Review on Recent Advances in Information Mining from Big Consumer Opinion Data for Product Design

Jian Jin

Department of Information Management, School of Government,

Beijing Normal University,

Beijing 100875, China

Ying Liu*

Mechanical and Manufacturing Engineering, School of Engineering,

Cardiff University,

Cardiff CF24 3AA, United Kingdom

Email: LiuY81@cardiff.ac.uk

Ping Ji

Department of Industrial and Systems Engineering

The Hong Kong Polytechnic University, Hong Kong SAR, China

C. K. Kwong

Department of Industrial and Systems Engineering

The Hong Kong Polytechnic University, Hong Kong SAR, China

In this paper, based on more than ten years' studies on this dedicated research thrust, a comprehensive review concerning information mining from big consumer opinion data in order to assist product design is presented. First, the research background and the essential terminologies regarding online consumer opinion data are introduced. Next, studies concerning information extraction and information utilization of big consumer opinion data for product design are reviewed. Studies on information extraction of big consumer opinion data are explained from various perspectives, including data acquisition, opinion target recognition, feature identification and sentiment analysis, opinion summarization and sampling, etc. Reviews on information utilization of big consumer opinion data for product design are explored in terms of how to extract critical customer needs from big consumer opinion data, how to connect the voice of the customers with product design, how to make effective comparisons and reasonable ranking on similar products, how to identify ever-evolving customer concerns efficiently, and so on. Furthermore, significant and practical aspects of research trends are highlighted for future studies. This survey will facilitate researchers and practitioners to understand the latest development of relevant studies and applications centered on how big consumer opinion data can be processed, analyzed and exploited in aiding product design.

1. INTRODUCTION

The continuous growth of e-business and Web 2.0 provides consumer more opportunities to share their opinions and experience gained from different products and services. It promotes an unprecedented volume of consumer data that are generated incessantly in various e-commerce and social network websites. For instance, customer transaction logs and click-through data are generated constantly in Amazon.com, Taobao.com, etc. Then, besides these e-commerce websites, customer opinions are posted from time to time on social networks, e.g., Twitter.com; review websites, e.g., Epinions.com, media websites, e.g., Cnet.com, and so on. Such consumer opinion data benefit potential customers to make wise purchase decisions with similar alternatives, provide informative messages to managers responsible for marketing strategy and planning facing a fierce competition, and facilitate product designers on a better understanding of customer concerns so to offer desirable products intended for a higher level of customer satisfaction.

Customer requirements (CRs), which play an important role in decision-making in the market-driven new product design (NPD), are conventionally obtained from a small number of pre-formatted high-quality survey data. However, the sheer volume and the generating velocity of online CRs data surpass human's ability to make sense out of them in a reasonable timeframe. Also, these online data, with different quality, are distributed in different resources and are embedded in informal free texts. It further burdens human's capacity to effectively obtain accurate CRs without much loss. Hence, internationally, a fast-growing interest is observed to research various algorithms and platforms to parse such a big volume of consumer opinion data. Many endeavors in this field are mainly reported in the academic disciplines of computer science and business management. Typical models are utilized to identify high-quality consumer opinion data, detect customer sentiment polarity and its corresponding opinion target intelligently at different levels, e.g., system level and product feature level. Besides, in order to provide a quick outline of customer opinions and grasp their major CRs, based on the recognized customer sentiments in online textual data, how to generate a summarization or a short list of representative opinions from a big volume of online consumer opinion data has also been investigated.

In addition, quality function deployment (QFD), as one of the widely applied methods that transform CRs into engineer characteristics (ECs), receives extensive discussions in the fields of business management and engineering design with an aim to fuel NPD pipelines from analyzing valuable consumer opinion data. Some welcome topics regarding QFD often involve the prioritization of a variety of CRs, how to integrate consumer opinion data into product design, etc. This is particularly important because today's NPD has to face an intensive competition globally since many alternative products and services are made available to 'choosy' consumers while both the pros and cons of alternatives can be easily found out and compared. Not to mention, many intensively competitive markets are also characterized by the fast evolution of CRs. It implies that scholars and practitioners in management should make full preparation to engage creative models to be innovative forces in the design industry, which are developed carefully for analyzing the competitive strategies and dynamics of CRs. All these endeavors will inevitably facilitate designers to exploit product affective information, customer behavior as well as customer concerns for NPD and integrate valuable consumer inputs in practical applications.

In Figure 1, an overview of the classic procedures concerning information mining from big consumer opinion data for product design is presented. Generally, three core components are included. First, a large volume of consumer opinion data can be obtained from analyzing the structures of various sites, e.g., e-commerce sites, social network sites, customer review sites, media sites, and so on, by web crawler programs. Next, consumer opinion data are processed and mined to acquire the customer concerns. Several critical tasks are involved, e.g., quality analysis, target identification, sentiment analysis, summarization, and sampling, etc. Customer concerns that are processed and mined from online big consumer opinion data formulate the valuable information of VOC (voice of the Customer). Then, extracted valuable VOCs from big consumer opinion data are expected to be utilized from different perspectives of product design. Due to the practical implication imposed by technological and economic concerns and the dynamics of fast-changing CRs, only essential VOCs can to be integrated into design and product offerings should also be compared to highlight their competitive advantages and targeted design and development strategies.

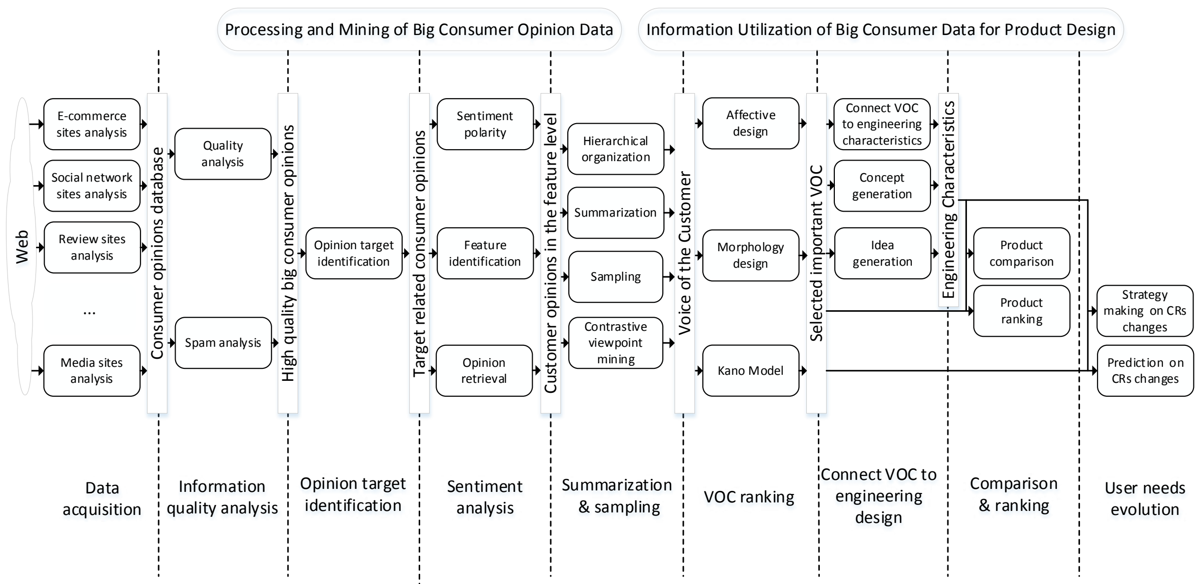

Figure 1. A framework of classical procedure regarding information mining from big consumer opinion data for product design

This paper is organized as follows. In Section 2, a comprehensive review is given focusing on how big consumer opinions are being processed and exploited to gain critical insights into product design. Next, in Section 3, existing approaches with respect to how big consumer opinions can be utilized in design are revealed. Lastly, some open research challenges, trends, potential future studies are presented to highlight the significance and value of information mining from big consumer opinion data for product design.

2. PROCESSING AND MINING OF BIG CONSUMER OPINION DATA

2.1 Information Quality

After a big volume of consumer opinion data are collected from different sites, as presented in the second lane of Figure 1, the very first concern about the introduction of big consumer opinion data into product design is the quality related issue. Relevant studies regarding information quality analysis mainly discuss how to evaluate the quality of big consumer opinion data and how to detect opinion spams.

2.1.1 The Quality of Customer Online Reviews

As a matter of fact, the quality of information available in a community is often inversely related to the size of its membership [1]. There exist some studies about evaluating the quality of online reviews. In these studies, the quality of online reviews is often defined in terms of online helpfulness voting ratio, x/y (x out of y people find a particular review helpful, e.g., "204 out of 209 people found the following review helpful"). The helpfulness voting ratio is regarded as the golden criterion in those publications to define the helpfulness of product reviews. In these literatures, the problem of helpfulness prediction is formulated as a regression problem [1 - 8] and a binary classification [9 - 12] with several categories of features, such as sentiment features, user reputation or expertise features, information quality based features, etc. For instance, Liu et al. investigated reviewers' expertise, writing styles and timeliness from observations in terms of how the helpfulness of online reviews is influenced [4, 5]. These factors were then combined linearly to estimate the percentage of online helpful votes. In their later research, this model was utilized to forecast the sales of a product [13]. Hong et al. focused on learning user preferences in three aspects, i.e., whether reviews meet their information needs, whether reviews is credible and whether reviews have the mainstreaming opinion [11]. Then, a binary classification algorithm was modeled to predict the helpfulness of online reviews. A sound analysis was conducted regarding several hypotheses that might influence the percentage of online helpful votes [7]. These hypotheses include the conformity hypothesis, the individual-bias hypothesis, the brilliant-but-cruel hypothesis and the straw-man hypothesis. Finally, the helpfulness voting ratio were said to not just depend on the content but also "on how the expressed evaluation score that relates to other evaluation scores of the same product." Besides, different feature selection methods were compared to evaluate the effectiveness of features in predicting online helpfulness voting ratio [14].

Meanwhile, as consumers are not obligated to vote such reviews, usually, only a small proportion of reviews eventually receive sufficient votes. Different from the above helpfulness voting ratio, Liu et al. argued that the helpfulness represented by the helpfulness ratio is not fair due to kinds of biases [15]. Next, human annotators were hired to evaluate the helpfulness. Annotators' evaluations were found to be different greatly from the helpfulness votes observed on Amazon.com and, given a guideline for the helpfulness evaluation, evaluations from annotators achieve highly consistent through a kappa statistic. Then, the helpfulness evaluation was modeled as a multiple class classification problem and categories of features were extracted to predict annotators' evaluation. Inspired by Liu's work [15], Zhang and Tran brought a probabilistic distribution model for the judgment of whether a review is helpful or not directly from review texts [16]. Chen and Tseng offered a review evaluation framework using information quality [17]. They classified reviews into five groups, i.e. "high-quality", "medium-quality", "low-quality", "duplicate" and "spam", and categories of features were utilized to detect the quality of product reviews. However, Li et al. argued that the annotated corpus for each domain of interest is not practical [18]. They devised a snippet based unsupervised learning approach to estimate the sentiment of online reviews. This approach was utilized to classify whether online reviews are classified as recommended or not recommended by analyzing reviews that receive unanimous judgments of two annotators only. Besides, it was found that, given a helpfulness guideline, helpfulness values for annotators are not aligned with online voting [19]. Then, the averaged helpfulness labeling value was analyzed by a regression-based approach, and which domain-independent features are significant for helpfulness analysis had also been analyzed.

2.1.2 Opinion Spam

Due to various reasons, fake reviews are widely appearing in e-business websites and they are highly possible to mislead consumers and review analyzers. Accordingly, different methods have been developed to identify opinion spamming activity.

Jindal and Liu defined three types of opinion spams, including untruthful opinions, reviews on brands only and non-reviews [20]. Then, a logistic regression model was built to identify the latter two types. For the first case, the approach of lift curves was utilized to visualize the prediction results and identify reviews whose ratings deviate from the average rating to a large extent. Later, they proposed a rule-based approach to identify unusual review patterns from online reviews [21]. From then, some researchers began to pay attention to the problem about opinion spam. Conventionally, they extracted categories of features from review texts and/or from reviewers and built a classification model to analyze whether a review is a suspicious one [22 - 27]. However, these studies mainly analyzed the content of online reviews, a sharp increase in online rating with a dramatically increasing or decreasing may imply a potential manipulation by newly arrived reviews [28, 29]. Accordingly, a template-based approach was developed to detect the burst of online rating.

Besides investigations about opinion spam detection, some scholars studied review spammer identification. Lim et al. described that spammers may post several reviews to the same product with similar ratings or give a specific group of products higher/lower ratings [30]. Accordingly, a regression method was applied to identify spam reviews and review spammers. Similarly, spamming clues were identified by a graph model-based approach and an iterative algorithm was developed to identify potential spammers [31]. Also, some utilized some intuitive features, such as a reviewer's rank, a reviewer's total number of posted reviews, and the number of a reviewer's badges, etc. to predict whether a reviewer is trustful or not [32]. Mukherjee et al. argued that a fake reviewer group might induce worse influence than individual fake reviewers [33, 34]. Then, eight group spam behavior indicators and four individual spam behavior indicators were utilized and an iterative approach was utilized to identify fake reviewer groups.

2.2 Opinion Target Identification

As presented in the third lane of Figure 1, after high quality customer opinion data are obtained, the next focus is on the opinion target identification. In particular, only one specific product might be focused in his/her reviews by a novice customer. However, for some experienced customers, different products might be referred to a single piece of online review and different alternatives are often compared in terms of various aspects. Then, an interesting problem is how to identify the exact product that the customer refers to automatically. It hence motivates some scholars to explore the interesting problem about opinion target identification.

One of early work in this category was conducted by Dalvi et al., which aims to match an online review to a possible object [35]. In their study, words in object attribute descriptions were translated to words in online reviews and the Expectation Maximization (EM) algorithm was utilized to estimate parameters. Similarly, another word-based transition model was utilized to estimate relations between nouns that potentially point to opinion targets [36]. Besides word based transition models, some employed a graph-based model [37] and the Centering Theory [38] to extract explicit and implicit opinion targets.

To identify names from consumer opinion data, a three-phase approach is introduced [39, 40]. First, the Brown Clustering algorithm was applied to obtain a cluster of words with similar meaning and the brand name variation is captured by linguistic rules. Next, a CRFs (Conditional Random Fields) based approach was adopted to analyze whether a word indeed refers to a specific product model name. Finally, a rule-based name normalization was utilized to map names to their formal names. Similarly, tokens, POS (Part of Speech) tags, the shortest dependency paths, word distances between pairs of opinion words and product features were extracted and they were utilized as features in a CRFs based approach for the detection of opinion target [41].

2.3 Sentiment Analysis

The fourth lane in Figure 1 points to sentiment analysis. Sentiment analysis or opinion mining on big consumer opinion data generally refers to a computational study of customer attitudes towards products on customer textual data. As presented in Figure 1, the input of sentiment analysis module often contains customer opinions with a target product. Some critical tasks in sentiment analysis involve how to extract sentiment polarity, how to identify different product aspects that customers care about, how to retrieve targeted products with representative opinions, etc. Also, in this section, cross-language opinion mining will be briefly reviewed.

2.3.1 Sentiment Polarity

The task of sentiment polarity extraction on online opinion textual data is to analyze whether a customer expresses a positive/negative/neutral opinion towards the target product or a particular feature of one product. In many studies about this task, sentiment words are often assumed to be adjectives or adverbs. Generally, these studies can be categorized into linguistic rules based algorithms [42 - 44], conventional classification based approaches [45 - 48], graph model based approaches [49 - 51], deep learning based approaches [52, 53], etc.

One of the most straightforward approaches is to apply some heuristic rules to obtain the sentiment polarity. Based on some intuitive heuristic hypotheses, researchers tried to score words for analyzing the sentiment polarity. For instance, a typical hypothesis can be that the ratings for an extensively discussed feature are more similar to the overall ratings [54]. Then, according to these heuristic hypotheses, a word based scoring scheme was proposed to estimate the customer sentiment polarity of a particular feature. Similarly, based on some intuitions, a frequency statistic term weighting method was built to judge the sentiment polarity of one sentence [55]. Ding and Liu built a holistic lexicon-based approach that applied an opinion aggregation function to estimate the sentiment polarities of product features [42, 43]. In this holistic approach, three linguistic rules were utilized to find contextual information that is helpful to infer the sentiment polarity of context-dependent opinion words. Besides, such intuitions based word scheme was also reported to analyze customer sentiment polarity in Chinese reviews [56] and learn domain-specific sentiment lexicons [57].

Note that, the objective of sentiment polarity extraction can be also regarded a typical binary classification problem (positive vs. negative) or trinary classification problem (positive, neutral, negative). Accordingly, some traditional methods of classification such as the techniques of support vector machine (SVM) were utilized to classify these customer textual data. For example, with SVM, categories of features were extracted for analyzing the sentiment polarity, e.g. bag of words [45], features about texts, affix similarity and word emotion [58], features about texts and appraisal scores [59], etc. Similarly, the least squares based linear model [60], an additive model [61], linguistic modalities with opinion holding predicates [44] were reported to estimate the sentiment polarities of online reviews in different levels. Also, word dependency relations were also utilized and different dependency tree based algorithms were designed for the analysis of sentiment polarity [46, 47, 62 - 64]. Besides these studies that analyzed the sentiment polarity using a single independent classifier, some employed a sequential model to gain the sentiment polarity in sentence level or product feature level, such as a Markov Model-based approach [48], a Hidden Markov Model (HMM) based approach [66], a CRFs based approach [67, 68], etc.

Customer online opinions, which are one important type of textual data regarding consumer information, attract researchers apply some widely applied probabilistic models such as Latent Dirichlet Allocation (LDA) to analyze customer sentiment polarity. For instance, a probabilistic mixture model was proposed to analyze topic and sentiment simultaneously [49]. In such model, a document was regarded to be generated by background words. Then, feature words were selected from one of many subtopics and a sentiment word was utilized to describe the topic. Similarly, a MaxEnt-LDA model which regards reviews as a combination of background words, general aspects, general opinions, aspects and aspects-specific words was presented to extract different aspects and the corresponding sentiment polarity [50]. Besides, graph models that consider relations between different sentences [51], relations between product features [69], relations between products and reviewers [70] were reported for the analysis of sentiment polarity.

Obviously, the construction of reliable corpus or a dictionary may highly affect the success of supervised learning approaches for the analysis sentiment polarity. Then, to disambiguate the sentiment polarities of some ambiguous words in a specific context, based on a large labeled corpus, conditional mutual information was defined and it was later utilized to evaluate the sentiment polarities of German reviews [71]. However, to build corpus manually is always time-consuming and labor-intensive. Accordingly, pros and cons reviews in Epinions.com were utilized as training data and three categories of features were extracted to predict the sentiment polarities [72]. Similarly, pros and cons reviews in cnet.com, viewpoints.com, reevoo.com, and gsmarena.com were integrated in a linear probabilistic model to determine the sentiment polarities in product feature level [73] and a fine-grained emotion dictionary was built by using typical news with predefined emotion list, which includes touching, empathy, boredom, anger, amusement, sadness, surprise and warmness [74]. To take the advantages of both approaches, an integration of lexicon-based and corpus-based approach was also reported for sentiment polarity analysis [75]. Based this approach, sentiment polarities in online opinions were incorporated in a collaborative filtering framework [76].

