership
School of Management and Governance
University of Twente

ANNIKA GORETZKI
Student number University of Twente: s1359215
Student number Technical University of Berlin: 338456
E-Mail: annika-goretzki@web.de

UNIVERSITY OF TWENTE.

OCT 15, 2013

Executive Summary

This thesis develops a new method to measure the performance of different types of online activities. It addresses the still quite immature, indecisive literature about web measurement and makes valuable contributions to theory and practice by answering research questions.

The thesis exposes that the ubiquity of the internet as well as the development of the information and communication technologies create new opportunities for marketers to address their target groups. In order to display the benefits of the resulting online activities, marketing departments of companies need to quantify the performance of their web applications. However, also the XX GmbH faces the common business problem that benefits, business relevance, efficiency and potential for improvement of used web applications remain uncertain. This is ascribable to a missing continuous monitoring, measurement, analysis and reporting procedure based on determined metrics for web performance measurement.

Even though online monitoring tools offer new ways for marketing measurement, this thesis argues that no standardized, accepted technique for web performance measurement exists and none of the present measurement approaches could fulfill all identified requirements to quantify web applications. Therefore, the thesis develops a method to assess the performance of different types of web applications based on a set of consistent web metrics. Thereby, theoretical knowledge and practical justification are combined iteratively in three research cycles. The data to develop the new web performance measurement tool is gathered from company documents, scientific literature, expert interviews and web monitoring. Based on the gained information, the structure of the new tool, named “web scoreboard”, is designed, the scores, subgoals and metrics are selected and a grading and weighting system is determined.

By using the single case study design with three selected web applications of XX as units of analysis, this thesis also proves the practical applicability of the new method. The web scoreboard indicates the performance of each web application regarding the benefits and the costs during the investigated period. Moreover, the thesis recommends a set of KPIs as well as improvement measures to address the key metrics and introduces a tool for strategic guidance.

Even though some recommended conditions still need to be met in order to guarantee an efficient application of the web scoreboard continuously, the outcomes of the thesis can support a firm’s management to exploit the full underlying potential of its web applications.

Table of content

List of figures

List of tables

Abbreviations

XX GmbH	Company which is made unrecognizable due to the confidentiality agreement
BVDW	Federal association of Digital Economy
BITKOM	Federal Association for Information Technology, Telecommunications and New Media
B2B	Business-to-Business
B2C	Business-to-Consumer
Eurostat	Statistical office of the European Union
EU27	Association of 27 countries of the European Union
FTE	Full-Time Equivalent
KPI	Key Performance Indicator
ROI	Return on Investment
R&D	Research & Development
SEM	Search Engine Marketing
SEO	Search Engine Optimization
SNP	Social Networking Potential
URL	Uniform Resource Locator
WCMS	Web Content Management System
WWW	World Wide Web

Statement of original authorship and confidentiality

Statement of original authorship

I certify that the present master thesis is the result of my own work except where otherwise stated. The use of material from other sources has been properly and fully acknowledged by giving explicit references in the text.

Stuttgart, 15th of October 2013

Annika Goretzki

Confidentiality statement

The present master thesis contains company internal information and confidential data of the XX GmbH. Therefore, the thesis can just be used for evaluation purpose and the contents of this thesis are to remain confidential. Without the explicit permission of the XX GmbH the thesis – neither completely nor in parts – is not allowed to be published or accessible for third parties.

Stuttgart, 15th of October 2013

Annika Goretzki

Acknowledgements and keywords

Acknowledgements

I would like to express my sincere gratitude to my first supervisor, Dr. E. Constantinides, for his professional guidance, useful critiques and enthusiastic encouragement. Despite his busy schedules, he has always made himself available to clarify my questions and I could not have thought of a better supervisor. Besides, a special thanks to Dr. M.L. Ehrenhard, for his good advice and useful comments.

My grateful thanks are also extended to my company supervisors, for their valuable support and constructive recommendations on this project. Finally, a big thank-you goes to the web application managers of XX for their patient guidance and assistance throughout the research process.

Keywords

Performance measurement, online activities, web application measurement, web metrics, method development

1 Introduction – an overview of the thesis

1.1 Online market and web applications as background to the research

Development online market. The ubiquity of the internet, the development of information and communication technologies (Eurostat, 2012) and the accompanied knowledge transparency (Kotler, 2002) are changing the marketing of companies (Hanna et al., 2011; Pires et al., 2006) as well as the way to make business (Wirtz et al., 2010). Those trends foster the customer empowerment, give rise to various Web 2.0 applications (Bernoff & Li, 2008; Constantinides & Fountain, 2008; Hoffman & Fodor, 2010) and drive the expansion of user-generated content (Kozinets et al., 2010; Merz et al., 2009). This results in threats as well as new opportunities for marketers (Hoffman & Novak, 2012; Pires et al., 2006). To face the challenges and benefit from the technological, economic and social drivers (Kaplan & Haenlein, 2010), new media activities are judged by executives as a core element of companies' marketing strategy (Hutter, 2012; McKinsey, 2012; Hanna et al., 2011). Online activities can increase the effectiveness of marketing and might have a powerful impact on the performance (Hanna et al., 2011; Kaplan & Haenlein, 2010; Engagementdb, 2009). Those facts highlight the business relevance of web applications but this increasingly important topic (BITKOM, 2012) still implicates several problematic issues (Kelly, 2013; Järvinen, 2012).

1.2 Web measurement as underlying management problem and research gap

General management problem. In general, to measure the impact of marketing activities on business performance has always been a very demanding task (Farris et al., 2010). Even though online monitoring tools offer new ways for marketing measurement, it is still a big challenge for the marketing departments of companies to quantify the performance and to show the financial outcome of online and especially social media activities (Kelly, 2013; Järvinen, 2012; Etlinger & Li, 2011; Visible, 2011; Hanna et al., 2011; Hoffman & Fodor, 2010; Mangiuc, 2009; Stelzner, 2012). Therefore, the specific effects, the influence on the customers' buying decisions and the return on investment (ROI) of various web applications are uncertain (Divol et al., 2012; Kaplan & Haenlein, 2010). However, a proof of the efficiency and effectiveness of marketing is nowadays a central task for marketers (Meffert et al., 2008), because they need figures justifying the business relevance (Farris et al., 2010) and demonstrating the benefits of their efforts in order to legitimate investments (Sterne, 2010). Thus, companies need to find an appropriate method in order to monitor, measure, analyze and report the performance of their web applications.

Management problem XX GmbH. Those problems are also often discussed topics in the "XX GmbH". Right now, "XX" uses various web applications – as websites, online-communities and web shops – to inform, communicate and make business with its target groups. The problem is that at the moment a continuous monitoring, measurement, analysis as well as reporting procedure of web activities does not take place. Through the lack of consistent measurement metrics and a missing set of defined KPIs, XX faces the common management problem that the benefits, business relevance, efficiency and potential for improvement of the used web applications remain uncertain.

Research gap. Research about web measurement is still quite immature (Murdough, 2009) and no standardized approach for a systematic and comprehensive assessment of the performance and financial impact of web applications is widely accepted (Kelly, 2013; Bremser & Chung, 2006; Divol et al., 2012; Welling & White, 2006a, 2006b). A literature review about web performance measurement shows that research remains indecisive about appropriate tools and indicators to evaluate the effectiveness and efficiency of different types of web applications. Thus, a universally applicable method needs to be designed.

1.3 Research questions answered in this thesis

Research questions. The management problem and research gap lead to the following central question: *"What is an appropriate method to measure the performance of different types of web applications?"*. In order to structure this broad, general question in more specific research issues, it can be divided into five research questions. Those subquestions help to narrow the focus of the thesis and by working on those questions, the answer to the central question is compiled step by step. Table 1 provides an overview of the five research questions. Their answer as well as their contribution to fulfill the research aim is explained in the respective chapter illustrated in Table 1.

Table 1: Research questions and the chapters in which they are explained and answered

Overview research questions	
Research question need to be answered	**Chapter**
1) Which of the existing performance measurement methods is appropriate for web measurement?	5
2) Which metrics are appropriate to measure benefits and costs of different types of web applications?	6
3) How can the identified metrics for web measurement be weighted and graded?	7
4) How is it possible to test the practical applicability of the new method?	8
5) What is an appropriate format for standardized reporting and for strategic guidance?	8

1.4 Research aim – designing a web performance measurement tool

Research aim. Performance measurement of online activities is important as only if the development is monitored, the business relevance can be justified, improvement measures can be derived and the "trial and error mentality" can be lived (Haberich, 2013). As there exists no standardized, widely accepted technique in theory or practice (Kelly, 2013; Bremser & Chung, 2006; Divol et al., 2012; Welling & White, 2006a, 2006b) and none of the established performance measurement methods could fulfill the identified requirements for a web performance measurement, this thesis aims at designing a new performance measurement tool for online activities. The target is to develop a universally applicable method to assess the performance of different types of web applications based on a set of consistent web metrics. With the help of this measurement tool, the management of companies might exploit the full underlying potential of web applications in order to support marketing and sales. In addition, it is aimed at testing the designed method as well as recommending a set of KPIs and introducing a tool for strategic guidance to provide a format for a standardized reporting.

1.5 Research design and methodology to develop the tool "web scoreboard"

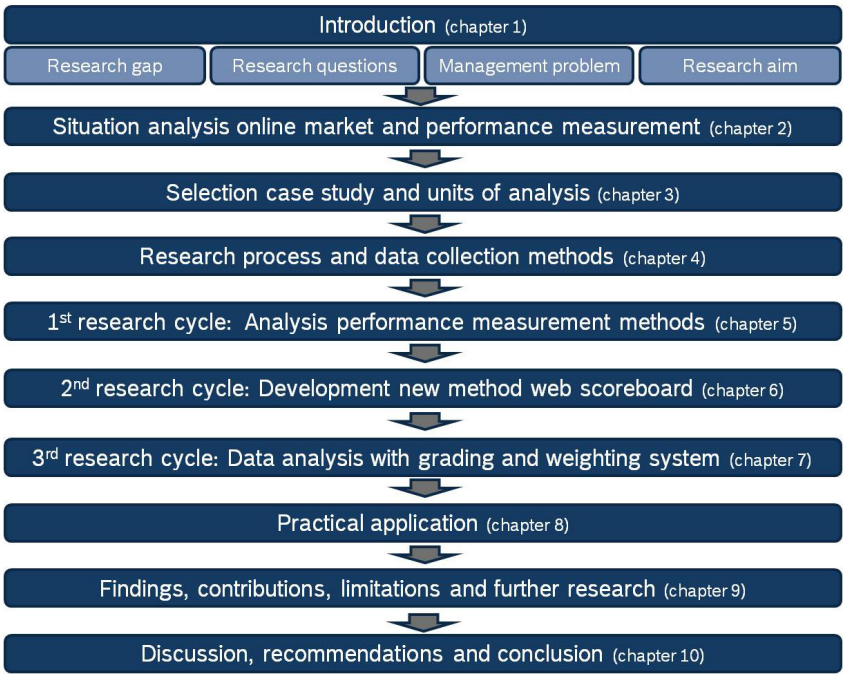
Single case study. The new performance measurement tool, named "web scoreboard", is developed based on the single case study about the XX GmbH. Hereby, three different web applications are used as multiple units of analysis. Those selected applications represent the different types of online activities (i.e. web applications with the major objective of information, engagement or conversion). Based on a single organizational context, the thesis aims at generating knowledge which can be applied to solve a common problem (i.e. web performance measurement).

Research process. The research process of this thesis is based on an iterative combination of theoretical knowledge and practical justification. In order to design a method which solves the identified common business problem, an interactive cycle between building/developing and evaluating/justifying is used. Thus, this thesis consists of three research cycles. In order to develop the new measurement tool with its categorization, scores, subgoals and metrics as well as its weighting and grading system, three different data collection methods are used to gather qualitative and quantitative data from primary and secondary data sources. Those are: reviewing of internal company documents and scientific literature, conducting of expert interviews and monitoring of web data.

1.6 Outline of the thesis

The thesis is composed of ten chapters which are illustrated in Figure 1. The first chapter gives an introduction to the most important issues of the thesis. By analyzing the situation of the online market and highlighting the importance and problems of performance measurement, the second chapter serves as a basis. Subsequently, the selection of the single case XX GmbH with the three web applications as units of analysis, is described. Chapter four gives an overview of the three research cycles which are used in the design process of the new method and explains the data collection methods. Based on identified requirements, existing measurement tools are evaluated in chapter five which verifies the necessity to develop a new method. The following chapter describes the categorization, scores, subgoals and metrics of the new method, labeled "web scoreboard". The process of data analysis with the development of a grading and weighting system is explicated in chapter seven. Subsequently, the practical application of the web scoreboard is tested by implementing it to the three units of analysis. In this context the data collection as well as the achieved scores with their key metrics and respective improvement measures, are described. Further, a set of KPIs with a standardized reporting format is introduced and a tool for strategic recommendations is illustrated. Chapter nine summarizes the findings and contributions of the thesis and comes up with limitations and further research issues. In the end, chapter ten discusses the method of the web scoreboard, recommends conditions for its further application and concludes the thesis.

Figure 1: Structure of the master thesis



1.7 Delimitations of scope, key assumptions and their justifications

Key assumptions. The target groups of XX are B2B and B2C customers. The first key assumption is that the differences of those target groups do not play a role in developing the web scoreboard, hence they can be neglected. Besides, the web applications used to design the new method are addressing different regions. In order to justify this selection, another key assumption is that the addressed users do not have relevant dissimilarities in culture and in their online user behavior (Li & Bernoff, 2008). Further, the web scoreboard is set up in order to provide a measurement method for various types of online activities. Thus, it is supposed that the selected web applications with their different strategic goals represent diverse web activities.

Scope. Even though the tool is developed based on a single case study, its aim is to be applicable also to other web applications, divisions and companies. The data of XX's web applications to test the web scoreboard was collected during one single month (1st until 30th of June 2013).

Writing style. Theoretical characterizations are given occasionally and the single case study about XX exemplifies the design and application of the new tool. The writing style of terms and the illustrations used in this thesis are geared to the way utilized at XX.

1.8 Conclusion of the introduction

As the management problem and research gap illustrate, unacknowledged questions regarding the performance assessment of web applications exist which cannot be answered with established measurement tools. Thus, this thesis aims at designing a new method to measure the performance of different types of web applications based on consistent metrics. Also XX faces the common business problem that benefits, business relevance, efficiency and potential for improvement of used web applications remain uncertain. Thus, three web applications operated by XX serve as units of analysis for the single case study. The data to develop and test the new web performance measurement tool is collected from internal company documents, scientific literature, expert interviews and web monitoring. The theoretical knowledge and practical justification is combined iteratively in three research cycles. Based on the gathered information, the structure of the so-called "web scoreboard" is designed, the scores, subgoals and metrics are selected and the grading and weighting system is determined.

2 Situation analysis deriving the management problem and research gap

2.1 Situation analysis online-market

2.1.1 Development of the medium internet, Web 2.0 and user behavior

Development of the medium internet. The medium internet has become a fixed component in the daily life of many people. In 2012, 72% of the EU27 households had access to the internet via a fast broadband connection (Eurostat, 2012). The faster internet connections and the ubiquity of the World Wide Web (WWW) through devices like Smartphones support the rapid progress of modern information and communication technologies (BVDW, 2013).

Web 2.0, social media and user behavior. In the framework of this progress terms as "Web 2.0"[1] and "Social Media" are used. Web 2.0 describes the way of internet usage that does not focus anymore on technical functions but rather on the interactivity of internet users (Hassler, 2012; Li & Bernoff, 2008). Kaplan and Haenlein (2010) see the Web 2.0 as a platform that is collaboratively modified by users and based on which social media evolved. Social media is defined by Hoffman and Novak (2012) "as Web-based applications that permit creation, sharing, manipulation and consumption of user generated content" (p. 202). The power of social media has its origin in the external desire of humans, as social entities, to connect with each other and thereby create content (Blanchard, 2012; Li & Bernoff, 2008; Qualman, 2011). User generated content means "various forms of media content that are publicly available and created by end-users" (Kaplan & Haenlein, 2010, p. 61). Customers are no longer passive participants but actively engaging and contributing (Kozinets et al., 2010; Merz et al., 2009) as well as influencing with their user-generated content other consumer's behavior (Kotler, 2002; Riegner, 2007). This can be done because people trust more in the evaluations of their peers (Qualman, 2011; Karpinski, 2005). Thus, recommendations of other users – often called Word-of-Mouth – are more effective than traditional marketing communication (Kozinets et al., 2010) since they are perceived by customers as a trustworthy and reliable source of information (Razorfish, 2008). This entails that the consumer now has the control and power over company brands (Qualman, 2011; Bernoff & Li, 2008; Constantinides & Fountain, 2008; Hoffman & Fodor, 2010). The internet has changed the buying and decision process of customers (TNS, 2011). Information about products is gathered through websites or published customer experiences in online-communities. After the buying decision is made, users can nowadays even purchase a lot of products online (Huizingh, 2002).

[1] The term "Web 2.0" was introduced by Dale Dougherty, a co-founder of the firm O'Reilly Media, Inc., in October 2004 to express that the internet was at a turning point (O'Reilly, 2005).

2.1.2 Impact of the internet developments on companies and managers

Impact of the changing technologies on marketing. Due to those rapid developments of the modern communication and information technologies, the transparency of markets and the change in user behavior, firms need to adjust their activities to stay competitive in the long run (Nielsen Company, 2009). The new technologies are and will be also very important for the role of marketers in the firm and the operation of marketing concepts (Tapp & Hughes, 2004). This is proven by the fact that marketing has changed less in the past 100 years altogether than during the last five years (Turner & Shah, 2011). Besides, web applications increase constantly in importance in the B2C as well as in the B2B area (BVDW, 2012a).

Advantages of web applications[2]**.** Web applications provide multiple advantages. First of all, they can generate consumer insights faster and at lower costs than traditional methods, make co-creation possible, develop new business models and bring the firm closer to the market (Gillin & Schwartzman, 2011; Stelzner, 2012). Furthermore, especially social media activities can help businesses to develop beneficial trust-based relationships to customers (Bulut & Mandaric, 2012; Constantinides & Zinck Stagno, 2012) and bring new clients to the firm (Kelly, 2013). In addition, key advantages of the interactive social media tools are the possibility to listen, interact and collaborate (Bulut & Mandaric, 2012) as well as to influence each stage in the buying decision process of the consumer (Divol et al., 2012). Overall, the internet can provide benefits regarding costs and time as well as competitive advantage by creating long-term customer relationships and loyalty of the market (Li & Bernoff, 2008; Nambisan, 2002). Even though the WWW has changed the way of interaction and offers various possibilities, the company needs to set the framework (Hoffman & Fodor, 2010) and need to listen and react (Qualman, 2011) in order to participate successfully in online activities (Hoffman & Novak, 2012; Kaplan & Haenlein, 2010). Finally, managers need to keep in mind that trust is crucial for the success of online applications (Sterne, 2010).

2.2 Situation analysis of the on- and offline marketing performance measurement

Performance measurement. According to Neely et al. (1995) performance measurement is defined as "the process of quantifying the efficiency and effectiveness of action" (p. 80). In general, performance measurement is a broad topic which crosses functional boundaries (Neely, 1999) and the literature about this important research field is diverse (Marr &

[2] This thesis uses the expression "web applications" in order to talk about activities which are undertaken in the WWW – regardless if those online activities are informative, interactive or commercial.

