Quantitative Analysis of Qualitative Data: Using Voyant Tools to Investigate the Sales-Marketing Interface

Gabor Hetenyi[1] (iD), Attila Dr. Lengyel[2] (iD), Magdolna Dr. Szilasi[3] (iD)

[1]Károly Ihrig Doctoral School of Management and Business, University of Debrecen (Hungary)
[2]Department of Tourism and Catering, John von Neumann University, GK (Hungary)
[3]Doctoral School of Health Sciences, University of Debrecen (Hungary)

gaborhetenyi@yahoo.com, guszfraba@gmail.com, drmagdolnaszilasi@gmail.com

Abstract:

Purpose: The present study aims to give a short introduction into the possibilities offered by Voyant Tools to quantitatively explore qualitative data on the Sales-Marketing Interface (SMI).

Design/methodology/approach: The study is exploratory in nature. The sample consists of sales and marketing employees of six manufacturing companies. Answers to three open-ended questions were analysed quantitatively and visualised in various ways using the online toolset of Voyant Tools. We experimented with four different tools out of the twenty-four offered by Voyant Tools. These tools were: Cyrrus tool, Correlation tool, Topics tool and Scatter plot tool. All four tools that were tested on the data have scalable parameters. Various settings were tested to demonstrate how input conditions influence modelling of the textual data.

Findings: Positive aspects of Voyant Tools: It was demonstrated that the four selected text analysis tools can yield valuable information depicted in the form of attractive visualisation formats. Negative aspects of Voyant Tool: It is also highlighted how rushed conclusions can be arrived at by falsely interpreting the visualised data. Limited aspects of Voyant Tools: It is shown how setting different input parameters can affect results. Out of the four examined tools the Scatter plot tool offering an analysis and modelling method based on t-SNE (t-Distributed Stochastic Neighbour Embedding) proved to yield the most complex information about the text.

Research limitations/implications: As the study aimed to be exploratory a sample of convenience was used to collect qualitative data. Although quantitative methods can be invaluable tools of preliminary analysis and hypothesis adjustment in the processing of qualitative data, their results should always be checked against the traditional content analysis techniques which are more sensitive to the complex structure of semantic units. These quantitative techniques are to help early exploration of textual data.

Practical implications: Professional implications: Managerial implications might be connected to the fact that in a fast changing global business environment managers and corporate decision makers in general might find the attractive visualisation outputs of Voyant Tool easy to analyse and interpret various aspects of business. Academic implications: As Voyant Tools is an open source, free online sofware not even requiring regsitration and at the same time has an impressive array of sophisticated statistical tools, it might be a cost-effective way of analysing qualitative data for low budget academic users.

Originality/value: As there is virtually no earlier literature on how quantitative data visualisation techniques can be used in marketing research, especially in the analysis of the SMI, utilisation possibilities

of Voyant Tools and other quantitative data analysis and visualisation software for handling qualitative data is definitely a worthwhile area for further research.

To cite this article:

1. Introduction

Qualitative research has evolved into an accepted and invaluable research method since it was first advocated by German sociologists Max Webber and Georg Simmel (Dey, 2003; Gummesson, 2005; Lapan, Quartaroli & Riemer, 2011; Mayer, 2015; Silverman, 2016; Flick, 2018). The false dichotomy between the two research methods has by now been resolved as there are more and more studies employing mixed research methodologies that take advantage of both qualitative and quantitative data collecting and data processing techniques (Molina-Azorin, Bergh, Corley & Ketchen, 2017; Bryman, 2017; Braun, Clarke, V., Hayfield, N., & Terry, 2019).

Marketing research can be traced back to the first part of the 19th century when the first market-focused data collections took place in the USA (Lockley, 1950). Since then marketing research has gone a long way and today there have accumulated an abundance of literature on both qualitative (Carson, Gilmore, Perry & Gronhaug, 2001; Belk, 2007) and quantitative research (Franses & Paap, 2001; Müller, Boda, Ráthonyi, Ráthonyi-Odor, Barcsák, Könyves et al., 2016; Lipowski, Pastuszak & Bondos, 2018) used in this field.

Qualitative research methods have been used in marketing research for decades (Bellenger, Bernhardt & Goldstucker, 2011; Wilson, 2018). However, analysis of qualitative data, usually in the form of the transcripts of various interview techniques or answers to open-ended questions in self-reported questionnaires, is usually restricted to quoting passages, typical sample answers or themes that emerge during some form of content analysis (Hsieh & Shannon, 2005).

Quantitative analysis of qualitative data (Young, 1981), other than word clouds, is extremely scarce in marketing research and not frequently used in other social sciences either (Bernard & Ryan, 1998). On the one hand, it is understandable, as the transformation of a coherent text, which is a complex, multi-layered information source with contextualised meaning, into smaller meaning units necessarily entails some loss of information (Krippendorff, 2018). It might be tempting to think that qualitative analysis of qualitative data does not result in information loss, however, as Bernard aptly points out. Quantitative analysis involves reducing people (as observed directly or through their texts) to numbers, while qualitative analysis involves reducing people to words (Bernard, 1996: page 10). Obviously, the validity and generalisability of the results depend on the research design including sampling methods as well as the form of analysis applied to the collected data.

In order to be able to apply quantitative statistical methods with qualitative data the answers of respondents are typically coded. Coding can be as complex as to include sixteen steps (Assarroudi, Heshmati-Nabavi, Armat, Ebadi & Vaismoradi, 2018). At the end of the coding process longer meaning units are reduced to one word. These one-word codes can then be analysed as categorical data using quantitative statistical methods. However, there are statistical methods, such as the latent Dirichlet allocation used for topic modelling (Jacobi, Van Atteveldt & Welbers, 2016; Toubia, Iyengar, Bunnell & Lemaire, 2019) to mention one, which can be used without any previous coding.