2.3.2 Feature Identification

A highly correlated task with the analysis of sentiment polarity is feature identification from customer online opinions. Take opinions about a mobile phone for instance. Different features, such as battery, screen, weight, etc. might be discussed by consumers. Automatic feature identification helps designers to locate exact customer concerns.

In some early studies, heuristic rules were often employed to identify features from customer concerns. For instance, Hu and Liu utilized the association mining algorithm to generate a set of frequent nouns or noun phrases and they were regarded as possible product features [77]. Similar techniques were reported to use heuristic frequent language rules to find words or phrases which match the rule patterns [78]. Later, linguistic rules were generated according to words relations. Popescu and Etzioni illustrated a system that was able to identify both components and properties of the given product, in which a rule-based relaxation labeling approach was applied to find product features and extract word sentiment orientation recursively [79]. Qiu et al. analyzed relations between sentiment words and product features in dependency trees and several rules were derived to extract both sentiment words and product features iteratively [47]. Besides relations between sentiment words and product features, opinion-opinion relations and feature-feature relations were considered in a bootstrapping framework [80], in which the Likelihood Ratio Tests for Bootstrapping and the Latent Semantic Analysis for Bootstrapping were applied. Additionally, the task about product feature extraction was regarded as a topic co-reference resolution problem [81]. It assumed that two opinions will share the same opinion topic if they are topic co-referent. Then, categories of features were utilized for pairwise topic coreference classification. Also, studies that analyze latent semantic relations between product features from pros and cons reviews were reported [82].

As aforementioned arguments, customer online opinions are one important type of textual data and some widely applied probabilistic models are often welcome for this particular task. To extract features and the corresponding sentiment polarity, an unsupervised probabilistic model called JST (joint sentiment/topic model) was proposed [83]. In this model, topics are generated dependent on sentiment and words are generated on sentiment as well as topic pairs. Later, a reverse model called Reverse-JST was presented where sentiments are generated dependent on topic distributions [84]. But it was found that the JST perform slightly better than the Reverse-JST. Another unsupervised model named AUSM (Aspect and sentiment unification model) was proposed [85]. In this model, topics are generated dependent on one sentiment and words in one sentence are generated according to one topic and corresponding sentiment. It was found that AUSM is more effective than JST. But some argued that AUSM fails to identify sentiment words that are specific to one aspect and it also does not separate sentiment words from factual words. Then, a joint aspect and sentiment model was proposed [86]. Also, some argued that AUSM does not consider multi-grain global topics and local topics and a topic model named MG-LDA (Multi-grain LDA) was proposed [87]. Similarly, a fine-grained labeled LDA was utilized to extract product features [88] However, MG-LDA was criticized that sentiment-oriented aspects are neglected and a joint multi-grain topic sentiment (JMTS) model was then developed [89]. Arjun Mukherjee and Liu argued that many approaches fail to cluster product features into the same category and, for this purpose, two probabilistic graphical models were presented to extract and cluster features simultaneously [90]. In addition, together with LDA, different probabilistic models were compared for the extraction of product features [91, 92].

Some researchers argued that words or phrases that refer to the same product features should be clustered together. Accordingly, an association rule method and a naive Bayesian method were utilized to classify words with similar meaning, which was represented as a product feature [93]. Similarly, a semi-supervised learning algorithm with two soft constraints [94], a Rocchio's algorithm based algorithm [95] and a multi-aspect bootstrapping algorithm [96] were proposed to cluster feature words. Also, a co-clustering algorithm was proposed to extract categories of product features and groups of opinion words by mutual reinforcement [97], in which the similarity between data objects was estimated by a linear combination of intra similarity and inter similarity.

2.3.3 Opinion Retrieval

The objective of opinion retrieve is to find documents with specific opinions. For this task, many scholars estimated a score for each document that linearly combines with both document relevance and document opinion score. For instance, Kim, Li and Lee estimated a document score [98], which linearly integrates three indicators, i.e. the probability that a word is opinion word which is estimated by the sentiWordNet [99] or the MPQA [100], the likelihood of a query given a word which is estimated by the Latent Semantic Analysis (LSA) and Point-wise Mutual Information (PMI), and the probability of a document generating a word which is estimated by conventional method, such as Vector Space model, BM25 or model. Similarly, an approach of opinion retrieval from blogs were proposed [101]. Ganesan and Zhai presented an opinion based entity ranking demo system to retrieve hotels that satisfy consumers in both structured preferences and unstructured preferences, in which two types of preferences were combined into a single score to evaluate how an entity meets the query [102].

However, some argued that the average over the opinion weights of terms to generate an opinion score was not optimal since multiple topics can be covered in a single document [103, 104]. Accordingly, these researchers estimated the opinion density at each position in one document and different opinion propagation functions were testified based on the position of each term in a document and the position of the query term. Finally, the opinion score was combined with the relevance score in a probabilistic retrieval model. Besides, an information retrieval framework was presented to facilitate consumers to identify high-quality reviews and make a comparison between products [105]. In this framework, an affinity opinion graph was built for ranking customer opinions according to information richness and information diversity. With this graph, opinions were ranked. Then, a greedy algorithm was utilized to present online reviews and products were compared using a radar graph.

2.3.4 Cross-lingual Sentiment Classification

Except for English, few manually built corpus in different languages is actually available. It interests researchers to borrow widely available corpus in English for cross-lingual sentiment analysis in other languages.

Wan proposed a lexicon based algorithm to identify sentiment of Chinese reviews [106]. In this algorithm, Chinese reviews were translated into English reviews by translation service and sentiment polarities were analyzed with the help of lexicons in both Chinese and English. Finally, several ensemble methods were proposed to combine results from different classifiers. Later, Wan utilized a co-training algorithm to classify Chinese reviews by using labeled English reviews and unlabeled Chinese reviews [107, 108]. In the training phase, each Chinese review was initially represented by both Chinese features and English features. Then two classifiers were built accordingly and instances without conflicting prediction from two classifiers were utilized as labeled data. In the testing phase, the average score from two classifiers was utilized as the final sentiment polarity. Also, four categories of methods for cross-lingual sentiment classification were analyzed comparatively [109]. Both lexicon-based methods and corpus-based methods were taken into consideration. Accordingly, with four categories of methods, three combination methods were proposed to boost the performance of cross-lingual sentiment classification.

With the help of the approach in [82] which analyzes latent semantic relations between product features from pros and cons reviews, a text mining system for cross-lingual opinion analysis was developed [110]. In this system, a product feature term was characterized as multi-dimension cross-lingual latent semantic clues and categories of features were extracted from reviews in a different language. Then, a topic model was applied to learn the latent topic structure about product features. Similarly, categories of syntactic and stylistic features were extracted from English and Arabic corpus and they were utilized to predict the sentiment polarities in sentence level [111]. Lin et al. argued that bilingual lexicon extraction methods tend to lead high complexity and unsatisfying precision [112]. Then, a mutual information based approach with a probabilistic pair alignment method and a word pair refinement algorithm was devised to extract a set of English-Chinese feature-opinion pairs.

2.4 Summarization and Sampling

Although product features as well as corresponding opinions can be identified, as shown in the fifth lane of Figure 1, the large volume of consumer opinion data still makes it impossible to read and comprehend the entire set. Then, a good approach is to provide a brief summarization on customer opinions. Hence, the next valuable tasks for product designers is to make a summarization about general customer concerns or sample representative customer feedbacks from big opinion data.

2.4.1 Hierarchical Organization of Consumer Reviews

Yu et al. built a hierarchical product structure by using online reviews [113, 114]. First, an initial hierarchy was constructed by parsing web pages which provide domain knowledge about products. Next, product features were extracted from Pros and Cons reviews with the help of synonym terms in thesaurus.com and the semantic distance between features were learned by analyzing underlying aspects. Finally, the aspect hierarchy was generated iteratively on the initial hierarchy according to three criteria, i.e., minimum hierarchy evolution, minimum hierarchy discrepancy and minimum semantic inconsistency. Based on this approach, their group presented an approach to answer opinion questions [115].

Zhai et al. proposed a method to cluster product features in online reviews [116]. In this method, two linguistic rules were utilized, i.e. feature expression sharing same words are likely to belong to the same group and the similarity of feature expression can be estimated by a WordNet distance-based approach. With these two rules, reviews can be self-labeled. Then, using both labeled and unlabeled instances, an EM algorithm was applied to cluster feature expressions into categories.

2.4.2 Review Summarization

To provide brief review summarization reports regarding customer description of different features will definitely facilitate both designers and consumers to grasp the main idea in big consumer opinion data. Generally, these reports are expected to summarize customer concerns about product features at the word level, phrase level or sentence level.

Four kinds of CRFs models were proposed to identify features and opinions from online concerns [117]. These CRFs were utilized to model dependencies between words and, according to identify features and opinions, a word level summarization report was presented. A summarization approach was proposed for rated aspects [118]. In this approach, different topic models were advised to identify aspects and a local prediction method and a global prediction method were employed to predict aspect ratings. Then, top three phrases with the highest frequency were selected to represent rated aspects. Also, an unsupervised learning approach was developed to provide a compact and informative summarization of online opinions [119]. In this approach, concepts about representativeness and readability were modeled and an optimization problem was formulated to seek concise and non-redundant phrases from high-frequency n-grams.

In these studies, a list of words/phrases/sentences that represent customer general concerns were provided. However, they neglect to group similar customer concerns together according to various features and present a higher level of opinion summarization for product comparison. Accordingly, Zhuang et al. identified product features and opinions with the help of manually labeled data and seed words in WordNet Then, a review summary report of sentence clusters was generated [120]. Another review summarization system was reported to cluster sentences with similar features in the same sentiment polarity [121]. In this system, both hierarchical groupwise-average clustering method and a non-hierarchical exchange method were applied to cluster features referred sentences. Then, those sentences with maximal information coverage were selected as representative sentences.

There also some studies utilizing the techniques of information visualization to summarize big volume of online opinions. For instance, a tag cloud based visualization approach was developed, which aims to help designers to understand a large number of customer online reviews [122]. Especially, adjective-noun word pairs were utilized to highlight the concentration of reviews. Also, a feature-based visualization system was built to analyze sentiment polarities of text document streams [123]. In this system, pixel map calendars and time density plots were applied to provide a global overview of the data distribution and to track the concrete temporal development of a single feature with high frequency.

2.4.3 Review Sampling and Recommendation

To gain the main concerns of consumers efficiently, some studies tried to sample representative sentences and recommend instructive reviews from a big volume of online comments.

An opinion summarization was proposed for Bengali news articles [124]. In this approach, a CRFs model was initially utilized to recognize theme words from subjective sentences and theme related sentences were clustered by the K-means algorithm. Then, a semantic graph was built to model document connections and a Page Rank like approach was applied to select representative sentences for each theme cluster. Also, to cluster reader comments in news articles, two probabilistic graph models were developed, in which representative sentences were selected by the approach of Maximal Marginal Relevance and the approach of Ranting & Length [125].

Ju et al. tried to evaluate the informativeness of a word and a document for effective sampling informative reviews [126]. Especially, the informativeness of a word was defined as the product of a certain POS proportion and the occurring frequency and the informativeness of sentences was defined as the normalized sum of informativeness of words. Then, an optimization problem was formulated to gain informative samples. Besides informativeness of sentences, other aspects were reckoned. For instance, product features, quality of reviews and sentiment polarities of reviews were extracted to select a small set of comprehensive reviews and various greedy algorithms were testified to balance different aspects with information coverage functions [127]. However, some argued that review sampling should consider the distribution of sentiment polarities. For this purpose, an optimization problem was formulated to gain a small set of reviews whose information closeness is smallest. Then, a greedy algorithm, an integer-regression algorithm, and an iterative-random algorithm were testified on a big review dataset [128]. Also, a review sampling strategy was suggested for analyzing corpus with imbalanced opinions by using two carefully designed classifiers [129].

Besides sampling a set of informative sentences or reviews from opinion data, it is also imperative to recommend a small set of reviews according to one's personal preferences. The rater-reviewer affinity and the rater-review affinity were modeled using different factors [130]. Then, an additive function was applied to model the mean rating given by a rater to a particular review. A matrix factorization model, which considered the information about rates and reviews, and a tensor factorization model, which considered the information about raters, reviewers, and products, were reported to recommend reviews for different raters [131]. Based on the tensor factorization model, one extended tensor factorization method, in which the overall rating of a product was incorporated as additional constraint, and another unbiased extended tensor factorization model, which captures the biases from raters/reviewers who tend to give a higher rating, were proposed [132].

2.4.4 Contrastive Viewpoints Mining

Although review summarization and review sampling provide a quick overview on CRs, perhaps to discriminate contrastive customer opinions will provide more valuable insights to both consumers and designers.

Some studies focused on the extraction of opposing viewpoints from online opinions. For instance, an unsupervised learning method was proposed to identify two groups with opposing opinion in a forum [133]. The sentiments of threads were determined by SentiWordnet and relations about agreement and disagreement were inferred by analyzing the consistency within reply-to relations and the consistency of user relations. A cross-perspective topic model was proposed for mining contrastive opinions in political documents [134]. The distance between different perspectives was evaluated by the Jensen-Shannon divergence. Besides, a similar topic model was developed to extract topics and expressions that indicate contention and agreement topics [135].

However, these studies neglect opinions with comparative and superlatives. Accordingly, customer context-dependent opinions in comparatives and superlatives were summarized and a rule-based approach was proposed to identify which entity in a comparative sentence is preferable [136]. Also, an algorithm based on sequential pattern mining with multiple minimum supports was applied on POS tags of review sentences and sentences with a small number of keywords to identify comparative patterns [137]. Xu et al. argued that many of these approaches fail to cover many cases of comparative sentences [138]. Then, a two-level CRFs model was built to identify comparative sentence in online reviews. The first level was utilized to model relations between entities and words, while the second level was to model product relations.

Some studies also targeted to summarize contrastive opinions. A two-stage method was proposed to summarize multiple contrastive viewpoints [139]. In the first stage, an extended LDA model was utilized to extract topics and viewpoints from review text. In the second stage, a modified PageRank method was employed to sample sentences with similar meaning and contrastive meaning. To generate comparative summaries of contradictory opinions, an optimization problem was formulated [140]. In this optimization problem, both content similarity with the same polarity and contrastive similarity with opposite polarities were considered for contrastive opinion summarization.

2.5 Others

Besides aforementioned research tasks, there also exist studies exploring how online big consumer opinion data can be utilized to recommend products, how big consumer affect product sales, etc.

Some researchers argued that consumer preferences may vary and it is necessary to recommend products according to ones' personal tastes. For this purpose, comparative sentences were analyzed to build a customer preference topological network and a PageRank based algorithm was applied for product recommendation [141]. But the problem is, generally, comparative sentences are few in online opinions. Accordingly, McAuley and Leskovec extracted latent topics extracted from online opinions and these topics were later utilized to estimate the correlation between customer preferences and products [142]. Besides latent topics, review quality [143], elapsed time of online reviews [144], customer suggestions [145] and tempo changes of customer preferences [146] were reckoned to recommend products. However, these studies failed to utilize the affluent information about customer similarities and product similarities. Xu et al. grouped users and items to hidden user communities and item groups respectively by a Bayesian Co-clustering approach. A Matrix Factorization model was then applied to recommend products in consideration of hidden item groups and user communities [147]. Similar studies that analyze customer demographic distributions [148] and evaluations from both opinion leaders, friends, and followers [149] were found for product recommendation.

As online reviews generally represent customer opinions and they are publicly available, some researcher guessed that these online opinions might affect product sales. According, different hypotheses were proposed to analyze potential factors of online opinions that might influence product sales [150]. Some exemplary hypotheses include that favorable reviews might raise product sales, reviews from higher quality reviewers might influence product sales, favorable/unfavorable news written by higher exposure reviewers are different from those written by lower exposure reviewers, etc. Then, categories of experiments were conducted to testify the availability of these hypotheses. According to the sentiments in online reviews, two sentiment divergence metrics were proposed to analyze how customer sentiment potentially affect product sales [151]. Similar studies can be also found to utilize online opinions to predict product sales rank [152 - 154].

3 INFORMATION UTILIZATION OF BIG CONSUMER OPINION DATA FOR PRODUCT DESIGN

3.1 VOC ranking

With sufficient VOC are collected, the next phase become how to make big consumer opinion data be utilized for product design. As shown in Figure 1, the first lane in this phase is the ranking of VOC. It essentially points to how to make effective analysis towards the put VOC. Actually, as a widely used tool to analyze VOC for product design, QFD has plentiful applications in the design community and different industry scenarios, from conceptual design to process planning, and from consumer product design to construction project management [155]. In this section, several aspects regarding customer affective need identification, product functional decomposition for conceptual design, the weighting of CRs will be discussed. Also, Kano's Model, which is one widely utilized modeling approach for customer concern classification and weighting, and its relevant applications are also summarized. Finally, in this section, how current studies integrate latent online consumers' opinions into product design will be briefly introduced.

3.1.1 Affective Design

Affective design, also known as Kansei engineering, is an effective methodology for the study of interactions between human affections and design of products to improve consumer satisfaction and has received much attention from both academia and industries. It targets at incorporating affective CRs into design elements that deliver customers' affective satisfaction and helps companies to develop new products that can better satisfy the emotional CRs.

An early outline conceptual framework was proposed for affective design, in which concepts from the Activity Theory was initially borrowed to understand affective CRs, emotion, and sentiment [156]. It motivated discussions regarding the interaction design and HCI (Human-Computer Interaction) communities concerning customer affections. Later, an improved framework was presented to assist the development of emotionally appealing products by eliciting CRs [157]. Several aspects of the Kansei engineering and relevant fields of linguistics, engineering, and psychology in affective design were explored for further functionality and user support. According to these theoretical studies, some prototype systems for affective design were built, i.e., a system for iterative product concept development [158], a system integrating interactions among affective design, engineering, and marketing issues [159].

Actually, before the deep discussion about complex models to gain optimal portfolios of ECs in affective design, designers have to make full understanding about customers' affective responses. For this purpose, a framework which reckoned characteristics of users, tasks, products, and environment was developed and several challenges on the definition about valid and reliable measurements of customer affect were discussed [160]. Hsu et al. collected Kansei adjectives from literature and utilized the Kawakita Jiro method to classify folding bikes into different groups [161]. With this approach, significantly different customers' descriptions on their affections were found in different groups. Similarly, emotion-related physiological responses were explained in the viewpoint of sympathetic and parasympathetic nervous systems [162]. Then, various sensor measurements were invited to capture human emotion reactivity in product design or systems engineering context. Additionally, a neural network-based approach was developed to analyze individual customer' affective responses to products [163]. Based on this model, different models were integrated for the analysis of a group of opinions on the basis of mean market affective responses. Also, based on studies that explored an individual's affective responses in several ways, some scholar made a combination of different affective reaction to analyze consumers' emotional stimuli [164].

