Journal of Intelligence Studies in Business

Journal of Intelligence Studies in Business

Publication details, including instructions for authors and subscription information: https://ojs.hh.se/index.php/JISIB/index

Integration of textual VoC into a CX data model for business intelligence use in B2C

Lucie Sperkova[a]*

[a]University of Economics, Prauge, Czech Republic;
*lucie.sperkova@vse.cz

To cite this article: Sperkova, L. (2019) Integration of textual VoC into a CX data model for business intelligence use in B2C. *Journal of Intelligence Studies in Business*. 9 (3) 39-55.

Article URL: https://ojs.hh.se/index.php/JISIB/article/view/476

Journal of Intelligence Studies in Business
Vol. 9, No. 3 (2019) pp. 39-55
Open Access: Freely available at: https://ojs.hh.se/

Integration of textual VoC into a CX data model for business intelligence use in B2C

Lucie Sperkova[a*]

[a]*University of Economics, Prague, Czech Republic*

**Corresponding author: lucie.sperkova@vse.cz*

Received 12 December 2019 Accepted 30 December 2019

ABSTRACT Customer experience (CX) focuses on customer feedback. CX is a holistic construct which contains different perceptual elements such as satisfaction and loyalty, but also emotions or personality. Customers share their opinions, which contain these elements also in textual expressions through different channels, known in research as Voice of Customer (VoC). Currently, VoC is collected mainly in customer surveys and manually evaluated, or through simple quantitative measurement from data scattered in various systems at the end of a customer journey. To bridge this gap, we designed a multidimensional CX data model for integrated storage of all customers' data from structured and textual sources. A consolidated CX measurement to monitor elements of CX during the entire customer journey from the customer perspective is proposed to serve as business intelligence. The artefact offers a self-contained expandable data mart affordable to implement in small and medium B2C enterprises. Companies can now manage customer relationships and future performance more automatically and effectively thanks to integrated information mined from texts, combined with other data from internal systems and shared across the company in unified reporting.

KEYWORDS Customer experience, data model, perceptual metrics, sentiment, voice of customer

1. INTRODUCTION

Customer experience (CX) can be understood in its holistic conception as a demonstration of experience through different elements, which originates in the customer. It encompasses cognitive, emotional, and social characteristics, as well as the user's quantitative interaction with the company (Verhoef et al., 2009) during the customer's entire purchase journey (Lemon and Verhoef, 2016). It is also an instrument to improve the value of the customer and the company. From the latter, it follows that the experience can be managed through the measurement of the elements connected as antecedents, succedents, or as parallel constructs to CX. These elements represent share-of-mind metrics, which are critical when managing to achieve better performance of an organisation.

The approaches for gathering data for CX are based mainly on the manual evaluation of questionnaires and surveys (e.g. Klaus and Maklan, 2013; Klaus, 2015; Khodadadi et al., 2016). Reading every text is time-consuming and resource-demanding (Nahili et al., 2019). In practice, the measurement of CX is currently dependent mostly on the evaluation of single metrics such as Net Promoter Score (NPS). These metrics are gathered with structured behavioural or transactional data as standalone quantitative indicators instead of metrics based on the text itself (e.g. Godes and Mayzlin, 2004; Liu, 2006; Wu and Zheng, 2012). The evaluators extract necessary data from specific tools or analytical CRM (e.g. Aziza, Oubrich and Søilen, 2015).

CX management emphasises value creation. From a managerial perspective, firms should pay attention to textual content when managing CX and, more importantly, focus on the right measures. The value cannot be calculated merely from structured data as it is impossible to set the probability of a rise of such surprising information from VoC. Successful CX management needs to systematically collect VoC, mine that VoC for insights, share the insights with the business, and incorporate the insights into business decisions. That requires the ability to design, implement, and manage CX in a disciplined manner as a business intelligence (BI) solution. As is seen from the results of qualitative research in (Šperková, 2019), the most challenging for marketers is the synthesis of information from VoC into useful reports. The model targets this synthesis of the information to gain new insight into the CX.

There are many barriers to achieving the full potential of VoC analysis within CX, as the author identified in the previous qualitative research (Šperková, 2019) which can be overcome with the CX data model (missing shared VoC insights across the organisation, struggle to prove financial results, textual VoC is not well-analysed, fragmented view of the customer and missing integration of data, missing action with individual customers, missing formalisation of the processes).

Since the underlying data for CX measurement are located in different internal and external sources of the company, the examination of VoC for managing CX from one source only is incomplete. Companies should collect both qualitative and quantitative data from these sources to acquire a holistic view of CX. When accessing data from separate systems, end-users are not able to interconnect the data according to their identifiers or metadata and find valuable information about individual customers resulting from various customer data and VoC interconnections. End-users need to access the data from one integrated physical place stored in a unified form. The unified storage ensures the accuracy and reliability of the following measurement with minimal manual effort.

The proposed CX data model follows and builds on the author's previous research on the integration of VoC into BI in the banking domain (Šperková, 2014; Šperková and Škola, 2015a, 2015b; Šperková, Škola and Bruckner, 2015; Vencovský, Bruckner and Šperková, 2016). The model exploits data from analytical CRM. The aim is not a complex platform based on CRM but a self-contained expandable and transferable data model containing the data from textual VoC among other data, which can be implemented in any BI solution that is also affordable for small and medium enterprises (SMEs). SMEs have not fully adopted big data analysis systems (Gauzelin and Bentz, 2017) as such applications are not primarily accessible to them (Papachristodoulou et al., 2017). This solution can facilitate timely decision making based on CX and improve relationships with customers.

2. PROBLEM STUDIED

Measurements of CX lack clear definitions of the constructs and dimensionalities. Research emphasises the need for the development of robust metrics for the CX measurement (Verhoef et al., 2009; Jain, Aagja and Bagdare, 2017; Lemon and Verhoef, 2016; Zaki and Neely, 2019). Gupta and Zeithaml (2006, p.735) stressed "the need for more studies that view customer metrics comprehensively, rather than examining only a few constructs at a time". Many conceptual models were designed (e.g. Parasuraman, Zeithaml and Berry, 1988; Lemke, Clark and Wilson, 2011; Grewal, Levy and Kumar, 2009; Klaus, 2015; Lemon and Verhoef, 2016; McColl-Kennedy et al., 2018) with different dimensions of research; for comparison see Khodadadi et al. (2016) and Havíř (2017). Prior research has suggested that the customer's assessment of experience influences not only the single share-of-mind metrics such as customer satisfaction, customer loyalty or word-of-mouth, but also customer profitability and customer lifetime value (e.g. Bolton et al., 2004; Verhoef, 2003).

Organisations tend to measure specific aspects of the CX as customer perceptions for a single transaction at a point in time, or as an overarching perception. Customer satisfaction is the dominant customer feedback for measuring perceptions; however, it typically does not capture the full CX as it is concentrated at the end of the customer journey while ignoring the underlying issues and concerns resulting from the experience during the customer journey. The idea of measurement of overall CX at each stage of the customer journey for every touchpoint (Lemon and Verhoef, 2016) is still in an early phase of development. There is no agreement on robust measurement approaches, and no rigorous assessment of metrics that should be collected has been developed to evaluate all aspects of CX across the customer journey (Lemon and Verhoef, 2016; Zaki and Neely, 2019). Existing scales (e.g. Brakus et al., 2009; Klaus, 2015) are aimed at specific research, and they are not understood as parts of the data model.

Researchers stress the importance of measurement of emotions, personality traits and sentiment detected in textual VoC as these CX elements accompany the customers' entire journey (Chen and Lin, 2015; Verhoef and Lemon, 2016; McColl-Kennedy et al., 2018). Personality, along with emotions, is a latent construct of CX. They are the main drivers of customer behaviour, and their determination can recognise behavioural patterns.

