Scientific Annals of Economics and Business
67 (SI), 2020, 57-67
DOI: 10.47743/saeb-2020-0025

Automated vs Manual Content Analysis – A Retrospective Look

Ionuț-Daniel Anastasiei[*], Mircea Radu Georgescu[**]

Abstract

Content Analysis, which is a part of qualitative analysis, has mainly been studied in scientific articles from health and medicine domains. With the emerge of social networks, there are new opportunities for content analysis, which can be used to analyse user generated content, from various sources. Nevertheless, the companies are investing millions of dollars in content analysis, which is often known as sentiment analysis. The discussion in this article helps to understand the main concepts of content analysis for those interested in the domain of qualitative analysis, with the help of automated and manual qualitative research. The overall conclusion is that automated qualitative analysis is dependent on how accurate is the tool used and this feature can be checked with the help of manual qualitative analysis.

Keywords: content analysis; automated qualitative research; manual qualitative research; sentiment analysis.

JEL classification: O33; O35.

1. INTRODUCTION

Nowadays, there are almost 3 billion people who regularly use at least one type of social network, regularly (Clement, 2020). This fact caused the appear of new emerging technologies that rely on the data created by users. All this data is known as *user generated data* and contributes a lot to the storage size of Big Data (Berthon *et al.*, 2015). Beyond selfies, personal biographies, shares and likes, there is a significant increase of opportunities for businesses to collect all this data, especially for marketing and economic purposes.

One of this opportunities is consisted of a new technology named *sentiment analysis*, which is based on the use of NLP (Natural Language Processing), text analysis and other qualitative analysis automated technologies. Currently, there are three approaches to sentiment analysis that can be grouped into four main categories: keyword spotting,

[*] Doctoral School of Economics and Business Administration, Alexandru Ioan Cuza University of Iași, Romania; e-mail: *ionut.anastasiei@feaa.uaic.ro*.

[**] Department of Accounting, Economic Informatics and Statistics, Faculty of Economics and Business Administration, Alexandru Ioan Cuza University of Iași, Romania; e-mail: *mirceag@uaic.ro* (corresponding author).

statistical methods, lexical affinity and concept-based techniques. The method of keword spotting technique was not used in this study because it is weak in two areas: it can't reliably recognize affect-negated words, and it relies on surface features. The statistical methods are based on Bayesian inference and support vector machines which is popular for affect text classification, the lexical affinity approach only detects sentiment based on the obvious affect words, which use a probable „affinity" to detect emotions and the last one, concept-based techniques that was used in this study, consisted of the analysis of key words, such as positive, negative, neutral words, which can help categorize a particular review based on the sentiment expressed by the user to a product or a company (Cambria *et al.*, 2013).

With the use of sentiment analysis, companies have the possibility to evaluate the social media health of a brand and to compare it to the business's competitor's brand which allows to capture trends and positive or negative tendencies (Zitnik, 2012). Trying to listen to the online voices of such users and gaining the understanding of their problems or feedbacks with the product, it can ultimately improve a brand's fame and help drive further research and strategy at both microeconomic and macroeconomic levels (Zitnik, 2012). Therefore, due to the importance of sentiment analysis, startup companies are looking too quickly and effectively to leverage social media (Zitnik, 2012). Also, another important role of this analysis for businesses is that it allows for actionable information to be extracted from any source, especially from online sources, providing competitive intelligence which allows for a better understanding of a business' environments (Fan and Gordon, 2014; Jayasanka *et al.*, 2014; Ribarsky *et al.*, 2014; Atwebembire, 2015).

Companies are now using a growing range of tools to develop this 360-degree vision, including social media monitoring tools to gather what customers say on platforms like Facebook and Twitter, predictive analytics tools for determine what customers can further research or buy, customer relationship management, and marketing automation software (Murphy, 2018). Customers are influenced by reviews and approvals so much that 85% read reviews online before a purchase, while 63% are more likely to buy from a site with user reviews (Murphy, 2018).