Schiuma, 2003; Neely et al., 1995). Having its origin in the beginning of the 20th century, the research field of performance measurement had a revolution in the mid 1990s (Neely, 1999).

Problems and development of marketing performance measurement. For all activities in the company, managers need to prove an enhancement in the value for the firm. As the budget is allocated to the activities that reach the determined goals in the most efficient way, marketers need to show their ROI as well. Apart from the general problem in marketing, namely to attribute success to single activities (Meffert et al., 2008), the problem is that ROI calculations do not work for on- and offline marketing activities, like social media (Gillin & Schwartzman, 2011; Kelly, 2013). This is the case because the results of those efforts are primarily no quantifiable, intangible assets – like market awareness, customer engagement, continuous feedback and product development (Farris et al., 2010; Gillin & Schwartzman, 2011; Welling & White, 2006a). Those long-term, nonmonetary benefits are not quantifiable but very valuable for a firm (Meier & Zumstein, 2013) as they might be the step before a customer generates financial outcome (Blanchard, 2012; Blanchard, 2009) and serve as a basis for long-term financial success and commitment (Kaplan & Norton, 2004). Performance measurement has been developed in order to display the effects of immaterial, intangible assets and make them measurable. Thereby, the business relevance of specific activities (Artz, 2010) can be justified. In the mid 1990s managers started to focus more on marketing performance management (Meffert et al., 2008; Neely, 1999). Nowadays, it is a widely discussed research field (Artz, 2010) and of keen interest in the e-business literature (Bremser & Chung, 2006).

Problems of marketing measurement in the online surrounding. The online environment provides novel possibilities to measure the effects of marketing activities (Gillin & Schwartzman, 2011; Kelly, 2013). Here the customer behavior is traceable and visible (Hennig-Thurau et al., 2010; Welling & White, 2006a) and real-time feedback on the effectiveness of marketing efforts is provided (Farris et al., 2010; Hassler, 2012). Nevertheless, no standardized, widely accepted approach for a systematic and comprehensive measurement of the performance and financial impact of web applications exists in theory and practice (Bremser & Chung, 2006; Divol et al., 2012; Welling & White, 2006a, 2006b).

Reasons for performance measurement of online activities. Web performance measurement continuously increases in importance. In scientific literature six main reasons

for this phenomena have be identified. Firstly, with the help of online technologies as web analytics and social media monitoring (see chapter 8.2), firms can document habits and preferences of customers (Phippen et al., 2004). Secondly, companies can gently influence the interactions and react to negative mentions, if they monitor what is going on in their web applications (Hassler, 2012; Sterne, 2010). Thirdly, the transformation of intangible assets into tangible, quantifiable figures in a structured way helps firms to focus their resources on those applications which have the greatest chance to acquire a return and to guarantee a long-term success (Gillin & Schwartzman, 2011; Meffert et al., 2008). Fourthly, by having a periodical data analysis, deviation in the monitored figures can be detected in order to identify measures for improvement (Meier & Zumstein, 2013) and to guide decision making (Haberich, 2013). Fifthly, as the efforts for developing, implementing and maintaining web applications can be very high, measurement can give managers the confidence they need to legitimate financial and human investments in online activities as marketing tools (Divol et al., 2012). Lastly, as the achievement of the determined goals can be monitored, the success of the improvement measures can be pursued afterwards (Meier & Zumstein, 2013).

2.3 Conclusion of the situation analysis

This chapter analyzed the situation in the online market and highlighted the importance for companies to apply web applications. Supported by the developments of the modern technologies, online activities offer a lot of benefits and are an indispensable tool for firms to address their target groups. In order to display the benefits of the online activities, marketing departments of companies need to quantify the performance of their web applications. Overall, the situation analysis highlighted the necessity to measure the performance of web applications in order to justify its business relevance. Apart from the importance of web measurement it also showed the occurring problems in performance measurement of online and offline marketing activities and thereby derives the general management problem. Simultaneously, the situation analysis disclosed the research gap of a missing standardized, widely accepted approach to measure the performance of web applications in a systematic and comprehensive way. Thus, this thesis aims to address this gap by developing a universally applicable method for web performance measurement. In the next chapter the XX GmbH will be explained because it faces the common business problem of web performance measurement. Therefore, this company case can exemplify the design of the new measurement method as a single case study.

3 Case study XX GmbH

3.1 Facts and figures about the XX GmbH

Excluded due to the confidentiality agreement

3.2 Management problem and selection of XX as company case study

This thesis aims at developing a new measurement method based on a single case study about the XX GmbH. A case study approach helps to get an overall understanding of the complexity of the underlying problem (Yin, 2009; Benbasat et al., 1987). Furthermore, case studies are used to study a problem in-depth and in a real-life context (Boyer & Swink, 2008; McCutcheon & Meredith, 1993; Meredith, 1998; Wacker, 1998). The case study is based on an embedded view to validate the results and to support the development of the new measurement method.

This company is sampled because it faces a common, important business problem: Measuring the performance of web applications and thereby quantifying the benefits and costs of online activities as well as justifying their business relevance (Kelly, 2013; Etlinger & Li, 2011; Visible, 2011; Hanna et al., 2011; Hoffman & Fodor, 2010; Mangiuc, 2009; Stelzner, 2012). More precisely, it approaches the novel opportunities offered by the internet ubiquity and developments of modern technologies these days so that it applies various web applications to reach its different target groups. Even though it is crucial to monitor if the return justifies the efforts, so far its online activities are not measured in a uniform way. This missing web measurement is a problem faced by a majority of firms nowadays. Based on a baseline study, supported by eleven expert interviews with the web application managers of XX, three web applications are selected as representative units of analysis to develop a new performance measurement. The interview questions and a list of the interview partners including their positions can be found in Appendix 3 and Appendix 4. The three web applications and their choice will be described in the following.

3.3 Web applications of XX and selection of the units of analysis

Strategic goals and different types of XX's online activities. XX exemplifies that the various types of online activities can be clustered according to their main objective in three different fields, namely:

1) Information; 2) Engagement; 3) Conversion

Meaning, XX categorizes all its online activities to the strategic goal which is most likely to be fulfilled by the respective web application (see Appendix 5). To guarantee that the newly developed method comprises various types of web applications so that it can be used to measure the performance of different online activities in the end, appropriate web applications are chosen which represent the three strategic objectives of online activities. Therefore, three representative applications are selected which are expected to be strong at fulfilling three different strategic goals. Firstly, the Website (A) which substitutes online activities aiming mainly at information. Secondly, the Online-community (B) which is expected to be strong in engagement and thereby represents web applications aiming at engagement. Thirdly, the Web Shop (C) which is expected to perform well in conversion.

A: Website. On this Website the potential customers can inform themselves about XX's products, services and the latest news. Therefore, the major strategic objective of "A" is information. This is important as if the customers know the brand and its benefits, a long-term pull effect can be created (Kotler & Pfoertsch, 2010). The relaunched version of this relatively established web application is online since spring 2013. In Appendix 6, the Homepage of the Website is pictured. So far no monitoring took place but it is urgently necessary in order to pursue the relaunch effects.

B: Online-community. An online-community is a social network which uses web tools for communication and collaboration of its members (Meier & Zumstein, 2013). "B" is such an online-community and thereby a social media activity which fulfils the basic idea of the Web 2.0: users are generating, sharing, collaborating and discussing content (Hassler, 2012; Kotler, 2002; Constantinides & Fountain, 2008; Hanna et al., 2011). The addressed B2B target group, can exchange information and discuss about problems in a forum, can raise ideas for product or process development, can join specific groups and ask questions. Therefore, the major strategic objective of B is engagement. Communities are judged by B2B marketers as a very effective social platform and are lately increasing in their importance, especially in B2B areas for product development. This is ascribable to the fact that B2B relationships are about informing and helping each other with expertise and advices (Gillin & Schwartzman, 2011). Since spring 2012, B is publicly available for every internet user. Appendix 7 shows a picture of B's landing page. The members get a weekly newsletter about the latest activities. Further, they are rewarded for their contribution with an activity sign. Thereby, key opinion leader can be detected and thereof a panel of key opinion leaders can be built up, which can foster

customers' loyalty. Besides, members can evaluate posted ideas of others with a "like" and "dislike" button.

C: Web Shop. A web shop is defined as a web-based software system which offers products, receive orders and coordinates payment and delivering modalities (Meier & Zumstein, 2013). The potential customers can order products directly via the online Web Shop. Therefore, the major strategic objective of "C" is conversion. In Appendix 8, a picture of the Web Shop's landing page is displayed.

3.4 Conclusion of the case study selection

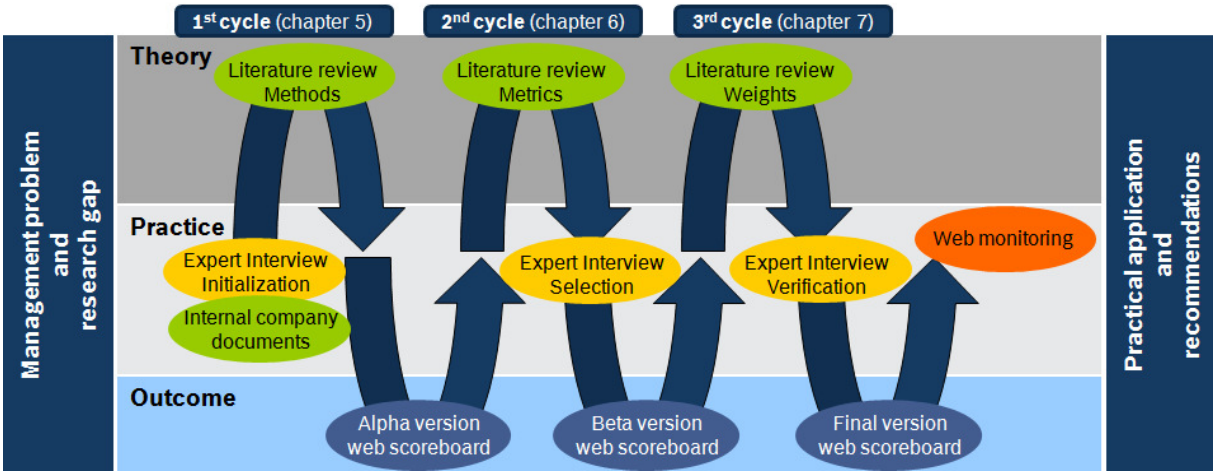
Chapter three analyzed the situation of the XX GmbH. The management of XX faces the common, important business problem that benefits, efficiency, business relevance and potential for improvement of its used web applications remain uncertain. This is ascribable to a missing measurement tool and procedure. Three of XX's web applications are selected to represent different types of online activities. By using those applications as units of analysis, it is aimed at developing an appropriate method to measure the performance of different types of online activities. Thereby, XX exemplifies the development and the initial design of a new measurement method. The following chapter provides assurance that eligible procedures are executed throughout the development of the new measurement tool.

4 Methodology leading to the new web measurement method

4.1 Introduction to the methodology and overview of the research process

Research process. The research process of this thesis is based on an iterative combination of theoretical knowledge and practical justification. This approach is used in the design-theory research (e.g. Cole et al. (2005), Hevner et al. (2004), Pries-Heje and Baskerville (2008), Sein et al. (2011)). The goal of this research field is to provide solutions to important business problems (Hevner et al., 2004; Pries-Heje & Baskerville, 2008). Action Design Research goes one step further and aims at generating knowledge which can be applied to solve a common problem, based on a single organizational context (Sein et al., 2011). In this thesis, XX exemplifies the common problem of web performance measurement. In order to design a method which solves the identified problem, an interactive cycle between building/developing and evaluating/justifying is necessary (Hevner, 2007; Pries-Heje & Baskerville, 2008). Sein et al. (2011) provide guidance for this process of building, intervention and evaluation. Based on this suggested process, the research cycles of the thesis are established. An overview of the thesis's research process is provided in Figure 2 and a timeline is displayed in Appendix 9.

Figure 2: Overview of the iterative research cycle process applied in this thesis



Source: Own illustration based on Pries-Heje and Baskerville (2008) and Sein et al. (2011).

4.2 Data collection methods

Data collection methods. Different data collection methods are combined to gather adequate data to develop and test the new method. The methodological triangulation to collect qualitative and quantitative data of primary and secondary data sources is typical for case studies and strengthens construct validity (Yin, 2009; Eisenhardt, 1989; Meredith, 1998).

Here, it consists of: (1) reviewing company documents and scientific literature, (2) conducting expert interviews, (3) monitoring web data (see green, yellow and orange fields in Figure 2).

Expert interviews. As interview partners, the web application managers of XX have been selected. The choice and the regional scope were recommended by the company supervisors. The interviews were conducted informally, i.e. an outlook invitation for a one hour meeting was sent, including the relevant interview questions. During the meeting the interviewee answered the questions orally and the collected answers were evaluated afterwards. Overall, the three rounds of interviews were conducted in the period of 11th of April until 22nd of July 2013. The respective questions are provided in Appendix 4, Appendix 11 and Appendix 14. In the first interviews, the aim has been firstly, to gain insights into XX's web applications in order to identify appropriate representatives for different types of online activities and secondly, to identify requirements for a web performance measurement tool (see chapter 5.2). In the second round, especially the preselection of the subgoals and metrics has been adapted and verified (see chapter 6.3.1). In the third round, particularly the weighting system has been justified (see chapter 7). Moreover, all issues were discussed and justified by the company supervisors in regular feedback and status-update meetings.

Literature review. A systematic literature review has been conducted to address three different positions in the research cycles. Details about the literature reviews which served as a theoretical foundation to develop the web scoreboard, will be explained in the respective chapters (for the first literature review see chapter 5.3, for the second see chapter 6.3.1 and for the third see chapter 7). The academic search service "Web of Knowledge" and the bibliographic database "Scopus" as well as "Google Scholar" were used for the literature reviews. The following keywords helped to identify relevant literature for the first, second and third research cycle: performance measurement, performance measurement systems, marketing measurement, performance measurement marketing, web measurement, online measurement, online marketing measurement, internet marketing measurement, Web 2.0, new media, social media, value of social media, effectiveness of social media, web analytics, web metrics, online performance metrics, weighting systems, grading systems, factor weights, evaluation systems, grading models, scoring models. Besides, in the references of the found literature it was also searched for additional relevant publications. The software tool "EndNote" was applied to manage the literature bibliography and the citations.

Web monitoring. This data collection method is described in detail in chapter 8.2.

4.3 Execution of the research cycles

First cycle. In the initial phase, insights resulting from expert interviews as well as generated in a web analytics workshop, were used to identify requirements for a web performance measurement tool. The interviews with eleven web application managers of XX as well as the workshop are further described in chapter 5.2. Besides, to guarantee that the web scoreboard can be used for different web applications in the end, eligible applications needed to be selected which represent various types of online activities. Thus, apart from internal company documents also the first round of expert interviews was used to gain insights about XX's web applications. The results of this baseline study are explained in chapter 3.3. In the next step, based on the identified requirements, a literature review was conducted in order to evaluate if an existing measurement method can be used to measure web performance (see chapter 5.3). The necessity to develop a new method was verified by the justification of the research gap and because none of the identified existing measurement approaches could fulfill all requirements. Thus, a first alpha version of the web scoreboard was designed.

Second cycle. After the initial structure of the web scoreboard has been set in its first version, an extensive literature review was conducted in order to find appropriate subgoals and metrics as well as to justify the categories and scores of the external and internal perspective. The identified metrics and subgoals have been preselected based on various criteria, resulting in a list of 55 metrics. In the second round of expert interviews, the preselection was evaluated and adapted in order to generate the final list of 38 metrics, displayed in the beta version of the web scoreboard. This second research cycle is described in chapter 6.3.1.

Third cycle. In the next step, existing literature was scanned in order to find adequate approaches to determine appropriate weights and grading intervals. Thereby, an eligible way to interpret the collected data of the web scoreboard was detected. After a weighting and grading system has been designed, it was verified by the company supervisors and web application managers in the third round of the expert interviews. This third research cycle which resulted in the final version of the web scoreboard is explained in chapter 7.

Practical application. Finally, in the last step, the web scoreboard was filled with data of the selected web applications in order to test its practical application. Therefore, in the inquiry period (1st until 30th of June 2013), monitoring was undertaken to collect data with the help of three tools. This is explained further in chapter 8.2.

5 Analysis of performance measurement methods – 1st research cycle

5.1 Introduction to performance measurement

Based on the identified management problem, chapter five explains the determination of requirements for a web performance measurement tool by conducting expert interviews and gaining insights from a web analytics workshop. According to those requirements, existing business and marketing performance measurement methods in the off- and online surrounding are investigated. Hereby, the aim is to answer research question number one (see Table 2).

Table 2: Research question which is answered in chapter five

Research question need to be answered
1) Which of the existing performance measurement methods is appropriate for web measurement?

By scanning the existing measurement methods in the literature, three of them – balanced scorecard, marketing dashboard and set of metrics – are detected as the most probable ones to fulfill the previously identified twelve conditions. Thus, they are analyzed in the following in detail. Furthermore, in the conducted literature review the existing measurement techniques are scanned in order to find appropriate elements as a basis for the new tool. Moreover, in this chapter a standardized measurement procedure is ascertained and the "General Electric Matrix" is described in order to provide a holistic theoretical foundation. Overall, answering research question number one contributes to the research aim by identifying what would be an optimal performance measurement method for online activities.

5.2 Requirements for a web performance measurement method

Identification of requirements. Based on the management problem, specific requirements for a web performance measurement method have been identified. Therefore, expert interviews with eleven web application managers of XX were conducted. The interview questions are provided in Appendix 4. Besides, a web analytics workshop with an external agency helped to gain insights into the required measurement tool. Apart from this, the company supervisors were asked about the conditions which a web measurement tool needs to meet. By asking experts in the initial step of the research process, the business problem and need was concretized in specific requirements for a web measurement method.

Requirements. In total, twelve requirements for a web performance measurement method have been identified. First of all, the tool needs to consist of understandable, standardized, financial and non-financial, quantitative and qualitative figures. Those precise, measureable,

differently weighted web metrics should enable a monitoring of the determined subgoals. Those subgoals need to be allocated to multiple determined scores. The defined scores should be aligned with the strategic business goals of online activities and represent the external (benefits) as well as the internal (costs) perspective. An appropriate weighting and grading system should help to evaluate and interpret the collected data in an eligible way. Everything needs to be displayed in one concise document. Besides, a standardized measurement process is necessary which can be continuously applied. In addition, the tool should provide the possibility to derive strategic recommendations.

Table 3: Requirements identified in the expert interviews and the web analytics workshop

Requirements for a web performance measurement tool
One concise document
External (benefits) and internal (costs) perspective
Alignment with strategic goals
Multiple scores
Subgoals
Measureable metrics
Financial and non-financial figures
Quantitative and qualitative figures
Weighting system according to strategic goals/relevance
Grading system to interpret collected data
Standardized, continuous measurement procedure
Illustration for strategic guidance

5.3 Existing performance measurement methods – advantages and limitations

Literature review about existing methods. Based on the direction which was given by the identified requirements, a literature review about existing business and marketing performance measurement systems in the off- and online surrounding was conducted in the next step. This first literature review aimed at different things. Firstly, the identified research gap of a lacking web performance measurement method needed to be justified. Secondly, it was aimed at checking if one of the existing performance measurement systems could fulfill the previous identified requirements. Thirdly, by reviewing existing techniques, basic elements for a new performance measurement tool should be detected.

