Re-imagining the Marketing Mix as a Tool for Developing Strategy (Supplementary Material)

Peter Prevos

ARTICLE HISTORY
Compiled November 12, 2018

1. Introduction

This document provides supplementary material for a paper submitted to the *Journal of Strategic marketing*. This supplementary material provides detailed technical notes about the research method, including computer code and data sources.

The methodology described in this document analyses a large corpus of journal abstracts by combining techniques from qualitative Grounded Theory and quantitative network analysis. This analysis reviews a corpus of over 200 journal abstracts published by the International Water Association. The results of this analysis were published in a book about the theory and practice water utility marketing (Prevos, 2017).

This document is written in Org Mode (9.1.14), which forms part of the GNU Emacs software (25.2.2). Org Mode enables reproducible research by combining prose with analytical computer code. Embedding computer code in research reports and publications allows readers to recreate the analyses and promotes reproducibility and peer review (Schulte, Davison, Dye, & Dominik, 2012; Peng, 2011). The Org Mode file used to compile this pdf version and the raw data are available from FigShare (Prevos, 2018).

2. Research Method

This report describes a method to systematically review of a body of literature, combining Grounded Theory coding techniques with mathematical analysis of networks (Wasserman & Faust, 1994; Wolfswinkel, Furtmueller, & Wilderom, 2011). This method is most effective for analysing a large number of short texts, such as abstracts or social media posts.

Wolfswinkel et al. (2011) described a five-stage approach to systematic literature review, following an approach based on Grounded Theory (Table 1). The first stage *defines* the research by identifying the criteria for inclusion, the fields of research, appropriate sources and the specific search terms. The second phase entails the actual *search* of databases, based on the criteria defined in the first phase. A representative sample of literature is *selected* in the third phase. The penultimate phase *analyses* the corpus of literature, after which the results can be *presented*.

The first three steps define a systematic method to select a sample of the literature. The analysis phase follows the three-step approach common to Grounded Theory. In this method the literature is analysed in the same way as qualitative Grounded Theory research. The method described in this paper introduces quantitative network analysis to detect patterns in the assigned codes.The final step involves presenting the results of the findings.

Table 1. Five-stage grounded-theory method for reviewing the literature in an area: to be used in an iterative fashion (Wolfswinkel, 2011).

Stage	Task
Define	Define the criteria for inclusion/exclusion
	Identify the fields of research
	Determine the appropriate sources
	Decide on specific search terms
Search	Search for a sample of literature
Select	Refine the sample
Analyse	Open coding
	Theoretical coding
	Selective coding
Present	Represent the structure and content
	Structure the article

2.1. Define

First defines the criteria for inclusion or exclusion of an article in the data set. Examples of these criteria are publications by a specific publisher, only journals with a minimum impact factor, articles published in a specific time period, specific news articles, online social networks or other sources that contain large volumes of short texts. The second step involves selecting the appropriate domain of research within the identified inclusion criteria. After the research have been defined, the targeted source can be identified, which can include multiple databases. Last step in the define phase involves formulating the search terms reflective of the selected research area (Wolfswinkel et al., 2011).

2.2. Search

Searching literature databases yields the initial sample of literature. This search will not only produce duplicates, but it also identify literature that requires the earlier sampling and search criteria, such as synonyms for the search terms, to be reviewed. The search phase is therefore an iterative process that is repeated until the search no longer yields additional relevant results (Wolfswinkel et al., 2011).

2.3. Select

The initial search results are refined through a filtering process. Duplicates are removed, after which each item is reviewed for fit within the research scope (Wolfswinkel et al., 2011). This phase results in the final sample of literature to be used in the analysis.

2.4. Analyse

Grounded Theory is premised on the idea that theory can be constructed through observation of the social world. A Grounded Theory is one that is inductively derived from the study of the phenomenon it represents. Concepts, categories and themes are identified and developed as the research is conducted, following an abductive reasoning model (Charmaz, 2006; Flick, 2009; Liamputtong & Eddy, 2005; Strauss & Corbin, 1990). Abductive reasoning is a form of inductive reasoning whereby the inference is based on qualitative instead of quantitative observations.

Grounded Theory was developed in the middle of the twentieth century, a time when qualitative research in sociology was under scholarly pressure. Positivist ideals of the scientific method stress objectivity, generalisability and replication of research and criticise qualitative research as being unscientific (Charmaz, 2006). In response to this criticism sociologists, Glaser and Strauss developed Grounded Theory as a method for qualitative research. Although Strauss and Glaser

originally worked together, they each developed individual interpretations of the method, resulting in an extensive polemic from which two distinctive schools emerged: the positivist Glaserian approach and the interpretive Straussian approach (Jones & Noble, 2007).

The Glaserian approach is a positivist-oriented method which strictly adheres to the principle that theory emerges from the data without interaction with previous ideas and concepts. The Glaserian school stays close to the positivist ideals of objectivity and emphasises that researchers should approach the topic with a *tabula rasa*, without specific knowledge of the subject (Jones & Noble, 2007). According to (Glaser, 2014), the purpose of qualitative research is to seek a set of interrelated hypotheses, organised around a core category, systematically generated from research. A Grounded Theory is a "conceptual theory that is abstract of time, place and people" (Jones & Noble, 2007). In the Glaserian view of Grounded Theory, rigorous procedures need to be followed in order to capture what is considered the true nature of reality and subsequently generate a theory that "fits, works, is relevant and readily modifiable" (Strauss & Corbin, 1990, p. 33).

