

**Development of a Half-Automated Product Database for Search Engine
Optimization**

Masterarbeit

zur Erlangung des akademischen Grades „Master of Science (M.Sc.)“ im Studiengang Wirtschaftswissenschaft der Wirtschaftswissenschaftlichen Fakultät der Leibniz Universität Hannover

vorgelegt von

Name: Neefe



Vorname: Christian



Prüfer: Prof. Dr. Michael H. Breitner

Hannover, den 09. September 2014

Table of content

Table of content	I
List of figures	IV
List of tables	VI
List of abbreviations	VII
1 Introduction	1
1.1 Motivation and relevance	2
1.2 Structure of the paper	3
1.3 Derivation of a research question.....	4
2 Research Design	5
2.1 Identification of relevant literature	5
2.2 Explanation of various analysis tools	7
3 Clarification of abstract terms related to IT-languages and databases	8
3.1 Hypertext Markup Language.....	9
3.2 Cascading Style Sheets	10
3.3 MySQL	11
4 Search Engine Marketing	12
4.1 Search Engine Advertising versus Search Engine Optimization	13
4.2 Importance of the search engine Google	16
4.3 Operation Method of search engines	17
4.4 Elements of the On-Page Optimization	18
4.4.1 Title-Tag.....	19
4.4.2 Meta-Tags.....	20
4.4.3 Uniform resource locator.....	21
4.4.4 Continuous text and keywords.....	22

4.4.5 Enumerations and text highlighting	25
4.4.6 Headlines	25
4.4.7 Links and Pictures.....	26
4.4.8 Unique content and search engine optimized writing.....	27
4.4.9 Google +	28
4.5 Search Engine Spamming	30
4.5.1 Keyword-Stuffing.....	31
4.5.2 Meta-Spam	32
4.5.3 Nepotistic Linking.....	32
4.5.4 Doorway-Pages	33
4.5.5 Bait-and-Switch.....	34
4.5.6 Blog and guestbook spam	34
4.6 Common SEO mistakes.....	35
5 Research Method: Development and Implementation of templates and databases.....	37
5.1 Profile of the company Energieheld GmbH	37
5.1.1 Cooperation, Promotion and Awards	39
5.1.2 Need for Search Engine Optimization.....	41
5.2 An adaption of the classical waterfall model	42
5.3 Requirements.....	45
5.4 Preliminary design.....	47
5.5 Detailed design	51
5.6 Technical Implementation	55
5.7 Web Analytics	59
5.7.1 Selection of relevant Web Analytic Tools.....	60
5.7.2 Selection of relevant Key Performance Indicators	63

6 Research results and discussion	66
7 Recommendations and adaptation	80
8 Limitations	83
9 Conclusion and outlook.....	85
References	88
Appendix	A1

Development of a Half-Automated Product Database for Search Engine Optimization

Christian Neefe

Leibniz University of Hanover

September 09, 2014

Abstract

This thesis deals with the development of a half-automated product database for energy related reconstructions, in particular gas heating systems, with regard to actual search engine optimization standards. Within this paper the most important factors for a good rating in the search engine “Google” will be described with the help of a comprehensive literature review. Afterwards several templates were created using an adaption of the classical waterfall model of a software development process. An analysis of the different templates with the support of various SEO tools pays regard to actual Key Performance Indicators of the search engine Google and also to the legibility and user orientation of the website. The most important elements to realize a positive ranking in search engines nowadays were examined in a practical context. The results of this search engine optimization project can be viewed as a significant part of an holistic online marketing approach.

Keywords: SEM, SEO, search engine optimization, product database, on-page-optimization, online marketing, template

1 Introduction

This master thesis was written during the summer term of 2014 at the institute of information systems research at the Leibniz University of Hanover and served to the achievement of the academic degree of Master of Science (M.Sc.).

1.1 Motivation and relevance

The search for information in the internet became a fixed component in our culture, in 2011 were about 5 billion search requests per day.¹ This extent indicates the strength and relevance of search engines and also a good ranking in the search engine result page (SERP).