Voyant Tools is a web-based, free, open source text analysis software package that offers versatile and sophisticated text manipulation capabilities useful for both the beginner and advanced humanities scholar (Welsh, 2014; Uboldi &

Caviglia, 2015; Bradley, 2018; Miller, 2018). It has already been used as a quantitative text analysis tool in several peer-reviewed articles (Steiner, Agosti, Sweetnam, Hillemann, Orio, Ponchia et al., 2014; Clouder & King, 2015; Williams, Inversini, Buhalis & Ferdinand, 2015; Zahedzadeh, 2017). Data visualisation in social sciences research is an under-researched area (Uboldi & Caviglia, 2015).

Via the analysis of data on SMI, the present paper demonstrates how Voyant Tools can be used to quantitatively analyse qualitative data. The harmonious, constructive and efficient cooperation between the Sales and Marketing (SM) departments are considered a key element of customer satisfaction and strategic success in a fast-changing global market. The Sales-Marketing Interface (SMI) can be burdened with various conflicts and dysfunctions which (Malshe, Friend, Al-Khatib, Al-Habib & Al-Torkistani, 2017; Cometto, Labadie & Palacios, 2017), if not investigated and properly managed, can undermine the overall performance of the company. Thus, it is of utmost importance to obtain a clear picture of the state of SMI and its possible problems. Even though the optimization of the SMI is obviously a crucially important challenge, the SMI is a seriously under researched area within business research and the application of quantitative techniques to qualitative data in connection with SMI has not been researched at all. Previous studies on the SMI typically applied qualitative data collection and processing methods such as personal interviews of sales and marketing managers with summaries of the main findings, but no coding of text (Matthyssens & Johnston, 2006), minimal coding of interview data (Hughes, Le Bon & Malshe, 2012) or a detailed and rigorous coding process (Malshe & Al-Khatib, 2017).

2. Sample and Methods

As it is an exploratory study a sample of convenience was used. Six different manufacturing companies (number of employees ≥ 250) were involved in the data collection process. The main criterium of qualifying into the research was the presence of a separate sales and marketing department within the company. Data collection was conducted via a self-reported online questionnaire which contained three open-ended questions. The link to the questionnaire was emailed to the Human Resources managers of the six companies and were forwarded to the SM employees by them. The date was gathered during a two-week period in March 2019. Out of the 352 questionnaires sent out to potential respondents we received 124 fully completed ones which served as the basis for our analysis. 75 of them were marketing employees and 49 sales employees. As there were Hungarian, Austrian, German and Austrian companies involved the questionnaires were distributed in three languages (German, Hungarian, English). As the first step in processing the data the returned questionnaires filled out in German or Hungarian were translated by a qualified translator into English. Respondents had to answer the following three questions:

1. Please describe your daily tasks in a few sentences.
2. What are the tasks of the other (sales or marketing) department?
3. How is sales-marketing cooperation managed in your company?

For limitations of space most method demonstrations are performed on the third question as it is the main focus of the analysis. As our survey contained only three questions and the number of completed questionnaires is small too, it was possible to compare the results of quantitative analysis carried out with the help of tools of Voyant and see how accurate quantitative results are. Obviously, Voyant Tools is especially useful with large textual data sets when content analysis methods are extremely time-consuming. Out of the twenty-four different text analysis tools this paper attempts to demonstrate the use of four.

2.1. Cyrrus Tool

It is a word cloud creation tool which positions the most frequent words centrally and in the biggest size in the cloud. It is possible to exclude words using the „Stop word„ function or specify the maximum number of words to be fetched from the corpus.

2.2. Correlation Tool

It allows the researcher to check which words tend to occur together within the text. Negative correlations signal words with an inverse occurrence pattern. In order to be able to perform Pearson correlation calculations the text is

divided into segments. The software examines how many times words appear in the various segments and the resulting numerical serves as the basis of the correlations. The significance level for each pair of words is also provided. Pearson correlation is typically applied with assumptions of normal distribution. However, several studies demonstrated that the Pearson correlation is robust enough to tolerate the violation of the above-mentioned typical assumption (Havlicek & Peterson, 1976; Fowler, 1987). Still, the results should be interpreted with caution.

2.3. Topics Tool

This tool uses a rather sophisticated algorithm called latent Dirichlet allocation (LDA). It is a topic model which assumes that words in the text belong to latent topics. It also assumes that there is a relatively small set of topics with a relatively small set of words used frequently by the topic. With the help of this tool term clusters and their distribution can be discovered. It is possible to set the number of topics to optimise modelling for.

2.4. Scatter Plot Tool

This is probably the most sophisticated tool among the text analysis tools of Voyant. The analysis functions of this tool include Principle component analysis, Correspondence analysis, document similarity check and t-SNE analysis. All four cluster plotting analyses uses algorithms that creates a 2 (or 3) dimensional representation of the data in a multidimensional space. The number of dimensions and the number of clusters to be created can be set by the analyst. Out of the four types of plotting methods t-SNE is discussed in this paper. t-SNE (t-Distributed Stochastic Neighbour Embedding) is a prize-winning method that can be applied especially well to high dimensional data sets such as qualitative textual data (Van Der Maaten & Hinton, 2008; Van Der Maaten, 2014). Cao and Wang define the method as follows, "*t-SNE tries to preserve local neighbourhood structure from high dimensional space in lowdimensional space by converting pairwise distances to pairwise joint distributions, and optimize low dimensional embeddings to match the high and low dimensional joint distributions.*" (Cao & Wang, 2017: page 1.)