Selection of examined methods. As the identified research gap already indicated, the amount of scientific measurement tools in the online surrounding was limited. The identified existing performance measurement systems published in the time frame from 1992 – 2013 have been preselected based on the predefined, specific requirements. This resulted in a list of 14 methods which is provided in Appendix 10. By having a closer look to the preselected

methods, three of them – balanced scorecard, marketing dashboard and set of metrics – seemed to be possible candidates to fulfill the requirements. Apart from a description of the methods, Appendix 10 also states briefly the reasons why the other eleven techniques are not further examined. Besides, by analyzing the methods, several elements which could be used as a basis for the design of a new tool have been detected. The methods might provide valuable input for the required comprehensive structure and performance indicators of a new web measurement method. Thus, the top three candidates with their reasons for selection, their advantages and their possibility to fulfill the conditions are examined in the following.

Balanced scorecard. One of the most influential performance measurement systems is the "balanced scorecard" (Evans, 2004; Marr & Schiuma, 2003).[3] It is developed by Kaplan and Norton (1992) and can be transformed to the e-business environment (Van Grembergen & Amelinckx, 2002). Thus, it has been selected for further examination as a possible framework for a web measurement tool. It considers quantitative figures and qualitative variables to guarantee a holistic judgment and balanced perspective of a firm's performance. It is a modern, comprehensive, capable system which can be adapted flexibly to the specific company requirements (Meffert et al., 2008) and has received significant attention in marketing (Homburg et al., 2012). This approach addresses the problem that the complexity of a firm's performance cannot be captured in one single performance indicator. Therefore, the system uses a classification into four categories (financial, customer, internal processes, learning and growth (Kaplan & Norton, 1992)) which consist of multiple criteria (Epstein & Manzoni, 1998). A limitation of the balanced scorecard is that it is internally focused and does not take competitors and customers into account. Besides, it is a snapshot of the current situation (Pauwels et al., 2009). Due to those characteristics, it fulfills six out of twelve requirements (see Table 4). Thus, the system is not sufficient as a single web measurement tool but provides valuable basic elements for a new method.

Dashboards. The recently introduced "marketing dashboard" can be defined as a performance management tool that provides managers a consistent overview of the firm's key marketing metrics to reach organizational goals (Farris et al., 2010; Pauwels et al., 2009; Velcu-Laitinen & Yigitbasioglu, 2012). Dashboards allow "the user to identify, explore, and communicate problem areas that need corrective action" (Yigitbasioglu & Velcu, 2012, p. 4) and thereby support marketers in monitoring, planning, assessing and reporting performance.

[3] In chapter 4.4.2.2 it will be explained how XX uses the balanced scorecard as a basis for its strategy map. But first the underlying principles of the balanced scorecard are of interest for the design the new method.

Besides, dashboards help to handle the increasing complexity of market data (Pauwels et al., 2009) and to solve problems (Velcu-Laitinen & Yigitbasioglu, 2012). Due to those reasons it has been selected to be examined in detail. In comparison to the balanced scorecard, a dashboard builds a bridge between the internal and external perspective by relating the benefits and the costs at one display (Pauwels et al., 2009). To consider both perspectives is important for managers (Bremser & Chung, 2006; Neely et al., 1995). This can be seen as a valuable input for the new method which also should be separated into those two perspectives. Another benefit of a dashboard is that it provides various managers of different departments and locations the chance to measure equally and share the data (Pauwels et al., 2009). Overall, this method fulfills eight out of the twelve required conditions. Thus, it cannot be used as a single web measurement method. However, due to the various advantages and the fact that a common recommendation for managers is to create a dashboard (Farris et al., 2010), dashboards provide precious input for the basic structure of a newly designed method.

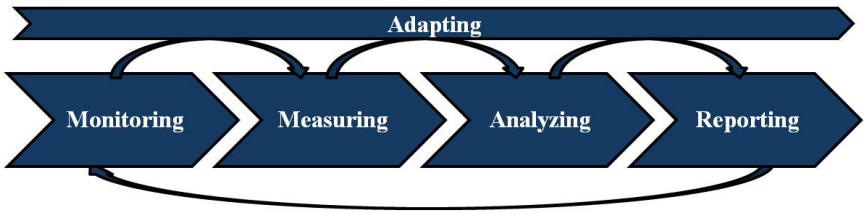
Metrics. "Metric" is a term that comes from the Greek language and means "measurement" (Meier & Zumstein, 2013). In general, a metric is a measuring system with standardized figures which facilitates to compare observations (Farris et al., 2010). It is "used to implement strategy, manage operations and track performance over time" (Bremser & Chung, 2006, p. 396). Metrics in the online environment are called "web metrics". The collection, measurement, analysis and interpretation of web metrics is called "web analytics" and is recently a booming business (Meier & Zumstein, 2013; Weischedel & Huizingh, 2006). Thus, this technique was chosen for further examination. Web metrics enable to measure the performance of web applications (Meier & Zumstein, 2013), to show the user behavior, to uncover potential for improvement (Weischedel & Huizingh, 2006), to compare (Straub et al., 2002) and to audit activities (Phippen et al., 2004). Due to those advantages a set of metrics needs to be determined for a newly designed measurement tool. The measured metrics need to be precise, understandable (Kelly, 2013) and aligned with the strategic goals (MarketingSherpa, 2010) to capture the interest of the managers and executives (Blanchard, 2012). Further, metrics should be weighted according to their relevance (Hassler, 2012). Overall, this technique fulfills seven out of the twelve required conditions. However, the elements and advantages of a set of metrics should be considered in the design process of a new method.

BCG and General electric matrix. The Boston Consulting Group (BCG) matrix has been developed in the early 1970s and is "a common approach to managing a portfolio of different

strategic business units (SBUs) or major product lines" (Zhou & Zuo, 2010, p. 120). An enhanced version of this four box "BCG matrix" is the "General Electric Matrix", also called McKinsey-Portfolio. It consists of nine blocks and has been developed also in the 1970s by the consulting company McKinsey and General Electric. This technique is used in strategy or product management in order to help firms with their product portfolio decisions. Its original version considers industry attractiveness and business strengths but as a variable and multifunctional model it can be configured in various ways (McKinsey, 2008). The general idea behind it is to classify the investigated product or services – as the web applications – according to their performance into nine different fields. Based on this classification three different strategic decisions can be recommended: "Invest", "Select", "Divest". In this thesis the underlying principle of the General Electric Matrix serves as a theoretical foundation and an adjusted version is used to interpret the results of the web performance measurement method (see chapter 8.5).

Measurement procedure. A standardization of the web measurement process is crucial to be successful in new media (Novak & Hoffman, 1997; Pauwels et al., 2009). Literature identified different issues that need to be undertaken in web measurement. Thus, in the development of a new method it needs to be considered that it enables managers to make five steps (see Figure 3). Firstly, to monitor, meaning to listen goal-oriented, collect and sort measureable data. Secondly, to measure, meaning to transform the collected data into a precise number (Blanchard, 2012). Thirdly, to analyze, meaning to interpret the figures in order to evaluate success and failure and determine the room for improvement (Pauwels et al., 2009). Fourthly, to report, meaning to indicate the results to managers on a periodical basis, optimally with the help of predefined key metrics or KPIs (Phippen et al., 2004). Fifthly, to adapt, meaning to constantly implement the derived improvement measures (Pauwels et al., 2009).

Figure 3: Five steps of a standardized measurement process



Source: Own illustration based on Blanchard (2012); Meier and Zumstein (2013); Pauwels et al. (2009).

5.4 Conclusion – justification to develop a new measurement method

Justification of new method. In general, a variety of methods is used to assess performance and various performance measurement systems are proposed in the marketing literature. Each of those techniques has its advantages and drawbacks. Three measurement methods – balanced scorecard, marketing dashboard and set of metrics – have been identified in the literature review as being the methods which are most likely to fulfill the previous detected requirements. They have in common that they contain a set of financial and non-financial figures (Artz, 2010) and summarize the performance of the investigated object in a single, concise document (Epstein & Manzoni, 1998). Nevertheless, none of the existing performance measurement methods reviewed could fulfill all identified requirements (see Table 4). Hereby, research question number one is answered: none of the existing measurement tools is appropriate to measure the performance of different types of web applications. Thus, under consideration of the required issues a new measurement method, labeled "web scoreboard"[4] is developed in an exploratory process. This will be explained in detail in the following chapter six.

Table 4: Checklist requirement fulfillment of the top three examined measurement tools

Checklist requirements for web measurement tool	Included in…			
Requirements for a web performance measurement tool	scorecard	...dash-	metrics	scoreboard
One concise document	✓	✓	✓	✓
External (benefits) and internal (costs) perspective	✗	✓	✗	✓
Alignment with strategic goals	✓	✓	✓	✓
Multiple scores	✓	✗	✗	✓
Subgoals	✗	✗	✗	✓
Measureable metrics	✗	✓	✓	✓
Financial and non-financial figures	✓	✓	✓	✓
Quantitative and qualitative figures	✓	✓	✓	✓
Weighting system according to strategic goals/relevance	✗	✗	✓	✓
Grading system to interpret collected data	✗	✗	✗	✓
Standardized, continuous measurement procedure	✗	✓	✓	✓
Illustration for strategic guidance	✓	✓	✗	✓
Appropriate web performance measurement tool	✗	✗	✗	✓

[4] The first idea of a web scoreboard was developed by Dr. Efthymios Constantinides (University of Twente) in 2000. Based on the theoretical foundation of the General Electric Matrix the original aim was to develop a performance measurement tool for different types of websites.

6 Development of a new web performance measurement method – 2nd research cycle

6.1 Introduction to the second research cycle

Chapter five resulted in the finding that a new performance measurement method needs to be developed; this chapter introduces the initial design of the new tool and thereby represents the second research cycle. The determination of the structure and scores used in the new method, labeled "web scoreboard" is explained. Besides, the selection and characteristics of suitable subgoals and metrics are described in detail. Hereby, research question number two (see Table 5) is answered which helps to fulfill the research aim by providing – with an appropriate set of metrics to measure the benefits and costs of different types of web applications – a main part of the newly designed method.

Table 5: Research question which is answered in chapter six

Research question need to be answered
2) Which metrics are appropriate to measure benefits and costs of different types of web applications?

6.2 Structure and categorization of the web scoreboard

Structure. According to the identified requirements, the new measurement tool, i.e. the web scoreboard, consists of an external perspective illustrating the return and benefits of the web applications and an internal perspective illustrating the investments and costs of the applications. In total seven scores, 36 subgoals and 38 metrics are chosen to measure the performance of online activities. The web scoreboard is set up in a concise Excel file.

External perspective. Defining scores according to the strategic goals of a web application is seen as a reasonable way to structure performance measurement (Evans & King, 1999; Welling & White, 2006a). The scores of the web scoreboard's external perspective are determined pursuant to the strategic aims of the different types of online activities because the verification of the business relevance can best be done by measuring the contribution to the determined business goals of each web application (Johnson & Misic, 1999). Thus, the resulting scores can be taken to check the goal achievement and thereby to justify the business relevance. Besides, a fourth score is included to address comprehensive objectives:

1) Information; 2) Engagement; 3) Conversion; 4) General benefits

Internal perspective. Additionally to the benefits also information regarding the costs is very important (Mangiuc, 2009). Therefore, the internal perspective of the web scoreboard is classified according to the expenses incurred with applying an online activity.

1) Investments; 2) Maintenance; 3) General costs

Those categories are defined as scores. Based on this, recommendations might be given how much efforts are needed to apply an application not just in the initial phase but also over time.

Scores. Scores are judged in literature as a good way to guarantee an overview of the results and to ease the measurement process, especially for management reporting. Further, it is recommended that the scores are broken down in subgoals and measurable metrics which are weighted differently (Hassler, 2012). Thus, also the determined scores of the web scoreboard are broken down in subgoals. To enable the monitoring of those subgoals appropriate, measureable metrics are selected which are represented by financial and non-financial, quantitative and qualitative figures. The selection is explained in the following.

6.3 Subgoals and metrics of the web scoreboard

6.3.1 Selection of metrics based on a literature review and expert interviews

Selection process. The ideal conception of web measurement is that each used web application can be measured in one uniform tool based on the same criteria (Hassler, 2012). To create such a measurement tool which is applicable for different types of web applications with different target groups and different business models, a careful selection of metrics is important. Since no single set of metrics – valid for every company – exists so far (Etlinger & Li, 2011), a review of the existing publications as journal articles, books, electronic articles, reports and web pages helped to select relevant metrics to measure the performance of web applications. Table 6 gives an overview of this literature. In the literature review a lot of metrics to assess the performance of different types of web applications have been identified. Those metrics have been scanned and preselected. To decide if a metric could be appropriate for the web scoreboard, several elements came into consideration. Firstly, the metric has to be SMART, i.e. specific, measurable, achievable, realistic and time-bound (Hassler, 2012); secondly, the measurement of the metric has to provide additional value (Etlinger & Li, 2011); thirdly, it has to suit to the different types of web applications. In the next step, this preselection of 55 metrics was discussed with the experts in order to choose the final list of metrics for the web scoreboard (see Appendix 11). The metrics which have not been considered as valuable by at least two out of the three experts have been excluded. Hereby,

another 17 metrics (see Appendix 12) have been removed. This selection process which combines theory and practice, represents the second research cycle of this thesis and its outcome is the beta version of the web scoreboard.

Table 6: Overview of the literature used to identify a set of relevant metrics

Overview literature metrics			
Authors	**Year**	**Topic**	**Type of source**
Bremser & Chung	2006	Provide a framework for developing performance measurement metrics for e-business	Journal Article
BVDW	2013	Provides metrics to measure social media performance	Report
Dickinger & Stangl	2011	Suggest a formative approach with eight dimensions to measure website performance	Journal Article
Divol et al.	2012	Provide metrics to measure social media performance	Journal Article
Evans & King	1999	Provides metrics to measure website performance	Journal Article
Evanschitzky et al.	2004	Provide metrics to measure e-business performance	Journal Article
Farris et al.	2010	Provide metrics to measure marketing performance	Book
Gillin & Schwartzman	2011	Provide metrics to measure social media performance and how they can be measured	Book
Happe	2008	Provides metrics to measure social media performance	Webpage
Hassler	2012	Provides metrics to measure websites and classifies them in four dimensions	Book
Hoffman & Fodor	2010	Provide metrics (based on the steps of the brand funnel) to measure social media ROI	Journal Article
Huizingh	2002	Provides metrics to measure website performance	Journal Article
Hung & McQuenn	2004	Provide metrics to measure e-business performance	Journal Article
Kelly	2013	Provides metrics to measure social media performance	Book
Lovett	2011	Provides metrics to measure social media performance	Book
MarketingSherpa	2013	Provides metrics to measure e-business performance	Webpage
Meffert et al.	2008	Provide metrics to measure marketing performance	Book
Meier & Zumstein	2013	Provide metrics to measure website performance and classifies them in four dimensions	Book
Metrics	2013	Provides metrics to measure website performance	Webpage
Phippen et al.	2004	Provide metrics to measure website performance	Journal Article
Qualman	2011	Provides metrics to measure social media performance	Book
Sterne	2010	Provides metrics to measure social media performance	Book
Turner & Shah	2011	Provide metrics to measure social media performance	Book
Van der Heijden	2003	Provides metrics to measure website performance	Journal Article
Visible	2012	Defines the Top 7 metrics to measure social media ROI	Electronic Article
Welling & White	2006	Review literature (27 studies) on the website measurement	Journal Article
Wolfinbarger & Gillin	2001	Provide metrics to measure e-business performance	Journal Article

In general, the literature review showed that a lot of Web 2.0 applications, e.g. online-communities and web shops, can be analyzed with the same metrics like Web 1.0 activities, e.g. websites. However, as it comes to specific requirements, additional metrics need to be monitored (Hassler, 2012). This leads to the fact that some of the metrics cannot be monitored for each web application as they measure a specific feature of a particular type of web application. Finally, each selected subgoal and metric is explained in the following. Here specific attention is given to the special attributes of websites, online-communities and web shops as those types of web applications are selected to develop and test the newly designed method. If specific information or an illustration of more complex metrics is necessary, the three units of analysis are exemplified representatively.

6.3.2 Subgoals and metrics of web scoreboard's external perspective

Information score and its respective subgoals and metrics.

a) Popularity measured by number of page views:
   Page views, also called page impressions, is defined as each call up of a page done by a visitor. The number of page views displays the traffic of a web application (Visible, 2012; Divol et al., 2012) and measures its popularity (Hassler, 2012; Sterne, 2010).
b) Interest measured by number of visits:
   Visits are defined as "the number of times individuals request a page on the firm's server for the first time" (Farris et al., 2010, p. 328). One visit is set-up as 30 minutes time span (Hassler, 2012) and measures the customers' interest (Hoffman & Fodor, 2010).
c) Awareness measured by number of visitors (Visible, 2012):
   Visitors are counted as "the number of individuals requesting pages from the firm's server during a given period" (Farris et al., 2010, p. 328) and often are identified by the usage of cookies (Hassler, 2012). This metric helps to give answer about the awareness which is the "percentage of total population that is aware of the brand" (Farris et al., 2010, p. 30) and the first step on the way to a profitable customer experience.
d) Development awareness measured by change in number of visitors compared to previous month:
   One of the major goals of marketing is to increase the brand awareness (Kelly, 2013). To pursue the development of awareness is crucial because the trend line is often more significant than the absolute number (Hassler, 2012; Sterne, 2010).
e) Traffic sources measured by search traffic, referral traffic and direct traffic:
   For a marketer it is important to know where the traffic of a web application comes from (Gillin & Schwartzman, 2011; Metrics, 2013). In general, traffic sources are classified in three categories: search traffic, referral traffic and direct traffic. As the objective of a web application is to keep a balance between those three traffic sources (Metrics, 2013), it is important to look at all of them. Firstly, search traffic is the traffic that is generated when people use a search engine as Google, Yahoo or Bing to access a web application (Hassler, 2012) and it is usually the largest traffic source (Gillin & Schwartzman, 2011). Secondly, referral traffic is the traffic that comes from links on other websites, e.g. forums, blogs, firms' web pages (Metrics, 2013) and can be seen as a quality indicator (Hassler, 2012). Thereby, referring sites are "sites that transit a visitor to (…) [a firm's

web application] via an inbound link"[5] (Gillin & Schwartzman, 2011, p. 82). Thirdly, direct traffic is generated by users typing the URL of the web application into a web browser (Metrics, 2013) or coming from an e-mail link (Gillin & Schwartzman, 2011). A high percentage of direct traffic means that a web application is memorable and the brand is strong (Hassler, 2012). This is especially important for interactive web applications like online-communities which need returning, active members. Nevertheless, it always needs to be considered that in the long-run new visitors are essential (Metrics, 2013).

f) Relevance in Google measured by page rank:
The page rank is available for homepages and single web pages and is determined by Google on a scale between one to ten (Hassler, 2012) by analyzing various factors in an unpublished algorithm (Gillin & Schwartzman, 2011). Google is the major source for search engine traffic (Metrics, 2013). Therefore, the relevance in this search engine measured by the page rank is crucial for the success of a web application as it is one major influencing factor on how easy it is found by internet users (Hassler, 2012).

