



Lappeenranta University of Technology
School of Engineering Science
Industrial Engineering and Management
Business Analytics

Improving content marketing performance measurement

Master's thesis
Laura Manninen

1.11.2019

Examiner: Professor Pasi Luukka, Mariia Kozlova

Instructors: Mariia Kozlova, Kimmo Katajamäki, case company

ABSTRACT

Author: Laura Manninen

Title: Improving content marketing performance measurement

Year: 2019

Location: Helsinki

Master's thesis. Lappeenranta University of Technology, Industrial Engineering and Management.

83 pages, 14 images, 6 tables and 2 appendices

Examiner: Professor Pasi Luukka, Mariia Kozlova

Keywords: content marketing, performance measurement, performance management, quality measurement, key performance indicators

The aim of this thesis is to examine how current practices of content marketing performance measurement are done at the case company and how reporting process could be improved. New reporting dashboard is created as part of practical execution of this study. Literature review and interviews of employees are methods used to understand the topics around performance measurement and to clarify how current reporting process is managed. Observations and phenomenological analysis are done to research how content marketing performance could be improved through this study. Comparison between reporting processes before and after this thesis are done in order to understand the improvement. Previously reports were created manually and after this study, the process is automated. Through new dashboard executed in this thesis, reporting is faster and more fluent, and content can be optimized and targeted to customers more efficiently.

TIIVISTELMÄ

Tekijä: Laura Manninen

Työn nimi: Sisältömarkkinoinnin suorituskyvyn mittaamisen kehittäminen

Vuosi: 2019

Paikka: Helsinki

Diplomityö. Lappeenrannan teknillinen yliopisto, tuotantotalous.

83 sivua, 14 kuvaa, 6 taulukkoa ja 2 liitettä

Tarkastajat: Professori Pari Luukka, Mariia Kozlova

Hakusanat: sisältömarkkinointi, suorituskyvyn mittaaminen, suorituskyvyn hallinta, laadunhallinta, tunnuslukumittarit

Tämän työn tavoite on tutkia, kuinka tämänhetkiset sisältömarkkinoinnin laadun mittaroinnin raportointikäytännöt hallinnoidaan yrityksessä ja kuinka raportointiprosessia voisi kehittää. Uusi raportointialusta luodaan osana tämän työn teknistä toteutusta. Kirjallisuuskatsaus ja työntekijöiden haastattelut selventävät, kuinka raportointi prosessi on hallinnoitu. Havainnointi ja fenomenologinen analyysi tutkivat, kuinka sisältömarkkinoinnin laadunhallintaa voisi kehittää. Ymmärtääkseen kehitystä vertaillaan raportointiprosesseja ennen ja jälkeen tämän tutkimuksen. Aikaisemmin raportit luotiin yrityksessä manuaalisesti, ja tämän tutkimuksen jälkeen prosessi on automatisoitu. Uuden raportointialustan myötä raportointi on nopeampaa ja sujuvampaa sekä sisältöä voidaan optimoida ja kohdentaa asiakkaille tehokkaammin.

ACKNOWLEDGEMENTS

Completing this thesis means finishing my studies at Lappeenranta University of Technology. The process of writing has been constant learning for me. Learning about the research process, new tools and techniques according to digital analytics and performance measurement reporting as well as working at a new company.

I would like to thank my instructor Kimmo Katajamäki for guiding me and giving new ideas. I would like to say thanks to my teammates and mentor Lauri Ajanki at the case company for always answering my questions and helping me out. Also, thank you for my instructor from LUT, Mariia Kozlova, for the guidance of academic writing and supporting this research.

LUT has offered me value and knowledge during my studies. Changing study program from Bachelor of Information Technology to Industrial Management and Business Analytics has offered me diverse skills and expertise. Changing from one program to another was not always easy, yet having different know-how accumulates my skillset. Thank you LUT for giving me the opportunity to study this new master's program about analytics and keep developing new study programs that support us students in our future careers.

Lappeenranta has offered the best times of my life so far. Thank you for my family for always being there to support me with my studies. During my studies I have got to know many amazing people that are still my best friends. I am grateful for each one of you.

Helsinki, 1st of November 2019

Laura Manninen

TABLE OF CONTENTS

1	<i>Introduction</i>	7
1.1	Project background.....	7
1.2	Objective and scope.....	8
1.3	Research questions and methods.....	8
1.4	Structure of the study	10
2	<i>Methodological background</i>	11
2.1	Qualitative research methodology	11
2.2	Data collection	12
2.3	Analysis and interpretation methods	13
3	<i>Literature review</i>	15
3.1	Content marketing	16
3.2	Content quality and performance measurement	19
3.2.1	Performance measurement	19
3.2.2	Performance management process	21
3.2.3	Measuring media content.....	23
3.3	Key performance indicators.....	24
3.4	Digital analytics	27
3.4.1	Process of analytics	29
3.4.2	Digital marketing tools	31
3.4.3	Digital marketing data	32
3.4.4	Example key performance indicators for digital analytics	34
3.4.5	Data visualization	39
4	<i>Project background</i>	41
4.1	Case company	41
4.2	Interviews.....	42
4.2.1	Content marketing reporting current situation at the company	42
4.2.2	Reporting as part of decision-making support.....	44
4.2.3	Collecting data	44
4.2.4	Reporting processes at the company	45
4.2.5	Interviews summary	47
5	<i>Project and results</i>	48
5.1	Process.....	48
5.2	Data collecting and processing.....	49
5.2.1	Advertisement data.....	50
5.2.2	Behavioral data.....	51
5.2.3	Google analytics and custom dimensions.....	52
5.2.4	Processing data	54
5.2.5	Quality of the data	56
5.3	Calculation	57
5.4	Reporting dashboard	60

5.5	Results and analysis	64
5.5.1	Current practices of content marketing performance management.....	65
5.5.2	Reporting dashboard.....	65
5.5.3	Improvement of reporting.....	67
6	<i>Discussion</i>.....	70
6.1	Measuring performance of content marketing.....	70
6.2	Indicators	71
6.3	Future implementations.....	72
7	<i>Conclusions</i>	75
	<i>References</i>.....	76
	<i>Appendices</i>.....	82
	Appendix 1 Data Structure - first draft	82
	Appendix 2 Power BI final data model	83

1 INTRODUCTION

Measuring quality of content is getting more essential for the reason that content marketing is an increasing field of marketing. Digitalization has had an impact on the growth of content marketing and measuring its performance has become feasible. This study focuses on building a reporting system to measure and report performance and quality of content marketing.

This study focuses on technical execution of reporting dashboard and literature research around the topic. Technical execution includes process of building reporting dashboard to automate content marketing reporting at the case company. Technical execution of the project from now on in this study is called CORE, content marketing reporting. The objective of this study is to research how content marketing measurement is managed at the case company currently and how the process can be improved through CORE.

1.1 Project background

This study is executed as a part of authors work in media corporation. Case company has dozens of services in the field of media. The metrics used to measure quality of those services are many. Analytics and reporting are used in measuring quality based on data. In this thesis the objective is to create a reporting dashboard for measuring performance of content marketing.

Company's current reporting and measuring processes of content marketing are not sufficient for the business unit and its customers. Reporting includes manual steps which now in this study are automated, and new reporting dashboard created. This thesis will include technical execution of content marketing reporting, calculations for key performance indicators (KPI) and analysis of development of reporting processes. Indicators are used in measuring the performance of content.

The purpose of new reporting dashboard is to offer value for the case company and

improve reporting processes. New business measurement process could empower, customer reporting gets higher qualification and expedition, manual steps of reporting can decrease and also processing of data and enhancement for usage of common indicators are created.

1.2 Objective and scope

The objective of this thesis is to study how practices of content marketing quality is currently measured at the case company and how performance measurement of content marketing can be improved and reporting automated. The aim is to create a reporting dashboard for measuring the quality and performance of content marketing.

The focus of this study is to research the current reporting processes of content marketing at the case company. This study includes technical execution of building content marketing reporting dashboard, CORE, which is a part of improvement process of measuring performance of content marketing. How CORE improves measurement processes at the case company is observed.

1.3 Research questions and methods

The aim of this research is to find an answer what are the current practices at the case company with respect to content marketing quality management and how measuring quality and performance of the content marketing could be improved. The objective of building the reporting dashboard is that the process of reporting would be automated, and manual steps reduced. The research is done around the development process of content marketing measurement at the company. Impact of new reporting dashboard is examined and observed.

Qualitative research methodology is used in this study and literature research and interviews are methods chosen. Literature review collects research information on how to measure quality of content and how to improve it. Interviewing experts at the company is done in order to collect information about the current status of content marketing reporting

processes at the company.

Observations and phenomenological method are used to analyze the current situation of reporting and how the process could be improved. Comparison is done to examine how the executed practical project improves the performance measurement and management processes. Research questions, their objectives, methods and expected results are shown in table 1.

Table 1 Research questions and methods

Research question	Objective	Methods	Expected results
What are the current practices of content marketing management at the case company?	To gain understanding and knowledge about topics related content marketing and measuring performance. Research how performance of content marketing reporting is done currently and how it could be improved.	Literature review Interviews	Concise overview of literature with clear directions on improving content marketing. Current company practices are revealed, problems and improvement areas are identified.
How performance measurement processes of content marketing can be improved?	Technical execution and analyzing how CORE improves performance measurement process.	Observations, phenomenological analysis and comparison	Analysis of the CORE performance and recommendations for its future development.

1.4 Structure of the study

This study begins with research methods explained following with literature research about content marketing, measuring content quality and performance, key performance indicators and digital analytics. This study contains background information of the company and knowledge of interviewed experts at the company is shared. The process of the practical project and results are discussed and analyzed. Lastly further discussion about future implementations and conclusions are described.

2 METHODOLOGICAL BACKGROUND

Research starts within an idea and a problem or a question to which the answer is searched. Research includes the questions and hypothesis to which the research is based on, justification and results. After the idea of the research and literature review, research strategy and methods are defined. (Järvinen & Järvinen 2000, p. 4-5)

This chapter includes definitions for each method and explanations on the specific research methods chosen in this study. Qualitative research was chosen therefore to understand the current situation and how the measurement processes are improved.

2.1 Qualitative research methodology

Mack et al. (2005, p. 1-2) summarizes that qualitative research is a scientific research which consist of seeking systematically chosen procedures to answer questions and afford findings that were not defined in advance and that are applicable within boundaries of the study. Qualitative research does not usually contain statistical procedures or other quantification. It focuses on understanding the nature of the research problem. Defining the research problem is the most important step in the entire research process. Every case study should begin with a literature review and carefully chosen research questions and objectives. (Baškarada 2014, p. 1-3)

Literature review is collecting knowledge from literature references and documents (Järvinen & Järvinen 2000, p. 163). In this thesis literature about content marketing, measuring performance, indicators and digital analytics is referred. Knowledge around these topics helps building reporting dashboard and understanding the importance of key indicators in quality measurement. Online documentations are used to collect information about different digital analytics tools and data warehousing systems used in the practical project, CORE. Documentations increase understanding those tools and technical processes.

Collecting information using interviews and literature methods supports understanding the phenomenon better. In this thesis it is researched how to improve measuring performance of content marketing at the company and how to develop reporting and measuring processes through literature, interviews and observing. To understand the current situation of reporting at the company, interviews of company's development manager, content marketing experts and data developers were conducted. Development manager was as well interviewed after deploying CORE in order to reflect its impact on reporting process at the case company.

2.2 Data collection

Interviews can be open, half-structured or structured depending on if the questions and answer alternatives are defined beforehand. Structured interviews have questions based on the hypothesis and those have answer alternatives which will be gone through similarly with all the participants. Open interviews are directed by the research themes, but the interview is not as structured, and their purpose is to understand the phenomenon comprehensively. Half-structured interviews conduct on both structured questions and open questions. Qualitative research methods ask mostly open-ended questions that are not formed necessarily on a similar way with each participant (Mack et al. 2005, p. 4). Open interview questions make the interview more flexible and the interviewees are able to answer with their own words. (Järvinen & Järvinen 2000, p. 153)

Järvinen and Järvinen (2000, p. 154) propose a classical definition of interviews as conversations between researcher and interviewee. The authors of this textbook outline that the aim of an interview is to collect right information and learn about the researched phenomenon. They also point out that interviews are effective way of collecting information and raise new aspects around the topic. Flexibility and effectiveness of open interviews were the main reasons to choose open interviews as a method in this study.

Interviews are a critical part of most qualitative studies. According to Northcutt and McCoy (2004, p. 196) there are different types of interviews, such as telephone interviews,

individual interviews, group interviews and therapeutic interviews. Individual interviews were chosen in this study, as the interviewees are experts in different fields and questions varied between them.

Information collecting methods were chosen as interviews and literature review around the research questions in this study. Interviews were open interviews and the interviewees experts at the company. Data developers, content marketing and sales managers and producers were interviewed. Questions for the interviewees were defined differently depending on the role of the interviewee.

Northcutt and McCoy (2004, p. 196) define which are the things interviewer should avoid. Interviewee should not be talking too much or asking just yes or no questions, which provide too little information. Questions that may divert the focus of the interview should be avoided and the focus of the questions should be around the topic. In the beginning of the interview or beforehand the purpose of the research should be explained to the interviewees, in order to get sufficient results from the interviews.

2.3 Analysis and interpretation methods

Different analysis and interpretation methods are used in this study. The purpose of interviews is to answer on research question, how current practices of reporting content marketing are done at the case company. The results of personal interviews are overviewed, and the current situation is defined based on the interviews. Interviews also support understanding better what is now needed to improve reporting process at the case company.

Phenomenological method is used to understand the current situation and also how the situation can be improved through the practical execution of this thesis. Phenomenological research is seeing phenomena and understanding experience (Munhall 2007, p. 146). In this case it is understanding the current situation at the company and observing how the practical execution, CORE, improves the reporting situation.

Literature research is supporting the phenomenological analysis and understanding how measuring and managing performance processes could be improved. Interpretation is done by observing and comparing results how the current situation of reporting was improved through this thesis. An additional interview of development manager is executed in the end of this study to support comparison analysis of CORE's effect on reporting process at the company. Methods and their objectives are described in table 2.

Table 2 Methods and their objectives

Method	Objective
Literature review	Understanding how performance measurement process could be improved and supporting phenomenological analysis.
Phenomenological analysis	Understanding the current situation better and how the situation can be improved through the practical execution of this thesis.
Interviews	Increasing knowledge about the current situation of content marketing reporting at the case company and what is now needed to improve reporting process. Also, to understand CORE's effect on reporting process.
Observation	Analyze current situation of reporting at the company and how CORE improves the reporting situation.
Comparison	Comparison between current reporting process and reporting process after this study is analyzed.

3 LITERATURE REVIEW

Literature review focuses on literature about content marketing, different key performance indicators, measuring quality of content, digital analytics and data visualization. These topics support the research about how to improve content performance and quality management and measurement processes. Articles and literature were found from academic library, LUT Finna and Google Scholar. Through literature increasing knowledge of the author of this thesis supports the practical project, CORE, executed at the company.

Topics were chosen around content marketing and key performance indicators as the reporting dashboard is measuring performance using indicators. Digital analytics processes, tools used, data collecting, and visualization of data are discussed. Chapters of literature review and aim for each chapter are explained in table 3.