There is a tuneable function of t-SNE in Voyant, the level of perplexity (0-100) which largely determines, what cluster model is plotted. If the data is very dense perplexity close to 100 might be the most suitable but with lower density data lower levels of perplexity will yield the best results, that is the most accurately identified clusters. The algorithm behind perplexity examines the "local" and "global" aspects of the data set, that is, it tries to determine the number of closest neighbours of each word (data points) or expressed differently, it can be "*measure of the effective number of neighbours*" (Van Der Maaten & Hinton, 2008: page 2582).

3. Results and Discussion

| Figure 1. Question 1 | Figure 2. Question 2 | Figure 3. Question 3 |

(Own editing using www.voyant.tools.org)

In Cyrrus tool there is a default "stop-word" list containing the most typical non-content words such as "the, and, but, etc". It was supplemented by other text-specific words of little significance such as "however, some, most, etc". The remaining words are mainly (92%) nouns. Words clouds are to be interpreted with caution, because they do not reflect collocations, co-occurrences or possible meaning variations. the word "management" in Figure 1 is a typical example as it can mean the board of leaders or the set of processes. However, some preliminary guesses can be made about these three qualitative data sets. The first question concerned the daily tasks of the SM employees. It

is apparent that "contact" "partners" and "management" are the most frequent terms. It suggests a typical daily work schedule of SM employees. As there were considerably more marketing department respondents than sales ones, it is not surprising that the term "marketing" is the most used in the answers to the first and the second question. As the first two questions were about the work done by SM employees, it is not hard to explain the considerable overlap between the word clouds in Figures 1 and 2. The third question aimed to gauge how the respondents thought the cooperation between SM was actually realised. The frequency of words "meetings, regular, common" signal the importance of face-to-face contact and sharing. It is interesting that sales and marketing appear with equal weight suggesting a possible drive to reach balance between them. It is also telling that the terms "organisation", "goals" and "company" appear with considerable weight for the first time. The underlying cause might be the realisation by both departments of the necessity of harmonising the SMI to foster organisational goals and benefit the company as a whole. Figure 4 shows some of the strongest correlations between words in the answers given to Question 3.

Obviously strong correlations can signal collocation of the pairs of words. In the above picture "regular" and "meetings" are collocated in the form of "regular meetings" in most segments (It was checked with the Collocates tool). Not all, because in that case the correlation would be one. The same applies to "weekly" and "meeting", "telephone" and "conference" and meetings (regular, weekly) seem to be a crucial factor in the optimisation of the SMI. Looking at the correlating pairs of words it seems apparent that the strongest correlations are present between words that refer to some form of communication (meetings, conference, communication, telephone).

Figure 5. shows three topics variations of the answers given to Question 3. It has to be noted that the LDA algorithm randomly assigns words (number can be set) to topics (number can be set) when it is started. Thus, each time the algorithm is run there will be slight differences in the results. Besides setting the number of topics and the number of words per topic to model the text on it is also possible to set the number of iterations for the algorithm. The default is 50, but the present results were obtained after 200 iterations. The more iterations are run, theoretically, the more accurate the topics will reflect clusters in the text.

Term 1	←	→	Term 2	Correlation (r)	Significance (p)
conference			telephone	0.89871705	0.0004067596
meetings			regular	0.8982954	0.00041335827
background			campaign	0.8849847	0.00066491275
campaign			materials	0.8849847	0.00066491275
effective			person	0.8849847	0.00066491275
goals			perform	0.875	0.0009160936
meetings			weekly	0.8728716	0.0009774802
campaign			organization	0.8728716	0.0009774802
common			communica…	0.8517522	0.0017599578
cooperation			help	0.8479983	0.0019358585

Figure 4. Correlations of words in the answers given to Question 3 (Own editing using www.voyant.tools.org)

Topic	Topic
meeting department managerial consensus provide stakeholders respond	department managerial consensus provide stakeholders conference respond
cooperation tasks joint company effective marketing organization	cooperation joint company marketing tasks effective organization
sales team marketing conflicts teams regular help	sales team marketing respect problem conflicts teams
company organization organizational units goals background materials	company organizational units organization background materials goals
meetings regular communication common sales telephone information	meetings regular communication common meeting sales telephone
management goals perform designing goal achieve corporate	goals perform corporate designing achieve management consultation
marketing email sales departments work newsletters department	marketing sales departments work newsletters department directives

Figure 5. Topics variations from the answers to Question 3 (Own editing using www.voyant.tools.org)

Each seven topics in Figure 5 contain all the words in the corpus, but only the top seven words are displayed. The order of the words is important. The first words in each topic contribute more to the topic than the other words, thus the seven words in each topics demonstrates an order of importance as well. There are some inferences that can be made from these topics. Both iterations yielded topics in which "meetings" appear as the organizing force of the topic. In the first iteration (on the left) it appears twice in first position. As the question was about how SM cooperation is realised and managed, it seems that common meetings for the two departments play an important role in optimising SMI. The fact that the word "cooperation" occurs in both topics sets in first position is probably attributable to the main focus of the question being the nature of cooperation between SM. The term "marketing" is in first position in both versions but "sales" is in first position in only one of them. This is a typical case to demonstrate why exercising caution with results is warranted. At first glance this occurrence pattern of "marketing" and "sales" might suggest that marketing is somewhat more important in these companies than sales. However, the numbers of respondents in the sample were considerable higher from the marketing departments than from sales, which might be the real cause of this occurrence pattern. The term "company" also appears in both sets as being the most important word of the topic together with other words such as "organization" or "goals". It might suggest that the harmonious relationship between SM significantly affects the company at large. At the same time, there might be a reverse interpretation, namely, the company and its goals have a significant influence on the relationship between SM. In order to come to valid conclusion a close consultation of the answers is unavoidable. Having consulted the answers, it is clear that both interpretations hold true at the same time. It has to be noted that the 124 answers to Question 3 represent a relatively small corpus, which can be read through in a relatively short time. In a different sampling scenario where there are thousands of answers to open-ended questions the topics tool of Voyant might become a much "heavier weapon" in the hand of the researcher.