Engagement score and its respective subgoals and metrics.

a) Interaction measured by number of members (since the existence of the application):
The number of members is a valuable metric for interaction (Lovett, 2011; Sterne, 2010) as it shows the number of users who are interested in a relationship with the company. The number of members is measured in total, meaning how many members are counted since the existence of the application. For a website this figure is not available. In an online-community a member is a user who registered successfully. A web shop counts those users who registered successfully for a customer account to buy products.

b) Development of interaction measured by change in number of members compared to previous month:
Aside from the overall number of members the member development is measured. This is done because persistently pursue the trend line is essential as the progress is more significant than the absolute number of actual members (Sterne, 2010).

c) Activity index measured by number of page views per visit:
Page views per visit is a useful indicator for engagement (Gillin & Schwartzman, 2011; Sterne, 2010) because it shows how profound a visitor uses a web application and how

[5] Inbound link is defined as "a hyperlink that transits from an external web domain to (…) [the] own" (Gillin & Schwartzman, 2011, p. 82).

capable it is for maintaining the interest of visitors. To calculate the activity index, the number of page views is divided by the number of visits (Hassler, 2012).

d) Duration measured by average minutes spent on application:
Visit duration is defined as the time that passes between calling up the first page and a visit's last page. The minutes spent on the application are a valuable indicator for engagement (Gillin & Schwartzman, 2011; Sterne, 2010). However, the interpretation of this figure might be different for the different types of web applications. As long as the intent of the application is to inform the visitor by providing a lot of content or to engage the user, a long duration is positive. If a web page provides just few content or the customer wants to finish a process quickly, e.g. an order in a web shop, a shorter visit duration is valuable (Hassler, 2012; Meier & Zumstein, 2013).

e) Attractiveness measured by returning visitors rate:
The attractiveness of a web application is measured by the percentage of visitors who returns to the web application because it is assumed that the application is appealing to the respective, returning user. It can be calculated with the help of cookies. Nevertheless, a firm has to keep in mind that acquiring new users is also very important (Hassler, 2012).

f) Bounce rate measured by visitors viewed just one single page:
The bounce rate measures the "fraction of Web site visitors who view a single page" (Farris et al., 2010, p. 293), i.e. entry page equals exit page. It is an important metric for engagement and makes a statement about the suitability of a web page as an entry point (Gillin & Schwartzman, 2011; Turner & Shah, 2011).

g) Loyalty measured by number of recommendations:
Loyalty is an "indication of base future revenue stream" (Farris et al., 2010, p. 31) and recommendation of users is a valuable metric to show customer loyalty. The willingness to recommend expressed in positive mentions like "get it!", "love", "like" (Divol et al., 2012; BVDW, 2013) can be seen as a strong marketing advantage (Farris et al., 2010).

h) Criticism measured by number of negative mentions:
Negative mentions displayed by terms as "hate" or "dislike" (BVDW, 2013) need to be considered (Turner & Shah, 2011). Some criticism increases the credibility (Sterne, 2010) and is a possibility for improvement (Qualman, 2011). However, if there are too many negative mentions, a firm needs to monitor and respond cautiously to each comment.

i) Social networking potential (SNP) measured by percentage of members with high social networking potential/multiple orders:
Members with high SNP can be identified by detecting the members who have created the

most content and the highest number of positive comments or who have the biggest audience and degree of connectedness (Sterne, 2010). These persons can be used as brand ambassador to influence the target group (BVDW, 2013) and a firm should try to engage them further to talk positively about the firm (Gillin & Schwartzman, 2011). Therefore, it is important for an online application to have the possibility to identify those influencers. On a website the amount of heavy users (Meier & Zumstein, 2013) could be taken to detect persons with high SNP. However, there is no possibility to get in contact with those heavy users because the visitors stay anonym. In an online-community, members with high SNP are those who have a high activity sign. In a web shop an influencer is a customer who ordered more than once since the existence of the shop. Those customers can be expected to be online-affine and satisfied which leads to a high probability of telling others about the shop.

Conversion score and its respective subgoals and metrics.

a) Conversion generated measured by number of downloads/posts/orders:
   Conversion is the transformation of a visitor into a user behaving according to the objectives of the web application. On the Website the number of downloads is defined as conversion goal. In the Web shop the conversion is displayed by the number of product orders (Meffert et al., 2008; Meier & Zumstein, 2013) and in the Online-community it is the amount of posts in the forum (Sterne, 2010).
b) Saving scenarios measured by amount of maximal saving (in euro) through downloads:
   Downloads which are ascertain by counting "the times a file was downloaded" (Farris et al., 2010, p. 293) are a useful web metric (Lovett, 2011). Due to the fact that they substitute printed material – what saves costs for: printing, distributing as well as the personal for those issues – downloads inherent a saving potential. Besides, it can help to identify which specific topics are particularly relevant for the users (Lovett, 2011).
c) Revenue measured by amount of generated revenue through orders:
   This metric can just be measured in a web shop and shows the actual revenue generated through orders placed at this web application.
d) Innovation measured by number of ideas:
   The number of novel product ideas displays the opportunity of a web application to foster innovation by including the user in the product development process (Happe, 2008). In the web scoreboard ideas are counted regardless whether they are realistic or not and how the idea is used in the end. This metric is just measurable for an open innovation community.

e) Lead generation measured by number of completed registrations:
"A lead is an inbound prospect who is interested in your product or service" (Turner & Shah, 2011, p. 247). Lead generation is one of the major marketing objectives (Kelly, 2013) in the B2C and B2B area (Gillin & Schwartzman, 2011) and a very common web metric (Sterne, 2010). For instance, in the Online-community and in the Web Shop users who have completed the registration process are counted as leads. Taking this definition of lead generation, the Website can not generate leads. Nevertheless, leads need to be converted into customers to earn money (Gillin & Schwartzman, 2011; Turner & Shah, 2011).

f) Generated data permission measured by number of e-mail addresses (since the existence of the application):
The number of generated e-mail addresses is a valuable metric to measure the success of a web application because the data permission can be a starting point of a profitable customer journey (Lovett, 2011). Even though it is just a potential customer who obtains information (Meier & Zumstein, 2013), it might lead to an online or offline transaction or a new business partnerships at a later point in time (Stelzner, 2012).

g) Abandonment rate measured by percentage of processes (downloads/registrations/orders) started but not completed:
The abandonment rate shows the percentage of users who have started a process but did not complete it. This metric is useful, if users have to pass through a number of steps until they complete a process. The considered processes are on a website the uncompleted downloads, in an online-community the discontinued registrations and in a web shop started but uncompleted purchases (Farris et al., 2010).

General benefit score and its respective subgoals and metrics.

a) Strategic fit measured by addressed goals of strategy map:
As every company needs to decide individually which applications fit best to its goals and target groups (Bulut & Mandaric, 2012; Kaplan & Haenlein, 2010), the strategic fit is of high importance. Taking the example of XX, the strategic goals are clustered in a strategy map adjusted from the "balanced scorecard" developed by Kaplan and Norton (1992) (see Appendix 13). Thus, the web applications need to address the goals displayed in the predefined strategy map in order to contribute to the strategic business objectives.

b) Usability measured by the usage of the new design:
   Simplicity, elegant features and a clear structure expressed by a high usability are central success factors of web applications (Meffert et al., 2008; Razorfish, 2008).
c) Reaction time measured by average hours until feedback reaction/order delivery:
   Speed of feedback (Qualman, 2011) measured by the average response time to customer posts (Sterne, 2010) is an important issue in an interactive online-community. In a web shop customers want to have a fast reaction regarding their order delivery. Thus, here the reaction time is measured by the hours until an incoming order is delivered.
d) Major topics measured by top three keywords:
   "Relevant keywords are terms and phrases that your customers use when they're looking for the products or services you offer" (Gillin & Schwartzman, 2011, p. 94). To know the relevant keywords leading to a firm's web application is essential for its success, as search engines are usually the largest traffic source. Thus, firms need to detect what search terms potential customers are using and consider this in Search Engine Optimization (SEO) and Search Engine Marketing (SEM) (see chapter 8.3.2) (Gillin & Schwartzman, 2011).

6.3.3 Subgoals and metrics of web scoreboard's internal perspective

Investments score and its respective subgoals and metrics.

a) Implementation costs to duplicate measured by euro spent:
   The euro spent for the implementation of the web applications differ. In the case of XX the duplication costs, i.e. the costs to "copy" the already implemented application to another region, are calculated. Those costs display the "pure duplication" and exclude the costs for specific regional adaptations, translations, hosting as well as internal and external efforts needed for the implementation. Unlike for B and C, no figures are available for the pure duplication of A. Thus, the costs for the Website's relaunch are considered.
b) Marketing spending at the launch of the application measured by euro spent:
   How many euro are spent for marketing activities in order to promote a new web application is an interesting metric (Farris et al., 2010) because it indicates what level of user activity can be expected in the initial phase.
c) Human resources to duplicate measured by full-time equivalent (FTE):
   Also personnel and time efforts have a monetary value (Blanchard, 2012). Thus, human resources needed for the implementation need to be calculated. This metric is measured in full-time equivalent (FTE) which is "a unit to measure employed persons or students in a way that makes them comparable" (Eurostat, 2013, paragraph 1). One full-time employee

is calculated with a 40 working hours contract. In all three examined web applications, employees with expertise in translation, communication and content issues as well as technical support are required for the duplication.

Maintenance costs score and its respective subgoals and metrics.

a) Technology measured in euro spent per month:
   To maintain a web application certain technology costs like IT support, hosting and Web Content Management System (WCMS) accrue.
b) Marketing spending measured by euro spent per month:
   The continuous marketing spending of a web application determine how much activity should be expected in this application (Farris et al., 2010). It is important to monitor whether the expected outcome of marketing activities is generated or not. Such activities can be for instance Google adwords campaigns.
c) Human resources for maintenance measured by full-time equivalent (FTE):
   In general one can say that it is less likely that an application is implemented, if it requires a lot of effort (MarketingSherpa, 2010). Thus, it is important to monitor the human resources necessary to maintain a web application. On a website human resources for content actualization are needed. Aside from the content actualization, an interactive online-community and a commercial web shop need human resources for the interaction process with the target groups.

General cost score and its respective subgoals and metrics.

a) Cost of information measured by costs per visitor:
   In general, just few literature about the link of web performance to marketing and financial performance exists (Welling & White, 2006b). As it is difficult to measure the ROI of web applications, marketers developed alternatives (Blanchard, 2012; Kelly, 2013). One of those alternative metrics to get information about the efficiency of the web applications is costs per visitor. It can be determined in dividing the cost side by the number of visitors in the respective month. In order to provide an objective basis for the calculation, the implementation costs (e.g. relaunch or duplication costs) are evenly allocated over the whole period of the web application's existence. Those partitioned costs are added to the monthly recurrent costs for technology and marketing. A detail explanation of the calculation is given in chapter 8.3.5.

b) Cost of engagement measured by costs per member:
   Return on engagement is another alternative metric (Kelly, 2013). By taking the number of members of the external perspective and compare it to the cost metrics one can calculate the cost of engagement. This metric cannot be calculated for a website. For an online-community the costs per member and for a web shop the costs per registered customer account can be determined. As the number of members since the existence of the application is considered, also the total amount of implementation costs is computed.
c) Cost of conversion measured by costs per download/post/order:
   To determine the costs of conversion one needs to consider the different defined conversion goals and values of the generated conversion of each application (Hassler, 2012). As this metric also considers the benefits per month, i.e. monthly downloads/posts/orders, the implementation costs are partitioned monthly as well.
d) Critical mass measured by number of downloads/posts/orders needed for amortization:
   The calculation of the amortization value is an interesting metric. In the case of XX, the critical mass is measured by the amount of downloads at which the investments in the relaunched version of the Website will amortize. For B the number of posts at which B will generate positive conversions is calculated. In C the number of orders at which a positive business result is achieved, determines the critical mass.

6.4 Conclusion and overview of the web scoreboard

Overview web scoreboard. The result of the second research cycle which is described in chapter six is the beta version of the web scoreboard, illustrated in Table 7. This version comprises the external perspective, meaning the benefits of the web applications, as well as the internal perspective, meaning the costs of the web applications. Displayed in the three columns, the external perspective consists of four scores, 26 subgoals and 28 metrics which are aligned with the strategic objectives of online activities. The internal perspective consists of three scores, 10 subgoals and 10 metrics which show the efforts necessary to implement and maintain web applications. By determining this initial design with appropriate subgoals and metrics, research question number two is answered.

Table 7: Web scoreboard with its structure, scores, subgoals and metrics

Web performance measurement tool "web scoreboard"		
External perspective		Metrics counted per month if nothing else mentioned
Information Score	**Information subgoals**	**Information metrics**
Metric a	Popularity	No. of page views
Metric b	Interest	No. of visits
Metric c	Awareness	No. of visitors
Metric d	Development awareness	Change in no. of visitors compared to previous month
Metric e	Traffic sources	Search traffic
		Referral traffic
		Direct traffic
Metric f	Relevance in Google	Page rank
Engagement Score	**Engagement subgoals**	**Engagement metrics**
Metric a	Interaction	No. of members (since existence of the application)
Metric b	Development interaction	Change in no. of members compared to previous month
Metric c	Activity index	No. of page views per visit
Metric d	Visit duration	Average minutes spent on application
Metric e	Attractiveness	Returning visitors rate
Metric f	Bounce rate	% of visitors viewed just one single page
Metric g	Loyalty	No. of recommendations
Metric h	Criticism	No. of negative mentions
Metric i	Social networking potential (SNP)	% of members with high SNP/multiple orders
Conversion score	**Conversion subgoals**	**Conversion metrics**
Metric a	Conversion generated	No. of downloads/posts/orders
Metric b	Saving scenarios	Amount of max. saving (in euro) through downloads
Metric c	Revenue	Amount of generated revenue (in euro) through orders
Metric d	Innovation	No. of ideas
Metric e	Lead generation	No. of completed registrations
Metric f	Generated data permissions	No. of e-mail addresses (since existence of the application)
Metric g	Abandonment rate	% of processes (registrations/orders) not completed
General benefit score	**General benefit subgoals**	**General benefit metrics**
Metric a	Strategic fit	Addressed goals of strategy map
Metric b	Usability	New design is used
Metric c	Reaction time	Average hours until reaction on feedback/order is delivered
Metric d	Major topics	Top 3 keywords
Results external perspective (100%)		
Internal perspective		
Investments score	**Investments subgoals**	**Investment metrics** (amount required one time)
Metric a	Implementation costs to duplicate	Euro spent
Metric b	Marketing spending (launch)	Euro spent
Metric c	Human ressources to duplicate	Full-time equivalent (FTE) (employees)
Maintenance costs score	**Maintenance costs subgoals**	**Maintenance metrics** (amount required recurrent)
Metric a	Technology	Euro spent/month
Metric b	Marketing spending	Euro spent/month
Metric c	Human ressources	Full-time equivalent (FTE) (employees)
General cost score	**General cost subgoals**	**General cost metrics**
Metric a	Costs of information	Costs per visitor
Metric b	Costs of engagement	Costs per member
Metric c	Costs of conversion	Costs per download/post/order
Metric d	Critical mass	No. of downloads/posts/orders needed for amortization
Results internal perspective (100%)		
Results overall scores		

7 Data analysis with a weighting and grading system – 3rd research cycle

7.1 Introduction to the necessity to develop a weighting and grading system

Monitoring can provide important outcomes but only if the data can be transformed in relevant implications and interpretations (BVDW, 2013). In order to analyze and interpret the monitoring data in an eligible way, a grading and weighting system is developed. This process represents the third research cycle by using the adjusted theory of the "Web Site Assessment Tool" presented in Evans and King (1999) and verifying the outcomes by expert interviews. Thereby, the research question shown in Table 8 is answered in this chapter.

Table 8: Research question which is answered in chapter seven

Research question need to be answered
3) How can the identified metrics for web measurement be weighted and graded?

In order to answer research question number three, the development of the weighting system is described in the following. Besides, the grading system with its score scales, the decision process to determine the intervals as well as the possibility to compare certain intervals, is explained subsequently. This process of data analysis is explanatory and contributes to the research aim by providing a valuable way to interpret the input data of the web scoreboard.

7.2 Weighting system – explanation and justification

Weighting process. In order to display that the scores and metrics of the web scoreboard are not equally important, they are weighted in a scale from 0% to 100%, i.e. the higher the percentages, the more important is the score/subgoal/metric. Recommendations for the weights are set up considering the reviewed literature (see Table 6) and knowledge in the field of web measurement as well as the adjusted theory of the "Web Site Assessment Tool" presented in Evans and King (1999). In order to avoid a subjective weighting, the suggested weights have been verified by the company supervisors and the web application managers. Those experts justified and adapted the weights in the third round of interviews (see Appendix 14). The resulting weights of the scores, subgoals and metrics exemplified for the three units of analysis are shown in Table 9.

Weighting scores. The importance of each score of the external perspective is determined according to the strategic objectives of each type of web application. Again, the three units of analysis act as a representative for the different types of informative, interactive or commercial web applications. For instance, as the main purpose of the Website is

information, the information score is weighted with 50%, whereas engagement accounts to 20% and conversion also is weighted with 20%. The major objective of the Online-community is engagement. Therefore, this score is weighted with 50% but information and conversion only with 20% each. In the Web Shop, conversion – according to its main strategic objective – is weighted with 50%, while information accounts to 20% as well as engagement. The universal general benefit score is always calculated with 10%. Hereby, the scores of the external perspective together achieve 100%. By applying the common principle to evaluate factors considering the strategic goals of an web application (Evans & King, 1999), it is aimed at making the results of different types of web applications comparable in the end. The exact percentages (50%, 20%, 20%, 10%) are determined based on the fact that each online activity needs to address the whole customer journey, i.e. information and engagement and conversion, without neglecting the focus on its main objective. An illustration of the weights is provided in Appendix 15. Also, the three scores of the internal perspective add up to 100%. Here the exact percentages (30%, 30%, 40%) are determined based on the fact that an application needs to be judge equally on its implementation (30%) as well as on the maintenance costs (30%) but its efficiency, i.e. comparing the benefits and the costs, is more important (40%).

Weighting metrics. Furthermore, as not all metrics are equally important (Hassler, 2012), each single metric is weighted differently as well. The weighting of each metric has been set up according to the impact of the subgoals and their respective metrics on the achievement of the strategic objective represented by the score. To judge this contribution of the subgoals, the reviewed literature (see Table 6) and gained knowledge in the field of web measurement has been used. Besides, the metrics which display the efficiency (i.e. general costs metrics) are weighted differently based on the respective strategic aim of each type of web application. The determined weights of the metrics vary between 5% and 50%. Those percentages have been verified by the experts in order to avoid a subjective weighting.

