How to Measure Quality of Service Using Unstructured Data Analysis: A General Method Design

Lucie Sperková, Filip Vencovský, Tomas Bruckner
Department of Information Technology, University of Economics Prague
lucie.sperkova@vse.cz, filip.vencovsky@vse.cz, bruckner@vse.cz

DOI: 10.20470/jsi.v6i4.242

Abstract: The aim of the paper is to design a general method usable for measuring the quality of the service from the customer’s point of view with the help of content analytics. Large amount of unstructured data is created by customers of the service. This data can provide a valuable feedback from the service usage. Customers talk among themselves about their experiences and feelings from consumption of the service. The design of the method is based on a systematic literature review in the area of the service quality and unstructured data analysis. Analytics and quality measurement models are collected and critically evaluated regarding their potential use for measuring IT service quality. The method can be used by IT service provider to measure and monitor service quality based on World-of-Mouth in order to continual service improvement.

Key words: Service Quality measurement, Quality in Use, Service Experience, unstructured data, Customer Voice, Word-of-Mouth, method design, content analysis, sentiment, semantic analysis

1 Introduction

A continual service improvement is needed to produce a high-quality IT service that supports stakeholders’ needs. This includes monitoring and evaluating service quality activities. Usually, technical parameters of IT service such as service availability, responsiveness or reliability are taken into account. In fact, the core of the service quality lies in a value for stakeholders. Stakeholders, especially customers, are the main reason for existence of the service. They perceive the service quality in a different way and therefore it is very difficult to manage it. They discuss their opinions of the service quality via social networks and message boards, interact with a service provider by using the same tools and use electronic questionnaires, email and voice communication as well. Collected data can be used in content analysis and identify existing problem areas. Outcomes of evaluation of the service quality can help determine whether the value is provided to customers correctly and also specify extent of the specific problem areas.

Most of the reviewed models in the Quality of Service (QoS) area does not use computer aided analysis of customer’s voice, although it could be useful and bring new information from customers. Many models use as a method of collection data survey questionnaire approach or interviews. Problem of those methods is lack of spontaneity which can be gained from the Internet resources as a Word-of-Mouth (WoM).

Lepmets et al. (2012) brings the comprehensive IT service quality measurement framework based on large literature research. They divided measures into intrinsic which are able to objectively measure and extrinsic which are subjective therefore difficult to measure and interpret. Intrinsic measures are assessable independently around the system, extrinsic measures depending on the system surrounding. As an example, intrinsic aspect is measured response of the system in exact time, extrinsic is the customer’s perception of the response - too slow. The authors of framework, so as ISO SQuaRE (ISO/IEC 25000, 2014) standard, claims that intrinsic aspects are close, or more precisely, prerequisite of extrinsic aspects. For evaluation of service quality, intrinsic metrics seems to be more useful because a provider of a service may detect the problem earlier. On the other hand, intrinsic metrics separately cannot predict customer behaviour sufficiently. Only customer feedback offers deeper understanding of problems perceived by the customers. Lepmets et al. (2012) proposes exclusively surveys as customer feedback measure in their framework.

The most relevant papers aimed at computer aided QoS evaluation uses Latent Semantic Analysis (Ashton, 2014) or Machine Learning Classification (Pai et al., 2012) as the method. These methods are well known but are uneasily to implement in service practice. For this purpose, the powerful tool Elasticsearch seems to be adequate. There are only a few academic articles, which use Elasticsearch.
in their research. These articles are focused on library science (Johnson 2013) and full-text searching (Divya & Goyal, 2013) or big log data (Bai, 2013). In a QoS area, Elasticsearch is used in Hierarchical Multi-Log Cloud-Based Search Engine (Singh and Velez, 2014). Although the aim of the paper is QoS, it operates with service log data. Textual data analysis was the part of theses elaborated at the Department of Information Technologies at the University of Economics in Prague this year. These thesis uses unstructured data as the input and Elasticsearch as a tool for data analysis. Methodology used in those theses are well-conceived and executed but lacks business and QoS context.