The scatter plots (Figures 6, 7, 8) were created by the t-SNE tool.

The tf-idf (term frequency-inverse document frequency) weighting method was used for the analysis. It is an option that can be set by the analyst besides the other two methods "raw frequencies" and "relative frequencies". It is a method that determines how important a word is to a document and is largely dependent on how often a word appears in a document. As there is only one document in our case, the algorithm divides the corpus into 10 segments and examines word frequencies in each segment. As it was noted earlier t-NSE is an award-winning method and the cluster plots that it is able to create can encourage jumping to conclusions that might not at all be sound. There are several reasons for this. The two that we consider the most important is discussed here. These two factors are the level of perplexity and the number of iterations. Figures 6, 7 and 8 show the results of the t-SNE algorithm run at three different levels (5, 50, 100) of perplexity. All three scatter plots bellow (Figures 6, 7, 8) were obtained after 5000 iterations. In order to test how the model changes at different levels of perplexity it was necessary to keep the number of iterations constant. Looking at the three scatter plots it is apparent that perplexity level 50 yielded the most convergent result, that is, the various clusters are the clearest in Figure 7. Perplexity levels 5 and 100 (minimum and maximum levels respectively) resulted in less convergent clusters. It seems obvious that the level of optimal perplexity is largely dependent on the data set. There is no fixed level that can be suggested to be used in general and beginner users of t-SNE in Voyant might need considerable time to get the best results (Wattenberg, Viégas & Johnson, 2016). Attempts have already been made to automate the selection of the perplexity parameter and thus make analysis much easier for the novice user. (Cao & Wang, 2017).

The above results gained with altering the level of perplexity seem to support the claim of the inventors of the t-SNE method who said that the t-SNE method is fairly robust to changes in the level of perplexity (Van Der Maaten & Hinton, 2008; Van Der Maaten, 2014). There are no dramatic differences between the models of the three different levels of perplexity.

The number of iterations the tool will use to create the model can be set between 100-5000. If we take a look at Figures 9, 10, 11, 12 (100, 600, 900, 5000 iterations respectively) the same can be stated as about the level of perplexity earlier. There is no linear relationship between.
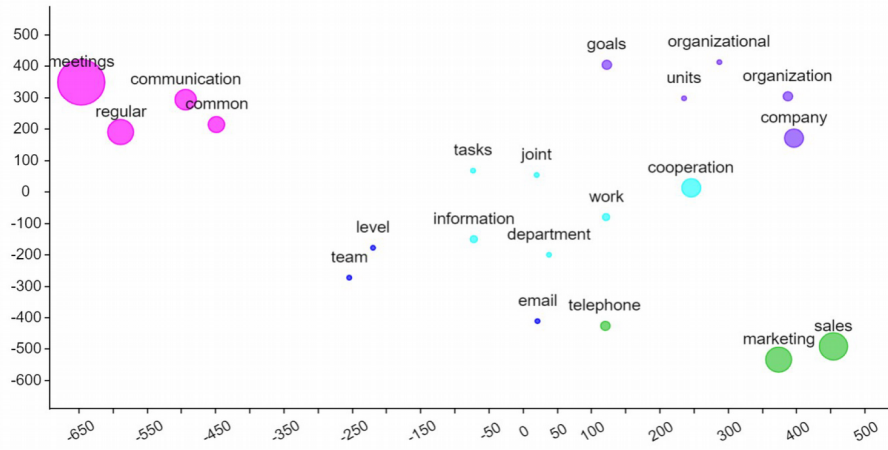
Figure 6. t-SNE generated clusters for the answers to Question 3 at perplexity level 5
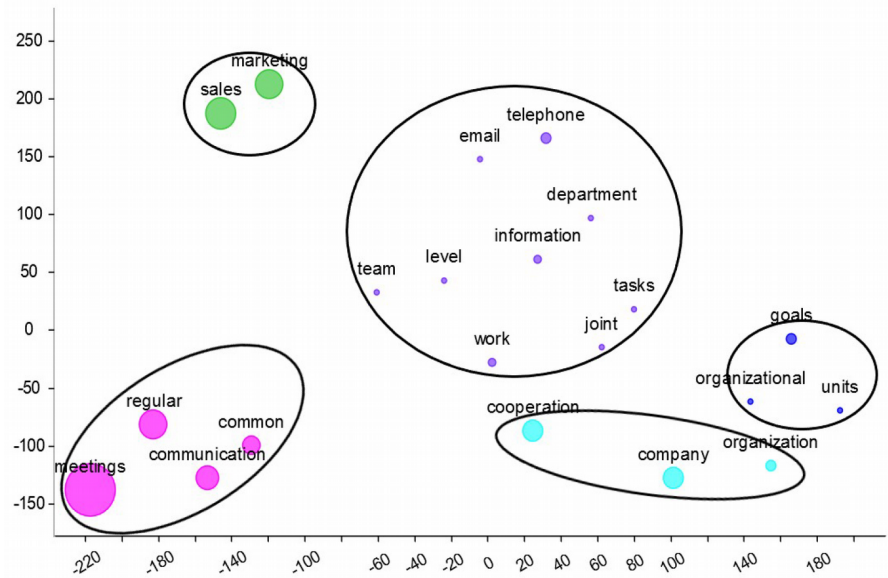(Own editing using www.voyant.tools.org)