The Straussian approach recognises that a fully objective study of social reality is impossible and follows a constructivist agenda in that "theories are not found ready made in reality, but must be constructed" (Strauss & Corbin, 1990, p. 58). This perspective allows for apriori coding to process the interview data, which entails the use of codes based on the literature review or existing theory (Furniss, Blandford, & Curzon, 2011; Liamputtong & Eddy, 2005). The Straussian approach to Grounded Theory is less restrictive and "procedures are neither mechanical nor automatic, nor do they constitute an algorithm guaranteed to give results" (Strauss and Corbin, 1990, p.59). Instead, Strauss and Corbin (1990, p.58) emphasise the importance of the "creative capabilities" and "theoretical sensitivity" of the researchers undertaking the investigation. In this school of thought, preconceived phenomena and theories are allowed to inform the analysis.

The divergence within Grounded Theory rendered the methodological landscape surrounding Grounded Theory so flexible that some researchers seem to regard it as encompassing any approach in which analysis is grounded in qualitative data. This flexibility can lead to a lack of research integrity because scholars can ignore violate the core analytic tenets of the methodology. Jones and Noble (2007) suggested guidelines to ensure research is properly grounded in reality:

- State the Grounded Theory school to which researchers subscribe
- Simultaneous collection, coding and analysis of data
- Systematic coding (open, theoretical and selective), memoing and sorting
- Generate one or more Core Categories

2.4.1. Open Coding

For this method, the social reality under consideration is the corpus of literature that resulted from the search phase. Open coding occurs by reviewing each text in the corpus and assigning codes to relevant sections of text. Each code represents a theme discussed in the discourse under consideration. Either of the coding methods described in the literature on Grounded Theory is suitable, provided that the intent of the researcher is clearly stated.

2.4.2. Theoretical Coding

In Grounded Theory, theoretical coding relies on the experience and insight of the coder (Strauss & Corbin, 1990). To reduce bias and to assist with processing large amounts of information, theoretical coding can also be undertaken by developing a semantic network of the codes assigned to the texts identified in the open coding phase.

When analysing large volumes of literature, theoretical coding can be replaced with defining a network of topics. The use of networks as an analytical tool emerged from the social sciences as a tool to visualise and analyse relationships between groups of people. The techniques of network analysis have also gained ground in many other fields of research, including information science, organisational studies, biology and practical applications such as counter-terrorism intelligence analysis (Wasserman & Faust, 1994). Networks are also commonly used to visualise information

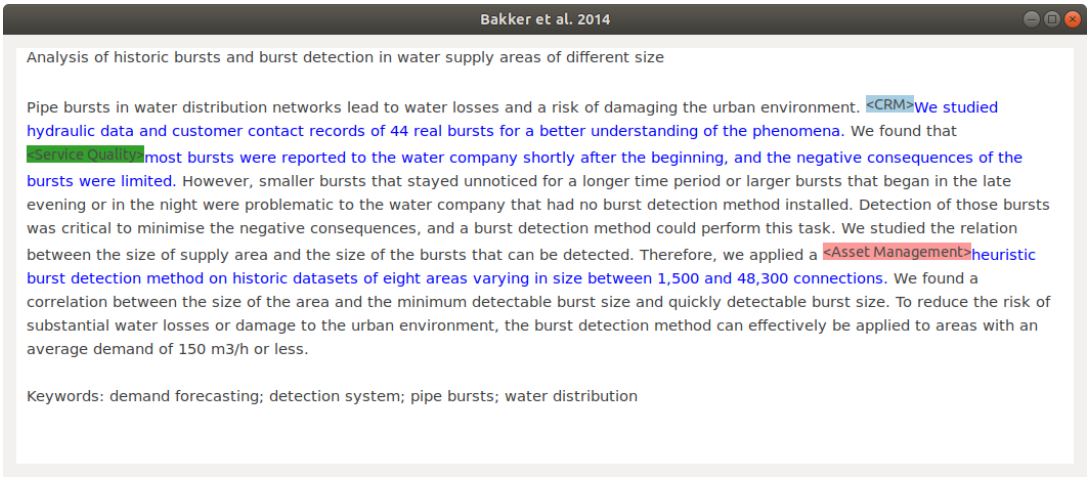

Figure 1. Open coding using the RQDA interface.

structures in the form of mind maps (Jesson, Matheson, & Lacey, 2011; Lings & Greenley, 2010) and are used to analyse discourse in both academic and non-academic environments (Moser, Groenewegen, & Huysman, 2013; Oshima, Oshima, & Matsuzawa, 2012). Using networks to analyse literature goes beyond citation networks as it addresses the content of each text instead of the relationships between the authors.

Graph theory can be used to develop a semantic network from a coded corpus of text. Graph theory is the study of mathematical structures to model relationships between objects or concepts (Wasserman & Faust, 1994). A network is a graph (G) that consists of a set of nodes (vertices), connected by a set of edges. Nodes can represent people, publications or abstract concepts. Edges represent relationships between nodes and can be either directed or undirected.

In this method for analysing text, nodes represent the themes identified in the open coding phase and edges signify that two themes were identified within the same text. The edges of a semantic network are undirected as the relationship between two topics is bidirectional; this is in contrast with citation networks where the relationship between two texts is unidirectional.

