

Exploration of text mining methodology through investigation of QMOD-ICQSS proceedings

Author: Daniel Carnerud, PhD-student

University: Mid Sweden University - Department of Quality Technology and Management, Mechanical Engineering and Mathematics

E-mail: daniel.carnerud@miun.se

Address: Roslagsgatan 45, 113 54 Stockholm

Telephone: 0046 (0) 735 24 49 00

Abstract

Purpose

The purpose of the study is to explore if and how text mining methodology can be applied to investigate and analyze large amounts of academic papers.

Methodology / Approach

The study uses probabilistic topic modeling as method to perform exploratory data analysis on QMOD-ICQSS proceedings.

Findings

The findings give reason for further exploration of text mining and probabilistic topic modeling as a relevant method to extract information from large amounts of academic papers. The generated topic models serve as a foundation for other researchers to conduct more detailed analysis of the QMOD-ICQSS proceedings depending on the specific area of interest.

Research Limitations

The study was limited to the use of one software and modeling technique, a multi-track approach would have strengthened the results.

Originality

There seems not to exist any other text mining study that has been conducted on conferenced proceedings in general nor on QMOD-ICQSS proceedings specifically.

Keywords

Text Mining, Probabilistic Topic Modeling, Quality Movement, Total Quality Management, QMOD-ICQSS

Paper type

Research paper

Introduction

During recent years there has surged a popular area of research labeled business intelligence and data mining (Larose, 2005). These methods and techniques are the consequence of a world that is becoming more and more digitalized and where new more powerful and efficient tools to handle the ever increasing amount of information are constantly developing (Ibid). Simply put, data is becoming cheaper and easier to collect as well as to store and to analyze. Traditionally, research in data mining within the fields of computer science, statistics, mathematics and artificial intelligence have focused on how large masses of numbers and figures can be used to gather valuable insights on the past, present and future (Ibid). However, a relatively new field of research is now sprouting out of data mining - text mining (Weiss *et al.*, 2012). Text miners use much of the knowledge and tools derived from traditional data mining but focus on linguistic data sets instead of numerical (Miner, 2012). This leap towards statistical processing of verbal data is driven by the urge to develop more able ways of generating knowledge from human records and data, which ultimately often are non-numerical (Aggarwal & Zhai, 2012).

The pioneers of total quality management (TQM) as Taylor (1911), Shewhart (1939), Fiegenbaum (1951), Juran (1964) and Deming (1986) all focused on the application of statistical tools and processes to develop efficient methods for inspection, control, assurance and continuous development of quality. Although not exclusively, much of their work revolved around statistical processing of numbers and figures. Much in the same way research combining data mining and quality issues has concentrated on numerical processing and in the cases where studies have focused on natural language and used text mining it has been industry oriented (Heim & Peng, 2010; Köksal *et al.*, 2011; Lo, 2008). However, in modern times the quality movement has seen an increased focus on research areas and issues of more intangible character such as theory, context, culture, leadership and learning which has led researchers as well as practitioners to request new ways of looking into and working with these subjects (Nair, 2006; Sila, 2007; Sousa & Voss, 2002; Zhang *et al.*, 2012). As researchers and practitioners of TQM are constantly looking for new efficient and reliable means to grasp the kind of intangible dimensions mentioned above, text mining can prove to be a valuable asset since both stand on the same philosophical foundation, statistics and mathematics, and both share the urge to learn more about how we can extract knowledge from human systems in resourceful and trustworthy ways (Finch, 1999; Khamis *et al.*, 2013). Hence, the purpose of this study is to explore if and how text mining methodology can be applied to investigate and analyze large amounts of academic papers.

Theoretical framework

Within the quality movement there exists a vivid debate concerning the historical developments and future direction of the research field, where many demand a more concrete theoretical grounding if the discipline is not to lose relevance as an academic field of study (Dahlgaard-Park, 1999; Dahlgaard-Park, 2011; Singh & Smith, 2006; Van Der Wiele *et al.*, 2000). Traditional research methods used for exploring the quality movement have to a large extent and successfully consisted of surveys, literature reviews and case studies (Dahlgaard-Park *et al.*, 2013; Kroslid, 1999; Mauléon, 2003; Montabon *et al.*, 2007; Sila & Ebrahimpour, 2002). However, as we now stand in an era where we with the help of digital tools are able to collect, process and analyze amounts of data in an earlier unprecedented way it could be of relevance to make use of these novel methods and data sources to shed new light on the quality movement (Angell & Corbett, 2009; Gattiker & Parente, 2007). This way, new epistemological knowledge could be generated and whether it confirms earlier conclusions or not is beyond the point, the mere exploration is in itself a

theory building activity which hopefully gives new energy to the research paradigm (Godfrey-Smith, 2003). Moreover, since there exists a lively debate concerning the fundamentals of TQM; what it actually consists of, its context dependence or independence and even how it properly should be named (Harnesk & Abrahamsson, 2007; Green; 2012, Hellsten & Klefsjö, 2000, Nair, 2006; Sila, 2007; Soltani *et al.*, 2008, Zhang *et al.*, 2012), it becomes imperative to stay broad-minded when conducting exploratory inquiries (Godfrey-Smith, 2003), which constitutes a good environment for successful data- and text mining (Miner, 2012; Weiss *et al.*, 2012). Nonetheless, there exists outspoken and largely accepted statements describing TQM as a holistic management approach which with the customer in focus aims for continuous improvement of operational effectiveness through involvement of the whole organization (Dahlgaard *et al.*, 2007) characterized by certain principles, practices and techniques (Dean & Bowen, 1994) or values, techniques and tools (Bergman & Klefsjö, 2010; Hellsten & Klefsjö, 2000). The practices, tools and techniques underpin the values or principles which are seen as the most profound part of TQM; the principles being “customer focus”, “continuous improvement” and “teamwork” (Dean & Bowen, 1994) and the values consisting of “leadership commitment”, “customer orientation”, “participation of everyone”, “process orientation”, “continuous improvements” and “base decisions on fact” (Bergman & Klefsjö, 2010; Hellsten & Klefsjö, 2000, Lagrosen *et al.*, 2012). In sum, the field of text mining stands firm on statistical and mathematical foundation and as such is could also show to be a relevant methodology for the quality movement to explore since it is also grounded on statistics and mathematics - basing decisions on facts. Hence, in one way text mining could be seen as a way of closing the circle, making intangible dimensions tangible with the help of statistical methodology.

Text mining

Purposeful text mining rests on the same assumption as data mining - looking for valuable patterns, correlations and trends in large datasets with the help of statistical and mathematical techniques, a process too complex and resource demanding for manual processing (Liu, 2011; Miner, 2012). Within the research field there has developed a Cross Industry Standard Process for Data Mining (CRISP-DM) consisting of six phases: Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation and Deployment (Chapman *et al.*, 2000; Marbán *et al.*, 2009; Kurgan & Musilek, 2006). The business understanding phase includes definition of the objectives, formulation of problem as well as of strategy to tackle the problem. Understanding data refers to the collection and initial exploration and evaluation of the data with possible change in scope and strategy. Preparing data includes cleaning the data of distorting information and values as well as narrowing down the elements and variables to be included and processed. In the modeling phase modeling techniques are chosen and calibrated for optimal results, sometimes if necessary, reversing to the data preparation phase depending on outcome. Then the result of the models are evaluated and compared with the objectives of the study, a part where human judgment is crucial (Chang *et al.*, 2009). Finally, depending of the outcome the models and results are applied in accordance with the objectives of the study.

Text mining can be conducted through different methods, however, among the most promising as of today are so called probabilistic topic models using the Latent Dirichlet Allocation (LDA) model as distributed algorithm, which offer a method to cluster words into topics on the basis of their most probable association (Blei, 2012; Charu *et al.*, 2012; Newman *et al.*, 2009; Xie & Xing, 2013). Topic modeling with the use of LDA was decided to be the principal method of processing the data.

Methodology

The study has been carried out according to the CRISP-DM process described in the previous section, thus, the methodological section is presented accordingly.

Business understanding

The objectives of the study were to test text mining methodologies on conference proceedings from the QMOD-ICQSS conference to see if relevant information could be generated. The QMOD-ICQSS conference has been organized for 16 consecutive years in places all around the world with participants from around the globe. Through the years it has developed into one of the largest scientific conferences in the world within the research fields of Quality, Service, Organizational development and related management issues. Thus, the QMOD-ICQSS constitutes a relevant object of study if one wishes to look into the quality movement and the areas of research that have been popular within the research community. Since the conference is focusing on the research areas of Quality, Service, Organizational development and related management issues there was no need to delimit the selection of conference proceedings.

There are many different software and programming languages used for data mining but it was decided that this study was to be conducted in the R, <http://cran.r-project.org/>, since it is a freeware that has proven powerful in its text processing capabilities with new text mining and topic model packages developing continuously (Feinerer, 2008; Theussl *et al.*, 2012).

Data understanding

Since the QMOD-ICQSS conferences proceedings are not available for downloading through any database manual collection was necessary and help from people who had attended and organized previous conferences was needed. After positive responses from key figures it was clear that a dataset consisting of previous QMOD-ICQSS could be gathered. This year, 2014, the 17th QMOD-ICQSS conference is organized which left proceedings from 16 earlier conferences to be potentially included in the study. Since data- as well as text mining is benefited from as large datasets as possible, the aim of the study was to include as many proceedings possible from previous QMOD-ICQSS conferences, estimating a dataset consisting of in-between 1500-2000 proceedings. In the end conference proceedings from 2003, 2006, 2007, 2008, 2009, 2010, 2011, 2012 and 2013 were successfully collected, rendering a dataset consisting of in total 1019 conference proceedings. Even though seven conferences were not included in the dataset and the initial goal to include all previous proceedings was not met, the amount of proceedings was considered large enough to proceed with the study.

Data Preparation

Since the proceedings were structured in different ways as well as in different digital formats a noteworthy proportion of time was dedicated into transferring the data to a digitally processable format, a so called corpus in text mining language. Since all conference papers are instructed to include certain information besides the text in the paper (paper title, author(s), academic title, university/affiliation, abstract, purpose, methodology/approach, findings, research limitation/implication, originality/value of paper, keywords and list of references) it was necessary to choose if any of the sections were to be separated and labeled. Initially, an attempt was made to break loose author(s), purpose, keywords and list of references since this seemed to constitute relevant parts to be studied and compared.