Note that the incorporation of affective CRs into affective design aims to optimize customers' affective satisfaction. With several practical constraints, a rule-guided search GA approach was applied to determine the optimal design attribute settings for affective design [165]. Also, relations of customer affections were considered to derive ECs values for optimized affective design [166, 167]. Jiang et al. argued that affective analysis and determination of engineering specifications were often conducted separately, which might induce different settings on the design attributes and engineering requirements [168]. Then, a multi-objective optimization model was formulated, which considered affective analysis and the determination of engineering specifications simultaneously. Besides, a structural equation modeling based approach that considers apparent usability and affective quality [169] and a fuzzy regression method for the analysis of affective quality and fuzziness [170] were reported to optimize customers' affective satisfaction.

Besides, several interesting research problems in affective design also explored, such as the identification and elicitation of affective CRs [171 - 173], the quantification of connection between CRs and ECs [174], management of design information [175], auditory intuitive emotions for the evaluation of products [176], mass personalization [177], influence analysis and evaluation of user experiences in product experience engineering [178 - 180].

3.1.2 Morphology Design

Over the last several decades, the conceptual design has been paid increasing attention by academia and industry and plays an important role in NPD. Specifically, functional decomposition and morphology become the most common conceptual design method and the morphology enables designers to analyze all solutions that occur at the same time.

Generally, the morphological matrix, as one a popular tool for conceptual design, is welcome in many approaches for NPD. For instance, a formal mathematical framework that integrates the morphological matrix in a computerized conceptual design was proposed [181]. In this framework, the matrix was quantified, which associated each solution principle with a set of characteristics, and an optimization problem was formulated for the selection of individual solutions. Also, the morphological matrix was utilized to select most appropriate solution on the manipulator travel frame design projection [182]. Also, based on space matrix, an integrated model of function repository and solution repository was built. Next, according to the built space matrix, a computational conceptual design process was proposed, which includes design synthesis algorithm based on space matrix, feature matching based on physical parameter, solutions constraints matching based on design catalog, functional structure design based on functional structure evolutionary model, and evaluation and selection approach based on design catalog [183]. However, Ma et al. argued that quantitative evaluations to each function solution and subjective evaluation uncertainties were seldom considered in a morphological matrix based conceptual design approaches, which induce potentially difficult to obtain the optimal conceptual design by combining various function solution principles [184]. Then, customer preferences with subjective uncertainties and the information of product failure data were utilized to evaluate function solution principles and, based on a fuzzy morphological matrix approach, a fuzzy multi-objective optimization model was developed for conceptual design.

Nonetheless, some criticisms appear toward conventional systematic design regarding functional decomposition and morphology, such as difficulty in unanimous function decomposition, poor diffusion in industry, etc. Then, scholars innovate different approaches to make improvements in current studies or supply some complementary solutions. For instance, it was found difficult to make automated functional decomposition since there is no consensus on the concept of "function" [185]. Then, an automated functional decomposition method was developed, in which a hierarchical material structure representation and a hierarchical shape graph were utilized to model the morphological changes of material flows. Some argued that many conventional design evaluation methods do not support multiple outcomes. Accordingly, a Bayesian Belief Network (BBN) was utilized to learn relations between design variables from previous morphological design chart and a user interface was developed for the dynamical search in the conceptual design [186]. Based on QFD, a 3D morphological chart which integrated CRs was proposed for design variance of simple and technically mature products [187]. It was argued that such 3D assembly facilitated marketing, design, and manufacturing teams for better visual stimuli. Besides, some new solutions were proposed as complements to those methods utilize functional decomposition and morphology. For instance, a new conceptual design approach that focuses on fundamental logic and main tools [188] and a parameter analysis approach which helps to identify dominant conceptual level issues and relations repeatedly [189] were reported.

Also, studies in recent years introduced biomimetic design for concept design. To provide a better representation model of objects in geometric model-based design, according to on the set theory and mathematical morphology, the machining process was integrated for designing objects and an analogy between the design and machining processes was established to provide a generic and robust geometry model [190]. In addition, it was found that design methods in the morphological domain fail to examine whether a solution is optimal. Then, a new approach based on systematic biology and evolutionary biology was suggested for the verification and validation [191]. Besides, adaptive morphology was introduced for robots design to reduce tradeoffs during locomotion. It helps to provide new functional materials and structures and allow to accommodate according to opposing dynamic requirements [192].

3.1.3 Technical Blueprint

In QFD research and development, the identification of CRs and their relative weights are probably the first focus since these tasks may affect the selection of optimal target values of ECs.

A hierarchical approach was introduced to extract CRs [193], which includes the verbatim construct, the superordinate construct, and the imposed construct. Then, an ART2 neural network was built to analyze customer segmentation and to conduct the market analysis. For the prioritization of CRs, the simplest method may be based on a numeric scale, e.g., from one to ten, where ten often means something like the indispensable ones [194]. Due to a substantial amount of human efforts involved in VOC interpretation, a conjoint analysis method was put forward to compare VOC in a pairwise manner for their relative weights [195]. Han et al. designed a linear partial ordering approach to obtain precision prioritization information about VOC [196]. It was argued to be capable to reduce the burden on the extraction of VOC weights and ECs' relations. However, CRs might change along the time. Accordingly, the grey theory was introduced into QFD to monitor the changes of VOC importance [197]. Similarly, a Markov chain model was integrated into QFD to analyze changes of VOC weights [198]. Some researchers also argued that the determination of the importance of VOC should consider not only the fulfillment degree of CRs but also the competitive products. Accordingly, a method that considering competitive products, current performance, and customer satisfaction was proposed to determine the importance of CRs [199].

To calculate the relative importance of VOC and to handle design concept variations, the analytic hierarchy process (AHP) has been introduced. AHP was originally developed for the sake of resource allocation and planning [200]. In product design, AHP basically establishes a design concept hierarchy with prioritized subordinates. Then AHP decomposes the linguistic-based CRs into different levels of subordinates and alternatives according to their correlations. Due to the mathematical rigor in AHP framework that facilitates prioritization, incorporating AHP into QFD for NPD provided many promising results to improve the conventional QFD limitation. Armacost et al. developed a framework for prioritizing VOC in QFD to improve industrialized housing design and manufacturing process [201]. An integration of AHP and QFD was also reported for location planning under some practical requirements [202]. A framework that incorporated fuzzy set and AHP was shown to prioritize VOC in target planning for QFD [203]. Then, an example from automotive product development was illustrated to verify the availability of this framework. Fung combined QFD, AHP and fuzzy set theory in a hybrid system to measure and prioritize the imprecise VOC [204]. Wang et al. compared the prioritization matrix method and AHP on several factors, namely, time, cost, difficulty, and accuracy [205]. They concluded that if time, cost and difficulty are the major concerns in product improvement, the prioritization matrix method is preferred, while where accuracy is the major requirement, the AHP method would be a better choice.

3.1.4 Kano's Model

Some researchers associated Kano's Model with QFD to obtain and understand CRs. Kano's Model is a useful tool for understanding CRs and their impacts on customer satisfaction. In Kano's Model, different CRs are categorized based on how well they are able to affect customer satisfaction. CRs are distinguished as must-be attributes, one-dimensional attributes, attractive attributes and indifferent attributes.

Matzler and Hinterhuber categorized CRs into different groups and evaluated their importance based on Kano's Model. This categorization of CRs was then utilized in QFD for NPD [206]. Shen et al. also took customer attributes analyzed by Kano's Model as the input of planning matrix in QFD to help designers for better CRs understanding [207]. But these efforts are qualitative to combine Kano's Model into QFD with little quantitative analysis. Lai et al. utilized a customer survey to estimate customer satisfaction and customer dissatisfaction values [208]. Different from existing studies, quantity values were utilized to integrate Kano's Model into the QFD by establishing a mathematical programming model to optimize product design. Later, Mu et al. proposed to bring Kano's Model into QFD to quantify CRs in an uncertain and vague environment [209]. A fuzzy multi-objective model was then suggested to balance between customer satisfaction and development cost. Kwong, Wong, and Chan utilized a neuro-fuzzy approach to generate a customer satisfaction model [210]. A concrete example was given to demonstrate that their model was better than a statistical regression approach. Some other researchers also proposed a method that integrated Kano's Model with QFD to recognize the importance of an attribute for customer satisfaction maximization [211, 212]. Ji et al. integrated Kano's Model with QFD in consideration of both discrete variables and continuous variables and an integer optimization problem was formulated to maximize customer satisfaction [212]. In addition, corporate decision making and engineering decision making were integrated for customer-driven quality improvement efforts [213]. Especially, potential attributes were classified and prioritized for further improvement using Kano's Model and QFD based on rigorous business analysis and trade-off studies.

Chen and Chuang present a robust design approach to achieve a higher level of customer satisfaction in aesthetic qualities [214]. In such robust design approach, grey relational analysis with the Taguchi method was proposed to optimize subjective qualities with multiple-criteria characteristics. Then, the Kano's model was employed to balance the weights of multiple-criteria for designers to understand relations between performance criteria and customer satisfaction. To decide weights of multiple-criteria, some studies applied regression methods with dummy variables to recognize critical attributes, but such kind of methods might lead to an inaccurate classification of multiple-criteria in some specific condition [215]. Accordingly, a moderated regression approach was proposed to improve the performance of the dummy regression method with dummy variables, which aims to produce more accurate attribute classification for Kano's Model. A similar study was conducted to measure and quantify the relationships between customer satisfaction and the fulfillment of customer requirements in Kano's Model [216].

3.1.5 Latent Online Users

As mentioned in previous sections, online reviews and social network sites provide valuable information about CRs. Many studies began to explore how to extract latent CRs from online users.

To invite online users for product design, the first task becomes how to obtain CRs efficiently. A text summarization approach was presented to assemble CRs from product reviews [217]. Then, a ranking strategy was proposed to find critical CRs. A Bayesian sampling approach was proposed to extract CRs from online reviews [218]. In this approach, critical terms were sampled from CRs related items. Then, the term disambiguation problem and the keyword recognition problem were investigated to maximize the overall performance of the feature identification. Tuarob and Tucker tried to learn CRs from tweets and they define the product favorability as the product of polarity, subjectivity, and popularity [219]. Then, the product favorability was utilized to select the most and the least favorite products. To obtain sufficient CRs and present an overall picture of a specific product, online opinions in multiple sources were reckoned. For instance, an unsupervised probabilistic graph model was proposed to extract and normalize attributes jointly from multiple web pages, which tackled both page-independent content information and page-dependent layout information of the text fragments in Web pages [220]. Four different types of online opinions, including blog postings, discussion board threads, user reviews and critic reviews, were compared in terms of POS distribution, information gain of different POS in different types and other information evaluation metrics [221].

A deep learning model based on an improved stacked denoising auto-encoder on sentence-level features was proposed. It aims to extract relations among various named entities (e.g., enterprises, products, demands, and capabilities) underlying the text-based context and such social interaction context is expected to be utilized to provide more information for cross-enterprise manufacturing demand-capability matchmaking [222]. Also, a pointwise mutual information based method and an unsupervised learning method were employ to identify relations between topic terms that were identified from multiple sources and then the topic hierarchy was generated according to the identified relations [223]. Besides, based on customer online opinions, a pairwise ranking approach and an integer programming model were proposed to prioritize CRs according to customer satisfaction [224].

Note that, one straightforward approach to describe customer preferences is to profile each consumer with some informative labels. For this purpose, some innovated models were present to generate labels to characterize customer demographic data. For example, a logistic regression based framework was proposed to predict the labels of users in social media, in which multiple relations in social network were reckoned in a semi-supervised method [225]. Similarly, categories in news media were projected into short texts in social media to model user's interests and both articles and relevant categories in Wikipedia were employed to reduce the semantic gap between social media and news media [226]. Also, customer demographic data were learned from purchased products [227]. Especially, customer profiling was formulated as a multi-task multi-class problem and a structured neural embedding model was proposed to learn the representations of products. However, consumers might present different interests on various product aspects. To cluster consumers with similar interests, a permutation-based structural topic model was proposed [228]. Using this model, the frequency of different product aspects and the occurrence ordering were presented. Some researchers also proposed a method to identify consumer clusters and corresponding opinion leaders within the specific consumer cluster [229].

Additionally, many studies talk about how products are diffused in the social network. Social influences from different topics were estimated by a factor graph model and an efficient affinity propagation algorithm was proposed to analyze latent associations between topics [230]. A decision-making process in E-commerce was analyzed by considering the social influence of online reviews [231]. Also, two posterior evaluation models were proposed to check whether online ratings were independent of others' recommendations [232]. In this study, influential friends were found by considering social positions of users in the friendship network and their personal characteristics were independent of other individuals. According to different assumptions of social interaction behavior, different models were proposed and compared to measure the social influence that one may receive from his or her friends [233]. It was found that a member's ego-centric network should be measured by a model that considers both the frequency of interaction and friends' evaluation. Similarly, the amount of influence was found to be moderated by both recipients' perception of their opinion leaders and the sources' volume of product usage [234]. Besides, both sociometric and self-reported measures of opinion leadership were found to be weakly correlated with different kinds of adoption-related behaviors.

3.2 Connect VOC to Engineering Design

The next critical step in the phase that makes use of big consumer opinion data is how to integrate selected important VOC into product design. The second lane of this phase in Figure 1 list some highly relevant studies regarding this topic, which include how to connect selected important VOC, how design concepts are generated, how design ideas are generated, etc.

3.2.1 General Idea

In customer-driven product design, after successfully identifying CRs, designers start to consider how to interpret CRs to improve their products. Especially, how to connect CRs with ECs is one important question in QFD. Several contributions are made visible in this area.

Generally, in the design area, studies about connecting CRs with ECs have to cope with the inherent vagueness of human language and subjective judgment in VOC [235]. This problem is often seen to be analyzed by introducing the fuzzy set theory [236, 237] or the rough set theory [238] into QFD. For instance, Harding et al. claimed that market-driven strategies encourage enterprises to launch products that customers want to buy [236]. Then, to meet CRs and facilitate information sharing between members of extended design teams, they developed a market-driven design system based on fuzzy logic for the interpretation of market information into product specifications. Fung et al. integrated the idea of the least squares regression into fuzzy linear regression and proposed an asymmetric fuzzy linear regression method [237]. This fuzzy linear regression method was utilized to estimate the uncertainty in the functional relations between CRs and ECs for product planning, which is one of an important process in NPD based on QFD. In their later research, a fuzzy expert system was proposed to identify important ECs [239]. In this fuzzy expert system, both the importance of ECs and their mutual impacts were considered in a fuzzy environment using QFD and the fuzzy relation measures between CRs and ECs were estimated. Similarly, a fuzzy regression-based approach [240] and a novel multi-objective genetic algorithm (GA) based rule-mining method [241] were developed for analyzing the connection between CRs and ECs. Besides, both past sales records and product specifications [242], creative thinking process [243] and customer online concerns [244] were taken into considerations for the connection between CRs and ECs. In addition, to balance affective and engineering concerns, a hybrid association mining and a refinement mechanism were applied to support affective mapping decisions [245]. Specifically, the rough set and a *K* optimal rule discovery approach were applied to identify hidden relations underlying forward affective mapping. Next, based on conjoint analysis, a weighted ordinal logistic regression was derived to model backward affective mapping in consideration of affective quality.

Linguistic variables were also found to be more appropriate to describe the inputs of QFD [246]. The method using linguistic variables is different from the previous efforts where the input data were assumed to be precise and treated as numerical data only [194, 195, 247]. But linguistic variables were found sometimes difficult to be handled for the subjective assessments [248]. To ease this problem, an integrated linguistic-based group decision-making approach was proposed to cope with multiple types and multi-granularity linguistic assessments given by multiple decision-makers in QFD planning. This approach processes words in CRs directly and minimizes the risk of loss of information, without translating linguistic information into various fuzzy numbers. In an uncertain and vague environment, Kano's Model was also reported to be integrated into QFD to quantify CRs [209]. A fuzzy multi-objective model was reported to be utilized to balance customer satisfaction and development cost. Recently, Lan, Liu and Lu developed a deep belief net based approach to discover design tasks and quantify their interactions from different design document archives [249]. The proposed model was utilized to discover design tasks by unfolding hidden units by sets of strongly connected words and estimate interactions among tasks on the basis of their co-occurrence frequency in a hidden topic space.

3.2.2 Concept Generation

Product conceptualization is regarded as a key activity in NPD and generally, it involves concept generation and evaluation, which play a crucial role in early stage of design process. In recent years, much research into the development of these two tasks has looked at the concept generation stage.

A process of repeated steps regarding both concept generation and conception evaluation was utilized for conceptual design [250]. In this research, different levels of solution abstraction and tasks at each level were described and extended solutions were presented which include kind synthesis, spatial synthesis, and physical synthesis. According to this study, the concept generation was regarded to involve both analysis and synthesis activities. Then, the concept generation process was analyzed by comparing it with a linguistic interpretation process and the concept synthesis process was examined by recognizing concept blending and non-alignable differences [251]. However, these two activities were argued to be interchangeable rather than independent [252]. Accordingly, a new method was proposed to treat concept generation as a proposition-making process and adapt the formal logic definitions of analytic and synthetic propositions to generate new concepts.

Design concept evaluation was often argued to one important step, which is believed to influence the later stages of concept design and the success of the whole design solutions. It hence attracts many researchers to innovate various approaches analyzing technical details about concept evaluation. Rondini et al. complained that approaches regarding Product-Service System (PSS) design only discussed requirement generation and identification for service design and only one or two phases were analyzed for concept development and evaluation [253]. Then, a Product-Service concept tree was designed to identify and evaluate PSS. Also, AHP based approach [254] and GA based approach [255] were reported for concept evaluation. But some argued that many studies regarding concept evaluation do not consider to satisfy design constraints and maximize customers' preferences [256]. Accordingly, a two-step approach was introduced to evaluate the best concept, which includes relative importance ranking of design criteria and elicitation of customers' preferences in the form of rough numbers.

There also some studies analyzing concept generation and evaluation at the same time. General sorting techniques were adapted for initial requirements acquisition and platform definition [257]. Then, a fuzzy c-Means algorithm was employed to cluster design options and select preferred product concepts. Similarly, a group of concepts was generated by a GA based approach and the concept evaluation was implemented using a fuzzy neural network to obtain an optimal concept [258]. Combined with the theory of inventive problem solving (TRIZ) methodology, a framework that utilized fuzzy linguistic evaluation was proposed to obtain design concepts from knowledge domain and to evaluate alternative concepts for the determination of promising product concepts [259].