Each topic within a text is conceptually related to each of the other topics within that text. The basic graph theory principle is that each text forms a complete graph. A graph is considered complete when a node is connected to every other node in the network by an undirected edge. Each complete graph has n nodes and $n(n-1)/2$ edges. Each text in the corpus is defined by a complete network with n themes, where each theme is represented by a node and the relationships between nodes with edges. The example in figure 2 shows the complete networks of four hypothetical texts ($A - D$). Three themes are assigned to texts A and B, two topics to text C and four topic to text D. Topics number 2, 3 and 5 are shared between more than one text (Figure 2).

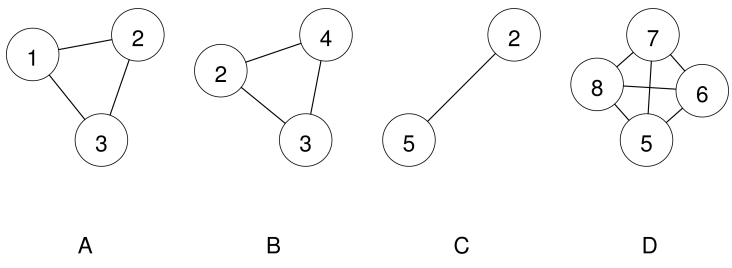

Figure 2. Example of four complete networks

The discourse network is defined by the union of the relationships between all themes discussed within the corpus. The discourse in the example is represented by graph G, which is defined by the union of the set of complete graphs ($G = A \cup B \cup C \cup D$, Figure 3). In cases

where there is more than one relationship between two codes, the weight of the edge is equal to the number of connections. In the example, the edge connecting nodes 2 and 3 has a weight of two because these topics both appear in the same two texts (A and B).

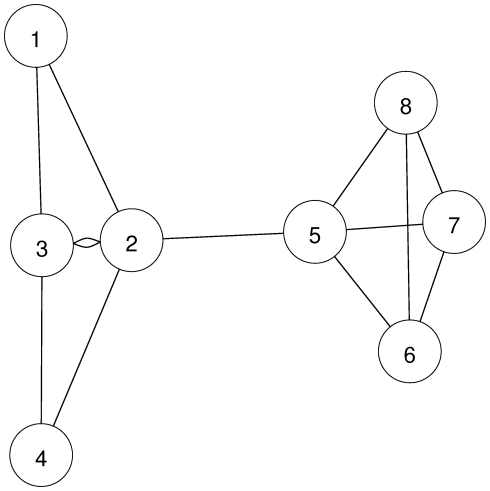

Figure 3. Union of complete networks

A graph can be described mathematically with an adjacency matrix that defines which nodes in a graph are adjacent to each other. The adjacency matrix $\mathbf{Q}$ of a finite graph G consisting of n nodes is the $n \times n$ matrix, where non-diagonal entries represent the number of edges from node i to node j. The adjacency matrix derives from the Document-Term-Matrix, which is a matrix that describes the frequency at which codes occur in the corpus. The adjacency matrix $\mathbf{Q}$ for graph G is defined by the crossproduct of the Document-Term-Matrix $\mathbf{P}$: $\mathbf{Q} = \mathbf{P} \times \mathbf{P^T}(diag(\mathbf{Q}) = 0)$. The diagonal of the adjacency matrix $\mathbf{Q}$ is set to zero to eliminate self-referencing nodes. The numbers in the adjacency matrix refer to the weight of each edge, which signifies the number of times two codes occurred in the same document. The example in formula 1 shows the Document-Term and adjacency and matrices for Figure 3.

$$\mathrm{P}= \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array} \qquad \mathrm{Q}= \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{array} \begin{array}{c} \begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{array} \\ \begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 2 & 0 & 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix} \end{array} \quad (1)$$

2.4.3. Selective coding

Using graph theory enables mathematical analysis of the relationships between topics in the corpus of text. Network analysis can be deployed to quantitatively generate the core categories of the codes defined in the open and theoretical coding phases.

The most common method is degree centrality, which is defined as the number of edges connected to each node. The more connected a node is, the higher its degree (Wasserman &

Faust, 1994). Within the context of a discourse network, the higher the degree of a node, the more central a topic is to the body of literature.

Graphs also have an underlying structure, such as communities and cliques. Complex graphs consist of one or more communities, which are cohesive sub-networks. A community is a collection of nodes that is densely connected with each other, but significantly less connected with nodes outside the community (Fortunato, 2010; Newman & Girvan, 2004; Newman, 2006). Expressed in probabilistic terms, two random nodes are significantly more likely to be connected to each other when they form part of the same community, than when they don't (Orman & Labatut, 2009).

Using the example network in Figure 3, two communities can be visually distinguished: nodes 1–4 and nodes 5–8. Texts A, B and C belong to the first community, while texts C and D belong to the second community. Text C (nodes 2 and 5) spans both communities. This solution is valid because each node has more connections to nodes within its own community, than to nodes outside its own community.

The communities in the example are easy to detect visually, but detecting community structure becomes more difficult as the network grows. Community detection is a mathematical process to cluster nodes into cohesive sub-networks (Fortunato, 2010; Orman & Labatut, 2009; Newman & Girvan, 2004; Newman, 2006). Community detection increases the parsimony of the network by identifying those groups of nodes that are most closely related to each other. Communities are analogous to latent variables that describe the discourse of the corpus. A community of discourse represents a collection of texts that discuss a similar underlying theme. Two topics belong to the same community of discourse when they are more likely to be used in the same text than two terms that do not form part of the same community.

