UNIVERSITY OF CAMBRIDGE
Cambridge Service Alliance

Textual Regularity Mining and Data-driven Customer Experience Analytics

Mohamed Zaki and Benjamin Lucas

This is a Working Paper

Why this paper might be of interest to Alliance Partners:

In this working paper, Mohamed Zaki and Benjamin Lucas show why the complex, intertwined nature of customer experience (CX) necessitates a data-driven analytics approach, but also why conceptual CX frameworks should be treated as part of such toolkits. Using three primary datasets; a social media customer feedback dataset for a large retailer, an online review dataset from a not-for-profit organization, and a national customer experience text survey from a B2B service provider, the authors retrieve and validate thematic groupings of customer experiences, and then explore these in greater detail using literal markers of the dimensions “Objects and Entities”, “Thoughts, Feelings and Sensations”, “Behaviors” and “Ownership and Relation”. Specifically, the approach combines topic modelling, multiclass classification and association rule mining; filtering unstructured textual data into a more structured picture of CX before distilling focused, granular regularities. In the case of results using undersampling on the test dataset only, a result of 65.85% was achieved for the social media dataset, 69.61% for the online review dataset, and 61.57% on the survey (B2B) dataset. In the case of results using undersampling on both the training and test-sets, a result of 66.87% was achieved for the social media dataset, 72.13% for the online review dataset, and 65.29% on the survey (B2B) dataset. Precision, recall and F-measure scores were sound and reasonably balanced across most classes, with the exception of some classes in the survey (B2B) datatset, owing to the relatively larger number of topics. Collectively these results give us certainty around the quality of the LDA procedure used here.

October 2018

Find out more about the Cambridge Service Alliance:
Linkedin Group: Cambridge Service Alliance
www.cambridgeservicealliance.org

The papers included in this series have been selected from a number of sources, in order to highlight the variety of service related research currently being undertaken within the Cambridge Service Alliance and more broadly within the University of Cambridge as a whole.

Textual Regularity Mining and Data-driven Customer Experience Analytics

Mohamed Zaki[1], Benjamin Lucas[2,3,1]

1. Cambridge Service Alliance, Institute for Manufacturing, University of Cambridge, Cambridge, United Kingdom.
2. Nottingham University Business School, University of Nottingham, Nottingham, United Kingdom.
3. N-LAB, University of Nottingham, Nottingham, United Kingdom

Acknowledgements: This research is funded by the Marketing Science Institute (MSI) Research Initiative on Customer Experience (Research Agreement: #4000004).

Introduction

Customer experience (CX) holistically embodies customer encounters with, and purchases of a firm's products or services, and the decision processes, and organization and consumer generated information flows that influence them, leading to present and future encounters and purchases (MSI, 2016) embedded within customer journeys, comprising firm-created touchpoints (Lemon and Verhoef, 2016). CX is thus critical to the design and marketing of services, and products, as well as management practice (Brakus, Schmitt and Zarantonello, 2009; Homburg, Jozić and Kuehnl 2015; Maklan, Antonetti, and Whitty, 2017; Weill and Woerner, 2018). In fact, research by Gartner finds that 89% of companies compete on a foundation of customer experience (Gartner, 2016). Moreover, unsatisfactory customer experiences result in around $83 billion in losses by US enterprises each year through abandoned purchases and defections (Forbes, 2013). This highlights the critical nature of understanding consumption experiences holistically, and by extension, the need to develop analytics solutions to aid in shaping future customer purchase decisions and experiences. The challenge is that CX, and the customer journeys in which they are embedded are complex; involving often intertwined dimensions, and extreme heterogeneity, as a function of individual characteristics and a vast array of contextual differences. This situation is further complicated by the fact that analyzing CX often involves textual customer recounts of their experiences (solicited or unprompted), further complicating the process of generating insights when differences in natural language are factored in.

This study addresses this challenge by investigating the use of theory-informed, data-driven methodologies for CX analytics. By approaching CX analytics in a data-driven manner, a picture more consistent with current CX theorizing, referring particularly to the intertwined nature of dimensions, can be obtained and packaged in a 'dashboard ready' manner. This can be summarized into four main contributions. Firstly, we summarize major developments in CX theorizing and embed the concept as a tool within a data-driven analytics toolkit. Secondly, we show how the intertwined nature of customer experience dimensions necessitates the use of a theory-informed, data-driven approach. Thirdly, we address the issue of emotional proxies and manifestations in CX analytics as part of this approach. This contribution involves letting the data *speak for itself*, and recognizing that emotions manifest in contextually defined descriptors of experiences, rather than via just "good" and "bad" words and expressions. Specifically, we argue why there should be no separation between vocabulary when evaluating the physical, cognitive, emotional, sensory, behavioral and social aspects of CX. Finally, we draw together different perspectives on text analytics; combining thematic extraction (Tirunillai and Tellis, 2014; Puranam, Narayan and Kadiyali, 2017), to summarize unstructured textual data, with a language pragmatics perspective (Abbasi et al., 2018) to distil granular regularities to provide highly-focused managerial insights that transcend CX dimensions.

Customer Experience and Customer Feedback

In general, soliciting customer feedback is recognized as having positive effects on marketer-consumer relationships (Bone et al., 2017). Functionally speaking, '*the voice of the consumer*' (Iacobucci, Ostrom and Grayson, 1995) has implications for marketing decisions, spanning incident resolution, PR and new product and service development. Thus sources of willingly provided post-hoc customer feedback offer marketers a valuable arena for better understanding consumption experiences with the goal of attracting and retaining customers. In particular, the analysis of unstructured and semi-structured web data, a key component of business analytics (Kune et al., 2016), is emerging as a means to generate CX insights for managers, owing to the abundance of customer recounts and reviews available through digital channels (e.g. online review platforms, social media), where customers actively share their experiences to allow such analysis in practice (Stephen et al., 2017). The difficulty for managers however, is how to navigate the complexity of scaled business analytics to measure customer experience in a commercially beneficial way. Thus the future of CX lies in translating the conceptualizations of CX developed within the marketing research discipline (e.g. Brakus, Schmitt and Zarantonello, 2009; De Keyser et al., 2015) into advanced measurement systems for deployment in diverse environments.

Along these lines, marketing researchers have produced a number of useful studies focused on using customer responses to, and recounts of their experiences to generate managerial relevant insights, and commercially relevant outcomes. Furthermore, researchers have called specifically for the use of text analytics and emerging analytics technologies to better measure experiences (Verhoef, Antonides and de Hoog 2004; Verhoef, Kooge and Walk 2016; Villarroel Ordenes et al., 2014). For example, Knox and Van Oest (2014) developed a model of recoveries and churn prevention following customer complaints, while Xiang et al. (2015) used text analytics to analyse customer reviews in hotels to understand the relationship between customer experience and satisfaction. The next steps however involve marketing pivoting as a discipline toward developments arising from the data analytics domain, with particular emphasis on data-driven marketing research (Rust and Huang, 2014; Ringel and Skiera 2016), and embracing new sources of digital marketing intelligence (Stephen and Lamberton, 2016).

Because recounts of customer experiences are deeper and more multifaceted than what can be captured by measuring volume and valence only (Archak, Ghose and Ipeirotis, 2011), researchers are beginning to embrace machine learning techniques to capture richer insights. For example, Tirunillai and Tellis (2014) used Latent Dirichlet Allocation (LDA) to measure latent satisfaction and quality dimensions. This focus on automatic data-driven "dimension" extraction is similar to that employed in research by Lee and Bradlow (2011) who automated product attribute identification for ranking and comparing brands, and Culotta and Cutler (2016) who measured brand perception estimates based on select focal attributes.

Customer Experience Frameworks

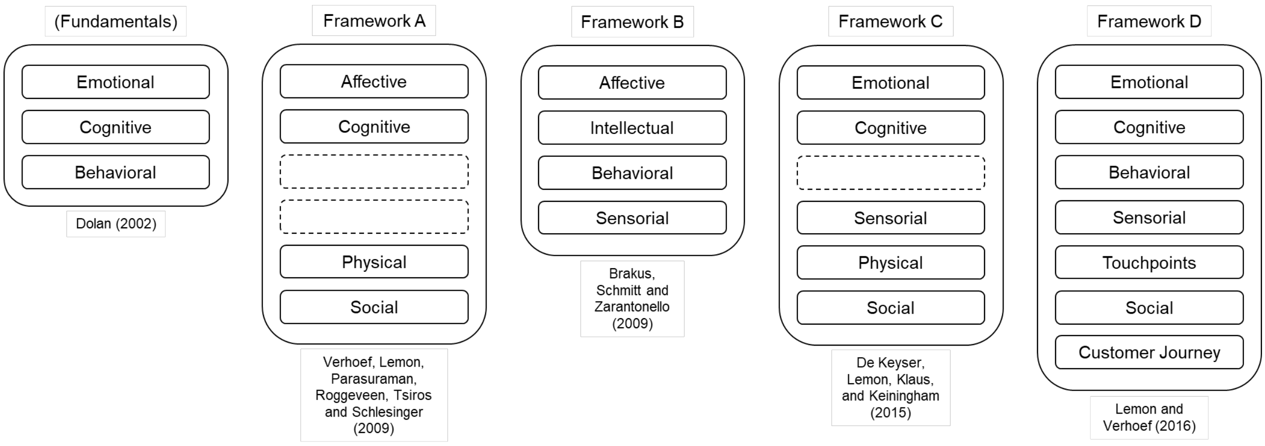
On the customer side, CX involves, different responses that the brand and its environment create, characterized by varying intensity, valence (e.g. positive and negative), spontaneity and duration (i.e. at every contact point between the customer and the organization) (Brakus, Schmitt and

Zarantonello, 2009; Grewal, Levy and Kumar, 2009). On the organization side, customer behavioral responses become identifiable and measurable factors or dimensions, each leading to different chains of effects within an interrelated system of CX dimensions. For example, Brakus, Schmitt and Zarantonello (2009) showed how different CX factors contribute to a higher order experience construct, and the subsequent relationships between this construct and outcomes such as loyalty.