Research in service quality analyzes sentiment as an indicator of satisfaction based on text analytics. However, they focus only on the product/service/organisation perspective (Song et al., 2016; Palese and Piccoli, 2016; James et al., 2017; Vencovský, 2018; Farhadloo et al., 2016) rather than CX quality. The sentiment is aggregated for single products or product features, but it is not possible to map the sentiment back to the customer who wrote the comment. Customer perspective is neglected. When appraising an interaction, it is essential to evaluate the polarity of the interactions in a particular context – not only from the viewpoint of a marketing objective, but also the customer perspective – if the company wants to contribute to customer retention.

Customer reviews predominate as a dominant source of VoC. The evaluation of reviews is primarily performed with structured Likert-type scale ratings (Tsang and Prendergast, 2009) and assessment of their effects on purchase decisions with some research focus on sentiment (Zhang et al., 2016; Li et al., 2019). Although companies typically possess quantitative CRM data on customer buying habits and classifications, there is little knowledge about the personality and emotions of these customers and their evaluations. CX is more complicated than simple CRM metrics alone (Zaki and Neely, 2019). Information like opinions and emotions, but also personality, which cannot be found in transactional and other structured data, are partly hidden in customers' written expressions. Metrics should focus more on perceptions and attitudes to gain a comprehensive understanding of customers from their perspective.

Some researchers have built frameworks for automatic analysis of single sources of textual VoC data for BI purposes (Chau & Xu 2012; Peng et al. 2012; Yulianto et al., 2018), however without any context to structure data within the multidimensional data model.

Earlier, Yaakub et al. (2012; 2015) proposed an enhancement to the customer analysis multidimensional data model for the ontology model to calculate and analyse the opinion orientation of some groups of customers for products in certain levels based on the ontology gained from textual customer reviews. The customer analysis model designed by Yaakub (2015) represents the starting point for the CX data model in this article. Yaakub's research misses an integration with other structured customer data on an individual level, so the designed model stands alone without any context to other tables from the customer dimension. However, Yaakub emphasises the importance of the integration of the opinion from textual data with other customer structured data. This research extends the opinion fact table by adding emotions and personality traits linked to the customer dimension. The CX data model significantly expands on Yaakub's model by adding other tables with textual information and references to tables with the structural data from other sources. The CX data model results from the need for CX measurement. The artefact of CX measurements primarily builds on and extends the research in CX conducted by Lemon and Verhoef (2016) and Zaki and Neely (2019).

3. MINING THE CX ELEMENTS FROM THE TEXTUAL VOC

Customer's opinions play a significant role in CX. These opinions are contained in VoC and contain "sentiments, appraisals, attitudes, and emotions toward entities and their attributes expressed in a written text" (Liu, 2015). Analysing of VoC requires text analytics due to the textual expression. Text analytics characterises the content of the unstructured text by subject matter - major and minor topics and by positive and negative sentiment or emotions. The aim of this article is not to find the specific methods of analysing the text, but the way to store the information gained by these methods in a unified data model.

The opinion target is an entity which represents the object of the CX and its aspects described in the customer textual contribution (comment) as depicted in Figure 1. The elements of the CX then reflect perceptions of the opinion targets and are represented by:

1) **Sentiment** expressed about the target objects (positive, negative, neutral) and its intensity.
2) **Discrete emotions** expressed about the target objects (e.g. joy, sadness, trust). The text can be multi-emotional.
3) **Personality traits** of the customer who expressed the opinion (e.g. extroversion, neuroticism) can be determined from the overall expression of the customer. A customer is considered to have more than just one personality, so its intensity must be tracked.

The object can be a product or service, topic, event, person or an issue related to the product, service or company itself about how the customer expresses their opinion. The object can be discussed from different perspectives, which represent different aspects. These aspects can be product/service attributes (features), components, functionality or the dimensions of quality. For example, if the customer buys a sightseeing flight, the object is the *flight* itself, and aspects can be the *plane* comfortability, *price* the customer paid for the flight, *weather* on the day of the flight, or the satisfaction with the *pilot*. All these aspects the customer can evaluate with different words, some of which carrying the sentiment of appraisal words. According to Song et al. (2016), only the crucial aspects should be considered features or components.

Some aspects can be close to each other on the same topic and clustered together into a aspect category under one term. This step reduces the number of different aspects with the same informative value. The aspect category is typically more general than the aspect term itself and does not necessarily occur as a term in the text. For example, if the customer talks about the aspect category *weather*, she or he can use a sentence like “It was a beautiful sunny day without any clouds”. The detected aspects are *sunny day* and *cloud*, and both terms fall under the aspect category *weather*.

For the purposes of CX measurements, satisfaction and expressed emotions about individual objects and their aspects are not only interesting but also the overall customer satisfaction. The customer can write many comments; each comment contains opinions

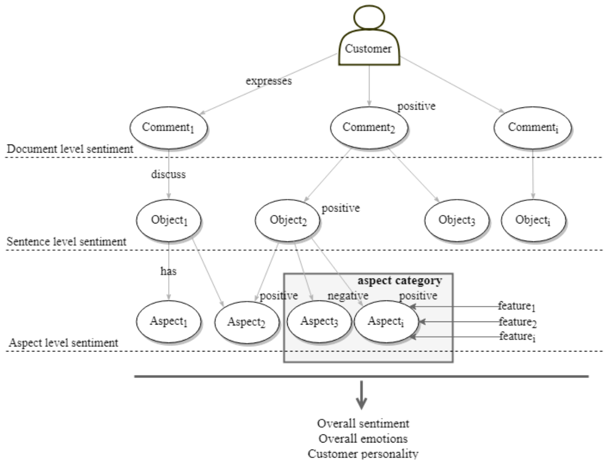

Figure 1 The parts of the VoC content from a single customer perspective.

about different objects with several aspects. The sentiment and emotions are assigned to every aspect for every object in every subjective comment (if there is some detected). The object’s sentiment is derived from the expressed aspects’ sentiment; the comment’s sentiment is derived from the sentiment of discussed objects in the comment. In other words, the overall satisfaction and expressed emotions are gained from the classification of the lower levels of analysis. This approach is consistent with the multidimensionality and enables us to add customer perspective to the CX with preservation of the product perspective by drilling and slicing at lower levels of granularity (for example to measure the average sentiment of a specific aspect of an object from the perspective of a chosen customer segment).

The personality traits are determined from all the comments the customer has written. More textual data ensures a better prediction of the personality. Determination of the emotional elements is possible only from the subjective and evaluative text with the emotional sentiment. The element of personality traits is also possible to determine from the text with a rational sentiment.

4. METHODOLOGY

The research is driven by the design science methodology (Wieringa, 2014). The solution design follows the preliminary research in CX and VoC.

Definition of CX measurement is based on a literature review, which puts existing constructs and metrics into mutual relationships and previous research (Šperková, 2019). Necessary metrics and indicators to be followed by companies were detected in order to measure complex CX.

Design of metrics for CX measurement is based on customer sentiment, customer emotions and personality traits extracted from textual VoC by text analytics methods. The design respects the criteria for the application of text analytics methods to gain the necessary elements, and specifications of metrics and indicators are defined. The metrics are expressed from the BI perspective, according to Kimball et al. (2015).

Design of the multidimensional CX data model is enhanced with stored information extracted from textual VoC. The model respects the principles of multidimensional modelling (Inmon, 2002), and uses the unified modelling language (UML) class-based approach. Based on the target metrics, a method is suggested to store the underlying data for measurement and reporting of CX elements.

The architectural framework for the solution design is depicted in Figure 2. The picture shows the integration process of textual VoC to the CX measurement from the data source collection to the reporting. The approach to the integration of textual VoC applies a textual extract/transformation/load (ETL) process onto the documents and extracts information as values of the entities and their attributes, as found in the text. Information is stored in the multidimensional model as structured data.

The process of integration requires more stages for the transformation of the data. Therefore, three stages were suggested as depicted in Figure 2. First, the textual pre-stage stores the results from a pre-processing phase, which includes data cleaning and feature extraction and selection. These results serve as a data layer for content analysis. Second, the textual stage then stores information gained by text analytics methods above the content. And third, the analytical stage is then linked to the textual stage for the calculation of CX metrics as a logic layer of the BI. The analytical stage stores the structured customers' data with the results of analytical processes (such as data mining models) where CX elements are modelled based on the extracted information from textual data. The textual stage must be loaded first to fill tables in the analytical stage. The loading phase in the model is a part of the textual ETL. The highest layer of the framework is the access layer, which encompasses proposed metrics for reporting.