Several algorithms for community detection have been developed, based on a range of mathematical principles. The problem of numerical validation of community structure has not yet been satisfactorily solved and no agreed method exists to assess the quality of communities (Fortunato, 2010). A comparative study between five such algorithms found that the Walktrap and Spinglass algorithms generally obtain the most reliable results (Orman & Labatut, 2009). The reliability of community selection can be assessed by calculating the modularity of the community structure.

The modularity of a community structure ranges between -$^1/_2$ and 1. A positive number indicates that the number of edges within communities exceeds the expected number. Modularity reflects the concentration of edges within communities compared with random distribution of links between all nodes, regardless of communities. A negative value indicates that the community structure is invalid. Low positive values indicate an overlap between communities and a modularity of one indicates an ideal structure without overlapping nodes. This procedure can be used to assess the mathematical validity of the community structure (Newman, 2006). The modularity of the example in Figure 3 is 0.45.

Modularity can provide insight into the validity of community structure. More important than the quantitative modularity measure is that the proposed community structure needs to be interpretable in order to extract meaning from the discourse. The final step in the analysis is thus to interpret the identified communities and describe the discourse.

2.5. Present

In the final phase of the process, the knowledge obtained from the analytical phase is structured and presented to communicate the findings. Using network analysis techniques allows for the graphical representation of the findings, providing intuitive overviews of knowledge, as shown in the above example.

3. Results

This method has been implemented using the R statistical programming language (R Core Team, 2016) and four specialised packages. The RQDA package was used for the open coding of the corpus of texts (Huang, 2014) and the iGraph package was used to visualise and analyse the network (Csardi & Nepusz, 2006). Data was loaded and transformed with the Tidyverse and reshape2 packages (Wickham & Grolemund, 2016). Tables are exported to LaTeX with the xtable package. All data and code are available through FigShare (Prevos, 2018). The method has also been published on the author's blog, The Devil is in the Data.

To transform a corpus of codes texts into a discourse network, the following approach was taken:

- All texts are manually coded using the RQDA software package.
- The Document-Term-Matrix is defined by determining the frequency that each code appears in each of the texts.
- The adjacency matrix is created by taking the cross product of the Document-Term-Matrix.
- The diagonal of the adjacency matrix is set to zero to remove self referencing nodes.
- The network is visualised and analysed using the iGraph package.

3.1. Define

The water utility sector publishes a large amount of literature on the specific problems faced by water managers. This knowledge is mainly disseminated in the form of reports commissioned by water utilities, water research organisations, government agencies, commercial research organisations and universities. These publications are mostly grey literature as they are neither formally published nor independently peer reviewed. Studying this type of literature is problematic because these documents are not indexed by bibliographic databases. Comprehensively reviewing grey literature is challenging due to the large number of repositories where this information resides (Mahood, Eerd, & Irvin, 2014). The water industry also publishes a number of peer reviewed journals, with the American Water Works Association and the International Water Association as the largest and most influential outlets. These journals are accessible through databases and are open to systematic research.

The field of research for this content analysis is marketing, which for the purpose of this analysis is defined as the discipline that views the firm from the customer's perspective (Drucker, 1954). Although water management literature is rarely directly based on marketing theory, it can nevertheless be viewed through a marketing lens. The broad definition of marketing promoted by the American Marketing Association (2013) covers the complete process of production, communication and supply of offerings. From a water supply perspective, this definition encompasses the complete value chain managed by water utilities, from extraction and supply of water to the provision of supplementary services. For example, discussions about how customers perceive the chemical and biological properties of water can be viewed as discussing product quality. Using the broad definition of marketing enables the technical discourse in water utility journals to be assessed from a marketing theory perspective.

The analyses was undertaken using all English language journals published by the International Water Association (IWA), a self-governing non-profit organisation which provides a worldwide network for water professionals and promotes standards of best practice in sustainable water management. The IWA publishes thirteen peer reviewed journals in the English language, covering a range of water utility management related aspects.

The journals were searched on occurrences of the word "marketing" in the title or abstract, plus a range of labels to denote users of water: "customer", "consumer", "user", "household", "community" and "public". This set of search terms selects only those articles that discuss matters pertaining to customers and thus are likely to fall within the domain of marketing. Abstracts were chosen in lieu of full articles because only the topic of the article needs to be identified,

rather than a full analysis of the arguments contained within.

3.2. Search

The journals were searched in May 2015 using the IWA and ScienceDirect databases, which contain over 40,000 articles. The search resulted in 5,103 entries (Table 2). Only the marketing, customer and consumer related search results were used to obtain a sample of the literature. This search resulted in a total of 590 entries, of which 23 entries contained more than one of the three keywords.

Table 2. IWA journal search results.

Keyword	Results
marketing	87
customer	207
consumer	296
user	251
household	555
community	1,282
public	2,425
TOTAL	5,103

3.3. Select

Unsuitable entries were removed for a variety of reasons. Some did not contain a journal abstract, but tables of contents, magazine articles or were written in languages other than English. Abstracts that discuss topics outside the scope of this research, such as bottled water, sanitation, agriculture or the wholesale of water, were also removed. The 590 search results were thus refined to a sample of entries. Some abstracts contained more than one keyword, resulting in 225 unique journal article abstracts available for analysis. The results were screened to assess their suitability for analysis. Unsuitable entries were removed for three reasons (Table 3):

- References to pages without journal article abstract, e.g. journal table of contents and magazine articles.
- Foreign language articles.
- Content not related to the retail of tap water, such as bottled water, sanitation, agricultural supply or wholesale of water.