However, due to the difference in proceeding structure and digital formats the ambition was deemed too time consuming and instead the data was structured according to year of conference, title and authors.

The final corpus consisted of 84.838 unique words which in total summed up to 3.618,120 words since many occur several times. Out of this amount of words a minority actually carry any real contextual meaning. Several unique words are in fact "non-sense" words that have been generated when converting the original documents into a processable format and others are so called stop-words that are used to simplify and clarify written communication but essentially bear no meaning in themselves. Since all words in the corpus influence the topic models output different cleaning methods were applied and tested, however, the goal is always to include as many words as possible in a topic model since this gives the algorithms more data to work and calibrate the topic models with. Firstly the corpus was cleaned out of English stop-words which do not add value to the analysis, the amount of unique stop-words only amounts to 175 but the amount of total words is substantially lowered. After initial modeling sessions, conference proceedings in French were identified and it was decided to clean the corpus of French stop-words as well. Secondly, two main approaches were made, cleaning words that occur less than a specified amount of times and applying the so called tf-idf formula. Sorting out words that occurred less than a specific amount of times is straight forward but it can be seen as quite roughly trimmed since there exists no clear way of establishing the word-limit. The other method of bringing out the most valuable and meaningful words is by a formulation that counts the perceived importance of a word is used, the so called tf-idf formulation. Tf-idf computes term frequency with inverse document frequency which calculates weightings or scores for words. Term frequency is the number of times a term occurs in a document. Inverse document frequency register whether a term is common or rare across all documents, checking the number of documents containing a word and reversing the scaling. The point of tf-idf is to capture how much information a word provides, a word that is relatively unique and appears in few documents will get a higher value than a word appearing everywhere. Even if the theoretical foundation of the tf-idf formulation has been under discussion for the last 30 years it was none the less judged as a more scientific mechanism for cleaning the corpus than by manually tampering with the occurrence level of a word.

Modeling

When modeling the parameter worked with was the number of topics in the model. Models for each specific year were also generated. The number of topics that can be generated in a model are indefinite, the issue is which amount of topics is optimal for answering the specific research question. In this case, the mission was to explore if and how text mining can be applied to study academic papers. For this reason it was judged that it would be valuable to generate and present several models on the same corpus to show how topics and topic models evolve. In all, more than 40 topic models were generated through the process ranging from 10 to 100 topics. Depending on the amount of words and topics that were included in a model the time for generating the results varied heavily, ranging from 20 minutes to 8 hours, with the simple relation; the more words and topics the longer time for processing.

As complement to the topic modeling, lists of the most common words were also generated since it was thought possible that they could bring about interesting findings.

Evaluation

When it comes to analysis and evaluation of topic models no machine learning models have yet outperformed that of human judgment. (Dacres. *et al.*, 2013; Chang *et al.*, 2009; Xie & Xing, 2013). Interpretation and evaluation was thus conducted by the author and are by nature subjective. The main advantage of this type of qualitative modus is that it is fairly easy for a human with a somewhat normal vocabulary to relatively quickly and efficiently recognize nonsense words. However, once sensible models are generated the difficulty of identifying patterns and structure arise. As the number of topics in a model increases the overall view of the topics decreases. Additionally, as every topic model is unique and the topics in the models are generated in random order it fast becomes difficult to contrast and evaluate results in large models as well as between models. Also the question of how many words should be included in each topic influences the interpretability, albeit it does not influence the models in themselves.

Deployment

Since the topic models generated fruitful results, continuous testing and evaluation on the same corpus is seen as possible and meaningful, leading to better understanding and development of text mining software and learning models.

Results

As expected different strategies influenced the models output, however, the sequential fashion in which these methods were applied and tested is for practical reasons not presented chronologically instead a summary of the process is made.

As work initially progressed the topic models were not showing any sensible results, with many non-sense word occurring and being given high significance. After some testing suspicions regarding the tf-idf formula arose why experimentation with word occurrence limits started. It turned out that when combining the two methods the models showed substantially more relevant results. Consequently the tf-idf formula was given less and less influence on the models and in the end the tf-idf formula was completely left out and only a word occurrence limit was used instead. Through experimentation it was judged that the models were performing best when words that occurred less than twenty times were excluded, creating a corpus of 9344 unique words.

When experimenting with different amount of topics it was assessed that 10 topics gave too little information on the corpus and 100 topics became too elusive. In the same way generating few models did not give a good overall picture of the difference between models and too many models led to information overflow. Hence, the best result was deemed to be represented with four models consisting of 20, 40, 60 respectively 80 topics, presented in appendix A. The models should be interpreted as the probability of words occurring together decreases further down the column.

The topics can be presented with different amount of words, this scope does not inflict on the model in itself but on the ability to evaluate the topics. On one side the more words that are shown the more nuanced and in depth the topics become, on the other side the more words the more difficult it becomes to compare and differentiate between topics and models. After having topics presented with eight to thirty words the impression was that topic models presented with twenty words gave the best overall ability to assess and compare

models and topic since the number of words gave depth to each topic without it becoming excessive.

A list of the most common words in the final corpus as well most common words per year was also generated, appendix B.

Discussion

The text mining methods have proven successful as four models with clear and relevant topics were generated. Which topic model is judged to hold the highest relevance depends on which subject or issue one wants to investigate. If an overall summary is of most interest the models with few topics is likely to be of most value, for more detailed examination of a specific word or subject the larger models are probably preferred. A way of combining the models and approaching a subject is to start off with a topic model consisting of few topics and then follow how the topic and the words evolve and disperse. By studying the topic models per year a chronological perspective is gained from the material. When looking at the topic model consisting of 20 topics, appendix A, one can see that topic 3 clusters around that the corpus consists of conference proceedings, topic 4 illustrate that the conference proceedings cite journals related to TQM and topic 8 show that the data-set is rooted in the academia. Topic 20 could be interpreted as QMOD-proceedings revolve around a belief that we have entered new times which require people and companies to change and differentiate with the use of creativity. Related to the theoretical discussions of TQM it is clear that the topics most easily fitted within existing TQM frameworks are those belonging to the superficial dimensions - practices, techniques, tools and methods (Bergman & Klefsjö, 2010; Dean & Bowen, 1994; Hellsten & Klefsjö, 2000, Lagrosen et al., 2012) such as Business excellence models (topic 7), Six-sigma (topic 10), Lean (topic 11), Kansei design (topic 12) and ISO (topic 17). Linkage to the principles "customer focus", "continuous improvement" and "teamwork" (Dean & Bowen, 1994) and the values "leadership commitment", "customer orientation", "participation of everyone", "process orientation", "continuous improvements" and "base decisions on fact" (Bergman & Klefsjö, 2010; Hellsten & Klefsjö, 2000, Lagrosen *et al.*, 2012) emerge in topic 5 covering process management, topic 9 containing customer service, topic 13 encompassing service and customer satisfaction, topic 14 revolving around innovation and product development, topic 18 dealing with supply chain systems and finally leadership incorporated in topic 19. Topics that are obviously treated by the conference proceedings but which do not clearly fit the above stated TQM frameworks are topic 6 concerning knowledge and learning, topic 2 and 16 focusing on social dimensions, sustainability and environment and topic 1 and 15 addressing public service and health care. In all, the results indicate that the QMOD-ICQSS conference proceedings and the research they reflect can be classified within the theoretical framework of TQM proposed by Dean & Bowen (1994) and Bergman och Klefsjö (2010). However, these results also imply that QMOD-ICQSS conference proceedings cover areas beyond the scope of these frameworks giving fuel to researchers like Green (2012), Harnesk & Abrahamsson (2007), Nair (2006), Sila (2007) and Zhang *et al.* (2012) who argue that TQM and the quality movement needs to further explore the role of context and culture to better understand the true nature of TQM and freeing it of ambiguity. Studying the more extensive topic models consisting of 40, 60 and 80 topics, appendix A, further underlines these dualities, as the amount of topics increase, their level of detail is more refined and thus can be interpreted and classified quite arbitrarily. Looking at the topic models per year no clear trend regarding the theoretical framework has successfully been identified and more analysis is needed. Nevertheless, the topic models clearly differ between each year confirming that the conference and its proceedings cover different topics each year. Also, each year has unique, year-specific, topics such as: *performance management* (topic 18,

2003), *benchmarking* (topic 7, 2006), *kansei design* (topic 4 & 17, 2007), *risk* (topic 8, 2007), *lean* (topic 11, 2008), *event, tourism and environment* (topics 14, 15 and 20, 2009), *transport services, change management and wellness services* (topics 3, 9 and 11, 2010), *health care and knowledge development* (topics 2 and 7, 2011) *organizational learning, csr, life quality and project quality* (topics 4, 13, 19 and 20, 2012) and *social media, leadership and online quality* (topics 13, 14 and 20, 2013). The year specific topics suggest that there are trends amongst the conference proceedings and could possibly confirm that the year-specific themes have a clear impact on the proceedings, although this claim cannot be supported by the study since the conference themes were not available in the data set.

Considering the nature of the papers it is not surprising that the most occurring words in the overall summary as well as per year, appendix B, consequently are words like quality, management, process and service. Some words that could be regarded as stop-words such as can, also, new and one are found and it is not clear how big influence they have on the modeling phase. The words could manually be removed in the data preparation phase, however, this would probably influence the modeling phase but further inquiry into this issue needs to be made before any conclusions can be drawn. Even so, when it comes to knowledge extraction it is questionable of how much importance lists of occurring words actually are. A list of the most occurring words state nothing more than this, that a word has occurred many times, which in turn does not really give that much substantial base for further analysis. In this sense topic models are of much higher relevance since they create a context which describes how words correlate with one another, giving a more solid and comprehensive base for analysis. This could imply that traditional research methods such as key-word are inferior to that of text mining and topic modeling as the later give more causal explanations of context and associations. A way of combining the two methods could show fruitful as key-word search and common words indicate in which direction a researcher should focus, using text mining and topic modeling as a way to extract more detailed information.