5. DESIGN OF CX MEASUREMENT

The proposal of the metrics represents the application of CX constructs (Parasuraman, Zeithaml and Berry, 1988; Lemke, Clark and Wilson 2011; Klaus, 2015; Lemon and Verhoef, 2016; McColl-Kennedy et al., 2018) with their constituent elements as the result of a thorough analysis of the current research in CX. The selection of the specific metrics and corresponding dimensions follows the

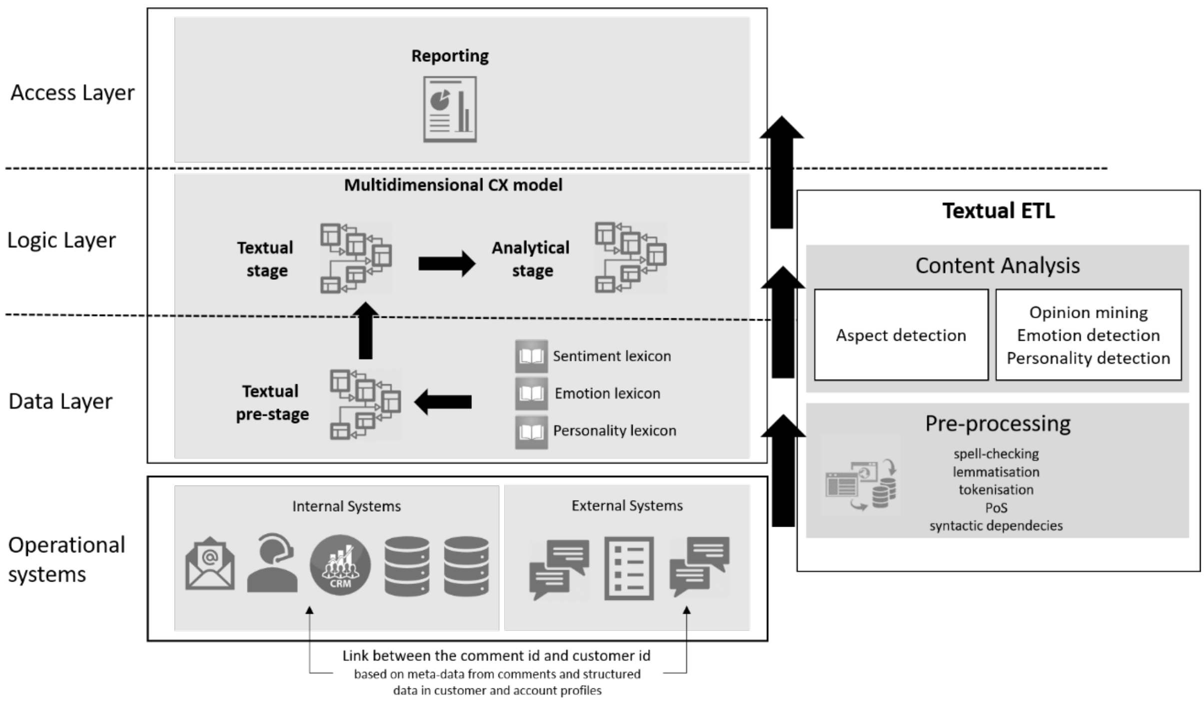

Figure 2 BI framework for the process of VoC integration to CX.

literature review and the results of the quantitative research in Šperková (2019). During the interviews, the participants were asked which metrics and indicators are important for the monitoring and evaluating the CX in their organisations. The measurement is enriched for new elements of customer sentiment, customer emotions and personality traits.

Since CX is in this research assumed a customer perspective during their decision journey, the following metrics were suggested for measurement, to evaluate elements of CX from a customer perspective during the customer journey. The metrics evaluating CX from a company perspective are omitted as the goal is to get to the individual level of the customer. Metrics summarise various aspects of the data in the multi-level aggregated form and are comparable to the surveyed dimensions in the CX research.

The metric is understood as a quantitative or qualitative indicator or an evaluation criterion to assess the level of CX with its constituent elements. The primary purpose is to highlight the relevant facts that the company needs to address and improve the level of CX. The underlying data for the evaluation of the metrics are stored in the data model, which is designed for the querying in order to gain the metrics results. The metrics are supposed to be visualized in reports and dashboards: applications that organise metrics in a clear and intuitive graphical form for further managing CX. The metrics listed in Table 1 are designed on a general level as they can be customised according to the business and available data. The table contains the definition and construction of metrics and related CX elements which ensure a placing of the metrics into the CX construct.

Table 1 Metrics and indicators in CX measurement.

Metric/Indicator	**Definition/Construction/Sub-metrics**	**Related CX elements**
Customer Effort Score (CES)	Determines how much effort a customer has to exert to get a result (issue solved, request fulfilled, product purchased, question answered) on a scale from very easy to very difficult.	Engagement (cognitive); Satisfaction; Personality
Customer Satisfaction Score	Determines satisfaction on a scale from very unsatisfied to very satisfied.	Engagement (affective - intimacy); Satisfaction (evaluative)
Discrete Emotion	Detects customer's primary emotions according to the model of (Plutchik 1980): anger, anticipation, disgust, fear, joy, sadness, surprise, trust. Emotions can be enhanced for other from Plutchik's Wheel of Emotions.	Emotions; Satisfaction; Loyalty (attitudinal)
Emotional Value	Detects the emotional value based on detected emotions on the scale: strongly negative, negative, rational, positive, strongly positive.	Emotions
First Response Time	Calculates the average amount of time elapsed until initial response to the customer's contact according to the type of the contribution: comment type = complaint, suggestion, requirement	Satisfaction
Involvement level	Indicates the involvement level based on the following metrics: number of unique site visits, number of advertising impressions and clicks, number of website page views, time spent per session, time spent per page, number of in-store visits, number of newsletter subscriptions	Engagement
Net Promoter Score (NPS)	Determines detractors, promoters and passive customers. The indicator represents the answer to the question "How likely you would recommend company/product/service to a friend or colleague?" on a scale from 0 = very unlikely to 10 = very likely). Detractors, for a score of 0–6, passives, for a score of 7 or 8, promoters, for a score of 9 or 10.	Satisfaction (evaluative)
Number of cancellations	Indicates the number of cancellations the customer made (i.e. cancel an ordered service).	Satisfaction
Number of complaints	Indicates the number of complaints the customer sent to the company by a summary of individual comments with the type = complaint.	Engagement (cognitive - interaction); Emotions Satisfaction
Number of compliments	Indicates the number of compliments the customer sent to the company by a summary of individual comments with the type = compliments.	Engagement (cognitive - interaction); Emotions Satisfaction
Number of public comments	Indicates the number of public contributions by summary.	Engagement (affective - interaction); Loyalty (attitudinal)
Number of requirements	Indicates the number of suggestions the customer sent to the company by a summary of individual comments with the type = requirement.	Customer expectation; Engagement (cognitive - interaction); Emotions