Although experience as an abstract concept can be broken into somewhat hierarchical layers (basic experience, aware experience and conscious experience – as in Hoffman and Novak (2017) for instance), separated into material, versus purely 'experiential' components (Schmitt, Brakus and Zarantonello, 2015.), or different stages of the customer journey, as in Lemon and Verhoef (2016), the exact timing and nature of such phases varies dramatically between different product and service consumption experiences, owing to, as introduced, widely investigated differences between humans and their characteristics and a vast array of contextual factors. Although this complexity is captured in the evolution of CX theorizing toward increasingly more holistic perspectives, the intertwined nature of CX dimensions has far more fundamental roots.

For instance, Dolan (2002) asserts that emotion, cognition and behaviour are inherently intertwined concepts and processes. These can be argued to be the 'building blocks' of customer experience; a necessary and universally-present set of dimensions, underlying the customer journey. Emotions alone are considered complex, intertwined psychological and physiological states that influence reasoning and judgement around events and actions in response to the physical and social world, influencing most dimensions of cognition (Dolan, 2002). This view is reflected in the evolution of CX frameworks. In Figure 1, we provide an overview of major developments, starting with Dolan (2002) as a fundamental basis. We then consider Verhoef et al. (2009), who view CX from a largely physical environment perspective, exploring the contextual dimensions; physical and social, within which behaviour is contained alongside affective and cognitive dimensions. This is followed by Brakus, Schmitt and Zarantonello (2009), from a brand experience perspective, and De Keyser et al's (2015) holistic view, and concluding with Lemon and Verhoef (2016), who introduced an emphasis on touchpoints, and the reinforced the link between customer experience and customer journey.

Figure 1: Major Developments in CX Theorizing (Chronological)



Brakus, Schmitt and Zarantonello's (2009) brand experience dimensions are; (1) sensory, (2) affective, (3) behavioral, and (4) intellectual, capturing human senses, human emotions, human behaviour and human cognition respectively. This captures both the means of receiving inputs, as well as subsequent responses to CX stimuli. As with the aforementioned translation of CX behavioral factors into measurement points for organizations to use in managing CX, these antecedent input factors perform the same important role. For instance, in related work, Schmitt (2011) discuss how in a consumption setting, experiential attributes can shorten the link between trigger and response, by inducing heightened processing fluency. Functional attributes by comparison, tend to be evaluated in a more deliberate manner by consumers.

Relatedly, De Keyser et al's (2015) customer experience framework includes *cognitive*, *emotional*, *physical*, *sensorial* and *social* factors (De Keyser et al., 2015). Within each of these factors, a number of sub-factors and alternative scales of measurement are theoretically possible. For example the emotional factor, though important, loses its value if only interpreted as capturing a simple bi-polar representation of emotional response arising from a customer experience. As with Brakus, Schmitt and Zarantonello's (2009) framework, it follows that the non-emotional CX factors are likely subject to shaping or influence from the emotional factor. Schmitt (2011) explicitly discusses this, giving the example of post purchase dissonance following hedonic consumption (i.e. where the consumer experience a mixed negative and positive emotional state), and also discusses the potential for multiplicative interaction between humanic and environmental sensory clues or signals received by the customer during contact with an organization (e.g. where one negative aspect of an experience compromises other factors, and thus the overall experience).

Lemon and Verhoef (2016) defined customer experience more recently as a holistic concept, encompassing emotional, cognitive, behavioural, sensorial and social dimensions, operating within and around a customer journey, comprising 'touchpoints'; any point of contact between the customer and the firm for any reason, at any stage. Returning full-circle to the fundamentally intertwined nature of CX dimensions, we thus propose a new CX framework to guide this study, informed by the five perspectives in Figure 1 involving a holistic picture of CX; encompassing the overall experience and highlights within the customer journey, then subdividing into four specific, but intertwined dimensions. These are (1) *Objects and Entities*; informed by the physical CX dimension and recent thinking around touchpoints, (2) *Thoughts, Feelings and Sensations*; combining the cognitive, emotional and sensory CX dimensions, (3) *Behaviors*; the behavioral CX dimension, and finally, we focus on (4) *Ownership and Relation*, a different take on the social CX dimension concerning attribution of experience and relations between social actors involved in the experience.

Emotions in Customer Experience

Emotion in the context of customer experience can be thought of as an experienced state, occurring during any stage of the customer journey, from decision making, to purchase, to consumption, and operating as a product of any given individual and number of aspects of the experience, thus manifesting in a number of different, explicit and non-explicit ways when recounted (i.e. expressed in literal or implicit terms). Research questions around emotions in CX tend to zoom in on operational or experience-engineering foci. For instance, Kumar et al. (2014) emphasize emotional reactions in their operationalization of CX, encompassing satisfaction, recovered service failures, and unrecovered service failures. Focusing on positive emotions, Chun, Diehl and MacInnis (2017) show how encouraging consumers to savor an upcoming (yet to be consumed) experience can heighten actual experience, and improve positive recollection of the experience after the fact. Sentiment, on the other hand, is a term usually used to denote the levels of valence and arousal (Jiang et al., 2014) present in a recount of a focal consumption experience.

In terms of the operationalization and measurement of CX, emotion is one of the more difficult dimensions. As a result, optimizing the operationalization and measurement of emotion in the context of consumption experience has been the subject of extensive academic debate for some time (Richins, 1997). In terms of manifestation, this pertains mostly to the challenge of isolating non-explicit (implicit) sentiment (see for example: Van de Kauter, Breesch and Hoste, 2015). Explicit sentiment on the other hand, involves the direct conveying of an emotion in a literal, or close to literal manner. Given the complexity of human thought and language however, this cannot always be relied upon, with humans making use of more complex linguistic devices, such as figurative language, sarcasm and different types of potentially abstract humor. Another complicating factor is the role of so-called emotional complexity, and the fact that human experiences are generally recognized as being characterized by mixed emotional states, even when one in particular seems dominant or defining (Aaker, Drolet and Griffin, 2008).

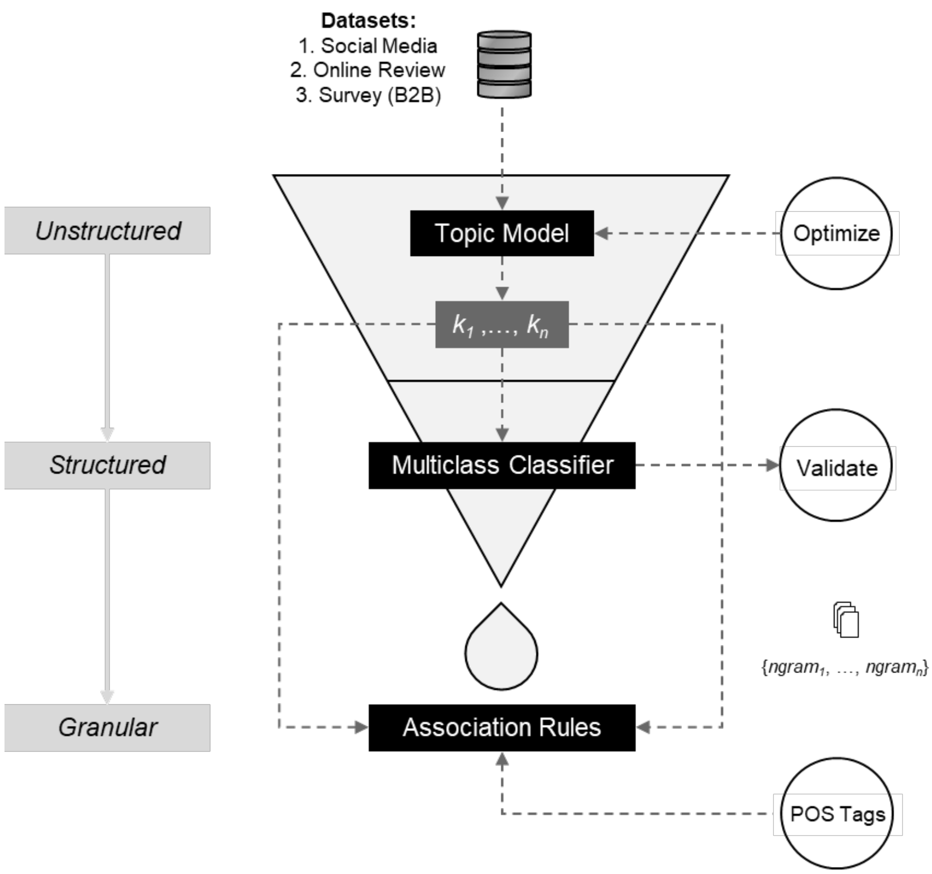
According to one view of *emotional complexity theory*, emotional states can be characterized more precisely by *emotional dialecticism* (simultaneous negative and positive feelings) and *emotional differentiation* (granular emotional experience with emotional states comprising varied discrete negative and positive states) (Grossmann, Huynh and Ellsworth,2015). Another view of emotional complexity corresponds with our view of the extreme heterogeneity of CX being accounted for by extreme potential differences between humans and their characteristics and a vast array of contextual factors. For example, Lindquist and Barrett (2008) and Kang and Shaver (2004) focus their perspectives more on characterizations of individuals as emotionally complex, on the one hand, concerning the ability to identify and experience certain emotions, and on the other, capturing the breadth of emotions an individual experiences and how they experience them. Both views however raise questions around simplifying sentiment analysis into ‘best fit’ categories, or even simplified polarity when exploring the emotions in a context such as CX analytics.