Table 3 Chapters of literature review

Literature	Aim
Content marketing	Background about content marketing to understand what is going on in the business field and what are the basic objectives of content marketing.
Content quality and performance measurement <ul style="list-style-type: none"> • Performance measurement • Performance management process • Measuring media content 	To understand performance measurement processes and how performance of media content is measured.
Key performance indicators	Background information and requirement of key performance indicators to gain more understanding of how KPIs are used in measurement processes.
Digital analytics <ul style="list-style-type: none"> • Process of analytics • Digital marketing tools • Digital marketing data • Example key performance indicators for digital analytics • Data visualization 	The process of digital analytics explained and divided into subsections. Process of analytics contains marketing tools used, digital marketing data, key performance indicators for digital analytics and finally data visualization. These subsections explain each part of the process of reporting performance, from tools to data and visualizing it.

3.1 Content marketing

Content marketing is creating content for customers. It is a form of marketing that is focused on publishing and distributing content for a targeted audience. Content marketing

is a marketing strategy of attraction. Content marketing offers educational, helpful, compelling, entertaining and engaging information for customers. (Lieb 2011, p. 1)

The purpose of content marketing is to drive profitable customer interactions for customer's needs and interests. It is crucial that businesses create content that is relevant and of high quality on an ongoing basis. Gaining recognition, trust, credibility, loyalty and authenticity are the pursued benefits of content-based marketing. The aim of content marketing is to create value, help customers and answer questions with the provided information. (Lieb 2011, p. 1-2)

Businesses have been doing content marketing formerly for example by publishing newsletters (Lieb 2011, p. 2). Technology has made content marketing easier and marketing is moving online on a fast pace (Pulizzi & Barret 2009, p. 98). Internet, digital channels and social media have lowered the costs for publishing content to attract clients. Blogs, eBooks, YouTube videos, tweets, search engines and Facebook posts are all example channels where content can be shared. Though publishing online is cheaper than physical publications, it certainly is not free. Delivering content to target audiences requires work, strategy, originality and persistence. (Lieb 2011, p. 2-3)

Key factors for marketers to know before publishing content are as follows: know target audience, define the key theme and message of the content, create an editorial calendar, create user-generated content and be present in different channels (Lieb 2011, p. 12-13). Creating content marketing strategy requires behavioral, essential and targeted plan. Communication must have a purpose and the content should deliver important information for the prospect. Content must be targeted to buyers and content marketing an integral part of the strategy. (Pulizzi & Barret 2009, p. 99)

Sharing entertaining stories through content is a central part of user attraction. Creating feelings, laugh and cry, makes customer share the story forwards (Lieb 2011, p. 19). Sharing relevant stories and valuable information for users is arguably an efficient way to interact with users and possible clients than just advertising. Targeting content for the audience and the specific channels is important. Creating both text, photo and video

content is more attractive than only a text article. Making a scheduled plan when to publish and in which channels will help since then the marketing strategy is more consistent.

Social media marketing is a researched topic nowadays due to increased consumption of it. There are many channels where content can be shared and marketed. Ashley and Tuten (2014, p. 15-17) composed a study using content analysis of the creative strategies present in the social media content shared by a sample of top brands. The study shows that branded social media content can be used to increase brand liking, awareness, customer engagement and promote loyalty, increase consumer communication about the brand and increase potential traffic to brand locations. These actions rely on social networks and may involve activities from users, such as dialogue between business and consumer. Branded social media content can be used to influence consumer's attitudes about the brand and also provide content that consumers can share with their own networks, which increases brand awareness even more.

Interactive content can be shared in social channels where consumers can interact with the brand. The study by Ashley and Tuten (2014, p. 16, 24) shows that consumers choose brands strategically and they will likely discuss in online communications to form positive self-images. Use and gratifications theory (Luo 2002, as cited in Ashley & Tuten 2014, p. 24) suggests that social media participants are willing to desire informativeness and entertainment, from which the latter one seems to be a stronger motivator for engagement with brands.

Content Marketing Producer (2019) at the case company stated that content marketing is a growing market and a major part of company's business. New companies are willing to market their products through this kind of marketing procedures. Content marketing at the case company is offering entertaining and informational content for users. Blog posts, articles and other online material is produced to promote brands and products. Articles purpose is to offer entertaining, educational content and stories to engage with the audience.

3.2 Content quality and performance measurement

Measuring quality is related to measuring performance and key performance indicators are metrics to measure content's quality. Key performance indicators are a group of indicators, which have been selected by management team that has the perspective of particular interests of the organizational unit and its critical and current factors. (Samsonowa 2012, p. 32)

Performance measurement processes are discussed paying regard to content management and measurement. The concept and characteristics of measuring media content are explained in this chapter. Challenges in measuring media content are also discussed.

3.2.1 Performance measurement

Performance measures and indicators are tools utilized to understand, improve and manage organization's activities. Measures allow understanding about how well a project or company is proceed and are the defined goals met. Customer satisfaction, process effectiveness and efficiency are measured. Performance measures define if and where correction of problems and process improvements are necessary. (Franceschini et. al. 2007, p. 110)

Measuring quality is related to measuring performance. The performance measurement literature has a variety of terms to describe metrics that measure goal attainment in organizations. Inter alia, following terms have been used to describe performance measurement: "performance metrics", "performance indicators", "performance criteria", "key result indicators", "performance measures", "strategic measures" and "key success indicators". (Samsonowa 2012, p. 27-28)

Performance measurement provides a structured approach about the strategic business plan, performance and goals. Measurement focus is on what needs to be accomplished and how time resources and energy is obtained at the company. Measurement provides feedback on progress towards company's aims. Measuring performance improves internal

communication between employees among with external communication between organizations and customers. Measurement helps to justify projects and their costs, whether a project has a valuable performance and positive impact on results, therefore measurement is done for supporting decision making processes at the organization. (Franceschini et. al. 2007, p. 111)

The purpose of performance measurement system is not controlling and managing process development. Yet, there are aspects that measurement does not show. The cause and effect of outcomes are not easily settled due to the fact that there are time differences between cause and effect and without collaborating data it is difficult to show that a specific project was the cause of a specific outcome. Poor results are not necessarily cause of poor execution. If the objectives are not reached, obviously something is wrong, but performance indicators do not provide the reason. Instead further investigation behind the reasons are raised. It is also essential to remember that measurement system is only an approximate of the actual system. (Franceschini et. al. 2007, p. 111)

A study shows that only when the reasons behind the chosen metrics is considered, company's efforts to use marketing measurement and the results of outcomes can be understood. Digital analytics is needed to track customers increasing interaction with brands trough digital channels to measure their performance. A few case studies demonstrate that using digital analytics to optimize performance measurement has had a positive impact on sales revenue and improved the efficiency of marketing. The use of digital analytics or performance measurement does not naturally improve performance. The value is rather gained by how company exploits the performance measurement system under certain circumstances. (Järvinen & Karjaluoto 2015, p. 117-118)

Järvinen and Karjaluoto describe that (2015, p. 121) human resources and specific skills are required from implementing performance measurement system. Motivating employees and communicating benefits to them towards using performance management increases positive attitude for using new system. Creating a culture where the use of performance management is beneficial and encourage to use those metrics in managing business and making decisions contributes to effective usage of indicators.

The most common mistakes in performance measurement systems are that there are too many variables measured or the opposite, there are too few. Too much or too little data is used or ignored and used ineffectively. Collecting inconsistent, conflicting and unnecessary data is a common mistake. Focusing on the long-term goals and data collection instead of the short-term measures is crucial. It is necessary to have a balance with measurement time-period, not too often but not too rarely. If measures are done too rarely, potential problems might not be realized until it is too late to take action on it. And measuring too often could result in extravagant costs and unnecessary effort. (Franceschini et. al. 2007, p. 112)

3.2.2 Performance management process

It is challenging to quantify performance precisely among that indicators are needed to measure performance and they are alternatives that provide approximations, not an absolute value. By using only one single indicator general performance cannot be quantified, rather a set of multiple performance indicators are needed. (Samsonowa 2012, p. 30)

Strategy, tactical and operational level goals are defined by management of an organization and performance of those goals is measured. Measuring performance of content is a part of the performance management process. As described in figure 1, the process starts with planning which includes planning of strategy, defining target goals and key performance indicators, which characterize timeframes for the strategy. The measurement phase includes the current status described, data collection and calculation of KPIs. Analysis includes activities that are above measurement actions, evaluating, projecting, interpreting and forecasting from the current situation's perspective. Analyzing the effects of corrective actions and how the goals were achieved. Improvement based on the conclusions drawn from analysis are done. Short-term decisions, such as resource reassignments or budget cuts, or long-term adjustments of organizational goals can be committed. The performance management process cycle can be applied in both long and shorter timeframe when the goal achievement is checked intermediately and actions to improve performance are done for the whole period of time. (Samsonowa 2012, p. 37)

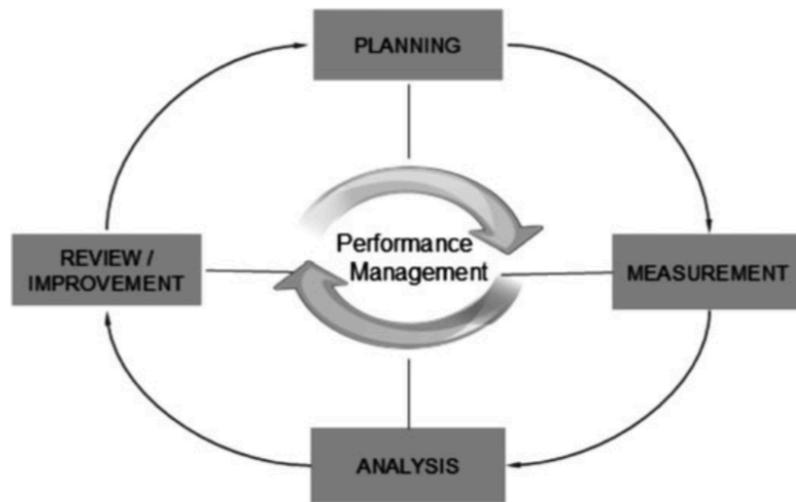


Figure 1 Process of performance management (Samsonowa 2012, p. 36)

Process improvement chain cited in Franceschini (2007, p. 5) is shown in figure 2. Definitions of indicators and measured parameters and data collecting are part of the implementation processes. What is measured and which indicators used should be defined before data collection. After implementations, measurement and analysis based on the results can be done. Decisions are made grounded on the analysis of results and there are three different levels of decisions, individual problem solving, incremental improvements and process reengineering. As shown in figure 2 performance measurement is connected to the automated process cycle. Process outputs are a spur for possible actions and decisions. The focus of process should be results rather than actions.

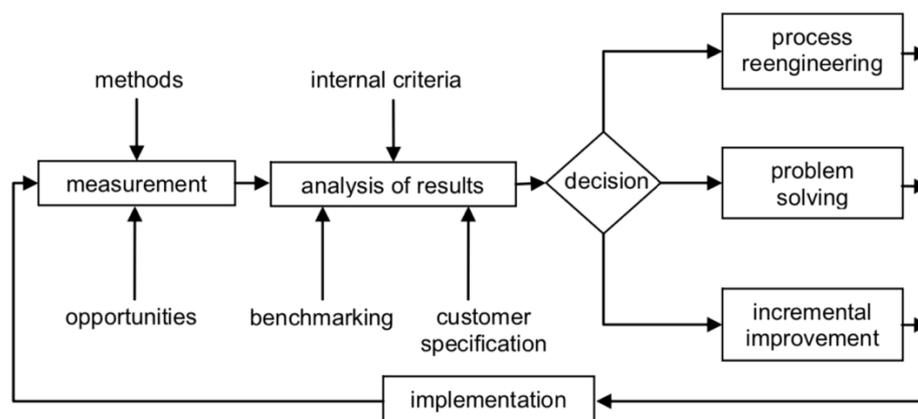


Figure 2 Process improvement chain (Barbarino 2001, as cited in Franceschini 2007, p. 6)

Various studies show that the performance measurement systems are implemented using the process steps such as gathering data, analysis and interpretation, reporting results, taking actions based on them and updating indicators. Gathering reliable data is challenging, although standardization and automation of data collecting are possible. Data is useless without proper interpretation and analysis, thus measuring performance outcomes should have an influence on managerial attitudes and behavior. Improving performance requires taking corrective actions toward existing practices. Therefore, modifying current metrics systems is vital to reflect changes to accomplish defined goals. (Järvinen & Karjoluoto 2015, p. 119-120)

3.2.3 Measuring media content

Robert Picard summarizes the concept of quality as follows: “The concept of quality involves providing value for the money or time expended by consumers to obtain and use a product or service” (Picard 2000, p. 97). Quality is a main factor of developing consumer trust and creating loyalty by making services or products with higher quality and value than those offered by competitors. When applied to journalism defining quality is problematic. The quality of journalism is often related to truth, fairness, completeness but also achieving social, cultural and political goals asserted in democratic societies. (Picard 2000, p. 97)

Sánchez-Taberner (1998, as cited in Picard 2000, p. 98-99) defines 10 characteristics of quality such as, exclusivity of uniqueness, adaption of content to human needs, veracity, company identity, originality, precision, pleasing content, imagination, timeliness, creativity and emotional or temporal proximity, comprehensibility, attractive presentation and physical base. To measure some of these is complex, for example measuring veracity and comprehensibility is problematic.

Measuring quality of content is necessary since the competition and amount of services rises as well as social networking services become increasingly important in consumer’s communication habits (Schivinski et. al. 2016, p. 64). Since most of the articles are nowadays online and digital analytics has been around, measuring certain indicators has

become easier. Impression and clicks on an article's page, average time spend on reading an article or the reading depth of an article, describe how well users engage with the content. Yet, measuring factors such as originality, reliability and truthfulness are more complex (Picard 2000, p. 99).

Schivinski et. al. (2016, p. 64-65) conducted a study about measuring consumer's engagement with brand-related social-media content. The study explains how nature of social media has changed consumers engagement with brands. Consumers interact with brands by writing, reading, commenting, liking and sharing. The study suggests that engagement with brand-related content in social media needs to be measured rather than engagement with the brand. Measurement should cover a large range of brand-related social media actions and also differentiate stages of media engagement from consumer's perspective. Schivinski et. al. (2016, p. 74-75) summarizes, that before managers can confidently utilize branding and marketing using social media, they need to gain an understanding on how consumers interact with brands and how they behave on those social channels. When consumers are engaged in a certain brand, they may even successfully begin to create user-generated content by writing product reviews and posting brand-related content themselves.

3.3 Key performance indicators

“Key performance indicators represent a set of measures focusing on those aspects of organizational performance that are the most critical for the current and future success of the organization” (Parmenter 2007, as cited in Samsonowa 2012, p. 31). Key performance indicators are a set of performance indicators, which have been selected by management teams that can reflect the important metrics that are interesting according to the organizational business unit (Samsonowa 2012, p. 32) and are meant to help finding out why certain things fail as well as clarify why things work out (Jackson 2016, p. 36).