The results do not give any help in resolving the difficulties of analyzing topic models but the study does give support to Chang *et al.* (2009) in that the human judgment is capable of identifying relevant topics and topic models, especially since the recognition of the tf-idf formula as the source of error in the initial topic models was completely due to human assessment. However, as larger topic models consisting of many topics are generated they fast get too big and difficult to overview for the human mind in an objective way, and this issue of evaluation of topic models is of crucial importance and needs to be handled if topic modeling is to develop as a practical and meaningful methodology.

Regarding the software used in this study, R is recognized as a powerful programming language with many possibilities for the user to write own solutions as well as to use packages used by others. Nevertheless, the user needs to possess substantial programing skills to make R justice and other more user friendly software probably have the upside of offering a less experienced programmer more time to spend on working with the material instead of reading and writing code.

Finally, if the organizers would like to encourage data and text mining activities on the QMOD-ICQSS proceedings there are some experiences to be drawn concerning the structure and presentation of the proceedings. Easily put, the conference proceedings have become more graphically and visually attractive during recent years which make them harder to transform into a workable data format. Since a graphically more appealing presentation of the conference proceedings is probably seen as something positive for the conference participants this should be kept up, instead the proceedings could also be saved

in a format that simplifies digital processing and in best case saved in a database easily accessible to researchers.

Conclusion

Text mining with the help of topic modeling has through the study proven to be one way of processing QMOD-ICQSS conferences proceedings. Text mining and topic modeling is seen as unique methodology compatible with values, tools and methods within the TQM community, offering new interesting ways of approaching large datasets adding a perspective to that of traditional methods such as content analysis, case studies, surveys or literature reviews.

Text mining in general and topic modeling specifically offer great opportunities to with the help of modern soft- and hardware explore large amounts of information and discover underlying structures, patterns and trends. However, attaining real-life datasets that are suited for this kind of processing is easier said than done. Hence, the crucial point is to successfully attain and experiment with relevant datasets from organizations if the methods are to develop into a relevant tool for people working with quality development and improvement. The other issue that needs to be resolved is interpretability, as a topic model grows larger it gets more difficult to overview and assess for a human mind. Traditional data mining has resolved this issue by applying machine learning models to evaluate results, text mining would also benefit from the development of similar learning models so that evaluation does not solely rest on human judgment.

Research Limitations

The study includes conference proceedings from 2003, 2006, 2007, 2008, 2009, 2010, 2011, 2012 and 2013. Since the QMOD-ICQSS conference has been organized on 16 prior accounts that leaves 7 conferences out of the dataset which limits the possibility to make generalizations of the QMOD-ICQSS conference as a whole based on the findings. The modeling was only made in the software R, other programming languages and software could give other perspectives on the material. In the analysis only topic models were applied, other learning methods could give valuable insights on the material.

Future Research

A corpus consisting of conference proceedings from the specified QMOD-ICQSS conferences is generated through this study and will be made available for other researchers which simplifies future research on the same dataset. If the data-set would be extended to include the remaining seven QMOD-ICQSS conferences the dataset would be unique in its kind and valuable resource when exploring text mining as methodology as well as the historical evolution of the quality movement.

A valuable contribution to the understanding and applicability of text mining could be a comparative study on the same dataset, contrasting text mining with traditional research methods. This way strengths and weaknesses of the methods could be highlighted, indicating when certain methods are to be preferred over others.

Acknowledgments

This work could not have been possible without the kind help of Jens Dahlgaard, Su Mi Dahlgaard-Park and Ingela Bäckström who all joined efforts to gather conference

proceedings stored in libraries and desk drawers around Sweden. The also author owe his greatest appreciation for the programming and statistical help, support and guidance from Måns Magnusson and Mikael Bengtsson at Linköpings University's department of computer and information science.

List of references

- Aggarwal, C. C. & Zhai, C.X. (2012). *Mining Text Data*. Boston, MA: Springer US
- Angell, L. C., Corbett, L.M., (2009). The quest for business excellence: evidence from New Zealand's award winners, *International Journal of Operations & Production Management*, 2009, Vol.29(2), p.170-199
- Bergman, B. & Klefsjö, B. (2010). *Quality: from Customer Needs to Customer Satisfaction* (3 ed.). Lund: Studentlitteratur
- Blei, D., (2012). Probabilistic Topic Models, *Communications of the ACM*, vol. 55, no. 4, 77-84
- Chakrabarti, S. (2003). *Mining the Web: discovering knowledge from hypertext data*. Amsterdam: Morgan Kaufmann
- Chang, J., Boyd-Graber, J., Wang, C., Gerrish, S., Blei, D. (2009). Reading Tea Leaves: How Humans Interpret Topic Models. *Neural Information Processing Systems*.
- Chapman, P.; Clinton, J.; Kerber, R.; Khabaza, T.; Reinartz, T.; Shearer, C. & Wirth, R. (2000). CRISPDM 1.0 step-by-step data mining guide. Technical report, CRISP-DM.
- Charu C Aggarwal and ChengXiang Zhai. (2012) A survey of text clustering algorithms. *Mining Text Data*, p. 77–128.
- Dacres, S.; Haddadi, H. & Purver, M. (2013). 'Topic and Sentiment Analysis on OSNs: a Case Study of Advertising Strategies on Twitter.', *CoRR abs/1312.6635*.
- Dahlgaard, J., J., Kristensen, K. & Kanji, G., K. (2007). *Fundamentals of total quality management : process analysis and improvement*. London: Taylor & Francis.
- Dahlgaard-Park, S., M., Chen, C.K. , Jang, J.Y. & Dahlgaard, J., J. (2013) Diagnosing and prognosticating the quality movement – a review on the 25 years quality literature (1987–2011), *Total Quality Management & Business Excellence*, 24:1-2, 1-18
- Dahlgaard-Park, S., M. (1999). The evolution patterns of quality management: Some reflections on the quality movement. *Total Quality Management*, 10(4&5), 473–480.
- Dahlgaard-Park, S., M. (2011). The quality movement – where are you going? *Total Quality Management & Business Excellence*, 22(5), 493–516.
- Dean, J. W. Jr., Bowen, D., E. (1994). Management theory and total quality: improving research and practice through theory development. *Academy of Management Review*, July, 1994, Vol.19(3), p.392(27)
- Deming, W., E. (1986). *Out of the Crisis*. Cambridge University Press, Cambridge
- Finch, B. J., (1999). Internet discussions as a source for consumer product customer involvement and quality information: an exploratory study, *Journal of Operations Management*, 1999, Vol.17 (5), pp.535-556

Gattiker, T. F., Parente, D. H., (2007). Introduction to the special issue on innovative data sources for empirically building and validating theories in Operations Management, *Journal of Operations Management*, 2007, Vol.25(5), pp.957-961

Juran, J.M. (1964). *Managerial Breakthrough*. McGraw-Hill, New York

Feigenbaum, A. (1951). *Total Quality Control*, McGraw-Hill, New York.

Feinerer, I., (2008). Text Mining Infrastructure in R. *Journal of statistical software* vol: 25 iss:5 [1548-7660]

Godfrey-Smith, P., (2003). *Theory and reality: an introduction to the philosophy of science*. Chicago, Ill.: University of Chicago Press.

Green, T., J., (2012). TQM and organisational culture: How do they link? *Total Quality Management & Business Excellence*, 23:2, 141-157

Harnesk, R., Abrahamsson, L., (2007). TQM: an act of balance between contradictions. *The TQM Magazine*, 2007, Vol.19(6), p.531-540

Heim, G. R., Peng, D. X., (2010). The impact of information technology use on plant structure, practices, and performance: An exploratory study, *Journal of Operations Management*, 2010, Vol. 28(2), pp.144-162

Hellsten, U., Klefsjö, B., (2000). TQM as a management system consisting of values, techniques and tools. *The TQM Magazine*, 2000, Vol.12(4), p.238-244

Khamis, N., Rilling, J., Witte, R., (2013). Assessing the quality factors found in in-line documentation written in natural language: The JavadocMiner. *Data & Knowledge Engineering*, Sept, 2013, Vol.87, p.19(22)

Kroslid, D., (1999). *In search of quality management: rethinking and reinterpreting*. Doctoral dissertation, Linköping University.

Kurgan, L.A., Musilek, P., (2006). A survey of Knowledge Discovery and Data Mining process models, ”, *The Knowledge Engineering Review*, Volume 21, Issue 01, 2006, pp 1-24.

Köksal, G., Batmaz, İ., Testik, M. C., (2011). A review of data mining applications for quality improvement in manufacturing industry, *Expert Systems With Applications*, 2011, Vol.38(10), pp.13448-13467

Lagrosen, Y., Travis, F. T., Lagrosen S., (2012). Brain integration as a driver for quality management success. *International Journal of Quality and Service Sciences*, 2012, Vol.4(3), p.253-269

Larose, D. T. (2005). *Discovering knowledge in data an introduction to data mining*. Hoboken, N.J.: Wiley-Interscience

Liu, B. (2011). *Web data mining: exploring hyperlinks, contents, and usage data*. 2nd ed. Heidelberg: Springer