The development of a conceptual model for creating a database containing an overview of customers (360 degrees customer view) as a real alternative to existing CRM solutions can be seen in Figure no. 1.

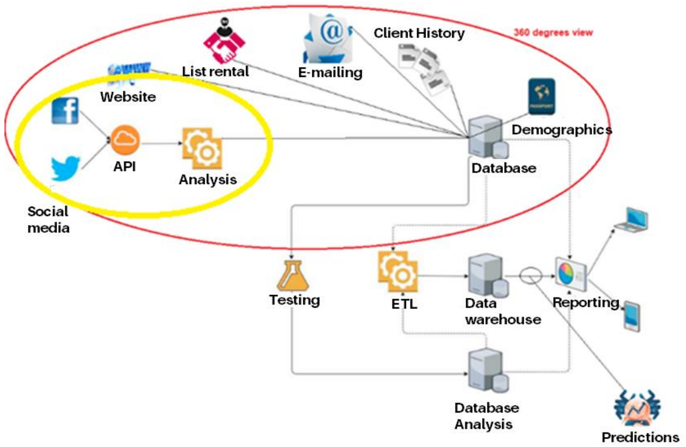

Figure no. 1 – Sentiment analysis in a 360 degrees view framework

Deriving from the model shown in Figure no. 1, the main purpose of our article is to demonstrate that automatic qualitative analysis is as reliable as manual qualitative analysis, if the recommendations and elements of good practice are followed by all those interested in this domain.

2. LITTERATURE REVIEW

In short, content or text analysis is a scientific research used to identify patterns in the way customers are talking about a product or a company (Luo, 2019). In order to effectively use the content analysis, the continue need of collecting data from texts is mandatory and can be applied to books, papers, web content, social media posts, reviews, speeches, photos or videos. Content analysis can be both quantitative, focused on counting and measurement, and qualitative, focused on interpretation and understanding. In both types, words, themes and concepts are categorized or "coded" in the texts and then the results are analysed (Luo, 2019).

Researchers use content analysis to find out the purposes, messages, and effects of communication content. They can also make inferences about the producers and the audience of the texts they analyse. Content analysis can be used to quantify the appearance of certain words, phrases, topics, or concepts in a set of historical or contemporary texts. Qualitative data are not exclusively in the field of qualitative research. Rather, the term can refer to anything that is not quantitative or rendered in numerical form. Many quantitative studies include open-ended survey questions, semi-structured interviews or other forms of qualitative data. What distinguishes data from a quantitative study from those generated in a qualitatively designed study is a set of assumptions, principles, and even values about truth and reality (Thorne, 2000).

An interesting part of qualitative analysis is the analysis of User-Generated Content (UGC), known as *user generated content* is any form of content, such as images, videos, text, and audio, that has been posted by users on online platforms, such as social media (Berthon *et al.*, 2015). At the same time, opinion monitoring, offers companies the chance to create virtual spaces where users can post and share opinions about the products or services offered. In an attempt to quantify the impact of these opinions on others, companies are trying to expand their means of finding information, such as giving customers a chance to discuss products in a forum (Fondevila-Gascon *et al.*, 2012). Unlike quantitative research methods, the analysis of qualitative content is not related to any science and there are fewer rules to follow. Therefore, the risk of confusion in matters relating to philosophical concepts and discussions is reduced. Throughout the process, the researcher must respect a qualitative perspective, and the main problem is to obtain the rigor and reliability that make the results as accurate as possible. There is a clash of the scientific world about the reliability, but it is agreed that generally, 80% is an acceptable margin of validity (Bengtsson, 2016).