According to Jackson (2016, p. 36), there are two types of KPI, the tactical and the visionary KPI. Visionary KPI emulates what company is trying to achieve. It is meant to

drive change and help the company build its culture. Visionary KPI needs to be defined by leaders as part of the strategy. Tactical KPI for itself depends on the targets and objectives of marketing.

The major difference between “metric” and “measure” is that metric consists of additional information about the referent, and measure is a quantifying value. Basic terms of indicators are shown in figure 3. Samsonowa (2012, p. 29) describes that a metric sets a measure into the context, defines a reference unit and a unit of measure. Performance indicator is an additional metric that is meant to reflect the performance of a business unit. Key performance indicators are selected by management to be representative and critical performance measures in a specific business case. A key performance indicator is a single element of this group of indicators. (Samsonowa 2012, p. 29, 32)

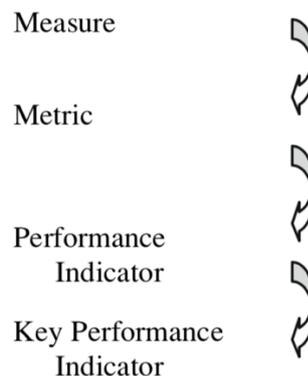


Figure 3 Indicators (Samsonowa 2012, p. 32)

A performance indicator consists of a number and a unit of measure. This number describes magnitude, how much, and the unit is the number of meaning. Indicators are always tied to the representation-goal. Three following types may be related to indicators: effectiveness, efficiency and customer care. Effectiveness describes if process outputs conform to requirements. Efficiency is a process feature that indicates whether the required output of a process is gained at minimum resource cost. Customer care indicates if users or customers appreciate the provided performances. (Franceschini et. al. 2007, p. 110)

Selecting indicators to measure performance is a major issue of reporting performance

measures. To select the indicators needed, the purpose of the measurement or for example a marketing campaign should be known. Based on the goals the indicators are chosen to measure whether goals are reached. Choosing enough but not too many indicators is a key since measuring only the crucial factors is needed (Franceschini et. al. 2007, p. 2). If too many indicators are chosen, the interpretation might be misleading.

Indicators should be accepted and well-understood by managers and employees of the organization. Each indicator has a target for what it is measuring. Indicator has basic requirements, for example it should be representative, simple and easy to interpret. It should be capable to indicate time-trends based on the indicator. Sensitivity to changes outside or within the organization is required for indicator. Data collecting, processing and updating should be easy in order to make calculations based on the indicator. (Franceschini et. al. 2007, p. 8)

Every KPI should have following attributes assigned to it, a metric that has a benchmark, a timescale, a reason to be reported and metric that has an associated enclosing if a problem occurs (Jackson 2016, p. 38). It is important that the indicators have a target what they are measuring and are answering to the questions specified by that target. “If the business question could be answered at scale then it was worth fixing the process” (Jackson 2016, p. 59). When the chosen indicators are relevant for the business, the needed answers should follow.

Rubinson and Pfeiffer (2005, p. 187, 189) conducted a study of key performance indicators for brand equity management. The study describes specific key performance indicators measures and how to reasonably set aims for each measure in order to track and manage the success of marketer’s brands. Without properly constructed measures, the strength of a brand can be easily misled and without targets a measurement system exists, but it is not a brand management system. Attributes that are correlated to the loyalty measure should be turned into KPIs.

The study of Rubinson and Pfeiffer (2005, p. 195) summarizes key steps for implementing KPIs as follows: organizing a marketing strategy and planning teams to coordinate the

selection of KPIs and setting targets, choosing which KPIs would be placed on the dashboards in each division, setting consistent business goals for each KPI, presenting KPIs so that they are linked with other metrics at the organization and implementing a tracking of KPIs to report on the effects of new marketing activities.

3.4 Digital analytics

Digital analytics procedures consist of analytics tools, collecting data and visualizing data. This chapter focuses on digital analytics and explains its definition and processes, examples of analytics tools are explained, digital marketing data is discussed, and visualization practices are explained.

Avinash Kaushik (2009, as cited in Cutroni 2010, p. 1) defines digital analytics as: “The analysis of qualitative and quantitative data from your website and the competition, to drive a continual improvement of the online experience that your customers, and potential customers have, which translates into your desired outcomes (online and offline)”. Cutroni (2010, p. 1) continues by defining three major tasks every business must handle when conducting digital analytics: measuring qualitative and quantitative data, improving continuously the website and aligning measurement strategy together with business strategy.

Jackson tells a story in publication *Cult of Analytics, Data Analytics for Marketing* (Jackson 2016, p. 1-4) about his work as a web developer for a media company. He had made a mistake when developing a login system. A voucher field was missing on a web page and without analytics they did not know about the problem until a colleague called about incorrect online subscription mails. This led into researching the problem more and founding important findings around the business. Additional marketing locally seemed to be wasted effort according to data, which then could be reduced, and website visitors could be served better.

Jackson (2016, p. 4) summarizes that everything was related to the business which made it easier to explain the situation to the management team, as speaking the same language is

an important skill to master. Importance of digital analytics can be realized in such cases, when the unnoticed problems can be corrected, and performance made better with the help of data. Jackson continues explaining the importance of communication skills and understanding each other. Speaking the same language between developers and management might be a challenge, since everyone is doing their own tasks and might not know about the terms spoken in another team.

Järvinen and Karjaluo (2015, p. 121-122) conducted a case study where they researched how industrial companies use digital analytics to measure digital marketing. The case companies in that study differed in terms of satisfaction towards the use of digital analytics. Greatest benefit seemed to be the ability to track how many users visit company's website and how much traffic marketing activities attract. Marketers were able to measure financial outcomes and be aware of the related effectiveness of various digital marketing channels and actions that attract visitors to their site. Analytics makes it possible to measure what kind of content attracts potential customers most to interact on their website and to observe which actions customers take at the web page. User activity on a website indicates their interest in company's services and offerings.

The goals of digital marketing are primarily aimed at increasing sales. Metrics to measure revenue are in different stages, traffic generation to the website, user interaction and behavior on the website. Revenue and profits are gained through online sales leads. By measuring all these different stages of customer's path to purchase decision can be understood better and digital marketing improved. The second important goal of digital marketing is to enhance customer relationships by providing customer service through different channels. Measuring quality of customer service using analytics is more complex, where customer feedback is the only indicator. The third goal of digital marketing is to improve brand awareness, which is more complex to measure than sales. (Järvinen & Karjaluo 2015, p. 122)

At the case company reports and analytics are used in analyzing results of marketing campaigns (Content Marketing Producer 2019). The benefits of analytics are that it is easier to measure what kind of content performs better and which websites are attracting

different types of users. Segmentation of ads or articles is more efficient when information about audience is known on a certain page.

3.4.1 Process of analytics

What matters most about data is not data itself but rather, knowledge and information it contains (Sebei et. al. 2018, p. 2). Cutroni (2010, p. 2) explains that data analysis must drive a continuous improvement process. It is crucial to take action based on data. The purpose of digital analytics is to keep improving. Such as in Jackson's story (Jackson 2016, p. 1-4), the problem could have been noticed earlier and improvement done if there was analytics used earlier on. The process of digital analytics is about measuring, analyzing and changing as described in figure 4.

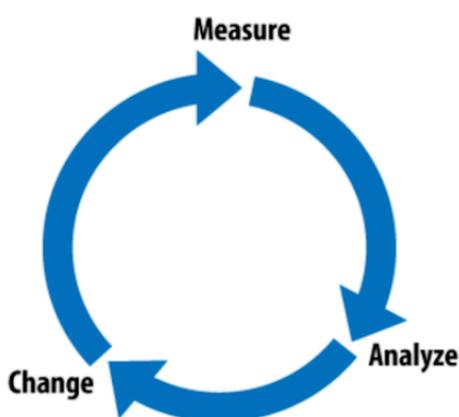


Figure 4 The process of digital analytics (Cutroni 2010, p. 3)

There are certain steps Sebei et. al. (2018, p. 11) study about big data and analytics process suggests for processing data. The first step is to collect and store data from different sources and then prepare and clean data for further usage. Data integration, aggregation and representation present data in suitable format for the analysis. Processing queries, data modeling and analysis are procedures of data mining processes and the aim of those is to analyze data through analytics techniques. Finally, data interpretation consists of visualizing data, understanding the analysis and making decisions out of it.

Most of the times data analysis shows that there is a problem, but it doesn't tell how to fix

it. Creating different solutions and testing them displays the potential solution to the problem. Measuring those tests real time generates the best results. The goal is no longer to measure traffic of an online business but to increase business outcomes, how well the website performs in business terms. (Cutroni 2010, p. 3)

Another model to describe the process of digital analytics is Jackson's (2016, p. 24) REAN model. REAN comes from words *Reach*, *Engage*, *Activate* and *Nurture*. Jackson explains that REAN is a framework which helps to visualize what is measured and why for planning purposes. All key performance indicators fall into at least one of the four dimensions of REAN. Services and products need to *reach* an audience of customers, to raise attention for the brand or a service. Sales teams and advertisers need to *engage* with the prospects, for example by talking, teaching, helping or persuading their product or service. Typically, multiple channels and activities are needed in order to engage audience. *Activate* is to take the first step, such as buy the first product. Activities are needed for prospects to take and actions are wished from them to complete. Finally, once a customer has completed their first action *nurture* is needed. Nurture is the activities needed for customers to come back again, consume more and build a relationship with them. (Jackson 2016, p. 25-26)

Jackson (2016, p. 26) continues with explaining how to apply REAN framework into measurement. Reach can be measures or methods used to attract traffic to website and measure how customers found that website or product. Engage measures can be such as time spend interacting on a website or click depth of the site. Effectiveness of visitors can be a measure of actions on a website, actions that were wanted from customers. Nurture is meant for measure and encourage visitors to come back to consume content of website.

Indicators attached to this framework could be, that reach is the traffic source, where visitors have found the site at the first place. Engage could be measured with reading time and a scroll depth of certain web page. Activation effectiveness could be measured with clicks or impressions and nurture could be measured with metric of returning visitors.

Comparing both Cutroni's (2010, p. 2-3) and Jackson's (2016, p. 24-16) models of digital

analytics processes there are similarities and differences. Cutroni's models of analytics process is simplified model. Before measuring and analyzing specifications what is needed to measure should be done. In order to make analysis, objectives and metrics should be defined. Change can happen toward the goals after analysis. Jackson's model for its part describes the whole cycle of the process and offers examples of actions and indicators to use.

3.4.2 Digital marketing tools

Digital marketing tools used in this thesis were chosen according to the tools used earlier at the case company. These are digital analytics tools that can be used in performance measurement and were as well used in CORE.

Google Analytics is a tool to measure a website and it was launched in 2005. It tracks and reports website traffic and it is a part of Google Marketing Platform. Google Analytics tracks various website metrics, such as visits, pageviews, unique visitors, bounce rate. But even more necessary, the tool can track business outcomes, goals, which are called as conversions. Email marketing, display advertising, social media and other types of ads can be tracked. (Cutroni 2010, p. 1, 4)

Another example of digital analytics tool is Chartbeat, which is a JavaScript based, real-time digital analytics product, which allows to monitor traffic and engagement in real time. Chartbeat can be installed to web page by using a small snippet of JavaScript code. The application allows user to install integrations to certain applications, such as communication tool Slack and a web component framework AMP. Chartbeat uses different metrics and filters to analyze webpage. Example filters are referrer domain, author of a page, section of a visited site and new users that visit the domain for the first time in 30 days. (Chartbeat Documentation A 2019, Chartbeat Documentation B 2019)

Another tool used with digital marketing campaigns is AppNexus, which is a cloud-based software platform for digital advertising. The products of AppNexus enable and optimize programmatic digital advertising and use machine learning tools to distribute better

advertising results. It is the world's largest independent marketplace for digital advertising and enterprise technology for sellers and buyers to help maximize revenue and improve campaign performance. (AppNexus 2019)

Example tools of data visualizations are Google Data Studio and Microsoft Power BI. Those two are also used at the case company and Power BI was chosen to be used at the content marketing report executed in this study. Google data studio is a visualization and reporting tool which is part of the Google Marketing Platform. With the help of Data Studio beautiful, informative and customizable reports and dashboards can be created and shared. Power BI Desktop in turn is a visualization tool offered by Microsoft. The tool allows to connect multiple sources of data, build visuals and share those as reports within organization. Common uses for Power BI are connecting to data, transforming and cleaning data to create a data model, create visuals and reports that are collections of different visuals and share reports with others. (Google Marketing Platform Documentation 2019, Microsoft Power BI documentation A 2019)

JavaScript code snippets, so called tags, are used to collect visitor's data in order to analyze it. Tags are integrated into all web and mobile sites to receive information data about user activity. Three different tags are used, counter, conversion and remarketing tags. Counter tags are tracking pixels that count visits of users, conversion tags determine the number of conversions on a site and remarketing tags identifies returning visitors on a website and addresses their interest in offered products usually through search engines. Tag management systems are used to integrate, edit and manage tags on websites. Flexibility, simplified workflow, fast reaction to changes and adaption of new campaigns in real-time are benefits of using tag management systems. (Digital Guide 2019)

3.4.3 Digital marketing data

There is an ongoing drive of development for digital marketing. Marketers recognize the need for developing in this field and a gap in skills of assessing marketing actions in digital marketing can be seen. Marketers need training and understanding the use of key performance indicators in this environment. Analysis is more effective when those

measures are integrated with traditional measures for marketing. Combining digital marketing indicators and traditional marketing measures creates wider understanding on the performance of a certain marketing campaign. Depending on which platforms are used in marketing, indicators to measure performance can be chosen. (Saura et al. 2017, p. 11)

Quantitative data describes what happens on web page while qualitative data describes why it happens. Collecting both of them is crucial to understand the reasons behind metrics. Qualitative data comes from sources as interviews, surveys and usability tests. Asking questions from website visitors lead to understanding why and what they are searching for from the website. To get a better understanding on users' behavior is good to combine results from different data collection tools. The only way to know why someone converted better on a page than on another is to add qualitative data. Both user surveys and heuristic analysis of the pages require qualitative data. (Cutroni 2010, p. 2, Jackson 2016, p. 47)

Behavioral data for its part refers to information produced as a result of actions. Behavioral data is not static data, rather it changes faster. Static data can refer to person's slowly changing characteristics, such as education and income. Behavioral data offers value that static data cannot provide. It contributes to target advertising and risk assessment. In other words, targeting audience means that there is no point in displaying car advertisement to someone who has no interest or likelihood of buying that specific product. (Greenstein 2015, p. 88)

There are many ways to measure user's behavior and optimize websites. One example of this is A/B Testing that assigns to two versions of a web page or an element of a page, such as image or a heading. These two different versions of a web page are published, and user's behavior is measured. The aim of A/B testing is to indicate site's effectiveness against performance indicators including conversions, revenue per visit and click through rates. A/B testing helps to measure which one of the two different versions of a site performs better and why. (Saura et al. 2017, p. 8)

Saura et al. (2017, p. 8) explains about different rating systems that can be used to classify

the type of users according to quality or merit or amount. Surveys and forms are used to apply the number of conversions or goals in a web page or campaign. Surveys are tools that allow users to send information to a web page. At the case company surveys are used to collect user information about users of websites. Data Developer B (2019) explained that information, such as gender of a user, is collected by surveys. After collecting information through surveys machine learning can be used in describing user's information based on behavior of other users. Machine learning algorithms help with targeting and optimizing advertising better.