Figure 7. t-SNE generated clusters for the answers to Question 3 at perplexity level 50
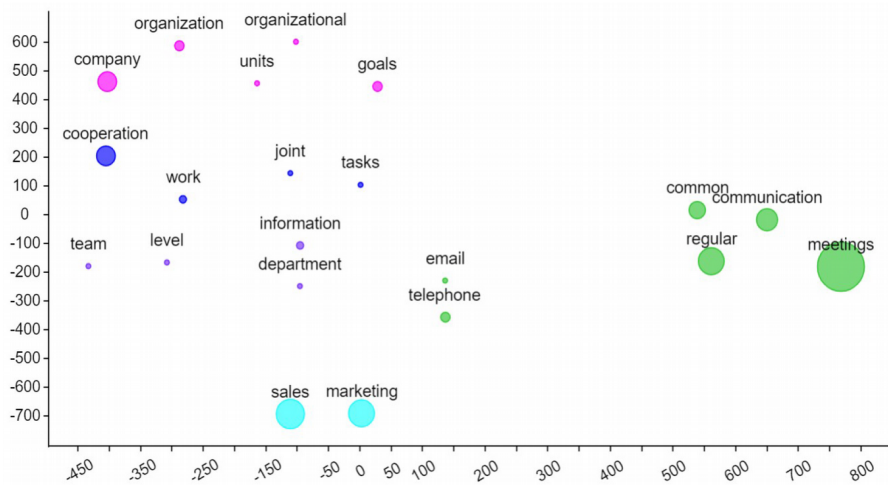(Own editing using www.voyant.tools.org)

Figure 8. t-SNE generated clusters for the answers to Question 3 at perplexity level 100
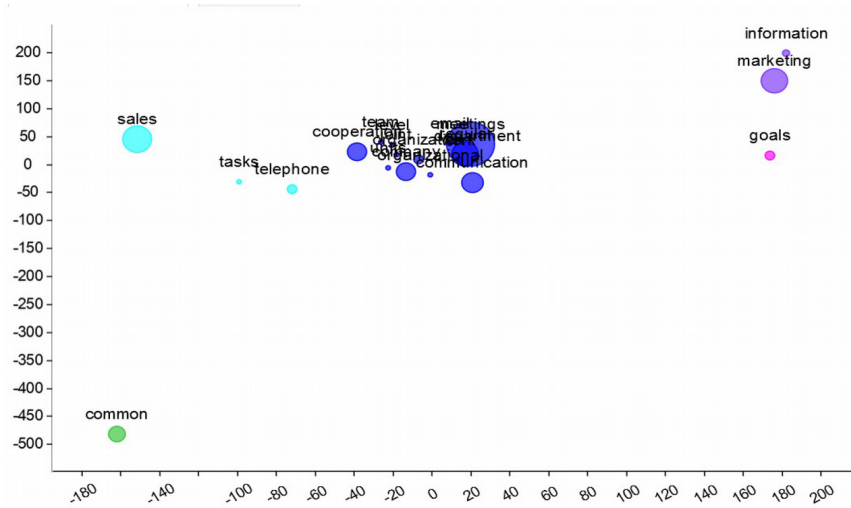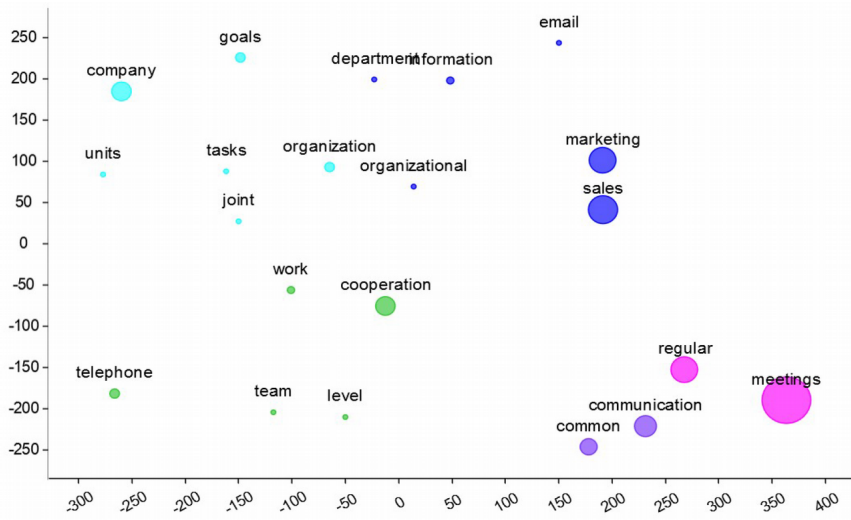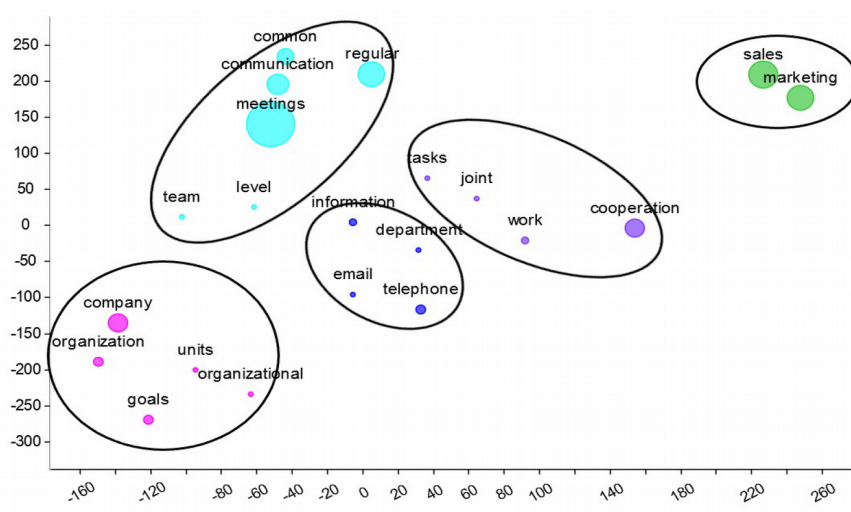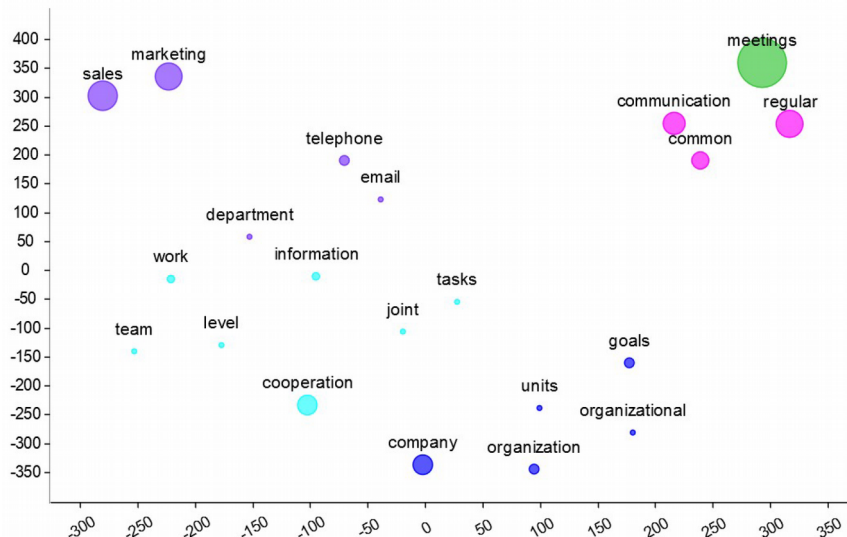(Own editing using www.voyant.tools.org)