Table 9: Weighting system of the web scoreboard

		Weights web applications		
		A	B	C
External perspective		Weight	Weight	Weight
Information subgoals	**Information metrics**	50%	20%	20%
Popularity	No. of page views	10%	10%	10%
Interest	No. of visits	10%	10%	10%
Awareness	No. of visitors	10%	10%	10%
Development awareness	Change in no. of visitors compared to previous month	25%	25%	25%
Traffic sources	Search traffic	5%	5%	5%
	Referral traffic	5%	5%	5%
	Direct traffic	5%	5%	5%
Relevance in Google	Page rank	30%	30%	30%
Engagement subgoals	**Engagement metrics**	20%	50%	20%
Interaction	No. of members (since existence of the application)	5%	5%	5%
Development interaction	Change in no. of members compared to previous month	10%	10%	10%
Activity index	No. of page views per visit	15%	15%	15%
Visit duration	Average minutes spent on application	10%	10%	10%
Attractiveness	Returning visitors rate	20%	20%	20%
Bounce rate	% of visitors viewed just one single page	15%	15%	15%
Loyalty	No. of recommendations	10%	10%	10%
Criticism	No. of negative mentions	10%	10%	10%
Social networking potential (SNP)	% of members with high SNP/multiple orders	5%	5%	5%
Conversion subgoals	**Conversion metrics**	20%	20%	50%
Conversion generated	No. of downloads/posts/orders	20%	20%	20%
Saving scenarios	Amount of max. saving (in euro) through downloads	10%	10%	10%
Revenue	Amount of generated revenue (in euro) through orders	20%	20%	20%
Innovation	No. of ideas	15%	15%	15%
Lead generation	No. of completed registrations	15%	15%	15%
Generated data permissions	No. of e-mail addresses (since existence of the application)	10%	10%	10%
Abandonment rate	% of processes (registrations/orders) not completed	10%	10%	10%
General benefit subgoals	**General benefit metrics**	10%	10%	10%
Strategic fit	Addressed goals of strategy map	50%	50%	50%
Usability	New design is used	15%	15%	15%
Reaction time	Average hours until reaction on feedback/order is delivered	15%	15%	15%
Major topics	Top 3 keywords	n.a.	n.a.	n.a.
Investments subgoals	**Investment metrics** (amount required one time)	30%	30%	30%
Implementation costs to duplicate	Euro spent	40%	40%	40%
Marketing spending (launch)	Euro spent	30%	30%	30%
Human ressources to duplicate	Full-time equivalent (FTE) (employees)	30%	30%	30%
Maintenance costs subgoals	**Maintenance metrics** (amount required recurrent)	30%	30%	30%
Technology	Euro spent/month	30%	30%	30%
Marketing spending	Euro spent/month	30%	30%	30%
Human ressources	Full-time equivalent (FTE) (employees)	40%	40%	40%
General cost subgoals	**General cost metrics**	40%	40%	40%
Costs of information	Costs per visitor	50%	20%	20%
Costs of engagement	Costs per member	20%	50%	20%
Costs of conversion	Costs per download/post/order	20%	20%	50%
Critical mass	No. of downloads/posts/orders needed for amortization	10%	10%	10%

7.3 Grading system – explanation and justification

Scores. In the grading system of the web scoreboard, the scores are evaluated in a scale from 1.0 to 5.0. Hereby, 1.0 is determined as the worst grade, 5.0 as the highest grade. In order to come to the grade, the value achieved by a web application is scheduled in the developed grading interval. The resulting grade is afterwards multiplied with the determined weight of

the metric. By adding up the single grade of the metrics, the total score is determined. Taking the total scores altogether leads to an overall score of the external and internal perspective. This calculation, shown in Figure 4, is derived from the Fishbein Attitude Model.

Figure 4: Calculation of the total scores based on the adjusted Fishbein Attitude Model

$$\text{Total score} = \sum_{i=1}^{n} \text{Rating of factor i} \times \text{Weight of factor i}$$

Source: Evans and King (1999)

Decision about intervals. It is important but difficult to decode the metrics in the right way and to decide which value should be assessed with which grade (BVDW, 2013). To determine the grade, five intervals are set up for each metric. Insights gained from three different sources in the third research cycle are considered in deciding about the values of the intervals. In the following examples for the information gathered from those three sources and used in determining the intervals are listed:

1) Literature review:
   - In a blog with around 2,000 visitors 15 - 20 messages per day can be evaluated as very impressive (Gillin & Schwartzman, 2011).
   - A page rank higher than six can be judged as very good (Hassler, 2012).
   - The three traffic sources should be balanced so that each source should bring up 33% of the total traffic. However usually search traffic accounts to more than 50% of the overall traffic of a webpage (Metrics, 2013).
   - Regarding the bounce rate it can be stated that the lower the bounce rate the better it is (Meier & Zumstein, 2013).
   - Concerning the percentage of high SNP persons exists a so called “90-9-1 rule”, i.e. in a community it can be judged as good if 90% of all users just read and observe but do not contribute, 9% contribute occasionally and 1% are highly active and responsible for most of the content (Li & Bernoff, 2008).
2) Expert interviews (third round):
   - One webpage aims for instance at 30% bounce rate.
3) Historic monitoring data:
   - The average number of visits from the last six months is graded with a “3”.

Comparison of intervals. To guarantee the comparability of the grades, the space used in the interval of one single metric (e.g. gap of 3%) is the same for each grade (e.g. < 3%; 3% - 6%; 6% - 9%; 9% -12%; > 12%). Certain metrics which are displayed in absolute numbers need separate intervals for each of the different types of web applications and developmental stages (see green colored metrics in Table 10). Those metrics are customized for the units of analysis and must be adapted for each investigated application. This critical issue is further discussed in chapter 10.1. The other metrics, mostly displayed in percentage or as an index, can be graded based on consistent intervals for all applications. To illustrate the grading system, Table 10 displays the grading intervals of B. The grading intervals of the Website and the Web Shop are shown in Appendix 16 and Appendix 17.

Table 10: Determined intervals Online-community as an example for the grading system

Excluded due to the confidentiality agreement.

Comparison results to benchmarks. One has to keep in mind that it is important to interpret the analytic data in context (Kent et al., 2011) and compare it to benchmarks in order to make a meaningful statement about it (Hassler, 2012). A benchmark which constitutes a reference value can be historic data, an industry benchmark or a competitor reference. In order to compare the own results with the results of the industry and competitors, the availability of input data is often problematic. Thus, a comparison with the own history is often the only way for recently introduced applications (Sterne, 2010).

At the present developmental stage of the web scoreboard, the actual scores can just be compared to the maximum achievable score in order to derive measures for improvement. By applying this measurement tool over a longer period, one can generate historic monitoring data that should be used as a benchmark. The development of benchmarks enables to set targets for each score and to analyze which measures were successful. However, until benchmarks for the different types and developmental stages of web applications are developed, the grading intervals must be determined for each investigated application anew. This problematic issue is further discussed in the critical analysis in chapter 10.1.

7.4 Conclusion of the data analysis

This chapter described the third research cycle. The outcome of this research cycle is the final version of the web scoreboard. By determining the different weights and grading intervals it is

possible to analyze and interpret the collected data of web applications in a useful way. Hereby, research question number three is answered. Besides, the designed weighting system implicates the identification of the most important subgoals and metrics. In the next step, the developed final version of the web scoreboard is tested by applying it to the three units of analysis. Hereby, the key metrics – identified by the determined weights – are examined in detail.

8 Practical application of the new performance measurement method

8.1 Introduction to the practical application

In order to judge the practicability of the newly designed performance measurement method, it is applied to the three units of analysis. This chapter explains the practical application of the web scoreboard and thereby answers the research questions shown in Table 11.

Table 11: Research questions which are answered in chapter eight

Research questions need to be answered
4) How is it possible to test the practical applicability of the new method?
5) What is an appropriate format for standardized reporting and for strategic guidance?

First of all, the data collection process via web analytics and social monitoring tools is exemplified. Based on the input data for the web scoreboard, each metric is graded pursuant to the determined interval scale and according to the performance of the respective application. Adding up the individual metrics which are multiplied by the defined weights, leads to scores for the external and internal perspective as well as to an overall score for each application. The resulting scores are explained in the following. Furthermore, the determined key metrics are highly relevant in order to improve the performance of the web applications. Therefore, specific measures to enhance the information, engagement and conversion scores by addressing their key metrics are explained and exemplified based on the units of analysis. Thereby research question number four is answered. In addition, to include a comparison of the costs and benefits of the web applications, the general cost score is also examined. Research question number five is answered by defining the discovered key metrics as a set of KPIs which is suggested for continuous reporting in a standardized PowerPoint chart. Moreover, as a tool for strategic guidance based on the results of the web scoreboard, an adjusted General Electric Matrix is introduced, applied and illustrated.

8.2 Data collection via web analytics and monitoring tools

Overview monitoring tools. Like most companies (Blanchard, 2012) also XX uses a combination of different monitoring tools to observe its online activities. Thus, to test the web scoreboard, quantitative and qualitative data was collected via the web analytics tools “webtrends” and “SEOquake” as well as via the social media monitoring tool “brandwatch”. With the help of those tools, data of the three web applications was collected in the inquiry period (1st until 30th of June 2013). Table 12 shows the legend of the web scoreboard which explains the different data sources of the metrics for XX. As it is aimed at that the web

scoreboard is also applicable to other firms which might not apply those three tools, a list with free monitoring tools for data collection is provided in Appendix 18.

Table 12: Different data sources shown by the legend of the web scoreboard

Data sources of metrics:	
	Webtrends
	Brandwatch
	SEOquake
	Admin account of application
	Information from contact person of investigated application
	Own calculation / evaluation based on examinations

Web analytics via webtrends. Web analytics consists of the definition, measurement, analysis and evaluation of web figures in order to derive measures to reach the determined goals of online activities (Haberich, 2013; Meier & Zumstein, 2013). In the case of XX this is basically done via the tool "webtrends". Webtrends is – as a premium service for sophisticated behavioral web analysis – an advanced company solution (Blanchard, 2012). This software package provides periodically statistical updates about web applications. All registered users of the tool can view, analyze and print the statistics online from a web browser. The evaluations are made on a 24-hour cycle which offers the possibility for a continuous analysis of the user behavior.

Web analytics via SEOquake. Concerning the web scoreboard, SEOquake is used in order to get information about the page rank of the investigated web applications. By installing this free software package in the own web browser one gets information about every website that is displayed on the browser (SEOquake, 2013).

Social media monitoring via brandwatch. Social media monitoring is the process of listening to online conversations in the Web 2.0 about a firm or brand. It is crucial to pay attention to users on a continuous basis (Divol et al., 2012) in order to learn from listening and to be successful in social media (Kumar & Mirchandani, 2012). Social media monitoring tools use software to aggregate comments from various sources. Thereby, it is possible to track the public opinion around a brand and to identify trends and needs of the potential customers (Turner & Shah, 2011). XX just started social media monitoring in 2012 and currently it is done via "brandwatch". This tool conducts a keyword monitoring which pursues and quantifies the online mentions (Blanchard, 2012). With the help of brandwatch, it is possible to get information about the tonality of content. However, one needs to keep in mind that due to sarcastic or ironic content, tools that automatically evaluate the tonality never reach 100%

accuracy but usually just around 65%. In general, sentiment tracking is not easy (Hassler, 2012) and often it is criticized in practice as being inadequate. For instance, in the three investigated tools a sentiment analysis via brandwatch was conducted, however, due to insufficient conversations about those three web applications in the WWW, the results were not yet satisfying and thus no statement about this metric could be made.

8.3 Resulting scores, key metrics and recommendations for improvement

8.3.1 Overview of resulting scores and overall score

Overview resulting scores. As described in chapter seven, a grade is determined for each metric and accompanied each score of the web scoreboard. The grades of the single scores in the external perspective allow making a statement which of the determined strategic goals of the online activities is reached in what amount by the investigated web applications in the evaluation period. The higher the score the better the particular goal is achieved. Looking at the total score of the external perspective in the inquiry period June 2013, the Web Shop performed best (2.7), followed by the Online-community (2.6) and the Website (2.3). Concerning the total score of the internal perspective, the opposite picture appeared. Here, the Website achieved the best result (3.6), followed by the Online-community (2.8) and the Web Shop (1.6). This allows making a statement that the Website achieved the best performance and the Web Shop the worst performance regarding the financial figures.

Overall score. While a direct comparison of the web applications based on the achieved single scores needs to be done with cautiousness, the overall grade – including the results of the external and internal perspective – allows an evaluation and comparison about the performance of each web application in the period of data collection. Here, the Website achieved with 3.0 the best value, followed by the Online-community (2.7) and the Web Shop (2.1). Table 13 gives an overview of the achieved scores.

Table 14 shows the detailed results of the three web applications.[6] Hereby, the different colors in the third column describe the different data sources to gather data of the determined metrics (see chapter 8.2).

[6] As Microsoft Office Excel 2007 is limited in its function to display the numbers of the web scoreboard in the English writing style (e.g. 1,000.15), the German way to display numbers is used (e.g. 1.000,15). Hereby, the thousands separator is indicated with a full stop "." while the decimal points are indicated by using a comma",".

Table 13: Overview of the resulting grades of the scores

Results web applications	A	B	C
External perspective	Grade	Grade	Grade
Information Score	2,9	2,8	2,6
Engagement Score	2,3	2,5	3,0
Conversion score	1,2	2,0	2,6
General benefit score	1,6	3,6	3,2
Results external perspective (100%)	2,3	2,6	2,7
Internal perspective			
Investments score	2,4	3,1	1,4
Maintenance costs score	4,4	2,7	2,3
General cost score	4,0	2,7	1,2
Results internal perspective (100%)	3,6	2,8	1,6
Results overall scores	3,0	2,7	2,1

Table 14: Detailed results of the web scoreboard

Excluded due to the confidentiality agreement.

8.3.2 Information score: key metrics and measures for improvement

Results information score. A was expected to be especially strong in reaching the strategic objective "information". This expectation was met as the Website achieves the highest score. This means, that it contributes to this strategic goal better than the other two web applications and thereby justifies its choice as a representative of a web application with information as its major goal. With the resulting grade of 2.9, the Website is closely followed by the Online-community (2.8) and the Web Shop (2.6). Improvement measures should be undertaken to perform better regarding this objective. Optimally this is done by addressing the two most important subgoals: "relevance in Google" and "development of awareness" with the respective metrics "page rank" (30%) and "change in number of visitors" (25%).

Page rank improvements. The page rank in search engines like Google can be addressed by Search Engine Optimization (SEO). SEO comprises all methods which aim at improving how easily a web application can be found in search engines when people search for terms that are related to a firm's brand or products (Hassler, 2012; Meier & Zumstein, 2013). For companies' web applications it is essential to get ranked on the first page of results because just very few searching users go beyond this first page (Gillin & Schwartzman, 2011). As it is the major source for search engine traffic, Google is especially important (Metrics, 2013). Even though no one beside Google knows how exactly the page rank algorithm works,

increasing the number and quality of (inbound) links which refer to the web application is widely accepted as the main impact factor (Gillin & Schwartzman, 2011; Hassler, 2012).[7] Meanwhile also Facebook likes or tweets on Twitter influence Google's page rank (BVDW, 2012b). Thus, XX needs to link its web applications to other high quality websites or social media sites in order to improve the page rank. Besides, XX should determine where traffic comes from (Metrics, 2013) and which keywords are used. Knowing this, XX can have a look at how the pages are structured that have a high ranking for those keywords and can use those keywords to optimize the content strategy of the page (Gillin & Schwartzman, 2011).

Improvements in change in number of visitors. By improving the page rank, an additional effect is that the number of visitor will increase because the web application can be found more easily in the WWW. Besides, Search Engine Marketing (SEM) can be conducted to increase the metric "change in number of visitors". The aim of SEM is to increase the number of visitors by setting-up online advertisements like Google adwords campaigns. Here monitoring, whether the activities are successful or not, is crucial (Meier & Zumstein, 2013).

8.3.3 Engagement score: key metrics and measures for improvement

Results engagement score. The Online-community did not reach – as originally expected – the highest grade in this score. On the contrary the Web Shop reached with 3.0 the highest value, followed by the Online-community (2.5) and the Website (2.3). This result might be explained by the limited activities undertaken in the Online-community in the period of data collection due to economic reasons. Thus, it should not be seen as a failed verification that B is a valuable representative of an online activity aiming at engagement. To improve the engagement score, targeted activities addressing the three most relevant subgoals with their metrics are recommended. The three major subgoals and metrics are: "attractiveness" measured by the "returning visitors rate" (20%), "bounce rate" measured by the "percentage of visitors viewed just one single page" (15%) and "activity index" measured by the "no. of page views per visit" (15%).

Improvements of the returning visitors rate. The returning visitors rate of all three web applications is not high. As it is important that existing (potential) customers come back for more (Kelly, 2013), this metric needs to be increased through targeted activities (Meier & Zumstein, 2013). On a website new, valuable content needs to be published continuously

[7] The Google page speed improvement tool "developers" can be used to identify relative simple measures for page rank improvement (e.g. reduce pictures). (Link: https://developers.Google.com/speed/pagespeed/insights).

(Weischedel & Huizingh, 2006) in order to give users a reason to come back. In an interactive application like an online-community it is important to respond fast (Divol et al., 2012), engage interesting discussions, be honest and respect the rules of social media (Kaplan & Haenlein, 2010) in order to improve this metric. Recommended activities for the Web Shop are to use the generated user e-mail addresses in order to send quarterly newsletters to the interested customers. Despite the reasonable goal to increase the returning visitors rate, a firm has to keep in mind that acquiring new users is also very important (Hassler, 2012).

Bounce rate improvements. All three investigated web applications have a high bounce rate. A high bounce rate often means that users cannot find on the web page what they have expected to find (Gillin & Schwartzman, 2011). In order to address this problem a consistent navigation according to the customer path and not according to the internal company structure needs to be configured. Further, to reduce the single-page-visits, the creation of entry pages with exact, descriptive titles as well as concise, actual, short sentences written in an easy language is important (Hassler, 2012).

Improvements of the activity index. The activity index, which displays how capable the web application is to maintain the interest of visitors, can often be improved through graphical optimization. A consistent page structure accompanied by few scrolling is needed (Hassler, 2012). Apart from an attractive design and a high usability, the activity index can be improved by emotional, relevant content and an easy navigation (Meier & Zumstein, 2013). To foster the user interactivity in the case of the Online-community, the possibility to express “like” and “dislike” should be established also in the forum of the community. Moreover, the number of page views per visit can be improved by getting more people talking about an application. In C and B this could be fostered by introducing user test reports of XX’s products. Here the identified key opinion leaders could be engaged as product testers.

8.3.4 Conversion score: key metrics and measures for improvement

Results conversion score. Regarding the conversion score the expected results were reached. C achieved the highest grade (2.6), followed by B (2.0) and A (1.2). These scores confirm the Web Shop as a representative application aiming at conversion. To improve this score, with the correspondent subgoal “conversions generated” and the metric “no. of posts/orders/downloads” (20%), targeted activities can be suggested.

Improvements in number of downloads. The number of downloads does not need improvement.

Improvements in number of posts. To increase the number of posts in the Online-community it should be considered to limit the readability of user comments to the registered community members.

