Quality Engineering

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/lqen20

Mining quality determinants of product-service systems from user-generated contents

Luca Mastrogiacomo, Federico Barravecchia, Fiorenzo Franceschini & Frederic Marimon

QUALITY ENGINEERING
https://doi.org/10.1080/08982112.2021.1877305

Mining quality determinants of product-service systems from user-generated contents

Luca Mastrogiacomo[a], Federico Barravecchia[a], Fiorenzo Franceschini[a], and Frederic Marimon[b]

[a]DIGEP (Department of Management and Production Engineering), Politecnico di Torino, Torino, Italy; [b]Faculty of Economic and Social Sciences, Department of Economy and Business Organization, Universitat Internacional de Catalunya, Barcelona, Spain

ABSTRACT

The concurrent availability of an ever-increasing volume of User-Generated Content (UGC) and text mining techniques has enabled methods to conduct both qualitative and quantitative research that was unimaginable a few years ago. These investigations have proved to generate interesting insights with potential implications in manufacturing and engineering design. This paper introduces a novel practical methodology to address the following research issue: *how to infer the quality determinants of a Product-Service System (PSS) from UGCs?* Analyzing a sample of UGCs by means of a topic modeling method, the idea of this study is to produce and examine the set of determinants relevant to the perception of the quality of a specific PSS. The application of the proposed approach is exemplified on a specific use-oriented PSS: car-sharing.

KEYWORDS

user-generated contents; quality determinants; product-service systems; topic modeling; car-sharing

Introduction

Product-Service Systems (PSS) are integrated bundles of products and services which aim at generating value and creating customer utility (Boehm and Thomas 2013). Business models combining the offering of products and services are becoming considerably more widespread (Mastrogiacomo, Barravecchia, and Franceschini 2019, 2020).

Several factors explain the recent success of the PSS-based business model on a global scale and in a wide range of economic activities: maturity of information technology, general user acceptance of PSS offers format, affordable and extended access to internet, wide range of suppliers available, etc. (Dyllick and Rost 2017; Neri et al. 2018).

Although PSS-based business models are widespread and of considerable interest to academics and professionals, only a handful of studies have focused on the quality of PSS. (Waltemode, Mannweiler, and Aurich 2012; Mert, Waltemode, and Aurich 2014).

Assessing the quality of PSS is a complex problem that goes beyond the conventional quality paradigms (Zonnenshain and Kenett 2020). This complexity is essentially due to their heterogeneous composition of various elements (products, services and infrastructures) that concurrently contribute to the overall quality of the systems (Beuren, Gomes Ferreira, and Cauchick Miguel 2013; Mert, Waltemode, and Aurich 2014; Marimon et al. 2019; Barravecchia, Mastrogiacomo, and Franceschini 2020; Robinson, Giles, and Rajapakshage 2020).

There are certainly general frameworks for assessing separately the quality of both services and products, but these structures are not adequately considering PSS (Mukherjee 2019). To date, research has not determined a complete model for measuring the quality of PSS yet, and each specific case requires the identification of a relevant framework for understanding and evaluating the quality perceived by customers. On one side traditional approaches require the involvement of experts and the use of questionnaires to assess quality determinants (DeVellis 2016), on the other they are sometimes criticized as inefficient, also due to their static nature, being difficult to adapt to the changing of customer needs over time (LaValle et al. 2011; Franceschini, Galetto, and Maisano 2019).

This paper attempts to overcome these limitations by proposing a novel data-driven approach for identifying the quality determinants of a PSS using unstructured User-Generated Content (UGC), i.e., information directly generated by customers (e.g.,

CONTACT Fiorenzo Franceschini fiorenzo.franceschini@polito.it DIGEP (Department of Management and Production Engineering), Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129, Torino, Italy

open-ended responses or on-line reviews). Although presented in an unstructured form, these data can constitute a low-cost solution, constantly and spontaneously updated, which – if correctly used – can generate new insights for the engineering design, management and control of PSS quality. In this sense, UGCs can offer a promising solution to the identification of PSS latent quality determinants, overcoming the limitations arising from the use of the traditional frameworks (Allen, Sui, and Akbari 2018).

Analysis of UGC is an increasingly popular area of research in several fields. While a few preliminary studies attempted to identify and analyze quality determinants using UGC, none of these concerned PSS, and a structured methodology has not been formalized yet.

To bridge these gaps, this paper aims at addressing the following research question: *How can quality determinants of a Product-Service System be determined from the analysis of UGCs?*

In order to highlight the benefits achieved by the proposed approach, a practical application of the methodology is presented. As PSS reference we considered the case of Car-sharing, for which an important volume of UGCs is available on-line. The study of a large dataset of on-line reviews is carried out by means of the Structural Topic Model (STM), i.e., a variant of Latent Dirichlet Allocation (LDA) technique which is able to identify underlying topics in a collection of documents also incorporating document metadata (i.e., provider, score, nationality, etc.). Document metadata can enhance inference and qualitative interpretability of the results. In a broad sense, the rationale behind this kind of analysis is that *if a topic is discussed, then it may be critical to the quality of the considered PSS*.

This study provides new methodological and empirical insights into the domain of quality management by providing two main contributions. Firstly, it defines a structured methodology for the definition of the determinants of the quality of a PSS through a data mining method. This approach opens wide research perspectives, bypassing the limitations of traditional approaches (questionnaires or customer interviews). Secondly, it provides a general framework for analyzing the quality of a PSS car-sharing. Despite the growing popularity and relative diffusion, studies concerning the quality of this specific family of PSS are few and limited to traditional approaches.

The rest of the paper is organized as follows. Section background and literature review presents the research background from different perspectives: the study and analysis of UGCs (Section user-generated content) and the use of text mining techniques in the field of product/service quality (Section quality determinants and topic models). The proposed methodology of analysis is described in Section methodology: from unstructured UGCs to quality determinants, while Section case study: car-sharing PSS presents an application to the car-sharing PSS family, with particular reference to how the experimental dataset was obtained and processed. Section result analysis: insights on car-sharing quality determinants highlights the results of the analysis. The main conceptual and technical implications of the study are discussed in Section implications. Finally, the concluding section explores the limitations and future directions of this research.

Background and literature review

User-generated content

Information available on the web is exponentially growing (Kambatla et al. 2014). According to the International Data Corporation (IDC), by 2025 there will be 175 Zettabytes of data in the world, with a significant percentage of business-relevant information originating from unstructured forms, mainly text (Reinsel, Gantz, and Rydning 2018). Some of the content of this huge pool of information is generated by businesses and organizations, while others are generated directly by users. User-Generated Contents (UGC) are defined as "creative works that are published on publicly accessible websites and are created without a direct link to monetary profit or commercial interest" (OECD 2001). Reviews, bulletin boards, forum discussions and social networks interactions are different forms of UGCs (Krumm, Davies, and Narayanaswami 2008). There are several on-line platforms, or review aggregators, that collect and share UGCs (such as: Yelp!, Amazon, Facebook, TripAdvisor, etc.). These platforms encourage users to contribute actively and spontaneously to the creation of content. Also, many product or service providers encourage their customers to share experiences and feedbacks about their offerings in exchange for promotions, gifts or incentives (Guo, Barnes, and Jia 2017).

Many authors have pointed out how UGCs affects business demand, performance and sales (Tirunillai and Tellis 2012). Several articles highlighted the influence of UGCs in the purchasing and decision-making process of customers (Park, Lee, and Han 2007). In particular, it was highlighted that purchasing

Table 1. Possible approaches for determining the quality determinants of a service/PSS.

Data Source	Data collection	Approach	Indicative studies
Customer perception	- Structured questionnaire	1. Survey Instruments (ServQual)	(Parasuraman, Zeithaml, and Berry 1988)
		1. Post-service structured feedback	(Dholakia, Singh, and Westbrook 2010)
		1. Consumer surveys	(Chang and Yeh 2002; Marimon et al. 2019)
Consumer reviews	- Unstructured text	1. Analysis of the word-frequencies	(Barreda and Bilgihan 2013)
		1. Sentiment Analysis	(Pang and Lee 2008)
		1. Topic Modeling (LDA)	(Allen and Xiong 2012; Allen, Xiong, and Afful-Dadzie 2016; Guo, Barnes, and Jia 2017; James, Calderon, and Cook 2017)

intentions are influenced by: (i) the amount of reviews available; (ii) the quality of the reviews; (iii) perceived informativeness and perceived persuasiveness of the reviews; (iv) credibility of the review source and aggregation platform (Park, Lee, and Han 2007; Lee, Park, and Han 2008; Zhang et al. 2014). It has also been found that offline purchasing decisions are often made on the basis of information found on-line (Lee, Park, and Han 2008).

Accordingly, UGCs can be a valuable, low-cost source of information for understanding the needs, desires, expectations and perceptions of products and services.

Quality determinants and topic models

Several approaches to identify the quality determinants of products and service have been proposed in the literature (Korfiatis et al. 2019). Table 1 summarizes some different types of approach, together with data sources, methods for data collection and bibliographical references.

Traditionally, the definition of determinants influencing quality perception has been supported by quantitative methods, mainly based on data obtained through questionnaires and interviews (DeVellis 2016). Following DeVellis (2016), four basic steps characterize these approaches: (i) specification of the domain of the construct; (ii) generation of the initial sample of items from literature research, interviews, surveys; (iii) refinement of the pool of items of the scale (data collection and Preliminary Factor Analysis); (iv) optimization of the measurement scale (data collection and confirmatory factor analysis). Although it is well established, this method is highly resource-intensive in terms of time and people involved. The quality of the outcomes resulting from the application of these methodologies are highly related to the willingness of respondents to participate and to the complexity of the questionnaire (Groves 2006). Furthermore, it presents several limitations, including: (i) the limited size of the sample of customers, (ii) the definition of the initial pool of items to measure quality perception is biased by the subjectivities of the domain experts, some determinants may be missed or underestimated, (iii) responses can include potential errors that might not be detectable (DeVellis 2016). With reference to Table 1, approaches 1 to 3 fall into this category.