3.4.4 Example key performance indicators for digital analytics

Numbers are important for any business when looking at how the business is doing on a high level, even the basic indicators, such as impressions and pageviews. Ideally KPIs measure per business unit how that specific unit is performing. There are three types of metrics in digital analytics, ratios, counts and KPIs. Count is the basic unit of measure, for example number of visits. Example from a ratio in this context would be pageviews per visit. KPI can be either a ratio or a count infused with business strategy. (Jackson 2016, p. 37, 162)

The Web Analytics Association defined unique visitor, pageviews and visit as the three big counts in measuring performance. This is due to that nearly all ratios and KPIs have one of these metrics included. Pageviews can be also replacement of clicks or events. The main point is that there are people, *unique visitors*, doing things, *pageview*, in a time frame, *session or visit*, and those are the metrics to measure the performance of content. One example of a ratio can be pageviews per visit. It could be also a KPI that the business logic would apply. (Jackson 2019, p. 50)

Type of users, type of sources, keywords, keyword ranking, conversion rate and goals conversion rate are known performance indicators in digital marketing. Organic and paid search can be used as indicators as well (Järvinen & Karjaluoto 2015, p. 123). According to Google Analytics Documentation (C 2019) every referral to a web page has a source, or also known as origin and a medium. Source includes the traffic source where the user has

originally come to the site. Source can be for example a search engine, the name of a site or direct. Direct source means that user has typed the URL directly into their browser or bookmarked that site. Medium can be “organic”, cost per click (CPC), referral, email and “none”. Organic means is unpaid search and direct traffic has a medium of none. Referral is the name of AdWords campaign or a custom campaign and email refers to an email campaign. Type of users are new visitors and returning visitors. Returning visitors visit the website more than once. New visitors can be also referred as unique pageviews and pageviews as returning visitors or visitors in total. (Saura et al. 2017, p. 9)

Indicators keywords and keywords ranking are based on keywords or also called as “tags” on a web page. Keywords in web content enables users to find sites through search engines. “Keywords are words or phrases that are used to match your ads with the terms people are searching for” (Google Analytics Documentation 2019 D). A non-branded keyword is a keyword that does not contain brand name of website as a target. Ranking for non-branded keywords allows web page gain new visitors who are not familiar with the brand. Keyword rank is an estimate of website’s position for specified search terms for search engines’ result pages. The lower the rank is the easier the website can be found as results on search engines for certain keyword. (Saura et al. 2017, p. 9)

Defining target goals is an important component of digital analytics measurement strategy. Having properly defined goals allows analytics to provide critical information, such as the conversion rate and the number of conversions for the website. Conversions are one indicator measured in digital and marketing analytics. “Goals measure how well your site or app fulfills your target objectives. A goal represents a completed activity, called a conversion, that contributes to the success of your business” (Google Analytics Documentation E 2019).

Conversion, a target goal, could be for example a purchase or click on an ad. Conversions depend on the marketing objective defined. It is the defined goal or objective of the campaign (Saura et al. 2017, p. 7, Google Analytics Documentation E 2019). Conversion rate is the average number of conversions per click. Conversion rates are calculated by number of conversions divided by the number of clicks or actions on ad. Goals conversion

rate represents a completed activity. Conversion rate could for example represent the purchase rate of a web store or a percentage of clicks on and displayed ad. (Saura et al. 2017, p. 9)

Increasing sales is often the goal of digital marketing. Revenue and profits can be measured using indicators such as sales revenue or profits through sales leads. Website behavior of users can be measured for example by number of sales leads, sales leads growth, product information sheet downloads or video views and sales lead per traffic source. Other metrics can be for example costs per traffic source, average costs aroused per sales lead and percentage of sales leads that lead to transaction. (Järvinen & Karjaluoto 2015, p. 123)

Jackson (2019, p. 82) summarizes that different KPIs are needed to measure performance and that instead of having tons of things to look at it is better to have three or four metrics. Examples of KPIs according to Jackson are visitor volume ratio (VVR), cost per visit (CPV), cost per engaged visit (CPEV) and content visit ratio (CVR). Each of these consist of costs or another metric to measure efficiency of a page. Cost per visit is different to cost per click, since it is a cost of a visitor arriving at website and cost per click is cost for someone clicking a paid link or banner, Jackson explains.

There is a difference between pageviews, sessions and users, which all can be included as an indicator to measure performance. A page view is defined as a tracked view of a web page. It can also be displayed as traffic on a certain page (Saura et al. 2017, p. 7). Google Analytics Documentation (A 2019) describes that there are four levels of scope, product, hit, session and user. Scope is a characteristic of each dimension and product the value applied to the product that has been set and hit is a value applied to the single hit for which it has been set. Session is a value applied to all hits in a single session and user value is applied to all hits in sessions, future and current, until value changes or custom dimension is inactivated. These levels are visualized in figure 5.

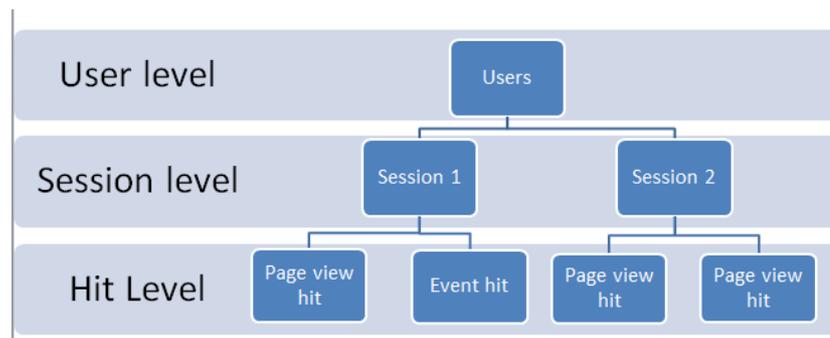


Figure 5 Levels visualized (Digishuffle blogs 2017)

When the page is reloaded a new view is tracked. A unique pageview presents the number of sessions where the page was viewed at least ones. Unique pageviews tracks users viewing the same page during same session. Both users and sessions are measured in Google Analytics. Sessions indicate the number of individual sessions by all users at the site. This is visualized in figure 6 where there are three sessions tracked and each session has three hits, H1, H2 and H3. If user is inactive for 30 minutes, a new session will be tracked. The session by one user during a specific timeframe is considered to be an additional user and an additional session. Future sessions from the same user within the selected dates are counted as additional sessions, but not as additional users. (Google Analytics Documentation B 2019)

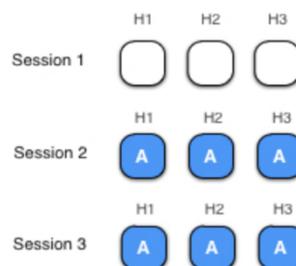


Figure 6 Sessions and hits visualized (Google Analytics Documentation A 2019)

Apart from sessions and hits, there is also difference between users and clicks. Click column in Google Ads reports represents how many times an ad was clicked by user. A user might click the ad multiple times. Within the same session Google Ads tracks multiple clicks from a specific user while Analytics recognizes a single user. Impression is a different indicator than pageviews and clicks. Impressions are an instance of an organic-

search listing or on an image displayed on a web page or sponsored ad (Saura et al. 2017, p. 7). Impressions are counted each time an ad is shown on a search result page or other website. (Google Analytics Documentation B 2019, Google Ads Documentation 2019)

Examples of key performance indicators are collected in table 4. Indicators are listed and their purpose in measurement summarized.

Table 4 Key performance indicators and examples

Key performance indicator	What is measured?
Pageviews, unique pageviews	Amount of traffic on a page, tracks when a site is viewed.
Impressions	Search result listings or an ad is shown.
Clicks, events	Interactions of users on a page.
Sessions, users	User creates a session, and a user can create multiple pageviews. Users or sessions are tracked on a certain page.
Traffic source	Source of traffic, where users have come to the site.
Search, paid or organic	Websites search listings and results.
Keywords	Keywords that users use when searching the website.
Conversion, conversion rate	Goals of marketing, usually a purchase or traffic on a page.
Sales lead per action	Profit of a marketing campaign is measured.
Cost per visit, cost per traffic source	Cost ratio per traffic source, can be revenue of a search marketing campaign is measured.
Cost per click, cost per engage visit	Cost ratio for visits or clicks, revenue of an ad campaign is measured.
Pageviews per visit	Ratio of user's pageviews in a visit.

3.4.5 Data visualization

Data graphics display visually measured quantities by means of combined use of points, numbers, lines, words, symbols, colors and shading. Graphics are instruments for reasoning quantitative information at their best. The most effective way to summarize and describe large sets of numbers is to look at a picture of those. Well-visualized graphics are usually the simplest and the most powerful ways of showing the meaning of data. (Tufte 2013, p. 9)

Tufte (2013, p. 13) explains that statistical graphics consist of complex ideas communicated with clarity, efficiency and precision. Graph should show the data, present many numbers in a small space, encourage the eye to compare pieces of data, reveal details and introduce the viewer to think about the substance rather than design or the technology behind the production. If viewer's attention goes to the design or weak visualizing techniques, the graph should be modified to draw attention to the core idea of visualized graph, the analysis behind the data.

Clear shared information through graphs does not necessary mean that data is simple or less complex. Tufte (1990, p. 51) explains that simpleness of data and design will not straight increase the reading clarity of visualized data. Quantity of detail is a different issue from the difficulty of reading. The less complex is the line the less interesting is the reading, simpleness is not a guide to clarity.

The tools of visualizing data are many. Choosing the right fit for a certain purpose is meaningful. When choosing the tools, it matters how large data sets are used and what are the aims of visualizing. Different data visualization tools have different limitations considering how much data it can handle. Interpreting a graph is in the eye of the viewer but it matters how the design of the graph is built. Choices of colors, typography and size of lines draw attention to the data that is shown. According to Tufte (2013, p. 81), human eyes are extremely sensitive with color variations. It is meaningful to choose colors in order to make certain parts of the visualization stand out or blend in.

Based on the purpose of a graph it is chosen which one of the graphs is used. Basic graph types can be such as bar plots, tables, maps and columns, gauge charts, funnel and line charts. Examples of charts are visualized in figure 7. Different kinds of matrixes, scatterplots and heatmaps can be also a way to visualize. Depending on what are the dimensions to visualize and which metrics the graph should present. (Microsoft Power BI Documentation B 2019)



Figure 7 Different chart types (Microsoft Power BI Documentation B 2019)

Kelleher and Wagener (2011, p. 824-826) explain guidelines for effective data visualization. Decision to use a certain graph depends on the dataset. According to the study those authors a basic choice selecting a plot is between displaying details or patterns. This choice requires a decision which type of a plot is on demand and which objects are used to encode values within the plot. When individual values are critical, bar or line graphs should be used, as position and length are easily quantitatively detected. For simultaneously visualizing multiple variables a dot plot would be an alternative to a bar chart. To visualize a difference between time steps, a change of time-series, can be enhanced by vertical axis transformations. Using a color scheme that matches the type of data as well supports the purpose of a graph.

4 PROJECT BACKGROUND

Background information of the case company and its services are presented though the company wishes to stay incognito. Findings of the interviews of experts at the case company are shown in this chapter. There is discussion around current situation and reporting processes of content marketing and measuring performance.

Interviews are divided into chapters based on the topic of the issue. Those topics are the current situation of content marketing reporting, reporting as a part of decision-making at the company, collecting data and reporting processes.

4.1 Case company

Case company is a media corporation that focuses on digital services and publishing. Company offers dozens of medias services, such as print media, digital services, trainings and events, business literature publications and distribution services. Company's products are both for companies and consumers. Services and products provide information around topics related to lifestyle, business and career.

Content marketing at the case company is offering informational and entertaining articles for users. Articles are created by case company's customer brands and companies, which are willing to create online material, such as blog posts and articles to promote their brand and stimulate interest towards their products and services. Clients who use content marketing services offered by case company are for example clothing brands, banks, telecommunications and software companies, health services, electricity and construction companies.

Author of this thesis is working in the web services department at the company as a developer trainee. This thesis is executed as a part of authors work at the company.

4.2 Interviews

The interviews were open interviews and they were conducted as personal interviews by the author of this thesis. Each interview was recorded in order the author of this thesis to interpret the answers of those. Questions for the interviewees were defined differently depending on the role of the interviewee. Interviewees were also given a chance to comment the topic freely.

Five experts were interviewed in this study. Three of them were content marketing experts at the case company and they were interviewed about the current situation of reporting content marketing and reporting as a part of decision-making support. Two of the experts were data developers and they explained about data collection and reporting processes at the company. The goals of interviews were to gain knowledge about current reporting processes and how reports are supporting decision-making at the company.

4.2.1 Content marketing reporting current situation at the company

Content marketing is an ascending trend at the moment. It is a crucial part of sales at the company since new customers and content are rising up continuously. The services that are being sold at the company are printed media, digital services, content marketing and event services. There are four content marketing teams at the company and all of them are reporting performance differently. The aim of building new reporting dashboard is to unify that reporting at the company and unify reports between different teams. (Content Marketing Producer 2019, Development Manager at Media Sales 2019)

Content Marketing Producer (2019) mentioned that every article is reported, and measuring content is an important part of marketing. Product Leader (2019) at the case company explained that content marketing reporting is used with reporting results of advertisement and campaigns to customers and also measuring results internally at the company. The reports are sent to customer companies and with presented in cooperation with the salesperson of specific campaign. Through reports customers can see how the campaign is performing.

The current reporting process of content marketing is done manually by collecting the metrics from Google Analytics and copying those into Power Point templates. Screenshots are used and power point reports are shown to customers. Main collected metrics have been pageviews, unique visitors, reading time of an article and traffic sources, network domains and user profiles. Reporting at the moment is not as comprehensive as it could be. (Development Manager at Media Sales 2019, Content Marketing Producer 2019)

This new reporting system that is built is meant to simplify and speed up the reporting process. Automated reporting reduces mistakes and the metrics are more precise since the measuring is not done manually. The workload that is done now with reporting will reduce and working hours are free for other tasks and actually analyzing the results. The goal is that the salesperson or person who needs the report would only need to export the dashboard and send it to customers. That workflow is also a benefit of content marketing report executed since the salesperson can export the report themselves and there is no other workload included in the process. (Development Manager at Media Sales 2019, Product Leader 2019)

Product leader (2019) at the company estimated that current manual reporting is taking approximately 100 full working days. There is a need for automated reporting system for the reporting situation has been manual for many years. Content Marketing Producer (2019) talked about the benefits of the automated reporting dashboard, the working time released could be focused on analysis of the performance of campaigns and improving services rather than reporting.

Another objective of this new reporting dashboard is to report metrics that were not reported before, for example reading depth, device distribution and more detailed information about users. Development manager at media sales stated that reports could be also more in-line with each other and more easily comparable. More comprehensive information is included to new report dashboard compared to current manual reports.