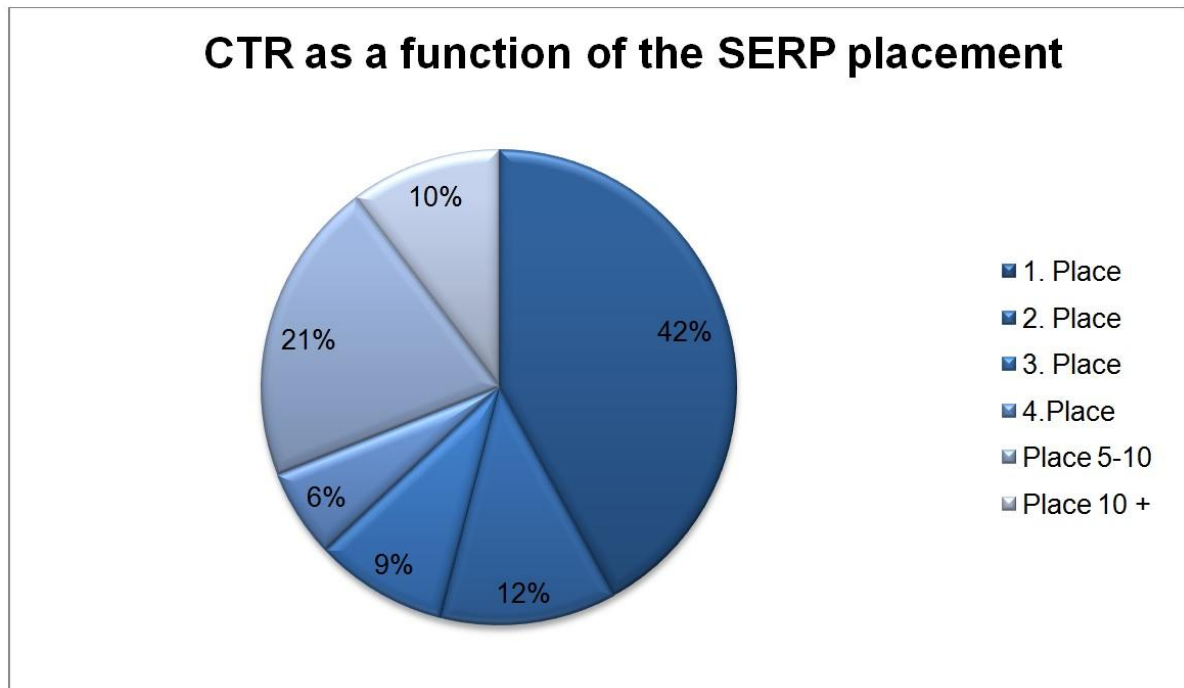


Figure 1: CTR as a function of the SERP placement
Source: Self-made according to Enge et al., 2012, p. 21

Figure one indicates the importance of a good ranking in the search engines. The click through rate (CTR) is a percentage rate which defines the relation between a click and an impression. In this study the result which was on the top of the SERP was clicked by 42 percent of the users, while all search results which are not on the first page achieved a CTR of 10% together. The search behavior is based on two facts. The first one is the recommendation of the search engine to rank a specific page on a high level; the second one is the fast visibility which also provides a high comfort.² Search Engine Optimization (SEO) can be quite powerful and can lead to a difference between hundreds or thousands of engaged and relevant visitors to a

¹ Cp. Enge et al., 2012, p. 1

² Cp. Enge et al., 2012, p. 22

website or nearly no visitors at all.³ These differences have huge impacts for companies who use the internet to deliver important information and offer services or products which also lead to financial transactions.⁴ In a tightened competition between several website-operators, the search engine optimization became essential to achieve a good position in the SERP.⁵ To increase the competitiveness of the website, there are two aspects which are of utmost importance in the course of the optimization: First one are the requirements of the search engine and the second aspect is the consideration of the needs of the user.⁶ Within this thesis the focus lies on the On-Page-Optimization of a company website with the help of a half-automated product database with the aim to shape a website which is characterized by user orientation and search engine friendly characteristics. The half-automated product database has two substantial advantages: On the one hand a consistent layout will be ensured, which is geared to the elements of an On-Page-Optimization. On the other hand, content can be created substantially faster and thus also cheaper. This master thesis rest upon a collaboration with a company, the web pages which are based on the half-automated product database are published on the domain "www.energieheld.de". The advantage for both sides, the company and the student, are evident. On the one side content can be created, published, analyzed and evaluated in a real environment. On the other hand, the company gains more customers and receives a half-automated product database which can be customized for future needs and thus delivers an extensive added value.