Besides traditional approaches, in recent years, some researchers tried to leverage information directly obtained from the web to gain deeper insights into perceived quality and customer satisfaction. An alternative approach to determine the quality determinants of an object (product, service or PSS) may be the analysis of UGCs and, more specifically, of on-line reviews which can offer a low-cost, unbiased and trustful source of information for understanding customer opinions, expectations, and requirements (Sweeney, Soutar, and Mazzarol 2008).

The identification of quality determinants is based on the in-depth analysis of such data, leveraging text mining tools able of obtaining information through text documents written in a natural language (Aggarwal and Zhai 2012). The fundamental logic of these methods is that if a topic is discussed (within the UGC) than it is critical to the definition of the quality of the PSS under investigation. Most of the previous studies based on UCG focused their analysis on keywords frequency and sentiment analysis (Liu 2012; Bi et al. 2019). Few researches applied Unstructured Topic Modeling, in particular Latent Dirichelet Allocation (LDA), to identify quality determinants through the analysis of UGC (Tirunillai and Tellis 2014; Guo, Barnes, and Jia 2017). One of the main limitations of these studies is that they do not consider some important metadata associated to the reviews (such as: ratings, geographical location, category of user, date, service provider name, …). These metadata may provide additional and valuable information to discriminate the specific role of each quality determinant in the general reference framework (Roberts, Stewart, and Tingley 2019).

Despite the full potential of UGC analysis for investigating customer opinions concerning a wide range of products, services and PSS, most of the

Table 2. Comparison of the proposed approaches with existing ones focused on the analysis of UGC.

Existing approaches	Main Limitations	Related advantages of the proposed methodology
Analysis of the word-frequencies	- Only word frequencies are considered relevant for the classification of the documents	- Word co-occurrence patterns and word frequencies are used to capture latent topics running through a collection of unstructured textual documents
Sentiment Analysis	- Dictionary-based approach	- Do not require specific lexicons or linguistic databases
	- Customer opinion is mined by analyzing positive or negative words contained in a textual document - Difficulties in recognition of things like sarcasm and irony, negations, jokes, and exaggerations	- Customer attitude is directly derived from customer rating
Topic Modeling (LDA)	- Difficulties in the estimation of input parameters of the model - Results highly sensitive to input parameters of the model	- STM algorithm does not require the setting of input parameters of the model
	- Metadata associated with the analyzed UGCs are not taken into account	- STM algorithm exploits metadata for the identification of quality determinants
	- Algorithm application	- Structured methodology indicating all the necessary steps from data extraction to results analysis

application of UGCs analysis focused their attention on gaining insights on tourism and hospitality services.

Thus far, some progress has been made in UGC analysis to identify the quality dimensions of an object, but there are still many gaps to be filled. In this view, Table 2 compares the proposed methodology with the other existing approaches highlighting advantages and elements of novelty.

Methodology: From unstructured UGCs to quality determinants

Usually, most score distributions of on-line reviews follow a J-shaped or a U-shaped distribution, where top and/or low values present the larger concentration of reviews, while middle values present lower concentrations (Matakos and Tsaparas 2016; Ho, Wu, and Tan 2017). Based on this observation, we can conclude that issues that somehow influence positively or negatively the difference between expectations and perceptions are discussed in UGCs. In a broad sense, if consumers are talking about a specific aspect, then that aspect is relevant in determining their degree of satisfaction. Knowing and understanding these determinants is the first step in improving the quality of an object, its engineering design and consequently customer satisfaction.

In this study, a topic modeling approach is used to identify quality determinants. The methodology is organized into seven steps: (i) dataset extraction; (ii) pre-processing; (iii) identification of the optimal number of topics; (iv) topic modeling (v) labeling; (vi) validation of results; (vii) Result analysis. Figure 1 relates each phase with the stages of analysis of the Text Mining Analysis Roadmap (TMAR) (Zaki and McColl-Kennedy 2020). The following subsections detail the aforementioned phases, while Section case study: car-sharing PSS proposes an application of the proposed methodology.

Dataset extraction

The first step of the proposed methodology consists of collecting UGCs from social media and other review aggregators, i.e., the so-called web-scraping activity. Despite many platforms (e.g., Google, Facebook) are attempting to limit the download of large amounts of UGCs, a variety of software applications are available for web scraping. Moreover, text mining programs often include libraries for web scraping (Zaki and McColl-Kennedy 2020). Besides the textual content of the reviews, such tools often allow the collection of relevant metadata. Information such as title, author, date, score, nationality can be valuable for the next steps of analysis. It is important to emphasize that the variety of data sources can influence the quality of the final results. Different sources ensure a greater diversification of the type of user and of the focus of the reviews.

Pre-processing

The downloaded text content needs to be pre-processed and unified in order to improve the efficiency of the topic modeling algorithm (Meyer, Hornik, and Feinerer 2008; Guo, Barnes, and Jia 2017). According to previous approaches, common pre-processing operations are (Meyer, Hornik, and Feinerer 2008; Tirunillai and Tellis 2014; Guo, Barnes, and Jia 2017):

- conversion of the text to lowercase in order to eliminate ambiguity with uppercase words;
- removal of the punctuation and numbers, since adding little topical content;

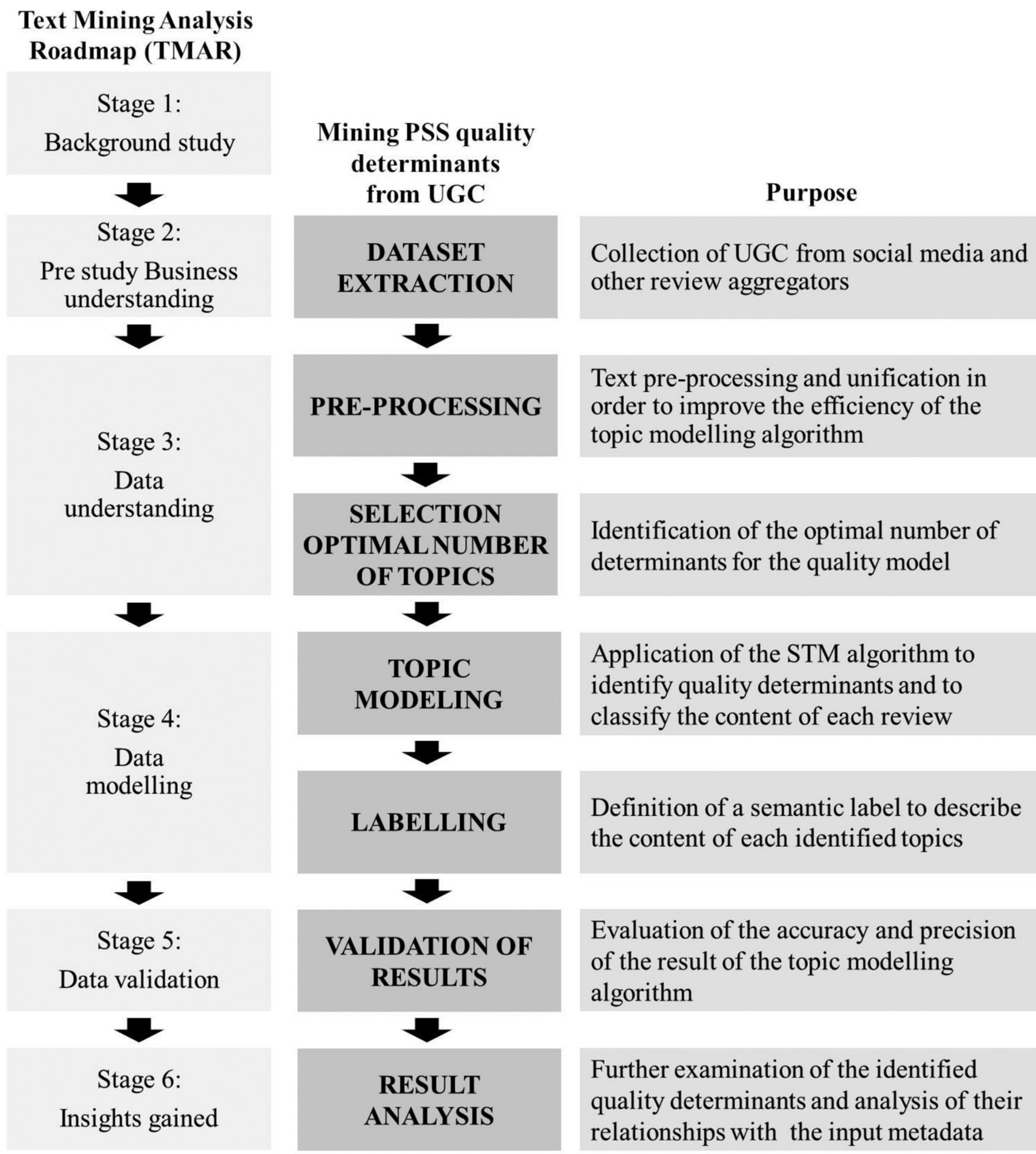

Figure 1. Activity flow of the proposed methodology and comparison with the Text Mining Analysis Roadmap (TMAR) (Zaki and McColl-Kennedy 2020).