Abstracts were also rejected when they only mention the customer's tap, without any further mention of the impact to customers themselves. This process resulted in abstracts for further analysis, some of which contain multiple keywords, leaving a sample of abstracts.

Table 3. Screening of search results.

Screening	n
Not English	4
No article	112
Not retail	230
Selected	244

The keyword "marketing" appeared least in the search results, with "customer" and "consumer" used equally. Interestingly, some abstracts used both customer and consumer to refer to the beneficiaries of water services. The number of times these keywords were mentioned sharply increased since they year 2000, reflecting the increased interest in customer centricity in this industry (Figure 4).

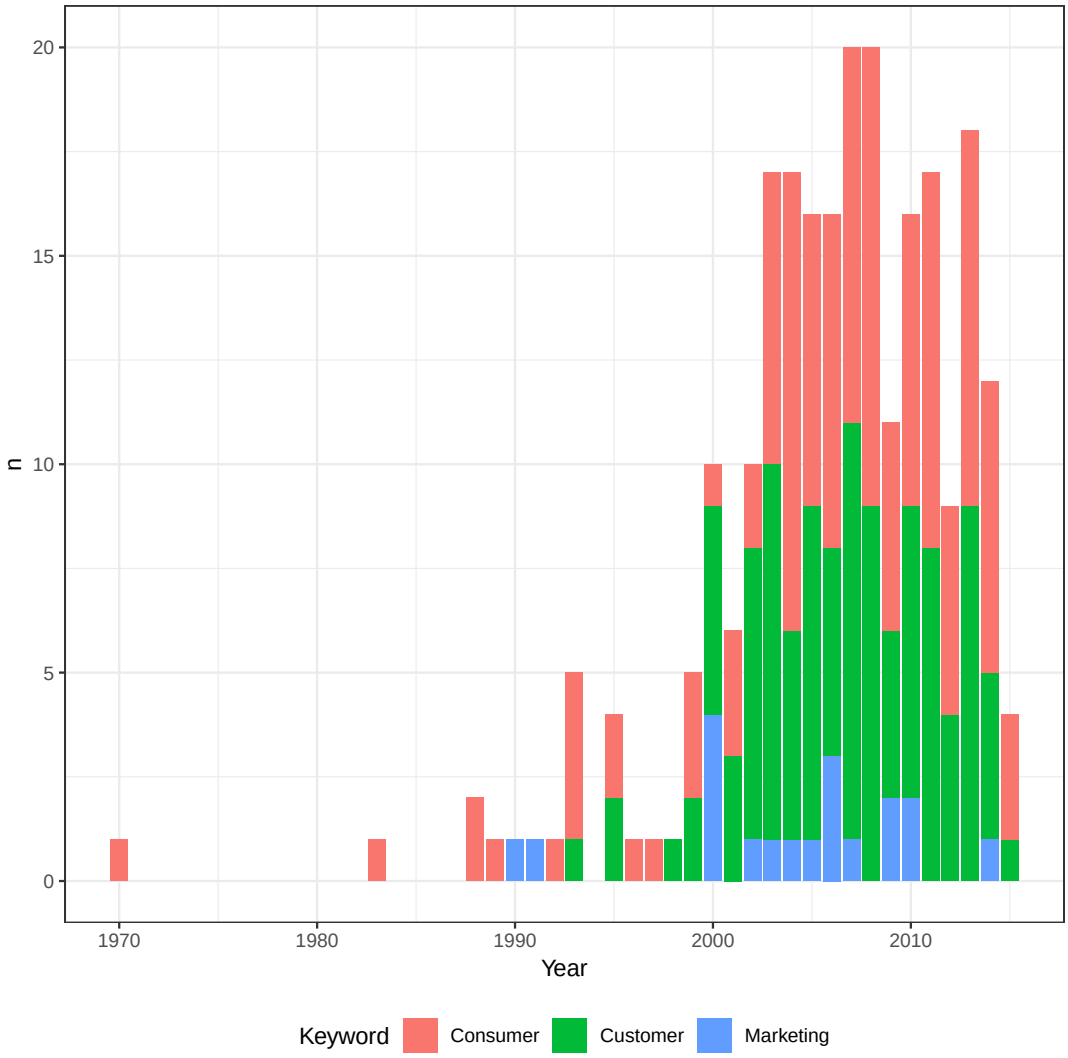

Figure 4. Longitudinal frequency of search keywords in IWA journals.

3.4. Analyse

Open coding was undertaken using an a-priori approach with the majority of codes sourced from marketing dictionaries to provide a marketing perspective on the water industry literature (Doyle, 2011; Hart & Stapleton, 2012). Each abstract was assessed for best fit with any of the lemmas from the dictionaries.

Any non-marketing related technical aspects, i.e. the design, construction, operation and maintenance of water supply systems, were coded as Asset Management. A total of 27 unique topics were identified through an iterative process with a total of 591 topics assigned to the 225 abstracts (Figure 5).