Some studies investigate sources for concept generation, e.g. design library, design patents, customer opinions, cloud sourcing, etc. Conventionally, conceptual solutions were generated from design library or design patents. For instance, a computational design tool was developed to create conceptual solutions to detailed functional specifications by expanding online design library in the form of procedural rules [260]. Liang et al. built a text analysis approach to discover design rationale from design patents for concept generation [261]. In particular, a semantic sentence graph was built to model sentence relations and a manifold ranking algorithm was utilized to highlight issue related sentences. Similarly, the techniques of natural language processing was also applied to conduct function dividing process in conceptual design. For instance, Yamamoto et al. extracted linguistic hierarchal structures to highlight hierarchal relationships between the upper- and lower-level functions [262].

A Convolutional Neural Network model was innovated to quantify the ability of a digital design concept to perform a function based on 3D convolutions that predict functional quantities of digital design concepts [263]. Another computational framework was described to measure the novelty, feasibility, and diversity of design concepts based on design patents [264]. Besides these conventional sources to extract design concepts, customer opinions were reckoned. It was found that a group of customers often present certain preferences that are related to the same product [265]. Customer descriptions and design requirements were regarded as higher level concepts and lower level concepts respectively. Then, customer preferences were translated to design domain ontologies for concept generation from customer preferences. Similarly, customer opinions in customer center were transformed to CRs. Then, the new product specification was built according to different groups of CRs [266]. Also, online customer opinions were reported to be utilized for design concept generation. According to customer online opinions, a design framework was built to abstract relevant design concepts and build a database of logic propositions [267] and an optimization problem was formulated for design concept selection to maximize potential profit by considering both design constraints and hierarchical customer preferences [268]. Moreover, some researchers explained that several practical challenges hinder the crowdsourcing to be widely utilized to acquire innovative concepts [269]. Then, a concept re-construction module for concept selection and evaluation was designed to organize word tokens, which were extracted from online crowd-sourced concepts. The module was later integrated into a unified frame using domain ontology and extended design knowledge.

There also exist studies conducting experiments to promote new approaches to visualize concepts [270, 271], to capture and classify concepts in maintenance process [272], to examine the effectiveness of a committee on conception generation and selection [273, 274], to make eco-related and sustainability-related decisions within the conceptual design phase [275], etc.

3.2.3 Idea Generation

The ability to develop innovative new products can be a source of competitive advantage for any firm. The generation of ideas for new products, as the very first step to promote the development of innovative products, is often a critical and creative activity for both management practitioners and scholars. In this area, much research efforts have been put on the process development of idea generation and evaluation, the sources of idea generation, factors on idea generation, etc.

A system that facilitates industrial designers' divergent thinking process was built [276]. It aims to avoid the dead end on generating fresh ideas and generate new concepts from the intersection of different groups of designer's idea sketches. Also, a system was presented to support idea generation in PSS [277]. In such system, designers' acquisition of new design solutions was provided by case-based knowledge base and ideas were evaluated considering both customer satisfaction and resource constraints. Based on the techniques of both Axiomatic Design and TRIZ, a hybrid model of the problem-solving was created for innovative product design by integrating problem analysis and idea generation approaches into the conceptual design stage [278]. However, these studies were criticized since incremental innovation is the only focus [279]. Then, a creative idea generation framework, including future envision, opportunity identification and analysis, idea generation, idea expansion and ideation control, was developed for the fuzzy front end of radical innovation from the user experience perspective. Besides, preferences about the last group of customers that adopt a product were integrated as an important source for innovative products and services and a specially designed approach was developed [280]. They advocated that such method will increase such group of customers' perception of different aspects and help radical or incremental innovations in NPD.

Generally, there exist several sources for idea generalization, such as internal perspectives from managers, discussions in meetings, opinions from social network sites, etc. For the effectiveness of idea generation meetings, a model that considers functions of sketching activities as interactions with the group's external memory was introduced and it was found that supporting a re-interpretive cycle in the individual thinking process and enhancing access to earlier ideas were more helpful [281]. For the effectiveness of idea generation team, a research study was conducted to investigate the specialization and diverse expertise of the idea generation team as well as goal constraint on the generation of new product ideas [282]. A deep learning model was to cluster multiple concepts for product development, which are designed by individual team members and descripted by natural language [283]. The model is expected to support design teams in identifying possible areas of "over-clustering" or "under-clustering" in order to enhance divergent concept generation processes. Zahay et al. suggested that online crowdsourcing will provide insights for idea generation on product line extension and help to improve the process of NPD [284]. Simon and Tellier argued that few studies were conducted regarding factors that encourage actors to shape social networks during the development of new ideas [285]. Then, a qualitative analysis was conducted and four factors explaining why actors turn to others during the idea-development process were identified. In addition, an exploratory survey was conducted to examine the profitability of different sources of new product ideas that are currently used by companies in the support of creativity [286].

Also, some research studies investigate how creative ideas are generated. For instance, an experiment was conducted to understand how newly acquired information was accepted and integrated into a design problem [287]. It was found that information that is more distantly related to the design problem may affect idea generation more if it is a new problem with an open goal, while information that is more obviously similar to the problem may impact idea generation more if it is an old one. Similarly, how creative ideas are promoted or filtered throughout the design process was examined for education on students major in engineering design [288]. Besides, four knowledge creation modes of socialization, externalization, combination, and internalization [289 - 292] were integrated to product idea generation [293]. It was found that socialization and internalization present positive relations with product idea generation.

Relevant studies regarding idea generation also include how to utilize semantics-based strategies and structure-based strategies on creative analogy retrieval for NPD [294], how to build different simulation analysis models and merge it into the existing prototype according to knowledge structures [295], how to develop metrics to evaluate the effectiveness of ideas [296], how different external representations in engineering design influence design fixation [297], etc.

3.3 Comparisons and Ranking

In a competitive market, many companies offer a variety of products to compete for market shares in different segments. Due to the rich information about competitive products is widely available, both designers and consumers face the challenges to understand pros and cons. It triggers the interests of many researchers to explore how to make effective comparisons or provide a ranking of different alternatives. Accordingly, as presented in Figure 1, the third lane of the phase that focusing on information utilization of big consumer opinion data for product design become product comparison and product ranking.

Some researchers employ product descriptions for comparisons and ranking. A set of product descriptions in natural language were utilized to build a product comparison matrix, which provides a condensed view regarding a family of products [298]. To support ontology development in design engineering, an information search and retrieval framework was built based on the semantically annotated multi-facet product family ontology [299]. In such framework, a document profile model was described to suggest semantic tags for annotation purpose and both faceted search and retrieval of product information were implemented for product ranking. Similarly, a general approach to assess product semantics was presented, which aims to help designers understand some semantic parts and rank product prototypes [300]. Also, a deep learning approach was developed to construct an unsupervised learning ontology network for discovering the various associations between individual knowledge concepts, in which the subsequent probability and velocity network analysis with different statistical behaviors were applied to evaluate the correlation degree between concepts for design information retrieval [301].

Besides, some scholars invite customer concerns for product comparison and recommendation. For instance, according to the Kansei studies, a Personal construct theory (PCT) based product configuration analysis method was proposed to extract customer emotion-related product attributes and a means-value chain was used to generate labels of comparable targets for competitor analysis [302]. But this method was found to fail to classify products based on the quantitative and qualitative characteristics [303]. Then, an optimization problem was formulated to cluster products according to market basket data. However, many studies neglect sustainability and environment-related criteria. Then, the concept about environment-related ECs was described and an environment-related comparison of products was conducted from the point of view of both manufacturers and consumers [304]. Some studies also made product ranking and recommendation from product feature level in the perspective of consumers. For instance, multi-format preference information was utilized in an integrated approach, which incorporates group decision-making, multi-format preference analyses, and three types of least square models, to achieve a higher level of higher customer satisfaction [305]. In addition, a fuzzy cognitive pairwise comparison method was built to evaluate consumer preferences for multiple features and techniques of fuzzy grading clustering were employed to group the product alternatives into different consumer preference grades [306]. Similarly, considering multiple criteria and alternatives, a primitive cognitive network was developed to build an effective product ranking strategy [307].

However, to obtain sufficient customer opinions is sometimes time-consuming. Note that, as aforementioned above, a large volume of online reviews provide valuable customer opinions, which induces that such online information was introduced to make product comparisons. In some earlier studies, comparative sentences only were analyzed for product comparisons and ranking. For instance, product names were extracted from online reviews by the techniques of CRFs and two applications were presented for product comparison [308]. A graph propagation method was proposed to compare products in considerations of online reviews and community-based question answering [309]. In this method, comparative sentences were first extracted from online reviews and the number of preference between products was then utilized to build a product comparison graph. Later, a product comparison network was reported by exploiting comparative sentences in online reviews [310]. Three different types of graphs were built according to the overall consumer sentiments and different regression models were utilized to analyze the factors that influence the product rank. According to [310], another product comparison network was built [311]. In this network, transitive sentiment, rather than averaged sentiment or overall opinions, was utilized to analyze the network influence. However, comparative sentences only account for a small proportion of online opinions. Then, a system was built to identify product weakness from online reviews and such information was analyzed for product comparison [312]. Also, according to the identified sentiment orientations, an intuitionistic fuzzy number was constructed to represent the performance of an alternative product at the feature level. Then, techniques of an intuitionistic fuzzy weighted averaging operator and preference ranking organization methods for enrichment evaluations II were borrowed for product comparisons [313]. However, it was arguably important to predict the rank of products in the near future. Then, affinity rank history, average ratings, and affinity evolution distance were extracted from online reviews and the AutoRegressive model with exogenous inputs was applied to rank products [314]. Besides, a digraph structure was built according to both descriptive online opinions and comparative online opinions, from which an overall eWOM score for each product and a ranking of products were derived [315].

3.4 User Needs Evolution

The last lane in Figure 1 refer to studies that concerns user needs evolution. Actually, the success of customer-driven products is highly dependent on whether CRs are satisfied. However, nowadays, enterprises have to capture the fast evolution of customer tastes and make the proper adjustment to respond consumers' requirements. Generally, to provide valuable information for enterprises, two sub-problems were widely discussed in the academic field, e.g. how to make corresponding engineering changes corresponding to the changes of CRs and how to make an effective prediction regarding future CRs.

For the first problem, different levels of changes in production planning were analyzed, such as changes in product family design, product line design, product configuration, etc. In some early studies considering CRs changes, to formalizing front-end processes was suggested for efficient understanding on CRs to cope with continuous changing markets. It was argued that such approach helps to and result in consistently more successful new products in NPD [316]. According to discussions on a multi-domain transmission mode of dynamic requirements, a product family flexible design method was proposed to adjust design parameters based on the uncertainty analysis of market demand changes in the future for customer satisfaction and mass customization manufacturing [317]. Based on the fitness evaluation of a product line from marketing and engineering, a multi-objective optimization problem was built for product line design to satisfy changing CRs and maintaining commonality in product platforms [318]. To meet changes in both CRs and new technologies, manufacturing companies have to pay attention to production planning and control in the planning and control of engineering changes in manufacturing systems. For this purpose, approaches to implementing production planning and control were compared on different tasks [319]. A three-phase evaluation model, that incorporates fuzzy theory, value engineering, and multi-criterion, was developed to find optimal strategies for product configuration changes and combination selection on product component suppliers [320]. Besides, to reduce the total process time for engineering changes in the complex product development, a Monte Carlo based simulation algorithm was proposed to select the most economic propagation path for design change in a practical product development process which involve multiple design tasks and different relationships among these tasks [321]. A similar study was conducted using process simulations to generate possible modes that analyze factors on the life cycle of the newly designed products for the forecast about the success of new product configuration [322].

For the second problem, a large volume of consumer preferences was collected to analyze the dynamic change pattern and make a proper forecast. A CRs analysis and forecast (CRAF) system that provide product development functions with quantitative and qualitative CRs information was developed to forecast dynamic CRs and lower the risk in NPD for fast changing markets [323]. A time series exponential smoothing technique was employed to forecast future attribute trend patterns and [324]. Then, a CRs demand model that reflects emerging product preferences and a classification approach that helps to identify attributes that have low predictive power was developed. Similarly, a new algorithm combining decision tree for large-scale data, discrete choice analysis for demand modeling, and automatic time series forecasting for trend analysis was proposed to capture hidden and upcoming trends of product demand [325].

However, Guo et al. argued that CRs should be classified to easily obtained CRs, predictable CRs and unpredictable CRs [326]. Then, unpredictable CRs were analyzed by introducing the core idea of design-driven innovation and such problem was formulated as a problem that CRs generation was triggered by resources variation in the super system. Besides, online opinions were reported to be utilized to predict product preference design trends [327] and to monitor changes in customer opinions [328].

4 RESEARCH TRENDS, CHALLENGES AND FUTURE STUDIES

As pointed out by MIT Technology Review [329 - 331], we are experiencing a revolution stage in business where new ways of collecting, analyzing, and organizing consumer opinion data are emerging, because innovative information technologies are helping marketing managers and product designers to make well-informed decisions, and to meet their desires for economic efficiency and customer satisfaction. However, it is at an early stage of the revolution despite being an obsession of valuable consumer opinion data in NPD. To really see what is happening, critical evaluation should be summarized based on what has been achieved and what remains far from being solved with respect to the exploitation and integration of such data in product and service design. Next, several challenges and open problems are highlighted to signify the importance of future studies in this trendy research thrust.

(1) In order to promote more competitive products, companies often make great efforts to investigate customer and customer behaviors, obtain and analyze their feedbacks, and in turn, provide effective responses in product offerings. Though a number of algorithms to extract customer concerns were developed in engineering design, only a small number of formatted high-quality customer opinions and some specific design problems were focused. Nonetheless, to constantly fuel innovative product design, creative design platforms that facilitate designers to conduct in-depth analysis on a large amount of geographically distributed CRs in a more systematic manner are vital for NPD in a fierce global market. As noted, in many e-commerce websites, consumers can spend hours and hours browsing hundreds of products, examining comparable alternatives, making final purchase decision, and posting a large quantity of opinions. In 2017 alone, Alibaba had recorded USD 25.4 billion of gross merchandise volume in Chinese singles' day and, according to a survey released by the China Internet Watch, more than 20M reviews and comments were posted every day in June 2016. To exploit the value from such a huge volume of online opinions, undoubtedly big consumer opinion data, for NPD will empower any companies to spot business opportunities ahead of others and launch desirable product and services accordingly. While designers in multinational corporations (MNCs) may have acquired the ability with coordinated resources to build complex IT systems for such a purpose, it surpasses the capacity of most small and medium enterprises (SMEs) to embark such a journey towards NPD owing to various economic and technical issues that SMEs may never be able to reach, although the appearance of such online customer opinions makes it feasible for SME to obtain sufficient critical CRs. Therefore, it invites more agile and can-be-tailor-made efficient algorithms and tools that can penetrate such technical and practical barriers, and cloud-based IT platform development that can actively engage SMEs if they wish to take advantage of such big opinion data and initiatives of data analytics.

(2) While many state-of-the-art algorithms are able to parse a large set of customer opinions and conduct analysis in order to support various decision makings, however, as aforementioned, such customer opinions are constantly being generated, which impacts customer's perception and understanding gained over the targeted products. It challenges the current studies and practice that take such opinion data offline and treat it in a static manner which fails to address its dynamic nature. This is fundamentally different from the conventional study where research data collected from customer surveys are used. Besides, outbreak news concerning product faults and malfunctions may suddenly burst into the public, and rumors or insights may be leaked in social nets. These messages can be propagated at an unprecedented speed nowadays to significantly affect both existing and potential consumers upon their perceptions over the products as well as their decisions in favoring them or not. In the era of Big Data, this further means that not only the volume of such opinion data matters, but its incremental velocity and dynamic nature should be taken into account during in-depth analysis, which is largely missing in the literature. Indeed, efficient algorithms and analytical models on deviation detection of customer dynamic opinions, rather than sentiment analysis and opinion identification alone that are discussed in existing studies in the field of computer science, are bound to aid designers in dealing with the dynamic nature of CRs and respond swiftly in anticipating the arrival of further expectations. Predictive analytics on customer insight discovery and management based on perhaps stochastic dynamic optimization may become more relevant in tackling such practical concerns and scenarios. To be more specific, it requires imbuing algorithms with the ability to make better use of background knowledge with respect to the business world where customers gain their understanding from, to model the dynamic nature of CRs, to forecast the emerging CRs, and to identify anomalous CRs by examining big opinion data in NPD.

(3) While it has become a general practice in many market-driven companies to collect customer feedback and then design and develop products and services according to the analytical results derived from such customer inputs, however, as pointed out in Kano's model, attributes that excite consumers are often difficult to foresee. This dilemma has led to a whole set of design research that focuses on customer understanding, design optimization, and so on, which attempt to address some interesting but challenging issues, e.g., whether it is always necessary to follow CRs in NPD; whether offering prompt responses to consumers will lead to a higher level of satisfaction; etc. To answering these questions will enrich our understanding of several fundamental principles established in the market-driven product design.

With the arrival of Big Data, IoT, smart mobile devices, etc, it has become possible that many more varieties of data and information concerning customer behavior, cognitive aspect and use context can be made available, which is hard to be obtained in the past. Therefore, it is crucial for the design community to further leverage design inputs from CR-based to UX-centric (user experience) in NPD, such as customer preferences, product usage context, use cases, customer emotional responses and their interconnected relations. Furthermore, to explore UX in NPD, several inspiring relevant issues need to be carefully dealt with first, e.g., how to extract UX factors as well as their interrelations from big consumer opinion data for NPD; how to build a conceptual design model that connects various emotional factors; how to build up UX related knowledge base in the context of big consumer opinion data, etc. Related studies on this dimension have just been kicked off. It is expected that the insights uncovered through such efforts may one day significantly update our understanding on product design.

ACKNOWLEDGEMENT
The authors would like to express their gratitude to a number of academics who have been the collaborators or served as the external advisors on this topic since 2003/4. They are: Prof. Han Tong Loh, Deputy President (Academic) & Provost, Singapore Institute of Technology (formerly with the Department of Mechanical Engineering, National University of Singapore); Prof. Ee-Peng Lim, School of Information Systems, Singapore Management University (formerly with the School of Computer Science and Engineering, Nanyang Technological University); Prof. Jenny Harding, Wolfson School of Mechanical, Electrical and Manufacturing Engineering, Loughborough University; Prof. Bing Liu, Department of Computer Science, University of Illinois at Chicago; Prof. Aixin Sun, School of Computer Science and Engineering, Nanyang Technological University; Prof. Chris McMahon, Engineering Design and Product Development, Technical University of Denmark (formerly with the University of Bristol and I*d*MRC at the University of Bath); Prof. Steve Culley, Department of Mechanical Engineering, University of Bath; Prof. Karthik Ramani, School of Mechanical Engineering, Purdue University; Prof. Wei Chen, Department of Mechanical Engineering, Northwestern University; Prof. Shuming Gao, the State Key Lab of CAD&CG, Zhejiang University; and Prof. Chunping Li, School of Software, Tsinghua University.