According to an important scientific paper (Bengtsson, 2016) on content analysis, there are five main categories of planning a comprehensive analysis: the purpose, units of analysis, choosing the method of data collection, choosing the method of analysis and practical implications. In many studies, content analysis has been used to analyze answers to open-ended questions in questionnaires (Kyngäs *et al.*, 2011). However, the contents are often brief and that causes a difficulty in using the content analysis effectively; reduction, grouping, and abstraction require rich data. In addition, reliability has often been difficult to evaluate, often because articles or scientific research have mainly focused on reporting the analysis of quantitative rather than qualitative data obtained in the study. This means that subjectivity is a very common thing. However, if researchers use content analysis to analyze answers to open-ended questions, like in our case the analysis of the reviews, they should provide an adequate description so that readers are able to readily evaluate its reliability (Elo *et al.*, 2014).

In order to conduct a qualitative content analysis, self-criticism and good analytical skills are required. Any qualitative analysis should include continuous reflection and self-criticism by the researcher from the beginning of the study (Pyett, 2003; Thomas and Magilvy, 2011). To avoid subjectivity, one study (Kyngäs *et al.*, 2011) stated that data are most often analyzed by one researcher, which is a common thing for novices, especially when using inductive content analysis. If there is a case, the reliability of the analysis can be confirmed by checking for the representativeness of the data (Thomas and Magilvy, 2011).

The trend of comparison between manual and automatic content analysis increased in the last years in various domains. For example, in one article (Alla *et al.*, 2018) a systematic review was conducted to state if word frequency really is useful or not and if it may assist understanding and facilitating evidence-informed policy more broadly. Also, according to the same article, automated content analysis provides potential analytical tools that could be utilized at various stages of policy development to examine, and facilitate, the diffusion of innovation into policy (Alla *et al.*, 2018).

3. RESEARCH METHODOLOGY

According to a well-known study in the literature (Kyngäs *et al.*, 2011), the most used method in content analysis studies is purposive sampling. This kind of analysis is suitable for qualitative studies where the researcher is interested in informants who have the best knowledge concerning the research topic, in this case, the users of Kindle product from Amazon. The methodology used in the comparison between manual and content analysis, from the end of the analysis, was also conducted in another study that identified same false-negative and true negative results and it also concluded that automatic content analysis is far more useful, even if the accuracy is 100% accurate (Byrne *et al.*, 2013). When using purposeful sampling, decisions need to be made about who or what is sampled, what form the sampling should take, and how many people or sites need to be sampled (Creswell, 2013). The data used in this case study is consisted of 200 opinions of some Amazon customers regarding the Kindle product, which is the most sold product of Amazon from the eBook products. The reviews were obtained from Kaggle (a subsidiary of Google LLC, is an online community of data scientists and machine learning practitioners). Obviously, thousands of reviews can be analyzed, but they were chosen precisely to comply with the analysis units.

The methodology used in this article is divided into three main categories:

1. Export reviews on the site, in Excel.
2. Sort the sentiment used automatically in the R language.
3. Analyze the content of positive and negative comments using Maxqda, a qualitative analysis software program.

The export of the reviews was made in Excel precisely to facilitate their import in R studio, the program used for their automatic interpretation. The 200 reviews were unioned into a column. Sorting the feeling of the reviews was automatically analyzed in R Studio, with the help of the R programming language. Both for this and for the creation of the wordclouds, several libraries were used.

Table no. 1 – Libraries used in R Studio

```
library(RSentiment)
library(rJava)
library(readtext)
library(RColorBrewer)
library(wordcloud)
library(wordcloud2)
library(NLP)
library(tm)
library(SnowballC)
```

```
# read all reviews
REVIEWS <- read.csv("review1.csv", stringsAsFactors=FALSE)

#calculate sentiments
calculate_total_presence_sentiment(REVIEWS)
SENTIMENT_calculat <- calculate_sentiment(REVIEWS)

# export sentiment
write.table(SENTIMENT_calculat, file ="my_data.csv")
```

Figure no. 2 – R Code used for processing the reviews

At the same time, the MAXQDA qualitative analysis program was used to detect errors caused by automation, by doing a manual content analysis. This program can be used for any type of qualitative analysis, including interviews or focus groups. In this context, it was used to interpret user-generated content.