- Lo, S., (2008). Web service quality control based on text mining using support vector machine, *Expert Systems With Applications*, 2008, Vol.34(1), pp.603-610
- Marbán, O., Mariscal, G., Segovia, J., A Data Mining & Knowledge Discovery Process Model (2009). *Data Mining and Knowledge Discovery in Real Life Applications*, Editors: Julio Ponce and Adem Karahoca, pp. 438, February 2009, I-Tech, Vienna, Austria
- Mauléon, C. (2003). Recapturing the spirit of quality. Lic.-thesis. . Department of Quality Sciences, Chalmers University of Technology, Göteborg, Sweden.
- Miner, G. (red.) (2012). *Practical text mining and statistical analysis for non-structured text data applications*. 1st ed. Waltham, MA: Academic Press
- Montabon, F., Sroufe, R., Narasimhan, R., 2007. An examination of corporate reporting, environmental management practices and firm performance. *Journal of Operations Management*, 2007, Vol.25(5), pp.998-1014
- Nair, A., (2006). Meta-analysis of the relationship between quality management practices and firm performance—implications for quality management theory development. *Journal of Operations Management*, 2006, Vol.24 (6), pp.948-975
- Newman, D., Asuncion, A., Smyth, P., Welling, M., (2009). Distributed Algorithms for Topic Models, *Journal of Machine Learning Research* 10 (2009) 1801-1828
- Shewhart, W.A. (1939). *Statistical Method. From the Viewpoint of Quality Control*. Graduate School of the Department of Agriculture, Washington, D.C.
- Singh, P. J., Smith, A., (2006) Uncovering the faultlines in quality management, *Total Quality Management & Business Excellence*, 17:3, 395-407
- Sila I., Ebrahimpour, M. (2002). An investigation of the total quality management survey based research published between 1989 and 2000; A literature review. *International Journal of Quality & Reliability Management*, 2002, Vol.19 (7), p.902-970
- Sila, I., (2007). Examining the effects of contextual factors on TQM and performance through the lens of organizational theories: An empirical study. *Journal of Operations Management*, 2007, Vol.25 (1), pp.83-109
- Sousa, R., Voss, C., (2002). Quality management re-visited: a reflective review and agenda for future research. *Journal of Operations Management*, 2002, Vol.20 (1), pp.91-109
- Theussl, A., Feinerer, I., Hornik, K. (2012) A tm Plug-In for Distributed Text Mining in R. *Journal of statistical software* vol: 51 iss:5 [1548-7660]
- Taylor, F.W. (1911). *The principles of scientific management*. Engineering & Management Press, Atlanta
- Van Der Wiele, A., Williams, A.R.T., & Dale, B.G. (2000). Total Quality Management: Is it a fad, fashion, or fit? *ASQ Quality Management Journal*, 7(2), 65–79.
- Weiss, S. M., Indurkha, N. & Zhang, T. (2012). *Fundamentals of Predictive Text Mining*. London: Springer London

Xie, P., Xing, E. (2013) Integrating Document Clustering and Topic Modeling., Uncertainty In Artificial Intelligence - Proceedings of the Twenty-Ninth Conference (2013), Bellevue, Washington, U.S.

Zhang, D., Linderman, K., Schroeder, R. G., (2012). The moderating role of contextual factors on quality management practices. Journal of Operations Management, 2012, Vol.30 (1-2), pp.12-23

Appendix A - Generated topic models
80 Topics (Topic 61- 80)

Topic 61	Topic 62	Topic 63	Topic 64	Topic 65	Topic 66	Topic 67	Topic 68	Topic 69	Topic 70	Topic 71	Topic 72	Topic 73	Topic 74	Topic 75	Topic 76	Topic 77	Topic 78	Topic 79	Topic 80	
requirements	product	information	services	system	service	value	control	environmental	improvement	development	management	system	quality	lean	tools	education	management	health	creativity	
process	products	quality	transport	design	quality	service	process	performance	process	leadership	strategy	competitiveness	management	six	information	students	business	care	governance	
development	brand	can	wellness	can	satisfaction	resources	measurement	management	team	new	organizations	index	system	sigma	quality	higher	quality	patients	blocks	
product	customers	control	time	analysis	customer	logic	data	practices	management	process	strategic	development	systems	call	process	university	companies	patient	transformation	
can	new	fuzzy	freight	engineering	customers	new	chart	studies	continuous	project	organization	management	iso	training	data	quality	global	healthcare	qmod	
management	customer	one	market	elements	attributes	customers	can	case	production	technology	organizational	quality	organization	management	system	universities	value	hospital	building	
methods	can	data	one	method	journal	network	value	ems	journal	teams	innovation	model	indicators	requirements	company	knowledge	student	ecosystem	medical	excellence
processes	market	sample	operating	components	performance	creation	analysis	study	implementation	industrial	framework	grant	processes	supply	techniques	course	corporate	hospitals	time	
customer	business	process	quality	model	marketing	public	system	firms	organization	fig	business	chaordic	organizations	chain	tool	institutions	new	staff	can	
information	quality	paper	czas	approach	services	system	statistical	measures	learning	ldi	performance	national	standards	demand	analysis	teaching	turnaround	improvement	measurement	
und	company	galletto	coopetition	decode	model	resource	values	measurement	processes	product	control	knowledge	products	customer	application	academic	siemens	services	process	
different	value	scale	enterprises	methods	vol	marketing	quality	organisations	study	engineering	based	leadership	safety	companies	manufacturing	study	science	process	measures	
model	marketing	use	order	using	management	customer	production	organisation	development	can	can	organizations	business	suppliers	method	research	capital	quality	productivity	
figure	may	method	improvement	functions	expectations	cocreation	charts	companies	improvements	one	competitive	based	standard	customers	user	learning	can	treatment	system	
method	also	used	value	figure	research	services	samples	organisational	company	change	approach	process	work	improvement	criteria	educational	approach	can	elearning	
well	experience	reliability	oeo	risk	servqual	innovation	parameters	impact	project	time	resources	using	product	value	will	courses	development	study	performance	
quality	will	design	competitive	task	gaps	actors	measurements	based	kaizen	enterprises	resource	next	activities	process	one	teachers	accessed	model	learning	
possible	needs	two	chain	surface	hotel	vargo	distribution	research	work	companies	change	used	can	also	used	programs	ability	clinical	new	
der	japanese	new	production	concept	gap	business	model	journal	factors	business	culture	flowers	improvement	work	problems	faculty	market	service	use	
enterprise	time	many	figure	one	staff	lusch	measured	systems	tools	nursing	research	qmodicqss	organizational	risk	problem	evaluation	recession	information	way	

Appendix A - Generated topic models

Topic models per year with 20 topics per year

2003

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8	Topic 9	Topic 10	Topic 11	Topic 12	Topic 13	Topic 14	Topic 15	Topic 16	Topic 17	Topic 18	Topic 19	Topic 20
quality	management	management	knowledge	quality	management	manufacturing	quality	qualit	quality	management	management	service	knowledge	environmental	performance	certification	brand	process	inspection
education	organizational	new	process	management	companies	development	tqm	plus	iso	companies	model	customer	work	management	management	iso	brands	information	process
siemens	culture	development	innovation	service	process	new	management	management	management	lever	quality	quality	company	excellence	tqm	performance	model	quality	quality
management	privatization	environmental	management	customer	processes	process	performance	processus	change	development	excellence	customers	quality	social	quality	quality	emotional	processes	control
business	organization	environment	practice	banking	quality	firms	practices	dune	work	training	people	satisfaction	can	fieldbeing	evaluation	sales	customer	improvement	dfss
higher	reform	waste	learning	internet	business	innovation	total	dun	organisational	managers	results	attributes	management	new	organisations	firms	associations	can	sigma
development	decision	product	vol	services	engineering	product	implementation	comme	system	using	organizations	technology	creativity	field	system	logistics	loyalty	order	design
university	thai	industrial	brand	qfd	strategy	products	vol	acteurs	process	new	improvement	vol	new	principle	diversity	management	rational	model	six
excellence	market	can	one	can	new	management	customer	ainsi	level	number	performance	hotel	competencies	organisation	research	value	quality	organization	can
tqm	factors	company	different	customers	company	vol	study	SANT	organisations	people	control	services	also	will	business	cost	strength	organizational	measurement
australia	action	economic	mgt	process	system	research	employees	cest	loyalty	process	selfassessment	marketing	lewis	development	organisational	logistic	value	enterprise	manufacturing
institutions	success	products	quality	system	project	business	research	direction	customers	also	organization	journal	learning	performance	criteria	impact	variables	management	figure
sustainable	industry	production	processes	requirements	development	organizations	change	joy	organisation	results	criteria	level	people	enterprises	systems	chain	relationships	simulation	different
top	will	sustainable	business	information	erp	organizational	progress	place	performance	performance	approach	research	creative	quality	total	supply	consumers	development	costs
corporate	research	one	managers	tqm	organisation	performance	manufacturing	aussi	important	group	organizational	management	will	one	appraisal	abnormal	product	person	vol
program	assumptions	innovation	practitioners	knowledge	systems	work	table	certification	data	practices	areas	providers	person	environment	survey	table	evaluations	following	may
implementation	telecommunica	companies	job	iso	results	strategy	award	dmarche	case	study	strategy	guests	possible	concept	employees	certified	will	sustainable	charts
australian	may	process	stage	organization	develop	market	analysis	dveloppement	employees	program	efqm	model	experience	can	literature	number	company	layer	value
improvement	government	systems	strategy	level	information	journal	journal	entre	can	information	process	order	ideas	principles	organisation	enterprises	consumer	employees	machining
also	telephone	ecological	organizational	systems	iso	small	used	sysme	new	smes	processes	factors	theory	philosophy	context	better	can	will	one