- removal of the stop words (e.g. "the", "and", "when", "is", "at", "which", "on", etc.);
- removal of the words shorter than 2 or longer than 15 characters;
- exclusion of the words with an extremely low frequency, since their inclusion would confound results or would not be representative of any specific topics.
- application of a stemming process to reduce inflected or derived words. Stemming removes the commoner morphological and inflectional endings from words in English (Jivani 2011). For example, the words “likes”, “liked”, “likely” and “liking” are reduced to the stem “like”;
- removal of the words generally not related to topical content (such as: "another", "mean", "problem", "review", "made", "did", "done", etc.);
- replacement of the most common n-grams, i.e. set of co-occurring words, with a single term (Durrani et al. 2015). For example, the n-gram “customer service” need to be replaced by the term “customerservice”.

Selection of the optimal number of topics

An essential parameter required by topic modeling algorithms is T, i.e., the number of topics able to describe the text corpus. The literature presents some possible alternatives to define T (Wallach, Mimno, and McCallum 2009). In most applications, the topic modeling algorithm is applied iteratively to identify the optimal value of T, evaluating the performance of the topic model at each iteration. The number of topics optimizing the performance of the topic model is then selected. With the purpose of identifying the optimal number of topics, we suggest the use the held-out likelihood. This metric quantifies the likelihood of the model on a subset of the UGCs (usually 10%) that were not used for the estimation of the model (Roberts et al. 2014). Held-out likelihood can be seen as a measure of how the developed topic model is able to explain the overall variability in the text corpus (Scott and Baldridge 2013).

Other alternative and complementary criteria can be also applied (Roberts et al. 2014), including: (i) *semantic coherence*; (ii) *exclusivity* and (iii) *residuals.*

Topic modeling

Topic modeling is an unsupervised class of machine-learning algorithms that can detect latent topics running through a collection of unstructured documents (Müller et al. 2016). Given a big set of documents, topic modeling algorithms deal with the problems of: (i) identifying a set of topics that describe a text corpus (i.e., a collection of text document from a variety of sources); (ii) associating a set of keywords to each topic and (iii) defining a specific mixture of these topics for each document (Blei, Ng, and Jordan 2003).

In detail, for the purpose of analyzing UGCs, a probabilistic topic modeling method, named Structural Topic Model (STM) can be used. It can be considered as an extension of established probabilistic topic models, such as Latent Dirichlet Allocation (LDA) (Blei, Ng, and Jordan 2003) or Correlated Topic Models (CTP) (Blei and Lafferty 2007). A significant advantage of STM over other methods is that it allows the connection of arbitrary information, in the form of covariates (such as customer score, date and place of publication of the review, provider, etc.), with the degree of association of a document with a topic (topical prevalence) as well as the degree of association of a word with a topic (topical content). Recent papers and overviews of the STM algorithm exist for readers unfamiliar with its background and use (Roberts et al. 2014, Roberts, Stewart, and Tingley 2019).

Labeling

For each topic, the STM approach identifies the most relevant keywords. However, to generate a relevant semantic label, the method still requires some human input (Blei 2012). The list of the top keywords that describe a topic to be used for the labeling phase can be obtained using different criteria, for instance (Roberts, Stewart, and Tingley 2019):

a. **Highest probability**: words with the highest probability within each topic;
b. **FREX**: words that are both frequent and exclusive.

Other approaches, such as the Score (i.e., words with the highest score according to LDA ranking) and Lift (i.e., a score calculated by dividing the topic-word distribution by the empirical word count probability distribution) can provide alternative and supplementary strategies to identify most representative keywords for each topics (Roberts, Stewart, and Tingley 2019).

The combined use of these different approached facilitates topic labeling by offering a more defined characterization of the topic content. In addition to the list of the top keywords, the most relevant reviews of each topic (i.e., the reviews with the highest weight related to the topics) can also be considered for the topic labeling.

Data verification

Due to the absence of a reference, the result verification of text mining algorithms has been the subject of several research studies (Zaki and McColl-Kennedy 2020). For the purpose of the proposed methodology, obtained results can be verified by comparing the assigned topics of a randomly selected sample of reviews with a manual topic assignment. For each of the compared reviews, manual evaluators are requested to agree in the association of one or more of the topics identified by the topic modeling algorithm (sets of keywords and label). The so-defined manual topic assignment needs to be considered as reference and compared to that obtained by topic modeling algorithm. For each review and topic, the four cases reported in Table 3 can occur.

According to Costa et al. (2007), four verification indicators can be calculated: accuracy, recall, precision and F-measure (see Table 4). *Accuracy* is the most intuitive performance measure and it is equal to the ratio of correctly predicted observation to the total

Table 3. Possible outcomes of the data verification test.

Validation result	Error type	Probability	Description	Identification of a topic: Topic modeling algorithm	Identification of a topic: Manual evaluator
True Positive (*tp*)	Correct inference	(1- β)	Agreement between authors and algorithm in the assignment of a review to a topic	Yes	Yes
True Negative (*tn*)	Correct inference	(1- α)	Agreement between authors and algorithm not to assign a review to a topic	No	No
False Positive (*fp*)	Type I error	α	Misalignment between the assignment of the review to a topic by STM and the non-assignment by the manual evaluator	Yes	No
False Negative (*fn*)	Type II error	β	Misalignment between the assignment of the review to a topic by the manual evaluator and the non-assignment by STM	No	Yes

Table 4. Verification indicators (Costa et al. 2007).

Name	Definition	Formula	Codomain
Accuracy	It evaluates the effectiveness of the algorithm by its percentage of correct prediction. It is a ratio of correctly predicted observation to the total observation.	$A = \frac{(tp + tn)}{(tp + tn + fp + fb}$	[0;1]
Recall	It is the ration between the correctly predicted positive observation and all observation in actual class. Also known as sensitivity or true positive rate.	$R = \frac{tp}{(tp + fn)}$	[0;1]
Precision	It is a measure which estimates the probability that a positive prediction is correct. It is the ratio between the correctly predicted positive observations and the total predicted positive observation.	$P = \frac{tp}{(tp + fp)}$	[0;1]
F measure	It is the weighted average of Precision and Recall indicators. This score takes both false positives and false negatives into account.	$F = 2 \times \frac{P \times R}{P + R}$	[0;1]

observations. It measures how often the algorithm produce a correct topic assignment. Accuracy assumes equal costs for both kinds of errors. Further metrics should be calculated in order to evaluate more accurately the performance of the applied method. To fully evaluate the effectiveness of a topic modeling algorithm, two indicators should also be considered: Recall and Precision. *Recall*, also known as sensitivity or true positive rate, can be defined as the ratio of the total number of correctly predicted observation (true positive) with the sum of true positive and false negative observations. Recall metric answers to the questions: "If a topic is present in a review, how often is the algorithm able to detect it??". *Precision*, also known as positive predictive value, is equal to the ratio between the total number of correctly classified positive examples by the total number of predicted positive prediction (true positive + false positive). This metric answers to the question: "What proportion of positive topic assignments was actually correct?". *F-measure*, also known as F_1 score, is defined as the weighted harmonic mean of the precision and recall of the test. This metric, measure how precise the topic model is, by taking into account how many topics has been correctly identified, as well as how robust it is, it does not miss a significant number of topic identification.

Nassirtoussi et al. (2014) reported that accuracy values above 55% could be accepted as "report-worthy". According to Zaki and McColl-Kennedy (2020), in most cases, accuracy is between 50% and 80%.

Results analysis

Topic modeling algorithms primarily produce two outputs: the latent quality determinants of the PSS under analysis (set of keywords and labels) and the proportion of each topic associated with each review. In addition, the STM application allows a variety of analyses, including:

i. analysis of the prevalence of the topics within the analyzed sample of reviews,
ii. analysis of the relationships between metadata and topical prevalence;
iii. analysis of the relationships between customer score and topical prevalence;
iv. analysis of the correlations between topics;
v. analysis of the time series of topic proportion.

The purpose of these investigations is to obtain a clearer picture of quality determinants useful to understand how to use these insights to develop

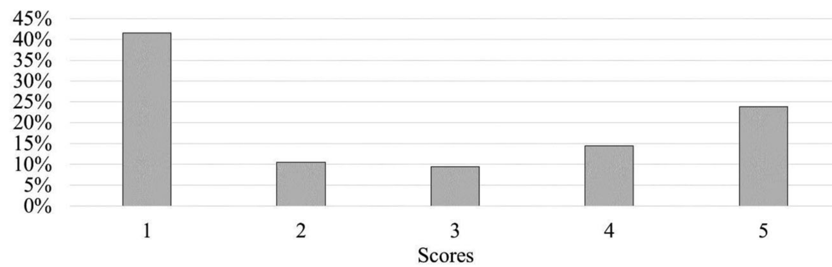

Figure 2. Distribution of the review scores (five level ordinal scale).

quality improvement strategies or to re-address engineering design.

Case study: Car-sharing PSS

Several previous studies have attributed car sharing as a PSS, a use-oriented PSS in particular (Williams 2007; Liu et al. 2014). Car-sharing is a form of shared mobility that has gained increasing popularity in recent years (Shaheen and Cohen 2007). Given its promise to reduce traffic congestion, parking demands and pollution, this mode of shared transportation has spread especially in urban contexts, so much so that several new competitors are recently entering this market designing and proposing new service solutions (Shaheen and Cohen 2013).

Despite the emerging importance of this type of mobility, no structured analysis has been performed to understand the most critical aspects of car-sharing quality determinants (Illgen and Höck 2019). Only a couple of interesting papers on the subject which relies on the analysis of UGCs. However, such studies are far from being complete and structured: the study by Guglielmetti Mugion et al. (2019) discusses the antecedents of the use of car-sharing through an analysis of UGCs limited to service users in the city of Rome; on the other hand, the work of Jeong et al. (2019), only investigates on-line car-sharing reviews on the google play website, which is designed to collect users' opinions mainly related to the mobile application of the service.