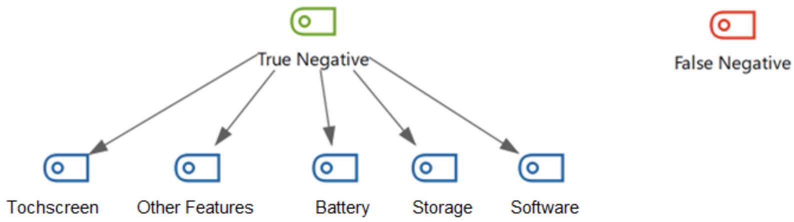

Figure no. 3 – Coding that was used for the qualitative analysis

All reviews identified as positive or negative by R Studio were taken in turn and, for greater accuracy and analysis purposes, were manually coded in Maxqda. For positive sentiment, there have been used the same codes.

After the export of the reviews analyzed in R Studio was extracted, it was used as a data source in Maxqda. The results were split in four main categories, as follows:

- True negative, which were the results interpreted as negative in both manual and automated analysis;
- False negative, which were the results interpreted as positive in the manual and negative in the automated analysis;
- True positive, which were the results interpreted as positive in both manual and automated analysis;
- False positive, which were the results interpreted as negative in the manual and negative in the automated analysis.

4. RESULTS AND DISSCUSSIONS

The results of the automatic interpretation of the R language code use for research can be seen in Table no. 2. The results obtained from all 200 reviews, 22 are treated as negative and 142 as positive.

Table no. 2 – The results of automated content analysis

[1,]	"Sarcasm"	"Negative"	"Very Negative"	"Neutral"	"Positive"	"Very Positive"
[2,]	"1"	"19"	"2"	"36"	"41"	"101"

Word clouds were also created in the R language, to give an idea of the words used by customers. At the same time, in Table no. 3 it can be seen most of the words used.

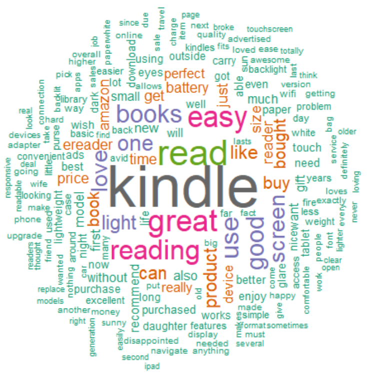

Figure no. 4 – The word-cloud of reviews

Table no. 3 – Frequency of words

	word	freq
kindle	kindle	119
read	read	65
great	great	57
easy	easy	53
reading	reading	46
use	use	44
books	books	42
good	good	40
love	love	37
screen	screen	36

Furthermore, word-clouds were made for reviews interpreted as negative. There can be easily observed that there was a limit of the content analysis: the word *great* appears, although its definition is known to be a positive word. Therefore, manual analysis is also required. In addition to this word, it can be observed other negative words, such as *small*, *dark*, *long* or general words, such as *battery*, *light* or *screen*.

Figure no. 5 – Negative reviews word cloud

Figure no. 6 – Positive reviews word cloud

Table no. 4 – Negative words frequency

	word	freq
kindle	kindle	13
easy	easy	8
light	light	8
one	one	8
great	great	7
reading	reading	6
read	read	5
ads	ads	4
battery	battery	4
first	first	4

Table no. 5 – Positive words frequency

	word	freq
kindle	kindle	74
great	great	42
easy	easy	37
use	use	35
reading	reading	33
read	read	30
good	good	28
books	books	23
like	like	23
love	love	20

In Table no. 4, the most used word is kindle, which is the product itself, then there come other neutral words like easy, read, first. The problem identified in the results is that the word great is identified in seven reviews. In the further steps of the analysis, all reviews with neutral or positive words, identified as negative reviews will be classified as false negative.