More specifically, while CX research tends toward conceptualising emotions as a core dimension of a multidimensional customer experience concept (Lemon and Verhoef, 2016), in practice, emotional proxies, or sentiment measures, are commonly used as KPIs in and of themselves (Schweidel and Moe, 2014). This latter perspective is logical, given its intuitive nature and accessibility, and the fact that thematic and topic extraction techniques (in the context of digital textual data) can run the risk of leaving too much to the imagination, rather than delivering precise managerial insights. However, although emotion is multi-faceted, arousal levels (Jiang et al., 2014) and convenience, along with the nature of certain customer feedback and sharing platforms (see Schweidel and Moe, 2014 for an extended discussion), may be factors in explaining why customers might only choose to spend their time sharing their experiences along the lines of one focal emotion (e.g. a complaint for a particularly bad experience, or a compliment for a particularly good experience), rather than engaging in recounts that exhibit greater emotional dialecticism. We thus address the issue of emotional complexity, comprising use of language (manifest), context heterogeneity (experienced), and individual heterogeneity by taking a data-driven approach, which does not actually separate emotion from its embedded context. Our underlying theoretical rationale is discussed next.

Data-driven Customer Experience Analytics

Here, we argue that given the present conceptualization of CX as a holistic rather than multidimensional concept (per-se), any attempt to "operationalize" CX should do so in a way that captures the intertwined nature of CX dimensions, and provides a bigger-picture summary of the customer journey and its constituent touchpoints. As introduced, any such approach should also be able to derive regularity from the chaos of natural language, accommodating extremely heterogeneous experiences and customer recount styles. In an analytics context, textual regularities and patterns that can be drawn from data must be strong enough to provide managerially useful insights, and must be able to be validated. Once unstructured textual data has been organized into a structure, and once this structure has been validated however, a foundation is then created for distilling more granular and focused regularities that represent managerially useful insights which may have otherwise been overlooked. Taking all this into account, we start by representing the broader picture of customer experience based around unsupervised machine learning in the form of a topic model. This approach is based on global corpus-level common and recurrent patterns of text, which are then validated using a supervised model. We use this as a foundation for "exploring experiences", for which we employ word tokenization (part-of-speech (POS) tagging), ngram condensing, and association rules at a local topic and document-level, in search of patterns which capture CX dimensions. The advantage here, is that vocabulary is not split and forced to load onto specific CX dimensions, but rather that textual regularities are derived to create insights around CX dimensions, given the certainty of; (a) an accurately decomposed unstructured textual dataset, (b) the use of literal markers of relevant terms, and, (c) retaining lexical and grammatical regularities in the textual dataset in the form of co-occurring word patterns. Our approach is summarized in Figure 2 and detailed in the proceeding sections of this working paper.

Figure 2: Methodological Framework for Data-driven CX Analytics



Datasets

We make use of three separate primary datasets. The first is a social media customer feedback dataset drawn from a major UK-based retailer (Facebook posts by customers), comprising 1700 posts. The second is an online review dataset for a UK-based not-for-profit patient experience review platform, comprising 3200 reviews from different health service providers. The third is the textual component of a nation-wide survey from a UK-based heavy industry B2B service provider, comprising 2800 responses. The textual component of this survey asked respondents to submit any additional comments they had after completing a customer satisfaction survey.

Method

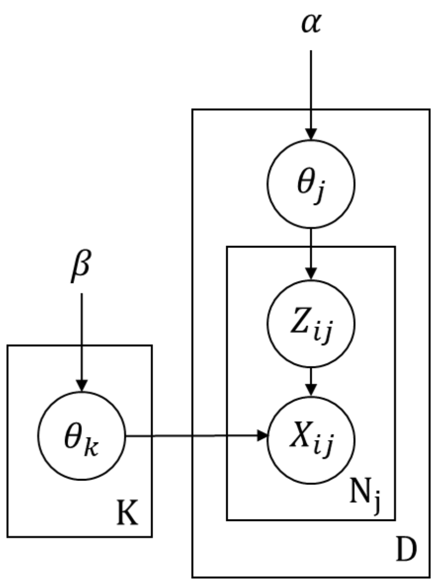
The method stages are discussed as follows. Additionally, we include a brief methodological appendix which is also introduced here.

Topic Model Optimization

The Latent Dirichlet Allocation (LDA) procedure (Blei, Ng and Jordan, 2003) assumes a theory of meaning deriving from co-occurrence patterns within and across documents (within a corpus) (Farrell, 2016), and assumes observable word co-occurrence patterns are determined according to latent variables (Lu, Wei and Hsiao, 2016). That is, LDA can be used to locate topics within a corpus and their constituent representative terms (Curme et al., 2014). In basic terms, this allows researchers to identify “linguistic regularities” associated with certain objects, activities and concepts (Roy et al., 2015). The technique has been proven highly useful for marketing research applications (Tirunillai and Tellis, 2014; Puranam, Narayan and Kadiyali, 2017).

In this section we detail the process of parameter optimization to optimize the performance of the Latent Dircihlet Allocation model on our corpora. LDA is speculated to encounter limitations on smaller texts, and the literature proposes a number of possible solutions. However, as in the case of all probabilistic techniques, the end goal is the quality and usefulness of the end information produced. Further, to counter this, we employ a rigourous optimization procedure, and by nature of the focal measures reported here, demonstrate the appropriate nature of this method in this context. Additionally, we sought to work with existing tools from the marketing research literature (Tirunillai and Tellis, 2014; Puranam, Narayan and Kadiyali, 2017), to increase the transparency and accessability of our approach. The LDA algorithm is summarized in the below plate notation (Figure 3), from Newman et al. (2009).

Figure 3: Plate Notation for Latent Dirichlet Allocation



In Figure 3; K = number of topics, N = number of words, D = number of documents, *X* = observed word, *Z* = assigned topic, *i* = ith word, *j* = jth document, *k* = a topic, θk = word probability given *k*, and θj = topic probability given *j*. Here, the parameters α, β and K were those used in optimization. We employed optimization based on hill-climb loops, given computational efficiency considerations (see appendix). α and β were given a search ranges between 0.001 and 1, with steps of 0.001, K was given a minimum value of 2 in all cases and a maximum value that sought to contain the maximum number of documents per topic (a consideration for the following stage of analysis). In addition to this search criteria, the individual models at each search interation were computed at 2000 iterations.

The results of the optimization procedure are represented in Figures 4, 5 and 6 respectively. The plots show the changes in parameter combinations over the number of combinations (horizontal axis), on a normalised scale (vertical axis). From left to right, these plots show the relatively small step combinations made to the alpha and beta parameters, as the number of topics is also adjusted. Log-likelihood (normalized) reaches a maximal value on the righthand side of the plots.