Metric/Indicator	Definition/Construction/Sub-metrics	Related CX elements
Number of returns	Indicates the number of returns the customer made (for example to cancel the service).	Satisfaction
Number of suggestions	Indicates the number of suggestions the customer sent to the company by a summary of individual comments with the type = suggestion.	Engagement (cognitive - interaction); Emotions
Personality	A mixture of personalities values (openness, agreeableness, conscientiousness, extraversion, neuroticism) according to the Five-Factor model (McCrae and John 1992). Can be visualised as a radar graph.	Personality
Problem Resolution Time	Calculates the average amount of time for resolution of the customer's complaint: between when the customer first creates an issue ticket to when the issue is solved.	Satisfaction
Recency, Frequency, Monetary (RFM)	Determines: Recency – How recently made the customer purchase (interval between the time of the last transaction and first day of each season); Frequency – How often the customer purchases (number of days which occur a transaction during each season; Monetary – How much the customer spent (the average amount of money spent on purchases during each season). The result is the customer's placement in the cube according to binning the scores of frequency, recency and monetary into five equal frequency bins (Kohavi and Parekh 2004).	State in the Customer Journey; Engagement (interaction); Loyalty (behavioural)
Referral Value	Indicates the customer referral value by the following metrics: Reach: Number of impressions/responses/shares (forwarded content) to the customer's contributions; Sentiment of the responses to the customer's contributions; Importance of the responded contacts to the customer's contribution; Number of sent invitations to join the community by customer; Number of public contributions; Sentiment of the shared contributions	Sentiment; Engagement (affective - influence)
Review Score	Indicates the quality of the subject of consumption at a numerical scale.	Satisfaction (evaluative)
Sentiment	Calculates the sentiment of the customer contribution as a value to determine the polarity of the sentiment: positive, if sentiment value > 0, negative, if sentiment value < 0, neutral, if sentiment value = 0.	Satisfaction (emotional); Loyalty (attitudinal); Engagement (affective - intimacy)
Share-of-Wallet	Determines how much of available budget customer spent at the company versus competitors.Customer's total revenue/total spend x 100	Engagement; Loyalty (behavioural)
Value of Knowledge	Indicator of demanding (high value) and difficult customers (low value). Demanding: willing to participate in finding problem solutions. Difficult: requires energy on solving issues without the support of knowledge.	Engagement

In reports, metrics can be viewed from many dimensions, for immediate use in decision-making processes in the organisation. The examples of considered dimensions for filtering, slicing and drilling the metrics are defined in Table 2. The results of metrics can also act as dimensions to filter/slice/drill other metrics. For example, customer satisfaction is determined by customer *sentiment*. Customer sentiment can be assigned both to all the customer's comments as a summary sentiment and to a specific comment, object or aspect only. Further, sentiment can be used to slice the measurement of *most active customers* (determined by a number of comments posted by the customer) and show only those with negative sentiment polarity. All metrics are related to the time dimension, and the customer dimension as the measurement is aiming to the customer perspective. The values in dimensions (e.g. particular segments) can change according to stakeholders‘ needs.

The classification to the segments that can serve further as views to some metrics as dimensions can depend on results of other metrics. For example, the classification into loyalty segments depends on the results of the RFM score and engagement level (spreading positive WoM). The RFM score ascertains if the customer is still alive and makes purchases, but it may be that customer has not purchased for an extended period of time, but still talks positively about the company, thus spreading positive WoM. Such a customer would be in the loyalty matrix more in the left top corner as a latent loyal. The understanding of the causes of weak and negative attitudes in a customer can help companies identify barriers to purchase.

The definition of the segments and the borderlines between the segments depends on the business case and the goals of the company. The right segments should fulfil characteristics of similarity within the segments, differences between the segments, sufficient size of the segment and verifiability over time.

Table 2 Dimensions in CX measurement.

Examples of dimensions resulting from the data model	Description
Channel dimension	Stores the different modes for interacting with customers. Represents the source of data of VoC (e.g. review, email, post on a social network)
Customer dimension	Stores the static information about the customer
Object dimension	Represents the product, service, topic, issue, person or event represented as an object detected in the text
Aspect dimension	Represents the aspects of the object (dimension of quality, functionality, component) detected in text
Comment dimension	Detect type of the comment (complaint, compliment, suggestion, requirement, need)
Time dimension	Universal periods used throughout the model (year, quarter, month, week, date, datetime)
Sentiment polarity	Detect the polarity of the sentiment (positive, negative, neutral)
Loyalty segment	Determines customer loyalty based on a two-dimensional model of (Dick and Basu 1994). The result is the customer's placement in the matrix: *No loyalty* (low repeat purchases, weak relative attitude); *Spurious loyalty* (high repeat purchases, weak relative attitude); *Latent loyalty* (low repeat purchases, strong relative attitude); *Loyalty* (high repeat purchases, strong relative attitude)
Recency, Frequency, Monetary (RFM) segment	Determines the RFM segment based on the measured RFM score. The segments can be refined according to stakeholders' needs. *Loyal customer* (highest recency, highest frequency, highest monetary); *Potential loyal* (high recency, high monetary, more than one purchase); *New customer* (high recency, low frequency); *Attention seeker* (high monetary, high frequency, low recency); *Sleeping customer* (lower recency, lower frequency, lower monetary); *Lost customer* (lowest recency, lowest frequency, lowest monetary)
Customer state in the journey	Dimension extends the RFM result for other information gained about the customer. Detects the customer state in his/her customer journey according to Buttle and Maklan (2015): *Suspect*: potential customer fit the target market; *Prospect*: the customer fits the target market profile and is being approached for the first time; *First-time customer*: the customer makes the first purchase; *Repeat customer*: the customer makes an additional purchase.; *Majority customer*: the customer selects the company as a supplier of choice.; *Loyal customer*: the customer is resistant to switching suppliers and has a strong positive attitude to the company or offer; *Recovered customer* (customer who was considered as lost in last defined period, but purchased in the current period)

6. TEXTUAL STAGE OF THE CX DATA MODEL

The textual stage of the designed data model is a result of the textual ETL and can be divided into a textual pre-stage for results from pre-processing and a textual stage. In the pre-stage, the lexicons are stored for applications of models for content analysis or other data suitable for further processing. Among these lexicons are sentiment, emotion and personality lexicons for detection of these elements, but also domain ontologies for result refinement.

The results of text analyses are stored in the textual stage – sentiment, emotion, personality traits and opinion targets (objects and aspects), which further serve for measurement of CX elements in the analytical stage.

6.1 Input data

The customer interacts with the company and its other customers or prospects through different channels. An example of the process of collection of initial data from two different sources through web crawling and application interface is described in Šperková, Škola and Bruckner (2015). The data are transferred in a clearly defined format suitable for storage in the relational database. The input to the ETL process for the CX model is a structured table with all the raw interactions containing the opinions of the customers. Every single interaction is stored in the database under its identifier. The meta-data of one interaction represents one row in the table, including the raw text of the content. The input table contains at least these attributes:

- *Interaction identifier*: the unique identification of the interaction
- *Contributor identifier*: the unique identification of the user who expresses the comment

- *Source identifier*: the unique identifier of the source of the interaction
- *Timestamp*: the exact time the comment was sent or posted
- *Comment text*: the content of the interaction in plain text

Other attributes can be added (if any exist) containing additional information, for example, other contributors' identifiers, if the comment is accompanied by a rating in Likert-type scale or if the comments belong to different dimensions of experience.

The table stores the raw textual data prior to any text analytics processes, so it is always possible to return to the original text. Every comment discusses at least one target object (opinion target) and not all the target objects discussed in one comment must be correlated. In the textual stage, this table represents the table *Comment* in the model and can be completed by other transformed input data from different channels. The determination of the opinion target is a task for text analytics.

6.2 Pre-processing Phase

The pre-processing of the textual comments is based on a standard feature extraction and selection processes (Liu, 2015). The comments are spell-checked and parsed to sentences based on punctuation, tokenised and lemmatised. After tokenisation, the morphological tags (attribute *Morpho_Tag*) are assigned to the words and store to the relational table *Entity* (Figure 3). The morphological tags are results of the morphological analysis, which works with isolated verbal forms, regardless of their context. Each tag is a string of 16 characters. Every position has its meaning. The first position determines PoS (N for noun, V for verb, A for adjective etc.), the second position contains the detailed determination of the word part. The 11th position determinates negation (Straka and Straková, 2018).

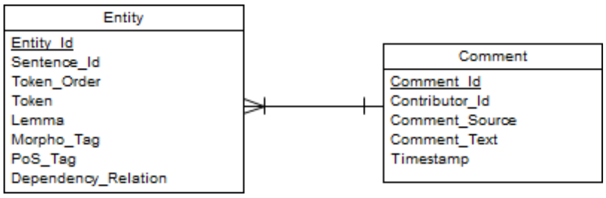

Figure 3 Relationship between the comment and entity tables.