REFERENCE
[1] Otterbacher, Jahna. 2009. "Helpfulness in Online Communities: A Measure of Message Quality." In *CHI'09*, 955–64. Boston, MA, USA.
[2] Kim, Soo-Min, Patrick Pantel, Tim Chklovski, and Marco Pennacchiotti. 2006. "Automatically Assessing Review Helpfulness." In *EMNLP'06*, 423–30. Sydney, Australia.
[3] Ghose, Anindya, and Panagiotis G. Ipeirotis. 2007. "Designing Novel Review Ranking Systems: Predicting the Usefulness and Impact of Reviews." In *ICEC'07*, 303–10. Minneapolis, MN, USA.
[4] Liu, Yang, Xiangji Huang, Aijun An, and Xiaohui Yu. 2008a. "HelpMeter: A Nonlinear Model for Predicting the Helpfulness of Online Reviews." In *WI-IAT'08*, 793–96.
[5] Liu, Yang, Xiangji Huang, Aijun An, and Xiaohui Yu. 2008b. "Modeling and Predicting the Helpfulness of Online Reviews." In *ICDM'08*, 443–52.
[6] Zhang, Richong, and Thomas Tran. 2008. "An Entropy-Based Model for Discovering the Usefulness of Online Product Reviews." In *WI-IAT'08*, 759–62.
[7] Danescu-Niculescu-Mizil, Cristian, Gueorgi Kossinets, Jon Kleinberg, and Lillian Lee. 2009. "How Opinions Are Received by Online Communities: A Case Study on Amazon.com Helpfulness Votes." In *WWW'09*, 141–50. Madrid, Spain.
[8] Miao, Qingliang, Qiudan Li, and Ruwei Dai. 2009. "AMAZING: A Sentiment Mining and Retrieval System." *Expert Systems with Applications* 36 (3):7192–8.
[9] O'Mahony, Michael P., and Barry Smyth. 2009. "Learning to Recommend Helpful Hotel Reviews." In *RecSys'09*, 305–8. New York, USA.
[10] Zhang, Richong, and Thomas Tran. 2011. "An Information Gain-Based Approach for Recommending Useful Product Reviews." *Knowledge and Information Systems* 26 (3):419–34.
[11] Hong, Yu, Jun Lu, Jianmin Yao, Qiaoming Zhu, and Guodong Zhou. 2012. "What Reviews Are Satisfactory: Novel Features for Automatic Helpfulness Voting." In *SIGIR'12*, 495–504. Portland, Oregon, USA.
[12] Zheng, Xiaolin, Shuai Zhu, and Zhangxi Lin. 2013. "Capturing the Essence of Word-of-Mouth for Social Commerce: Assessing the Quality of Online E-Commerce Reviews by a Semi-Supervised Approach." *Decision Support Systems* 56 (0):211–22.
[13] Yu, Xiaohui, Yang Liu, Xiangji Huang, and Aijun An. 2010. "A Quality-Aware Model for Sales Prediction Using Reviews." In *WWW'10*, 1217–8. Raleigh, North Carolina, USA.
[14] Ngo-Ye, Thomas L., and Atish P. Sinha. 2012. "Analyzing Online Review Helpfulness Using a Regressional ReliefF-Enhanced Text Mining Method." *ACM Transactions on Management Information Systems* 3 (2):10:1–10:20.
[15] Liu, Jingjing, Yunbo Cao, Chin-Yew Lin, Yalou Huang, and Ming Zhou. 2007. "Low-Quality Product Review Detection in Opinion Summarization." In *EMNLP-Conll'07*, 334–42.
[16] Zhang, Richong, and Thomas Tran. 2010. "A Novel Approach for Recommending Ranked User-Generated Reviews." *Advances in Artificial Intelligence* 6085:324–27.
[17] Chen, Chien Chin, and You-De Tseng. 2010. "Quality Evaluation of Product Reviews Using an Information Quality Framework." *Decision Support Systems* 50 (4):755–68.
[18] Li, Yijun, Qiang Ye, Ziqiong Zhang, and Tienan Wang. 2011. "Snippet-Based Unsupervised Approach for Sentiment Classification of Chinese Online

Reviews." *International Journal of Information Technology and Decision Making* 10 (6):1097–1110.
[19] Liu, Ying, Jian Jin, Ping Ji, Jenny A. Harding, and Richard Y. K. Fung, "Identifying Helpful Online Reviews: A Product Designer's Perspective", Computer-Aided Design, Volume 45, Issue 2, pages 180–194, February 2013
[20] Jindal, Nitin, and Bing Liu. 2008. "Opinion Spam and Analysis." In *WSDM'08*. Palo Alto, California, USA.
[21] Jindal, Nitin, Bing Liu, and Ee-Peng Lim. 2010. "Finding Unusual Review Patterns Using Unexpected Rules." In *CIKM'10*, 1549–52. Toronto, ON, Canada.
[22] Wu, Guangyu, Derek Greene, and Padraig Cunningham. 2010. "Merging Multiple Criteria to Identify Suspicious Reviews." In *RecSys'10*, 241–44. Barcelona, Spain.
[23] Ott, Myle, Yejin Choi, Claire Cardie, and Jeffrey T. Hancock. 2011. "Finding Deceptive Opinion Spam by Any Stretch of the Imagination." In *HLT'11*, 309–19.
[24] Ott, Myle, Claire Cardie, and Jeff Hancock. 2012. "Estimating the Prevalence of Deception in Online Review Communities." In *WWW'12*, 201–10.
[25] Lau, Raymond Y. K., S. Y. Liao, Ron Chi-Wai Kwok, Kaiquan Xu, Yunqing Xia, and Yuefeng Li. 2012. "Text Mining and Probabilistic Language Modeling for Online Review Spam Detection." ACM Transactions on Management Information Systems 2 (4):25:1–25:30.
[26] Morales, Alex, Huan Sun, and Xifeng Yan. 2013. "Synthetic Review Spamming and Defense." In *WWW'13 Companion*, 155–56. Rio de Janeiro, Brazil.
[27] Song, Long, Raymond Yiu Keung Lau, Ron Chi-Wai Kwok, Kristijan Mirkovski, and Wenyu Dou. 2017. "Who Are the Spoilers in Social Media Marketing? Incremental Learning of Latent Semantics for Social Spam Detection." *Electronic Commerce Research* 17 (1):51–81.
[28] Xie, Sihong, Guan Wang, Shuyang Lin, and Philip S. Yu. 2012b. "Review Spam Detection via Time Series Pattern Discovery." In *WWW'12 Companion*, 635–36. Lyon, France: ACM.
[29] Xie, Sihong, Guan Wang, Shuyang Lin, and Philip S. Yu. 2012a. "Review Spam Detection via Temporal Pattern Discovery." In *KDD'12*, 823–31. Beijing, China: ACM.
[30] Lim, Ee-Peng, Viet-An Nguyen, Nitin Jindal, Bing Liu, and Hady Wirawan Lauw. 2010. "Detecting Product Review Spammers Using Rating Behaviors." In *CIKM'10*, 939–48. Toronto, ON, Canada.
[31] Wang, Guan, Sihong Xie, Bing Liu, and Philip S. Yu. 2012. "Identify Online Store Review Spammers via Social Review Graph." *ACM Transactions on Intelligent Systems and Technology* 3 (4). ACM:61:1–61:21.
[32] Kokkodis, Marios. 2012. "Learning from Positive and Unlabeled Amazon Reviews: Towards Identifying Trustworthy Reviewers." In *WWW'12 Companion*, 545–46. Lyon, France.
[33] Mukherjee, Arjun, Bing Liu, Junhui Wang, Natalie Glance, and Nitin Jindal. 2011. "Detecting Group Review Spam." In *WWW'11*, 93–94. Hyderabad, India.
[34] Mukherjee, Arjun, Bing Liu, and Natalie Glance. 2012. "Spotting Fake Reviewer Groups in Consumer Reviews." In *WWW'12*, 191–200. Lyon, France.
[35] Dalvi, Nilesh, Ravi Kumar, Bo Pang, and Andrew Tomkins. 2009. "A Translation Model for Matching Reviews to Objects." In *CIKM'09*, 167–76. Hong Kong, China.
[36] Liu, Kang, Liheng Xu, and Jun Zhao. 2012. "Opinion Target Extraction Using Word-Based Translation Model." In *EMNLP-Conll'12*, 1346–56. Jeju Island, Korea.
[37] Zhang, Qi, Yuanbin Wu, Yan Wu, and Xuanjing Huang. 2011. "Opinion Mining with Sentiment Graph." In *WI-IAT'11*, 249–52. Washington, DC, USA: IEEE Computer Society.
[38] Ma, Tengfei, and Xiaojun Wan. 2010. "Opinion Target Extraction in Chinese News Comments." In *COLING'10*, 782–90. Beijing, China.
[39] Yao, Yangjie, and Aixin Sun. 2014. "Product Name Recognition and Normalization in Internet Forums." In *SIGIR'14*. Gold Coast, Australia.
[40] Yao, Yangjie, and Aixin Sun. 2015. "Mobile Phone Name Extraction from Internet Forums: A Semi-Supervised Approach." *World Wide Web*, 1–23.
[41] Jakob, Niklas, and Iryna Gurevych. 2010. "Extracting Opinion Targets in a Single- and Cross-Domain Setting with Conditional Random Fields." In *EMNLP'10*, 1035–45. Cambridge, Massachusetts.
[42] Ding, Xiaowen, and Bing Liu. 2007. "The Utility of Linguistic Rules in Opinion Mining." In *SIGIR'07*. Amsterdam, The Netherlands.
[43] Ding, Xiaowen, Bing Liu, and Philip S. Yu. 2008. "A Holistic Lexicon-Based Approach to Opinion Mining." In *WSDM'08*. Palo Alto, California, USA.
[44] Kobayakawa, Takeshi S., Tadashi Kumano, Hideki Tanaka, Naoaki Okazaki, Jin-Dong Kim, and Jun'ichi Tsujii. 2009. "Opinion Classification with Tree Kernel SVM Using Linguistic Modality Analysis." In *CIKM'09*, 1791–4. Hong Kong, China.
[45] Polpinij, Jantima, and Aditya K. Ghose. 2008. "An Ontology-Based Sentiment Classification Methodology for Online Consumer Reviews." In *WI-IAT'08*, 518–24.
[46] Hassan, Ahmed, and Dragomir Radev. 2010. "Identifying Text Polarity Using Random Walks." In *ACL'10*, 395–403. Uppsala, Sweden.
[47] Qiu, Guang, Bing Liu, Jiajun Bu, and Chun Chen. 2009. "Expanding Domain Sentiment Lexicon Through Double Propagation." In *IJCAI'09*, 1199–1204. Pasadena, California, USA.
[48] Hassan, Ahmed, Vahed Qazvinian, and Dragomir Radev. 2010. "What's with the Attitude?: Identifying Sentences with Attitude in Online Discussions." In *EMNLP'10*, 1245–55. Cambridge, Massachusetts.
[49] Mei, Qiaozhu, Xu Ling, Matthew Wondra, Hang Su, and ChengXiang Zhai. 2007. "Topic Sentiment Mixture: Modeling Facets and Opinions in Weblogs." In *WWW'07*, 171–80. Banff, Alberta, Canada.
[50] Zhao, Wayne Xin, Jing Jiang, Hongfei Yan, and Xiaoming Li. 2010. "Jointly Modeling Aspects and

Opinions with a MaxEnt-LDA Hybrid." In *EMNLP'10*, 56–65. Cambridge, Massachusetts.
[51] Wu, Yuanbin, Qi Zhang, Xuanjing Huang, and Lide Wu. 2011. "Structural Opinion Mining for Graph-Based Sentiment Representation." In *EMNLP'11*, 1332–41. Edinburgh, United Kingdom.
[52] Bespalov, Dmitriy, Bing Bai, Yanjun Qi, and Ali Shokoufandeh. 2011. "Sentiment Classification Based on Supervised Latent N-Gram Analysis." In *CIKM'11*. Glasgow, Scotland, UK.
[53] Liu, Tao, Minghui Li, Shusen Zhou, and Xiaoyong Du. 2011. "Sentiment Classification via L2-Norm Deep Belief Network." In *CIKM'11*, 2489–92. Glasgow, Scotland, UK.
[54] Long, Guodong, Ling Chen, Xingquan Zhu, and Chengqi Zhang. 2012. "TCSST: Transfer Classification of Short & Sparse Text Using External Data." In *CIKM'12*, 764–72. Maui, Hawaii, USA.
[55] Hu, Xinghua, and Bin Wu. 2009. "Classification and Summarization of Pros and Cons for Customer Reviews." In *WI-IAT'09*, 73–76.
[56] Zagibalov, Taras, and John Carroll. 2008. "Automatic Seed Word Selection for Unsupervised Sentiment Classification of Chinese Text." In *COLING'08*, 1073–80. Manchester, United Kingdom.
[57] Lau, Raymond Yiu Keung, Chun Lam Lai, Peter B. Bruza, and Kam F. Wong. 2011. "Leveraging Web 2.0 Data for Scalable Semi-Supervised Learning of Domain-Specific Sentiment Lexicons." In *CIKM'11*, 2457–60. Glasgow, Scotland, UK: ACM.
[58] Lin, Kevin Hsin-Yih, Changhua Yang, and Hsin-Hsi Chen. 2008. "Emotion Classification of Online News Articles from the Reader's Perspective." In *WI-IAT'08*, 220–26.
[59] Sindhwani, Vikas, and Prem Melville. 2008. "Document-Word Co-Regularization for Semi-Supervised Sentiment Analysis." In *ICDM'08*, 1025–30.
[60] Hu, Xia, Lei Tang, Jiliang Tang, and Huan Liu. 2013. "Exploiting Social Relations for Sentiment Analysis in Microblogging." In *WSDM'13*, 537–46. Rome, Italy.
[61] Liu, Jingjing, and Stephanie Seneff. 2009. "Review Sentiment Scoring via a Parse-and-Paraphrase Paradigm." In *EMNLP'09*, 161–69. Singapore.
[62] Wu, Yuanbin, Qi Zhang, Xuanjing Huang, and Lide Wu. 2009. "Phrase Dependency Parsing for Opinion Mining." In *EMNLP'09*, 1533–41. Singapore.
[63] Zhang, Qi, Yuanbin Wu, Tao Li, Mitsunori Ogihara, Joseph Johnson, and Xuanjing Huang. 2009. "Mining Product Reviews Based on Shallow Dependency Parsing." In *SIGIR'09*, 726–27. Boston, MA, USA.
[64] Wei, Wei, and Jon Atle Gulla. 2010. "Sentiment Learning on Product Reviews via Sentiment Ontology Tree." In *ACL'10*, 404–13. Uppsala, Sweden.
[65] Cataldi, Mario, Andrea Ballatore, Ilaria Tiddi, and Marie-Aude Aufaure. 2013. "Good Location, Terrible Food: Detecting Feature Sentiment in User-Generated Reviews." *Social Network Analysis and Mining* 3 (4):1149–63.
[66] Jin, Wei, Hung Hay Ho, and Rohini K. Srihari. 2009. "OpinionMiner: A Novel Machine Learning System for Web Opinion Mining and Extraction." In *KDD'09*, 1195–1204. Paris, France.
[67] Chen, Li, Luole Qi, and Feng Wang. 2012. "Comparison of Feature-Level Learning Methods for Mining Online Consumer Reviews." *Expert Systems with Applications* 39 (10):9588–9601.
[68] Rentoumi, Vassiliki, George A. Vouros, Vangelis Karkaletsis, and Amalia Moser. 2012. "Investigating Metaphorical Language in Sentiment Analysis: A Sense-to-Sentiment Perspective." *ACM Transactions on Speech and Language Processing* 9 (3). ACM:6:1–6:31.
[69] McAuley, Julian John, Jure Leskovec, and Dan Jurafsky. 2012. "Learning Attitudes and Attributes from Multi-Aspect Reviews." In *ICDM'12*.
[70] Moghaddam, Samaneh, and Martin Ester. 2013. "The FLDA Model for Aspect-Based Opinion Mining: Addressing the Cold Start Problem." In *WWW'13*, 909–18. Rio de Janeiro, Brazil.
[71] Saad, Farag, and Brigitte Mathiak. 2013. "Revised Mutual Information Approach for German Text Sentiment Classification." In *WWW'13 Companion*, 579–86. Rio de Janeiro, Brazil.
[72] Kim, Soo-Min, and Eduard Hovy. 2006. "Automatic Identification of Pro and Con Reasons in Online Reviews." In *COLING'06*, 483–90. Sydney, Australia.
[73] Yu, Jianxing, Zheng-Jun Zha, Meng Wang, and Tat-Seng Chua. 2011a. "Aspect Ranking: Identifying Important Product Aspects from Online Consumer Reviews." In *ACL'11*, 1496–1505. Portland, Oregon.
[74] Rao, Yanghui, Jingsheng Lei, Liu Wenyin, Qing Li, and Mingliang Chen. 2013. "Building Emotional Dictionary for Sentiment Analysis of Online News." *World Wide Web* June:1–20.
[75] Qiu, Likun, Weishi Zhang, Changjian Hu, and Kai Zhao. 2009. "SELC: A Self-Supervised Model for Sentiment Classification." In *CIKM'09*, 929–36. Hong Kong, China.
[76] Zhang, Weishi, Guiguang Ding, Li Chen, Chunping Li, and Chengbo Zhang. 2013. "Generating Virtual Ratings from Chinese Reviews to Augment Online Recommendations." *ACM Transactions on Intelligent Systems and Technology* 4 (1):9:1–9:17.
[77] Hu, Minqing, and Bing Liu. 2004. "Mining and Summarizing Customer Reviews." In *KDD'04*, 168–77.
[78] Liu, Bing, Minqing Hu, and Junsheng Cheng. 2005. "Opinion Observer: Analyzing and Comparing Opinions on the Web." In *WWW'05*. Chiba, Japan.
[79] Popescu, Ana-Maria, and Oren Etzioni. 2005. "Extracting Product Features and Opinions from Reviews." In *EMNLP'05*, 339–46.
[80] Hai, Zhen, Kuiyu Chang, and Gao Cong. 2012. "One Seed to Find Them All: Mining Opinion Features via Association." In *CIKM'12*, 255–64. Maui, Hawaii, USA.
[81] Stoyanov, Veselin, and Claire Cardie. 2008. "Topic Identification for Fine-Grained Opinion Analysis." In *COLING'08*, 817–24. Manchester, United Kingdom.
[82] Guo, Honglei, Huijia Zhu, Zhili Guo, XiaoXun Zhang, and Zhong Su. 2009. "Product Feature