Figure 4: Latent Dirchlet Model Optimization – Social Media Dataset

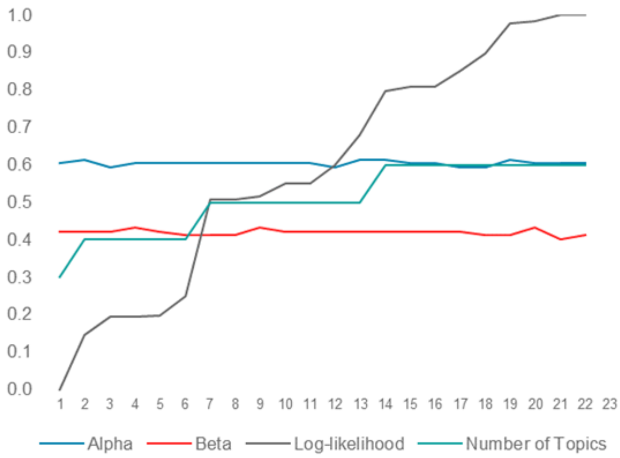


Figure 5: Latent Dirchlet Model Optimization – Online Review Dataset

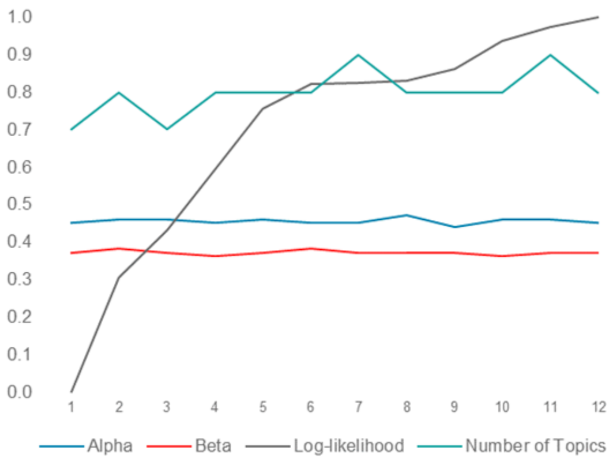
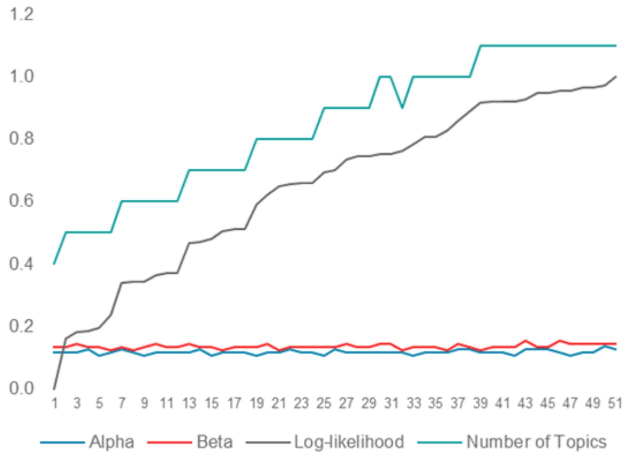
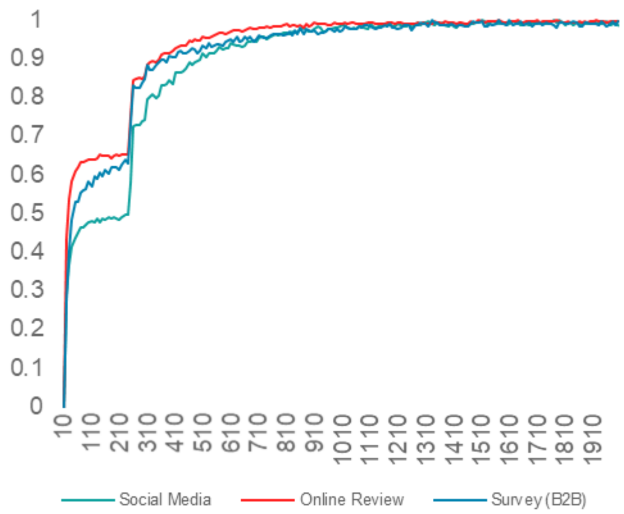


Figure 6: Latent Dirchlet Model Optimization – Survey (B2B) Dataset



The optimal parameters were then used to compute the LDA models for each dataset as an additional performance check. Log-likelihood was maximized at 6 topics, with α 0.605 and β 0.412 for the social media dataset, 8 topics, with α 0.452 and β 0.373 for the online review dataset, and 11 topics, with α 0.125 and β 0.142 for the survey (B2B) dataset. Figure 7 shows results for maximized log-likelihood attainment (normalized, vertical axis) over 2000 iterations (horizontal axis) for each re-computed model. Model behavior was similar across datasets, with all models beginning to palteau around 800 iterations.

Figure 7: Latent Dirichlet Behavior Over Iterations



Topic Labels and Mappings

Next, we explore the optimized topic solutions for each dataset. Normally, the top-weighted terms for topics (in our case, the top five), are used as a definition for those topics. Here, we use these terms as 'labels' for our topics. The left side of the following Tables (1 to 3) shows these labels, the right side, shows the average topic mapping quality achieved for each retrieved topic. Mapping values, which range between 0 and 1 are also mostly sound and consistent across topics. Identifying information such as brand names and locations have been removed from the reported results only as a data usage best-practice measure.

Table 1: Retrieved Topics and Mappings – Social Media Dataset

	Topic Labels	k1	k2	k3	k4	k5	k6
k1	(loyalty card promo), (loyalty card), car, customers, store	0.85	0.02	0.04	0.04	0.02	0.02
k2	bags, food, packaging, plastic, please	0.02	0.84	0.04	0.07	0.02	0.02
k3	customer, delivery, online, service, shopping	0.02	0.02	0.84	0.05	0.03	0.03
k4	bought, chicken, date, found, store	0.01	0.02	0.05	0.85	0.04	0.03
k5	free, please, range, stock, store	0.01	0.01	0.05	0.08	0.83	0.02
k6	customer, service, shopping, staff, store	0.03	0.02	0.05	0.04	0.02	0.85

Table 2: Retrieved Topics and Mappings – Online Review Dataset

	Topic Labels	k1	k2	k3	k4	k5	k6	k7	k8
k1	feel, health, help, mental, support	0.71	0.03	0.04	0.06	0.04	0.04	0.05	0.03
k2	baby, centre, children, staff, start	0.03	0.84	0.02	0.02	0.01	0.04	0.02	0.02
k3	involvement, people, service, staff, trust	0.05	0.03	0.72	0.05	0.03	0.04	0.06	0.02
k4	care, help, staff, team, thank	0.05	0.02	0.03	0.77	0.03	0.03	0.03	0.03
k5	appointment, hospital, service, time, told	0.06	0.02	0.03	0.06	0.71	0.03	0.03	0.06
k6	confidence, course, help, helped, people	0.04	0.07	0.04	0.04	0.02	0.72	0.02	0.04
k7	food, patients, staff, time, ward	0.06	0.03	0.05	0.05	0.03	0.03	0.72	0.03
k8	clinic, dental, dentist, service, staff	0.02	0.04	0.02	0.03	0.03	0.02	0.02	0.84

Table 3: Retrieved Topics and Mappings – Survey (B2B) Dataset

	Topic Labels	k1	k2	k3	k4	k5	k6	k7	k8	k9	k10	k11
k1	hours, machine, service, time, wrong	0.81	0.00	0.01	0.02	0.01	0.02	0.01	0.03	0.03	0.02	0.04
k2	communication, engineer, relationship, service, spot	0.01	0.86	0.01	0.02	0.01	0.02	0.01	0.01	0.02	0.02	0.01
k3	bit, cost, expensive, fine, little	0.01	0.01	0.85	0.02	0.01	0.02	0.01	0.02	0.02	0.02	0.01
k4	brilliant, excellent, fault, helpful, service	0.01	0.01	0.01	0.84	0.01	0.01	0.01	0.02	0.03	0.02	0.02
k5	guy, issues, nice, online, store	0.01	0.01	0.02	0.02	0.85	0.02	0.01	0.02	0.03	0.02	0.01
k6	cheaper, expensive, price, prices, reduce	0.01	0.01	0.02	0.01	0.01	0.87	0.01	0.01	0.02	0.01	0.02
k7	(brand name), day, deal, issue, machine	0.01	0.01	0.01	0.02	0.01	0.01	0.83	0.02	0.03	0.02	0.02
k8	available, day, days, stock, wait	0.01	0.01	0.02	0.01	0.01	0.02	0.02	0.84	0.03	0.02	0.01
k9	delivery, happy, pretty, satisfied, service	0.02	0.03	0.05	0.04	0.02	0.06	0.02	0.03	0.64	0.04	0.04
k10	call, customer, improve, phone, time	0.01	0.02	0.02	0.01	0.01	0.02	0.01	0.03	0.03	0.81	0.02
k11	doing, guys, job, moment, time	0.01	0.02	0.02	0.02	0.02	0.01	0.01	0.02	0.03	0.03	0.82

For all three datasets, the topic mapping scores were relatively consistent, in terms of mappings for documents to topics, and overlap into other topics. For the social media dataset, the retrieved topics reveal thematic groupings around travel to the store and a loyalty card program touchpoint (k1), a theme around plastic packaging and bags (k2), and a theme around an online shopping touchpoint and deliveries (k3). Themes were also retrieved around purchasing fresh food and expiry dates (k4), around free range products (k5), and around the customer service staff touchpoint (k6).

For the online review dataset, the retrieved topics reveal thematic groupings around mental health support (k1), infant care (k2), the service staff touchpoint (k3 and k4), appointment times and service (k5), reactions to a patient course (k6), staff and food touchpoints in a hospital context (k7), and service and staff in a dental clinic context (k8). For the survey (B2B) dataset, the retrieved topics reveal thematic groupings around service hours and timing (k1), communication and relationships with a staff related touchpoint (k2), pricing (k3, k6), satisfaction with customer service (k4, k9), online store touchpoints (k5), deals and issues around a particular supplier's brand (k7), stock availability and wait time (k8), the telephone touchpoint and timing (k10), general satisfaction and general comments on timing (k11).