Improvements in number of orders. For a web shop the most important metric is, apart from the generated revenue, the number of orders. In the case of C, those figures are relatively low and thus monitoring which should result in optimization of the application, is essential (Haberich, 2013). In general, literature identified specific issues that should be considered in a web shop: display price and additional costs (e.g. delivery costs) at the first sight, include a detailed product description and add product pictures with 360° view. Moreover, cross-selling, high usability in the order process and a high value of the shopping basket is important. Those issues can be addressed via ad words campaigns, by displaying "suggestions for additional goods" and by having personalized product offers (Hassler, 2012). Further, user test reports should be introduced because peer recommendations and experiences increase trust, which is important in e-commerce.

8.3.5 General cost score: results and measures for improvement

Results general cost score. Regarding the general cost score, the Website achieved the highest score (4.0), followed by the Online-community (2.7) and the Web Shop (1.2). This shows that the Website performs best as it comes to a comparison of benefits and costs. However, one needs to keep in mind that for the relatively established Website, which addresses a wider target group than the Online-community and Web Shop, the relaunch costs are considered while for the relatively new applications, B and C, the duplication costs are calculated.

Costs of information (costs per visitor). In order to calculate a realistic figure of this metric, the implementation costs (i.e. relaunch or duplication costs) are evenly allocated over the whole period of the web application's existence. This means that they are partitioned to the number of month since the web application – or in the case of the Website the relaunched version – is available in the WWW. The monthly maintenance costs (technology and marketing) are also included in the calculation. To improve this metric, on the one hand, the

number of visitors needs to be increased by the above explained measures and on the other hand the implementation costs need to amortize over time.

Costs of engagement (costs per member). As the number of members since the existence of the application is considered, also the total amount of the implementation and maintenance costs since its existence must be accounted. Costs of engagement measured by costs per member cannot be calculated for the Website. To improve this metric, measures to increase the number of members need to be realized (see chapter 8.3.3).

Costs of conversion (costs per download/post/order). The costs of conversion are defined differently for all three web applications. As the benefit side (number of downloads/posts/orders) is considered on a monthly basis, also the duplication costs are partitioned to the number of month since the web application is available in the WWW. To improve this metric, measures explained in chapter 8.3.4 need to be undertaken.

Critical mass (number of downloads/posts/orders until amortization). To calculate the critical mass the maintenance costs are added to the total implementation costs (i.e. relaunch or duplication costs). The maintenance costs are multiplied beforehand by the number of months since the web application exists. This total amount of costs is compared to the respective benefit side, i.e. downloads/posts/orders.

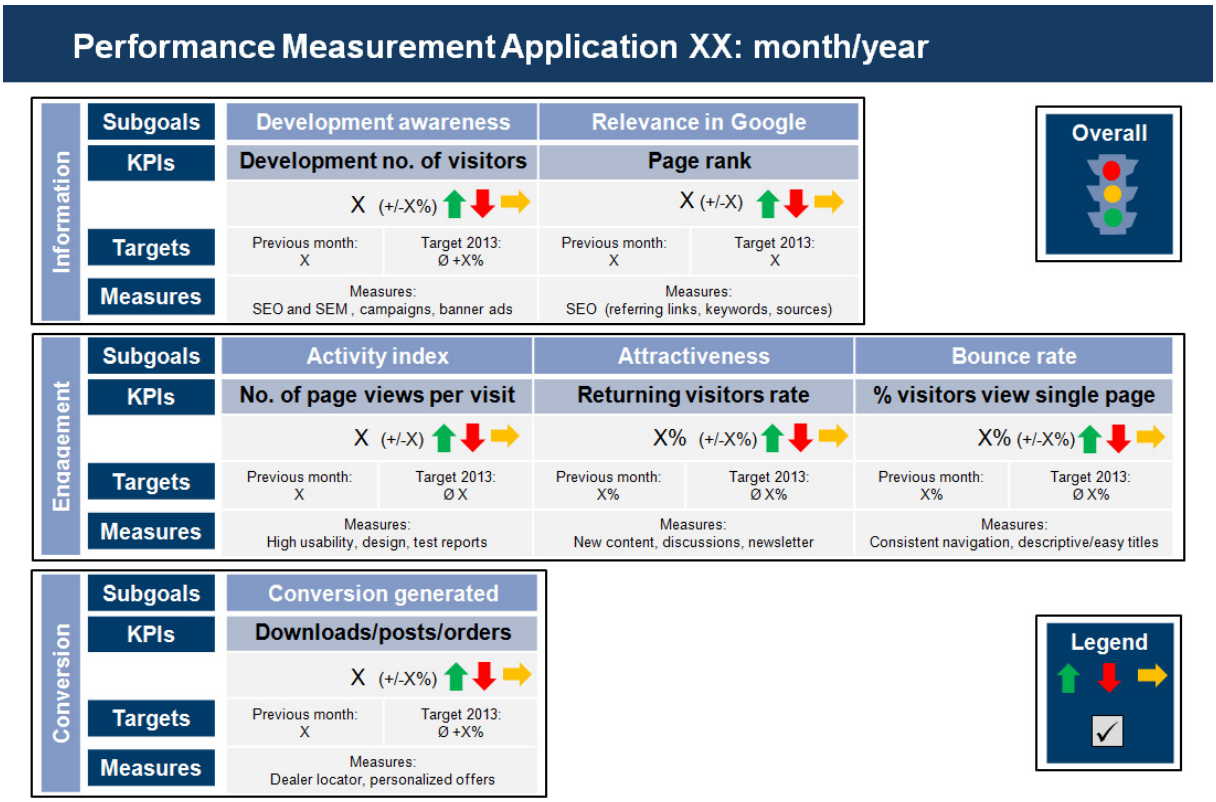
8.4 Application of a KPI set for continuous, standardized reporting

Selection KPI set. In the end, performance measurement should result in the determination of Key Performance Indicators (KPIs) which should be used for continuous measurement and standardized reporting in a consistent overview (Divol et al., 2012; Phippen et al., 2004). Hereby, KPI is defined as a metric that "is central to the well-being of the organization" (Sterne, 2010, p. 4). The KPI set should consist of five to ten carefully chosen key metrics. They either need to be aligned to strategic business objectives (Haberich, 2013) or goals must be defined for each KPI (Hassler, 2012) in order to evaluate their development (BVDW, 2012a). Based on those requirements, the six most important metrics (see chapter 8.3) of the three scores which are aligned to the strategic goals of the different online activities are defined as KPIs for web performance.

Standardized reporting format. In order to provide a consistent overview, it is suggested to use a standardized PowerPoint chart for reporting to the management. A template of this chart

is provided in Figure 5. Its usage is exemplified by the examined applications (see Appendix 19, Appendix 20 and Appendix 21). It structures the KPIs, shows the development compared to the previous month, displays the target and status and suggests improvement measures.

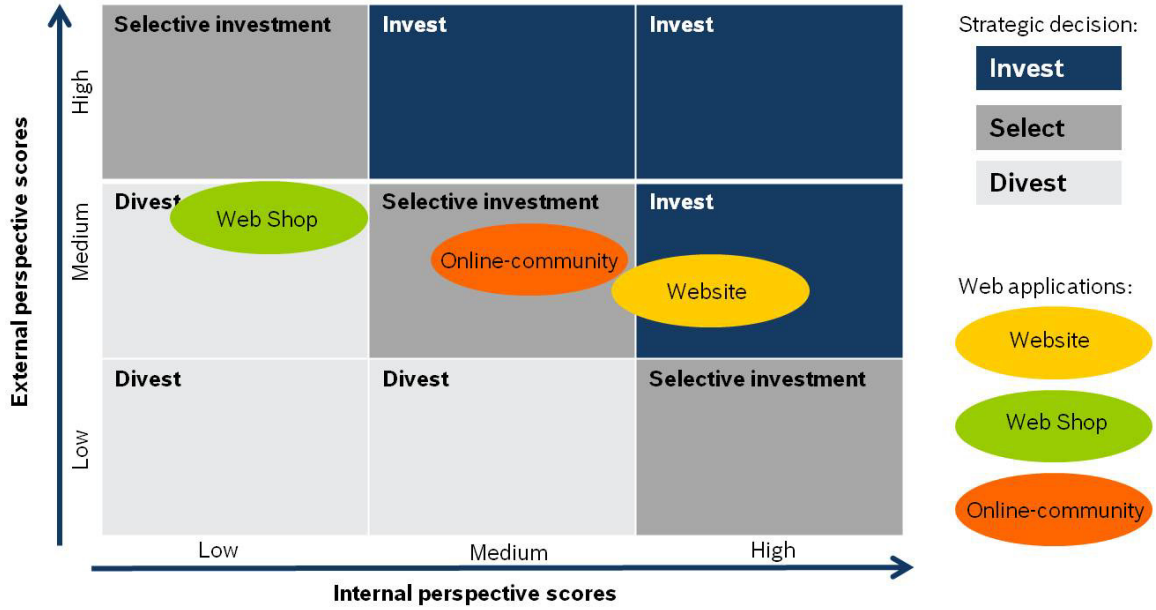
Figure 5: Draft of KPI PowerPoint chart for monthly reporting



8.5 Application of the General Electric Matrix for strategic recommendations

Adjusted General Electric Matrix. The General Electric Matrix (see chapter 5.3) is adjusted in order to interpret the results reached by each web application in the external and internal perspective. Figure 6 illustrates the adjusted General Electric Matrix and exemplifies the results of XX. The horizontal axis shows the achieved score of the internal perspective; the vertical axis shows the score of the external perspective. Each web application is mapped according to its results.

Figure 6: Illustration of the adjusted General Electric Matrix



By applying this instrument, strategic recommendations can be given concerning which web application should be pushed and which web application should be eliminated. This is illustrated by the three different colors of the nine blocks which display the recommended strategic decision: dark blue = invest; dark grey = select; bright grey = divest. In the case of XX, the Website performs very well in the internal perspective and medium in the external perspective and therefore can be classified in the "invest" area. The Online-community achieves medium results in the external as well as internal perspective and can be classified into the box "selective investment". The Web Shop renders well in the external perspective but achieves bad results in the internal perspective. Thus, this web application is categorized in the "divest" box. As the results of the internal perspective are based on some assumptions about the accrued costs, it is important to be critical when interpreting the results. Thus, the displayed recommendations of this tool about the three investigated web applications should be taken as an indication to take action and to improve the performance.

9 Findings and contributions as well as limitations and further research

9.1 Summary of findings and contributions of the thesis

Summary of the thesis. By answering the five research questions as well as the central question, the research aim of this thesis, i.e. to develop a universally applicable performance measurement method for different types of web applications, is fulfilled. The situation analysis of the online market and the introduction to the importance and problems of performance measurement in the second chapter served as a basis. None of the existing performance measurement methods – reviewed in chapter five – could fulfill all previous identified requirements for a web measurement tool; hence a new method was developed. In order to combine theoretical foundation and practical justification in the design process for the new method, three research cycles – illustrated in chapter four – were applied. Hereby, the single case study design of the XX GmbH with three representative web applications – explained in chapter three – was used. Based on a literature review and verified by expert interviews, appropriate scores, subgoals and metrics for the new method, named "web scoreboard", have been identified in chapter six. After adequate weighting and grading systems, to interpret the gathered data, have been designed in chapter seven; the practical application of the web scoreboard was tested in chapter eight. The resulting scores with their key metrics and respective improvement measures showed that the web scoreboard is an appropriate web measurement method. Besides, the results can be transformed in an adjusted General Electric Matrix to derive strategic recommendations. Moreover, the detected key metrics can be used in form of a standardized reporting KPI chart. The research process resulted in the identification of limitations and further research issues stated in this chapter. Besides, the test of the practical application serves as a basis for the critical analysis about the new method and the given recommendations for a further practical application in chapter ten.

Summary of findings and research questions. The main findings of the thesis are summarized in Table 15. This table gives an overview of the five research questions and the chapters in which they have been answered. Furthermore, it summarizes the answers to those research questions in one sentence. As already stated at the beginning of the thesis, the five research questions can be expressed in one general, central question, i.e. *"What is an appropriate method to measure the performance of different types of web applications?"*. The findings achieved by answering the subquestions, contributed in a structured, progressive way to answer the main question in the end. With the web scoreboard, an appropriate method to measure the performance of the different types of web applications has been developed. The

appropriateness of the new tool was proven by testing the practical application of the web scoreboard based on the collected data of the three selected web applications.

Table 15: Research questions and their summarized answers

Overview findings	
Research questions with their summarized answers	**Chapter**
1) Which of the existing performance measurement methods is appropriate for web measurement?	5
➡ None of the existing ones. Thus, a new web performance measurement tool "web scoreboard" is developed.	
2) Which metrics are appropriate to measure benefits and costs of different types of web applications?	6
➡ In total 38 web metrics categorized in 4 external (benefits) and 3 internal (costs) scores are selected.	
3) How can the identified metrics for web measurement be weighted and graded?	7
➡ Scores and metrics are weighted according to their relevance and the web applications' strategic goals.	
➡ Grades are set in a scale from 1 (very bad) to 5 (very good) according to predefined intervals.	
4) How is it possible to test the practical applicability of the new method?	8
➡ By applying the method to the three selected web applications and derive improvement measures.	
5) What is an appropriate format for standardized reporting and for strategic guidance?	8
➡ One PowerPoint chart which shows the developments of a set of KPIs is an appropriate reporting format.	
➡ The adjusted General Electric Matrix is an appropriate format for strategic guidance.	

Contribution of the thesis. This thesis makes several, valuable contributions to theory and practice. More precisely, with the aim to solve a common business problem (i.e. web performance measurement) by looking at a specific organizational context, this thesis develops a universally applicable method.

Overall, by applying the continuous interchange between theoretical foundation and practical justification during the research process, the thesis keeps the balance between the interests of practitioners and researchers and thereby delivers value to both. The thesis contributes to address the common management problem by providing a method to measure the benefits, business relevance, efficiency and potential for improvement of online activities. Besides, the thesis sheds light on the research gap by establishing a standardized approach for a systematic performance measurement of different types of web applications.

The new method itself makes valuable contributions. The results of its external perspective show which strategic goals can be reached in what amount by which web application. Besides, by comparing the external perspective, i.e. benefits, and the internal perspective, i.e. costs, of the respective web applications, the web scoreboard provides alternatives to the classical ROI metric in order to prove efficiency. Hereby, decisions about future investments might be simplified. Further, the overall scores indicate the performance of each web application during the investigated period. Additionally, the developed weighting and grading system suggests a way to give meaning to the gathered data of web applications and identifies

key metrics. Moreover, by uncovering and suggesting improvement measures for the key metrics, the web scoreboard helps to increase the performance of a firm's online activities.

For XX, the development process of the new method was beneficial as it raised the awareness of the web application managers and responsible persons regarding the importance of web measurement. In the practical application, the initial design based on the theoretical findings was customized to XX. Thus, the achieved results are implemented in the KPI chart and in the General Electric Matrix and suggestions to improve the web applications are given. Besides, the results of the web scoreboard can be used in order to give recommendations to other regions or divisions of XX. When they are searching for a web application to reach a specific target group or strategic objective one can give precise recommendations about necessary financial and human resources as well as expected benefits.

In general, by suggesting uniform measurement metrics and a set of KPIs as well as a format for consistent reporting and a tool for strategic decision-making, this thesis assists to standardize the measurement process. Moreover, it helps to quantify the performance and thereby provides figures to demonstrate the benefits and to justify the business relevance.

9.2 Limitations and further research

Limitations of the thesis. Like every study also this thesis has its limitations. First of all, investigating the single case about the XX GmbH in a case study design implicates a limited generalizability of the outcomes. Besides, the selection of the investigated web applications as units of analysis entails limitations. It is assumed that the three selected web applications with their different strategic goals represent diverse types of online activities in order to guarantee that the developed tool provides a measurement method for diverse types of web activities. Moreover, as the data collection took place at a single period of time (June 2013), the results of the web scoreboard represent a snap shot picture.

The comparability of single achieved scores of the web scoreboard is limited. Comparing different types of applications with different strategic objectives (information, engagement, conversion), different target groups coming from different regions is problematic. Even though it is assumed that the addressed users do not have relevant dissimilarities in culture and in their online user behavior those issues need to be kept in mind when comparing the results of the web scoreboard. In any case, unlike the separate scores, the overall scores are comparable because – looking at it in its entirety – the designed weighting system considers

the different strategic goals. However, also the weights inherent limitations as they have been set up based on the assumption that the online activities of each company can be categorized in three strategic goals. In addition, the web scoreboard offers a limited possibility to examine the user structure and characteristics. In sum, those limitations need to be admitted but they do not affect the importance of the outcomes and some of them can be addressed by further research.

Further research. As the designed web performance measurement tool is based on a single case study about the XX GmbH, its applicability for other web applications, other divisions or firms, needs to be proven in the future. Normally, by selecting the structure, scores, subgoals and metrics of the web scoreboard based on various types of online activities its general applicability should be guaranteed. However, to check this general implementation and to get comparable results, it is recommended to apply the web scoreboard in the next step to web applications whose results can be directly compared to each other (e.g. two online communities, two websites or two web shops within the same development stage). Based on this also benchmarks for the different types of web applications could be developed.

To overcome the limitation of a snap shot picture and to analyze the development of the web applications it is recommended to gather data over a longer period of time. Thereby, goals can be set up based on historical, panel data. In addition, in a future research it would be very interesting to expand the method to web applications of other firms from the same industry in order to develop industry benchmarks. The big challenge here would be the data collection. Overall, the developed method needs to be applied over time also to other web applications, business divisions and companies in order to verify its generalizability and function. This leads to the creation of benchmarks and ensures the comparability of the results.

In addition, after establishing a continuous measurement and reporting based on the outcomes of this thesis, it can be recommended to further conduct an online user survey in order to gain insights into the structure and characteristics of the web applications' users. As knowledge is power, the more a company knows about the customer preferences the better it can adjust its activities and strategic position (Blanchard, 2012). For XX, an online survey to generate insights – based on the brand funnel – is explained in detail in Appendix 22. Besides, an English and German survey draft is provided in Appendix 23 and Appendix 24.

.

10 Discussion, recommendations and conclusion

10.1 Critical discussion and recommendations to apply the new method

Critical analysis and overall recommendations. Even though the newly designed web scoreboard already proved its practical applicability, certain issues regarding the application of the new method are still critical. Those issues need to be discussed in order to provide continuously an effective way of performance measurement for different web applications, different business divisions and different companies. In general, the configuration of an appropriate infrastructure, the definition of clear roles and responsibilities, the development of eligible processes (Divol et al., 2012) as well as an open, knowledge-sharing framework (McKinsey Global Institute, 2012) are necessary conditions in order to support a successful application of web performance measurement. Thus, based on the identified critical issues, recommendations regarding technical, procedural, cultural and organizational issues as well as suggestions concerning standardization and responsibilities are given in the following.