In this paper we bring business and QoS context through a method design for unstructured data analysis of service experience WoM. The method uses freely available tools and techniques. We select the methodological frame first. Into the frame, we put the specific techniques related to QoS area.

2 Service Quality models

Many service quality models have been proposed since the eighties (e.g. Grönroos, 1984; Heywood-Farmer, 1988; Ennew et al., 1993; Philip & Hazlett, 1997; Robledo, 2001). Most of them are based on and revisit SERVQUAL model (e.g. Asubonteng et al., 1996) developed by Parasuraman et al. (1985, 1988). This approach applies in particular customer relationship and understands service quality as the gap between customers’ expectations and their overall assessment or perception of quality of provision. The gap is related to different levels of expectation and perceptions result from customers and service’s point of view. The aim of those models is to recognize client’s requirements and measure their satisfaction. Results are useful in the process of service performance improvement towards a more complete fulfillment of the clients’ expectations in the context of the service value analysis. Customers’ feedback allows identifying the strengths and weaknesses of the service.

Seth et al. (2005) reviewed and evaluated 19 service quality models developed till that time. They defined two main categories of the models, the first set of models are developed using the already mentioned model SERVQUAL, also known as the Gap Model and are based on expectation-confirmation theory. The other set of models are different from the Gap Model. A performance-based approach to the measurement of service quality SERVPERF (Cronin & Taylor, 1992) model is the well-known representative of the second category. In contrast to SERVQUAL, SERVPERF uses for evaluation just perceived service performance rating. Cronin and Taylor (1994) argue, that the expectations of customer are implicit in perceived performance and it is unnecessary to measure them separately.

IT focuses more on the technical parameters of the service (Voříšek & Basl 2008; Taylor et al. 2007a). IT service is such a service which is delivered by IT provider to support the business of its customers, created by IT processes that consume IT resources during the course. Grönroos (2000) distinguishes between the services that are directly based on the use of IT artefacts (hardware, software, data) and services where the main role play staff. The service is implemented on the basis of service level agreement. Comprehensive model for IT service should respect both specifics of IT services and principles of customer approach utilized in the theory of services.

In addition to this research, it is necessary to take in account to whom is the service delivered. The definition above is more accurate for service delivered to business (2B), but IT service is also consumed by individual end customer (2C), which ordered the service for his own needs. Some services 2C are delivered through 2B middle chain (internet banking, television broadcasting). The importance of this detection will be discuss later in this paper.

The first IT related service quality model were published by Berkley and Gupta (1994). This model describes in detail where IT had been used or could be used to improve specific service quality dimensions. The model shows that IT strategic alignment is needed, but it only highlights the impact of IT on service quality. The model does not offer a way to measure and monitor service quality and is silent about the level of IT use for particular service setting. Pitt et al. (1995) fitted SERVQUAL for use in IT service area. Jiang et al. (2002) used perceived quality questionnaire from the IT personnel point of view to gain overall service quality. Conceptual model of e-service quality by Santos (2003) extends the SERVQUAL dimensions to reflect more high-tech service issues. This model is not based on any statistical analysis and does not provide any specific measurement scales. IT-based model (Zhu et al., 2002) focuses on the linkages among the service dimensions as measured by SERVQUAL, the constructs representing the IT-based service quality, preferences towards traditional services, experiences in using IT-based services, and perceived IT policies. This model does not provide a measure of service quality of IT-based transactions.
Seath et al. (2005) identified some research issues in relation to IT and service quality:

- What type of information system architecture is needed for effective delivery of quality service?
- How to listen to the voice of customer through information systems?
- How frequently the information systems need collect data related to customer perceptions and his/her possible behaviour?

The last two we are taking account in our method design as they relate to customer feedback and measurement of the service quality.

3 Customer perspective in Service Quality

One of the factors entering the service quality is customer satisfaction and perception of the service. As was given above, service quality is the gap between customer’s expectations and their overall assessment of the service. If we emphasize the word “overall”, we can claim that this assessment can include customers’ experience with the service or the usability of the service for the customer.