4.2.2 Reporting as part of decision-making support

Reports are a part of decision-making support at the company. Through report it is seen which articles are successful and what needs to be improved. The indicators are a key issue when pricing for example advertisement or blogs that include content marketing. The average reading time of an article at a certain site can have an effect on the prices on that site. Reporting and measuring content's quality support the sales process. Profile information about users is collected as well and those help to understand user segments of articles. (Product Leader 2019, Content Marketing Producer 2019)

Content marketing producer (2019) explained about reporting content marketing at the case company. Indicators at reports can show descending and ascending trends. It shows what topics are interesting at the moment and what are the user segments or target audiences to consume specific topics or sites. It is a way to show customers what is successful and how much an article is viewed and how customer's brand awareness is raised.

Reporting content marketing is challenging as the campaign and customer relationships are more long-term based. Interview of Content Marketing Producer (2019) brought up the need of paying attention to long-term benefits as well. Daily updates of reporting dashboards are sufficient, but real-time reporting is not crucial yet at the case company.

4.2.3 Collecting data

Data developer A (2019) at the case company brought up the issue of how data is collected and how custom dimensions are used at the company to collect data into Google Analytics. Custom dimensions are used to collect data that Google Analytics does not track automatically. Article's metadata is not collected automatically, so those are additional dimensions (custom dimensions) tracked to Google Analytics. In the reporting dashboards there is needed information such as what is the topic of the article and publishing date. As a media corporation there is a lot of data to be used which Google Analytics does not track automatically, therefore custom dimensions are an important part of reporting.

The custom dimensions are collected from Google Analytics using BigQuery, which is a data warehousing platform offered by Google Cloud Platform. Data from Google Analytics is collected automatically from different property views into BigQuery. There are challenges with BigQuery and Google Analytics connection, such as the use of custom dimensions and several accounts. At the case company different sites have separate accounts in Google Analytics and the data is collected with different dimensions and those have different index numbers. In BigQuery the index number of dimensions need to be defined and it varies between different services and dimensions. When there is more than one custom dimension needed the SQL query at BigQuery will need to be joined with other additional queries. Every service has their own metrics and dimensions and those go into different data sets which are aggregated in BigQuery with unique query for every media service. (Data Developer A 2019)

Data Developer A (2019) explained about the challenges with BigQuery, it has a delay with collecting data. It is not capable to collect data in real-time, the data flows in every 15 minutes or even in smaller delays. Google's real-time API could be used to get data more real-time, and this feature could be used in the future at the company.

There are challenges with aggregating data from different systems, as running queries cost, it is important to aggregate data to reduce the costs. Planning has to be done when aggregating data, for example which dimensions aggregate in which order. Aggregation is used for example when data is aggregated from session level to user level. Big amount of data is challenging as the tools have limitations concerning the amount of data (Data Developer B 2019). Data needs to be formatted so that the visualization tools can produce it and that is one reason for data preprocessing. (Data Developer A 2019)

4.2.4 Reporting processes at the company

Getting to know the data is a challenge when building reporting systems, stated Data Developer B (2019). Different terms and language of advertisers need clarification and familiarization since there might be a gap between developers and the management team. The massive amount of data is also challenging. The visualization tool Power BI has

restrictions on how much data the application can handle. Data Developer B (2019) explained that some of the unnecessary dimensions need to be removed in order to reduce the size of data. The structure needs to be built so that the system would be cost-efficient. Running queries in BigQuery costs therefore the data is collected once in a day.

Similar reporting dashboard to report digital display advertising was built at the company last year. Display advertising report was reporting for example clicks and views on ads, click-through-percentages and demography information about users. Display advertising report and content marketing report different measures, although the purpose is to build visualized dashboards similarly in order to unify diverse reports at the company.

Private browsers were a challenge while trailing data in the display advertising report. Those effect on tracking unique visitors on the page, for the reason that a private browser creates a new visit. A challenge is when a data model is built and some additional features are asked, it is possible that data model does not support adding those features. Then the model needs rechanging in order to modify and add features. (Data Developer B 2019)

Some of the reports at the case company are visualized with Google Data Studio, as it cooperates well with Google's products (Data Developer A 2019). Other visualization tool used is Microsoft Power BI, which is more diverse than Data Studio. BigQuery, a serverless cloud data warehouse tool offered by Google, is used in collecting and aggregating data, since it is a tool that can handle massive amounts of data in the timeframe and it is cost-efficient. (Data Developer B 2019)

Microsoft Power BI has more features and also building the data model is more structured than in Google Data Studio. Power BI was chosen to be the visualization tool used in content marketing reporting, cause of its diversity in features and graphs. Salespersons at the company are also more familiar with reports made with Power BI. Display advertising report was done using Power BI as well and using the same software to visualize reports is more convenient for the company and the users of those reports.

Data Developer B (2019) explained that the benefits of building reporting systems inside

the company are lighter repairing and maintenance. The knowledge and developing the systems for future usage are within the team at the which makes the development process more convenient when there are no other companies in between. Reacting on changes is faster and adding features more fluent. This is due to the fact that the knowledge and skills to develop is within the team at the case company.

4.2.5 Interviews summary

The aim of interviews was to gain knowledge about current reporting situation at the company and research how reporting supports decision-making processes at the company. Among these, knowledge from developers was shared and know-how about technical process of building dashboards and gathering data was raised.

To conclude the results of interviews, content marketing that is produced at the case company are articles that advertise company's customer's services and products. Currently reporting is done manually by copying measured values from Google Analytics into Power Point templates. Reports support decision making at the company within sales and management teams. The most interesting topics of articles at the moment are examined, user segments and target audiences are identified, and content optimized based on those segments.

Different content marketing teams create reports differently. The aim of this study is to unify reporting with the new reporting dashboard. The goal is to simplify, automate and speed up reporting process. New metrics are as well needed in reporting and those are additional indicators added into reporting dashboard.

Metadata of articles is collected using custom dimensions in Google Analytics. Reporting dashboards have been created earlier at the case company using visualization tool Microsoft Power BI, which is as well used in CORE to unify reports at the company. The benefit of building reporting systems within the company makes repairing and maintenance of report lighter, when the knowledge is inside the company.

5 PROJECT AND RESULTS

The objective of the practical project was to build a reporting dashboard to measure the quality and performance of content marketing services offered by the case company. In this case reported content marketing services are articles that include content marketing. The practical project is called CORE in this study. This chapter is about the process of CORE explained. Preprocessing and aggregating data, calculating indicators and building report dashboard is described and challenges are discussed.

Analysis of the results to research questions is done and the results of those are presented. Current practices of content marketing quality management at the case company are explained and the improvement process through CORE is evaluated.

5.1 Process

In the beginning of the project a planning meeting at the case company was organized, and the objective was discussed. After specifications weekly meetings were organized and the process for every week was discussed. CORE project was conducted as a part of author's trainee program and work within web services at the case company.

The communication between team members happened between different communication platforms. Team collaboration tool Slack was used in communicating daily and Microsoft Teams was used in sharing files and having meeting calls. Project management tool JIRA was used to keep up with different tasks and follow the progression of the project.

The aim for CORE was to create a reliable reporting system for measuring quality and performance of content marketing with different key performance indicators. Content marketing services measured in this dashboard are articles in case company's sites. Different websites that offer content marketing articles are around ten and those all are included to reporting dashboard. Indicators that measure performance of articles are

visualized using different graphs and tools on the dashboard. The purpose was to reduce manual workflow of reporting and make user friendly dashboard from where reports could be shown for not only workers at the company but for company's customers as well. The dashboard is meant to be updated daily and its usage would reduce the risks that manual reporting contains, such as calculation mistakes.

5.2 Data collecting and processing

At first, information model was built, and the structure of data was defined. Data types were defined and the source for different data points were specified. Data was collected from three different data sources, Google Analytics, AppNexus and Chartbeat and the structure of data is shown in *Appendix 1*. The source type for data sources were defined. For example, data source types differ between segment feed or standard feed for AppNexus, log level data for Chartbeat and data layer for custom dimensions of Google Analytics. Data from those three sources was aggregated using article ID.

Data from digital advertising platform AppNexus had been available and used in previous reporting processes at the case company. For example, AppNexus data was used in the Display advertising report (Data Developer B 2019). Data from website analytics tool Google Analytics had been used mostly in traffic management reports (Data Developer A, 2019). Data of digital analytics tool Chartbeat was a new source for the case company which made it possible to have new indicators in reporting. Data from Google Analytics and AppNexus were mostly used alone, combining these datasets was a new process, which caused some challenges. Combination of different data sources creates new value for reporting and new indicators that can be measured.

Tools used in CORE were chosen for the needs of the case company and those are tools that are used currently within author's team at the case company. Collecting and aggregating data was proceed using Google Cloud's service, data warehousing platform BigQuery. Running data daily and setting it into production was done using python script and Airflow scheduler. Connecting data into database management system, Google Cloud,

an OCBD connector was used. And finally, reporting dashboard was visualized using visualization tool Microsoft Power BI.

This chapter explains the process of CORE. Data collection process of three different sources is explained. Processing and aggregating data is described, and challenges discussed. Quality of data and calculating indicators are defined and finally visualizing data and building reporting dashboard are proceeded.

5.2.1 Advertisement data

Advertisement data in this study is defined as user information and data of different user segments. Data from AppNexus had been available previously for the case company's use. Data from this source is user information and user segments. Impressions are used calculating different segments and their quantity and percentage of the total. As an example of user segment is education groups of users, those are elementary school, high school, university and trade school. Impressions show how many of article's users are from the education group of each of these and a percentage of the total is then calculated.

Data was collected from AppNexus using Google Cloud's services by using SQL queries in BigQuery. Data was inside different feeds, standard feed and segment feed, therefore those needed to be aggregated using queries and SQL 'join' function. AppNexus Console API uses different feeds for tracking data. According to AppNexus Wiki (A 2019, B 2019) segment feed stores data of network and advertiser-level segments and feed contains one row per segment load. Standard feed provides log-level data on managed publishers' and advertisers' transacted impressions and clicks and conversions. This feed contains one row per tracked impression, click or conversion.

Different user segments of profile information were defined in one table which was joined into the main data table using line item ID. The aggregation procedure of user segments by the author of the thesis is based on the previous work of data developers. By using segments different information could be defined such as user's age group, gender, education group, profession and type of housing. The data of different user segments was

collected from segment feed. Advertiser's ID, device type and regional information about user were collected from standard feed.

AppNexus line items were used getting article ID, article URL and salesperson's name. When salesperson at the company creates a campaign in AppNexus, line items are created and could be used for reporting dashboard. AppNexus Wiki (C 2019) defines that line items represent the agreed upon strategies executed for the advertiser under an insertion order. After setting up a line item and optionally associating it to one or more insertion orders, campaigns can be created, and money usage can be defined to fulfill agreement.

There were some restrictions when using AppNexus line items. There was a limit of 100 characters, which affected to that some of the page URLs were too long for the line item. This could not be corrected, therefore page path from Google Analytics was used as article's URL. Page path collected from Google Analytics has additional URL parameters, which were parsed in order to show the shortest URL for that specific article. If the page path has many different parameters new row for each article is created in the data table, which was not in demand.

5.2.2 Behavioral data

Two other data sources of CORE are offered by web analytics tools Google Analytics and Chartbeat. These tools offer behavioral data of users. Google Analytics had been used in reporting previously at the case company and Chartbeat was a new data source available.

Websites measured have different views and properties in Google Analytics platform, as such data needed to be aggregated and processed by SQL queries in BigQuery. BigQuery is a serverless cloud data warehouse offered by Google and it is a part of Google Cloud Platform. It supports standard SQL dialect which reduces the need for rewriting code. Real-time analytics, data storing and transferring services are available in BigQuery. (Google Cloud Documentation A 2019)

A uniform data table was executed with the help of other data developers at the case

company. Uniform table collected Google Analytics data from different sites and property views of the case company's services. That uniform table was used in CORE where collecting custom dimensions and other dimensions from Google Analytics data. Uniform table is useful, due to the fact that there are many services offered by case company and each of them has their own views in Google Analytics and custom dimensions are tracked differently. Data from mobile applications were in different property views in Google Analytics and was added to the datasets as well. Mobile application data was collected from three services only, as other services do not have mobile applications.

Reading depth, total reading time and average reading time of articles were dimensions collected from Chartbeat's log level data. Reading depth, so called as scroll depth, is a new indicator for content marketing reporting and it shows the percentage of how far user has scrolled on that page. For Chartbeat's data another SQL query was created, and data collected. Data was in more detailed level compared to other data.

There were some challenges when combining data from Chartbeat with other data sets. Chartbeat's JavaScript constantly listens for acts of engagement on pages within an active browser and checks user's engagement (Chartbeat Documentation C 2019). This caused that Chartbeat's data is in more detailed level and it needed to be changed into user level, so that every user has a row of their engagement information in data. In reporting dashboard data was needed in article level, and this is changed while visualizing data in Power BI Desktop.

5.2.3 Google analytics and custom dimensions

In Google Analytics custom dimensions were used collecting data that Google Analytics does not generate automatically. Custom dimensions are default dimensions and metrics in a specific analytics account that the user of the account creates themselves. The purpose of those dimensions is that specific data which Google Analytics does not track automatically can be collected. Custom dimensions allow the combinations of Analytics data and other data. Custom dimensions can appear as primary dimensions in custom reports, but they can also be used as segments and secondary dimensions in standard reports. (Google Analytics

Documentation A 2019)

Custom dimensions in Google Analytics were set up beforehand by a subcontract company of the case company that is working with analytics. Some of the custom dimensions were missing and those needed to be included in reporting dashboard and a request for connecting those to Google Analytics for the subcontract company was committed. Data from missing dimensions starts tracking after the dimensions are added to Google Analytics, therefore Google Analytics cannot track historical data before the connection is made.

Custom dimensions used in CORE were as follows:

- Article ID
- Publishing date
- Author of an article

These dimensions were needed in order to calculate key performance indicators of the dashboard. Missing custom dimensions of some of case company's websites were added through Tag management system using Java Script code snippets.

Custom dimension index numbers were needed in order to collect specific dimensions of the data using SQL queries. Those index numbers were collected from Google Analytics custom dimensions. A uniform data table collecting crucial dimensions and custom dimensions of all services was created at the case company. Required dimensions were collected from that uniform table.

Regular expression was used to define the page type of website and to define article's ID. Regular expression is an expression that describes a set of strings and defines search patterns (Mitkov 2003, p. 754). For example, it is possible to extract strings, in this case article IDs, from website URL. Based on the page path and article ID also regular expression was used to specify the page type to be an article or other. Page type other contains front pages and other pages which are not articles. Article ID was also collected as a custom dimension from Google Analytics and it was as well parsed from article's URL.

Article ID parsed from the URL was used in aggregating data.

5.2.4 Processing data

Aggregating data of Google Analytics, AppNexus and Chartbeat was proceed using article ID, which was parsed from page URL of each article page. If there is not an article ID in the URL, the page is not an article and it was not included to the report dashboard, for the reason of only articles wanted to be reported. Data preprocessing process and aggregation was done by the author of the thesis. Data from these three different sources was collected using SQL queries in Google Cloud's platform BigQuery. At the case company data is collected into Google Cloud's services and from there exported into reporting tools. Those three data tables were all connected to AppNexus' data using 'join' function, to limit the data only to content marketing articles. This was by cause of that only content marketing campaigns were connected to AppNexus at the case company.