Categorization with Multilevel Latent Semantic Association." In *CIKM'09*, 1087–96. Hong Kong, China.
[83] Lin, Chenghua, and Yulan He. 2009. "Joint Sentiment/Topic Model for Sentiment Analysis." In *CIKM'09*, 375–84. Hong Kong, China.
[84] Lin, Chenghua, Yulan He, and Richard Everson. 2010. "A Comparative Study of Bayesian Models for Unsupervised Sentiment Detection." In *CONLL'10*, 144–52. Uppsala, Sweden.
[85] Jo, Yohan, and Alice H. Oh. 2011. "Aspect and Sentiment Unification Model for Online Review Analysis." In *WSDM'11*, 815–24. Hong Kong, China.
[86] Xu, Xueke, Songbo Tan, Yue Liu, Xueqi Cheng, and Zheng Lin. 2012. "Towards Jointly Extracting Aspects and Aspect-Specific Sentiment Knowledge." In *CIKM'12*, 1895–9. Maui, Hawaii, USA.
[87] Titov, Ivan, and Ryan McDonald. 2008. "Modeling Online Reviews with Multi-Grain Topic Models." In *WWW'08*, 111–20. Beijing, China.
[88] Chen, Runyu, and Wei Xu. 2017. "The Determinants of Online Customer Ratings: A Combined Domain Ontology and Topic Text Analytics Approach." *Electronic Commerce Research* 17 (1):31–50.
[89] Alam, Md. Hijbul, and SangKeun Lee. 2012. "Semantic Aspect Discovery for Online Reviews." In *ICDM'12*, 816–21.
[90] Mukherjee, Arjun, and Bing Liu. 2012a. "Aspect Extraction Through Semi-Supervised Modeling." In *ACL'12*, 339–48. Jeju Island, Korea.
[91] Moghaddam, Samaneh, and Martin Ester. 2011. "ILDA: Interdependent LDA Model for Learning Latent Aspects and Their Ratings from Online Product Reviews." In *SIGIR'11*, 665–74. Beijing, China.
[92] Moghaddam, Samaneh, and Martin Ester. 2012. "On the Design of LDA Models for Aspect-Based Opinion Mining." In *CIKM'12*, 803–12. Maui, Hawaii, USA.
[93] Yang, Christopher C., Y. C. Wong, and Chih-Ping Wei. 2009. "Classifying Web Review Opinions for Consumer Product Analysis." In *ICEC'09*, 57–63.
[94] Zhai, Zhongwu, Bing Liu, Hua Xu, and Peifa Jia. 2010. "Grouping Product Features Using Semi-Supervised Learning with Soft-Constraints." In *COLING'10*, 1272–80. Beijing, China.
[95] Yang, Chin-Sheng, Chih-Ping Wei, and Christopher C. Yang. 2009. "Extracting Customer Knowledge from Online Consumer Reviews: A Collaborative-Filtering-Based Opinion Sentence Identification Approach." In *ICEC'09*, 64–71. Taipei, Taiwan.
[96] Zhu, Jingbo, Huizhen Wang, Benjamin K. Tsou, and Muhua Zhu. 2009. "Multi-Aspect Opinion Polling from Textual Reviews." In *CIKM'09*, 1799–1802. Hong Kong, China.
[97] Su, Qi, Xinying Xu, Honglei Guo, Zhili Guo, Xian Wu, Xiaoxun Zhang, Bin Swen, and Zhong Su. 2008. "Hidden Sentiment Association in Chinese Web Opinion Mining." In *WWW'08*, 959–68. Beijing, China.
[98] Kim, Jungi, Jin-Ji Li, and Jong-Hyeok Lee. 2009. "Discovering the Discriminative Views: Measuring Term Weights for Sentiment Analysis." In *ACL'09*, 253–61. Suntec, Singapore.
[99] Esuli, Andrea, and Fabrizio Sebastiani. 2006. "Determining Term Subjectivity and Term Orientation for Opinion Mining." In *EACL'06*. Stanford, California.
[100] Wiebe, Janyce, Theresa Wilson, and Claire Cardie. 2005. "Annotating Expressions of Opinions and Emotions in Language." *Language Resources and Evaluation* 39 (2/3):165–210.
[101] Zhang, Wei, Clement Yu, and Weiyi Meng. 2007. "Opinion Retrieval from Blogs." In *CIKM'07*, 831–40. Lisbon, Portugal.
[102] Ganesan, Kavita, and Chengxiang Zhai. 2012. "FindiLike: Preference Driven Entity Search." In *WWW'12 Companion*, 345–48. Lyon, France.
[103] Gerani, Shima, Mark James Carman, and Fabio Crestani. 2010. "Proximity-Based Opinion Retrieval." In *SIGIR'10*, 403–10. Geneva, Switzerland.
[104] Gerani, Shima, Mark Carman, and Fabio Crestani. 2012. "Aggregation Methods for Proximity-Based Opinion Retrieval." *ACM Transactions on Information Systems* 30 (4):26:1–26:36.
[105] Huang, Shen, Dan Shen, Wei Feng, Catherine Baudin, and Yongzheng Zhang. 2009. "Improving Product Review Search Experiences on General Search Engines." In *ICEC'09*, 107–16. Taipei, Taiwan.
[106] Wan, Xiaojun. 2008. "Using Bilingual Knowledge and Ensemble Techniques for Unsupervised Chinese Sentiment Analysis." In *EMNLP'08*, 553–61. Honolulu, Hawaii.
[107] Wan, Xiaojun. 2009. "Co-Training for Cross-Lingual Sentiment Classification." In *ACL'09*, 235–43. Suntec, Singapore.
[108] Wan, Xiaojun. 2011. "Bilingual Co-Training for Sentiment Classification of Chinese Product Reviews." *Computational Linguistics* 37 (3):587–616.
[109] Wan, Xiaojun. 2012. "A Comparative Study of Cross-Lingual Sentiment Classification." In *WI-IAT'12*, 24–31.
[110] Guo, Honglei, Huijia Zhu, Zhili Guo, Xiaoxun Zhang, and Zhong Su. 2010. "OpinionIt: A Text Mining System for Cross-Lingual Opinion Analysis." In *CIKM'10*, 1199–1208. Toronto, ON, Canada.
[111] Abbasi, Ahmed, Hsinchun Chen, and Arab Salem. 2008. "Sentiment Analysis in Multiple Languages: Feature Selection for Opinion Classification in Web Forums." *ACM Transactions on Information Systems* 26 (3):12:1–12:34.
[112] Lin, Zheng, Songbo Tan, and Xueqi Cheng. 2012. "A Fast and Accurate Method for Bilingual Opinion Lexicon Extraction." In *WI-IAT'12*, 50–57.
[113] Yu, Jianxing, Zheng-Jun Zha, Meng Wang, and Tat-Seng Chua. 2011b. "Hierarchical Organization of Unstructured Consumer Reviews." In *WWW'11*, 171–72. Hyderabad, India.
[114] Yu, Jianxing, Zheng-Jun Zha, Meng Wang, Kai Wang, and Tat-Seng Chua. 2011. "Domain-Assisted Product Aspect Hierarchy Generation: Towards Hierarchical Organization of Unstructured Consumer Reviews." In *EMNLP'11*, 140–50. Edinburgh, United Kingdom.
[115] Yu, Jianxing, Zheng-Jun Zha, and Tat-Seng Chua. 2012. "Answering Opinion Questions on Products by Exploiting Hierarchical Organization of Consumer

Reviews." In *EMNLP-Conll'12*, 391–401. Jeju Island, Korea.
[116] Zhai, Zhongwu, Bing Liu, Hua Xu, and Peifa Jia. 2011. "Clustering Product Features for Opinion Mining." In *WSDM'11*, 347–54. Hong Kong, China.
[117] Li, Fangtao, Chao Han, Minlie Huang, Xiaoyan Zhu, Ying-Ju Xia, Shu Zhang, and Hao Yu. 2010. "Structure-Aware Review Mining and Summarization." In *COLING'10*, 653–61. Beijing, China.
[118] Lu, Yue, ChengXiang Zhai, and Neel Sundaresan. 2009. "Rated Aspect Summarization of Short Comments." In *WWW'09*, 131–40. Madrid, Spain.
[119] Ganesan, Kavita, ChengXiang Zhai, and Evelyne Viegas. 2012. "Micropinion Generation: An Unsupervised Approach to Generating Ultra-Concise Summaries of Opinions." In *WWW'12*, 869–78. Lyon, France.
[120] Zhuang, Li, Feng Jing, and Xiao Yan Zhu. 2006. "Movie Review Mining and Summarization." In *CIKM'06*, 43–50. Arlington, Virginia, USA.
[121] Ly, Duy Khang, Kazunari Sugiyama, Ziheng Lin, and Min-Yen Kan. 2011. "Product Review Summarization from a Deeper Perspective." In *JCDL'11*, 311–14. Ottawa, Ontario, Canada.
[122] Yatani, Koji, Michael Novati, Andrew Trusty, and Khai N. Truong. 2011. "Review Spotlight: A User Interface for Summarizing User-Generated Reviews Using Adjective-Noun Word Pairs." In *CHI'11*, 1541–50. Vancouver, BC, Canada.
[123] Rohrdantz, Christian, Ming C. Hao, Umeshwar Dayal, Lars-Erik Haug, and Daniel A. Keim. 2012. "Feature-Based Visual Sentiment Analysis of Text Document Streams." *ACM Transactions on Intelligent Systems and Technology* 3 (2):26:1–26:25.
[124] Das, Amitava, and Sivaji Bandyopadhyay. 2010. "Topic-Based Bengali Opinion Summarization." In *COLING'10*, 232–40. Beijing, China.
[125] Ma, Zongyang, Aixin Sun, Quan Yuan, and Gao Cong. 2012. "Topic-Driven Reader Comments Summarization." In *CIKM'12*, 265–74. Maui, Hawaii, USA.
[126] Ju, Shengfeng, Shoushan Li, Yan Su, Guodong Zhou, Yu Hong, and Xiaojun Li. 2012. "Dual Word and Document Seed Selection for Semi-Supervised Sentiment Classification." In *CIKM'12*, 2295–8. Maui, Hawaii, USA.
[127] Tsaparas, Panayiotis, Alexandros Ntoulas, and Evimaria Terzi. 2011. "Selecting a Comprehensive Set of Reviews." In *KDD'11*, 168–76. San Diego, California, USA.
[128] Lappas, Theodoros, Mark Crovella, and Evimaria Terzi. 2012. "Selecting a Characteristic Set of Reviews." In KDD'12, 832–40. Beijing, China.
[129] Li, Shoushan, Shengfeng Ju, Guodong Zhou, and Xiaojun Li. 2012. "Active Learning for Imbalanced Sentiment Classification." In *EMNLP-Conll'12*, 139–48. Jeju Island, Korea.
[130] Agarwal, Deepak, Bee-Chung Chen, and Bo Pang. 2011. "Personalized Recommendation of User Comments via Factor Models." In *EMNLP'11*, 571–82. Edinburgh, United Kingdom.
[131] Moghaddam, Samaneh, Mohsen Jamali, and Martin Ester. 2011. "Review Recommendation: Personalized Prediction of the Quality of Online Reviews." In *CIKM'11*, 2249–52. Glasgow, Scotland, UK.
[132] Moghaddam, Samaneh, Mohsen Jamali, and Martin Ester. 2012. "ETF: Extended Tensor Factorization Model for Personalizing Prediction of Review Helpfulness." In *WSDM'12*, 163–72. Seattle, Washington, USA.
[133] Lu, Yue, Hongning Wang, Chengxiang Zhai, and Dan Roth. 2012. "Unsupervised Discovery of Opposing Opinion Networks from Forum Discussions." In *CIKM'12*, 1642–6. Maui, Hawaii, USA.
[134] Fang, Yi, Luo Si, Naveen Somasundaram, and Zhengtao Yu. 2012. "Mining Contrastive Opinions on Political Texts Using Cross-Perspective Topic Model." In *WSDM'12*, 63–72. Seattle, Washington, USA.
[135] Mukherjee, Arjun, and Bing Liu. 2012b. "Mining Contentions from Discussions and Debates." In *KDD'12*, 841–49. Beijing, China.
[136] Ganapathibhotla, Murthy, and Bing Liu. 2008. "Mining Opinions in Comparative Sentences." In *COLING'08*, 241–48.
[137] Jindal, Nitin, and Bing Liu. 2006. "Identifying Comparative Sentences in Text Documents." In *SIGIR'06*, 244–51. Seattle, Washington, USA.
[138] Xu, Kaiquan, Stephen Shaoyi Liao, Jiexun Li, and Yuxia Song. 2011. "Mining Comparative Opinions from Customer Reviews for Competitive Intelligence." *Decision Support Systems* 50 (4):743–54.
[139] Paul, Michael J., Chengxiang Zhai, and Roxana Girju. 2010. "Summarizing Contrastive Viewpoints in Opinionated Text." In *EMNLP'10*, 66–76. Cambridge, Massachusetts.
[140] Kim, Hyun Duk, and Chengxiang Zhai. 2009. "Generating Comparative Summaries of Contradictory Opinions in Text." In *CIKM'09*, 385–94. Hong Kong, China.
[141] Zhang, Kunpeng, Ramanathan Narayanan, and Alok Choudhary. 2010. "Voice of the Customers: Mining Online Customer Reviews for Product Feature-Based Ranking." In *WOSN'10*, 1–9. Boston, MA.
[142] McAuley, Julian John, and Jure Leskovec. 2013b. "Hidden Factors and Hidden Topics: Understanding Rating Dimensions with Review Text." In *RecSys'13*, 165–72. Hong Kong, China.
[143] Raghavan, Sindhu, Suriya Gunasekar, and Joydeep Ghosh. 2012. "Review Quality Aware Collaborative Filtering." In *RecSys'12*, 123–30. Dublin, Ireland.
[144] Zhang, Kunpeng, Yu Cheng, Wei-keng Liao, and Alok Choudhary. 2012. "Mining Millions of Reviews: A Technique to Rank Products Based on Importance of Reviews." In *ICEC'11*, 12:1–12:8. Liverpool, United Kingdom.
[145] Stavrianou, Anna, and Caroline Brun. 2012. "Opinion and Suggestion Analysis for Expert Recommendations." In *Proceedings of the Workshop on Semantic Analysis in Social Media*, 61–69. Avignon, France.

[146] McAuley, Julian John, and Jure Leskovec. 2013a. "From Amateurs to Connoisseurs: Modeling the Evolution of User Expertise Through Online Reviews." In *WWW'13*, 897–908. Rio de Janeiro, Brazil.
[147] Xu, Yinqing, Wai Lam, and Tianyi Lin. 2014. "Collaborative Filtering Incorporating Review Text and Co-Clusters of Hidden User Communities and Item Groups." In *CIKM'14*, 251–60. Shanghai, China.
[148] Zhao, Wayne Xin, Jinpeng Wang, Yulan He, Ji-Rong Wen, Edward Y. Chang, and Xiaoming Li. 2016. "Mining Product Adopter Information from Online Reviews for Improving Product Recommendation." *ACM Transactions on Knowledge Discovery from Data* 10 (3):29:1–29:23.
[149] Ma, Yue, Guoqing Chen, and Qiang Wei. 2017. "Does Big Data Mean Big Knowledge? Integration of Big Data Analysis and Conceptual Model for Social Commerce Research." *Electronic Commerce Research* 17(1): 3–29
[150] Hu, Nan, Ling Liu, and Jie Jennifer Zhang. 2008. "Do Online Reviews Affect Product Sales? The Role of Reviewer Characteristics and Temporal Effects." *Information Technology and Management* 9 (3):201–14.
[151] Zhang, Zhu, Xin Li, and Yubo Chen. 2012. "Deciphering Word-of-Mouth in Social Media: Text-Based Metrics of Consumer Reviews." *ACM Transactions on Management Information Systems* 3 (1):5:1–5:23.
[152] Archak, Nikolay, Anindya Ghose, and Panagiotis G. Ipeirotis. 2007. "Show Me the Money!: Deriving the Pricing Power of Product Features by Mining Consumer Reviews." In *KDD'07*, 56–65. San Jose, California, USA.
[153] Archak, Nikolay, Anindya Ghose, and Panagiotis G. Ipeirotis. 2011. "Deriving the Pricing Power of Product Features by Mining Consumer Reviews." *Management Science* 57 (8):1485–1509.
[154] Liu, Yang, Xiaohui Yu, Xiangji Huang, and Aijun An. 2010. "S-PLASA+: Adaptive Sentiment Analysis with Application to Sales Performance Prediction." In *SIGIR'10*, 873–74. Geneva, Switzerland.
[155] Chan, Lai-Kow, and Ming-Lu Wu. 2002. "Quality Function Deployment: A Literature Review." *European Journal of Operational Research* 143 (3):463–97.
[156] Aboulafia, Annette, and Liam J. Bannon. 2004. "Understanding Affect in Design: An Outline Conceptual Framework." *Theoretical Issues in Ergonomics Science* 5 (1). Taylor & Francis:4–15.
[157] Barnes, Cathy, and Stephen Paul Lillford. 2009. "Decision Support for the Design of Affective Products." *Journal of Engineering Design* 20 (5):477–92.
[158] Chen, Chun-Hsien, Li Pheng Khoo, and Wei Yan. 2006. "An Investigation into Affective Design Using Sorting Technique and Kohonen Self-Organising Map." *Advances in Engineering Software* 37 (5):334–49.
[159] Kwong, C.K., Huimin Jiang, and X.G. Luo. 2016. "AI-Based Methodology of Integrating Affective Design, Engineering, and Marketing for Defining Design Specifications of New Products." *Engineering Applications of Artificial Intelligence* 47 (Supplement C):49–60.
[160] Khalid, Halimahtun M. 2006. "Embracing Diversity in User Needs for Affective Design." *Applied Ergonomics* 37 (4):409–18.
[161] Hsu, Fang-Chen, Yi-Hsien Lin, and Chien-Nan Chen. 2015. "Applying Cluster Analysis for Consumer's Affective Responses Toward Product Forms." *Journal of Interdisciplinary Mathematics* 18 (6):657–66.
[162] Balters, Stephanie, and Martin Steinert. 2017. "Capturing Emotion Reactivity Through Physiology Measurement as a Foundation for Affective Engineering in Engineering Design Science and Engineering Practices." *Journal of Intelligent Manufacturing* 28 (7):1585–1607.
[163] Diego-Mas, Jose A., and Jorge Alcaide-Marzal. 2016. "Single Users' Affective Responses Models for Product Form Design." *International Journal of Industrial Ergonomics* 53 (Supplement C):102–14.
[164] Abegaz, Tamirat, Edward Dillon, and Juan E. Gilbert. 2015. "Exploring Affective Reaction During User Interaction with Colors and Shapes." *Procedia Manufacturing* 3 (Supplement C):5253–60.
[165] Fung, Chris K.Y., C.K. Kwong, Kit Yan Chan, and Huimin Jiang. 2014. "A Guided Search Genetic Algorithm Using Mined Rules for Optimal Affective Product Design." *Engineering Optimization* 46 (8). Taylor & Francis:1094–1108.
[166] Lo, Cheng-Hung, and Chih-Hsing Chu. 2009. "Experimental Study for Computer Aided Affective Product Styling." *Computer-Aided Design and Applications* 6 (4):471–82.
[167] Jiang, Huimin, C.K. Kwong, K.W.M. Siu, and Y. Liu. 2015. "Rough Set and Pso-Based Anfis Approaches to Modeling Customer Satisfaction for Affective Product Design." *Advanced Engineering Informatics* 29 (3):727–38.
[168] Jiang, Huimin, C.K. Kwong, Ying Liu, and W.H. Ip. 2015. "A Methodology of Integrating Affective Design with Defining Engineering Specifications for Product Design." *International Journal of Production Research* 53 (8):2472–88.
[169] Seva, Rosemary R., Katherine Grace T. Gosiaco, Ma. Crea Eurice D. Santos, and Denise Mae L. Pangilinan. 2011. "Product Design Enhancement Using Apparent Usability and Affective Quality." *Applied Ergonomics* 42 (3):511–17.
[170] Chan, K. Y., and U. Engelke. 2017. "Varying Spread Fuzzy Regression for Affective Quality Estimation." *IEEE Transactions on Fuzzy Systems* 25 (3):594–613.
[171] Jiao, Roger J., Qianli Xu, Jun Du, Yiyang Zhang, Martin Helander, Halimahtun M. Khalid, Petri Helo, and Cheng Ni. 2007. "Analytical Affective Design with Ambient Intelligence for Mass Customization and Personalization." *International Journal of Flexible Manufacturing Systems* 19 (4):570–95.
[172] Ling, S.H., P.P. San, K.Y. Chan, F.H.F. Leung, and Y. Liu. 2014. "An Intelligent Swarm Based-Wavelet Neural Network for Affective Mobile Phone Design." *Neurocomputing* 142 (Supplement C):30–38.
[173] Kim, Hyun K., Sung H. Han, Jaehyun Park, and Joohwan Park. 2016. "Identifying Affect Elements Based on a Conceptual Model of Affect: A Case Study on a