According to ISO 9241-11 usability is a way to measure a service ability to help a user to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use. It is dependent on the service, task, user and circumstances. In ISO 9241-210 definition states that user experience is a person’s perceptions and responses that result from the use and/or anticipated use of a product, system or service. In this article we will use the term Quality in Use presented in ISO/IEC 25000.

The concept of service experience is one of the most important references for consumer decision making and the issue is how to effectively extract valuable information from it (Pai et al., 2012). One of the most common ways of expressing a service experience is Word-of-Mouth (WoM). Huang et al. (2012) define WoM as a process when individual transmits his personal attitude to another person. According to Hu, Paul and Zhang (2006) it involves all the informal communication directed other consumers about the ownership, use or characteristics of the product or service. It could be any claim with the sentiment - positive, neutral or negative.

Already Grönroos (1984, p. 42) observed that Word-of-Mouth has a more substantial impact on potential customers then traditional marketing activities, and also highlighted the need for service quality research based on consumers’ views (Seath et al., 2005). Also Parasuraman et al. (1985) used WoM as a key contributor to the expected service in their SERVQUAL model.

3.1 Measurement of Quality in Use

One of the most popular methods to evaluate service experience and usability (or Quality in Use) is satisfaction scale. Indeed, the approach to the theory satisfaction is ambiguous. Seth et al. (2005) claim that in literature prevails the concept of satisfaction as the more stable metric. Perceived quality is an antecedent and a component of customer satisfaction. In addition to customer satisfaction, other one-dimensional rating scales can be used. Overall feeling or “user rating” may be one of the examples. Hu, Paul and Zhang (2004) found that these assessments have bimodal anomalous U-shaped distribution, thus the average score does not necessarily reflect the actual quality of the product and can provide false information. Furthermore, these data are difficult to interpret when identifying shortcomings of operated IT services (Lepmets, 2012).

Multi-dimensional rating scales offer better insight because they involve more than one quality attributes. Evaluators can select predefined attributes described by SQuaRE standard (ISO/IEC 25000, 2014), if they consider them as important. The commercial equivalent of the SERVQUAL “Rater” uses 22-items scale whose dimensions are reliability, assurance, tangibility, empathy, and responsiveness. Firms ask customers to assess the dimensions against their prior expectations using a 5-point scale (Coulthard, 2004). According to Cronin and Taylor (1992, 1994) its use is problematic due to formalisation of customer expectations. In relation to that, SERVPERF is more unambiguous. To evaluate service quality more reliably, service climate questionnaire (Jia & Reich, 2011) should be used in addition.

Another Investigation of Quality in Use typically requires expensive experimentation like focus groups, telephone surveys, user-generated reports or forms (questionnaires) inquiring about satisfaction in case of customer research and expert evaluation and mystery shopping in case of external evaluation,
when the gained content is later manually extracted and investigated (e.g. Olsson & Salo, 2012; Bruun & Stage, 2012). They are time, financially and personally challenging tasks.

The last category of the Quality in Use measurement contains semantic analysis. As we can see from customer research methods, a large amount of user created unstructured data exist in the relation of IT services. And more of them we can extract from the Internet. We suppose, that analysis of the user created unstructured data can help to more complex and effective service quality evaluation.

Tab. 1: Methods of QinU measurement (authors)

<table>
<thead>
<tr>
<th>Structured data</th>
<th>Unstructured data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SCALES</strong></td>
<td><strong>EXTERNAL</strong></td>
</tr>
<tr>
<td><strong>EVALUATION</strong></td>
<td><strong>CUSTOMER</strong></td>
</tr>
<tr>
<td><strong>RESEARCH</strong></td>
<td><strong>SEMANTIC</strong></td>
</tr>
<tr>
<td><strong>(CONTENT)</strong></td>
<td><strong>ANALYSIS</strong></td>
</tr>
<tr>
<td>Single-attribute scales</td>
<td>Expert Evaluation</td>
</tr>
<tr>
<td>Multi-attribute scales</td>
<td>Mystery shopping</td>
</tr>
<tr>
<td>Complex scales</td>
<td>Telephone survey</td>
</tr>
<tr>
<td></td>
<td>Forms (questionnaires)</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As we mentioned rating scales can be difficult to interpret. User rating of IT service can be also hardly separable from the core of the service, which is the main business activity of the service (Kaitovaara, 2004). By this claim we are turning back to the categorization of IT services according to recipient as we already mentioned above. Some services can be delivered to the end customer through the third site, mainly 2B customer in this delivery chain (internet banking). Additional commentary to the scale rate can clarify the reason of satisfaction or dissatisfaction with the IT service.