Using article ID as a common dimension caused challenges as the data was messy. Custom dimension article ID from Google Analytics had many titles for one single ID, which caused that same article had different dimensions. There should be only one URL and title for each article, but the impurity of data caused miscalculations. This issue is caused possibly by incorrect tag management of dimensions on websites. The problem was corrected using partition functions in SQL, this function takes the article ID that has the most pageviews and supposes its title to be the correct title of that page, and this is then named as a formatted title. Partition functions splits the data based on a similar value, orders it by another value (Google Cloud Documentation B 2019). In this case data splits by the ID and is ordered by pageviews. An example data is shown in table 5 where first value of the title is used formatting a correct title for the article. Similar partitioning was done to article date and page path so that impurity of those dimensions was corrected.

Table 5 Example of aggregating title using partition function

Article ID	Page title	Pageviews	Formatted page title
123	Title example 1	75	Title example 1
123	Another title 1	3	Title example 1
123	Another title 2	1	Title example 1
123	Another title 3	2	Title example 1

Article ID was parsed from website URL and it was used instead of Google Analytics custom dimension article ID, for the reason that custom dimension had incorrect hits, which caused the problem that articles that had wrong ID or title information and pages which were not articles were shown in the report. This issue was corrected when parsed ID from the URL was used due to the fact that each article has its ID on the page path of the article.

Aggregating data from three different data sources was challenging, by reason of data being in different levels. In reporting dashboard article level data is visualized, though log level data from Chartbeat was at more detailed level. By aggregation it was possible to get these datasets to similar level. Impurity of data caused challenges, which needed to be corrected as well as article ID connecting challenges.

Advertising data and Google Analytics data were aggregated using SQL join-function and article ID. The data was set into production using Airflow, which is a platform that programmatically can schedule and monitor workflows. Airflow scheduler executes tasks on an array of workers following specified dependencies. Workflows can be defined as code, which makes the process more testable and maintainable. (Airflow Documentation 2019)

A python script was created to set data into production using Airflow and production deployment was done by a lead programmer at author's team at the case company. SQL queries which collect the data from Google Analytics, AppNexus and Chartbeat were

parametrized in python script. This means, all dates were parametrized so that the script is using a date from the current date and airflow runs the script once in a day. Data is collected from each day and added to the data tables. These tables were then connected to Microsoft Power BI using ODBC connector, which is an interface by Microsoft that allows applications to access data in database management systems using SQL (Progress FAQs 2019). Power BI is the visualization tool used where the dashboard was built, and data visualized. Some of the calculations were done as well in Power BI. Expiration time of three months was set for the data, for the reason of limitation in visualization tool's capacity to handle massive amounts of data.

Data processing is visualized in figure 8. Article ID was used to connect three data sources and then data was set into Airflow in order to schedule queries. These data tables were then taken into Google Cloud Platform, where data was processed using SQL in BigQuery. Finally, data was imported into visualization tool Power BI using ODBC connector.

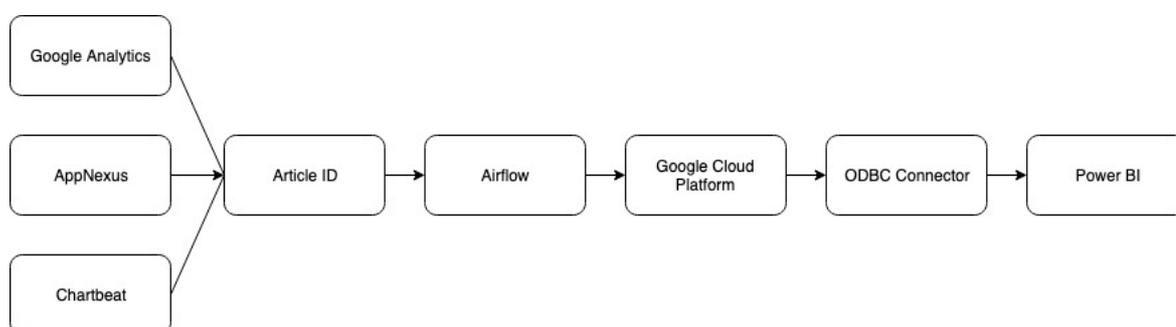


Figure 8 Data processing flow

5.2.5 Quality of the data

The quality of data was measured by calculating different indicators of the data. Measuring quality of Google Analytics custom dimensions is important since it affects the quality of the whole dataset. A report of those metrics was built in Google Data Studio.

The data was collected using SQL and then imported to Data Studio and metrics were calculated. Websites were categorized to articles and other pages. The article recognition was proceeded with regular expression functions, so called regex functions, where the page

path of different pages was parsed so that the articles' page type was categorized as an article.

One of the quality metrics was views where custom dimension article ID was found from the view. The metrics was a calculated percentage of the total amount of views of the distinct count of article IDs. The results were around 99 percentages. Although this result is good, measuring whether the article ID is exactly correct ID of that specific article is more challenging. These cases happened where Google Analytics custom dimension was presenting article's ID incorrectly. To clarify why it happens needs more research around the topic.

A challenge was also noticed in the calculations of pageviews and impressions. Impressions should be less than pageviews, since a pageview is tracked every time a user visits a page and impressions only when the ad is seen (Google Ads Documentation 2019, Google Analytics Documentation B 2019). Pageviews and impressions were compared between articles and miscalculations were tracked and corrected.

5.3 Calculation

Unique pageviews information could not be collected straight from Google Analytics into BigQuery, because of the structure of Google Analytics data (Data Developer A 2019). Therefore, article's pageviews and unique pageviews were calculated within the SQL query. Since the data was structured per sessions, unnesting for the data needed to be done. Unnesting creates a row from every hit that has been in the page. Pageviews number would duplicate for each row per article. To avoid duplicates, instead of using pageviews dimension, pageviews were calculated as a sum of those rows per article. Unique pageviews are calculated as distinct visitors of a site using visitor ID from Google Analytics data.

Impurity of data caused challenges in calculations. Multiple titles needed to be corrected using partition function which takes the first value titles to be the real title of an article and

then summarizes pageviews of that article. Similar partitioning was done to page paths and article date as well to fix the incorrect dimensions.

Impressions were calculated in the query by a sum of the events in the data points. Information of impressions were used to measure profile data, which came from segment and standard feeds. Impressions were used instead of pageviews because AppNexus collects impressions and clicks instead of pageview information. Impressions were defined in AppNexus data as event type “imp”, as the events can also include clicks and conversions, which in this case were not included as impressions. AppNexus collects 11 different impression types, and in CORE the impression type “Kept” was used. AppNexus Wiki (B 2019) defines this impression type as follows: “Kept: One of your managed advertisers served a creative.”

There were challenges with calculating impressions from AppNexus data. Line items were earlier used differently by the sales team at the company so when the change to use line items as article ID and URL of an article happened, there were still previous line items in the data, which caused a miscalculation in impressions. This was corrected by using filtering in the data. Only articles, which have a correct URL as line item were collected and therefore only content marketing articles were included to calculations.

Calculations for Chartbeat data were done in visualization tool Power BI. Data was collected from every day and added daily into data table. Average scroll depth and average and total reading time of an article were metrics needed in reporting. Averages could not be calculated straight from queries since more data is added daily. Total reading time is a sum of all user’s reading time and average of those was as well calculated. KPI metrics are presented in figure 9.

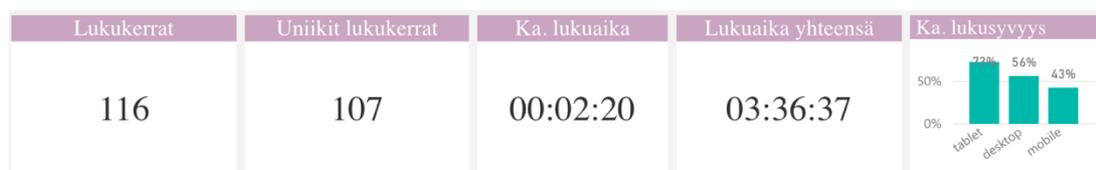


Figure 9 KPI bar, screenshot from dashboard - visualized KPIs: pageviews, unique pageviews, average reading time, total reading time and reading depth of different devices

There were challenges calculating article's scroll depth, as all the articles and web pages have different length, which varies on different devices. Average scroll depth could be only calculated for each article individually, for the reason of changes in article's length. It was calculated in Power BI dividing article's scroll depth with article's maximum scroll depth. Calculations were done for desktop, mobile and tablet and they were shown in KPI bar as a bar plot shown in figure 9.

Gauge graphs need maximum values in order to measure performance. Article's pageviews, reading depth and reading time are measured. This graph visualizes the goals of each service or a campaign and the arrow shows where article's situation is currently. If it is on the right side, it is reaching the target goal. Maximum values and middle steps of target goals for gauge graph were collected from an excel where the information was gathered from sales teams of different services. The goals were set by the sales teams and those were based on the knowledge about former marketing campaigns. A gauge graph of article's pageviews shown in figure 10.

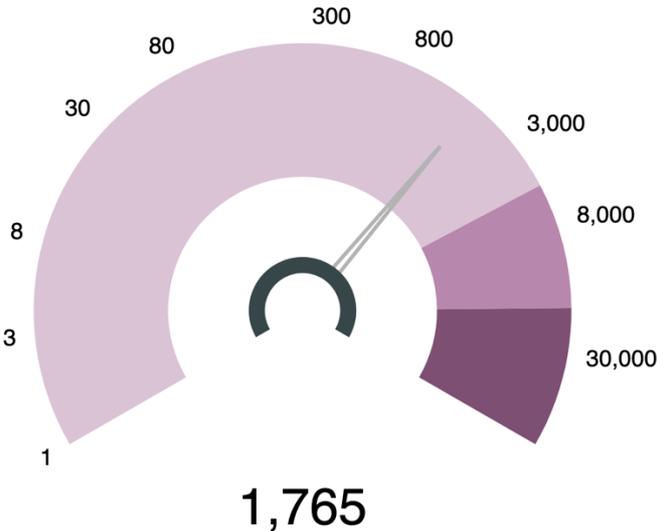


Figure 10 Gauge of pageviews, screenshot form dashboard

5.4 Reporting dashboard

Microsoft Power BI was chosen for the tool to visualize reporting dashboard. To run Power BI on Mac computer a virtual computer server needed to be installed, as Power BI Desktop only works in Windows computers. Finnish was chosen for reporting language considering that most of case company's services are reported in Finnish.

At first a draft dataset was used, and draft of the report was built. Missing dimensions were noticed when the data was visualized using Power BI. Those dimensions were added to the data tables and the report could be developed from draft to a prototype report. Data model was built and aggregations between different tables were made in Power BI, as shown in the *Appendix 2*. Aggregations between Google Analytics and AppNexus, Chartbeat, could be done using article ID.

User information was in a segment table which was separated based on the different user data, for example information about user's gender, age, education and housing. Those segments were separated into different tables. Additional tables were used to get titles of each segment and those tables were aggregated with AppNexus data using segment IDs. Through the draft report all the miscalculations of pageviews and impressions were tracked and corrected. Miscalculations are still after this possible due to the impurity of data.

Different visualization charts were used in the report. Tables, bar charts, donut charts, funnel charts, gauges, tree maps and word clouds were used. Dial gauge, HTML viewer and a word cloud are custom visuals downloaded from Power BI marketplace. Filters were used in each visualization and page. For example, blank values were filtered. Visualized graphs were chosen by the management team at the case company. These indicators are most measured at the company and the graphs chosen are most used when reporting performance to company's customers.

Ten pages were used in the dashboard. All of the pages visualized different topics. The pages were as follows: homepage, article page, summary, barometers, traffic sources, geographical information, audience, B2B, traffic sources as visualized and domain

sources. Every page visualizes article's KPIs such as pageviews, unique pageviews, reading time, reading depth and total reading time. Homepage has filter boxes of advertisers and articles and a filter for article's publishing dates. Filtering at the homepage will remain through the whole report. KPI measures were also be filtered by graphs on each page, so that impressions were filtered according to the geographical information of users. For example, impressions from female and male users of an article could be presented.

Article page summarizes KPI information about articles on a table and summary page has a HTML viewer of an article, so that the selected article can be shown. Barometers/Gauges visualize KPIs and their maximum values. Traffic sources page has a donut graph of both all traffic sources and traffic sources from social media. Traffic sources and domain sources are then visualized as word cloud figures as well. Geographical information page shows the impressions of an article in different cities and regions. An example graphs of traffic source from social media and a tree map of cities measured by article's impressions are shown in figures 11 and 12.

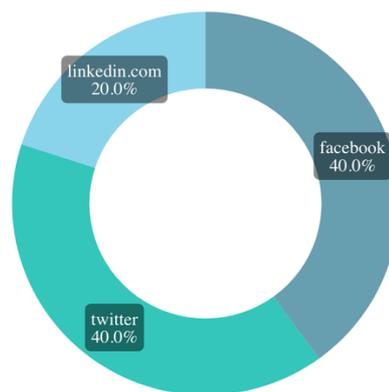


Figure 11 Social media traffic source, screenshot from report dashboard

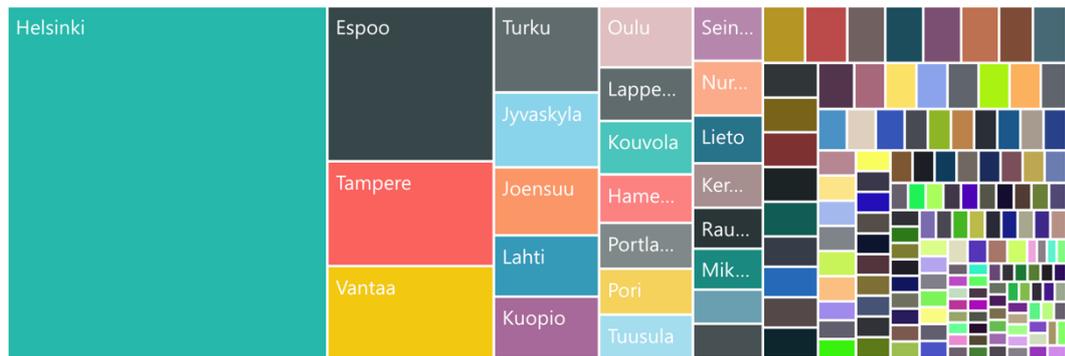


Figure 12 Figure 10 A tree map of cities, screenshot from dashboard

Audience page has geographical information about users of an article. Gender, age, education, salary groups and device types are visualized and also compared to the citizens average values in Finland. B2B section visualizes impressions of an article based on the profession and size of the company the users are working. Example graphs from report dashboard are presented on figures 13 and 14. Figure 13 shows the percentage of different age groups compared to average values of citizens and figure 14 represents distribution based on education groups.