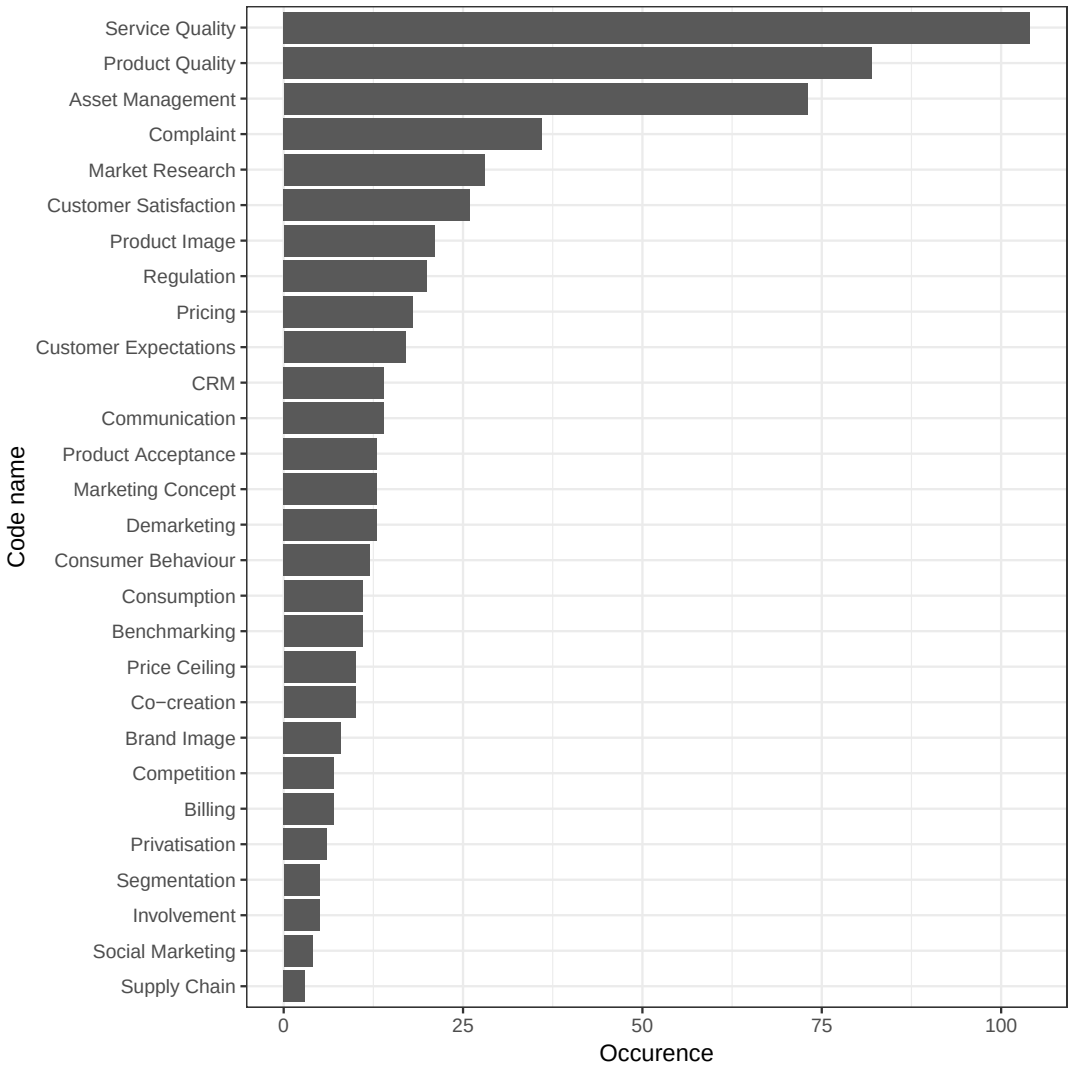

Figure 5. Frequency or abstract codes.

Each abstract contains an average of 2.60 codes, with a minimum of two codes per abstract to ensure a complete network can be formed. To visualise the network, the number of relationships between codes is represented by varying the width of the edges with their weight, the wider the line, the more frequently the two topics occurred in the same abstract (Figure 6). The size of the nodes is proportional to its degree. The larger the node, the more it is connected to other nodes. The graph is organised using the Fruchterman-Reingold method, which places nodes with the highest degree in the centre and the lowest degree on the perimeter (Wasserman and Faust, 1994).

The asset management, Service Quality and Product Quality topics have the highest degree

of all nodes in the network, forming a prominent triad in the centre of the graph. These three topics dominate the discourse in water utility management. The least discussed topics are on the periphery of the network, such as social marketing, consumer involvement and supply chain management.