Validation

To validate the performance of our LDA procedure, we predicted topics as a target using bit vector representations of the original documents in our corpora. This representation was chosen to reduce the document vocabulary into its simplest possible form, to assess the impact on topic mapping. We used a Naïve Bayes multi-class classifier for this task. The notation (Mitchell, 2005) in (1) below addresses the problem; Given $X^{new\_instance} = \langle X_1 \ldots X_n \rangle$ calculate the probability of Y taking on a given value, where $P(Y)$ and $P(X_i|\mathrm{Y})$ are estimated distributions (training dataset), and where X = attributes, Y = variables, y = values, and where X_i is assumed to be conditionally independent from the other $X_k s$ given Y, as well as independent from each subset of other $X_k s$ given Y. The notation in (2) below is the same, but for determining the most probable value.

Y: Any given value:

$$P(Y = y_k|X_1 \ldots X_n) = \frac{\mathrm{P}(\mathrm{Y} = y_k) \prod_i P\,(X_i|Y = y_k)}{\sum_j \mathrm{P}(\mathrm{Y} = y_j) \prod_i P(X_i|Y = y_j)} \quad (1)$$

Y: Most probable value:

$$Y \leftarrow \mathrm{argmax}_{yk} \frac{\mathrm{P}(\mathrm{Y} = y_k) \prod_i P(X_i|\mathrm{Y} = y_k)}{\sum_j \mathrm{P}(\mathrm{Y} = y_j) \prod_i P(X_i|Y = y_j)} \quad (2)$$

Here, we employed Naïve Bayes as a one-against-all multi-class classifier, such that the prediction target is identiyfing one given topic against all others. Tables 4, 5 and 6 show validation results for all extracted topics, across all datasets usign multi-class classification based on Naïve Bayes. Figure 4 shows classification results using undersmapling on the test-set only. Figure 5 shows results using undersampling on both the training and testing sets (undersampled indenpendently). 50/50 partitioning was used to conserve data usage. These figures show model accuracy under the dataset header, as well as results for precision and recall measures, sorted by F-measure. In the case of results using undersampling on the test-set only, a result of 65.85% was achieved for the social media dataset, 69.61% for the online review dataset, and 61.57% on the survey (B2B) dataset. In the case of results using undersampling on both the training and test-sets, a result of 66.87% was achieved for the social media dataset, 72.13% for the online review dataset, and 65.29% on the survey (B2B) dataset. Precision, recall and F-measure scores were sound and reasonably balanced across most classes, with the exception of some classes in the survey (B2B) datatset, owing to the relatively larger number of topics. Collectively these results give us certainty around the quality of the LDA procedure used here. Additional notes are provided in the appendix.

Table 4: Validation Results for Extracted Topics – Social Media Dataset

(USampTest-set Only)

Social Media (65.85%)	Precision	Recall	F Measure
k3	0.65	0.76	0.70
k6	0.74	0.65	0.69
k2	0.80	0.60	0.69
k1	0.90	0.54	0.67
k5	0.67	0.60	0.63
k4	0.48	0.82	0.60

(Usamp Train and Test)

Social Media (66.87%)	Precision	Recall	F Measure
k2	0.81	0.72	0.76
k3	0.72	0.74	0.73
k6	0.73	0.65	0.68
k1	0.67	0.61	0.64
k5	0.55	0.68	0.61
k4	0.59	0.61	0.60

Table 5: Validation Results for Extracted Topics – Online Review Dataset

(USampTest-set Only)

Online Review (69.61%)	Precision	Recall	F Measure
k2	0.77	0.82	0.79
k3	0.69	0.81	0.75
k8	0.68	0.83	0.74
k4	0.73	0.67	0.70
k7	0.80	0.59	0.68
k6	0.78	0.59	0.67
k1	0.52	0.75	0.62
k5	0.74	0.52	0.61

(Usamp Train and Test)

Online Review (72.13%)	Precision	Recall	F Measure
k2	0.88	0.84	0.86
k8	0.72	0.83	0.77
k7	0.82	0.71	0.76
k3	0.68	0.84	0.75
k6	0.73	0.66	0.70
k5	0.76	0.59	0.66
k4	0.73	0.57	0.64
k1	0.55	0.72	0.62

Table 6: Validation Results for Extracted Topics – Survey (B2B) Dataset

(USampTest-set Only)

Survey (B2B) (61.57%)	Precision	Recall	F Measure
k6	0.75	0.92	0.83
k3	0.79	0.70	0.74
k4	0.71	0.67	0.69
k8	0.66	0.68	0.67
k11	0.66	0.67	0.66
k2	0.81	0.52	0.63
k10	0.64	0.59	0.61
k1	0.63	0.44	0.52
k7	0.68	0.39	0.50
k9	0.34	0.89	0.49
k5	0.67	0.30	0.42

(Usamp Train and Test)

Survey (B2B) (65.29%)	Precision	Recall	F Measure
k6	0.87	0.91	0.89
k3	0.74	0.83	0.79
k9	0.70	0.77	0.73
k8	0.66	0.73	0.69
k2	0.73	0.65	0.69
k11	0.69	0.67	0.68
k1	0.67	0.59	0.63
k4	0.75	0.50	0.60
k10	0.52	0.52	0.52
k7	0.41	0.61	0.49
k5	0.55	0.41	0.47

To further illustrate prediction quality, we visualized the confusion matrices (normalized values) for each of the validation classifiers. These mostly show high quality predictions for each of the target classes across datasets (Figure 8). These can also be contrasted with the topic mapping quality tables (Tables 1b, 2b and 3b) to shed light on some of the 'more difficult to predict' topics.

Figure 8: Confusion Matrices for Multi-class Topic Classification

Social Media - USampTest-set Only

	k3	k4	k5	k6	k1	k2
k3	0.62	0.07	0.03	0.03	0.03	0.04
k4	0.07	0.67	0.04	0.03	0	0.01
k5	0.03	0.23	0.49	0.05	0	0.02
k6	0.11	0.11	0.03	0.53	0.02	0.02
k1	0.08	0.13	0.08	0.06	0.44	0.03
k2	0.05	0.2	0.06	0.02	0	0.49

Social Media - Usamp Train and Test

	k3	k4	k5	k6	k1	k2
k3	0.61	0.04	0.06	0.03	0.05	0.03
k4	0.05	0.5	0.14	0.06	0.04	0.03
k5	0.02	0.14	0.56	0.04	0.03	0.03
k6	0.08	0.04	0.06	0.53	0.09	0.02
k1	0.06	0.03	0.14	0.06	0.5	0.03
k2	0.03	0.1	0.05	0.01	0.04	0.59

Online Review - USampTest-set Only

	k7	k8	k6	k2	k4	k1	k3	k5
k7	0.64	0.04	0.01	0.04	0.05	0.16	0.12	0.03
k8	0	0.9	0.03	0.04	0.04	0.02	0.03	0.03
k6	0.01	0.06	0.64	0.09	0.04	0.11	0.11	0.03
k2	0	0.04	0.06	0.89	0.01	0.05	0.02	0.02
k4	0.04	0.08	0.02	0.03	0.73	0.15	0.02	0.02
k1	0.03	0.05	0.03	0.03	0.05	0.82	0.03	0.05
k3	0.07	0.01	0.02	0.03	0.02	0.04	0.88	0.02
k5	0.01	0.15	0.01	0.01	0.06	0.22	0.06	0.57

Online Review - Usamp Train and Test

	k7	k8	k6	k2	k4	k1	k3	k5
k7	0.77	0.05	0	0.02	0.04	0.07	0.13	0.01
k8	0.01	0.91	0.06	0.01	0.03	0.01	0.02	0.04
k6	0.03	0.04	0.72	0.04	0.02	0.12	0.08	0.04
k2	0.02	0.05	0.05	0.92	0.02	0.02	0.01	0
k4	0.04	0.09	0.07	0.02	0.62	0.13	0.07	0.05
k1	0.02	0.04	0.03	0.03	0.08	0.79	0.05	0.05
k3	0.04	0.03	0.03	0	0.02	0.04	0.92	0.01
k5	0.01	0.06	0.02	0.01	0.02	0.26	0.07	0.64

Survey (B2B) - USampTest-set Only

	k9	k11	k6	k3	k8	k10	k4	k7	k2	k1	k5
k9	0.51	0	0.02	0.01	0.02	0.05	0	0.02	0	0.02	0.01
k11	0.04	0.44	0.01	0.01	0.03	0.01	0.01	0.05	0.01	0.02	0.03
k6	0	0	0.6	0.04	0	0.01	0	0.01	0	0	0
k3	0	0	0.01	0.55	0.01	0.03	0.01	0.01	0.01	0.01	0.02
k8	0	0.01	0	0.02	0.48	0.08	0	0.04	0.03	0	0
k10	0.01	0.04	0	0.02	0.02	0.34	0	0.1	0.06	0.02	0.05
k4	0.07	0.06	0	0	0.03	0.02	0.33	0.09	0.01	0.04	0.01
k7	0.02	0.03	0.02	0.02	0.03	0.05	0.02	0.4	0.02	0.02	0.03
k2	0.04	0.02	0	0.02	0.02	0	0.04	0.06	0.43	0.01	0.02
k1	0.01	0.01	0.02	0	0.01	0.04	0.02	0.09	0.02	0.39	0.05
k5	0.03	0.03	0.01	0.05	0.08	0.02	0.01	0.11	0	0.05	0.27