2006

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8	Topic 9	Topic 10	Topic 11	Topic 12	Topic 13	Topic 14	Topic 15	Topic 16	Topic 17	Topic 18	Topic 19	Topic 20
model	value	social	management	public	requirements	benchmarking	quality	quality	time	innovation	indicators	knowledge	smes	design	sigma	product	time	change	health
quality	service	governance	project	contract	project	management	management	process	excellence	management	work	development	supply	process	six	quality	quality	performance	care
information	quality	enterprise	stakeholder	process	improvement	process	tqm	education	problem	people	schools	learning	chain	sigma	quality	design	strategic	management	patient
communication	customer	enterprises	maturity	service	figure	organizations	iso	management	production	process	culture	competence	quality	six	tqm	kansei	planning	measurement	patients
efqm	customers	quality	model	pqi	projects	business	excellence	improvement	building	bpi	quality	work	scm	dfss	problem	can	system	business	model
excellence	ikea	public	stakeholders	relationship	results	can	companies	can	can	employees	values	one	management	customer	improvement	competitive	school	process	management
management	added	key	excellence	ontology	quality	success	organizations	university	demand	managers	tqm	can	results	development	management	engineering	building	measures	medical
decision	can	education	process	figure	network	may	business	universities	creativity	business	service	process	customer	phase	japanese	products	model	project	process
implementation	product	policy	within	management	neural	organization	improvement	level	transformation	organizations	management	manager	product	product	approach	feeling	excellence	research	activities
internal	management	management	system	relationships	model	ideas	systems	performance	manufacturing	work	state	different	performance	new	risk	analysis	transformation	improvement	can
use	values	staff	interests	domain	analysis	one	implementation	excellence	system	study	sector	research	research	can	can	building	work	system	service
method	will	stakeholders	organisation	transport	can	factors	organization	system	will	within	two	management	small	variation	tools	control	information	systems	flow
can	costs	institutional	organisational	authority	impacts	creativity	process	based	flow	model	can	transformation	companies	quality	lean	transformation	creativity	excellence	occupational
need	food	organization	also	new	data	will	human	output	process	excellence	bahraini	new	planning	processes	project	advantage	blocks	organisations	time
conflict	time	stakeholder	quality	parties	organization	transformation	selfassessment	used	blocks	development	results	building	focus	tools	system	qmod	qmod	strategic	also
case	products	ethical	organisations	quality	different	workers	can	model	tools	new	development	sharing	lack	tools	system	qmod	qmod	strategic	also
organization	experience	excellence	governance	building	excellence	also	internal	using	one	leadership	research	people	practices	used	activities	excellence	jit	action	services
selfassessment	production	include	may	also	requirement	building	governance	processes	qmod	figure	oasi	common	financial	concept	systems	blocks	enhancement	organisation	quality
group	also	theory	level	one	blocks	qmod	total	potential	governance	results	maria	creativity	marketing	engineering	variation	reliability	law	strategy	clinical
different	building	renal	performance	governance	transformation	blocks	vol	work	also	customer	new	blocks	study	control	may	creativity	governance	processes	information

2007

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8	Topic 9	Topic 10	Topic 11	Topic 12	Topic 13	Topic 14	Topic 15	Topic 16	Topic 17	Topic 18	Topic 19	Topic 20
business	values	quality	kansei	design	iso	product	risks	quality	embed	process	service	research	service	management	reliability	kansei	quality	team	quality
management	police	management	design	can	quality	used	chain	management	equation	ikea	customers	process	customer	quality	engineering	design	management	development	university
quality	organizational	system	model	kansei	certification	figure	design	customers	can	orientation	patients	innovation	value	leadership	development	product	education	teams	course
information	uncertainty	award	rough	mystery	management	surface	demand	tqm	method	improvement	nurses	may	production	control	analysis	engineering	excellence	process	management
social	life	model	decision	label	sectors	data	supply	customer	process	processes	data	work	customers	business	system	value	higher	quality	universities
organizations	face	cost	data	human	sites	design	risk	service	inspection	treatment	care	approach	services	process	experiences	shape	university	product	process
sustainability	information	new	table	analysis	certified	visual	can	activities	order	international	also	one	lean	statistical	design	figure	model	management	healthcare
csr	organization	process	engineering	also	excellence	using	experiments	process	figure	organization	waiting	change	management	work	methods	concept	process	can	degree
corporate	uics	results	product	process	sector	system	knowledge	within	picking	variation	health	knowledge	business	analysis	technical	products	performance	new	can
organization	quality	efqm	variables	aibo	organisations	kansei	computer	ideas	time	management	satisfaction	can	marketing	results	can	analysis	vol	kaizen	students
public	public	excellence	set	user	qms	products	factors	new	different	project	quality	new	can	model	product	method	learning	six	teaching
also	prediction	product	rules	data	implementation	results	model	focus	equationsmt	warfarin	patient	also	design	performance	countermeasure	quality	research	group	managers
companies	enterprises	performance	colour	results	diffusion	semantic	management	production	values	also	entrepreneursh	way	new	new	commercial	experience	knowledge	sigma	level
responsibility	scatter	one	sets	embed	benefits	study	simulation	experience	using	local	worth	learning	using	change	production	concepts	improvement	health	research
stakeholders	model	immune	can	experimental	business	analysis	production	different	data	firms	outpatient	innovations	research	methods	structure	can	processes	research	study
stakeholder	service	costs	printing	users	companies	engineering	process	research	visiodrawing	patients	services	will	principles	organizations	failure	customers	system	dfss	table
company	can	market	using	different	bhel	surfaces	may	total	requirements	marketing	mean	development	processes	can	maintenance	shapes	academic	work	programs
crisis	vision	manufacturing	factor	time	number	can	example	can	one	china	public	action	case	improvement	experience	affective	institutions	employees	number
environment	work	figure	analysis	emotional	commodity	flowers	one	one	used	one	office	organisational	journal	ergonomic	components	new	organization	decision	improvement
communication	analysis	firms	entropy	words	different	stimuli	decision	product	therefore	six	customer	first	process	services	process	customer	efqm	company	total

Appendix A - Generated topic models

2008

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8	Topic 9	Topic 10	Topic 11	Topic 12	Topic 13	Topic 14	Topic 15	Topic 16	Topic 17	Topic 18	Topic 19	Topic 20
kansei	management	design	value	work	service	requirements	quality	patient	quality	lean	quality	service	process	product	system	business	service	university	process
design	integration	analysis	management	management	values	design	education	care	employee	management	management	quality	quality	complexity	management	sustainability	quality	education	management
engineering	iso	results	quality	local	customers	product	research	health	customer	sigma	excellence	orientation	failure	can	process	csr	factors	universities	processes
new	competences	can	business	system	value	requirement	will	quality	attractive	six	model	patients	production	customer	model	sustainable	model	research	business
affective	integrated	control	innovation	authorities	customer	development	organization	process	can	production	values	organizational	product	perspective	concept	social	customer	programs	model
product	organizations	model	customers	quality	business	may	level	satisfaction	ideas	supply	tqm	vol	decision	stakeholders	systems	businesses	customers	activities	control
factor	systems	used	vol	wage	edvardsson	information	also	service	employees	companies	organizational	services	development	demands	engineering	also	tolerance	higher	strategy
feelings	processes	experimental	relationship	environmental	ikea	can	learning	sigma	satisfaction	chain	vol	customer	call	value	requirements	environmental	zone	students	quality
analysis	quality	figure	can	development	experience	experience	development	research	customers	practices	business	health	manufacturing	facility	approach	performance	bank	new	key
can	level	kansei	customer	sustainability	valuesbased	visual	can	six	organizations	improvement	principles	care	problem	system	different	corporate	research	study	armistead
experiences	karapetrovic	ahp	system	employees	enquist	interior	industry	patients	work	quality	organization	dimensions	mode	development	design	development	performance	can	based
service	innovation	regression	service	standard	dimensions	customer	factors	improvement	will	business	can	journal	evaluation	necessary	can	starbucks	vol	phase	strategic
quality	important	methods	map	weibull	vol	marketing	students	will	pakistan	tqm	one	provided	methods	one	quality	value	satisfaction	management	framework
servicescape	companies	samples	model	systems	company	criteria	individual	services	political	also	performance	hospitals	customer	different	generic	responsibility	important	entrepreneurial	performance
products	procedures	models	iso	ems	research	products	pakistan	industries	companies	concept	company	public	time	culture	step	company	importance	faculty	improvement
needs	results	factors	improvement	organization	logic	new	leadership	hospital	company	can	vision	expectations	cases	sustainability	steps	new	hysteresis	number	can
interactions	msss	fft	product	swedish	environment	stakeholder	sector	can	microwave	results	will	sector	reliability	important	verlag	can	journal	also	company
environment	competence	simulation	new	one	marketing	framework	skills	acceptability	people	two	yes	mean	can	cement	one	companies	kano	business	journal
model	implemented	evaluation	excellence	responsibility	focus	research	problems	centeredness	instability	company	mission	factor	approach	also	winzer	management	banking	florence	tools
feeling	different	table	creation	emss	model	concept	software	results	ethical	performance	results	perceptions	one	company	figure	environment	table	entrepreneursh	vol

2009

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8	Topic 9	Topic 10	Topic 11	Topic 12	Topic 13	Topic 14	Topic 15	Topic 16	Topic 17	Topic 18	Topic 19	Topic 20
performance	service	product	value	innovation	pakistan	quality	embed	process	business	service	quality	project	event	tourism	education	quality	quality	quality	environmental
model	lean	process	can	business	quality	students	equationdsmt	management	management	quality	service	public	price	hotel	university	management	education	board	social
results	quality	development	housing	new	software	learning	conflict	business	leadership	marketing	satisfaction	procurement	customers	tourist	students	improvement	higher	patients	water
management	public	can	public	network	productivity	elearning	quality	mergeformat	employees	customer	journal	can	experience	data	knowledge	production	iso	hospital	public
trust	research	design	community	company	can	kansei	group	processes	work	services	vol	costs	value	can	technical	work	vol	care	sustainable
leadership	vol	risk	partnerships	research	will	course	decision	improvement	employee	customers	customer	services	research	requirements	system	tqm	management	health	development
process	customer	requirements	customer	management	engineering	can	ref	strategy	can	market	model	analysis	expectations	governance	development	managers	institutions	information	information
can	contracts	quality	knowledge	logistics	people	process	information	implementation	change	orientation	analysis	cost	consumers	corporate	universities	approach	academic	management	network
call	management	risks	process	also	information	one	management	system	development	management	marketing	graduates	perception	research	quality	tools	standards	patient	university
customer	transport	project	companies	enterprise	industry	knowledge	process	iso	people	gaps	research	results	also	law	research	total	leadership	lego	also
research	customers	customer	online	market	time	data	groups	systems	organizational	business	loyalty	quality	consumer	different	skills	manufacturing	research	process	new
figure	business	dfss	services	enterprises	research	use	seq	performance	organization	research	factor	used	management	italian	higher	lean	performance	new	problems
indicators	can	analysis	social	technology	also	sample	vol	vol	corporate	can	bank	contract	different	management	process	study	excellence	education	communication
based	journal	model	vol	park	tgp	galetto	task	information	practices	journal	variables	agencies	information	verona	entrepreneursh	analysis	university	services	environment
measurement	services	different	management	can	government	new	level	approach	values	higher	scale	one	may	tourists	workers	techniques	assurance	change	responsibility
quality	contract	one	new	competitive	management	control	tool	work	human	expectations	customers	level	products	also	market	process	institutional	healthcare	people
brand	relationship	safety	service	firms	vip	didactic	satisfaction	key	ethical	relationship	results	efficiency	marketing	facility	work	knowledge	universities	group	can
data	study	well	citizens	development	system	used	mteqn	strategic	ethics	gap	services	value	positive	model	courses	research	international	research	organizations
used	value	concept	local	companies	organization	teachers	time	model	company	company	items	information	customer	quality	needs	also	roles	one	facebook
one	environment	used	also	economic	hyperlink	scientific	tools	order	vol	zong	banks	process	can	italy	will	can	journal	focus	case