The focus of this case study is the evaluation of the quality of one-way car-sharing (i.e., members are allowed to begin and end their trip at different locations) and roundtrip car-sharing (i.e., members are required to begin and end their trip at the same location) using information available on-line in the form of user reviews.

Each phase of the proposed methodology, applied to the case of car sharing, is detailed in the following subsections

Dataset extraction and pre-processing

Analyzed data are reviews and relevant metadata (car-sharing providers, nationality, score, date, source) retrieved in December 2019 from different review aggregators: Yelp, Google, Trustpilot, Facebook and Playstore. Reviews were published from January 2010 to December 2019. Only English-language reviews were selected, with a total of almost 17,000 reviews from 22 car-sharing providers (Car2go, DriveNow, Maven, Zipcar, Goget, etc.), distributed in 3 countries (US, Canada and UK). Each provider was related to the type of car sharing (station-based or free-floating). The average length of the considered reviews is about 500 characters, characterized by a U-shaped distribution of the review scores (see Figure 2).

Following the procedure reported in Section pre-processing, the text content of the 17,000 reviews was preprocessed and normalized. Pre-processing activities resulted in a reduction of 76% of the text corpus.

Selection of the optimal number of topics and topic modeling

The analysis herein presented was carried out using the Structured Topic Modeling (STM) package of R software. The information concerning review scores, types of car-sharing (station-based or free-floating) and countries coupled with the text content of each review were used to define the topical prevalence (i.e., the proportion of document devoted to a given topic) and the topical content (i.e., the rate of word use within a given topic) in the STM model.

In order to identify the optimal number of quality determinants to describe the PSS under analysis, the STM algorithm was run by varying the number of topics and evaluating held-out likelihood indicator presented in Section selection of the optimal number of topics. The graphs in Figure 3 show the values of the held-out likelihood as a function of T (from 5 to 100). From the graph, it can be observed that starting from the value of T equal to 20 there is an almost

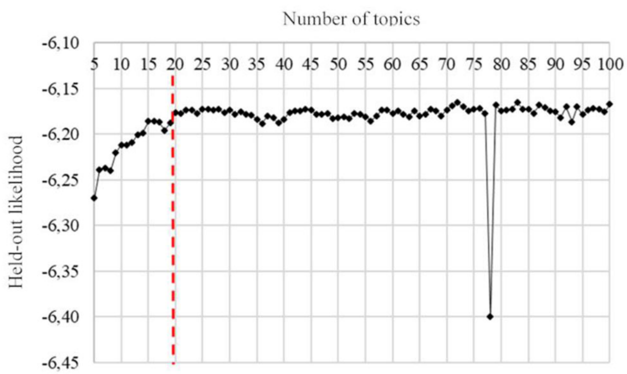

Figure 3. Held-out likelihood by number of topics (from 5 to 100).

stationary Held-out Likelihood. Considering this, an optimal number of $T = 20$ topics was identified.

The results of the topic modeling algorithm are shown in Table 5. For each of the 20 identified topics, Table 5 also reports the list of keywords according to the two alternative criteria (highest probability and FREX) discussed in Section labeling.

Labeling

In addition to the relevant lists of keywords, Table 5 shows the labels defined by the authors. After a first phase of independent analysis which led to partially different labels, a brainstorming among the authors allowed to settle the differences and obtain the final list of labels listed in Table 5. Both the independent analysis and brainstorming were done taking into account the lists of keywords as well as the most relevant reviews for each topic. Finally, the defined topic labels were submitted for confirmation to an external panel familiar with management, engineering design and car-sharing systems.

Data verification

Obtained results were verified by comparing the assigned topic of a randomly selected sample composed of 100 reviews with a manual topic assignment performed by the authors. For each of the 100 reviews, the authors were requested to agree in the association of one or more of the 20 topics identified by STM. Based on the comparison between manual and STM topic assignment (see Table 6), the four verification indicators presented in Section data verification were calculated (see Table 7). These metrics show a generally good correspondence between the assignment produced by STM and the authors. The accuracy of 94% proves good effectiveness of the method to predict the content of the reviews, correctly identifying true positive and true negative. The Recall and Precision indicators, respectively equal to 73% and 65%, show that the method performs well in terms identification of the topics (true positive).

Result analysis: Insights on car-sharing quality determinants

Figure 4 shows the proportion (i.e., the average topical prevalence) of the 20 identified topics in the analyzed reviews. The most discussed topics are topic 6 (app reliability), 9 (convenience) and 15 (customer service responsiveness). The less discussed topics are those related to the tangible component of the car-sharing service: topic 2 (accidents and damage management) and topic 8 (car condition). Note that this does not mean these topics are less or more "critical to quality" than others. The difference in proportions may depend on a number of factors, including the review aggregators used for the analysis, which may be more (or less) oriented toward collecting specific information on certain topics. For example, the Playstore commonly collects information related to the user experience with respect to the mobile applications.

Topics vs. Car sharing schemes

Figure 5 shows the association between identified topics and the scheme of car sharing (station-based or free-floating). In detail, the figure shows the marginal effects in the change of expected proportions of topical prevalence (i.e., the proportion of document devoted to a given topic): dots on the chart depict expected difference in topic proportions, while the horizontal bars represent uncertainty ($\alpha = 95\%$) in these estimates. As an example, see how topic 14 is more discussed (about 4%) for station-based than for free-floating car-sharing services.

The following considerations can be made by analyzing the figure:

- topic 12 (car availability) and 17 (car startup issues) do not seem to be influenced by the type of car-sharing scheme;
- reviews related to topic 3 (registration process), 5 (parking areas), 6 (app reliability), 7 (end trip issues), 11 (car proximity), 13 (efficacy) and 16 (intermodal transportation) are prevalent for free-floating scheme. It is noteworthy how topic 6 (app reliability) has an important prevalence of 15%.

Table 5. Keywords and labels of the identified topics.

	Criterium	Keywords	Label
1	*Highest Prob*	help, phone, call, person, office, answer, number	CUSTOMER SERVICE (PHYSICAL OFFICE)
	FREX	help, office, staff, answer, phone, person, question	
2	*Highest Prob*	damage, report, accident, fault, member, enterprise, claim,	ACCIDENT & DAMAGES MANAGEMENT
	FREX	damage, accident, minor, claim, deduct, fault, scratch	
3	*Highest Prob*	sign, process, website, license, drive, driver, registration	REGISTRATION PROCESS
	FREX	license, application, registration, process, driver, website, sign	
4	*Highest Prob*	charge, fee, late, return, time, pay, hour	CHARGES & FEES
	FREX	fee, charge, late, return, dollar, extra, extend	
5	*Highest Prob*	park, lot, spot, find, ticket, street, space	PARKING AREAS
	FREX	park, spot, ticket, space, street, lot, mete	
6	*Highest Prob*	app, work, update, book, map, reserve, time	APP RELIABILITY
	FREX	app, map, crash, feature, load, slow, version	
7	*Highest Prob*	trip, end, time, make, actual, take, system	END TRIP ISSUES
	FREX	trip, end, life, stuck, make, connect, actual	
8	*Highest Prob*	gas, dirty, rent, clean, tank, card, tire	CAR CONDITION
	FREX	dirty, smell, smoke, tank, hair, seat, dog	
9	*Highest Prob*	need, convenient, quick, recommend, awesome, clean, perfect	CONVENIENCE
	FREX	awesome, excel, amazing, perfect, quick, convenient, super	
10	*Highest Prob*	hour, price, rate, cost, expense, mile, cheaper	USE RATES
	FREX	price, rate, expense, cost, daily, tax, rental	
11	*Highest Prob*	minute, reservation, walk, wait, home, time, away	CAR PROXIMITY
	FREX	minute, walk, home, wait, figure, block, stand	
12	*Highest Prob*	car, available, location, vehicle, area, change, time	CAR AVAILABILITY
	FREX	available, vehicle, car, select, search, choose, date	
13	*Highest Prob*	use, time, now, far, user, review, star	EFFICACY
	FREX	use, user, far, happy, review, love, code	
14	*Highest Prob*	city, year, insurance, member, gas, need, month	SHARING BENEFITS
	FREX	errand, hybrid, live, city, suv, insurance, variety	
15	*Highest Prob*	service, custom, issue, company, terrible, problem, experience	CUSTOMER SERVICE RESPONSIVENESS
	FREX	service, custom, terrible, issue, support, resolve, company	
16	*Highest Prob*	way, drive, little, take, get, town, bus	INTERMODAL TRANSPORTATION
	FREX	town, bus, airport, taxi, bike, store, run	
17	*Highest Prob*	time, start, location, turn, lock, pick, key	CAR STARTUP ISSUES
	FREX	key, turn, battery, lock, door, start, waste	
18	*Highest Prob*	call, member, cancel, ask, rep, refund, manage	CUSTOMER SERVICE COURTESY
	FREX	representative, supervisor, agent, rude, manage, speak, conversation	
19	*Highest Prob*	account, card, email, credit, month, day, membership,	BILLING AND MEMBERSHIP
	FREX	account, email, payment, bank, credit, card, address	
20	*Highest Prob*	reservation, plan, time, need, book, cancel, advance	CAR RESERVATION
	FREX	plan, advance, entire, reservation, chance, screw, ruin	

Table 6. Examples of verification procedures. Total number of topics equal to 20.

STM topic assignment	Manual topic assignment	True Positive	True Negative	False Positive	False Negative
7 – 10	7 – 10 – 11	2	17	0	1
4 – 13 – 18	4 – 13	2	17	1	0
4 – 11	4 – 15	1	17	1	1
6	6	1	19	0	0

Table 7. Indicators values for the car-sharing case study.