Figure 9. t-SNE generated clusters for the answers to Question 3, 100 iterations
(Own editing using www.voyant.tools.org)

Figure 10. t-SNE generated clusters for the answers to Question 3, 600 iterations
(Own editing using www.voyant.tools.org)

Figure 11. t-SNE generated clusters for the answers to Question 3, 900 iterations
(Own editing using www.voyant.tools.org)

Figure 12. t-SNE generated clusters for the answers to Question 3, 5000 iterations
(Own editing using www.voyant.tools.org)

the number of iterations and the convergence of the model, even though as the number of data points grow (bigger data sets) the number of iterations required for the model to converge will grow too (Linderman & Steinerberger, 2017). 900 iterations yielded the best result with the clusters being the tightest (Figure 11). This model version is even better than the model in Figure 7 with the same level of perplexity but a much higher number of iterations (5000). The colours reflect data points (words in this case) that belong to the same cluster, while the size of the points is proportionate to the relative frequency of words. SM seems to be strongly related, which might be attributable to the nature of the question. There is a clearly detectable cluster that is about communication. Regular meetings and appropriate communication in general can significantly improve cooperation (Madhani, 2016) and reduces conflict (Snyder, McKelvey & Sutton, 2016). Communication and information sharing between SM is considered to be one of the keys to an effective management of the SMI (Biemans, Brenčič & Malshe, 2010). In Figure 11 there is a separate cluster that contains the most important information sharing methods in the two departments (telephone, email). Points marked in lilac signal the importance of joint tasks and work as well as cooperation between SM departments. It is interesting that there is a "corporate level" cluster with terms such as "company", "organisation", "goals" which highlights the significance of how corporate goals and vision can influence the efficiency of SM. It also supports earlier literature emphasizing corporate vision (Kumar, 2016; Groysberg, Lee, Price & Cheng, 2018).

These tools might be valuable for professional and academic purposes for different reasons. In academic settings, where time constraints are not as pressing as in the business world they might serve as means of preliminary analysis prior to more conservative and traditional methods of qualitative data analysis such as directed text analysis or grounded theory techniques. In business settings where being time-effective directly impacts cost-effectiveness these tools can be invaluable to save time and energy. It is especially true in the case of large data sets such as thousands of pages of comments from a corporate page. The tools that this paper presented vary in degree of sophistication and explanatory power. The Cyrrus tool or the Correlation tool can reveal limited interactions whithin the answers. The Topics tool provides a higher level of intimacy with the text as besides frequencies ranking is also taken into account. The t-SNE tool provides the highest level of sophistication and the deepest analytical possibilities revealing how groups of terms are related to each other.

4. Conclusions

As Soltani, Ahmed, Ying-Liao and Anosike (2014) point out qualitative methodologies in oparations management has been gaining significance in recent decades especially for fileds like interfacing. One such interface challenge is the SM interface which the present paper uses as an example fpr the demonstration of the possibilities Voyant Tools can offer. Qualitative methods resulting in large textual data sets in the operations management paradigm include in-depth interviews, anthropological studies, participant observations, case studies or etnographies. As

operations management is increasingly dependent on Big Data analytics (Choi, Wallace & Wang, 2018; Guha & Kumar, 2018) like data mining, Voyant Tools can serve as useful and valuable supplementary technique. Integrating qualitative and quantitative analysis techniques in the analysis of qualitative data can result in a more solid foundation to build research conclusions on. Voyant-Tools offers an impressive array of tools to visualise the results of quantitatively analysed qualitative data. Visualisation tools might tempt the researcher to read suppositions into the data that do not reflect the true relationships of meaning units existing in the data set. As textual data is a coherent system of meaning units, care must be taken with interpreting results especially because there is a danger that quantitative analysis of qualitative data necessarily leads to considerable loss of information. However, these quantitative methods can be invaluable tools of preliminary analysis and hypothesis adjustment. Their results should always be checked against the traditional content analysis techniques which are more sensitive to the complex structure of semantic units. These quantitative techniques are to help early exploration of textual data. As there is virtually no earlier literature on how quantitative data visualisation techniques can be used in marketing research, especially in the analysis of the SMI, utilisation possibilities of Voyant Tools and other quantitative data analysis and visualisation software for handling qualitative data is definitely a worthwhile area for further research.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