Jakauma ikäryhmittäin väestö vs kampanja

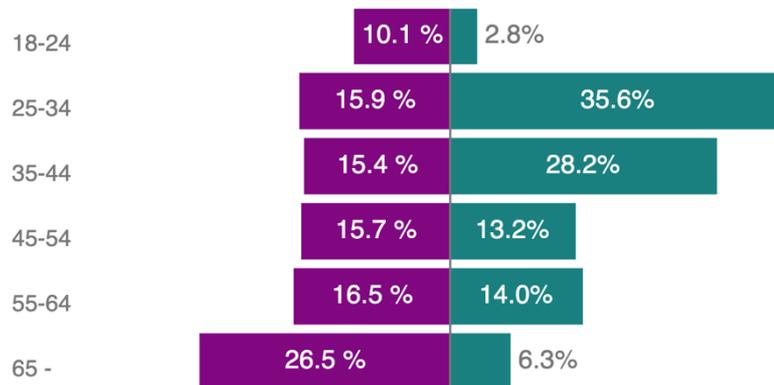


Figure 13 Age distribution, screenshot from dashboard

Jakauma koulutusluokittain

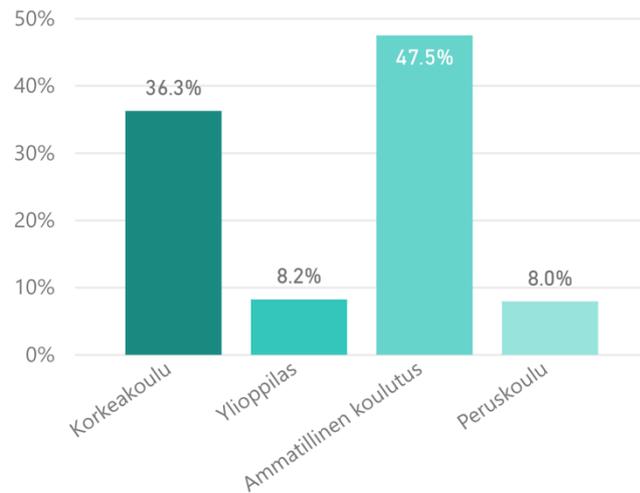


Figure 14 Education groups, screenshot from dashboard

Data analysis expression (DAX) and Power Queries were used on Power BI to format and calculate new values. Microsoft Power BI Documentation (2019) explains DAX as a collection of functions, constants, operators that can be used in a formula or expression to calculate and return values. In other words, DAX helps to create new information from data that is already in the model. Power Queries use M language, which is explained as follows: “Power Query M formula language is a function that is a mapping from a set of input values to a single output value. A function is written by first naming the function parameters, and then providing an expression to compute the result of the function” (Microsoft Power Query Documentation 2019).

Power Queries were used converting reading time from seconds to days in order to format it in the report. Also, last updating time of the dashboard needed to be shown in the report, and that was done using power queries as well as changing data sets from draft data into production data sets on the dashboard.

Formatting KPI metrics of Chartbeat, such as average and total reading time, was challenging as those values were presented as seconds in data and in the dashboard needed to be shown as hours and minutes. Formatting seconds into data a DAX code in Power BI was used to present values.

DAX was as well used to calculate reading time of articles with summary function for total reading time and average function for average. Calculating profiled pageviews, a percentage of impressions per pageviews, was calculated using DAX. This indicator shows how many of article's users could be profiled in AppNexus user segments information. Not all of the users can be profiled for example by the reason of private browsers, which do not collect enough data (Data Developer B 2019). DAX was also used formatting report, for example showing advertiser in the top bar, if an advertiser is chosen on the filter box. Calculations of article's average scroll depth of desktop, mobile and tablet were proceed using DAX functions and maximum scroll depth of an article.

An ODBC connector was used to connect data to database management system, which was Google Cloud's service in this case. An ODBC connector is defined as follows: "An ODBC driver uses the Open Database Connectivity (ODBC) interface by Microsoft that allows applications to access data in database management systems (DBMS) using SQL as a standard for accessing the data" (Progress FAQs 2019). After connection was made, the report dashboard could be refreshed daily.

When the report was finalized, more data of content marketing articles were collected and imported into Power BI. Data was collected daily and added into data tables and reporting dashboard was scheduled to refresh once in a day. UX visualization was finalized by the manager of the project at the case company.

5.5 Results and analysis

At the current moment of writing this study, new dashboard is in the process of deployment at the case company. Reporting dashboard has been tested and data is being collected from current content marketing campaigns. The full deployment of new dashboard will happen after data has been collected from couple of months and a training session for its users organized.

Result section explains how current practices of content marketing performance is managed at the case company and how the process of measuring performance is improved through this thesis. Interviewing employees at the company and literature review around topics concerning performance measurement were methods used to gain knowledge about current practices of content marketing reporting and how to improve the process. Phenomenon around current reporting process is analyzed in order to improve the process of reporting. Improvement of content marketing measurement process through this study is observed and comparison is done between reporting processes before and after this study. Process before new reporting dashboard executed in this thesis is compared with the reporting process after using new reporting dashboard in measuring performance of content marketing articles.

5.5.1 Current practices of content marketing performance management

Reporting processes before CORE were examined in this study through interviews of employees who work with content marketing and reporting. Reports were collected manually using indicators from Google Analytics and copying those into Power Point templates (Development Manager at Media Sales 2019). Collected metrics were pageviews, unique pageviews, reading time and traffic sources, network domains of articles according to Content Marketing Producer (2019).

Product leader (2019) at the company estimated that current reporting process is taking approximately 100 full working days. Reduced working hours and simplified reporting process are an advantage of CORE. And as Content Marketing Producer (2019) stated that the working time released could be focused on analysis of the performance of campaigns and improving services rather than making reports. Manually created reports using Power Point templates might be as well not consistent between each other and mistakes are yet more possible when values are copied manually.

5.5.2 Reporting dashboard

New reporting dashboard created in CORE is the automated reporting system created in

this study. Reporting dashboard collects data from three different data sources and its purpose is to report content marketing articles. Content is measured using key performance indicators chosen by management team at the case company. Geographical information of users is visualized in order to support measuring performance and target audience better and visualize how goals of certain campaigns are reached.

Reporting dashboard has ten pages and each page has its indicators. There is a homepage that collects all the article's and advertiser companies and report can be filtered so that only certain article's or company's campaign information is visible. Filtering supports using the dashboard for all content marketing reports at the case company, when there are many campaigns and articles reported at the same time. Each page shows indicators on a top bar such as pageviews, unique pageviews, average reading depth and average reading time and total reading time. Reading depth is shown as a bar plot of user's reading depth with each device desktop, mobile and tablet, for the reason of different devices variety in page length of an article.

Article's page included similar KPI information than KPI top bar on a table format. Table makes it possible to compare article's performance between each other. Also, one page shows the article through HTML viewer. Traffic sources and domain source pages describe where users have come to the site, both traffic source, social media sources and domains.

Dashboard has a page for gauge graphs, which measure pageviews, reading time and reading depth of an article. Those visualize the goals of each service. An arrow shows the current situation of an article and color codes visualize target goals and are those met in that specific campaign. There are goals for reading depth, pageviews and reading time for articles that are measured. Every service has their own target goals of two stages and those are written in an excel table from where the reporting dashboard visualizes the goals. Excel sheet of those goals can be updated by sales managers who follow these reports. This gauge graph is shown in pageviews gauge in figure 10, where the two target goals are in darker color and the arrow shows the current situation of the article.

In the report dashboard there are pages that describe geographical information of audience,

such as cities, gender, age, education and salary group, company's size and profession group. Audience page has information of the audience, and age, gender and education as well compared to the citizens average. Business to business (B2B) page visualizes user's company's size and profession groups using bar plots. These audience pages help to gain understanding of the audience of certain articles and campaigns and target advertising better.

5.5.3 Improvement of reporting

Improvement of reporting process was observed and compared to the process before CORE. An additional interview of development manager at media sales was conducted after CORE was finished and the process of deploying new reporting dashboard had begun. Through additional interview knowledge about improvement process is gained.

Development Manager at Media Sales (B 2019) explained that CORE accounts for its purpose. After the deployment of new dashboard, the usage will be followed during couple of months. Then, additional features will be planned and projected. When reporting process is manual and values are copied by the user of report, there is a higher change for mistakes in copied values. More precise metrics are a benefit of the new reporting dashboard, stated Development Manager at Media Sales (2019).

A salesperson who uses reports can export the dashboard as a PDF or Power Point straight from Power BI and send it to customer or use it as an internal report within the company. This process is fast and fluent and there is more time left for analyzing results of performance. Independent reporting of results in a desirable way is a faster lane to compose results than parsing manually (Development Manager at Media Sales B 2019). There is no other workload included in between which makes the process even more fluent. When salesperson can independently make reports that releases many resources for other tasks, stated Development Manager at Media Sales (B 2019).

Combining different data sources created new value for reporting at the case company. Dashboard could include more versatile information when there were more data sources

available. Measuring performance of articles improved when there was more data and indicators on hand. New indicators, such as reading depth of an article and renewed gauge graphs, were included to the dashboard (Development Manager at Media Sales B 2019). Gauge graphs visualizes pageviews and average reading time of an article and reading depth of an article with different devices. These gauge graphs support reporting how the goals of certain campaign have been reached.

Gauge graphs measure how targets have been reached. Audience pages of reporting dashboard for their part measure and visualize audience groups of an article. Audience groups such as education, age and gender group, help to get more understanding of what kind of users those articles attract to. Gaining knowledge of audience groups strengthens targeting advertising.

Development Manager at Media Sales (B 2019) explained that new reporting dashboard enables not only reporting articles but customer and service segments. Indicators and average values can be measured of different services and customer companies using filters in dashboard. This enables measuring how services and customers content has performed.

Content marketing performance measurement process before deploying new dashboard and after deploying it is summarized in table 6.

Table 6 Performance measurement process before and after CORE

Before CORE	After CORE
<ul style="list-style-type: none"> • Manual data collection • Process takes many working hours • Mistakes are possible • Every user makes different report • Only articles are measured 	<ul style="list-style-type: none"> • Automated data collection • Process is fast, only needs exporting report • Reduced miscalculations • Unified reporting dashboard • Customer and service segments are measured in addition to articles • New indicators • Gauges that visualise how targets have been met • Three data sources information combined

After deploying CORE, research about using the new reporting dashboard could be done. Which indicators are the most suitable for their use and which indicators are still needed in order to get the most suitable results and analysis of content's performance.

6 DISCUSSION

This chapter includes discussion about measuring performance of content marketing, indicators that are used in measurement processes and future implementations of CORE. These topics are reviewed considering the new reporting dashboard executed in this study.

6.1 Measuring performance of content marketing

Increased use of digital services and social media causes the need for brands to be present in those channels where customers are interacting. If before the purchase decision customers have been offered value by the company, they are more likely know the brand or a product and appreciates the company. Offering value for customers can be created using content marketing.

Measuring performance of content is important in order to create more valuable and attractive content in the future. If content is not performing well, problematic placement or irrelevant articles can be noticed. Through the reports decision makers at the company can see where the improvement is needed and what kind of content is performing better.

CORE brings value for measurement processes at the case company by making the performance management process faster and more efficient. Measuring content marketing content is valuable for the case company, when the services and social media presence has become critical for brands.

Jackson (2016, p. 11-12) describes that changing the culture is a challenge. Changing old ways of acting within the organization can be challenging. It is meaningful to create an atmosphere where change towards new ways of working and getting better results are possible. When leaders can describe the purpose behind new actions, and goals of the whole team or a certain project are known, it is probably easier to adapt the changes.

Changing the culture can be a challenge as well when CORE reporting dashboard is taken

into use at the case company. Development Manager at Media Sales (B 2019) stated that this new dashboard enables independent reporting by the sales teams, and this is a cultural change. When the reporting processes have been manual for many years, it is always difficult to change customary habits and ways of working. To make the change easier, the purpose of this new dashboard was explained by the management teams and information about new reporting process increased. Trainings to use new reporting system will be organized for the sales teams.

6.2 Indicators

As Cutroni (2010, p. 1) defines three main tasks for every business to handle around digital analytics: measuring data, continuously improving their website and aligning measurement strategy together with business strategy. When business strategy is known, choosing correct indicators to measure whether certain goals are reached, and in which factors the organization has most to improve.

According to the study of Järvinen and Karjaluoto (2015, p. 123) it seems that rarely the lack of skills is a major obstacle to understand the opportunities of digital analytics and its usage in performance measurement. Selecting metrics of company's needs is commonly a challenge where understanding marketing strategy would make the selection process more fluent.

The purpose of CORE is not to measure performance of a website or to improve website's performance. CORE is targeted to customers and stakeholders of the case company. Indicators are constant for all of the services. The purpose of the dashboard is to measure content's performance and improve content based on the reports and analysis. Targeting and optimizing articles better are the major aims of measuring content. If the focus of measuring was to improve website's performance, the indicators might be continuously changing and aligning the business strategy of the company.

The focus in the dashboard executed in this study is more in the content marketing business and customers of the case company rather than company's business itself. Content

marketing business at the case company has its own performance indicators and those were selected by management team at the company. Indicators were chosen to the need of reporting content marketing articles for the customers of case company. The use of indicators has an effect on choosing them. Different indicators can be chosen to measure content internally and externally to show performance to customers.

Indicators are many to choose from and as Jackson (2019, p. 82) described that instead of having tons of things to measure it is better to have three or four metrics. Those metrics should be chosen according to the aims of business. Within different marketing campaigns metrics can vary. When the strategy is well defined and goals known, choosing right indicators and metrics becomes easier.

CORE dashboard has more indicators than just three or four. Instead the dashboard has ten pages presenting different metrics. All of those pages have a different purpose of measuring performance, for example there is a page that visualizes the audience groups and another page that visualizes the B2B indicators. Audience page visualizes information of user segments, such as age, gender, education and device type of users. B2B page visualizes user's professional status and the size of a company. Also, measuring target goals using barometers and showing traffic sources have their own pages. Depending on a marketing campaign and customer, these report pages can be used. Not all of the indicators and metrics are examined each time.

6.3 Future implementations

At the current moment of writing, new reporting dashboard is in the process of deployment at the case company. The usage of this dashboard will be followed during couple of months by the sales teams who use reports. Additional features will be planned, and a second version of the dashboard produced later on based on those changes desired by management and sales teams.

Some of the indicators of CORE could be developed further. For example, reading depth of

an article is currently measured from the whole page. Reading depth percentage shows how far user has scrolled on that page. This is taken into account while looking at the gauge graphs which measure targeted goals of reading depth. This indicator would be more accurate if measuring only article's reading depth would be accessible. After the article on a web site there can be empty space or other ads below, while if only article's reading could be measured, the indicator would be more relevant. Measuring only article's scrolling depth instead of the whole page was considered in CORE, but within the project's timeframe it was not applicable to execute. This addition would be meaningful when measuring the performance of a reading depth of an article and will be projected to the future implementations of the dashboard.

Some of the graphs and visual aspects have got feedback from the sales teams who measure and analyze the performance of content marketing campaigns. The chosen colors of visualizations have been changed a bit according to the feedback. Also, a second version of the dashboard has been created. The second version is similar to the original one, but without the reading depth indicator, which still needs improvement that would apply in the future versions of this dashboard. It is essential to improve the new measurement dashboard and measurement process according to the feedback from its users.