Name	Value
Accuracy	0.94
Recall	0.73
Precision	0.65
F measure	0.69

- topic 1 (customer service – physical office), 2 (accident and damages management), 4 (charges and fees), 8 (car condition), 9 (convenience), 10 (use rates), 14 (sharing benefits), 15 (customer service responsiveness), 18 (customer service courtesy), 19 (billing and membership) and 20 (car reservation) are mainly discussed for station-based car-sharing scheme.

In general, the association between the topic proportion and the type of PSS (in this case, the car sharing scheme) allows to highlight the most critical determinants. This information can guide improvement and investment decisions in order to increase quality and customer satisfaction.

Topics vs. scores

Figure 6 shows the marginal effects in the change of the expected proportion of topical prevalence based on low (1 and 2) and high (4 and 5) review score. As the distance from the dashed axis increases, the probability of specific topics becomes more dominant. Three different clusters of topics can be qualitatively identified:

- topic 1 (customer service – physical office), 8 (car condition), 17 (car startup issues) and 20 (car reservation) are "neutral" with respect to the review score;
- topic 5 (parking areas), 6 (app reliability), 7 (end trip issues), 9 (convenience), 10 (use rates), 11(car proximity), 12 (car availability), 13 (efficacy), 14 (sharing benefits) and 16 (intermodal transportation) seem to be drivers of high scores.
- topic 2 (accident & damages management), 3 (registration process), 4 (charges & fees), 15 (customer service responsiveness), 18 (customer service courtesy) and 19 (billing and membership) are critical factors that lead to customer dissatisfaction when occurring.

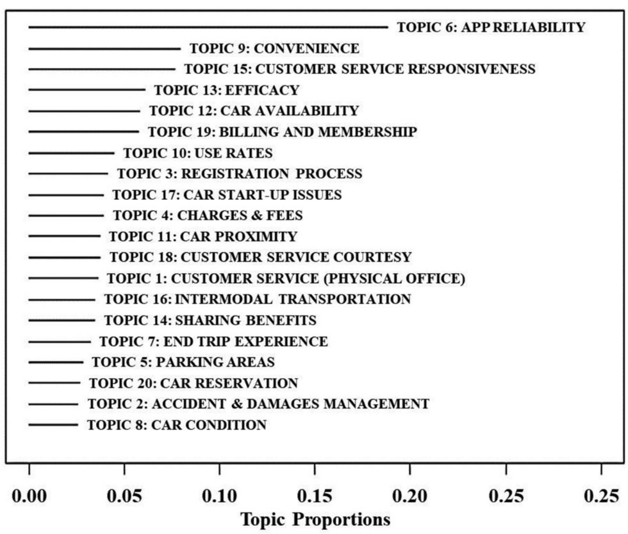

Figure 4. Topic proportions. Note that a greater proportion of discussion about one topic does not necessarily mean that topic is more critical to quality than others.

The relationship between quality determinants and review scores is a critical information for the management and design of PSS. It can address managers to identify the PSS critical issues from the customer satisfaction point of view.

Topics trend analysis

The evolution of topic proportion over time is a further point of interest. Analyzing this element allows producing a broader picture of how user needs and expectations are evolving. Figure 7 shows the trend of the quality determinants identified in the period between 2006 and 2019. Three different groups of evolutionary behaviors can be identified:

- *Stationary topics*, i.e. those topics whose proportion remains almost stable over time (see Figure 7A). Topic 1 (customer service - physical office); 2

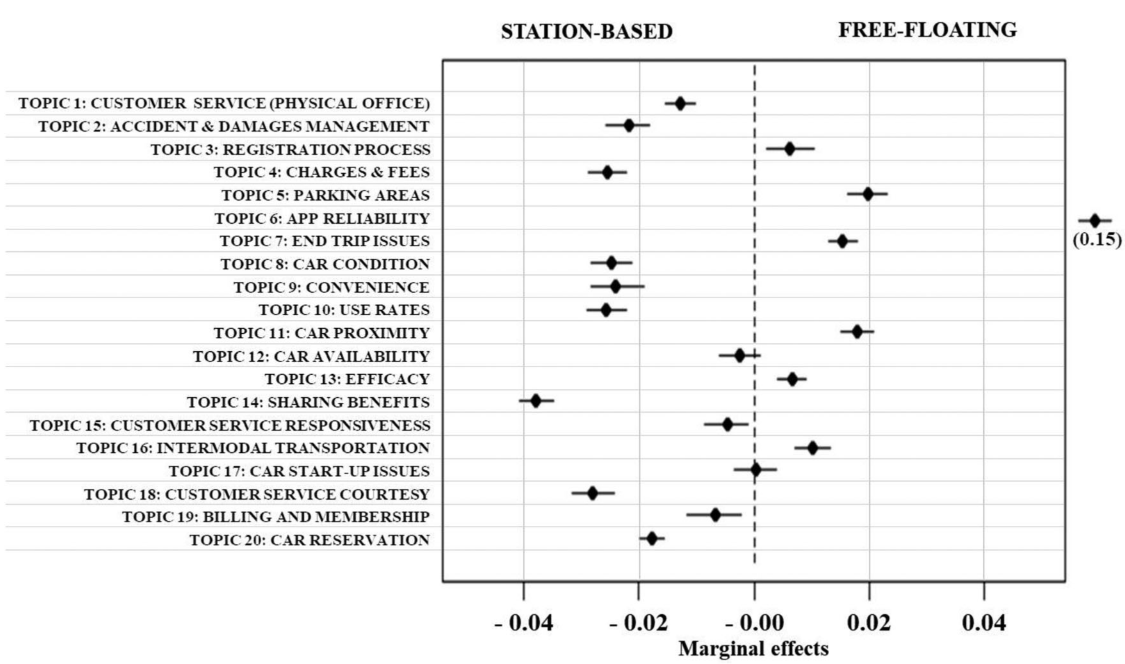

Figure 5. Marginal effects in the change of expected proportions of topical prevalence based on the scheme of car-sharing.

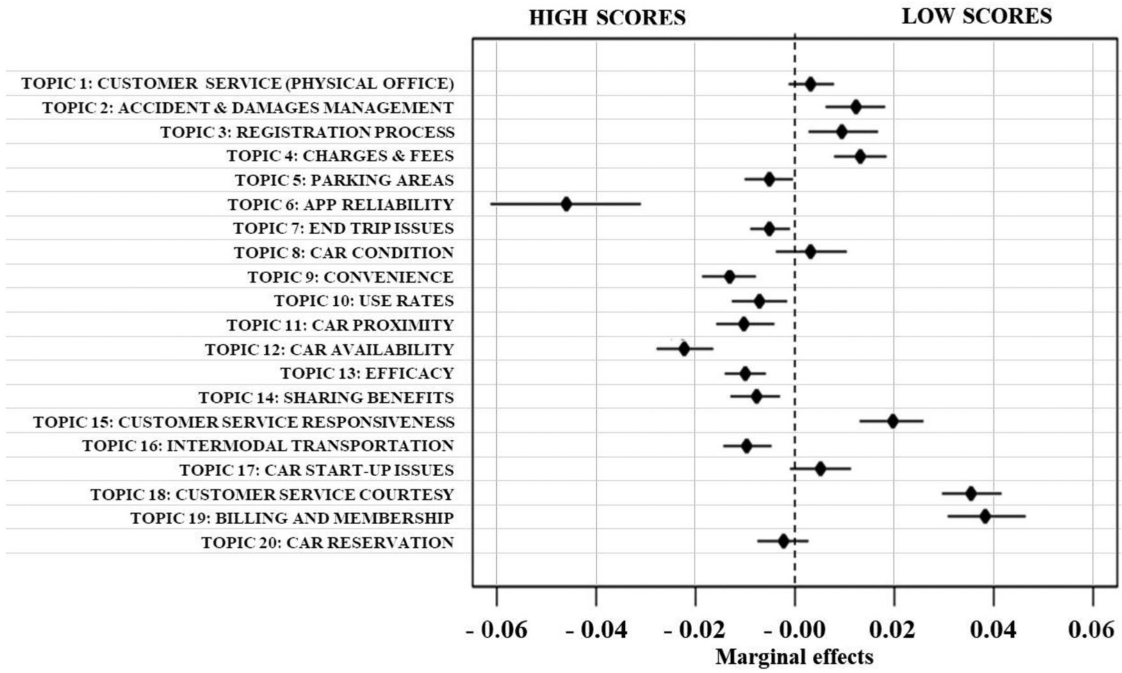

Figure 6. Marginal effects in the change of expected proportions of topical prevalence based on low and high review score. The dotted line represents the zero effect.

(accident & damages management); 5 (parking areas); 7 (end trip issues); 11 (car proximity); 12 (car availability); 17 (car startup issues); 18 (customer service courtesy) fall into this category. This pattern highlights the determinants of quality that have kept their interest unchanged for users over time.

- *Decreasing topics*, i.e. those topics that exhibit a clear decrease in topic proportion over time (see Figure 7B). Topic 4 (charges & fees); 8 (car condition); 10 (use rates); 14 (sharing benefits); 16 (intermodal transportation); 20 (car reservation) fall into this category. A possible explanation for this pattern is the loss of interest over time. As users became more familiar with some PSS aspects, these determinants gradually lost importance. Consider, for example, the "sharing benefits" determinant, it was highly discussed in the early years, becoming marginal in recent years.
- *Increasing topics*, i.e. those topics whose proportion has grown over time (see Figure 7C). Topic 3 (registration process); 6 (app reliability); 9 (convenience); 13 (efficacy); 15 (customer service responsiveness); 19 (billing and membership) fall into this category. The observed increase in topic proportion can be attributed to the emergence of new customer needs or to the greater importance given to initially little considered determinants.