In Figure no. 6 and Table no. 5, the positive interpretation of the results is far better than the negative one. Of course, the word *great* appears right after the product, this states that at least one quarter of the reviews were positive. Furthermore, the results obtained in R Studio were exported in Excel, each by its own category, positive or negative. From this point of the analysis, a parallel qualitative analysis of the reviews was conducted, starting with negative reviews and then with the positive ones. Also, in Maxqda, the results were split in a Boolean format, true or false negative reviews.

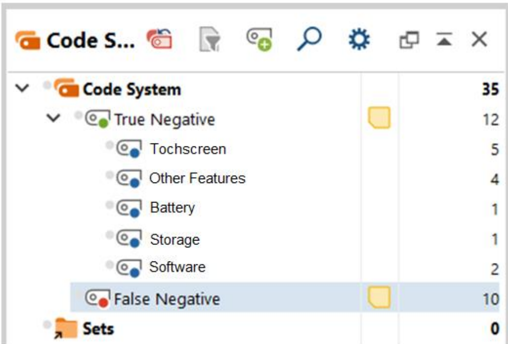

Figure no. 7 – Coding system of the negative reviews

Table no. 6 – Frequency and percentages of the negative codes

Parent code	Code	Cod. seg. (all do...	Cod. seg. (activ...	% Cod. seg. (...	% Cod. seg. (ac...	Documents
	True Negative	12	0	34.29	0.00	1
	False Negative	10	0	28.57	0.00	1
True Negative	Tochscreen	5	0	14.29	0.00	1
True Negative	Other features	4	0	11.43	0.00	1
True Negative	Software	2	0	5.71	0.00	1
True Negative	Storage	1	0	2.86	0.00	1
True Negative	Battery	1	0	2.86	0.00	1

In the current analysis, with the help of Maxqda, there were identified 12 posts that were indeed negative, and 10 were identified as false negative. This proves that R Studio has an accuracy of about 55% when it comes to identification of negative reviews. As an initial conclusion, most of the complaints were related to the screen of Kindle or to other functionalities of the product.

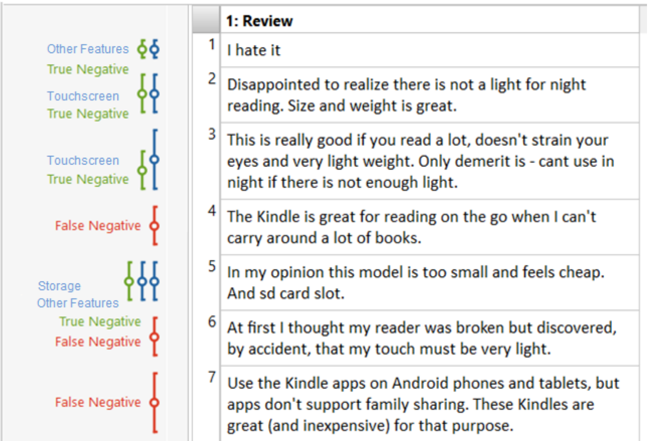

Figure no. 8 – Screenshot of content analysis in Maxqda

A screenshot of the coding program can be seen in Figure no. 8, where the user feedback on the product can be easily identified and, depending on the content, a code can be applied via the drag & drop function of Maxqda. Furthermore, one review can have more than one code, for example, the review with the identification number as three, present in Figure no. 8, has the code as true negative, with the addon of the touchscreen functionality. This fact boosts the improvement of the quality of the content analysis. On the other hand, in the review number six, the initial negative sentiment was changed to false negative. At first look, it indeed appeared as a problem with the touchscreen, but it was just the end user which did not discovered the light settings of the product.