Data stream and technical recommendations. A specific issue which needs to be addressed in the implementation of the new method is the organization of an automated, dynamic data stream to feed the web scoreboard with the required information. In the case of XX, it is recommended to use the already implemented monitoring tool webtrends. For each web application a functioning profile must be set up which contains the metrics of the web scoreboard. Besides, a webtrends account needs to be created for every employee in the online surrounding. In order to guarantee an efficient, target-oriented usage of the monitoring tool, web-based trainings about its functions should be provided. By using the possibility of webtrends to configure personalized dashboards, each web application manager can generate and export an (almost)[8] real-time overview within a few clicks. The data export should be connected to the Excel template of the web scoreboard in the next step. Once implemented, the newly collected data via webtrends results directly in scores of the web scoreboard. If another business division or company wants to apply the web scoreboard, the monitoring tool webtrends is probably not implemented. Therefore, a list of other free monitoring tools is provided in Appendix 18. Those tools need to be installed in order to collect data for the web scoreboard. Common tools (e.g. google analytics) provide also an export function which enables to set up an automated data feed for the web scoreboard.

[8] The evaluations are made on a 24-hour cycle. This offers a continuous analysis of the user behavior. All registered users of the tool can view and print the statistics online from a web browser.

Weighting and grading system and recommendations for standardization. However, before the web scoreboard can indicate the performance, one critical issue is the adaption of the grading system to the specific requirements of each investigated online activity. Here, suitable grading intervals for the absolute metrics and the cost calculation have to be adjusted. This customized determination of the grading intervals to the respective (type) of web application needs to be addressed and should get standardized with the help of benchmarks which evolve over time by investigating comparable web applications. Furthermore, each investigated web application has to be classified into a specific, standardized type of web application, i.e. it has to be assessed if a web application's main objective is information, engagement or conversion. This is necessary due to the fact that the determined weights are based on the respective strategic objectives of the different types of web applications. In general, standardization of procedures and tools is important in order to ease the daily business (Meier & Zumstein, 2013; Pauwels et al., 2009). Thus, the application of the web scoreboard, the KPI chart and the General Electric Matrix as standardized tools in a regular measurement procedure (5.3) should be fostered.

Continuous application and procedural recommendations. Continuous monitoring is essential (Divol et al., 2012). However, as shown by the weights not all metrics of the web scoreboard are equally important (Hassler, 2012) and thus different measurement periods are necessary. In the case of XX, so far no automated data stream is established, meaning the data for the web scoreboard needs to be collected manually via the available monitoring tools. Due to this time-consuming process, a quarterly application should be sufficient until this situation has changed. By implementing the web scoreboard to other business divisions or companies, it should be aimed at having an automated data feed right from the beginning and collecting data for the web scoreboard on a monthly basis. Hereby, clear goals for each score should be set in advance in order to monitor the development and track whether the improvement measures were successful or not (Hassler, 2012). Apart from the web scoreboard the set of key metrics needs to be monitored and analyzed more often in order to pursue the main developments of the web applications. As the online environment is changing very quickly and requires immediate action, web tracking should be a fixed component in the daily life of a web application manager as it acts as an early warning system. Thus, the identified KPIs should be monitored at the best on a daily or weekly basis, however, at least on a monthly basis. At XX, five of the six identified key metrics (indicated by the green color in Table 16) can be easily monitored with the help of personalized dashboards configured in the monitoring tool webtrends. For the remaining KPI (page rank), the tool SEOquake (see

chapter 8.2) offers an easy way for monitoring. By using the template of the PowerPoint chart introduced in chapter 8.4 a standardized format for continuous reporting is provided. Table 16 gives an overview of the KPIs. Here, the green color indicates webtrends as data source; the grey color depicts the monitoring tool SEOquake.

Table 16: Determined KPI set for XX with data sources

Information Score	**Information subgoals**	**Information metrics**
Metric d	Development awareness	Change in no. of visitors compared to previous month
Metric f	Relevance in Google	Page rank
Engagement Score	**Engagement subgoals**	**Engagement metrics**
Metric c	Activity index	No. of page views per visit
Metric e	Attractiveness	Returning visitors rate
Metric f	Bounce rate	% of visitors viewed just one single page
Conversion score	**Conversion subgoals**	**Conversion metrics**
Metric a	Conversion generated	No. of downloads/posts/orders

Application of results and cultural recommendations. Utilizing the results of the web scoreboard is a critical subject as well. In the web environment it is a common approach to experiment and test which activities work for which target group (Hassler, 2012). Thus, a "trial and error mentality" of customer-focused activities supported by continuous measurement and adaption needs to be applied. The collected data is of no use until this data is understood and the outcomes are applied (Phippen et al., 2004). Therefore, after the web scoreboard has identified potential for improvement of single web applications, the recommended measures have to be realized in order to improve the performance. This execution has to be done by each web application manager on its own initiative. To start the process at XX, the identified improvement measures for the three investigated web applications have been discussed with the respective web application managers.

Coordination issue and organizational recommendations. The coordination and integration aspect of web application management should not be underestimated. Especially web measurement is an interdisciplinary issue that needs a lot of communication effort. An easy and central information access is essential. For instance, a community for the web application managers in the internal communication tool can foster the internal collaboration and the knowledge-sharing culture.

Management support and recommendations for responsibilities. It is crucial to raise the awareness of the management that web applications need to be seen as an integral part of today's business. The support of the management is necessary, in order to implement the developed performance measurement method and the identified set of KPIs in the daily

business. Further, the responsibility of performance measurement should be allocated to the respective web application managers and should not be outsourced to a third party. Via monitoring tools as webtrends and by applying the outcomes of this thesis, the necessary foundations for an internal web measurement are given. Moreover, if the web application managers are responsible for web performance measurement, they get directly confronted with the developments of their online activities.

10.2 Conclusion

Nowadays it is increasingly important to justify the business relevance and legitimate investments in business activities by providing quantitative performance figures. In order to address the common problem of identifying the benefits, business relevance, efficiency and potential for improvement of various online activities, a new performance measurement method has been developed in this thesis. Different types of online activities represented by three web applications, operated by the XX GmbH, have been used as units of analysis in this development process. By answering the research questions and providing a universally applicable web performance measurement tool, this thesis makes several valuable contributions to theory and practice.

The new measurement tool “web scoreboard” provides an appropriate method to measure the performance of different types of online activities based on consistent metrics. The test of the practical application verified that the web scoreboard indicates the performance of each web application regarding the benefits and the costs during the investigated period. By implementing the web scoreboard as well as the recommended KPI set and the General Electric Matrix for strategic recommendations, web application managers can track the developments of their online applications and derive improvement measures.

Nevertheless, the discussion identified some critical issues which still need to be addressed in order to provide appropriate conditions for a continuous and efficient application of the new method. Those challenges need to be tackled in a next step by implementing the given recommendations regarding technical, procedural, cultural and organizational issues. Besides, in order to overcome the limitations of the thesis, further research issues are proposed.

To conclude, with sufficient management support and adequate human as well as financial resources, web performance measurement – and thereby the outcomes of this thesis – can be used in order to exploit the full underlying potential of different types of web applications.

11 References

Book.

Artz, M. (2010). *Controlling in Marketing und Vertrieb*. Wiesbaden, Germany: Gabler Verlag

Blanchard, O. (2012). *Social Media ROI*. Munich, Germany: Addison-Wesley Verlag.

Farris, P. W., Bendle, N. T., Pfeifer, P. E., & Reibstein, D. J. (2010). *Marketing Metrics: The definitive guide to measuring Marketing performance* (Second Edition ed.). New Jersey, USA: Pearson Education, Inc..

Gillin, P., & Schwartzman, E. (2011). *Social Markting to Business Customer: listen to your B2B market, generate major account leads, and built client relationships*. New Jersey, USA: John Wiley & Sons, Inc..

Haberich, R. (2013). *Future digital business* (First edition ed.). Heidelberg, Germany: mitp.

Hassler, M. (2012). *Web Analytics*. Heidelberg, Germany: mitp.

Kaplan, R. S., & Norton, D. P. (2004). *Strategy maps: Converting intangible assets into tangible outcomes*. Boston, USA: Harvard Business Press.

Kelly, N. (2013). *How to measure social media: A Step-by-Step Guide to Developing and Assessing Social Media ROI*. Boston, USA: Pearson Education Inc.

Kotler, P., & Pfoertsch, W. (2010). *Ingredient Branding: Making the Invisible Visible*. Berlin, Germany: Springer-Verlag.

Li, C., & Bernoff, J. (2008). *Groundswell: Winning in a world transformed by social technologies*. Boston, USA: Harvard Business Press.

Lovett, J. (2011). *Social media metrics secrets*. Hoboken, USA: John Wiley & Sons, Inc..

Meffert, H., Burmann, C., & Kirchgeorg, M. (2008). *Marketing: Grundlagen marktorientierter Unternehmensführung* (10. revised edition ed.). Wiesbaden, Germany: GMV Fachverlage GmbH.

Meier, A., & Zumstein, D. (2013). *Web Analytics & Web Controlling* (First edition ed.). Heidelberg, Germany: dpunkt.verlag GmbH.

Qualman, E. (2011). *How Social Media transforms the way we live and do business*. Hoboken, USA: John Wiley & Sons, Inc..

Sterne, J. (2010). *Social media metrics: How to measure and optimize your marketing investment* (Vol. 3). Hoboken, USA: John Wiley & Sons, Inc..

Turner, J., & Shah, R. (2011). *How to make money with social media: An insider's guide on using new and emerging media to grow your business* (Third printing ed.). New Jersey, USA: Pearson Education, Inc..

Yin, R. K. (2009). *Case study research: Design and methods* (Vol. 5). California, USA: Sage Publications.

Zhou, Z. P., & Zuo, Z. P. (2010). *BCG Matrix's Problems and Its Reconstruction*. New York, USA: St Plum-Blossom Press Pty, Ltd.

Book Section.

Hoffman, D. L., & Novak, T. P. (2012). Social Media Strategy. V. Shankar, G. S. Carpenter, J. Farley & B. A. Hamilton (Eds.), In *Handbook of Marketing Strategy* (pp. 198-215): Edward Elgar Publishing.

Conference Paper.

Bulut, I., & Mandaric, M. (2012). *Researching the role of social media marketing in contemporary business practices.* Paper presented at the symorg 2012, Zlatibor, Serbia.

Cole, R., Purao, S., Rossi, M., & Sein, M. K. (2005). *Being Proactive: Where Action Research Meets Design Research*. Paper presented at the ICIS, Las Vegas, USA.

Constantinides, E., & Zinck Stagno, M. C. (2012). *Measuring Social Media Activity: An Index Development and Validation*. Paper presented at the European Marketing Academy Conference, Lisbon, Portugal.

Hutter, K. (2012). *The impact of social media on brand awareness and purchase intention: The case of MINI on Facebook*. Paper presented at the European Marketing Academy, Lisbon, Portugal.

Järvinen, J. (2012). *How Online Environment Improves Industrial Marketers' Ability to Measure Marketing Communication Impacts?* Paper presented at the European Marketing Academy, Lisbon, Portugal.

Van Grembergen, W., & Amelinckx, I. (2002). *Measuring and managing e-business projects through the balanced scorecard*. Paper presented at the 35th Annual Hawaii International Conference on System Sciences, Hawaii, USA.

Weischedel, B., & Huizingh, E. K. R. E. (2006). *Website optimization with web metrics: a case study*. Paper presented at the 8th international Conference on Electronic Commerce, USA, New York.

Electronic Article.

Etlinger, S., & Li, C. (2011). A framework for Social Analytics. 1-40. Retrieved from http://de.slideshare.net/setlinger/altimeter-social-analytics081011final?from=embed

Eurostat. (2012). News release: Internet access and use. 1-3. Retrieved from http://epp.eurostat.ec.europa.eu/cache/ITY_PUBLIC/4-18122012-AP/EN/4-18122012-AP-EN.PDF

McKinsey. (2012). McKinsey Global Survey results: Minding your digital business. 24-31. Retrieved from http://www.mckinsey.com/~/media/mckinsey/dotcom/client_service/High%20Tech/PDFs/MOBT_27_Minding_your_digital_business-McKinsey_Global_Survey_results.ashx

Razorfish. (2008). Consumer experience report. 1-84. Retrieved from http://de.slideshare.net/NakedTokyo/feed-the-razorfish-consumer-experience-report

Visible. (2011). Finding ROI in social media. 1-2. Retrieved from http://www.visibletechnologies.com/blog/finding-the-roi-in-social-media/#more-10149

Visible. (2012). Top 7 metrics for social media roi. 1-3. Retrieved from http://www.visibletechnologies.com/wp-content/uploads/2011/12/top-7-metrics-for-social-media-roi.pdf

Electronic Book.

Kotler, P. (2002). Marketing Management, Millenium Edition (pp. 465). Retrieved from VNU-UEB database Retrieved from http://dl.ueb.edu.vn/bitstream/1247/2250/1/Marketing_Management_-_Millenium_Edition.pdf

Journal Article.

Benbasat, I., Goldstein, D. K., & Mead, M. (1987). The case research strategy in studies of information systems. *Mis Quarterly, 11*(3), 369-386.

Bernoff, J., & Li, C. (2008). Harnessing the power of the oh-so-social web. *MIT Sloan Management Review, 49*(3), 36.

Boyer, K. K., & Swink, M. L. (2008). Empirical elephants - why multiple methods are essential to quality research in operations and supply chain management. *Journal of Operations Management, 26*(3), 338-344.

Bremser, W. G., & Chung, Q. (2006). A framework for performance measurement in the e-business environment. *Electronic Commerce Research and Applications, 4*(4), 395-412.

Constantinides, E., & Fountain, S. J. (2008). Web 2.0: Conceptual foundations and marketing issues. *Journal of Direct, Data and Digital Marketing Practice, 9*(3), 231-244.

Divol, R., Edelman, D., & Sarrazin, H. (2012). Demystifying Social Media. [Consultancy Report]. *McKinsey Quarterly 2*, 1-11.

Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of Management Review, 14*(4), 532-550.

Epstein, M., & Manzoni, J.-F. (1998). Implementing corporate strategy: from tableaux de bord to balanced scorecards. *European Management Journal, 16*(2), 190-203.

Evans, J. R. (2004). An exploratory study of performance measurement systems and relationships with performance results. *Journal of Operations Management, 22*(3), 219-232.

Evans, J. R., & King, V. E. (1999). Business-to-business marketing and the world wide web: planning, managing, and assessing web sites. *Industrial Marketing Management, 28*(4), 343-358.

Hanna, R., Rohm, A., & Crittenden, V. L. (2011). We're all connected: The power of the social media ecosystem. *Business Horizons, 54*(3), 265-273.

Hennig-Thurau, T., Malthouse, E. C., Friege, C., Gensler, S., Lobschat, L., Rangaswamy, A., & Skiera, B. (2010). The impact of new media on customer relationships. *Journal of Service Research, 13*(3), 311-330.

Hevner, A. R. (2007). The three cycle view of design science research. *Scandinavian journal of information systems, 19*(2), 87.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *Mis Quarterly, 28*(1), 75-105.

Hoffman, D. L., & Fodor, M. (2010). Can You Measure the ROI of Your Social Media Marketing? *MIT Sloan Management Review, 52*(1), 41-49.

Homburg, C., Artz, M., & Wieseke, J. (2012). Marketing Performance Measurement Systems: Does Comprehensiveness Really Improve Performance? *Journal of Marketing, 76*, 56-77.

Huizingh, E. K. (2002). The antecedents of web site performance. *European Journal of Marketing, 36*(11/12), 1225-1247.

Johnson, K. L., & Misic, M. M. (1999). Benchmarking: a tool for web site evaluation and improvement. *Internet Research, 9*(5), 383-392.

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons, 53*(1), 59-68.

Kaplan, R. S., & Norton, D. P. (1992). The balanced scorecard–measures that drive performance. *Harvard Business Review, 70*(1), 71-79.

Karpinski, R. (2005). The next phase: bottom-up marketing. *BtoB Magazine, 90*(5), 38.

Kent, M. L., Carr, B. J., Husted, R. A., & Pop, R. A. (2011). Learning web analytics: A tool for strategic communication. *Public Relations Review, 37*(5), 536-543.

Kozinets, R. V., de Valck, K., Wojnicki, A. C., & Wilner, S. J. S. (2010). Networked Narratives: Understanding Word-of-Mouth Marketing in Online Communities. *Journal of Marketing, 74*(2), 71-89.

Kumar, V., & Mirchandani, R. (2012). Increasing the ROI of Social Media Marketing. *MIT Sloan Management Review 54*(1), 54-61.

Mangiuc, D. M. (2009). Measuring Web 2.0 Efficiency. *Annales Universitatis Apulensis Series Oeconomica, 1*(11), 74-87.

Marr, B., & Schiuma, G. (2003). Business performance measurement–past, present and future. *Management Decision, 41*(8), 680-687.

McCutcheon, D. M., & Meredith, J. R. (1993). Conducting case study research in operations management. *Journal of Operations Management, 11*(3), 239-256.

Meredith, J. (1998). Building operations management theory through case and field research. *Journal of Operations Management, 16*(4), 441-454.

Merz, M. A., He, Y., & Vargo, S. L. (2009). The evolving brand logic: a service-dominant logic perspective. *Journal of the Academy of Marketing Science, 37*(3), 328-344.

Murdough, C. (2009). Social media measurement: It's not impossible. *Journal of Interactive Advertising, 10*(1), 94-99.

Nambisan, S. (2002). Designing virtual customer environment for new product development: toward a theory *Academy of Management Review, 27*(3), 392-413.

Neely, A. (1999). The performance measurement revolution: why now and what next? *International Journal of Operations & Production Management, 19*(2), 205-228.

Neely, A., Gregory, M., & Platts, K. (1995). Performance measurement system design: a literature review and research agenda. *International Journal of Operations & Production Management, 15*(4), 80-116.

Novak, T. P., & Hoffman, D. L. (1997). New metrics for new media: toward the development of Web measurement standards. *World Wide Web Journal, 2*(1), 213-246.

Pauwels, K., Ambler, T., Clark, B. H., LaPointe, P., Reibstein, D., Skiera, B., Wiesel, T. (2009). Dashboards as a Service Why, What, How, and What Research Is Needed? *Journal of Service Research, 12*(2), 175-189.

Phippen, A., Sheppard, L., & Furnell, S. (2004). A practical evaluation of Web analytics. *Internet Research, 14*(4), 284-293.

Pires, G. D., Stanton, J., & Rita, P. (2006). The internet, consumer empowerment and marketing strategies. *European Journal of Marketing, 40*(9/10), 936-949.

Pries-Heje, J., & Baskerville, R. (2008). The design theory nexus. *Mis Quarterly, 32*(4), 731-755.

Riegner, C. (2007). Word of mouth on the web: The impact of Web 2.0 on consumer purchase decisions. *Journal of Advertising Research, 47*(4), 436-447.

Sein, M., Henfridsson, O., Purao, S., Rossi, M., & Lindgren, R. (2011). Action design research. *Mis Quarterly, 35*(2), 37-56.

Straub, D. W., Hoffman, D. L., Weber, B. W., & Steinfield, C. (2002). Measuring e-commerce in Net-enabled organizations: An introduction to the special issue. *Information Systems Research, 13*(2), 115-124.

Tapp, A., & Hughes, T. (2004). New technology and the changing role of marketing. *Marketing Intelligence & Planning, 22*(3), 284-296.

Velcu-Laitinen, O., & Yigitbasioglu, O. M. (2012). The Use of Dashboards in Performance Management: Evidence from Sales Managers. *The International Journal of Digital Accounting Research, 12*(18), 39-58.