To encompass entire phrases or n-grams in the text as subjects of mining methods (for example an aspect *wellness weekend*) and not only features represented by frequent nouns, adjectives and adverbs, the syntactic dependencies depicting opinion and target relations are also assigned (for example, with the Universal Dependencies treebanks) and stored in the column *Dependency_Relation*. The relational storage of a Czech sentence *"Flying a plane could be cheaper"* after pre-processing is depicted in Table 3.

The pre-processing phase of the textual data serves directly for the text analytics methods, and it is not critical to store the intermediate results to the database. Nevertheless, these results can serve for further improvement and adjustment of the methods or as a domain knowledge corpus which can be enhanced for other attributes, for example, entity type (number, person, organisation and similar) or even sentiment. Such results can be stored in a relational table, which has a relation to the transformed input table *Comment,* as demonstrated in Figure 3. One comment contains many tokenised entities for the recognition of opinion targets and appraisal words in the next steps.

Only the opinion targets (objects and aspects) enter the next phase. Features and other words or phrases – which are representative words of aspects or appraisal words – do not enter the model. They serve only as evaluative words for the modelling of sentiment, emotions or personality traits. The aspect extraction is already a result of content analysis after pre-processing.

Table 3 Relational table with pre-processing results.

Comment_ Id	**Sentence_ Id**	**Token_ Order**	**Token**	**Lemma**	**Morpho_Tag**	**PoS_Tag**	**Dependency _Relation**
2899	4	1	Létání	létání	NNNS1-----A----	NOUN	nsubj
2899	4	2	by	být	Vc--------------	AUX	aux
2899	4	3	mohlo	moci	VpNS---XR-AA---	VERB	root
2899	4	4	být	být	Vf--------A----	AUX	cop
2899	4	5	levnější	levný	AANS1----2A----	ADJ	xcomp
2899	4	6	.	.	Z:--------------	PUNCT	punct

6.3 Conceptual Data Model of the Textual Stage

The textual stage represents the entities capturing the tacit knowledge available in textual comments. Figure 4 depicts the underlying conceptual model proposed to capture customers' opinions. The model shows only the necessary attributes for storing the textual data. The relation to ontology tables (emotion/personality/sentiment lexicons) and history tables is not depicted due to the readability of the model. For simplification, only concepts related to *object* and *aspect* are shown, as they are considered sufficient for the model's needs.

The model extends and builds on the knowledge of Yaakub (2015). The issue of Yaakub's model for opinion is that the fact table can store only one feature (aspect) per comment. This research adds the fact table *Opinion* into the conceptual model to gain a whole feature hierarchy.

In contrast to Yaakub's multidimensional model, the CX model needs to relate satisfaction and sentiment to relevant customers. Therefore, it is not sufficient to keep the sentiment value only for the particular aspect in the *Aspect* table, but it is needed to determine which customer holds this opinion as opinions differ from customer to customers.

In the conceptual model, the *Contributor* table as the contribution (comment) can also be written by a person who is not a customer of the company yet (i.e. potential customer or detractor). In a snowflake schema, it is possible to link the *Contributor* entity through the *Comment* table to the fact table *Opinion* and filter to the specific contributor. Then it is possible to find out the overall satisfaction of the contributor or their satisfaction with a particular comment, object or aspect.

The *Comment* class stores the full content of each customer's contribution, including the date when the comment was written. The source of each comment is stored in the class *Comment_Source* with values like 'email', 'call centre', 'social media', 'review' and similar. The *Comment_Type* determines the type of the comment based on the detected information in the text – whether it is a 'requirement', 'complaint', 'compliment', 'suggestion' or 'need'. If the comment is a review accompanied by the rating in a Likert-type scale, the attribute *Rating* gets its value.

Each *Comment* is written by a *Contributor,* and the relationship is many to one, as one contributor can write many comments. The

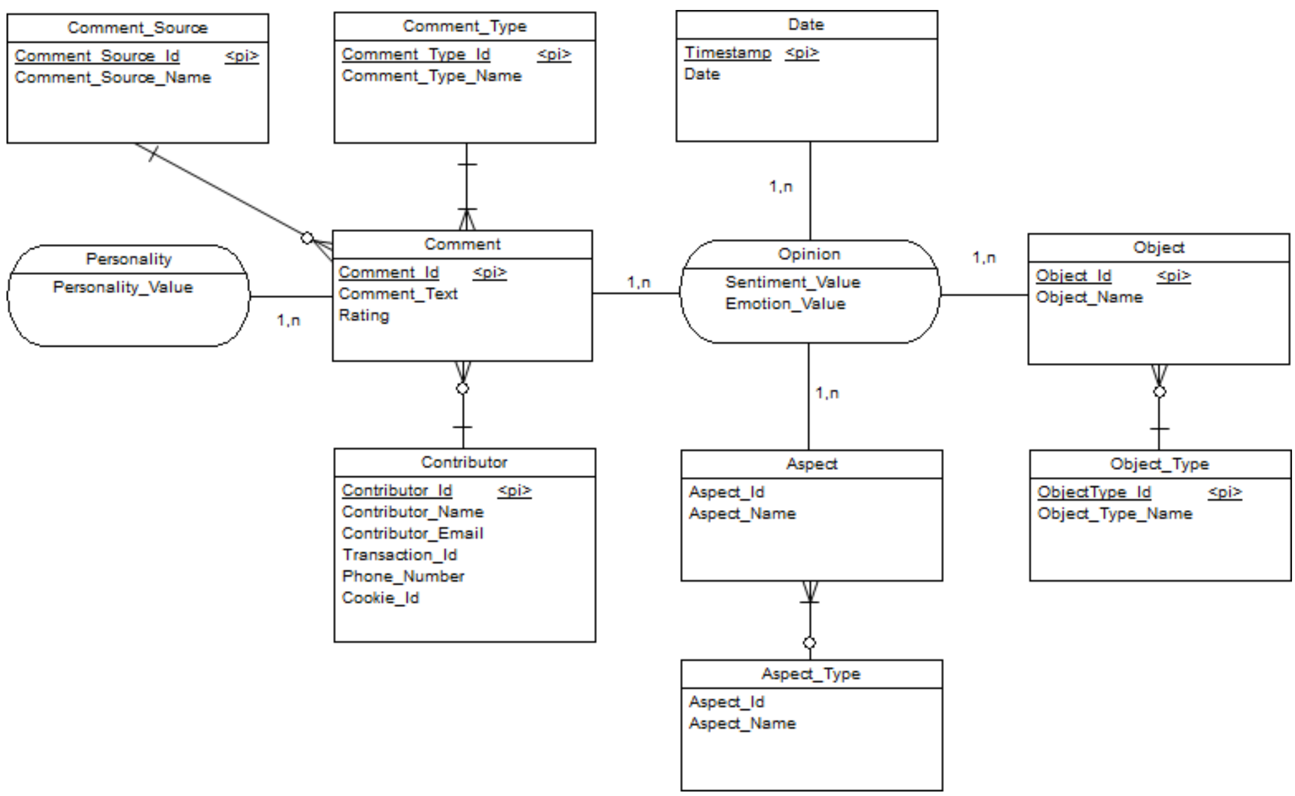

Figure 4 Conceptual model of the textual stage.

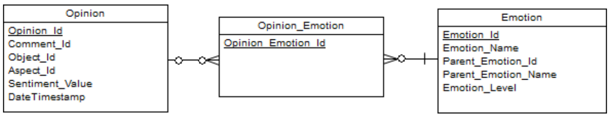

Figure 5 The m:n relation between opinion and emotion.

Contributor table comes with attributes identifying the contributor based on cookies, email, name, telephone number, transaction number or other identifiers which can distinguish the author of the comment based on meta-data gained from the comment. Since some comments are not tagged with customer or account identifiers, there is no direct way to link such comments with a particular customer or account in the database. During ETL, these attributes are further matched with existing information presented in the customer and account profiles, and the best match is then linked with the *Customer* table at the analytical stage. Contributors without a match get the attribute *Flag_Active* with the value 'non-active' in the *Customer* table. The linking of customer profiles with customer interactions is an essential step in ETL as it brings together the factual information about the customer gained from structured data with the factual information gained from the textual interaction used later in the CX data model.