Topic correlations

The analysis of the correlations between the topics can provide additional insights. Figure 8 reports a map of the marginal (positive) correlations among the topics: strongly related topics are represented as close to each other. Correlations were calculated considering topical prevalence of each review (i.e., the proportion of review devoted to a given topic). Positive correlations between topics suggest that the two topics are likely to be discussed within the same reviews (Roberts, Stewart, and Tingley 2019).

This type of visualization allows to identify different clusters of topics:

- topic 5 (parking areas), 7 (end trip issues), 10 (use rates), 11 (car proximity), 14 (sharing benefits), 16 (intermodal transportation), 17 (car startup issues) are positively correlated and can hence be grouped in a macro-category related to the ease of use of the car-sharing service;
- being positively related, topic 1 (customer service – physical office), 3 (registration process), 4 (charges & fees), 15 (customer service responsiveness), 18 (customer service courtesy), 19 (billing and membership) and 20 (car reservation) can be clustered into a single group – concerning customer interaction – as they largely refer to the relationships between customers and the car-sharing company;

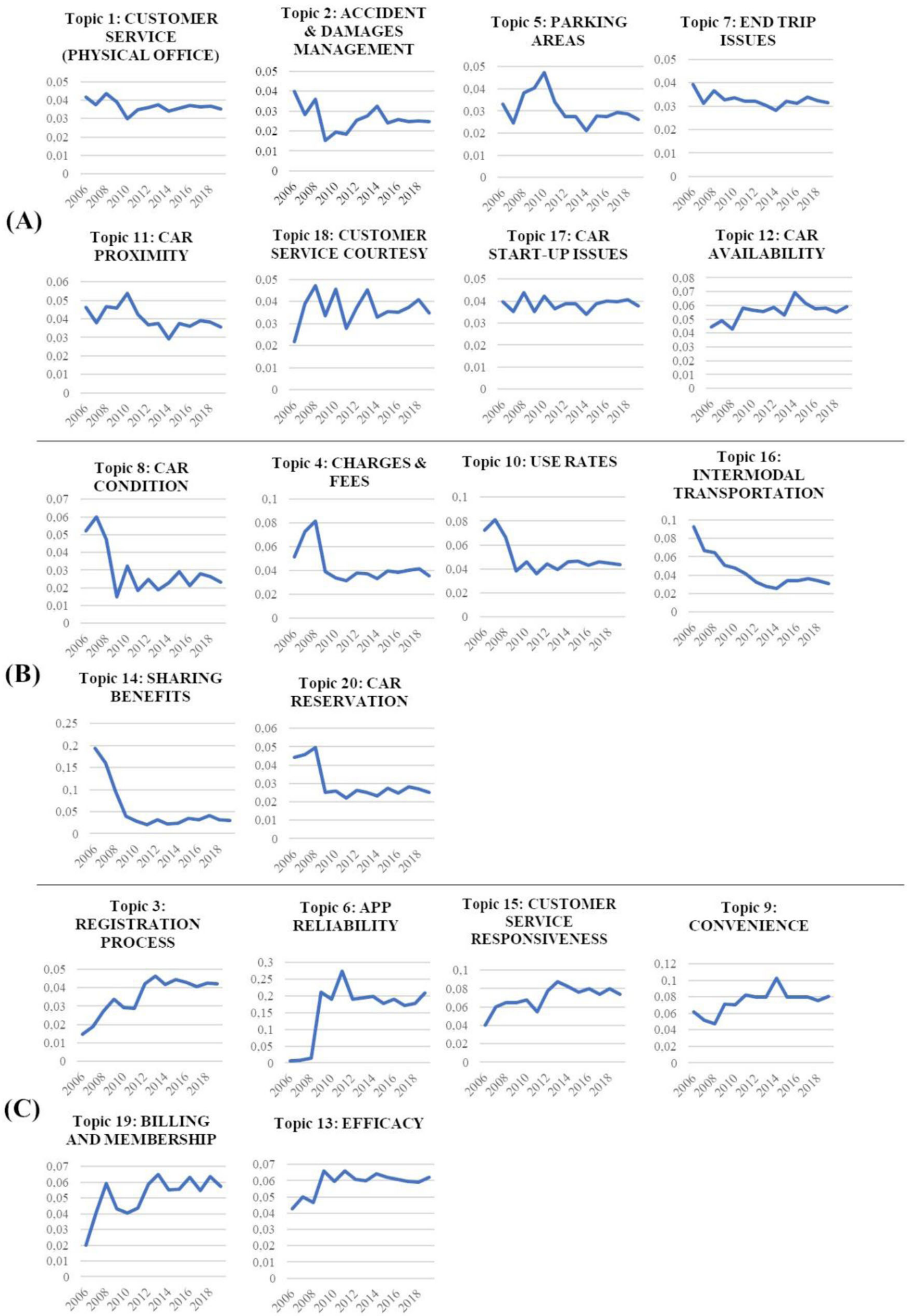

Figure 7. Evolution of topic proportion over time (2006 to 2019). (A) Stationary topics; (B) Decreasing topics; (C) Increasing topics.

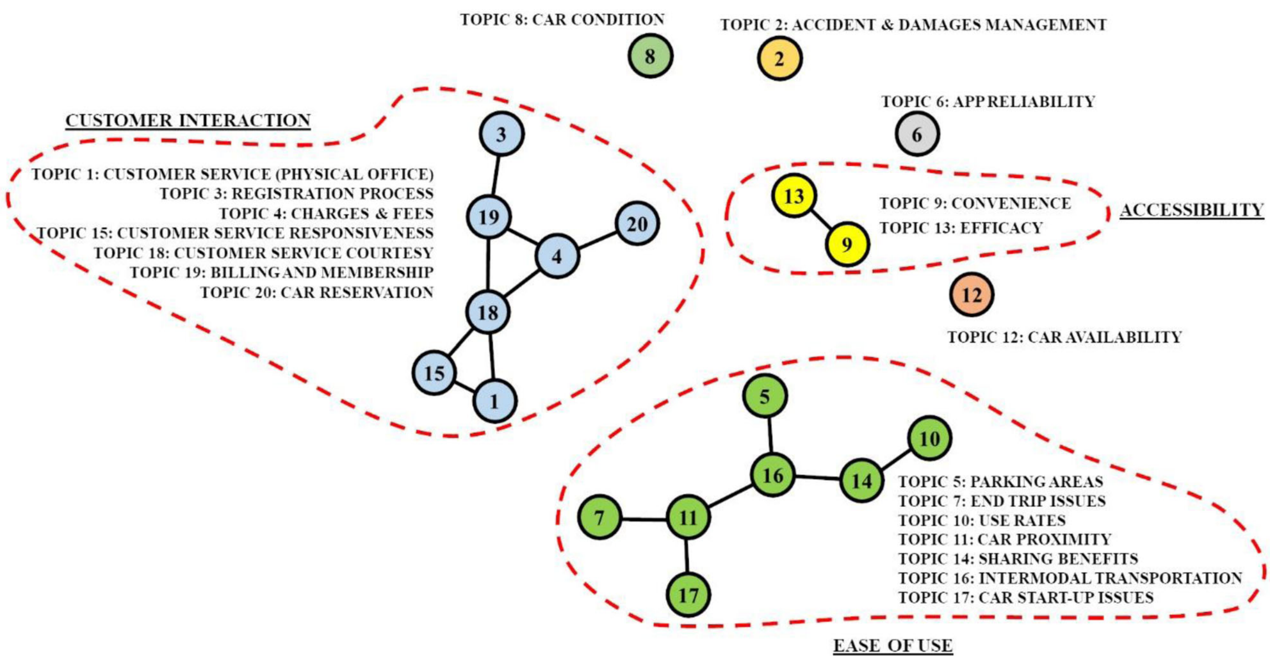

Figure 8. Map of the marginal (positive) correlations among topics. Only correlations greater than 0.2 are shown in the form of a link between topics.

- The last cluster, related to the accessibility of the service, includes topic 9 (convenience) and 13 (efficacy).

Other topics, such as topic 2 (accident & damages management), 6 (app reliability), 8 (car condition) and 12 (car availability) are not significantly related to each other.

A structured comparison with service quality dimensions

Table 8 compares the resulting quality determinants of car sharing with those of the ServQual model (Parasuraman, Zeithaml, and Berry 1988). The ServQual model only partially matches the results of this investigation. The differences can be traced back to several reasons, including:

- The focus of the two frameworks is slightly different. The ServQual model is focused on services, whereas the aim of the proposed methodology is to identify quality determinants of a PSS, where both products and services are offered to concurrently fulfill customer needs.
- The research approaches are different. While the ServQual model investigates customer perceptions and expectancy through the filling of questionnaires, the herein proposed methodology attempts to identify PSS quality determinants through the analysis of User-Generated Content.

Implications

The results of this preliminary investigation may have multiple implications for different stakeholders, including practitioners and academics:

- *Implications for quality practitioners*: the described methodology allows the definition of specific quality determinants in different business contexts. Its application in all those contexts where the use of standardized quality tools – such as questionnaires or surveys – may be limited is of particular importance. There are several sources of UGCs that can be used for identification of quality determinants. Specifically, reviews generated by users were considered in this study. Interesting (and potentially complementary) insights can result from the application of the proposed methodology to other different UGCs, such as complaints or usage reports.

The practical application of the proposed methodology may provide several benefits, including: (i) lower costs in terms of time and resources compared to traditional methodologies (such as survey, questionnaires, interviews); (ii) development of ad-hoc

Table 8. Comparison between quality determinants of the ServQual model (Parasuraman et al., 1988) and car-sharing quality determinants.