References

Assarroudi, A., Heshmati-Nabavi, F., Armat, M.R., Ebadi, A., & Vaismoradi, M. (2018). Directed qualitative content analysis: the description and elaboration of its underpinning methods and data analysis process. *Journal of Research in Nursing,* 23(1), 42-55. https://doi.org/10.1177/1744987117741667

Belk, R.W. (2007). *Handbook of qualitative research methods in marketing.* Northampton, MA: Edward Elgar Publishing Inc.

Bellenger, D.N., Bernhardt, K.L., & Goldstucker, J.L. (2011). *Qualitative research in marketing.* Decatur: Marketing Classics Press.

Bernard, H.R. (1996). Qualitative data, quantitative analysis. *CAM Journal*, 8(1), 9-11. https://doi.org/10.1177/1525822X960080010401

Bernard, H.R., & Ryan, G. (1998). Text analysis: Qualitative and quantitative methods. In Bernard, H.R. (Ed.), *Handbook of methods in cultural anthropology* (595-645). Walnut Creek, CA: AltaMira Press.

Biemans, W.G., Brenčič, M.M., & Malshe, A. (2010). Marketing-sales interface configurations in B2B firms. *Industrial Marketing Management*, 39(2), 183-194.

Bradley, J. (2018). Digital tools in the humanities: Some fundamental provocations?. *Digital Scholarship in the Humanities,* 34(1), 13-20. https://doi.org/10.1093/llc/fqy033

Braun, V., Clarke, V., Hayfield, N., & Terry, G. (2019). Thematic analysis. In Liamputtong, P. (Ed.), *Handbook of Research Methods in Health Social Sciences* (843-860). Singapore: Springer.

Bryman, A. (2017). Quantitative and qualitative research: further reflections on their integration. In Brannen, J. (Ed.), *Mixing methods: Qualitative and quantitative research* (57-78). London: Routledge.

Cao, Y., & Wang, L. (2017). Automatic selection of t-SNE Perplexity. *arXiv preprint arXiv:1708.03229.*

Carson, D., Gilmore, A., Perry, C., & Gronhaug, K. (2001). *Qualitative marketing research.* London: Sage.

Choi, T.M., Wallace, S.W., & Wang, Y. (2018). Big data analytics in operations management. *Production and Operations Management*, 27(10), 1868-1883. https://doi.org/10.1111/poms.12838

Clouder, L., & King, V. (2015). What works? A critique of appreciative inquiry as a research method/ology. In Huisman, J., & Tight, M. (Eds.), *Theory and method in higher education research* (169-190). Bingley, UK: Emerald Group Publishing Limited.

Cometto, T., Labadie, G.J., & Palacios, M. (2017). The determinants of the quality of the sales-marketing interface in a multinational customer brand focused company: the Latin American branches. In Campbell, C.L. (Ed.), *The Customer is NOT Always Right? Marketing Orientations in a Dynamic Business World* (900-908). Springer, Cham.

Dey, I. (2003). *Qualitative data analysis: A user friendly guide for social scientists*. London: Routledge.

Flick, U. (2018). *An introduction to qualitative research* (6th ed.). Sage Publications Limited.

Fowler, R.L. (1987). Power and robustness in product-moment correlation. *Applied Psychological Measurement*, 11(4), 419-428. https://doi.org/10.1177%2F014662168701100407

Franses, P.H., & Paap, R. (2001). *Quantitative models in marketing research*. Cambridge, UK: Cambridge University Press.

Groysberg, B., Lee, J., Price, J., & Cheng, J. (2018). The leader's guide to corporate culture. *Harvard Business Review*, 96(1), 44-52.

Guha, S., & Kumar, S. (2018). Emergence of big data research in operations management, information systems, and healthcare: Past contributions and future roadmap. *Production and Operations Management*, 27(9), 1724-1735. https://doi.org/10.1111/poms.12833

Gummesson, E. (2005). Qualitative research in marketing: Road-map for a wilderness of complexityand unpredictability. *European journal of marketing*, 39(3/4), 309-327. https://doi.org/10.1108/03090560510581791

Havlicek, L.L., & Peterson, N.L. (1976). Robustness of the pearson correlation against violations of assumptions. *Perceptual and Motor Skills*, 43(3 suppl.), 1319-1334. https://doi.org/10.2466/pms.1976.43.3f.1319

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative health research*, 15(9), 1277-1288. https://doi.org/10.1177/1049732305276687

Hughes, D.E., Le Bon, J., & Malshe, A. (2012). The marketing-sales interface at the interface: Creating market-based capabilities through organizational synergy. *Journal of Personal Selling & Sales Management*, 32(1), 57-72. https://doi.org/10.2753/PSS0885-3134320106

Jacobi, C., Van Atteveldt, W., & Welbers, K. (2016). Quantitative analysis of large amounts of journalistic texts using topic modelling. *Digital Journalism*, 4(1), 89-106. https://doi.org/10.1080/21670811.2015.1093271

Krippendorff, K. (2018). *Content analysis: An introduction to its methodology*. Sage publications.