The *Object* table stores the title of the discussed object. This table can represent any entity such a 'product', 'issue', 'service', 'event', 'person' determined in table *Object_Type*. The *Object* table allows the model to be multi-domain since it is possible to store comments on a wide range of topics.

The *object* can be recognised from the contribution based on the relation to the product/service it belongs to: metadata (i.e. review submitted to a particular product, an email regarding the product the customer purchased), the discussed domain or based on a dominant topic in the text. The *object* is represented by a finite set of its aspects $A = \{a_1, a_2, \dots a_n\}$. A customer contribution (stored in the model in the *Comment* table) contains opinions about a finite set of objects {*o1*, *o2*, …, *ov*} and a subset of aspects of each object.

The m:n relation between the tables *Object* and *Aspect* replaces the idea that the product or service, which is the subject of the comment, is always classifiable into a hierarchy or family of products or services (Yaakub 2015; Lau et al. 2009). The basic idea of m:n relations between the tables *Object*, *Aspect* and *Comment* is that a negative comment about the object (product) does not mean that the customer gives a negative opinion on its aspects, or negative comment about the aspect, as this does not necessarily mean that whole comment is negative too. Also, the same aspect can be assigned to different objects (the aspect 'battery' is associated with both the object 'mobile phone' and object 'laptop'). If the statement does not mention any aspect at all (overall experience), then the evaluation stored in the table *Opinion* is assigned to the *object* in table *Object*.

The fact table *Opinion* contains the information detected by text analytics methods and transformed into structured data. The table contains an identifier to dimension tables *Comment*, *Object* and *Aspect* as the relation between these tables is m:n. The table stores the calculated *Sentiment_Value,* which serves in reporting for sentiment polarity determination:

- positive, if Sentiment_Value > 0,
- negative, if Sentiment_Value < 0,
- neutral, if Sentiment_Value = 0

If the *Sentiment_Value* of all objects and aspects in the comment is zero and no emotions are detected, the comment is considered rational. The attribute *Rationality_Flag* in the *Comment* table determines the character of the comment based on detected opinion – if the comment is 'rational' or 'evaluative'. The rationality is determined by detected emotions and sentiment in the text.

The *Opinion* table also contains flags for eight primary discrete emotions according to Plutchik (1980). If the model were expanded to more emotions from the Plutchik's Wheel of Emotions, the schema would have to change, and the relational table *Opinion_Emotion* would extend the model (Figure 5) as the relation is m:n – one opinion can contain more emotions.

As personality detection is based on whole comments, the fact table *Personality* is related to the *Comment* table only. The detection of the personality traits depends on the expression as a whole and does not relate to an aspect or object. The table stores a calculated value for every personality trait according to the Five-Factor Model (McCrae and John 1992). The values are then aggregated on the customer level through related comments.

The *Date* table inserts the dynamic character of CX into the model and enables tracking information over time. *Timestamp* is an essential attribute for the reporting, considering that a customer can have an inconsistent experience during the iterative customer journey. The *Personality* table is not linked to the *Date* table due to the assumption that the personality does not change with time. The personality prediction can be refined with additional textual data.

7. ANALYTICAL STAGE OF CX DATA MODEL

The analytical stage of the data model builds mainly on the knowledge of customer intelligence and exploits tables used in analytical CRM following the designed metric. The analytical stage modelled in Figure 6 as a physical model depicts the interconnection of the textual stage with the tables typical in analytical CRM (e.g. personal information, sociodemographic data, product preferences), but also with the tables resulting from other sources of EIS:

- transactional data (orders, sales, etc.)
- campaign data (campaign costs, budgets, plans from campaign management systems)
- web data (click-stream data and other data from web analytics platforms)
- results of data mining and other analytical processes as aggregated data (e.g. CLV).
-

These aggregated data serve as underlying data for metric reporting. For this reason, the data model is denormalised. The denormalisation also enables easier querying for analytical purposes. It is emphasised that not all tables are depicted in the model as the complexity changes based on available sources of data and elements detected in the text. The model can contain several dimensions depending on the granularity level of the measured CX. The model corresponds to a part of a complex analytical model, which is linked to several other entities. It provides a data mart for CX measurement, which can, in turn, be extensible for new entities and attributes. The model in Figure 6 presents the fact and dimension tables necessary for metric reporting with examples of attributes.

The display of relations to the *Timestamp* and *Date* dimensions are omitted to keep the clarity of the model. Only the relation between the *Opinion* fact table and the *Timestamp* dimension table are kept to demonstrate the time dimension of the model. In reality, all fact tables have links to the *Timestamp* or *Date* dimension table to ensure the history maintenance with snapshots and storage in historical tables.

The relationship to the dimension tables *Customer* and *Product* is depicted in the physical model. Since the table *Object* can represent any entity, such as a product or service, it is desirable to map values to the right tables according to the dimension table *Object_Type*. Following this information, the *Object* table has links to other appropriate tables. The mapping to the right object type is built on the similarity rules. If the value of the attribute *Object_Name* in the table *Object* is founded in the attribute *Product_Name* of the *Product* table, the row containing this value also gets the attribute *Object_Id* mapped to the *Object* table. The principle with other classes would be similar. Comparably, the *Aspect_Type* table determines the mapping of the *Aspect* table to other internal dimension tables.

The *Customer* table replaces the *Contributor* table from the textual stage, and the unidentified customers (contributors without a match) get the 'non-active' value to the attribute *Flag_Active* as such a contributor has not made a transaction with the company.

Except for the *Product* and *Product Category* tables, the analytical stage presented in Figure 6 expands the textual stage for other tables loaded from internal systems listed in Table 4**Error! Reference source not found.**.

8. IMPLICATIONS

The CX data model and the subsequent measurement bring significant benefits for CX measurement and management. The artefact mitigates the barriers in achieving the full potential of analysing VoC within CX, detected in Šperková (2019).

The model represents the application of the CX construct. The model brings a certain formalisation to CX measurement and management. The CX data model can help with a scope definition of measures needed to be monitored in a company.

The model enables the necessary integration of textual VoC from various channels and links this data to operational data, data from web analytics and other sources at one consolidated place accessible to all stakeholders. The model is extensible and transferable to any business environment. New entities, attributes and related metrics and dimensions can always be defined. Connectors for new sources of data can be added.

The model is multidimensional and enables one to monitor elements from different viewpoints; dimensions allow querying specific subsets of data. Data which contain the specific forms of searched objects, aspects or comments are then displayed. The textual part of the model for storing the information from textual content leverages the use of insight gained from textual VoC within the share-of-mind metrics and significantly simplify the sharing of knowledge throughout the organisation. Thanks to dimensionality and collection of data across different touchpoints with the time dimension, the model enables one to measure the experience during the customer journey.

Due to the consistency with other trusted data in unified storage, the integrated data model guarantees higher credibility and accuracy of textual VoC and its subsequent measurement, which in the Internet environment may not be satisfied. Consolidation enables the reduction of random, time-consuming and error prone processes with less human effort, which is challenging to scale with the growing data. The connection to other financial data such as purchases, marketing costs, the performance of the channels and similar help to prove the financial results of CX actions.

The model reflects the customer perspective of the opinion target while the product perspective is not omitted. It is possible to aggregate the sentiment according to the particular object or aspect based on all comments from all customers who mentioned that aspect in these comments. The view on customers becomes unified, and their data stored in fine granularity at the individual level enables targeted one-to-one actions. The model enables employees to communicate with the customer consistently through all channels based on shared knowledge within the organisation.

Long-term monitoring of metrics within the consolidated reports allows finding patterns in CX and taking the corresponding approach or prevent certain situations. For business users, the reporting of metrics on dashboards brings the visibility and clarity of all monitored metrics and their instant overview of improving or deteriorating. The close co-operation of analysts and end-users is necessary. End-users must understand the essence of the metrics to be able to work with them correctly. This approach leads to continual improvement of CX and growth of agility, profitability and orientation to the customers.