2010

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8	Topic 9	Topic 10	Topic 11	Topic 12	Topic 13	Topic 14	Topic 15	Topic 16	Topic 17	Topic 18	Topic 19	Topic 20
service	company	transport	quality	dmu	innovation	quality	sigma	change	management	services	customer	maturity	environmental	quality	performance	work	quality	process	management
quality	development	services	management	performance	promotor	project	six	management	quality	wellness	service	development	ems	management	management	development	management	projects	employees
vol	dfss	freight	iso	organisation	networks	construction	improvement	culture	customer	customer	value	process	management	contracts	company	sustainable	process	tools	systems
research	product	market	organizations	efficiency	network	customer	quality	organizational	improvement	value	customers	can	studies	commodity	process	change	processes	phase	standards
customer	team	enterprises	tqm	companies	role	can	care	quality	focus	customers	crises	requirements	study	science	benchmarking	sustainability	improvement	project	iso
journal	project	competitive	companies	results	roles	value	healthcare	leadership	system	chart	professional	model	benefits	products	knowledge	different	iso	sigma	organizations
model	process	organic	principles	healthcare	tourism	work	different	best	process	providers	services	quality	companies	development	business	one	system	six	job
customers	also	farmers	certification	management	potential	experiences	managers	practice	isaksson	service	can	reference	organisations	can	study	also	care	sme	system
services	lean	competitiveness	customer	quality	enterprise	model	one	tqm	customers	ewma	organizational	class	iso	commercial	customer	human	health	used	retention
dimensions	can	one	vol	factors	project	system	statistical	based	development	data	competence	system	unknown	new	improvement	university	approach	analysis	turnover
marketing	tools	lithuania	values	organisational	process	management	county	organization	cement	control	crisis	set	survey	also	development	can	results	realization	factors
study	information	lithuanian	results	can	evaluation	projects	work	within	recession	chain	experience	value	systems	public	quality	capital	business	control	will
measurement	new	network	health	organisations	management	approach	council	organizations	performance	nonnormal	product	processes	case	business	indicators	social	customer	level	new
satisfaction	design	requirements	can	model	research	lowest	applications	corporate	business	one	harm	methods	also	transport	total	new	analysis	optimization	integrated
servqual	model	get	journal	care	concerning	attractive	within	implementation	block	distribution	management	indicators	small	etc	system	concept	focus	results	integration
process	focus	agriculture	company	members	performance	procurement	program	project	can	process	business	figure	performance	system	excellence	approach	one	quality	safety
indicator	group	produces	organization	improvement	development	will	organization	can	supply	can	marketing	forming	journal	research	information	quality	used	concept	employee
university	one	service	implementation	health	new	kano	physicians	one	potential	activities	creation	design	improvements	managed	human	universities	principles	phases	organization
experience	different	science	service	education	cooperation	one	can	research	measurement	mean	competencies	management	countries	example	product	swedish	company	management	industry
issue	projects	sector	leadership	excellence	people	tqm	organizational	methodologies	building	order	provider	product	impact	established	satisfaction	management	articles	measurement	companies

Appendix A - Generated topic models

2011

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8	Topic 9	Topic 10	Topic 11	Topic 12	Topic 13	Topic 14	Topic 15	Topic 16	Topic 17	Topic 18	Topic 19	Topic 20
innovation	care	quality	management	iso	education	knowledge	quality	system	performance	performance	value	process	work	satisfaction	method	product	quality	service	process
performance	health	customer	quality	management	market	development	management	efficiency	prevention	management	performance	project	students	customer	methods	quality	service	customers	management
companies	patient	kano	model	environmental	orientation	university	tqm	functions	target	practices	service	change	environment	quality	risk	products	research	satisfaction	vol
management	patients	qfd	excellence	quality	higher	education	service	fuzzy	public	human	practices	improvement	improvement	service	can	development	customers	customer	production
business	service	management	results	companies	vol	management	improvement	dmu	directors	relationship	customer	management	working	customers	process	management	analysis	quality	research
strategy	ideas	design	leadership	integration	students	research	model	can	targets	business	management	processes	consciousness	journal	measurement	tools	variables	job	relationships
system	development	requirements	approach	systems	research	sustainable	inspection	decode	year	hrm	research	lean	reliability	attributes	figure	costs	services	experience	lean
success	can	product	organization	system	marketing	change	process	components	branches	effectiveness	services	employees	group	items	analysis	system	table	commitment	supply
value	services	development	health	firms	university	social	fuzzy	variables	counter	studies	case	six	page	value	page	customers	factor	journal	website
model	knowledge	model	public	standard	journal	sustainability	page	page	sales	study	based	continuous	project	marketing	evaluation	process	study	respondents	adoption
technology	choice	group	process	benefits	management	business	journal	elements	information	line	journal	bpr	mean	services	element	analysis	groups	company	quality
suggestions	hospital	new	page	vol	business	vision	performance	model	independent	organizational	measurement	sigma	results	loyalty	development	data	regression	research	page
development	also	organizations	organizations	standards	learning	students	planning	kitves	indicators	managers	public	new	engineering	model	system	tool	shopping	ltd	processes
product	customer	project	organizational	study	institutions	process	systems	approach	reinigung	model	new	business	personality	can	used	customer	satisfaction	ethical	information
export	page	team	can	certification	academic	new	learning	quality	pharmacy	employee	transport	time	participants	research	data	cost	group	management	journal
page	different	customers	award	implementation	universities	mining	results	number	rts	companies	measures	company	survey	perceived	software	page	page	services	kaizen
yes	healthcare	satisfaction	also	international	qfd	also	processes	using	ratcheting	resource	page	page	entrepreneurshi	level	human	application	customer	employees	factors
based	information	case	development	performance	international	work	principles	processes	employees	analysis	system	team	research	page	possible	can	vol	organizational	also
organizations	vol	certification	improvement	company	approach	one	can	engineering	results	strategic	network	people	quality	dimensions	one	production	model	transteda	business
processes	medical	needs	implementation	journal	course	environmental	based	product	setting	journal	marketing	work	time	online	using	requirements	journal	behavior	innovation

2012

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8	Topic 9	Topic 10	Topic 11	Topic 12	Topic 13	Topic 14	Topic 15	Topic 16	Topic 17	Topic 18	Topic 19	Topic 20
health	process	sigma	learning	customer	social	data	satisfaction	education	management	process	index	capital	service	customer	lean	process	quality	quality	quality
care	management	six	organizational	patients	index	index	service	students	satisfaction	control	process	csr	innovation	supply	leadership	project	management	life	project
quality	processes	management	change	satisfaction	european	analysis	quality	quality	index	product	incidents	social	value	chain	work	research	innovation	tourism	index
healthcare	business	index	improvement	patient	local	information	customer	research	company	quality	time	knowledge	resources	lean	management	development	tqm	index	software
private	quality	process	culture	service	innovation	authors	index	course	excellence	index	blocks	index	resource	management	production	innovation	iso	cultural	process
system	index	improvement	employees	index	development	quality	customers	healthcare	tools	development	figure	business	new	quality	index	index	vol	excellence	management
index	organizations	factors	management	care	firms	decision	journal	system	quality	innovation	data	corporate	quality	research	improvement	system	research	case	ppap
measurement	development	implementation	organization	staff	new	production	model	index	ben	measurement	improvement	management	process	service	implementation	companies	index	qmod	effectiveness
sector	model	lean	index	data	knowledge	can	authors	safety	business	new	error	intellectual	index	customers	employees	product	performance	authors	development
services	organization	quality	new	process	cultural	set	research	staff	model	design	block	can	customers	index	case	new	study	can	papers
regulation	system	tools	work	can	food	back	factor	learning	criteria	manufacturing	medication	qmod	employees	authors	also	knowledge	companies	kaizen	research
public	level	projects	projects	quality	research	rules	marketing	student	process	data	analysis	people	customer	vol	can	planning	journal	vol	authors
medical	also	company	organizations	health	project	papers	study	skills	organization	phase	cement	practice	logic	involvement	behaviors	different	business	change	metaanalysis
institutions	information	authors	process	improvement	conference	home	scale	higher	techniques	multivariate	authors	model	testdriving	journal	different	will	principles	home	back
service	systems	project	can	healthcare	economic	conference	used	management	employees	authors	factors	human	firm	process	authors	study	organizations	personal	home
analysis	performance	success	model	icu	can	qmod	factors	method	employee	can	conference	back	journal	services	enterprise	vol	service	back	conference
average	can	back	quality	authors	back	process	data	training	customer	back	model	home	marketing	articles	system	time	authors	sustainable	qmod
indicators	iso	companies	creativity	back	papers	figure	results	hospital	improvement	variation	papers	authors	law	specifications	organization	activities	organizational	research	production
values	conference	home	different	claims	home	table	home	university	organizational	papers	distribution	papers	firms	can	papers	value	improvement	people	effect
authors	work	papers	knowledge	performance	companies	service	back	also	framework	conference	home	conference	market	implementation	home	can	continuous	dovista	model