ServQual dimensions (Service)		Car-sharing quality determinants (Product-Service System)
Reliability	The ability to perform the promised service dependably and accurately	Topic 6: App reliability Topic 7: End trip issues Topic 17: Car startup issues Topic 20: Car reservation
Assurance	The knowledge and courtesy of employees and their ability to convey trust and confidence	Topic 2: Accident & damages management Topic 18: Customer service courtesy
Tangibles	The appearance of physical facilities, equipment, personnel and communication materials	Topic 8: Car condition
Empathy	The provision of caring and individualized attention to customer	Topic 19: Account management
Responsiveness	The willingness to help customers and to provide prompt service	Topic 1: Customer service (physical office) Topic 3: Registration process Topic 12: Car availability Topic 15: Customer service responsiveness
–	–	Topic 4: Charges & fees Topic 5: Parking areas Topic 9: Convenience Topic 10: Use rates Topic 11: Car proximity Topic 13: Efficacy Topic 14: Sharing benefits Topic 16: Intermodal transportation

solutions for the context under analysis.Moreover, the identified quality determinants may provide a reference framework for the development of traditional questionnaires for monitoring and measuring customer satisfaction.A further starting point for analysis could come from the temporal analysis of the topics discussed. This survey can provide valuable insights for quality tracking: the presence of discontinuities in the proportion of topics discussed in different periods of analysis may highlight the occurrence of issues in the use of a particular PSS. Also, the connection between review scores and topics can provide information on the evolution of customer perception (satisfaction or dissatisfaction) of specific elements of the PSS.

- *Implications for academics*: the results of this study show that the quality determinants traditionally used to measure the quality of services or products are not sufficient to describe complex value offerings such as PSS. Therefore, the need for new tools and a new theoretical framework to describe the quality determinants of a PSS is highlighted. The proposed results can be the first step of a wider study to analyze different contexts to find a common theoretical framework.
- *Implication for engineering design*: despite being preliminary, the findings of this article highlight the potentialities that data-driven methodologies may have in engineering design. Understanding the quality determinants of PSS provides some support for their engineering design. UGCs may serve as primary source for the identification of customer needs, also analyzing the temporal dynamics of the quality determinants and eventually predicting patterns and anticipating customer needs. Further researches might explore how to integrate the application of data-driven tools into existing engineering design methodology.
- *Implications for car-sharing providers*: the analysis in this paper produces 20 topics, grouped in 8 clusters (customer interaction, ease of use, accessibility, car availability, app reliability, accident and damage management, car condition), that can influence the perception of quality and satisfaction of car-sharing services customers. This information can be used by car-sharing providers to build specific PSS quality control tools (audits, questionnaires, interviews, etc.) and to develop innovative car-sharing PSS increasingly aligned with customer needs and expectations.

Conclusions

By analyzing a large sample of unstructured UGCs, spontaneously produced by users, this study has shown how to extract valuable information about the quality determinants of a Product-Service System. To this end, a practical methodology is proposed and exemplified on a specific case study, i.e., car-sharing. The investigation unveils 20 determinants of quality that – although being specific to the case of car-sharing – are partially matching with those of the

ServQual framework (Parasuraman, Zeithaml, and Berry 1988).

The results of this research support the idea that the application of data-driven techniques and big data analytics may represent a great opportunity for the quality engineering and management fields. Overall, this analysis presents several aspects of novelty: (i) it is one of the first attempts to define a framework for the quality evaluation of PSS; (ii) the inference of quality determinants was possible given the use of Structural Topic Modeling techniques that allowed the concurrent analysis of both the UGCs and relevant metadata, such as the scores assigned by the users; (iii) the study of the quality determinants of a car-sharing service is a scarcely discussed research topic although the car-sharing is an increasingly important part of the transport industry and represents a virtuous example of circular economy.

Despite producing multi-faceted insights, the adopted method can potentially be subject to some distortions due to the composition of the analyzed sample of UGC. Moreover, the analysis introduces some elements of subjectivity (e.g., in the labeling operation). Future developments will be directed to face these limitations as well as to the use the proposed approaches for investigating the quality of other PSSs. In general, the aim of these studies could be the identification of a shared framework for the quality assessment of a generic PSS. Further efforts will also be aimed to the understanding of how temporal information can be used for a quality monitoring through appropriate control charts.

Funding

This work has been partially supported by "Ministero dell'Istruzione, dell'Università e della Ricerca". Award "TESUN-83486178370409 finanziamento dipartimenti di eccellenza CAP.1694 TIT. 232 ART. 6.

About the authors

Luca Mastrogiacomo is an Associate Professor at the Department of Management and Production Engineering at Politecnico di Torino. His main scientific interests currently concern the areas of Service Quality and Quality Engineering. Mastrogiacomo may be contacted at luca.mastrogiacomo@polito.it.

Federico Barravecchia is a Postdoc at the Department of Management and Production Engineering at Politecnico di Torino. His main scientific interests currently concern the areas of Service Quality and Quality Engineering. Barravecchia may be contacted at federico.barravecchia@polito.it.

Fiorenzo Franceschini is a Professor of Quality Engineering at Politecnico di Torino (Italy) – Department of Management and Production Engineering. He is author or co-author of nine books and more than 270 published papers in prestigious scientific journals, and international conference proceedings. His current research interests are in the areas of Quality Engineering, Performance Measurements Systems and Service Quality. Franceschini may be contacted at fiorenzo.franceschini@polito.it

Frederic Marimon Viadiu is a Professor of Quality & Innovation at Universitat Internacional de Catalunya (Spain) – Department of Economy and Business Organization. He is author or co-author of three books and more than 200 published papers in prestigious scientific journals, and international conference proceedings. His current research interests are in the areas of Quality Management, quality assessment in services, and diffusion analysis of different management standards. Marimon may be contacted at fmarimon@uic.es

ORCID

Frederic Marimon http://orcid.org/0000-0002-5572-7341

References

Aggarwal, C. C., and C. Zhai. 2012. *Mining text data*. New York: Springer-Verlag.

Allen, T. T., Z. Sui, and K. Akbari. 2018. Exploratory text data analysis for quality hypothesis generation. *Quality Engineering* 30 (4):701–12. doi: 10.1080/08982112.2018.1481216.

Allen, T. T., and H. Xiong. 2012. Pareto charting using multifield freestyle text data applied to Toyota Camry user reviews. *Applied Stochastic Models in Business and Industry* 28 (2):152–63. doi: 10.1002/asmb.947.

Allen, T. T., H. Xiong, and A. Afful-Dadzie. 2016. A directed topic model applied to call center improvement. *Applied Stochastic Models in Business and Industry* 32 (1):57–73. doi: 10.1002/asmb.2123.

Barravecchia, F., L. Mastrogiacomo, and F. Franceschini. 2020. The Player-Interface method: An approach to support Product-Service Systems concept generation and prioritization. *Journal of Engineering Design* 31 (6):331–48. doi: 10.1080/09544828.2020.1743822.

Barreda, A., and A. Bilgihan. 2013. An analysis of user-generated content for hotel experiences. *Journal of Hospitality and Tourism Technology* 4 (3):263–80. doi: 10.1108/JHTT-01-2013-0001.

Beuren, F. H., M. G. Gomes Ferreira, and P. A. Cauchick Miguel. 2013. Product-service systems: A literature review on integrated products and services. *Journal of Cleaner Production* 47:222–31. doi: 10.1016/j.jclepro.2012.12.028.

Bi, J.-W., Y. Liu, Z.-P. Fan, and E. Cambria. 2019. Modelling customer satisfaction from online reviews using ensemble neural network and effect-based Kano model. *International Journal of Production Research* 57 (22):7068–88. doi: 10.1080/00207543.2019.1574989.

Blei, D. M. 2012. Probabilistic topic models. *Communications of the ACM* 55 (4):77–84. doi: 10.1145/2133806.2133826.

Blei, D. M., and J. D. Lafferty. 2007. A correlated topic model of science. *The Annals of Applied Statistics* 1 (1): 17–35. doi: 10.1214/07-AOAS114.

Blei, D. M., A. Y. Ng, and M. I. Jordan. 2003. Latent dirichlet allocation. *Journal of Machine Learning Research* 3:993–1022.

Boehm, M., and O. Thomas. 2013. Looking beyond the rim of one's teacup: A multidisciplinary literature review of Product-Service Systems in Information Systems, Business Management, and Engineering Design. *Journal of Cleaner Production* 51:245–60. doi: 10.1016/j.jclepro.2013.01.019.

Chang, Y.-H., and C.-H. Yeh. 2002. A survey analysis of service quality for domestic airlines. *European Journal of Operational Research* 139 (1):166–77. doi: 10.1016/S0377-2217(01)00148-5.

Costa, E., A. Lorena, A. Carvalho, and A. Freitas. 2007. A review of performance evaluation measures for hierarchical classifiers. In *Evaluation Methods for machine Learning II: Papers from the AAAI-2007 Workshop*, 1–6.

DeVellis, R. F. 2016. *Scale development: Theory and applications*. 4th ed. Thousand Oaks: Sage publications.

Dholakia, U. M., S. S. Singh, and R. A. Westbrook. 2010. Understanding the effects of post-service experience surveys on delay and acceleration of customer purchasing behavior: Evidence from the automotive services industry. *Journal of Service Research* 13 (4):362–78. doi: 10.1177/1094670510375105.

Durrani, N., H. Schmid, A. Fraser, P. Koehn, and H. Schütze. 2015. The operation sequence model—combining N-gram-based and phrase-based statistical machine translation. *Computational Linguistics* 41 (2):185–214. doi: 10.1162/COLI_a_00218.

Dyllick, T., and Z. Rost. 2017. Towards true product sustainability. *Journal of Cleaner Production* 162:346–60. doi: 10.1016/j.jclepro.2017.05.189.

Franceschini, F.,. M. Galetto, and D. Maisano. 2019. *Designing Performance Measurement Systems: Theory and Practice of Key Performance Indicators*. Cham: Springer International Publishing.