Wacker, J. G. (1998). A definition of theory: research guidelines for different theory-building research methods in operations management. *Journal of Operations Management, 16*(4), 361-385.

Welling, R., & White, L. (2006a). Web site performance measurement: promise and reality. *Managing Service Quality, 16*(6), 654-670.

Welling, R., & White, L. (2006b). Measuring the value of corporate web sites. *Journal of Internet commerce, 5*(3), 127-145.

Wirtz, B. W., Schilke, O., & Ullrich, S. (2010). Strategic Development of Business Models: Implications of the Web 2.0 for Creating Value on the Internet. *Long Range Planning, 43*(2–3), 272-290.

Yigitbasioglu, O. M., & Velcu, O. (2012). A review of dashboards in performance management: Implications for design and research. *International Journal of Accounting Information Systems, 13*(1), 41-59.

Report.

BITKOM. (2012). Social Media in German companies (pp. 15). Germany, Berlin: Bundesverband Informationswirtschaft, Telekommunikation und neue Medien e.V. .

BVDW. (2012a). B2B Social Media (pp. 37). Germany, Düsseldorf: Bundesverband Digitale Wirtschaft e.V..

BVDW. (2012b). Social Search: Eine Einführung in die Welt der sozialen Suche (pp. 40). Germany, Düsseldorf: Bundesverband Digitale Wirtschaft e.V..

BVDW. (2013). Social Media Kompass 2012/2013 (pp. 168). Germany, Düsseldorf Bundesverband Digitale Wirtschaft e.V.

MarketingSherpa. (2010). Social Media Marketing Benchmark Report (pp. 15). USA, Warren: MarketingSherpa Inc..

McKinsey Global Institute. (2012). The social economy: Unlocking the value and productivitiy through social technologies (pp. 184). USA, New York McKinsey Global Institute.

Nielsen Company. (2009). Global Places and Networked Places (pp. 16). USA, New York: Nielsen Company.

Stelzner, M. A. (2012). Social Media Marketing Industry Report (pp. 42). n.a.: Social Media Examiner.

TNS. (2011). Digital Life Report (pp. 109). n.a.: TNS Infratest.

Web page.

Blanchard, O. (2009, 24th of August 2009). Social Fresh Conference: Basics of social media ROI Retrieved 9th of April, 2013, from http://de.slideshare.net/thebrandbuilder/olivier-blanchard-basics-of-social-media-roi

Engagementdb. (2009, July 2009). The world's most valuable brands. Who's most engaged? Retrieved 10th of April, 2013, from http://de.slideshare.net/charleneli/engagmentdb-report-july-2009

Eurostat. (2013, 24th of July 2013). Glossary: Full-time equivalent Retrieved 23rd of July, 2013, from http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/Glossary:Full-time_equivalent

Happe, R. (2008, 1st of April 2008). Social Media Metrics Retrieved 10th of April, 2013, from http://www.thesocialorganization.com/social-media-metrics.html

McKinsey. (2008, 2013). Enduring Ideas: The GE-McKinsey nine-box matrix Retrieved 15th of August, 2013, from http://www.mckinsey.com/insights/strategy/enduring_ideas_the_ge_and_mckinsey_nine-box_matrix

Metrics, W. (2013, n.d.). Website Traffic Sources Retrieved 9th of July 2013, from http://www.wwwmetrics.com/website-traffic.htm

O'Reilly, T. (2005, 2013). What is Web 2.0 Retrieved 25th of June, 2013, from http://oreilly.com/web2/archive/what-is-web-20.html

SEOquake. (2013, 2012). SEOquake parameters Retrieved 17th of August, 2013, from http://www.seoquake.com/pages/guide.php#l_parameters

12 Index of appendix

Appendix

Appendix 1: Illustration of XX's distribution process

Excluded due to the confidentiality agreement.

Appendix 2: Web applications used by XX with their target groups and URLs

Excluded due to the confidentiality agreement.

Appendix 3: Interview partners of the expert interviews

Excluded due to the confidentiality agreement.

Appendix 4: Interview questions for the expert interviews (first round)

The following interview questions have been asked in the oral expert interviews conducted with eleven web application managers of XX in the period of 11th of April until 21st of May 2013. Apart from general information regarding the respective web application, the definition of requirements for a web performance measurement tool was focused.

1) Which target groups are addressed with your web application?
2) Since when is your web application available in the WWW for every internet user?
3) What is the major strategic objective you pursue with your web application?
4) Do you monitor and measure the performance of your web applications already? If yes, based on what criteria does this take place? Do you use "webtrends" or "brandwatch" or other monitoring tools (e.g. etracker, google analytics) to gather monitoring data? How and in what format the gathered data is used?
5) Imagine, a new web performance measurement method shall be developed and implemented, what requirements need to be fulfilled by this tool? As a proposal from your colleagues, I can show you the following requirements for performance measurement. From your point of view, do you miss some necessary requirements? Or does the list contain some requirements which you would eliminate?

Appendix 5: The strategic objectives of XX's online activities

Excluded due to the confidentiality agreement.

Appendix 6: Homepage of the Website

Excluded due to the confidentiality agreement.

Appendix 7: Homepage of the Online-community

Excluded due to the confidentiality agreement.

Appendix 8: Homepage of the Web Shop

Excluded due to the confidentiality agreement.

Appendix 9: Timeline of the research process of the whole thesis

Excluded due to the confidentiality agreement.

Appendix 10: List of performance measurement methods after preselection

Overview preselected measurement methods		
Method	**Description**	**Source**
Balanced Scorecard	Description and reasons to select for further examination see chapter 5.3	Kaplan & Norton (1992)
Dashboard	Description and reasons to select for further examination see chapter 5.3	e.g. Farris et al. (2010); Pauwels et al. (2009); Velcu-Laitinen & Yigitbasioglu (2012)
Set of metrics	Description and reasons to select for further examination see chapter 5.3	e.g. Meier & Zumstein (2013); Weischedel & Huizingh (2006)
Performance Pyramid	Hierarchy of financial and non-financial performance measures which illustrates the link between strategy and day-to-day business. As the hierarchical approach is not suitable for web performance measurement, the pyramid is not considered in the thesis.	Lynch & Cross (1990)
Performance Prism	Innovative performance measurement framework which includes all stakeholders. It was developed at the Cranfield University. As the more popular balanced scorecard is prefered over the value-based Performance Prism, this method is neglected in the thesis.	Neely et al. (2002)
Tableau de bord	Similar to a dashboard. Aims at helping to understand relationships between processes and actions. Came up in France at the turn of this century. As the more popular and recommended dashboard technique was chosen for further examination, this approach is neglected in the thesis.	Epstein & Manzoni (1998)
Malcolm Baldrige Model	Influential, widely-accepted framework for organizational performance measurement. Established in 1987 in the US Congress and named after Malcom Baldrige (former Secretary of Commerce), it shows criteria for Performance Excellence in seven categories. As the more popular balanced scorecard is identified as more appropriate to fulfill the requirements, this model is not considered in the thesis.	Baldrige (2012)
DuPont	In the beginning ot the 20th centrury, Du Pont already applied almost all basic techniques that are still used to manage businesses nowadays. One famous example is the DuPont analysis which breaks down the ROE (Return On Equity) into three parts. As it is more a financial performance calculation it is neglected in the thesis.	Neely et al. (1999); Chandler (1977)
ABC analysis	Popular and effective method to classify inventory items into three categories (A, B, C). Thereby, it provides a mechanism to identify items that will highly influence the overall inventory cost. As the General Electric Matrix gives the possibility to derive strategic recommendations based on the costs as well as benefits of the web applications, the ABC analysis is not considered in the thesis.	Innes & Mitchell (1990); Yu (2011)
Web Site Assessment Tool	States that every assessment tool needs to consists of five components: "categories (broad areas to be investigated); factors (specific elements comprising each category); weights (importance placed on each category and factor); ratings (scores assigned to each category and factor); and total score (an overall compilation based on both weights and ratings)." The underlying structure as well as the adjusted Fishbein Attitude Model is used in the thesis.	Evans & King (1999)
WebSCORE	Systematic, scientific-based method to evaluate and optimize web applications. The integrated reference model enables to compare websites according to four basic components: purpose and strategy, content and functionality, navigation and interaction, media design and presentation. As it just considers four components, it can be neglected in the thesis.	Heidmann & Ziegler (2002)
Web scorecard	A small company located in the USA uses an approach called “web scorecard”. It promotes itself by converting “those scattered, inconsistent, tracking tools into a comprehensive, web-based performance management system that can deliver the answers you need to stay on track and improve results” (Web-Scorecard, 2006, paragraph 1). As the last update of its website was made in 2006, it can be assumed that the business is not running anymore so that this approach can be neglected.	Web-Scorecard (2006)
E-Commerce website evaluation method	Helps managers to assess and quantify the performance of a website in comparison to other websites. Thereby, it provides the possibility to judge the performance and identify potential for improvement. As this method needs two weeks to evaluate a website based on 100 criteria, it is not applicable for daily life business and thus, can be neglected.	Van der Merwe & Bekker (2003)
Website Benchmarking tool	The study takes a specific organization (Collegue of Business) and by evaluating different metrics over time, this tool provides the possibility to measure the investigated website on a continuous basis. Becasue it is not universally valid, the tool is not considered in the thesis.	Misic & Johnson (1999)

Appendix 11: Interview questions for the expert interviews (second round)

The following interview questions have been asked in the oral expert interviews conducted with the web application managers of the three selected web applications in the period of 21st of May until 31st of May 2013. Hereby, especially the preselection of the subgoals and metrics has been verified and adapted in order to generate the beta version of the web scoreboard.

1) The categorization, subgoals and metrics of the "web scoreboard" are configured based on a literature review on performance and web measurement. Looking at the web scoreboard do you agree with the chosen subgoals and metrics? Is something missing? Which of those metrics would you eliminate?

2) How do you define a generated "lead"? Do you have the possibility to generate leads or data permissions (e.g. e-mail address) with your web application?

3) How do you define "conversion" in your web application?

4) I would define members with a high Social Network Potential in your web application as…. Do you agree with this definition?

Appendix 12: List of metrics after preselection and before expert interviews

The 17 metrics shown in the table below have been part of the preselected metrics but have been eliminated during the second round of the expert interviews due to several reasons.

Aditional metrics after preselection	
Metrics	**Source**
Information metrics	
Click stream / path analysis	Phippen et al., 2004
Reach	Sterne, 2010
Frequency	Sterne, 2010
Exit pages	Gillin & Schwartzman, 2011
Entry pages	Gillin & Schwartzman, 2011
Alexa Rank	Alexa, 2013
Engagement metrics	
Most visited sites	Gillin & Schwartzman, 2011
Speed	Evans & King, 1999
Audio visual elements	Evans & King, 1999
Likes / Retweets	Visible, 2012
Enjoyment	Van der Heijden, 2003
Network centrality	BVDW, 2013
Conversion metrics	
Privacy/secrurity	Wolfinbarger & Gillin, 2001
General benefit metrics	
Managerial satisfaction	Huizingh, 2002
Ease-of-use	Hung & McQuenn, 2004
Quality of content	Sterne, 2010
General cost metrics	
Cost-per-Action (CPA)	Meier & Zumstein, 2013

Appendix 13: Strategy map of XX

Excluded due to the confidentiality agreement.

Appendix 14: Interview questions for the expert interviews (third round)

The following interview questions have been asked in the oral expert interviews conducted with the web application managers of the three selected web applications in the period of 01st of June until 22nd of July 2013. Hereby, especially the weighting system has been verified.

1) In our first meeting you named … as the major strategic objective of your web application. Expressed in percentages, how would you determine the relevance of the three strategic objectives "information", "engagement" and "conversion" for your web application?

2) I adapted the subgoals and metrics of the web scoreboard according to the input you gave me in our last meeting. Based on this new version, I made suggestions about the weights of the metrics. Let us quickly go through those metrics. Could you please tell me if you agree with the determined weights?

3) I already tested if the data for the selected metrics can be collected for your web application. Based on the found literature as well as some historical data, I made a first draft for a grading system. Do you have some suggestions for this?

4) The period of data collection for the web scoreboard is the month June 2013. I would gather the data of the external perspective for this period with the help of different data sources. Concerning two figures I need your help: What is the number of members in your web application? How many new members do you have in comparison to May 2013? How many users completed the registration process since the existence of the application?

5) Regarding the internal (costs) perspective I need your expert knowledge. Would you please be so kind to provide me information regarding the following six cost positions:

Internal perspective		
Investments score	**Investments subgoals**	**Investment metrics** (amount required one time)
Metric a	Implementation costs to duplicate	Euro spent
Metric b	Marketing spending (launch)	Euro spent
Metric c	Human ressources to duplicate	Full-time equivalent (FTE) (employees)
Maintenance costs score	**Maintenance costs subgoals**	**Maintenance metrics** (amount required recurrent)
Metric a	Technology	Euro spent/month
Metric b	Marketing spending	Euro spent/month
Metric c	Human ressources	Full-time equivalent (FTE) (employees)

Appendix 15: Illustration of the weighting system of the external perspective

Excluded due to the confidentiality agreement.

Appendix 16: Grading intervals of the Website

Since the web scoreboard consists of different types of web applications which are in different developmental stages, certain metrics displayed in absolute numbers need separate customized intervals for each web application. The other metrics, mostly displayed in percentage or an index, can be graded based upon uniform intervals for all web applications. This leads to a mixture between separate and common intervals in the grading system. The following table shows the grading system of the Website.

Excluded due to the confidentiality agreement.

Appendix 17: Grading intervals of the Web Shop

Excluded due to the confidentiality agreement.

Appendix 18: List of free monitoring tools for different online activities

Overview free monitoring tools for online activties		
Type	**Name**	**URL**
Web Analytics	Google Analytics	http://www.google.com/analytics/
Web Analytics	Feedburner	http://www.feedburner.com/
Online Domain Tools	Webscore	http://webscore.online-domain-tools.com/
Social Media Monitoring	Socialmention	http://www.socialmention.com/
Social Media Statistics	Simply Measured	http://simplymeasured.com/free-social-media-tools
Social Search tools	Socialseek	http://socialseek.com/app/
User Experience Software	Userzoom	http://www.userzoom.de/
Facebook Statistics	Minilytics	http://www.minilytics.com/
Twitter Statistics	Twittercounter	http://twittercounter.com/
Overview Measurement tools	MeasurmentCamp	http://measurementcamp.wikidot.com/tools-for-measurement

Appendix 19: KPI performance measurement of the Website (June 2013)

Excluded due to the confidentiality agreement.

Appendix 20: KPI performance measurement of the Online-community (June 2013)

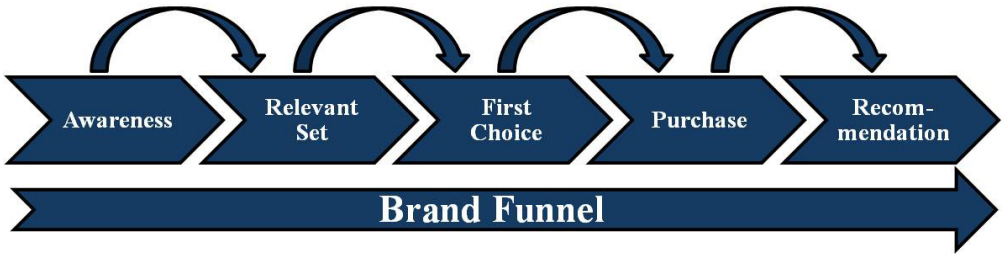
Excluded due to the confidentiality agreement.

Appendix 21: KPI performance measurement of the Web Shop (June 2013)

Excluded due to the confidentiality agreement.

Appendix 22: Online survey based on the brand funnel to analyze the user structure

After a continuous measurement based on the web scoreboard and the set of KPIs is established, it is important to gain insights about the user structure of the web applications. Web application provider often just can guess who their visitors are but they do not know it. However, information regarding the user structure is essential to address the target group and its needs in the right way. This information cannot be detected by web analytics but can be generated via online surveys (Hassler, 2012). Here facts about demographic factors like age, gender, profession can be acquired. Besides, insights about the amount of B2B and B2C customers reached as well as the user intersection of different web applications can be gained. Moreover, it is possible to get answers about the technical equipment of the users or their motivation for visit. Those customer insights help marketers to optimize their online activities. As an underlying structure of the online survey the "brand funnel" is recommended. The brand funnel[9], shown in the figure below, consists of the five steps Awareness, Relevant Set, First Choice, Purchase, Recommendation and displays the customer journey.



Source: Own illustration based on Farris et al. (2010); Li and Bernoff (2008).

Also, XX uses this approach to monitor the customer's purchase decision and to analyze at which steps most of them get lost on their customer journey. It is also a valuable approach in the online environment (Hassler, 2012). By asking the user especially about the brand funnel steps, the results of the online survey would show the percentage of users who proceed on their customer journey with XX products or competitor products. Based on this information, XX can address the identified gaps in the brand funnel with targeted activities in order to keep the users on their way through the funnel (Gillin & Schwartzman, 2011). A draft of the online

[9] This often used instrument is well known under various expressions as AIDA model (Hassler, 2012), Consumer decision making process (Lavidge & Steiner, 1961). Communication Effect Pyramid or Marketing funnel (Li & Bernoff, 2008) and Sales funnel (Farris et al., 2010).

survey in English and German is provided in Appendix 23 and Excluded due *to the confidentiality agreement.*

Appendix 24 below. It is recommended to conduct this online survey in the web applications for which further information about the user structure needs to be gathered. The participants of the online survey should be a simple random sample, i.e. the members of the population all have the equal chance to get selected (Kotler, 2002). The users need to be activated for instance via a newsletter which is sent to the registered e-mail addresses of the respective web application. In the newsletter a link to the online survey should be included. If no data permissions are generated with the investigated web application, a possibility to conduct an online survey is to display a pop-up on the web page which leads the users to the survey questions. An external agency should be included in the research process.

Appendix 23: Draft online survey (English) to generate insights about user structure

Excluded due to the confidentiality agreement.

Appendix 24: Draft online survey (German) to generate insights about user structure

Excluded due to the confidentiality agreement.

13 References appendix

Book.

Farris, P. W., Bendle, N. T., Pfeifer, P. E., & Reibstein, D. J. (2010). *Marketing Metrics: The definitive guide to measuring Marketing performance* (Second Edition ed.). New Jersey, USA: Pearson Education, Inc.

Gillin, P., & Schwartzman, E. (2011). *Social Markting to Business Customer: listen to your B2B market, generate major account leads, and built client relationships*. New Jersey, USA: John Wiley & Sons, Inc.

Hassler, M. (2012). *Web Analytics*. Germany, Heidelberg: mitp

Li, C., & Bernoff, J. (2008). *Groundswell: Winning in a world transformed by social technologies*. Boston, USA: Harvard Business Press.

Electronic Book.

Kotler, P. (2002). Marketing Management, Millenium Edition (pp. 465). Retrieved from VNU-UEB database Retrieved from http://dl.ueb.edu.vn/bitstream/1247/2250/1/Marketing_Management_-_Millenium_Edition.pdf

Journal Article.

Lavidge, R. J., & Steiner, G. A. (1961). A model for predictive measurements of advertising effectiveness. *The Journal of Marketing, 25*(6), 59-62.