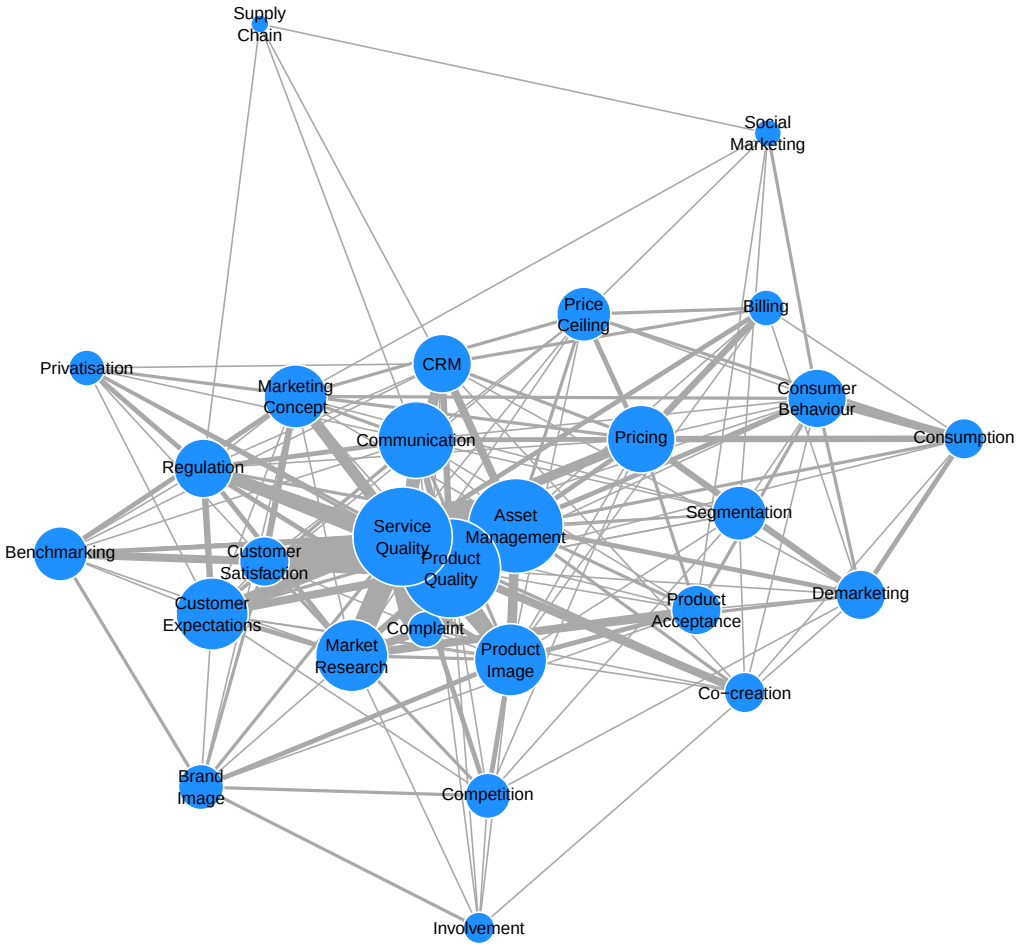

Figure 6. IWA journal abstract discourse network.

Selective coding was undertaken by identifying communities using the Walktrap and Spinglass algorithms (Newman & Girvan, 2004). Comparative analysis showed that the Spinglass method using 100 spins resulted in four communities of enquiry. This solution was chosen because it provided the best interpretable model with the highest modularity. The modularity for this solution is 0.19, which implies that communities are valid, but community membership is not certain which indicates a high level of interdependence between the communities. The four communities are, however, highly interpretable (Figure 7 and Table 4)

4. Conclusion

This report showed how a combination of qualitative Grounded Theory coding techniques, combined with quantitative network analysis can be used to analyse a large corpus of journal abstracts. The outcomes of this research were used to structure a book for water utility marketing practitioners (Prevos, 2017).

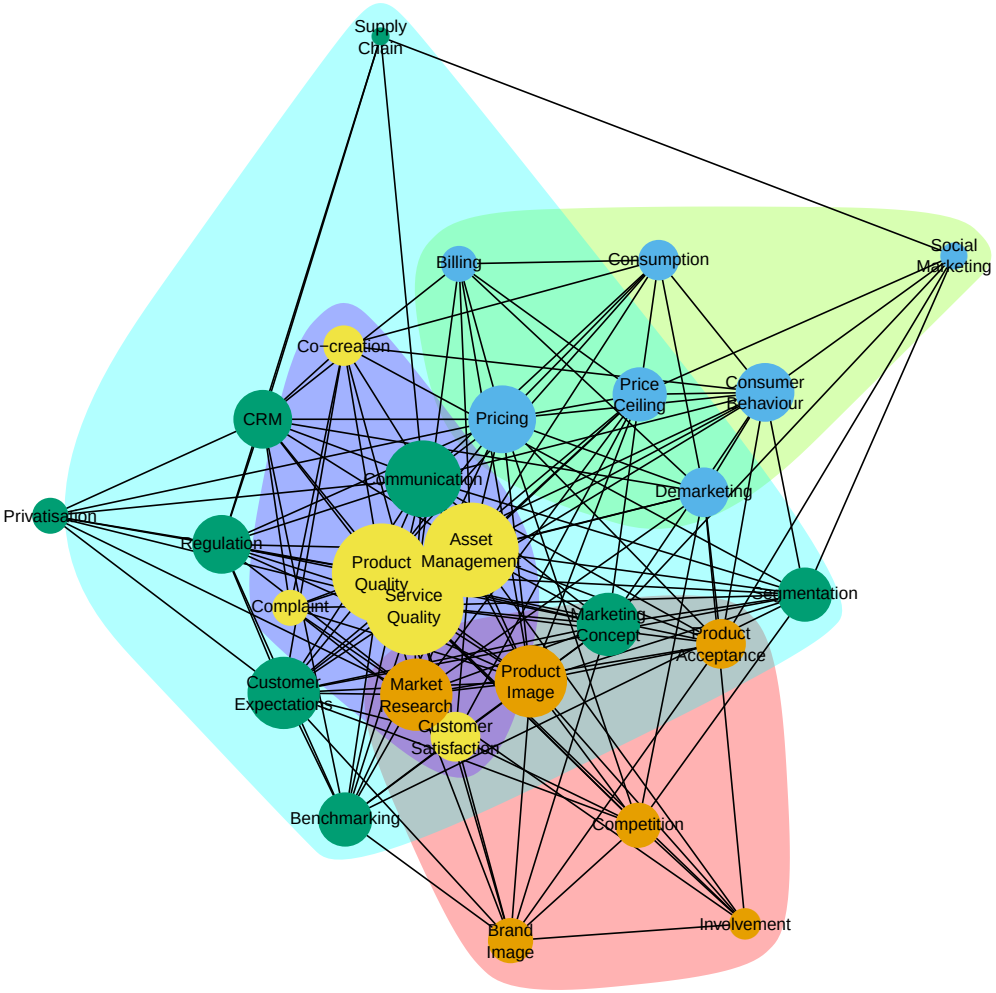

Figure 7. IWA journal abstract communities.