Gauge graphs, that visualize how targets have been met of reading time, reading depth of different devices and pageviews, could be developed further. For example, a time dimension could be added to the dashboard. Currently only article's publishing date is visible on the dashboard. Time dimension would present on which date and time pageviews and other information has come to that specific article. Time dimension would help to measure performance so that it is known from how long is the time period of data that has been gathered. Gauge graphs would present more accurate values when time could be limited for example from one to three months. Then report would show different goals for different time periods instead of showing only one target value which is shaped for three months of data.

One additional feature that could be developed in CORE is that when the salesperson creates campaigns in AppNexus, they need to copy article's ID and URL as AppNexus line

items. If the URL or ID is mistakenly copied, the article will not show correctly on the dashboard. Matching URL and ID between article, to ensure those belonging to the correct article, would be a future addition. Correcting line items, where the article ID is incorrect, might not be a simple task to develop. Now data is mapped using article ID, and when it is mistakenly copied, the data does not connect to the right article.

7 CONCLUSIONS

The purpose of this study was to research how current practices of content marketing performance measurement were done at the case company and how the process of performance management could be improved. This study was executed as part of authors work at the case company and experience of author about reporting processes and tools of development and analytics increased through this thesis. A practical project where new reporting dashboard was produced to improve performance measurement process.

Interviews of employees were accomplished to research how the practices of reporting were managed at the case company. Literature review was conducted to support authors knowledge about the topics around measurement processes. The improvement of performance measurement process at the case company was analyzed through observations, phenomenological analysis and comparison between the earlier content marketing process and process after executed reporting dashboard.

The first research question was to clarify what the current practices of content marketing management at the case company are. Before this study reports were created manually by collecting metrics from Google Analytics and copied into Power Point templates. This process was taking many working hours from sales teams.

The second research question was to observe how performance measurement process of content marketing can be improved. After deploying new reporting dashboard, the reporting process has been automated and sales teams are able to export reports straight from the dashboard. This automated process reduces many working hours for actually analyzing the results of campaigns. Reports are more precise when the metrics are not collected manually. New dashboard enables not only reporting articles but customer and service segments as well. This dashboard improves optimizing and targeting campaigns for customer segments. Usage of this new dashboard will be followed for couple of months and future implementations will be executed in the future versions of the dashboard. Reading depth indicators will be improved, and possible additional indicators enhanced.

REFERENCES

Airflow Documentation [online document]. 2019. Apache Airflow Documentation. [Accessed 1.7.2019]. Available at: <https://airflow.apache.org/>

AppNexus Documentation [online document]. 2019. AppNexus About. [Accessed 12.6.2019]. Available at: <https://www.appnexus.com/about>

AppNexus Wiki A [online document]. 2019. Segment Feed. [Accessed 28.6.2019]. Available at: <https://wiki.appnexus.com/display/api/Segment+Feed>

AppNexus Wiki B [online document]. 2019. Standard Feed. [Accessed 28.6.2019]. Available at: <https://wiki.appnexus.com/display/api/Standard+Feed>

AppNexus Wiki C [online document]. 2019. Line Item Service. [Accessed 1.7.2019]. Available at: <https://wiki.appnexus.com/display/api/Line+Item+Service>

Ashley, C., Tuten, T. 2014. Creative Strategies in Social Media Marketing: An Exploratory Study of Branded Social Content and Consumer Engagement. *Psychology and Marketing*, Vol. 32(1): 15–27 (January 2015). p. 15-17, 24. East Caroline University: Wiley Periodicals, Inc.

Başkarada, S. [online document]. 2014. Qualitative Case Study Guidelines. *The Qualitative Report*, 19(40). p. 1-3. [Accessed 25.9.2019]. Available at: <https://nsuworks.nova.edu/tqr/vol19/iss40/3>

Chartbeat Documentation A [online document]. 2019. Implement Chartbeat for Publishing. [Accessed 13.6.2019]. Available at: <http://support.chartbeat.com/docs/>

Chartbeat Documentation B [online document]. 2019. Integrations. [Accessed 28.6.2019]. Available at: <http://support.chartbeat.com/docs/integrations.html>

Chartbeat Documentation C [online document]. 2019. Description of Methodology. [Accessed 27.9.2019]. Available at: <http://support.chartbeat.com/docs/methodology.html>

Content Marketing Producer. 2019. Interview about current situation of reporting. Interviewer: Laura Manninen. 5. June 2019. Personal interview.

Cutroni, J. 2010. Google Analytics: Understanding Visitor Behavior. p. 1-3. USA: O'Reilly Media, Inc.

Data Developer A. 2019. Interview about data collection. Interviewer: Laura Manninen. 27. May 2019. Personal interview.

Data Developer B. 2019. Interview about data collection. Interviewer: Laura Manninen. 3. June 2019. Personal interview.

Development Manager at Media Sales. 2019. Interview about current situation of reporting. Interviewer: Laura Manninen. 5. June 2019. Personal interview.

Development Manager at Media Sales B. 2019. Additional interview about improvement of content marketing process. Interviewer: Laura Manninen. 4. October 2019. Personal interview.

Digishuffle blogs [online document]. 2017. Google Analytics – Sessions & Hits – Metrics Simplified Series. [Accessed 3.10.2019]. Available at <https://www.digishuffle.com/blogs/google-analytics-sessions-hits-metric-simplified/>

Digital Guide [online document]. 2019. Tag management: basics & systems. [Accessed 17.9.2019]. Available at <https://www.ionos.com/digitalguide/online-marketing/web-analytics/how-does-tag-management-work/>

Franceschini, R., Galetto, M., Maisano, D. 2007. Management by Measurement – Designing Key Indicators and Performance Measurement Systems. p. 2, 8, 110-112. Italy: Springer. ISBN 978-3-540-73211-2.

Google Ads Documentation [online document]. 2019. Impressions: Definition. [Accessed 24.6.2019]. Available at <https://support.google.com/google-ads/answer/6320?hl=en>

Google Analytics Documentation A [online document]. 2019. Custom dimensions & metrics. [Accessed 20.5.2019]. Available at <https://support.google.com/analytics/answer/2709828?hl=en>

Google Analytics Documentation B [online document]. 2019. The difference between Google Ads Clicks, and Sessions, Users, Entrances, Pageviews, and Unique Pageviews in Analytics. [Accessed 23.6.2019]. Available at <https://support.google.com/analytics/answer/1257084?hl=en>

Google Analytics Documentation C [online document]. 2019. Traffic source dimensions. [Accessed 15.8.2019]. Available at <https://support.google.com/analytics/answer/1033173?hl=en>

Google Analytics Documentation D [online document]. 2019. About Keywords in Search Network campaigns. [Accessed 15.8.2019]. Available at <https://support.google.com/google-ads/answer/1704371?hl=en>

Google Analytics Documentation E [online document]. 2019. About goals. [Accessed 17.8.2019]. Available at <https://support.google.com/analytics/answer/1012040?hl=en>

Google Cloud Documentation A [online document]. 2019. BigQuery. [Accessed 22.5.2019]. Available at <https://cloud.google.com/bigquery/>

Google Cloud Documentation B [online document]. 2019. Analytic Functions Concepts in Standard SQL. [Accessed 30.7.2019]. Available at <https://cloud.google.com/bigquery/docs/reference/standard-sql/analytic-function-concepts>

Google Marketing Platform Documentation [online document]. 2019. Product overview – Data Studio. [Accessed 22.5.2019]. Available at <https://marketingplatform.google.com/about/resources/data-studio-product-overview/>

Greenstein, S. 2015. Behind the Buzz of Behavioral Data. *IEEE Micro*, 35(2). p. 88-c3.

Jackson, S. 2019. *Cult of Analytics: Driving online marketing strategies using web analytics*. First edition. p. 50, 82. Great Britain: Butterworth-Heinemann. ISBN 978-1-85617-611-8.

Jackson, S. 2016. *Cult of Analytics: Data Analytics for marketing*. Second edition. p. 1-4, 11-12, 24-26, 36-38, 47, 59, 162. Great Britain: Routledge. ISBN 978-1-315-73474-3.

Järvinen, J., Karjaluoto, H. 2015. The use of Web analytics for digital marketing performance measurement. *Industrial Marketing Management*, 50(C), p. 117-123.

Järvinen, P., Järvinen, A. 2000. Tutkimustyön metodeista. p. 4-5, 153-154, 163. Tampere, Finland: Opinpaja Oy. ISBN 951-97113-8-4.

Kelleher, C., Wagener, T. 2011. Ten guidelines for effective data visualization in scientific publications. *Environmental Modelling and Software*, 26(6). p. 824-826.

Lieb, R. 2011. *Content Marketing: Think Like a Publisher – How to Use Content to Market Online and in Social Media*. p. 1-3, 12-13, 19. USA: Que Publishing. ISBN 978-0-7897-4837.

Mack, M., Woodsong, C., MacQueen, K., Guest, G., Namey, E. 2005. Qualitative Research Methods: A Data Collector's Field Guide. p. 1-2, 4. USA: Family Health International. ISBN 0-939704-98-6.

Microsoft Power BI Documentation A [online document]. 2019. What is Power BI Desktop? [Accessed: 12.6.2019]. Available at: <https://docs.microsoft.com/en-us/power-bi/desktop-what-is-desktop>

Microsoft Power BI Documentation B [online document]. 2019. Visualization types in Power BI. [Accessed: 30.7.2019]. Available at: <https://docs.microsoft.com/en-us/power-bi/visuals/power-bi-visualization-types-for-reports-and-q-and-a>

Microsoft Power BI Documentation C [online document]. 2019. DAX basics in Power BI Desktop. [Accessed: 26.9.2019]. Available at: <https://docs.microsoft.com/en-us/power-bi/desktop-quickstart-learn-dax-basics>

Microsoft Power Query Documentation [online document]. 2019. Understanding Power Query M functions. [Accessed: 26.9.2019]. Available at: <https://docs.microsoft.com/en-us/powerquery-m/understanding-power-query-m-functions>

Mitkov, R. 2003. The Oxford Handbook of Computational Linguistics. Oxford University Press. p. 754. ISBN 978-0-19-927634-9.

Munhall, P. 2007. Nursing Research – A Qualitative Perspective. Fourth Edition. p. 146. USA: Jones and Bartlett Publishers, Inc. ISBN 978-0-7637-3864-8.

Northcutt, N., McCoy, D. 2004. Interactive qualitative analysis. p. 196. USA: SAGE Publications, Inc. ISBN 9781412984539.

Picard, R. 2000. Measuring Media Content, Quality and Diversity – Approaches and Issues in Content Research. p. 97 -99. Turku, Finland: Suomen Akatemia. ISBN 951-738-827-6.

Product Leader. 2019. Interview about current situation of reporting. Interviewer: Laura Manninen. 5. June 2019. Personal interview.

Progress FAQs [online document]. 2019. What Is an ODBC driver? [Accessed 25.9.2019]. Available at: <https://www.progress.com/faqs/datadirect-odbc-faqs/what-is-an-odbc-driver>

Pulizzi, J., Barrett, N. 2009. “Get Content Get Customers” – Turn Prospects into Buyers with Content Marketing. NSB Management Review, Volume 2, No. 2, February 2010. p. 98-99. ISBN 978-0-07-067093-8.

Rubinson, J. Pfeiffer, M. 2005. Brand Key Performance Indicators as a Force for Brand Equity Management. Journal of Advertising Research 45, no. 2 (2005). p. 187, 189, 195.

Samsonowa, T. 2012. Industrial Research Performance Management - Key Performance Indicators in the ICT Industry. p. 27-32, 37. Germany: Springer. ISBN 978-3-7908-2761-3.

Saura, J., Palos-Sánchez, P., Suárez, L. 2017. Understanding the Digital Marketing Environment with KPIs and Web Analytics. Future Internet 2017, 9, 76. p. 7-9, 11.

Schivinski, B., Christodoulides, G, Dabrowski, D. 2016. Measuring Consumers' Engagement with Brand-Related Social-Media Content: Development and Validation of a Scale That Identifies Levels of Social-Media Engagement with Brands. Journal of Advertising Research 1 March 2016, Vol.56(1). p. 64-65, 74-75.

Sebei, H., Taieb, M., Aouicha, M. 2018. Review of social media analytics process and Big Data pipeline. Social Network Analysis and Mining, 8(1). p. 2, 11.

Tufte, E. 1990. Envisioning Information. p. 51. USA: Graphics Press LLC.

Tufte, E. 2013. The Visual Display of Quantitative Information. Second edition. p. 9, 13, 81. USA: Graphics Press LLC.

APPENDICES

Appendix 1 Data Structure - first draft

Index	Subindex	Name / Location	Location	Type	Description	Source	Source type	Generated by	Tab
1		Advertiser ID		int	The ID of the advertiser	Appnexus	Standard feed	Appnexus	[not visible]
2		Advertiser name		text	Name of the advertiser	Appnexus	Service	Appnexus	Front
3		Article ID		int	The ID of the article			User	[not visible]
	3.1	Article ID	GA Custom Dimension	int	The ID of the article	Google Analytics	BigQuery	Datalayer	[not visible]
	3.2	Article ID	Line item reporting label	int	The ID of the article	Appnexus	Service	User	[not visible]
	3.3	Article ID	Line item	int	The ID of the article	Chartbeat	LLD	Chartbeat	[not visible]
4		Article Name	Line Item name	text	Topic of the article	Google Analytics	BigQuery	User	Front
5		Pageviews		sum	Number of pageviews	Google Analytics	BigQuery	GA	Report
6		Unique Pageviews		sum	Number of unique pageviews	Google Analytics	BigQuery	GA	Report
7		Reading time		secs	Average reading time of article	Chartbeat	LLD	Chartbeat	Report
8		Reading depth		int	Reading depth 25,50,75,100%	Chartbeat	LLD	Chartbeat	Report
9		Publishing date	GA Custom Dimension	date	Date of publishing	Google Analytics	BigQuery	Datalayer	Report
10		Age group				Appnexus	Segment feed	Appnexus	Audience
11		Education				Appnexus	Segment feed	Appnexus	Audience
12		Household income				Appnexus	Segment feed	Appnexus	Audience
13		Gender				Appnexus	Segment feed	Appnexus	Audience
14		Professional status				Appnexus	Segment feed	Appnexus	Audience
15		Company size				Appnexus	Segment feed	Appnexus	Background
19		City		int		Appnexus	Standard feed	Appnexus	Background
20		Region		int		Appnexus	Standard feed	Appnexus	Background
21		Device type		int		Appnexus	Standard feed	Appnexus	Background
22		Traffic	GA Source/medium		Traffic source	Google Analytics	BigQuery	GA	Report
23		URL			URL of article				Report
	23.1	URL		text		Google Analytics	BigQuery	GA	Report
	23.2	URL	Line item reporting label	text	Full url of the article	Appnexus	Service	User	Report
24		Author	GA Custom Dimension	text	Article author	Google Analytics	BigQuery	Datalayer	Report
25		Impressions		int	Impressions	Appnexus	Event type	Appnexus	
26		Seller	Line item reporting label	text	Name of the seller	Appnexus	Service	User	
27		Social Traffic	GA hits social referral		Social traffic source	Google Analytics	BigQuery	GA	Report
28		Network Domain	GA network domain	text	Domain of user	Google Analytics	BigQuery	GA	Report

Appendix 2 Power BI final data model