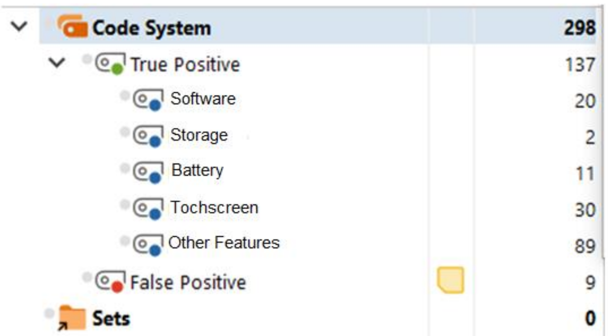

Figure no. 9 – Coding system of the positive reviews

For the positive reviews, a code system was developed in Maxqda that was divided into two main categories: true positive, and false positive.

Table no. 7 – Frequency and percentages of the positive codes

Parent code	Code	Cod. seg. (all do...	Cod. seg. (activ...	% Cod. seg. (...	% Cod. seg. (ac...	Documents
	True Positive	137	0	45.97	0.00	1
True Positive	Other features	89	0	29.87	0.00	1
True Positive	Tochscreen	30	0	10.07	0.00	1
True Positive	Software	20	0	6.71	0.00	1
True Positive	Battery	11	0	3.69	0.00	1
	False Positive	9	0	3.02	0.00	1
True Positive	Storage	2	0	0.67	0.00	1

In Table no. 7, the analysis identified only nine comments as being negative, which states that R Studio had an accuracy of 90% in terms of the identification of positive comments, compared with the 55% from the initial content analysis of negative reviews. At the same time, most of the positive comments were about the functionalities of the product, the screen and the software.

The results of both automated and manual content analysis can be seen in Table no. 8, where number the synthesis of the results was obtained from R Studio and Maxqda.

Table no. 8 – Results of the content analysis (manual vs automated)

Reviews	Results obtained in R Studio	Results obtained in Maxqda
Negative reviews	22	12
Positive reviews	142	133

5. CONCLUSIONS AND FUTURE STUDIES

In conclusion, content analysis can be automated in order to obtain the interpretation of sentiment expressed in reviews, but manual interpretation is also necessary for greater accuracy only if the time saved is worth or not. Qualitative data - derived, for example, from interviews, open-ended questions and written images - are expressed in words so can be easily interpreted. Consequently, the researcher cannot use statistical analysis to give meaning to the data and therefore needs other methods of analysis. Content analysis is one of these methods. Our paper shows how the general principles of the method can be used and how the validity and reliability of the whole process can be increased. The positive results interpreted with the automated analysis has an acceptable accuracy, which is great, but when it comes to identify the negative results, the capabilities are limited and the concept-based techniques needs a lot of improvement, with the integration of a qualitative analysis tool sensible to lexical affinity.

Although there are both advantages and disadvantages to performing a content analysis, it is an easy-to-understand analysis process that can be emulated even by those new to this domain. We hope that this paper can help others to give meaning to textual data and maintain the quality of the analysis.

In the following studies, the R language algorithm can be improved by adding more keywords, precisely to facilitate interpretation and improve accuracy or even with using other programming languages, like python. At the same time, the number of reviews analyzed can be much higher, reaching even thousands. This would bring it closer to a real case because, in general, Amazon-sized companies have millions of customers, and Customer Relationship Management systems are huge, with investments in the hundreds of millions of dollars.

ORCID

Mircea Radu Georgescu https://orcid.org/0000-0002-7022-3715

References

Alla, K., Oprescu, F., Hall, W. D., Whiteford, H. A., Head, B. W., and Meurk, C. S., 2018. Can automated content analysis be used to assess and improve the use of evidence in mental health policy? A systematic review. *Systematic Reviews, 7*(1), 194. http://dx.doi.org/10.1186/s13643-018-0853-z

Atwebembire, M., 2015. *Sentiment Analysis: Uses in Business.* (BCom), University of Cape Town. (ATWMAR001)

Bengtsson, M., 2016. How to plan and perform a qualitative study using content analysis. *NursingPlus Open, 2*, 8-14. http://dx.doi.org/10.1016/j.npls.2016.01.001