Table 4. IWA journals communities of discourse.

Community	Topics
A	Brand Image, Competition, Involvement, Market Research, Product Acceptance, Product Image
B	Billing, Consumer Behaviour, Consumption, Demarketing, Price Ceiling, Pricing, Social Marketing
C	Benchmarking, Communication, CRM, Customer Expectations, Marketing Concept, Privatisation, Regulation, Segmentation, Supply Chain
D	Asset Management, Co-creation, Complaint, Customer Satisfaction, Product Quality, Service Quality

References

Charmaz, K. (2006). *Constructing Grounded Theory.* Sage Publications.

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal*, *Complex Systems*, 1695. Retrieved from `http://igraph.org`

Doyle, C. (2011). *A Dictionary of Marketing.* Oxford and New York: Oxford University Press.

Drucker, P. (1954). *The Practice of Management.* New York: Perennial Library.

Flick, U. (2009). *An Introduction to Qualitative Research.* Sage Publications.

Fortunato, S. (2010, February). Community detection in graphs. *Physics Reports*, *486*(3-5), 75–174.

Furniss, D., Blandford, A., & Curzon, P. (2011). Confessions from a grounded theory PhD: Experiences and lessons learnt. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 113–122). ACM.

Glaser, B. G. (2014). Applying grounded theory. *The Grounded Theory Review*, *13*(1), 46–50.

Hart, N., & Stapleton, J. (2012). *The CIM Marketing Dictionary.* Hoboken: Taylor and Francis.

Huang, R. (2014). *RQDA: R-based Qualitative Data Analysis.* Retrieved from `http://rqda.r-forge.r-project.org/` (R package version 0.2-3)

Jesson, J., Matheson, L., & Lacey, F. M. (2011). *Doing Your Literature Review: Traditional and Systematic Techniques.* Los Angeles and London: SAGE.

Jones, R., & Noble, G. (2007). Grounded theory and management research: A lack of integrity? *Qualitative Research in Organizations and Management: An International Journal*, *2*(2), 84–103.

Liamputtong, P., & Eddy, D. (2005). *Qualitative research methods* (2nd ed.). Oxford University Press.

Lings, I. N., & Greenley, G. E. (2010). Internal market orientation and market-oriented behaviours. *Journal of Service Management*, *21*(3), 321–343.

Mahood, Q., Eerd, D. V., & Irvin, E. (2014, September). Searching for grey literature for systematic reviews: Challenges and benefits. *Research Synthesis Methods*, *5*(3), 221–234.

Moser, C., Groenewegen, P., & Huysman, M. (2013). Extending Social Network Analysis with Discourse Analysis: Combining Relational with Interpretive Data. In T. Özyer, J. Rokne, G. Wagner, & A. H. Reuser (Eds.), *The Influence of Technology on Social Network Analysis and Mining* (Vol. 6, pp. 547–561). Vienna: Springer Vienna.

Newman, M. E. J. (2006, June). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, *103*(23), 8577–8582.

Newman, M. E. J., & Girvan, M. (2004, February). Finding and evaluating community structure in networks. *Physical Review E*, *69*(2), 026113–1 – 026113–15. Retrieved 2015-05-31, from `http://link.aps.org/doi/10.1103/PhysRevE.69.026113`

Orman, G. K., & Labatut, V. (2009). A Comparison of Community Detection Algorithms on Artificial Networks. In D. Hutchison et al. (Eds.), *Discovery Science* (Vol. 5808, pp. 242–256). Berlin, Heidelberg: Springer Berlin Heidelberg.

Oshima, J., Oshima, R., & Matsuzawa, Y. (2012, June). Knowledge Building Discourse Explorer: A social network analysis application for knowledge building discourse. *Educational*

Technology Research and Development, *60*(5), 903–921.

Peng, R. (2011, December). Reproducible research in computational science. *Science*, *334*(6060), 1226–1227.

Prevos, P. (2017). *Customer Experience Management for Water Utilities: Marketing urban water supply*. London: International Water Association. (OCLC: 975033056)

Prevos, P. (2018). *Water utility marketing mix data and methods*. Figshare. Retrieved 2018-11-01, from `https://figshare.com/articles/IWA_journal_abstracts_database/7273061` (type: dataset)

R Core Team. (2016). *R: A Language and Environment for Statistical Computing*. Vienna, Austria. Retrieved from `www.R-project.org`

Schulte, E., Davison, D., Dye, T., & Dominik, C. (2012). A Multi-Language Computing Environment for Literate Programming and Reproducible Research. *Journal of Statistical Software*, *46*(3).

Strauss, A., & Corbin, J. (1990). *Basics of Qualitative Research. Grounded Theory Procedures and Techniques*. Newbury Park: Sage.

Wasserman, S., & Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press.

Wickham, H., & Grolemund, G. (2016). *R for data science: import, tidy, transform, visualize, and model data* (First edition ed.). Sebastopol, CA: O'Reilly. (OCLC: ocn968213225)

Wolfswinkel, J. F., Furtmueller, E., & Wilderom, C. P. M. (2011, November). Using grounded theory as a method for rigorously reviewing literature. *European Journal of Information Systems*, *22*(1), 45–55.