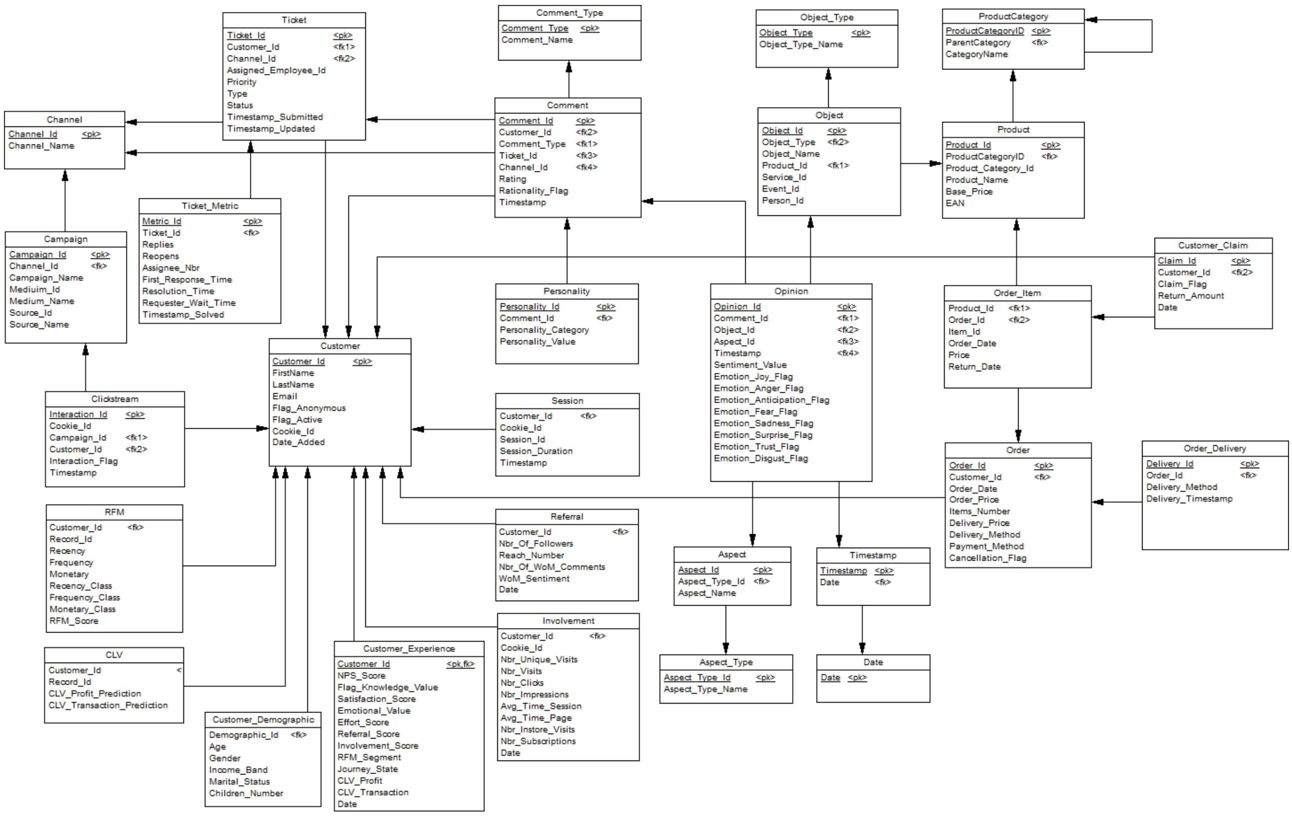

Figure 6 Physical model of the analytical stage of the CX data model.

Table 4 Tables added in the analytical stage.

Table name	**Description**
Channel	The Channel dimension table replaces the Comment_Source table from the textual stage. It contains all possible channels through customer interactions, not only with textual expression. This table represents the interconnection with other sources of the data.
Campaign	The Campaign dimension table extends the Channel for another granularity which represents campaigns the customer interacts through the web. The table can have foreign keys to tables Medium, Source, Placement or Banner depending on the granularity level. If it is possible to react to campaigns with textual data, the reference from the table Comment would be modelled.
Clickstream	The Clickstream fact table collects data from web analytics tools. It represents the interactions of individual cookies with individual campaigns. The attribute Interaction_Flag determines if the interaction was click or impression. If the Cookie_Id is recognised and linked to the Customer_Id, the reference with the Customer table is linked.
CLV	In the CLV fact table predictions are stored from the CLV modelling by different CLV models – the prediction of the transactions and profit for the next period for every customer.
Customer_ Claim	The Customer_Claim table stores the data about the customer's claims on purchased items. The Claim_Flag attribute determines if the claim is a replacement of the item, compensation or money return.
Customer_ Demographic	The table expands the table Customer for demographical data. This table serves for segmentation customers based on demographical data.
Customer_ Experience	The Customer Experience table serves for storage the results of different metrics or classifications to different segments which represent constituent elements of CX. For example, RFM_Segment is based on the results from table RFM. This table serves for easier querying to gain the results faster and preserving history values. Otherwise, these values can be found in other tables.
Involvement	The Involvement table stores the aggregated data from web analytics tools which serve as metrics. If the Cookie_Id is recognised and linked to the Customer_Id, the reference with the Customer table is linked.
Order	The Order fact table contains information about customer's orders. If the order was cancelled during the process of the purchase, the Cancellation_Flag gets the positive value. The table can contain many attributes regarding the prices, methods of payment, delivery and similar.
Order_Item	The Order_Item table represents items purchased within the Order. The table has a reference to the table Order. The attribute Return_Date represents the date of return if the customer returned the item. The table can have many attributes with references to additional tables like Service, if the item is coming with additional services.
Order_ Delivery	The Order_Delivery table stores the information about the timestamp when the order was delivered to the customer. Based on this information, the average time of delivery can be measured.
Referral	The Referral table stores aggregated data from social networks analysis as metrics.
RFM	The RFM table stores the information for RFM calculation – frequency, recency. Monetary values together with the assigned bin and segment. This table is a result of calculations and modelling based on the table Order.
Session	The Session table is based on data from web analytics tools and stores the information about the customer's visits on the company's websites. If the Cookie_Id is recognised and linked to the Customer_Id, the reference with the Customer table is linked.
Ticket	The Ticket table represents the customer's claim, requirement, need or complaint submitted to the company. The ticket can be composed of a thread of comments. The table has a reference to the Channel table, representing the channel through the customer submitted the ticket; the Assigned_Employee_Id can represent the foreign key to the Employee table (not shown in the model). The Timestamp_Submitted shows the time when the ticket was sent to the company and Timestamp_Updated records every update in the ticket.
Ticket_ Metric	The Ticket_Metric table serves for a calculation of metrics based on the information from attributes of the Ticket table. The table records number of replies to every ticket (number of comments), number of reopens, how many employees were assigned to the ticket till his solution, the time the customer had to wait to the first response to the ticket is stored in First_Response_Time, the Resolution_Time stores the total time from the first contact to the solution of the ticket. Requester_Wait_Time records the total time the customer spent waiting for the response.

9. CONCLUSION

The designed CX data model represents the application of CX constructs from previous research. The data model can be understood as a data mart of the data warehouse. The model builds on the knowledge from the customer analysis model designed by Yaakub (2015). It is divided into textual and analytical stages, where the analytical stage is dependent on results from the textual stage. The metrics are expressed from the BI perspective based on dimensional modelling as indicators and their characteristics, analytical dimensions and their characteristics, and the relationship between dimensions and indicators. The elements of emotions, sentiment and personality traits are automatically detected from textual VoC data with text analytics methods, which are the subject of further research. The mined information is joined with other operational, transactional and behavioural structured data from various systems in the unified multidimensional data model.

The artefact solves the complexity of understanding the customer base by implementing a sophisticated data-driven approach to the comprehensive measurement of overall customer experience. The collection of all customer data from multiple channels, with which the customer interacts during their journey, into singular storage with unified access enables one to look at all customer data from the customer perspective according to the time dimension and evaluate their experience in a timely manner and their state in the journey. The artefact brings customer experience measurement to a new level of customer insight which drives loyalty across different channels. This contributes to retention marketing efforts in companies and provides a direct customer-oriented approach, replacing mass marketing conducted on aggregated data.