Berthon, P., Pitt, L., Kietzmann, J., and McCarthy, I. P., 2015. CGIP: Managing Consumer-Generated Intellectual Property. *California Management Review, 57*(4), 43-62. http://dx.doi.org/10.1525/cmr.2015.57.4.43

Byrne, M. D., Jordan, T. R., and Welle, T., 2013. Comparison of manual versus automated data collection method for an evidence-based nursing practice study. *Applied Clinical Informatics, 4*(1), 61-74. http://dx.doi.org/10.4338/ACI-2012-09-RA-0037

Cambria, E., Schuller, B., Xia, Y., and Havasi, C., 2013. New avenues in opinion mining and sentiment analysis. *IEEE Intelligent Systems, 28*(2), 15-21. http://dx.doi.org/10.1109/MIS.2013.30

Clement, J., 2020. Number of social network users worldwide from 2010 to 2021. *Statista*. https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/.

Creswell, J. W., 2013. Qualitative inquiry and research design: Choosing among five approaches. Thousand Oaks, CA: ed. *Sage (Atlanta, Ga.)*.

Elo, S., Kääriäinen, M., Kanste, O., Pölkki, T., Utriainen, K., and Kyngäs, H., 2014. Qualitative Content Analysis: A Focus on Trustworthiness. *SAGE Open, 4*(1), 2158244014522633. http://dx.doi.org/10.1177/2158244014522633

Fan, W., and Gordon, M. D., 2014. The power of social media analytics. *Communications of the ACM, 57*(6), 74-81. http://dx.doi.org/10.1145/2602574

Fondevila-Gascon, J. F., Carreras-Alcalde, M., and Del Olmo-Arriaga, J. L., 2012. Impacto de las tecnologias de la informacion y la comunicacion en la eleccion de Universidad: El caso de Internet y las redes sociales. *Edutec. Revista Electrónica De Tecnología Educativa, 39*(March), a195. http://dx.doi.org/10.21556/edutec.2012.39.378

Jayasanka, R. A. S. C., Madhushani, E.R., Marcus, I. A. A. U., Aberathne, I. and Premaratne, S.C. 2014. Sentiment analysis for social media. *International Journal of Advanced Research in Computer Science and Software Engineering, 3*(7), 216-222.

Kyngäs, H., Elo, S., Polkki, T., Kaariainen, M., and Kanste, O., 2011. The use of content analysis in Finnish nursing science research. *Hoitotiede, 23*(2), 138-148.

Luo, A., 2019. What is content analysis and how can you use it in your research. *Scribbr*. https://www.scribbr.com/methodology/content-analysis/.

Murphy, R., 2018. Local Consumer Review Survey | Online Reviews Statistics & Trends. *Brightlocal*. https://www.brightlocal.com/research/local-consumer-review-survey/.

Pyett, P. M., 2003. Validation of qualitative research in the "real world". *Qualitative Health Research, 13*(8), 1170-1179. http://dx.doi.org/10.1177/1049732303255686

Ribarsky, W., Xiaoyu Wang, D., and Dou, W., 2014. Social media analytics for competitive advantage. *Computers & Graphics, 38*, 328-331. http://dx.doi.org/10.1016/j.cag.2013.11.003

Thomas, E., and Magilvy, J. K., 2011. Qualitative rigor or research validity in qualitative research. *Journal for Specialists in Pediatric Nursing, 16*(2), 151-155. http://dx.doi.org/10.1111/j.1744-6155.2011.00283.x

Thorne, S., 2000. Data analysis in qualitative research. *EBN notebook*. https://ebn.bmj.com/content/ebnurs/3/3/68.full.pdf

Zitnik, M., 2012. Using sentiment analysis to improve business operations. XRDS: Crossroads. *18*(4), 42-43. http://dx.doi.org/10.1145/2173637.2173655

Copyright