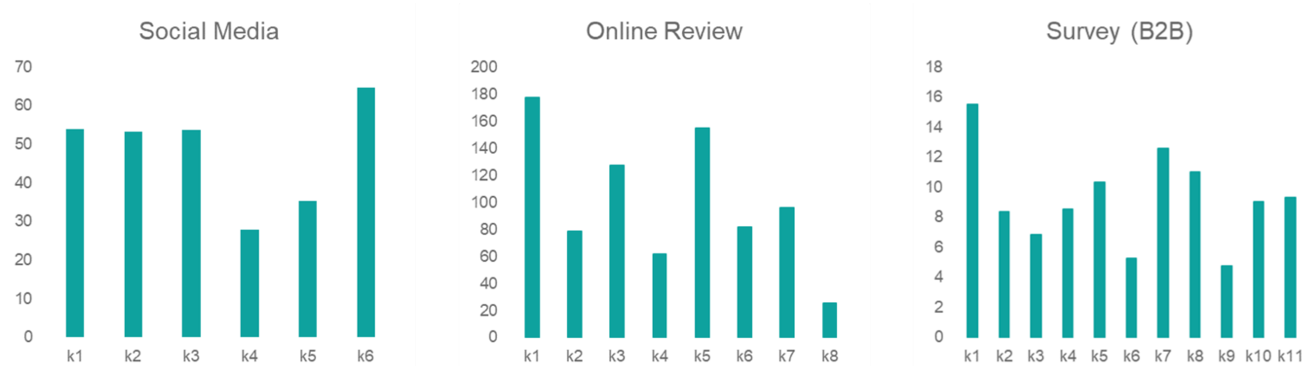
Survey (B2B) - Usamp Train and Test

	k9	k11	k6	k3	k8	k10	k4	k7	k2	k1	k5
k9	0.59	0	0.01	0.02	0	0	0.03	0	0	0.01	0
k11	0.12	0.44	0.02	0	0.03	0.01	0.01	0.02	0	0.01	0
k6	0.02	0	0.61	0	0	0.01	0	0.01	0	0.01	0
k3	0.05	0.02	0.07	0.46	0	0.04	0.01	0	0	0.01	0
k8	0.05	0.05	0	0	0.45	0.06	0	0.01	0.02	0.01	0.01
k10	0.09	0.05	0.01	0.02	0.02	0.39	0.01	0.03	0.03	0.01	0
k4	0.16	0.02	0	0	0	0.01	0.44	0.01	0.01	0.01	0
k7	0.11	0.04	0.05	0.01	0.04	0.05	0.03	0.26	0.02	0.02	0.03
k2	0.15	0.02	0	0.01	0.02	0	0.06	0.01	0.34	0.03	0.02
k1	0.18	0.02	0.02	0.01	0.05	0.02	0.02	0.01	0	0.29	0.04
k5	0.21	0.01	0.02	0.05	0.07	0.02	0.01	0.02	0	0.05	0.2

Exploring Experiences

In this section, we conduct a language pragmatics-based (Abbasi et al., 2018) informational exploration of customer experiences. Having gained a general free-vocabulary summary of experiences in the previous stages, we now turn our attention to distilling more granular insights pertaining to textual regularities representing the aforementioned *Objects and Entities*, *Thoughts, Feelings and Sensations*, *Behaviors* and *Ownership and Relation* CX dimensions. First, we take the retrieved topics for each dataset and apply tokenization in the form of POS tagging to all documents, using Apache OpenNLP (Apache OpenNLP, 2018), and the Penn Treebank (Penn Treebank, 2018). This approach allows us to capture a more literal representation of CX dimensions. Regarding the specific tags used; (1) Objects and Entities, employs the tags “NN” and “NNS” (nouns, singular and plural, excluding proper nouns). (2) Thoughts, Feelings and Sensations, employs the tags “JJ” (adjective), “JJR” (comparative adjective), “JJS” (superlative adjective), “RB” (adverb), “RBR” (comparative adverb), “RBS” (superlative adverb). (3) Behaviors, employs the verb tags “VB” (verb), “VBD” (past tense verb), “VBG” (present participle / gerund) and “VBN” (past participle). Finally, (4) Ownership / Relation employs the PRP$ (possessive pronouns) tag, and the PRP (personal pronouns) tag. Next, we filter terms to include only those specified in our framework (Figure 2). We then represent the remaining terms as ngrams. Doing so condenses the possible combinations of words, and eliminates the possibility of non-semantic rulesets which are not the focus here. We use 2grams (word pairs) for the social media and survey (B2B) datasets as they contain fewer words per document. We use 3grams for the online review dataset as they are generally longer documents. The average number of ngrams per topic is summarized in Figure 9.

Figure 9: Average Number of Ngrams per Document



Association rules are then derived for the tagged ngram documents as sets. Rule extraction and metrics are summarized in (3), adapted from D'Angelo, Rampone and Palmieri (2017). Confidence gives the percentage of documents containing *antecedent* that also contain *consequence*. Support gives the percentage proportion of documents containing *antecedent* ∪*consequence*. Lift for $antecedent \Rightarrow consequence$ is $\frac{\text{confidence}}{\text{expected confidence}}$ (See: Park et al., 2014). Results are presented in Tables 7, 8 and 9. Identifying information such as brand names and locations have been removed from the reported results only as a data usage best-practice measure. Additionally, the procedure used here treats contractions (e.g. "didn't") as two separate terms (e.g. "did" and "nt" "[not]").

Corpus (collections of posts, reviews or surveys):

$$corpus = \{document_1, document_2, \ldots, document_n\}$$

Contains:

$$document = \{ngram_1, ngram_2, \ldots, ngram_n\} \tag{3}$$

With a rule defined as:

$$antecedent \Rightarrow consequence$$

Where:

$$antecedent, consequence \subseteq corpus \quad \text{-and-} \quad antecedent \cap consequence = \emptyset$$

$$\text{-and-} \quad confidence > 0.60$$

In the social media dataset, the example rule from topic k1 highlights a remark made about "your loyal customers", directed at the brand, while the example rules from topic k2 and k5 highlight mentions of local branches of the retailer. The example rule from topic k3 highlights an issue around telephoning a customer service line, and in topic k4, the expression "was looking forward [to]", is indicative of a disappointing experience, separate from the direct interaction with the retailer. The rule "my […] year old" in topic k6 reflects remarks by a number of parents with young children, referring to different customer service interactions.

Table 7: Example Rules - Social Media Dataset

	Antecedent:	Consequence:	Confidence (%)	Set Support (%)	Lift
k1	loyal_customers	your_loyal	100%	1.84%	54.33
k2	local_(store brand name)	my_local	100%	3.59%	20.88
k3	called_customer_service	i_called_customer	100%	1.35%	74.25
k4	looking_forward	was_looking	62.50%	0.80%	77.88
k5	local_(store brand name)	my_local	100%	1.71%	19.47
k6	my_year, customer_service	year_old	100%	2.09%	11.94

In the survey (B2B) dataset, the example rule from topic k1 “I mentioned previously” highlights a point the customer has already communicated to the organisation. Example rules from topics k2, k10 and k11, highlight different aspects of interactions with the company (“good working relationship”, “response time” and “[the] service engineer kept me informed”). Example rules from topics k3, k4, k7 and k9 highlight the prominence of relatively explicit expressions of satisfaction (“everything was fine”, “they [were] very helpful”, “they [were] very good”, “everything was perfect”), while the example rule in topic k5 denoting the expression “there have been issues” is reflective of possible problems. The example rules from topics k6 and k8 respectively (“make it cheaper” and “stock more parts”) highlight issues around pricing and availability of physical aspects of the service provision.

We provide additional example rules for the online review dataset given that the constituent documents were generally longer, meaning there were more ngrams, and thus more opportunities to explore the rulesets. Here, we show some rule examples where similar expressions appear on both sides of the rule to further reinforce the existance of these text combinations.

Table 8: Example Rules - Survey (B2B) Dataset

	Antecedent:	Consequence:	Confidence (%)	Set Support (%)	Lift
k1	mentioned_previously	i_mentioned	100%	2.24%	26.80
k2	working_relationship	good_working	60.00%	1.60%	37.60
k3	everything_was	was_fine	93.80%	5.58%	10.51
k4	they_very	very_helpful	71.40%	1.92%	5.33
k5	been_issues, have_been	there_have	100%	1.50%	66.50
k6	it_cheaper	make_it	88.90%	4.32%	17.31
k7	very_good	they_very	80.00%	2.92%	27.40
k8	parts_stock	more_parts	80.00%	1.97%	18.04
k9	was_perfect	everything_was	62.50%	0.75%	23.23
k10	time_be	response_time	66.70%	0.99%	26.93
k11	kept_me, me_informed	service_engineer	66.70%	0.89%	75.00

Viewed from the perspective of the example rules in Table 9, topic k1 shows textual regularities pertaining to the personal circumstances of review writers (e.g.”I lost my job”), their feelings (e.g. “I didn’t feel comfortable”), as well as service outcomes (e.g “helped me so much”). The latter two example rules in topic k2 provide an interesting example of a contextually embedded regularity denoting parernts with young children (e.g. “my baby / daughter was weeks old”). Example rules for topic k3 showed a number of textual regularities pertaining to the mental health unit of one particular location and region, while for topic k4, we show two example rules generated from the one phrase “nothing was too much trouble”, used a number of times in documents in this topic as part of a recount of very helpful and considerate service received. The example rules for topic k5 show recurrent expressions about a “memoy assessment service” in a particular location, at a particular community hospital. For topic k6, the example rules encompass the mention of a “social inclusion and well being service” and some examples of more explicit positive sentiment (e.g. “have really enjoyed the course”). In topic k7, the ruleset included the text “transcribed on behalf of the patient”, which exists in the corpus as a pretext in the case of such. This inclusion is useful, as it shows when an intermediary is involved, and also that these transcriptions are linked to a mental health service. The example rules here also show expressions pertaining to potential staff shortages ("not enough staff") and the expression "found it hard" which was used in both first and third person in the documents to describe condition symptoms. The example rules in topic k8 capture one particular dental clinic in one particular location, reinforced by the presence of the same expression on both sides of the rule. The remaining two examples show the expressions "staff were very helpful" and "made me feel relaxed".