2013

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8	Topic 9	Topic 10	Topic 11	Topic 12	Topic 13	Topic 14	Topic 15	Topic 16	Topic 17	Topic 18	Topic 19	Topic 20
service	requirements	management	quality	information	vol	students	time	performance	learning	sigma	process	social	leadership	iso	lean	service	quality	process	online
quality	management	tqm	improvement	healthcare	service	education	production	management	quality	quality	management	media	value	management	management	quality	management	innovation	quality
services	development	quality	work	safety	study	quality	alliances	attributes	work	six	quality	product	university	quality	system	customer	project	management	product
model	new	business	training	management	journal	higher	success	systems	development	service	customer	can	public	companies	care	satisfaction	system	maturity	business
satisfaction	purchase	excellence	health	quality	knowledge	student	strategic	system	index	management	processes	communication	development	innovation	health	customers	systems	level	shopping
ecosystem	knowledge	organization	study	learning	quality	study	tps	integration	university	control	product	development	research	performance	production	journal	research	business	platform
attributes	smes	organizational	can	hospital	queue	satisfaction	alliance	value	based	process	value	design	leaders	vol	work	vol	projects	service	control
value	network	performance	knowledge	system	consumer	performance	factors	importance	competitiveness	monthly	project	use	strategic	certification	working	loyalty	risk	organizations	research
customer	strategy	culture	process	medical	waiting	results	improvement	organizations	back	method	case	marketing	business	study	organizations	marketing	requirements	processes	system
management	international	strategy	fmea	organizational	consumers	analysis	failure	case	national	culture	information	next	excellence	practices	process	market	model	operational	can
core	different	model	care	model	will	university	oeo	service	new	performance	implementation	grid	social	research	ethics	orientation	sustainability	results	information
index	research	results	case	hospitals	factors	teachers	czas	chaordic	knowledge	exploration	can	order	employees	data	swedish	maintenance	can	organization	index
business	product	employees	order	patient	sharing	course	quality	analysis	within	papers	manufacturing	page	stakeholders	effects	page	services	page	orientation	sellors
can	study	satisfaction	company	organization	information	teaching	operating	research	exam	creativity	system	companies	new	organizations	also	study	approach	chain	uncertainty
hospital	case	people	group	next	practices	back	performance	integrated	life	exploitation	new	back	vol	standards	value	model	index	supply	buyers
factors	strategic	employee	one	care	time	papers	next	study	papers	analysis	business	competences	local	analysis	implementation	research	next	bpo	entrepreneurshi
needs	licensing	efqm	next	education	research	next	system	innovations	research	index	next	index	theory	standard	leadership	perceived	turnaround	turnover	social
next	engineering	strategic	different	method	can	qmodicqss	losses	next	next	firms	index	qmodicqss	next	index	people	index	papers	research	education
patients	project	organisational	time	page	customer	page	implementation	improvement	page	articles	back	papers	model	business	next	page	safety	new	chinese
study	can	organizations	back	index	management	index	total	ipa	qmodicqss	improvement	papers	networks	approach	next	principles	management	level	next	factors

Appendix B – Most common words

The 100 most common words in the complete dataset

	Word	Frequency			Word	Frequency
1	quality	34084		51	important	5238
2	management	27665		52	factors	5122
3	process	17192		53	use	5082
4	service	15401		54	systems	5014
5	can	13267		55	table	4967
6	research	11210		56	iso	4927
7	customer	10916		57	two	4926
8	business	10468		58	papers	4915
9	model	9237		59	education	4888
10	performance	9236		60	order	4879
11	development	9224		61	employees	4779
12	also	9147		62	tqm	4775
13	new	9140		63	organizational	4761
14	system	8899		64	first	4642
15	value	8178		65	excellence	4590
16	study	8052		66	production	4503
17	one	7971		67	people	4347
18	results	7734		68	paper	4297
19	improvement	7505		69	implementation	4289
20	vol	7485		70	international	4238
21	analysis	7320		71	well	4209
22	work	7213		72	using	4208
23	different	7112		73	control	4184
24	product	7090		74	number	4154
25	information	7080		75	lean	4151
26	journal	7054		76	total	4138
27	customers	7011		77	back	4126
28	processes	7003		78	within	4126
29	used	6514		79	social	4125
30	knowledge	6455		80	page	4077
31	based	6425		81	may	4070
32	time	6292		82	market	3869
33	will	6254		83	marketing	3858
34	organization	6213		84	learning	3851
35	index	6118		85	health	3827
36	data	5972		86	way	3795
37	level	5938		87	activities	3794
38	satisfaction	5921		88	values	3776
39	companies	5870		89	requirements	3773
40	services	5742		90	leadership	3700
41	company	5716		91	however	3675
42	project	5566		92	focus	3640
43	approach	5554		93	many	3625
44	design	5550		94	needs	3617
45	change	5442		95	higher	3530
46	innovation	5397		96	products	3521
47	figure	5344		97	method	3514
48	organizations	5338		98	authors	3494
49	case	5269		99	public	3478
50	university	5241		100	six	3458

Appendix B – Most common words

The 100 most common words 2003

	Word	Frequency			Word	Frequency
1	quality	2942		51	employees	393
2	management	2687		52	approach	389
3	process	1455		53	value	386
4	performance	1136		54	number	380
5	can	935		55	table	380
6	customer	925		56	products	378
7	new	885		57	use	378
8	tqm	844		58	technology	376
9	development	836		59	excellence	374
10	service	834		60	total	373
11	knowledge	766		61	organizations	363
12	business	759		62	case	360
13	model	688		63	manufacturing	359
14	processes	664		64	practices	357
15	companies	663		65	implementation	355
16	also	655		66	two	351
17	information	648		67	analysis	343
18	vol	642		68	services	342
19	one	623		69	data	334
20	customers	616		70	organisation	332
21	results	583		71	design	330
22	company	579		72	university	328
23	research	570		73	managers	325
24	will	564		74	environment	324
25	system	547		75	factors	318
26	improvement	539		76	environmental	317
27	iso	518		77	many	315
28	work	506		78	market	315
29	control	504		79	however	313
30	organization	484		80	certification	304
31	different	481		81	way	303
32	level	478		82	organisations	299
33	product	474		83	figure	298
34	study	469		84	practice	293
35	systems	464		85	international	291
36	organizational	446		86	production	288
37	innovation	442		87	organisational	287
38	satisfaction	433		88	industry	275
39	people	430		89	criteria	273
40	based	426		90	focus	273
41	change	422		91	well	273
42	brand	421		92	must	270
43	used	418		93	possible	270
44	time	413		94	evaluation	269
45	important	411		95	following	263
46	journal	411		96	key	260
47	may	411		97	structure	260
48	order	409		98	activities	256
49	strategy	409		99	learning	252
50	first	399		100	needs	252

Appendix B – Most common words

The 100 most common words 2006

	Word	Frequency			Word	Frequency
1	quality	2857		51	figure	488
2	management	2450		52	systems	487
3	process	1741		53	control	472
4	excellence	1233		54	companies	468
5	can	1223		55	care	433
6	governance	1066		56	benchmarking	431
7	building	1004		57	people	428
8	transformation	962		58	customers	426
9	creativity	945		59	case	424
10	model	941		60	approach	419
11	qmod	940		61	journal	417
12	blocks	930		62	level	416
13	performance	874		63	two	416
14	knowledge	841		64	order	415
15	improvement	824		65	company	414
16	business	804		66	however	413
17	new	794		67	implementation	413
18	system	781		68	activities	390
19	service	775		69	vol	390
20	design	765		70	need	389
21	customer	739		71	way	386
22	value	726		72	requirements	385
23	also	723		73	learning	382
24	one	723		74	organisations	382
25	tqm	722		75	data	381
26	development	716		76	managers	378
27	time	691		77	organisation	377
28	product	679		78	factors	376
29	research	654		79	education	369
30	project	650		80	social	361
31	processes	640		81	needs	357
32	different	638		82	problem	354
33	work	622		83	many	353
34	change	615		84	strategic	351
35	will	608		85	products	350
36	health	601		86	measurement	348
37	used	586		87	focus	337
38	organization	580		88	well	336
39	information	572		89	table	334
40	six	558		90	first	333
41	study	545		91	using	333
42	sigma	543		92	employees	330
43	results	540		93	total	327
44	may	525		94	key	326
45	organizations	523		95	values	326
46	important	513		96	university	325
47	analysis	510		97	iso	323
48	based	508		98	tools	322
49	within	504		99	number	321
50	use	495		100	paper	320

Appendix B – Most common words

The 100 most common words 2007

	Word	Frequency			Word	Frequency
1	quality	3470		51	important	472
2	management	2209		52	products	471
3	can	1507		53	number	470
4	process	1432		54	case	469
5	design	1416		55	vol	468
6	service	1153		56	excellence	466
7	product	1085		57	journal	444
8	new	983		58	experience	442
9	customer	982		59	use	440
10	research	982		60	systems	436
11	system	935		61	level	433
12	also	924		62	concept	431
13	model	920		63	tqm	430
14	one	918		64	education	429
15	business	853		65	organizations	427
16	development	833		66	well	421
17	customers	825		67	services	418
18	figure	807		68	control	417
19	analysis	806		69	international	414
20	data	789		70	methods	414
21	value	761		71	company	406
22	different	756		72	many	405
23	used	754		73	way	405
24	results	745		74	activities	398
25	kansei	740		75	within	398
26	information	677		76	higher	397
27	study	675		77	iso	397
28	work	659		78	total	397
29	processes	645		79	three	387
30	university	624		80	technology	382
31	embed	620		81	example	381
32	engineering	610		82	change	380
33	approach	591		83	factors	380
34	two	590		84	evaluation	371
35	table	588		85	focus	366
36	production	577		86	part	360
37	time	577		87	people	357
38	using	562		88	however	355
39	will	561		89	reliability	353
40	based	554		90	social	353
41	performance	550		91	high	352
42	method	544		92	team	342
43	knowledge	538		93	see	338
44	improvement	529		94	needs	337
45	first	525		95	leadership	334
46	order	506		96	set	334
47	paper	493		97	decision	330
48	may	483		98	companies	329
49	organization	482		99	according	322
50	values	475		100	possible	319