Kumar, D. (2016). *Enterprise growth strategy: vision, planning and execution*. New York, USA: Routledge.

Lapan, S.D., Quartaroli, M.T., & Riemer, F.J. (2011). *Qualitative research: An introduction to methods and designs* (vol. 37). San Francisco: John Wiley & Sons, Inc.

Linderman, G.C., & Steinerberger, S. (2017). Clustering with t-sne, provably. *arXiv preprint arXiv:1706.02582*.

Lipowski, M., Pastuszak, Z., & Bondos, I. (2018). Synergy of Quantitative and Qualitative Marketing Research−Capi and Observation Diary. *Econometrics*, 22(1), 58-67. https://doi.org/10.15611/eada.2018.1.04

Lockley, L.C. (1950). Notes on the history of marketing research. *Journal of Marketing*, 14(5), 733-736.

Madhani, P.M. (2016). Competitiveness and sustaining performance: integrating sales and marketing. *SCMS Journal of Indian Management*, 13(1), 19-36.

Malshe, A., & Al-Khatib, J.A. (2017). A Repertoire of Marketers' Trust-Building Strategies Within The Sales-Marketing Interface. *Journal of Personal Selling & Sales Management*, 37(3), 213-227. https://doi.org/10.1080/08853134.2017.1345316

Malshe, A., Friend, S.B., Al-Khatib, J., Al-Habib, M.I., & Al-Torkistani, H.M. (2017). Strategic and operational alignment of sales-marketing interfaces: Dual paths within an SME configuration. *Industrial Marketing Management*, 66, 145-158.

Matthyssens, P., & Johnston, W.J. (2006). Marketing and sales: optimization of a neglected relationship. *Journal of Business & Industrial Marketing*, 21(6), 338-345. https://doi.org/10.1108/08858620610690100

Mayer, I. (2015). Qualitative research with a focus on qualitative data analysis. *International Journal of Sales, Retailing & Marketing*, 4(9), 53-67.

Miller, A. (2018). Text Mining Digital Humanities Projects: Assessing Content Analysis Capabilities of Voyant Tools. *Journal of Web Librarianship*, 12(3), 169-197. https://doi.org/10.1080/19322909.2018.1479673

Molina-Azorin, J.F., Bergh, D.D., Corley, K.G., & Ketchen, D.J. (2017). Mixed Methods in the Organizational Sciences: Taking Stock and Moving Forward. *Organizational Research Methods*, 20(2), 179-192. https://doi.org/10.1177/1094428116687026

Müller, A., Boda, E.J., Ráthonyi, G., Ráthonyi-Odor, K., Barcsák, B., Könyves, E. et al. (2016). Analysis of the Cavebath of Miskolctapolca's brand elements and guests satisfaction. *APSTRACT: Applied Studies in Agribusiness and Commerce*, 10(4/5), 155-160. https://doi.org/10.19041/APSTRACT/2016/4-5/20

Snyder, K., McKelvey, S., & Sutton, W. (2016). All together now? Exploring sales and marketing integration. *Sport, Business and Management: An International Journal*, 6(1), 2-18. https://doi.org/10.1108/SBM-08-2013-0027

Soltani, E., Ahmed, P.K., Ying-Liao, Y., & Anosike, P.U. (2014). Qualitative middle-range research in operations management: the need for theory-driven empirical inquiry. *International Journal of Operations & Production Management*, 34(8), 1003-1027. https://doi.org/10.1108/IJOPM-11-2012-0486

Steiner, C.M., Agosti, M., Sweetnam, M.S., Hillemann, E.C., Orio, N., Ponchia, C. et al. (2014). Evaluating a digital humanities research environment: the CULTURA approach. *International Journal on Digital Libraries*, 15(1), 53-70. https://doi.org/10.1007/s00799-014-0127-x

Toubia, O., Iyengar, G., Bunnell, R., & Lemaire, A. (2019). Extracting Features of Entertainment Products: A Guided Latent Dirichlet Allocation Approach Informed by the Psychology of Media Consumption. *Journal of Marketing Research*, 56(1), 18-36. https://doi.org/10.1177%2F0022243718820559

Uboldi, G., & Caviglia, G. (2015). Information visualizations and interfaces in the humanities. In Bihanic, D. (Ed.), *New Challenges for Data Design* (207-218). London, Springer.

Van Der Maaten, L. (2014). Accelerating t-SNE using tree-based algorithms. *The Journal of Machine Learning Research*, 15(1), 3221-3245.

Van Der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research*, 9(26), 2579-2605.

Wattenberg, M., Viégas, F., & Johnson, I. (2016). How to use t-SNE effectively. *Distill*, 1(10), e2. https://doi.org/10.23915/distill.00002

Welsh, M.E. (2014). Review of Voyant tools. *Collaborative Librarianship*, 6(2), 96-98.

Williams, N.L., Inversini, A., Buhalis, D., & Ferdinand, N. (2015). Community crosstalk: an exploratory analysis of destination and festival eWOM on Twitter. *Journal of Marketing Management*, 31(9-10), 1113-1140. https://doi.org/10.1080/0267257X.2015.1035308

Wilson, A. (2018). *Marketing research: Delivering Customer Insight*. London: Macmillan International Higher Education.

Young, F.W. (1981). Quantitative analysis of qualitative data. *Psychometrika*, 46(4), 357-388. https://doi.org/10.1007/BF02293796

Zahedzadeh, G. (2017). Overt attacks and covert thoughts. *Aggression and Violent Behavior*, 36, 1-8. https://doi.org/10.1016/j.avb.2017.06.009