Smartphone." *International Journal of Industrial Ergonomics* 53 (Supplement C):193–204.
[174] Akay, Diyar, and Mustafa Kurt. 2008. "A Neuro-Fuzzy Based Approach to Affective Design." *The International Journal of Advanced Manufacturing Technology* 40 (5):425.
[175] Bruch, Jessica, and Monica Bellgran. 2013. "Characteristics Affecting Management of Design Information in the Production System Design Process." *International Journal of Production Research* 51 (11):3241–51.
[176] Lu, Weihua, and Jean-Francois Petiot. 2014. "Affective Design of Products Using an Audio-Based Protocol: Application to Eyeglass Frame." *International Journal of Industrial Ergonomics* 44 (3):383–94.
[177] Zhou, Feng, Yangjian Ji, and Roger Jianxin Jiao. 2013. "Affective and Cognitive Design for Mass Personalization: Status and Prospect." *Journal of Intelligent Manufacturing* 24 (5):1047–69.
[178] Zhou, Feng, Yangjian Ji, and Roger J. Jiao. 2014. "Prospect-Theoretic Modeling of Customer Affective-Cognitive Decisions Under Uncertainty for User Experience Design." *IEEE Transactions on Human-Machine Systems* 44 (4):468–83.
[179] Jiao, Roger J., Feng Zhou, and Chih-Hsing Chu. 2017. "Decision Theoretic Modeling of Affective and Cognitive Needs for Product Experience Engineering: Key Issues and a Conceptual Framework." *Journal of Intelligent Manufacturing* 28 (7):1755–67.
[180] Zhou, Feng, Baiying Lei, Yitao Liu, and Roger Jianxin Jiao. 2017. "Affective Parameter Shaping in User Experience Prospect Evaluation Based on Hierarchical Bayesian Estimation." *Expert Systems with Applications* 78 (Supplement C):1–15.
[181] Olvander, Johan, Bjorn Lunden, and Hampus Gavel. 2009. "A Computerized Optimization Framework for the Morphological Matrix Applied to Aircraft Conceptual Design." *Computer-Aided Design* 41 (3):187–96.
[182] Ostertag, Oskar, Eva Ostertagova, and Robert Hunady. 2012. "Morphological Matrix Applied Within the Design Project of the Manipulator Frame." *Procedia Engineering* 48 (Supplement C):495–99.
[183] He, Bin, Wei Song and Yangang Wang. "Computational Conceptual Design Using Space Matrix". *Journal of Computing and Information Science in Engineering*, 15(1), 011004
[184] Ma, Hongzhan, Xuening Chu, Deyi Xue, and Dongping Chen. 2017. "A Systematic Decision Making Approach for Product Conceptual Design Based on Fuzzy Morphological Matrix." *Expert Systems with Applications* 81 (Supplement C):444–56.
[185] Yuan, Lin, Yusheng Liu, Yunfeng Lin, and Jianjun Zhao. 2017. "An Automated Functional Decomposition Method Based on Morphological Changes of Material Flows." *Journal of Engineering Design* 28 (1):47–75.
[186] Matthews, P. C. 2008. "A Bayesian Support Tool for Morphological Design." *Advanced Engineering Informatics* 22 (2):236–53.
[187] Lo, Cheng-Hung, Kevin C. Tseng, and Chih-Hsing Chu. 2010. "One-Step Qfd Based 3D Morphological Charts for Concept Generation of Product Variant Design." *Expert Systems with Applications* 37 (11):7351–63.
[188] Fiorineschi, Lorenzo, Federico Rotini, and Paolo Rissone. 2016. "A New Conceptual Design Approach for Overcoming the Flaws of Functional Decomposition and Morphology." *Journal of Engineering Design* 27 (7):438–68.
[189] Kroll, Ehud. 2013. "Design Theory and Conceptual Design: Contrasting Functional Decomposition and Morphology with Parameter Analysis." *Research in Engineering Design* 24 (2):165–83.
[190] Jimeno-Morenilla, Antonio, Rafael Molina-Carmona, and Jose-Luis Sanchez-Romero. 2011. "Mathematical Morphology for Design and Manufacturing." *Mathematical and Computer Modelling* 54 (7):1753–9.
[191] Kim, Sun-Joong, and Ji-Hyun Lee. 2015. "Parametric Shape Modification and Application in a Morphological Biomimetic Design." *Advanced Engineering Informatics* 29 (1):76–86.
[192] Mintchev, Stefano, and Dario Floreano. 2016. "Adaptive Morphology: A Design Principle for Multimodal and Multifunctional Robots." *IEEE Robotics Automation Magazine* 23 (3):42–54.
[193] Chen, Chun-Hsien, Li Pheng Khoo, and Wei Yan. 2002. "A Strategy for Acquiring Customer Requirement Patterns Using Laddering Technique and ART2 Neural Network." *Advanced Engineering Informatics* 16 (3):229–40.
[194] Griffin, Abbie, and John R. Hauser. 1993. "The Voice of the Customer." *Marketing Science* 12:1–27.
[195] Gustafsson, Anders, and N. Gustafsson. 1994. "Exceeding Customer Expectations." In *Proceedings of the 6th Symposium on Quality Function Deployment*, 52–57.
[196] Han, Chang Hee, Jae Kyeong Kim, and Sang Hyun Choi. 2004. "Prioritizing Engineering Characteristics in Quality Function Deployment with Incomplete Information: A Linear Partial Ordering Approach." *International Journal of Production Economics* 91 (3):235–49.
[197] Wu, H.-H., A.Y.H. Liao, and P.-C. Wang. 2005. "Using Grey Theory in Quality Function Deployment to Analyse Dynamic Customer Requirements." *The International Journal of Advanced Manufacturing Technology* 25 (11):1241–7.
[198] Wu, Hsin-Hung, and Jiunn-I Shieh. 2006. "Using a Markov Chain Model in Quality Function Deployment to Analyse Customer Requirements." *The International Journal of Advanced Manufacturing Technology* 30 (1):141–46.
[199] Lai, Xin, Min Xie, Kay-Chuan Tan, and Bo Yang. 2008. "Ranking of Customer Requirements in a Competitive Environment." *Computers and Industrial Engineering* 54 (2):202–14.
[200] Saaty, Thomas L. 1980. *The Analytic Hierarchy Process*. McGraw-Hill.
[201] Armacost, Robert L., Paul J. Componation, Michael A. Mullens, and William W. Swart. 1994. "An AHP Framework for Prioritizing Customer Requirements

in QFD: An Industrialized Housing Application." *IIE Transactions* 26 (4). Taylor; Francis:72–79.

[202] Chuang, P.-T. 2001. "Combining the Analytic Hierarchy Process and Quality Function Deployment for a Location Decision from a Requirement Perspective." *The International Journal of Advanced Manufacturing Technology* 18 (11). Springer London:842–49.

[203] Nepal, Bimal, Om P. Yadav, and Alper Murat. 2010. "A Fuzzy-AHP Approach to Prioritization of CS Attributes in Target Planning for Automotive Product Development." *Expert Systems with Applications* 37 (10):6775–86.

[204] Fung, Richard Y.K., K. Popplewell, and J. Xie. 1998. "An Intelligent Hybrid System for Customer Requirements Analysis and Product Attribute Targets Determination." *International Journal of Production Research* 36 (1):13–34.

[205] Wang, H., M. Xie, and T. N. Goh. 1998. "A Comparative Study of the Prioritization Matrix Method and the Analytic Hierarchy Process Technique in Quality Function Deployment." *Total Quality Management* 9 (6):421–30.

[206] Matzler, Kurt, and Hans H. Hinterhuber. 1998. "How to Make Product Development Projects More Successful by Integrating Kano's Model of Customer Satisfaction into Quality Function Deployment." *Technovation* 18 (1):25–38.

[207] Shen, X. X., K. C. Tan, and M. Xie. 2000. "An Integrated Approach to Innovative Product Development Using Kano's Model and QFD." *European Journal of Innovation Management* 3 (2):91–99.

[208] Lai, Xin, Kay-Chuan Tan, and Min Xie. 2007. "Optimizing Product Design Using Quantitative Quality Function Deployment: A Case Study." *Quality and Reliability Engineering International* 23 (1):45–57.

[209] Mu, Li-Feng, Jia-Fu Tang, Yi-Zeng Chen, and Chun-Kit Kwong. 2008. "A Fuzzy Multi-Objective Model of QFD Product Planning Integrating Kano Model." *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 16 (6):793 813.

[210] Kwong, C. K., T. C. Wong, and K. Y. Chan. 2009. "A Methodology of Generating Customer Satisfaction Models for New Product Development Using a Neuro-Fuzzy Approach." *Expert Systems with Applications* 36 (8):11262–70.

[211] Chaudha, Ankur, Rajeev Jain, A. Singh, and P. Mishra. 2011. "Integration of Kano's Model into Quality Function Deployment (QFD)." *The International Journal of Advanced Manufacturing Technology* 53 (5):689–98.

[212] Ji, Ping, Jian Jin, Ting Wang, and Yizeng Chen. 2014. "Quantification and integration of Kano's model into QFD for optimising product design", *International Journal of Production Research* 52(21):6335–6348

[213] Yadav, Om Prakash, and Parveen S. Goel. 2008. "Customer Satisfaction Driven Quality Improvement Target Planning for Product Development in Automotive Industry." *International Journal of Production Economics* 113 (2):997–1011.

[214] Chen, Chun-Chih, and Ming-Chuen Chuang. 2008. "Integrating the Kano Model into a Robust Design Approach to Enhance Customer Satisfaction with Product Design." *International Journal of Production Economics* 114 (2):667–81.

[215] Lin, Shu-Ping, Chen-Lung Yang, Ya-hui Chan, and Chwen Sheu. 2010. "Refining Kano's 'Quality Attributes-Satisfaction' Model: A Moderated Regression Approach." *International Journal of Production Economics* 126 (2):255–63.

[216] Wang, Ting, and Ping Ji. 2010. "Understanding customer needs through quantitative analysis of Kano's model", *International Journal of Quality & Reliability Management* 27(2):173–184

[217] Zhan, Jiaming, Han Tong Loh, and Ying Liu. 2009. "Gather customer concerns from online product reviews – A text summarization approach", *Expert Systems with Applications* 36(2) Part 1: 2107–2115

[218] Lim, Sunghoon, and Conrad S. Tucker. 2016. "A Bayesian Sampling Method for Product Feature Extraction from Large-Scale Textual Data." *Journal of Mechanical Design* 138 (6):061403–3–9.

[219] Tuarob, Suppawong, and Conrad S. Tucker. 2015. "Quantifying Product Favorability and Extracting Notable Product Features Using Large Scale Social Media Data." *Journal of Computing and Information Science in Engineering* 15 (SEPTEMBER):031003–1–12.

[220] Wong, Tak-Lam, Wai Lam, and Tik-Shun Wong. 2008. "An Unsupervised Framework for Extracting and Normalizing Product Attributes from Multiple Web Sites." In *SIGIR'08*, 35–42. Singapore, Singapore.

[221] Na, Jin-Cheon, Tun Thura Thet, and Christopher S.G. Khoo. 2010. "Comparing Sentiment Expression in Movie Reviews from Four Online Genres." *Online Information Review* 34 (2):317–38.

[222] Leng, Jiewu and Pingyu Jiang. 2016 "A deep learning approach for relationship extraction from interaction context in social manufacturing paradigm", *Knowledge-Based Systems*, 100:188 – 199

[223] Zhu, Xingwei, Zhao-Yan Ming, Xiaoyan Zhu, and Tat-Seng Chua. 2013. "Topic Hierarchy Construction for the Organization of Multi-Source User Generated Contents." In *SIGIR'13*, 233–42. Dublin, Ireland.

[224] Jin, Jian, Ping Ji, and Ying Liu. 2014. "Prioritising engineering characteristics based on customer online reviews for quality function deployment", *Journal of Engineering Design* 25 (7-9):303-324

[225] Wang, Pengfei, Jiafeng Guo, Yanyan Lan, Jun Xu, and Xueqi Cheng. 2016. "Your Cart Tells You: Inferring Demographic Attributes from Purchase Data." In *WSDM'16*, 173–82. San Francisco, California, USA.

[226] Kang, Jaeyong, and Hyunju Lee. 2017. "Modeling User Interest in Social Media Using News Media and Wikipedia." *Information Systems* 65 (April):52–64.

[227] Oentaryo, Richard Jayadi, Ee-Peng Lim, Freddy Chong Tat Chua, Jia-wei Low, and David Lo. 2016. "Collective Semi-Supervised Learning for User Profiling in Social Media." *Computing Research Repository* abs/1606.07707.

[228] Si, Jianfeng, Qing Li, Tieyun Qian, and Xiaotie Deng. 2013. "Users' Interest Grouping from Online

Reviews Based on Topic Frequency and Order." *World Wide Web* July:1–22.
[229] Miao, Qingliang, Shu Zhang, Yao Meng, and Hao Yu. 2013. "Domain-Sensitive Opinion Leader Mining from Online Review Communities." In *WWW'13 Companion*, 187–88. Rio de Janeiro, Brazil.
[230] Tang, Jie, Jimeng Sun, Chi Wang, and Zi Yang. 2009. "Social Influence Analysis in Large-Scale Networks." In *KDD'09*, 807–16. Paris, France.
[231] Kim, Young Ae, and Jaideep Srivastava. 2007. "Impact of Social Influence in E-Commerce Decision Making." In *ICEC'07*, 293–302. Minneapolis, MN, USA.
[232] Huang, Junming, Xue-Qi Cheng, Hua-Wei Shen, Tao Zhou, and Xiaolong Jin. 2012. "Exploring Social Influence via Posterior Effect of Word-of-Mouth Recommendations." In *WSDM'12*, 573–82. Seattle, Washington, USA.
[233] Xu, Yunjie, Cheng Zhang, and Ling Xue. 2013. "Measuring Product Susceptibility in Online Product Review Social Network." *Decision Support Systems*.
[234] Iyengar, Raghuram, and Christophe Van den Bulte. 2011. "Opinion Leadership and Social Contagion in New Product Diffusion." *Marketing Science* 30 (2):195–212.
[235] Kim, Kwang-Jae, Herbert Moskowitz, Anoop Dhingra, and Gerald Evans. 2000. "Fuzzy Multicriteria Models for Quality Function Deployment." *European Journal of Operational Research* 121 (3):504–18.
[236] Harding, J. A., K. Popplewell, Richard Y.K. Fung, and A. R. Omar. 2001. "An Intelligent Information Framework for Market Driven Product Design." *Computers in Industry* 44 (1):49–63.
[237] Fung, Richard Y.K., Yizeng Chen, and Jiafu Tang. 2006. "Estimating the Functional Relationships for Quality Function Deployment Under Uncertainties." *Fuzzy Sets and Systems* 157 (1):98–120.
[238] Zhai, Lian-Yin, Li-Pheng Khoo, and Zhao-Wei Zhong. 2009. "A Rough Set Based Decision Support Approach to Improving Consumer Affective Satisfaction in Product Design." *International Journal of Industrial Ergonomics* 39 (2):295–302.
[239] Kwong, C.K., Y. Chen, H. Bai, and D.S.K. Chan. 2007. "A Methodology of Determining Aggregated Importance of Engineering Characteristics in QFD for New Product Design." *Computers and Industrial Engineering* 53 (4):667–79.
[240] Chan, K. Y., C. K. Kwong, T. S. Dillon, and K. Y. Fung. 2011. "An Intelligent Fuzzy Regression Approach for Affective Product Design That Captures Nonlinearity and Fuzziness." *Journal of Engineering Design* 22 (8):523–42.
[241] Fung, K.Y., C.K. Kwong, K.W.M. Siu, and K.M. Yu. 2012. "A Multi-Objective Genetic Algorithm Approach to Rule Mining for Affective Product Design." *Expert Systems with Applications* 39 (8):7411–9.
[242] Jiao, Jianxin (Roger), Yiyang Zhang, and Martin Helander. 2006. "A Kansei Mining System for Affective Design." *Expert Systems with Applications* 30 (4):658–73.
[243] Shen, Hui-Cun, and Kun-Chieh Wang. 2016. "Affective Product Form Design Using Fuzzy Kansei Engineering and Creativity." *Journal of Ambient Intelligence and Humanized Computing* 7 (6):875–88.
[244] Jin, Jian, Ping Ji, Ying Liu, and S. C. Johnson Lim. 2015. "Translating Online Customer Opinions into Engineering Characteristics in QFD: A Probabilistic Language Analysis Approach", *Engineering Applications of Artificial Intelligence* 41:115–127
[245] Zhou, Feng, Jianxin Roger Jiao, Dirk Schaefer and Songlin Chen. "Hybrid Association Mining and Refinement for Affective Mapping in Emotional Design". *Journal of Computing and Information Science in Engineering,* 10(3), 031010
[246] Chen, Yizeng, Richard Y.K. Fung, and Jiafu Tang. 2006. "Rating Technical Attributes in Fuzzy QFD by Integrating Fuzzy Weighted Average Method and Fuzzy Expected Value Operator." *European Journal of Operational Research* 174 (3):1553–66.
[247] Akao, Yoji. 1990. *Quality Function Deployment: Integrating Customer Requirements into Product Design*. Productivity Press.
[248] Wang, Xiao-Tun, and Wei Xiong. 2011. "An Integrated Linguistic-Based Group Decision-Making Approach for Quality Function Deployment." *Expert Systems with Applications* 38 (12):14428–38.
[249] Lan, Lijun, Ying Liu and Wenfeng Lu. 2017. "Automatic Discovery of Design Task Structure Using Deep Belief Nets" *Journal of Computing and Information Science in Engineering* 17(4): 041001-8
[250] Liu, Y.-C., A. Chakrabarti, and T. Bligh. 2003. "Towards an 'Ideal' Approach for Concept Generation." *Design Studies* 24 (4):341–55.
[251] Nagai, Yukari, Toshiharu Taura, and Futoshi Mukai. 2009. "Concept Blending and Dissimilarity: Factors for Creative Concept Generation Process." *Design Studies* 30 (6):648–75.
[252] Liu, Ang, and Stephen C.-Y. Lu. 2014. "Alternation of Analysis and Synthesis for Concept Generation." *CIRP Annals* 63 (1):177–80.
[253] Rondini, Alice, Giuditta Pezzotta, Fabiana Pirola, Monica Rossi, and Paulo Pina. 2016. "How to Design and Evaluate Early Pss Concepts: The Product Service Concept Tree." *Procedia CIRP* 50 (Supplement C):366–71.
[254] Gironimo, G. Di, D. Carfora, G. Esposito, C. Labate, R. Mozzillo, F. Renno, A. Lanzotti, and M. Siuko. 2013. "Improving Concept Design of Divertor Support System for Fast Tokamak Using Triz Theory and Ahp Approach." *Fusion Engineering and Design* 88 (11):3014–20.
[255] Hilmann, J., M. Paas, A. Haenschke, and T. Vietor. 2007. "Automatic Concept Model Generation for Optimisation and Robust Design of Passenger Cars." *Advances in Engineering Software* 38 (11):795–801.
[256] Tiwari, Varun, Prashant Kumar Jain, and Puneet Tandon. 2016. "Product Design Concept Evaluation Using Rough Sets and Vikor Method." *Advanced Engineering Informatics* 30 (1):16–25.
[257] Yan, Wei, Chun-Hsien Chen, and Meng-Dar Shieh. 2006. "Product Concept Generation and Selection Using Sorting Technique and Fuzzy c-Means Algorithm." *Computers & Industrial Engineering* 50 (3):273–85.