Appendix B – Most common words

The 100 most common words 2008

	Word	Frequency		Word	Frequency
1	quality	2100	51	use	335
2	management	1664	52	order	324
3	service	1400	53	production	324
4	process	1270	54	first	323
5	can	1057	55	information	323
6	business	870	56	concept	321
7	customer	814	57	environment	317
8	research	813	58	engineering	310
9	model	798	59	organizational	310
10	value	745	60	control	306
11	design	701	61	two	306
12	also	693	62	lean	303
13	product	661	63	social	301
14	new	629	64	services	299
15	system	618	65	data	290
16	development	609	66	may	286
17	customers	595	67	number	284
18	will	582	68	excellence	283
19	one	575	69	table	282
20	processes	566	70	strategy	276
21	work	552	71	part	274
22	results	550	72	complexity	268
23	vol	536	73	focus	263
24	performance	520	74	case	262
25	values	512	75	sustainability	262
26	different	507	76	needs	261
27	based	505	77	international	260
28	study	490	78	within	260
29	analysis	488	79	people	258
30	organization	488	80	products	250
31	requirements	484	81	using	250
32	important	474	82	environmental	249
33	company	455	83	total	248
34	figure	449	84	industry	243
35	used	447	85	tqm	243
36	level	442	86	according	242
37	factors	418	87	stakeholders	242
38	journal	414	88	many	241
39	organizations	401	89	activities	240
40	satisfaction	392	90	marketing	240
41	companies	389	91	way	240
42	paper	388	92	innovation	237
43	university	385	93	six	237
44	approach	376	94	experience	236
45	systems	373	95	review	236
46	improvement	366	96	leadership	235
47	time	352	97	kansei	234
48	well	348	98	knowledge	234
49	education	347	99	need	232
50	employees	340	100	three	229

Appendix B – Most common words

The 100 most common words 2009

	Word	Frequency			Word	Frequency
1	quality	4725		51	people	726
2	management	3432		52	social	724
3	service	2769		53	processes	722
4	can	2100		54	project	717
5	process	2084		55	two	717
6	research	1907		56	well	714
7	business	1832		57	organization	696
8	customer	1496		58	change	693
9	also	1492		59	factors	688
10	vol	1426		60	international	683
11	new	1330		61	needs	677
12	education	1310		62	paper	657
13	services	1293		63	may	651
14	development	1224		64	students	649
15	one	1221		65	way	644
16	journal	1192		66	leadership	605
17	information	1188		67	design	590
18	model	1149		68	number	584
19	different	1143		69	table	584
20	will	1097		70	employees	579
21	analysis	1079		71	figure	578
22	value	1074		72	relationship	570
23	marketing	1048		73	using	565
24	results	1042		74	embed	562
25	customers	1032		75	however	561
26	study	1029		76	requirements	550
27	time	1019		77	group	548
28	system	1018		78	learning	547
29	work	1013		79	activities	545
30	university	1010		80	many	545
31	performance	976		81	high	542
32	level	946		82	possible	539
33	public	946		83	iso	532
34	knowledge	904		84	review	532
35	data	893		85	systems	518
36	used	890		86	literature	512
37	use	863		87	hyperlink	509
38	approach	856		88	lean	509
39	based	844		89	experience	508
40	case	842		90	total	508
41	improvement	835		91	part	504
42	satisfaction	828		92	tools	503
43	product	821		93	within	501
44	order	814		94	organizations	499
45	higher	811		95	studies	497
46	important	779		96	strategy	496
47	company	777		97	according	495
48	first	757		98	environment	494
49	market	755		99	innovation	487
50	companies	753		100	production	486

Appendix B – Most common words

The 100 most common words 2010

	Word	Frequency			Word	Frequency
1	quality	3085		51	culture	347
2	management	2276		52	knowledge	345
3	service	1513		53	total	344
4	process	1210		54	two	341
5	customer	1016		55	employees	338
6	can	947		56	factors	335
7	development	903		57	tqm	321
8	change	858		58	requirements	319
9	services	745		59	well	319
10	research	711		60	data	318
11	also	686		61	paper	313
12	one	655		62	environmental	311
13	business	638		63	studies	311
14	company	632		64	systems	311
15	new	631		65	implementation	303
16	performance	630		66	first	298
17	value	614		67	people	293
18	model	611		68	tools	290
19	study	595		69	use	289
20	system	589		70	projects	287
21	vol	556		71	six	285
22	improvement	554		72	health	283
23	work	522		73	table	280
24	different	521		74	international	278
25	results	515		75	order	277
26	customers	502		76	however	275
27	journal	491		77	methods	274
28	processes	488		78	production	273
29	product	486		79	transport	273
30	project	484		80	sustainable	268
31	analysis	459		81	satisfaction	266
32	companies	453		82	sigma	265
33	based	446		83	concept	263
34	used	446		84	needs	263
35	organization	444		85	using	262
36	iso	438		86	activities	260
37	approach	433		87	design	258
38	focus	422		88	evaluation	256
39	organizations	421		89	many	256
40	will	418		90	excellence	252
41	organizational	415		91	industry	251
42	innovation	400		92	best	250
43	time	392		93	values	249
44	within	381		94	case	248
45	market	373		95	number	248
46	level	367		96	resources	246
47	university	363		97	review	244
48	important	358		98	social	242
49	information	354		99	control	241
50	figure	353		100	leadership	241

Appendix B – Most common words

The 100 most common words 2011

	Word	Frequency			Word	Frequency
1	quality	4150		51	systems	763
2	management	3804		52	company	756
3	service	2234		53	organizations	743
4	process	2064		54	international	734
5	page	1874		55	organization	734
6	performance	1820		56	education	732
7	customer	1767		57	organizational	728
8	research	1755		58	will	723
9	can	1754		59	case	711
10	model	1491		60	market	709
11	journal	1420		61	employees	707
12	value	1392		62	use	707
13	system	1364		63	marketing	691
14	business	1332		64	first	679
15	development	1314		65	order	675
16	also	1234		66	important	663
17	improvement	1220		67	method	655
18	customers	1217		68	using	653
19	satisfaction	1197		69	within	647
20	study	1195		70	implementation	645
21	new	1184		71	number	645
22	results	1174		72	production	637
23	vol	1107		73	practices	628
24	one	1078		74	health	607
25	analysis	1068		75	environmental	606
26	iso	1064		76	public	595
27	papers	1015		77	well	592
28	book	1014		78	group	584
29	project	994		79	total	583
30	work	983		80	measurement	580
31	processes	974		81	methods	577
32	product	952		82	higher	576
33	used	951		83	environment	562
34	based	948		84	human	560
35	different	921		85	design	551
36	services	911		86	activities	549
37	companies	902		87	relationship	545
38	change	886		88	students	542
39	knowledge	873		89	care	517
40	factors	862		90	resources	517
41	information	860		91	according	515
42	level	832		92	people	512
43	time	830		93	tqm	512
44	table	828		94	focus	510
45	approach	827		95	high	502
46	figure	821		96	review	489
47	two	821		97	way	487
48	university	808		98	reliability	486
49	innovation	785		99	products	479
50	data	777		100	excellence	478

Appendix B – Most common words

The 100 most common words 2012

	Word	Frequency			Word	Frequency
1	quality	4198		51	production	865
2	management	3739		52	implementation	858
3	index	3682		53	customers	831
4	process	2841		54	factors	807
5	service	2271		55	learning	799
6	authors	2087		56	important	796
7	research	1842		57	companies	782
8	back	1823		58	product	781
9	papers	1817		59	organizations	780
10	home	1810		60	figure	777
11	can	1797		61	table	776
12	conference	1755		62	change	758
13	qmod	1725		63	systems	753
14	customer	1566		64	approach	728
15	lean	1417		65	use	723
16	business	1411		66	order	712
17	system	1390		67	social	702
18	also	1353		68	services	686
19	development	1323		69	two	683
20	improvement	1313		70	paper	671
21	new	1308		71	number	657
22	study	1280		72	three	638
23	innovation	1260		73	using	635
24	analysis	1187		74	control	632
25	data	1177		75	first	626
26	work	1164		76	international	615
27	satisfaction	1155		77	people	613
28	processes	1127		78	leadership	609
29	one	1099		79	total	602
30	value	1095		80	tools	598
31	different	1088		81	six	596
32	results	1072		82	well	596
33	organization	1064		83	within	590
34	performance	1060		84	university	589
35	model	1055		85	culture	588
36	information	1048		86	many	585
37	journal	1044		87	activities	583
38	vol	1026		88	healthcare	573
39	used	1022		89	iso	572
40	project	937		90	values	565
41	time	933		91	resources	563
42	company	927		92	method	559
43	based	920		93	way	552
44	knowledge	898		94	sigma	546
45	organizational	891		95	studies	544
46	case	887		96	focus	534
47	will	885		97	tqm	534
48	health	882		98	methods	524
49	level	875		99	care	517
50	employees	874		100	however	515

Appendix B – Most common words

The 100 most common words 2013

	Word	Frequency			Word	Frequency
1	quality	6557		51	organizational	928
2	management	5404		52	employees	919
3	process	3095		53	lean	917
4	service	2452		54	table	915
5	index	2048		55	systems	909
6	next	2022		56	implementation	897
7	research	1976		57	iso	880
8	business	1969		58	use	852
9	can	1947		59	tqm	825
10	back	1932		60	social	822
11	page	1929		61	six	817
12	papers	1906		62	will	816
13	qmodicqss	1876		63	students	813
14	study	1774		64	services	809
15	performance	1670		65	university	809
16	system	1657		66	excellence	786
17	customer	1611		67	paper	786
18	model	1584		68	education	783
19	results	1513		69	figure	773
20	development	1466		70	important	772
21	information	1410		71	company	770
22	new	1396		72	sigma	770
23	also	1387		73	total	756
24	value	1385		74	order	747
25	analysis	1380		75	learning	743
26	vol	1334		76	production	738
27	improvement	1325		77	people	730
28	based	1274		78	strategic	730
29	organization	1241		79	control	723
30	journal	1221		80	method	707
31	innovation	1202		81	first	702
32	work	1192		82	two	701
33	organizations	1181		83	using	696
34	processes	1177		84	requirements	692
35	satisfaction	1164		85	international	687
36	product	1151		86	related	655
37	level	1149		87	culture	650
38	companies	1131		88	values	646
39	project	1113		89	strategy	634
40	time	1085		90	health	630
41	one	1079		91	according	628
42	case	1066		92	higher	622
43	different	1057		93	change	612
44	knowledge	1056		94	well	610
45	data	1013		95	within	604
46	used	1000		96	practices	600
47	customers	967		97	marketing	597
48	factors	938		98	tools	578
49	approach	935		99	communication	577
50	leadership	928		100	activities	573