1.2 Structure of the paper

At first it is inevitable to explain the structure of this thesis. The first chapter starts with an overview about the importance of the topic, this also includes the derivation of a research question, which will be answered in the course of this paper. The second chapter explains the used research design, in particular the identification of the relevant literature and the explanation of various analysis tools, which will be

³ Cp. David, 2011, p. 8

⁴ Cp. Bischo-pinck/ Ceyn, 2009, p. 3

⁵ Cp. Erlhofer, 2011, p. 21; Cp. Justiz, 2011, p. 9

⁶ Cp. Justiz, 2011, p. 168

used to evaluate the success of the applied half-automated product database in the web. The ensuing chapter is dealing with the clarification of abstract terms; this contains different technical aspects like Hypertext Markup Language, Cascading Style Sheets and also MySQL. This IT knowledge is fundamental for the following parts of the thesis. Chapter four is based on an extensive literature review about search engine marketing. The topic of the search engine optimization will be defined, also containing the importance of the search engine Google and the operation method of search engines. The most important part are the elements of the On-Page-Optimization, these different aspects are the basis of the templates which will be created in chapter five of this thesis with the aim to achieve a good ranking in search engines. Furthermore common mistakes which should be avoided will be shown next to negative search engine spamming techniques. Beside the templates and the technical background of the database, chapter five also deals with the presentation of the company, where these web pages were created. To develop the templates, an adaption of the classical waterfall model of the software development process was applied. This contains the detection of requirements, the design, the implementation, verification and adaptations. The verification is based on part six of the paper, the analysis of the success of the project. Therefore different software solutions which are freely accessible like Google Analytics were used. The following chapters examine the research results, give recommendations and limitations. The thesis will end with a conclusion and an outlook.

1.3 Derivation of a research question

In order to achieve good results with regard to the issue of the search engine optimization, it is indispensable to phrase a research question. This research question will be answered with the support of the literature analysis realized in chapter four and the creation of the half-automated product database which is transferred to templates designed to actual search engine optimized standards. In detail, this requires the identification of relevant products and suitable indicators of heating, the identification of actual elements of the On-Page-Optimization, a navigation to the specific products on the website and writing appropriate content. After the development and implementation, which are geared to an adoption of the waterfall model of

the software development process, the success of this project can be analyzed with the help of various tools which will be described in chapter 2.2. For the evaluation it is indispensable to answer the following research question.

Research question: *“Which factors and criteria are used by Google for the ranking of half-automated product databases in the context of heating and how does the design impact the behavior of the user?”*

2 Research Design

The research design represents the order of the methodical research steps. In this master thesis, different types of methodical instruments found their application. A comprehensive literature research on the basis of the Webster and Watson with the emphasis on search engine optimization is the base of this paper. To analyze these theoretical findings and perceptions in a practical context, an empirical approach was made. With respect to the most important aspects of the On-Page-Optimization, several templates were created with the help of selected programming languages and published on the homepage of the company “Energieheld GmbH”. With the help of different tools, the success of the performed procedures can be observed and contribute to answer the research question of this paper.

2.1 Identification of relevant literature

For an improved understanding of this paper and the empirical section, it is inevitable to declare various abstract terms, in particular SEO and necessary IT-languages. To achieve an overview of the available literature, a reliable method of selection has to be implemented. To increase the relevance and timeliness of this thesis, an already analyzed topic will be reprocessed.⁷ This approach is in context of frequent updates of the algorithms of various search engines essential. Furthermore the accuracy and precision of a paper can be improved by considering exist-

⁷ Cp. Baker, 2000, p. 219

to reach a good ranking in SERPs.²⁶⁷ In addition, the pages just give a small added value for the user, as most of the content is already published in the web. Thus it was the aim of the project to write at least 300 words on each of the 94 web pages and to avoid duplicate content, as these pages would be confronted with problems of relevancy.

Also the pages on section 2-4 exhibit some limitations, especially the analysis method of the ranking of these 12 pages. For the analysis of the ranking for specific keywords, just the best ranking in the SERP of Google was determined, irrespective of the desired URL. An analysis with the exact URL would strengthen the fact, that these pages have an unsatisfactory ranking for the specific keywords, as the exact pages ranked even worse. Thus this methodical mistake has no influences in the evaluation of this analysis. Rather in contrary it strengthens the impression that these pages definitely do not serve as an entrance channel from search engines.