Table 9: Example Rules – Online Review Dataset

	Antecedent:	Consequence:	Confidence (%)	Set Support (%)	Lift
k1	i_lost_my	lost_my_job	75.00%	0.97%	77.00
	helped_me_so	me_so_much	75.00%	0.97%	46.20
	did_n't[not]_feel, i_did_n't[not]	n't[not]_feel_comfortable	60.00%	0.97%	61.60
k2	thank_you_so	you_so_much	100%	1.35%	73.88
	my_daughter_was, was_weeks_old	daughter_was_weeks	83.30%	0.85%	98.50
	my_baby_was	was_weeks_old	70.00%	1.18%	19.70
k3	(location 1)_mental_health	mental_health_unit	100%	3.66%	19.50
	(organisation 1)	(region 1)_(organisation 1)	92.90%	4.76%	13.34
	(region 1)_(organisation 1)	(organisation 1)	68.40%	4.76%	13.34
k4	nothing_was_too, was_too_much	too_much_trouble	100%	1.42%	21.39
	nothing_was_too, too_much_trouble	was_too_much	100%	1.42%	61.50
	you_very_much	thank_you_very	80.00%	2.44%	30.28
k5	attended_memory_assessment	memory_assessment_service	100%	3.21%	16.77
	service_(location 2)_community, assessment_service_(location 2), memory_assessment_service	(location 2)_community_hospital	100%	1.83%	43.60
	service_(location 2)_community, assessment_service_(location 2), (location 2)_community_hospital	memory_assessment_service	100%	1.83%	16.77
k6	you_very_much	thank_you_very	100%	1.17%	68.20
	it_helped_me, inclusion_wellbeing_service	social_inclusion_wellbeing	100%	1.17%	22.73
	have_really_enjoyed	really_enjoyed_course	66.70%	1.17%	18.94
k7	transcribed_behalf, mental_health	behalf_patient	100%	2.08%	22.40
	enough_staff	not_enough	85.70%	1.79%	20.57
	it_hard	found_it	71.40%	1.49%	20.00
k8	(location 3)_dental	(location 3)_clinic	88.20%	4.32%	20.44
	were_very_helpful	staff_were_very	87.50%	1.01%	38.01
	me_feel_relaxed	made_me_feel	71.40%	0.72%	33.10

Concluding Remarks and Managerial Insights

Collectively, the results of this working paper demonstrate the potential application of data-driven techniques in CX analytics, the usefulness of combining data-driven methods with theoretical CX frameworks, and the possibilities and advantages of not making a separation between constituent terms in a vocabulary when investigating the dimensions of CX. Even on relatively short documents, retrieved topic labels were shown to be a useful starting point for understanding CX. By employing theoretically informed POS tagging and filtering, association rules are also able to produce useful insights along the lines of different CX dimensions (*Objects and Entities*, *Thoughts, Feelings and Sensations*, *Behaviors* and *Ownership and Relation)*. More specifically, by fully leveraging textual regularities, managers can explore customer experiences as they manifest in a highly focused manner, and gain a better understanding of the factors underlying customer experiences, beyond just manifestations of select pre-determined words and expressions in recounts. The combined approach presented here makes use of a broader vocabulary to paint a more holistic but succinct picture of CX. Additionally, our approach shows promising results across three different kinds of primary datasets, broadly reflecting possible future usage applications, informed by popular channels for gathering textual CX data, and directly addressing industry trends towards leveraging new data sources and embracing technology-enabled, commercial purpose driven customer insight (MSI, 2016; MSI, 2018).

At a micro-level, consider the following implications for mangers. First, a manager of a service organization identifies that a customer had a positive overall experience (Thoughts, Feelings and Sensations), that included a positive social component (e.g. positive interactions with service staff, a pleasurable outing with friends), but also that the customer engaged in needlessly complex cognition during their decision process. Knowing this, the manager would need to know in the future which specific aspect of the service experience to leverage further (social), and which aspect needs improvement (facilitating more fluid customer cognition during the purchase decision process). Second, a retail manager may be aware of a number of dissatisfied customers based on a collection of online reviews or social media posts. In such a case, the manager needs to be aware not only of the recurrent nature of a particular issue (e.g. not enough staff), but also about regularities pertaining to specific touchpoints in the customer journey (*Objects and Entities*), and which social actors are involved (*Ownership and Relation*). To target resolution actions, a manager would be able to distinguish for example between the physical (e.g. store) versus digital (e.g. website) touchpoints involved and between a generally dissatisfying experience versus a catastrophic service failure (*Objects and Entities*), which has lead the customer to seriously consider competitor alternatives. The manager would also need to distinguish between a sole decision maker and purchaser versus a purchase made by someone on the behalf of another social group (*Behaviors*, e.g. family and friends), and with whom they came into contact during their experience (*Behaviors*, e.g. different staff members). Third, in the case of more complex services, with more human contact points involved (e.g. in a health service context), the manager here needs to know, among many other issues, which aspects of a patient's thoughts and feelings pertain to which interaction, with which human contact point, and at which stage in the patient's recovery journey such thoughts and feelings emerged. As a final example, take a manager in a B2B service setting. This manager would need to understand in particular how human contact points facilitate the delivery of service satisfaction to clients. Additionally, for B2B services linked or related to physical manufactured products, satisfaction or dissatisfaction may arise surrounding a broad range of partners, suppliers and components.

Limitations and Future Research

This study has the following limitations which should be noted, but which also open avenues for future research. Firstly, relatively small datasets are used. This was intentional, to focus the purpose of the project on theoretically informed methodology development, and we intend for future research to develop the approach further at scale. Second, our approach does not formally take temporal dimensions into account. This being said, by separating corpora into time windows, or formally into customer journey stages and touchpoints, our approach remains applicable. Third, future research will deepen the investigation into the use of association rules for text mining. Future research will explore variable rule confidence levels and other metrics, along with ngrams and term sets of varying sizes.

References

Aaker, J., Drolet, A. and Griffin, D., 2008. Recalling mixed emotions. *Journal of Consumer Research*, 35(2), pp.268-278.

Abbasi, A., Zhou, Y., Deng, S. and Zhang, P., 2018. Text Analytics to Support Sense-making in Social Media. *MIS Quarterly*, 42(2), pp.427-464.

Apache OpenNLP (2018). Retrieved from: https://opennlp.apache.org/

Archak, N., Ghose, A. and Ipeirotis, P.G., 2011. Deriving the pricing power of product features by mining consumer reviews. *Management Science*, 57(8), pp.1485-1509.

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

Brakus, J.J., Schmitt, B.H. and Zarantonello, L., 2009. Brand experience: what is it? How is it measured? Does it affect loyalty?. *Journal of Marketing*, 73(3), pp.52-68.

Bone, S.A., Lemon, K.N., Voorhees, C.M., Liljenquist, K.A., Fombelle, P.W., Detienne, K.B. and Money, R.B., 2017. “Mere Measurement Plus”: How Solicitation of Open-Ended Positive Feedback Influences Customer Purchase Behavior. *Journal of Marketing Research*, 54(1), pp.156-170.

Chun, H.H., Diehl, K. and MacInnis, D.J., 2017. Savoring an Upcoming Experience Affects Ongoing and Remembered Consumption Enjoyment. *Journal of Marketing*, 81(3), pp.96-110.

Culotta, A. and Cutler, J., 2016. Mining brand perceptions from twitter social networks. *Marketing Science*, 35(3), pp.343-362.

Curme, C., Preis, T., Stanley, H.E. and Moat, H.S., 2014. Quantifying the semantics of search behavior before stock market moves. *Proceedings of the National Academy of Sciences*, 111(32), pp.11600-11605.

D’Angelo, G., Rampone, S. and Palmieri, F., 2017. Developing a trust model for pervasive computing based on Apriori association rules learning and Bayesian classification. *Soft Computing*, 21(21), pp.6297-6315.

De Keyser, A., Lemon, K.N., Klaus, P. and Keiningham, T.L., 2015. A framework for understanding and managing the customer experience. Marketing Science Institute working paper series, pp.15-121. Retrieved from: http://www.msi.org/reports/a-framework-for-understanding-and-managing-the-customer-experience/

Dolan, R.J., 2002. Emotion, cognition, and behavior. *Science*, 298(5596), pp.1191-1194.

Farrell, J., 2016. Corporate funding and ideological polarization about climate change. *Proceedings of the National Academy of Sciences*, *113*(1), pp.92-97.