[258] Huang, Hong-Zhong, Ruifeng Bo, and Wei Chen. 2006. "An Integrated Computational Intelligence Approach to Product Concept Generation and Evaluation." *Mechanism and Machine Theory* 41 (5):567–83.

[259] Chou, Jyh-Rong. 2014. "An Ideation Method for Generating New Product Ideas Using Triz, Concept Mapping, and Fuzzy Linguistic Evaluation Techniques." *Advanced Engineering Informatics* 28 (4):441–54.

[260] Kurtoglu, Tolga, Matthew I. Campbell, and Julie S. Linsey. 2009. "An Experimental Study on the Effects of a Computational Design Tool on Concept Generation." *Design Studies* 30 (6):676–703.

[261] Liang, Yan, Ying Liu, Chun Kit Kwong, and Wing Bun Lee. 2012. "Learning the "Whys": Discovering design rationale using text mining — An algorithm perspective", *Computer-Aided Design* 44(10):916–930

[262] Yamamoto, Eiko, Toshiharu Taura, Shota Ohashi and Masaki Yamamoto. "A Method for Function Dividing in Conceptual Design by Focusing on Linguistic Hierarchal Relations". *Journal of Computing and Information Science in Engineering*, 10(3), 031004

[263] Dering, Matthew L. and Conrad S. Tucker. 2017 "A Convolutional Neural Network Model for Predicting a Product's Function, Given Its Form". *Journal of Mechanical Design*, 139(11): 111408-14

[264] Hao, Jia, Qiangfu Zhao, and Yan Yan. 2017. "A Function-Based Computational Method for Design Concept Evaluation." *Advanced Engineering Informatics* 32 (Supplement C):237–47.

[265] Cao, Dongxing, Zhanjun Li, and Karthik Ramani. 2011. "Ontology-Based Customer Preference Modeling for Concept Generation." *Advanced Engineering Informatics* 25 (2):162–76.

[266] Park, Yongtae, and Sungjoo Lee. 2011. "How to Design and Utilize Online Customer Center to Support New Product Concept Generation." *Expert Systems with Applications* 38 (8):10638–47.

[267] Liu, Ang, and Stephen C.-Y. Lu. 2016. "A Crowdsourcing Design Framework for Concept Generation." *CIRP Annals* 65 (1):177–80.

[268] Wang, L., B. D. Youn, S. Azarm, and P. K. Kannan. 2011. "Customer-Driven Product Design Selection Using Web Based User-Generated Content." In *IDETC'11*, 405–19. Washington, DC.

[269] Chang, Danni, and Chun-Hsien Chen. 2015. "Product Concept Evaluation and Selection Using Data Mining and Domain Ontology in a Crowdsourcing Environment." *Advanced Engineering Informatics* 29 (4):759–74.

[270] Won, P.-H. 2001. "The Comparison Between Visual Thinking Using Computer and Conventional Media in the Concept Generation Stages of Design." *Automation in Construction* 10 (3):319–25.

[271] Kang, Jinsheng, Kang Zhong, Shengfeng Qin, Hongan Wang, and David Wright. 2013. "Instant 3D Design Concept Generation and Visualization by Real-Time Hand Gesture Recognition." *Computers in Industry* 64 (7):785–97.

[272] Farnsworth, M., and T. Tomiyama. 2014. "Capturing, Classification and Concept Generation for Automated Maintenance Tasks." *CIRP Annals* 63 (1):149–52.

[273] Tsenn, Joanna, Olufunmilola Atilola, Daniel A. McAdams, and Julie S. Linsey. 2014. "The Effects of Time and Incubation on Design Concept Generation." *Design Studies* 35 (5):500–526.

[274] Dong, Andy, Dan Lovallo, and Ronny Mounarath. 2015. "The Effect of Abductive Reasoning on Concept Selection Decisions." *Design Studies* 37 (Supplement C):37–58.

[275] Bracke, Stefan, Shuho Yamada, Yuki Kinoshita, Masato Inoue, and Tetsuo Yamada. 2017. "Decision Making Within the Conceptual Design Phase of Eco-Friendly Products." *Procedia Manufacturing* 8 (Supplement C):463–70.

[276] Noguchi, Hisataka. 1997. "An Idea Generation Support System for Industrial Designers (Idea Sketch Processor)." *Knowledge-Based Systems* 10 (1):37–42.

[277] Akasaka, F., Y. Nemoto, R. Chiba, and Y. Shimomura. 2012. "Development of PSS Design Support System: Knowledge-Based Design Support and Qualitative Evaluation." *Procedia CIRP* 3 (Supplement C):239–44.

[278] Ko, Yao-Tsung. 2017. "Modeling a Hybrid-Compact Design Matrix for New Product Innovation." *Computers & Industrial Engineering* 107 (Supplement C):345–59.

[2769] Moon, Heekyung, and Sung H. Han. 2016. "A Creative Idea Generation Methodology by Future Envisioning from the User Experience Perspective." *International Journal of Industrial Ergonomics* 56 (Supplement C):84–96.

[280] Jahanmir, Sara F., and Luis Filipe Lages. 2015. "The Lag-User Method: Using Laggards as a Source of Innovative Ideas." *Journal of Engineering and Technology Management* 37 (Supplement C):65–77.

[281] Lugt, Remko van der. 2005. "How Sketching Can Affect the Idea Generation Process in Design Group Meetings." *Design Studies* 26 (2):101–22.

[282] Hirunyawipada, Tanawat, and Audhesh K. Paswan. 2013. "Effects of Team Cognition and Constraint on New Product Ideation." *Journal of Business Research* 66 (11):2332–7.

[283] Zhang, Chengwei, Youngwook Paul Kwon, Julia Kramer, Euiyoung Kim and Alice M. Agogino. 2017 "Deep Learning for Design in Concept Clustering." In *IDETC* '17, 68352-12

[284] Zahay, Debra, Nick Hajli, and Debika Sihi. 2017. "Managerial Perspectives on Crowdsourcing in the New Product Development Process." *Industrial Marketing Management*.

[285] Simon, Fanny, and Alberic Tellier. 2011. "How Do Actors Shape Social Networks During the Process of New Product Development?" *European Management Journal* 29 (5):414–30.

[286] McAdam, Rodney, and John McClelland. 2002. "Sources of New Product Ideas and Creativity Practices in the UK Textile Industry." *Technovation* 22 (2):113–21.

[287] Tseng, Ian, Jarrod Moss, Jonathan Cagan, and Kenneth Kotovsky. 2008. "The Role of Timing and Analogical Similarity in the Stimulation of Idea Generation in Design." *Design Studies* 29 (3):203–21.

[288] Starkey, Elizabeth, Christine A. Toh, and Scarlett R. Miller. 2016. "Abandoning Creativity: The Evolution of Creative Ideas in Engineering Design Course Projects." *Design Studies* 47 (Supplement C):47–72.

[289] Nonaka, I., P. Byosiere, C. C. Borucki and N. Konno. 1994. "Organizational Knowledge Creation Theory – A First Comprehensive Test". *International Business Review* 3 (4), 337–351.

[290] Nonaka, I., R. Toyama and N. Konno. 2000. "SECI Ba and Leadership: A Unified Model of Dynamic Knowledge Creation". *Long Range Planning* 33 (1), 5–34.

[291] Nonaka, I., G. Krogh and S. Voelpel. 2006. "Organizational Knowledge Creation Theory: Evolutionary Paths and Future Advances". *Organization Studies* 27 (8), 1179–1208.

[292] Nonaka, I. and H. Takeuchi. 1995. *The Knowledge-Creating Company*. Oxford University Press, New York.

[293] Schulze, Anja, and Martin Hoegl. 2008. "Organizational Knowledge Creation and the Generation of New Product Ideas: A Behavioral Approach." *Research Policy* 37 (10):1742–50.

[294] Gomes, Paulo, Nuno Seco, Francisco C. Pereira, Paulo Paiva, Paulo Carreiro, Jose L. Ferreira, and Carlos Bento. 2006. "The Importance of Retrieval in Creative Design Analogies." *Knowledge-Based Systems* 19 (7):480–88.

[295] Jauregui-Becker, J.M., and W.W. Wits. 2012. "Knowledge Structuring and Simulation Modeling for Product Development." *Procedia CIRP* 2 (Supplement C):4–9.

[296] Verhaegen, Paul-Armand, Dennis Vandevenne, Jef Peeters, and Joost R. Duflou. 2013. "Refinements to the Variety Metric for Idea Evaluation." *Design Studies* 34 (2):243–63.

[297] Atilola, Olufunmilola, Megan Tomko, and Julie S. Linsey. 2016. "The Effects of Representation on Idea Generation and Design Fixation: A Study Comparing Sketches and Function Trees." *Design Studies* 42 (Supplement C):110–36.

[298] Nasr, Sana Ben, Guillaume Becan, Mathieu Acher, Joao Bosco Ferreira Filho, Nicolas Sannier, Benoit Baudry, and Jean-Marc Davril. 2017. "Automated Extraction of Product Comparison Matrices from Informal Product Descriptions." *Journal of Systems and Software* 124 (Supplement C):82–103.

[299] Lim, Soon Chong Johnson, Ying Liu, and Wing Bun Lee. 2010. "Multi-Facet Product Information Search and Retrieval Using Semantically Annotated Product Family Ontology." *Information Processing & Management* 46 (4):479–93.

[300] Petiot, Jean-Francois, and Bernard Yannou. 2004. "Measuring Consumer Perceptions for a Better Comprehension, Specification and Assessment of Product Semantics." *International Journal of Industrial Ergonomics* 33 (6):507–25.

[2301] Shi, Feng, Liuqing Chen, Ji Han and Peter Childs. 2017. "A Data-Driven Text Mining and Semantic Network Analysis for Design Information Retrieval". *Journal of Mechanical Design* 139(11):111402-14

[302] Huang, Yuexiang, Chun-Hsien Chen, I-Hsuan Cindy Wang, and Li Pheng Khoo. 2014. "A Product Configuration Analysis Method for Emotional Design Using a Personal Construct Theory." *International Journal of Industrial Ergonomics* 44 (1):120–30.

[303] Holy, Vladimir, Ondrej Sokol, and Michal Cerny. 2017. "Clustering Retail Products Based on Customer Behaviour." *Applied Soft Computing* 60 (Supplement C):752–62.

[304] Bracke, Stefan, and Christoph Rosebrock. 2016. "Contribution for Product Analyses to Quantify and Predict Similar or Diverse Eco-Related Product Perception in the Usage Phase." *Procedia CIRP* 40 (Supplement C):68–72.

[305] Li, Yan-Lai, Jia-Fu Tang, Kwai-Sang Chin, Xing-Gang Luo, Yun Pu, and Yu-Shi Jiang. 2012. "On Integrating Multiple Type Preferences into Competitive Analyses of Customer Requirements in Product Planning." *International Journal of Production Economics* 139 (1):168–79.

[306] Yuen, Kevin Kam Fung. 2017. "The Fuzzy Cognitive Pairwise Comparisons for Ranking and Grade Clustering to Build a Recommender System: An Application of Smartphone Recommendation." *Engineering Applications of Artificial Intelligence* 61 (Supplement C):136–51.

[307] Yuen, Kevin Kam Fung. 2013. "Toward a Ranking Strategy for E-Commerce Products in an E-Alliance Portal Using Primitive Cognitive Network Process." *Procedia Computer Science* 17 (Supplement C):1091–6.

[308] Netzer, Oded, Ronen Feldman, Jacob Goldenberg, and Moshe Fresko. 2012. "Mine Your Own Business: Market-Structure Surveillance Through Text Mining." *Marketing Science* 31 (3):521–43.

[309] Li, Si, Zheng-Jun Zha, Zhaoyan Ming, Meng Wang, Tat-Seng Chua, Jun Guo, and Weiran Xu. 2011. "Product Comparison Using Comparative Relations." In *SIGIR'11*, 1151–2. Beijing, China.

[310] Zhang, Zhu, Chenhui Guo, and Paulo Goes. 2013. "Product Comparison Networks for Competitive Analysis of Online Word-of-Mouth." *ACM Transactions on Management Information Systems* 3 (4):1–20.

[311] Chen, Kun, Peng Luo, and Huaiqing Wang. 2017. "Investigating Transitive Influences on Wom: From the Product Network Perspective." *Electronic Commerce Research* 17 (1):149–67.

[312] Zhang, Wenhao, Hua Xu, and Wei Wan. 2012. "Weakness Finder: Find Product Weakness from Chinese Reviews by Using Aspects Based Sentiment Analysis." *Expert Systems with Applications* 39 (11):10283–91.

[313] Liu, Yang, Jian-Wu Bi, and Zhi-Ping Fan. 2017. "Ranking Products Through Online Reviews: A Method Based on Sentiment Analysis Technique and Intuitionistic Fuzzy Set Theory." *Information Fusion* 36 (Supplement C):149–61.

[314] Li, Hui, Sourav S. Bhowmick, and Aixin Sun. 2010. "Affinity-Driven Prediction and Ranking of Products in Online Product Review Sites." In *CIKM'10*, 1745–8. Toronto, ON, Canada.

[315] Yang, Xian, Guangfei Yang, and Jiangning Wu. 2016. "Integrating Rich and Heterogeneous Information

to Design a Ranking System for Multiple Products." *Decision Support Systems* 84 (Supplement C):117–33.
[316] Flint, Daniel J. 2002. "Compressing New Product Success-to-Success Cycle Time: Deep Customer Value Understanding and Idea Generation." *Industrial Marketing Management* 31 (4):305–15.
[317] Wei, Wei, Jun Ji, Thorsten Wuest, and Fei Tao. 2017. "Product Family Flexible Design Method Based on Dynamic Requirements Uncertainty Analysis." *Procedia CIRP* 60 (Supplement C):332–37.
[318] Chen, S.L., R.J. Jiao, and M.M. Tseng. 2009. "Evolutionary Product Line Design Balancing Customer Needs and Product Commonality." *CIRP Annals* 58 (1):123–26.
[319] Cichos, Daniel, and Jan C. Aurich. 2016. "Support of Engineering Changes in Manufacturing Systems by Production Planning and Control Methods." *Procedia CIRP* 41 (Supplement C):165–70.
[320] Wang, H.S., Z.H. Che, and M.J. Wang. 2009. "A Three-Phase Integrated Model for Product Configuration Change Problems." *Expert Systems with Applications* 36 (3, Part 1):5491–5509.
[321] Li, Yuliang, Wei Zhao, and Xinyu Shao. 2012. "A Process Simulation Based Method for Scheduling Product Design Change Propagation." *Advanced Engineering Informatics* 26 (3):529–38.
[322] Azadeh, A., Sh. Sadri, M. Saberi, J. H. Yoon, E. Chang, O. Khadeer Hussain, and N. Pourmohammad Zia. 2017. "An Integrated Fuzzy Trust Prediction Approach in Product Design and Engineering." *International Journal of Fuzzy Systems* 19 (4):1190–9.
[323] Chong, Yih Tng, and Chun-Hsien Chen. 2010. "Management and Forecast of Dynamic Customer Needs: An Artificial Immune and Neural System Approach." *Advanced Engineering Informatics* 24 (1):96–106.
[324] Tucker, Conrad S., and Harrison M. Kim. 2011. "Trend Mining for Predictive Product Design." *Journal of Mechanical Design* 133 (11):111008–18.
[325] Ma, Jungmok, Minjung Kwak, and Harrison M. Kim. 2014. "Demand Trend Mining for Predictive Life Cycle Design." *Journal of Cleaner Production* 68 (Supplement C):189–99.
[326] Guo, Jing, Runhua Tan, Jianguang Sun, Jianliang Ren, Shengxuan Wu, and Yang Qiu. 2016. "A Needs Analysis Approach to Product Innovation Driven by Design." *Procedia CIRP* 39 (Supplement C):39–44.
[327] Tucker, Conrad, and Harrison M. Kim. 2011. "Predicting Emerging Product Design Trend by Mining Publicity Available Customer Review Data." In *ICED'11*, 43–52. Copenhagen, Denmark.
[328] Goorha, Saurabh, and Lyle Ungar. 2010. "Discovery of Significant Emerging Trends." In *KDD'10*, 57–64. Washington, DC, USA.
[329] Malone, Thomas W. 2014. "A Revolution in Business." *MIT Technology Review*. April 7, 2014.
[330] Schiller, Dan. 2015 "The Internet and Business." *MIT Technology Review*. February 10, 2015
[331] Evans, Philip. 2015. "From Deconstruction to Big Data: How Technology is Reshaping the Corporation. " *MIT Technology Review*. May 13, 2015