Groves, R. M. 2006. Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly* 70 (5):646–75. doi: 10.1093/poq/nfl033.

Guglielmetti Mugion, R., M. Toni, L. Di Pietro, M. G. Pasca, and M. F. Renzi. 2019. Understanding the antecedents of car sharing usage. An empirical study in Italy. *International Journal of Quality and Service Sciences* 11 (4):523–41. doi: 10.1108/IJQSS-02-2019-0029.

Guo, Y., S. J. Barnes, and Q. Jia. 2017. Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent dirichlet allocation. *Tourism Management* 59:467–83. doi: 10.1016/j.tourman.2016.09.009.

Ho, Y.-C., J. Wu, and Y. Tan. 2017. Disconfirmation effect on online rating behavior: A structural model. *Information Systems Research* 28 (3):626–42. doi: 10.1287/isre.2017.0694.

Illgen, S., and M. Höck. 2019. Literature review of the vehicle relocation problem in one-way car sharing networks. *Transportation Research Part B: Methodological* 120:193–204. doi: 10.1016/j.trb.2018.12.006.

James, T. L., E. D. V. Calderon, and D. F. Cook. 2017. Exploring patient perceptions of healthcare service quality through analysis of unstructured feedback. *Expert Systems with Applications* 71:479–92. doi: 10.1016/j.eswa.2016.11.004.

Jeong, Y., Y. Yang, J. Suk, and K. Kim. 2019. A text-mining analysis of online reviews on car-sharing services. In *International Conferences on Internet Technologies & Society 2019*, 167–9.

Jivani, A. G. 2011. A comparative study of stemming algorithms. *International Journal of Computer Applications in Technology* 2 (6):1930–8.

Kambatla, K.,. G. Kollias, V. Kumar, and A. Grama. 2014. Trends in big data analytics. *Journal of Parallel and Distributed Computing* 74 (7):2561–73. doi: 10.1016/j.jpdc.2014.01.003.

Korfiatis, N., P. Stamolampros, P. Kourouthanassis, and V. Sagiadinos. 2019. Measuring service quality from unstructured data: A topic modeling application on airline passengers' online reviews. *Expert Systems with Applications* 116:472–86. doi: 10.1016/j.eswa.2018.09.037.

Krumm, J., N. Davies, and C. Narayanaswami. 2008. User-generated content. *IEEE Pervasive Computing* 7 (4):10–1. doi: 10.1109/MPRV.2008.85.

LaValle, S., E. Lesser, R. Shockley, M. S. Hopkins, and N. Kruschwitz. 2011. Big data, analytics and the path from insights to value. *MIT Sloan Management Review* 52 (2): 21–32.

Lee, J., D.-H. Park, and I. Han. 2008. The effect of negative online consumer reviews on product attitude: An information processing view. *Electronic Commerce Research and Applications* 7 (3):341–52. doi: 10.1016/j.elerap.2007.05.004.

Liu, B. 2012. Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies* 5 (1):1–167. doi: 10.2200/S00416ED1V01Y201204HLT016.

Liu, A., T. Wuest, W. Wei, and S. Lu. 2014. Application of prospect theory on car sharing product service system. *Procedia CIRP* 16:350–5. doi: 10.1016/j.procir.2014.02.022.

Marimon, F., J. Llach, M. Alonso-Almeida, and M. Mas-Machuca. 2019. CC-Qual: A holistic scale to assess customer perceptions of service quality of collaborative consumption services. *International Journal of Information Management* 49:130–41. doi: 10.1016/j.ijinfomgt.2019.03.009.

Mastrogiacomo, L., F. Barravecchia, and F. Franceschini. 2019. A worldwide survey on manufacturing servitization. *The International Journal of Advanced Manufacturing Technology* 103 (9-12):3927–42. doi: 10.1007/s00170-019-03740-z.

Mastrogiacomo, L., F. Barravecchia, and F. Franceschini. 2020. Enabling factors of manufacturing servitization: Empirical analysis and implications for strategic positioning. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture* 234 (9):1258–70. doi: 10.1177/0954405420911781.

Matakos, A., and P. Tsaparas. 2016. Temporal mechanisms of polarization in online reviews. In 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM). IEEE, 529–532. doi: 10.1109/ASONAM.2016.7752286.

Mert, G., S. Waltemode, and J. C. Aurich. 2014. Quality assessment of technical product-service systems in the machine tool industry. *Procedia CIRP* 16:253–8. doi: 10.1016/j.procir.2014.06.002.

Meyer, D., K. Hornik, and I. Feinerer. 2008. Text mining infrastructure in R. *Journal of Statistical Software* 25 (5): 1–54.

Mukherjee, S. P. 2019. *Quality, domains and dimensions*. Singapore: Springer.

Müller, O., I. Junglas, J. Vom Brocke, and S. Debortoli. 2016. Utilizing big data analytics for information systems research: Challenges, promises and guidelines. *European Journal of Information Systems* 25 (4):289–302. doi: 10.1057/ejis.2016.2.

Nassirtoussi, A. K., S. Aghabozorgi, T. Y. Wah, and D. C. L. Ngo. 2014. Text mining for market prediction: A systematic review. *Expert Systems with Applications* 41 (16):7653–70. doi: 10.1016/j.eswa.2014.06.009.

Neri, A., E. Cagno, G. Di Sebastiano, and A. Trianni. 2018. Industrial sustainability: Modelling drivers and mechanisms with barriers. *Journal of Cleaner Production* 194: 452–72. doi: 10.1016/j.jclepro.2018.05.140.

OECD2001. *Understanding the digital divide*. Available at: http://www.oecd.org/dataoecd/38/57/1888451.pdf.

Pang, B., and L. Lee. 2008. Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval* 2 (1–2):1–135. doi: 10.1561/1500000011.

Parasuraman, A., V. A. Zeithaml, and L. L. Berry. 1988. Servqual: A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing* 64 (1):12.

Park, D.-H., J. Lee, and I. Han. 2007. The effect of on-line consumer reviews on consumer purchasing intention: The moderating role of involvement. *International Journal of Electronic Commerce* 11 (4):125–48. doi: 10.2753/JEC1086-4415110405.

Reinsel, D., J. Gantz, and J. Rydning. 2018. 'The Digitization of the World From Edge to Core', *IDC White Paper*. Available at: https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf.

Roberts, M. E., B. M. Stewart, and D. Tingley. 2019. STM: R package for structural topic models. *Journal of Statistical Software* 91 (2):1–40. doi: 10.18637/jss.v091.i02.

Roberts, M. E., B. M. Stewart, D. Tingley, C. Lucas, J. Leder-Luis, S. K. Gadarian, B. Albertson, and D. G. Rand. 2014. Structural topic models for open-ended survey responses. *American Journal of Political Science* 58 (4):1064–82. doi: 10.1111/ajps.12103.

Robinson, T. J., R. C. Giles, and R. U. Rajapakshage. 2020. Discussion of "Experiences with big data: Accounts from a data scientist's perspective. *Quality Engineering* 32 (4): 543–9. doi: 10.1080/08982112.2020.1758333.

Scott, J., and J. Baldridge. 2013. A recursive estimate for the predictive likelihood in a topic model. *Journal of Machine Learning Research* 31:527–35.

Shaheen, S. A., and A. P. Cohen. 2007. Growth in worldwide car-sharing: An international comparison. *Transportation Research Record: Journal of the Transportation Research Board* 1992 (1):81–9. doi: 10.3141/1992-10.

Shaheen, S. A., and A. P. Cohen. 2013. Carsharing and personal vehicle services: Worldwide market developments and emerging trends. *International Journal of Sustainable Transportation* 7 (1):5–34. doi: 10.1080/15568318.2012.660103.

Sweeney, J. C., G. N. Soutar, and T. Mazzarol. 2008. Factors influencing word of mouth effectiveness: Receiver perspectives. *European Journal of Marketing* 42 (3/4):344–64. doi: 10.1108/03090560810852977.

Tirunillai, S., and G. J. Tellis. 2012. Does chatter really matter? Dynamics of user-generated content and stock performance. *Marketing Science* 31 (2):198–215. doi: 10.1287/mksc.1110.0682.

Tirunillai, S., and G. J. Tellis. 2014. Mining marketing meaning from online chatter: Strategic brand analysis of big data using latent dirichlet allocation. *Journal of Marketing Research* 51 (4):463–79. doi: 10.1509/jmr.12.0106.

Wallach, H. M., D. M. Mimno, and A. McCallum. 2009. Rethinking LDA: Why priors matter. *Advances in Neural Information Processing Systems* 23:1973–81.

Waltemode, S., C. Mannweiler, and J. C. Aurich. 2012. Life cycle oriented quality assessment of Technical Product-Service Systems. In *Leveraging Technology for a Sustainable World - Proceedings of the 19th CIRP Conference on Life Cycle Engineering*.

Williams, A. 2007. Product service systems in the automobile industry: Contribution to system innovation? *Journal of Cleaner Production* 15 (11-12):1093–103. doi: 10.1016/j.jclepro.2006.05.034.

Zaki, M., and J. R. McColl-Kennedy. 2020. Text Mining Analysis Roadmap (TMAR) for Service. *Journal of Services Marketing* 34 (1):30–47. doi: 10.1108/JSM-02-2019-0074.

Zhang, K. Z. K., S. J. Zhao, C. M. K. Cheung, and M. K. O. Lee. 2014. Examining the influence of online reviews on consumers' decision-making: A heuristic–systematic model. *Decision Support Systems* 67:78–89. doi: 10.1016/j.dss.2014.08.005.

Zonnenshain, A., and R. S. Kenett. 2020. Quality 4.0—the challenging future of quality engineering. *Quality Engineering* 32 (4):614–26. doi: 10.1080/08982112.2019.1706744.