Forbes. 2013. Four Ways To Improve The Buyer Experience Starting Today. Retrieved from: http://www.forbes.com/sites/lisaarthur/2013/06/26/four-ways-to-improve-the-buyer-experience-starting-today/#4faf9c1730ce

Gartner. 2016. Customer Experience Primer for 2016. Retrieved from: https://www.gartner.com/binaries/content/assets/gml/ki-pages/research-primers/a925c1dd-9211-44bf-95af-77163d4b540b_gartner_for_marketers_customer_experience_primer.pdf

Grewal, D., Levy, M. and Kumar, V., 2009. Customer experience management in retailing: An organizing framework. *Journal of Retailing*, 85(1), pp.1-14.

Grossmann, I., Huynh, A.C. and Ellsworth, P.C., 2016. Emotional complexity: Clarifying definitions and cultural correlates. *Journal of Personality and Social Psychology*, 111(6), p.895.
Homburg, C., Jozić, D. and Kuehnl, C., 2017. Customer experience management: toward implementing an evolving marketing concept. *Journal of the Academy of Marketing Science*, 45(3), pp.377-401.

Hoffman, D.L. and Novak, T.P., 2017. Consumer and object experience in the internet of things: An assemblage theory approach. *Journal of Consumer Research*, 44(6), pp.1178-1204.

Iacobucci, D., Ostrom, A. and Grayson, K., 1995. Distinguishing service quality and customer satisfaction: the voice of the consumer. *Journal of Consumer Psychology*, 4(3), pp.277-303.

Jiang, Z.Q., Li, W.H., Liu, Y., Luo, Y.J., Luu, P. and Tucker, D.M., 2014. When affective word valence meets linguistic polarity: Behavioral and ERP evidence. *Journal of Neurolinguistics*, 28, pp.19-30.

Kang, S.M. and Shaver, P.R., 2004. Individual differences in emotional complexity: Their psychological implications. *Journal of Personality*, 72(4), pp.687-726.

Knox, G. and Van Oest, R., 2014. Customer complaints and recovery effectiveness: A customer base approach. *Journal of Marketing*, 78(5), pp.42-57.

Kumar, V., Umashankar, N., Kim, K.H. and Bhagwat, Y., 2014. Assessing the influence of economic and customer experience factors on service purchase behaviors. *Marketing Science*, 33(5), pp.673-692.

Kune, R., Konugurthi, P.K., Agarwal, A., Chillarige, R.R. and Buyya, R., 2016. The anatomy of big data computing. *Software: Practice and Experience*, 46(1), pp.79-105.

Lee, T.Y. and Bradlow, E.T., 2011. Automated marketing research using online customer reviews. *Journal of Marketing Research*, 48(5), pp.881-894.

Lemon, K.N. and Verhoef, P.C., 2016. Understanding customer experience throughout the customer journey. *Journal of Marketing*, 80(6), pp.69-96.

Lindquist, K.A. and Barrett, L.F., 2008. Emotional complexity. Handbook of emotions, 4, pp.513-530.

Lu, H.M., Wei, C.P. and Hsiao, F.Y., 2016. Modeling healthcare data using multiple-channel latent Dirichlet allocation. *Journal of Biomedical Informatics*,*60*, pp.210-223.

Maklan, S., Antonetti, P. and Whitty, S., 2017. A Better Way to Manage Customer Experience: Lessons from the Royal Bank of Scotland. *California Management Review*, 59(2), pp.92-115.

MSI - Marketing Science Institute. 2016. Marketing Science Institute 2016-2018 Research Priorities. Retrieved from: http://www.msi.org/research/2016-2018-research-priorities/

MSI - Marketing Science Institute. 2018. Marketing Science Institute 2018-2020 Research Priorities. Retrieved from: http://www.msi.org/research/2018-2020-research-priorities/

Mitchell, T.M., 2005. Generative and discriminative classifiers: Naive bayes and logistic regression. Machine Learning. McGraw Hill.

Newman, D., Asuncion, A., Smyth, P. and Welling, M., 2009. Distributed algorithms for topic models. *Journal of Machine Learning Research*, 10(Aug), pp.1801-1828.

Ordenes, F.V., Theodoulidis, B., Burton, J., Gruber, T. and Zaki, M., 2014. Analyzing Customer Experience Feedback Using Text Mining A Linguistics-Based Approach. *Journal of Service Research,* p.1094670514524625.

Park, S.H., Jang, S.Y., Kim, H. and Lee, S.W., 2014. An association rule mining-based framework for understanding lifestyle risk behaviors. *PloS One*, 9(2), p.e88859.

Penn Treebank (2018). Retrieved from:
https://www.ling.upenn.edu/courses/Fall_2003/ling001/penn_treebank_pos.html

Puranam, D., Narayan, V. and Kadiyali, V., 2017. The Effect of Calorie Posting Regulation on Consumer Opinion: A Flexible Latent Dirichlet Allocation Model with Informative Priors. *Marketing Science*, 36(5), pp.726-746.

Richins, M.L., 1997. Measuring emotions in the consumption experience. *Journal of Consumer Research*, 24(2), pp.127-146.

Ringel, D.M. and Skiera, B., 2016. Visualizing asymmetric competition among more than 1,000 products using big search data. *Marketing Science*, 35(3), pp.511-534.

Roy, B.C., Frank, M.C., DeCamp, P., Miller, M. and Roy, D., 2015. Predicting the birth of a spoken word. *Proceedings of the National Academy of Sciences*, 112(41), pp.12663-12668.

Rust, R.T. and Huang, M.H., 2014. The service revolution and the transformation of marketing science. *Marketing Science*, 33(2), pp.206-221.

Schmitt, B., 2011. Experience marketing: concepts, frameworks and consumer insights. Now Publishers Inc.

Schmitt, B., Brakus, J.J. and Zarantonello, L., 2015. From experiential psychology to consumer experience. *Journal of Consumer Psychology*, 25(1), pp.166-171.

Schweidel, D.A. and Moe, W.W., 2014. Listening in on social media: A joint model of sentiment and venue format choice. *Journal of Marketing Research*, 51(4), pp.387-402.

Stephen, A.T., Dover, Y., Muchnik, L. and Goldenberg, J., 2017. Pump It Out! The Effect of Transmitter Activity on Content Propagation in Social Media. Saïd Business School WP 2017-01. Available at SSRN: https://ssrn.com/abstract=2897582 or http://dx.doi.org/10.2139/ssrn.2897582

Stephen, A.T. and Lamberton, C., 2016. A thematic exploration of digital, social media, and mobile marketing research's evolution from 2000 to 2015 and an agenda for future research. *Journal of Marketing*, 80(6), pp.146-172.

Tirunillai, S. and Tellis, G.J., 2014. Mining marketing meaning from online chatter: Strategic brand analysis of big data using latent dirichlet allocation. *Journal of Marketing Research*, 51(4), pp.463-479.

Van de Kauter, M., Breesch, D. and Hoste, V., 2015. Fine-grained analysis of explicit and implicit sentiment in financial news articles. *Expert Systems with Applications*, 42(11), pp.4999-5010.

Verhoef, P.C., Antonides, G. and de Hoog, A.N., 2004. Service encounters as a sequence of events the importance of peak experiences. *Journal of Service Research*, 7(1), pp.53-64.

Verhoef, P.C., Lemon, K.N., Parasuraman, A., Roggeveen, A., Tsiros, M. and Schlesinger, L.A., 2009. Customer experience creation: Determinants, dynamics and management strategies. *Journal of Retailing*, 85(1), pp.31-41.

Verhoef, P.C., Kooge, E. and Walk, N., 2016. Creating Value with Big Data Analytics: Making Smarter Marketing Decisions. Routledge.

Weill, P. and Woerner, S.L., 2018. Is Your Company Ready for a Digital Future?. *MIT Sloan Management Review*, 59(2), pp.21-25.

Appendix: Runtime Tests

This working paper employed a hill-climb optimization procedure for LDA parameter optimization and tuning. This approach takes input search boundaries and steps, set for focal parameters and uses a reduced permutation strategy based on a random starting point, versus iteratively trialling every parameter combination within given boundaries and steps. The computational efficency gains are illustrated below in Appendix Figure 1, where “or” denotes the online review dataset, “sm” denotes the social media dataset, and “sr” denotes the survey (B2B) dataset. BF denotes brute force (all parameter combinations), HC denotes the hill-climb procedure.

Appendix Figure 1: Parameter Optimization Configurations and Runtime

Here, the number of parameter combinations and runtime (normalized) and validation accuracy (using Naïve Bayes multi-class classification) are shown for a simple runtime experiment, with LDA given the same sets of parameter boundaries and steps to optimize between 2 and 3 forced topic counts. Interestingly, runtime was shown to be around 14x faster using the hill-climb optimization approach versus the brute force ‘all’ approach, with parameter combinations, correspondingly, around 14.5x less. Most importantly however, in this experiment, in the case of all dataset, better prediction accuracy upon validation was achieved using the parameters achieved using the hill-climb approach. Although likely to vary when the approaches are afforded a larger range of topic steps, this provides us with additional confidence in employing this substantially more computationally efficient optimization approach.