The aim of further research is to implement this CX data model and measurement in different business domains to validate its usefulness and portability in practice. The development of the application based on the artefact that enables collaboration, management of marketing activities, or alerting can also be a topic for further research.

10. REFERENCES

Aziza, A., Oubrich, M., & Søilen, K. S. (2015). The impact of CRM on QoE: An exploratory study from mobile phone industry in Morocco. *Journal of Intelligence Studies in Business*, *5*(2).

Bolton, R. N., Lemon, K. N., & Verhoef, P. C. (2004). The theoretical underpinnings of customer asset management: A framework and propositions for future research. *Journal of the Academy of Marketing Science, 32*(3), 271-292.

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

Buttle, F., & Maklan S. (2015). *Customer relationship management: concepts and technologies*. Third edition. Routledge.

Chau, M. & Xu, J. (2012). Business Intelligence in Blogs : Understanding Consumer Interactions and Communities. *MIS Quarterly*, 36(4), 1189–1216.

Chen, S. C., & Lin, C. P. (2015). The impact of customer experience and perceived value on sustainable social relationship in blogs: An empirical study. *Technological Forecasting and Social Change, 96*, 40-50.

Farhadloo, M., Patterson, R. A., & Rolland, E. (2016). Modeling customer satisfaction from unstructured data using a Bayesian approach. *Decision Support Systems, 90*, 1-11.

Gauzelin, S. and Bentz, H. (2017). An examination of the impact of business intelligence systems on organizational decision making and performance: The case of France. *Journal of Intelligence Studies in Business*, *7*(2).

Godes, D. & Mayzlin D. (2004). Using Online Conversations to Study Word-of-Mouth Communication. *Marketing Science*, 23(4), 545-60.

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

Inmon, W. H. (2002). *Building the Data Warehousing.*

Gupta, S., & Zeithaml, V. (2006). Customer metrics and their impact on financial performance. *Marketing science*, *25*(6), pp. 718-739.

Havíř, D. (2017). A comparison of the approaches to Customer Experience Analysis. *Economics and Business*, *31*(1), 82-93.

Jain, R., Aagja, J., & Bagdare, S. (2017). Customer experience–a review and research agenda. *Journal of Service Theory and Practice*, *27*(3), 642-662

Khodadadi, P., Abdi, F., & Khalili-Damghani, K. (2016). An Integrated Model of Customer Experience, Perceived Value, Satisfaction, and Loyalty in Electronic Stores. *International Journal of Enterprise Information Systems (IJEIS), 12*(4), 31-46.

Kimball, R., Ross, M., Mundy, J., & Thornthwaite, W. (2015). *The kimball group reader: Relentlessly practical tools for data warehousing and business intelligence remastered collection*. John Wiley & Sons.

Klaus, P. (2015). *Measuring customer experience: How to develop and execute the most profitable customer experience strategies*. Springer.

Klaus, P., & Maklan, S. (2013). Towards a Better Measure of Customer Experience. *International Journal of Market Research, 55*(2), 227–246.

Lau, R. Y., Lai, C. C., Ma, J., & Li, Y. (2009). Automatic domain ontology extraction for context-sensitive opinion mining. *ICIS 2009 Proceedings*, 35-53.

Lemke, F., Clark, M., & Wilson, H. (2011). Customer Experience quality: an exploration in business and consumer contexts using repertory grid technique. *Journal of the Academy of Marketing Science*, *39*(6), 846-869.

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

Li, X., Wu, C., & Mai, F. (2019). The effect of online reviews on product sales: A joint sentiment-topic analysis. *Information & Management*, *56*(2), 172-184.

Liu, B., 2015. *Sentiment analysis: Mining opinions, sentiments, and emotions*. Cambridge University Press.

Liu, Y. (2006). Word of mouth for movies: Its dynamics and impact on box office revenue. *Journal of marketing*, *70*(3), 74-89.

McColl-Kennedy, J.R., Zaki, M., Lemon, K.N., Urmetzer, F. and Neely, A. (2019). Gaining customer experience insights that matter. *Journal of Service Research*, *22*(1), 8-26.

McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of personality, 60*(2), 175-215.

Nahilia, W., Rezega, K., & Kazara, O. (2019). A new corpus-based convolutional neural network for big data text analytics. *Journal of Intelligence Studies in Business*, *9*(2).

Papachristodoulou, E., Koutsaki, M. and Kirkos, E. (2017). Business intelligence and SMEs: Bridging the gap. *Journal of Intelligence Studies in Business*, *7*(1).

Parasuraman, A., Zeithaml, V. & Berry, L. (1988). SERVQUAL: a multiple-item scale for measuring consumer perceptions of service quality, *Retailing: critical concepts*, 64(1), 140.

Park, B. K., & Song, I. Y. (2012). Incorporating Text OLAP in Business Intelligence. In *Business Intelligence Applications and the Web: Models, Systems and Technologies*, 77-101.

Peng, W., Sun, T., Revankar, S., & Li, T. (2012). Mining the “voice of the customer” for business prioritization. *ACM Transactions on Intelligent Systems and Technology (TIST)*, *3*(2).

Plutchik, R. (1980). Emotion: A Psychoevolutionary Synthesis, Harper and Row. *New York.*

Šperková, L., 2014. Word of Mouth analysis on facebook in banking. In: *Marketing identity*. Smolenice, pp. 236-252. Trnava: Univerzita sv. Cyrila a Metoda v Trnava.

Šperkova, L., 2019. Qualitative Research on Use of Voice of Customer in Czech Organisations. *Journal of Systems Integration*, *10*(2), pp.9-18.

Šperková, L. and Škola, P., 2015, September. Design of Metrics for e-Word-of-Mouth Evaluation From Unstructured Data for Banking Sector. In *European Conference on Knowledge Management* (p. 717). Academic Conferences International Limited

Šperková, L., & Škola, P., 2015. E-WoM Integration to the Decision-Making Process in Bank Based on Business Intelligence. *Proceedings of the 23nd Interdisciplinary Information Management Talks*, pp. 207-216. Springer.

Šperková, L., Škola, P. and Bruckner, T., 2015. Evaluation of e-Word-of-Mouth through Business Intelligence processes in banking domain. *Journal of Intelligence Studies in Business*, *5*(2).

Tsang, A. & Prendergast, G., 2009. Is a 'star' worth a thousand words? *European Journal of Marketing*, 43, 1269–1280.

Verhoef, P. C. (2003). Understanding the effect of customer relationship management efforts on customer retention and customer share development. *Journal of Marketing, 67*(4), 30-45.

Verhoef, P. C., & Lemon, K. N. (2016). Advances in customer value management. *Handbook of Research in Relationship Marketing*, 75-103.

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

Wieringa R. (2014). *Design science methodology for information systems and software engineering.*

Wu, W., & Zheng, R. (2012). The impact of word-of-mouth on book sales: review, blog or tweet?. In *Proceedings of the 14th Annual International Conference on Electronic Commerce,* 74-75.

Yaakub, M. R. (2015). *Integration of opinion into customer analysis model*. PhD thesis. Queensland University of Technology.

Yaakub, M. R., Li, Y., Algarni, A., & Peng, B. (2012). Integration of opinion into customer analysis model. In *Proceedings of the The 2012 IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technology, 03,* 164-168.

Yulianto, M., Girsang, A.S. and Rumagit, R.Y. (2018). Business intelligence for social media interaction in the travel industry in Indonesia. *Journal of Intelligence Studies in Business*, *8*(2).

Zaki, M., & Neely, A. (2019). Customer Experience Analytics: Dynamic Customer-Centric Model. In *Handbook of Service Science, Volume II.* 207-233.

Zhang, X., Yu, Y., Li, H., & Lin, Z. (2016). Sentimental interplay between structured and unstructured user-generated contents: An empirical study on online hotel reviews. *Online Information Review, 40*(1), 119-145.