9 Conclusion and outlook

In a time where more and more companies use the internet as a distribution channel, the competitiveness of a company's website is becoming increasingly important. A majority of the internet users utilize search engines as an entrance channel for the research of information, products and services.²⁶⁸ Thus SEO is a quite powerful instrument of the online marketing and can lead to a difference between hundreds or thousands of engaged and relevant visitors to a website or nearly no visitors at all.²⁶⁹ As discovered in chapter 5.1.2 especially the search engine Google has a high relevance for energieheld, as close to 99 per cent of all users who applied a search engine came from the market leader to the company's website. Next to the fulfillment of the requirements of the search engine, it is also important to meet the requirements of the customers. This paper thus tried to answer the question, *"Which factors and criteria are used by Google for the ranking of half-automated product databases in the context of*

²⁶⁷ Cp. David, 2011, pp. 11, 14

²⁶⁸ Cp. Tamblé, 2013, p. 22

²⁶⁹ Cp. David, 2011, p. 8

heating and how does the design impact the behavior of the user?". Therefore a half-automated product database was applied and transferred to templates which were designed to actual search engine optimization standards. First of all it was necessary to clarify abstract terms about IT-languages and the database "MySQL". This IT knowledge is fundamental for the following parts of this thesis. The sequent chapter comprises a detailed literature review about Search Engine Marketing. This implicit the comparison of SEO and SEA, the importance of the search engine Google and the operation method of search engines in general. As the focus of thesis lay on the On-Page-Optimization, the most important elements were constituted and represent the basis for the development of templates with the aim to achieve a good ranking in the search engine Google. Furthermore common mistakes which should be avoided were shown next to negative search engine spamming techniques. In chapter 5 the previous gained information were applied. For the creation of the half-automated product database with the transfer to templates which were designed to SEO standards, an adaption of the classical waterfall model of software development process was applied. In detail this contained the detection of requirements, a preliminary and detailed design, implementation, verification and necessary adaptations. The half-automated product database has two substantial advantages: On the one hand a consistent layout will be ensured, which is geared to the elements of an On-Page-optimization. On the other hand, content can be created substantially faster and thus also cheaper. This project also required the identification of relevant products and suitable indicators of heating, a website navigation to the specific products and writing of appropriate content. Therefore 94 pages were written and published on the domain of energieheld. To detect differences between the success of these pages in context of the SERP and user behavior, different templates were developed and implemented. The success of the applied half-automated product database in the web was evaluated with the help of various SEO tools in chapter 6. The most important findings were that web pages which have not integrated a picture of the product faced serious problems to get indexed in the search engine of Google. Furthermore the strategy to publish pages whose keywords have a low competition and search volume worked out. The product pages on section five received impressive results in the SERP of google, especially with information about prices and the integration

of an image. Nevertheless, the pages on the sections above exhibit better results in the behavior of the user, but did not act as an entrance channel to the website. The following chapters dealt with recommendations respective limitations. Especially a new landing page in the context of SEA and retargeting with the accentuation of the service of energieheld seems meaningful. Also various Off-Pages measures should be conducted. The major problems of this project were the small seasonal demand for gas heating and the sample size of product pages without an image.

All in all, nothing develops and changes so fast like the internet. Also SEO is a process which is characterized by continuous changes through technical developments and changing user behavior.²⁷⁰ In the future, the usage of mobile internet will increase significantly, therefore SEO activities must be geared towards mobile users and hence the presentation of a website also has to guarantee a high functionality on mobile devices.²⁷¹ Furthermore, social signals seem to be an important aspect for search engines and their rankings. The more often a website is shared in social networks, the more it can be assumed by search engines that this content is important and interesting for a large number of users.²⁷² One really important aspect in the SERP are rich snippets. Rich snippets are extracts which appear in the SERP and give additional benefits.²⁷³ The author-tag was an example for rich snippets, but unfortunately has not increased the CTR of a popular web page of energieheld, as approved in chapter 6 of this thesis. With its elimination, other possibilities were searched to change the appearance of the own web pages in the search result page with the aim/aspiration to increase the own CTR. One possibility which was also implemented by energieheld after the observation period of this project are rating stars. With the help of this kind of rich snippets, the attention of the user should be gained. Weinand expects that in the near future rich snippets will receive a high weighting in the assessment criteria of search engines and thus are an important aspect of modern search engine optimization.²⁷⁴

²⁷⁰ Cp. Hübener, 2009, pp. 100, 101

²⁷¹ Cp. Frederick/ Lal, 2009, pp. 280, 284, 285

²⁷² Cp. Smyth, 2012, pp. 17, 29; Cp. Weinand, 2013, p. 394

²⁷³ Cp. Yu, 2011, p. 319

²⁷⁴ Cp. Weinand, 2013, p. 375