3.2 Customer voice sources

Customer voice regarding service expectations, experience and quality are coming from different sources which is possible to measure with the methods from Tab. 1:

1) In structured form as
   a. evaluation using scales
   b. results of structured questionnaires

2) In unstructured form as
   a. direct feedback to company (mails, calls)
   b. results from a driven research (narrative stories)
   c. spontaneous WoM (reviews) from the Internet sources such as forums, blogs, social networks etc.

In essence, WoM content generally includes an introduction, explanation, description of experience with a product or service compared with other products or services, evaluations and recommendations. This content can be either positive or negative and, therefore, either favourable or detrimental to business operations. A review is considered to be of a good quality if it can help the service provider identify mistakes and learn possible ways of fixing them. Those reviews follow certain conventions which is different from the written narratives in specific surveys like giving a recommendation to purchase. Customer reviews on Internet sites are also written with the motivation of customer own use of the service, usually in conjunction with some small reward, tangible if the review site offers credit for reviews, intangible in the form of community recognition because of the perceived help afforded by a review (Hadegaard & Simonen, 2013).

The advantage of Internet resources is the fact that users describe their experience with a product or service as they really feels. eWoM (electronic Word of Mouth) which comes from the internet from the communication between end users of service affords a potential cheap and rich source of information concerning user experience and usability.
3.3 Customer needs

Internet discussions and reviews can be also a source of another important component of the service quality model - customer needs, also known as Voice of the Customer (VoC). Customer need is a description, in the customer's own words, of the benefit to be fulfilled by the product or service (Griffin & Hauser, 1993). For the quality purposes there were used already in Quality Function Deployment (QFD), developed in the seventies by Mitsubishi and adopted by many manufacturing industries, especially automotive. In QFD customer needs are linked to designed attributes thus encouraging the joint of marketing issues and engineering issues.

VoC is term used in two areas, in IT and marketing. In the IT field it is a process for capturing and evaluation of customer’ requirements and expectations concerning a service or product related to the software. The marketing concept is more general, but it is always a survey method, where the output is a hierarchical list of needs of customers and their satisfaction with the current status quo (Nečas & Marc, 2011; Gaskin et al, 2010). In other words the status quo is a customers’ perception of the service. Each need or set of needs has assigned to it a priority which indicates its importance to the customer. In QFD customer needs are then linked to design attributes and these attributes to actions the firm can take to implementation and product planning (Griffin & Hauser, 1993).

SERVQUAL focuses largely upon customers’ assessment of the service process and human interactions (Cronin & Taylor, 1992). Individual components of the interaction with the service can be assessed as a good quality, but it does not mean that overall experience is evaluated as a good quality, nor does measuring the components of service quality ensure that customers will be satisfied. Customer experience with the service is strictly personal and implies rational, sensorial, emotional, physical and spiritual levels of involvement. It also emerged already before the purchase thanks to advertising, promotion and WoM (Klaus, 2015). Experience also occur across different channels. Klaus (2015) divided service experience into 4 themes:

1. Process experience including items such as process ease and the challenge of using multiple channels in dealing with the provider
2. Direct evaluations of encounters with personnel
3. Influence of the physical environment
4. Situational and consumer moderators such as task orientation or location

We can conclude that service experience is based more on customers’ cognitive and emotional assessment of value rather than evaluation against the benchmarks or expectations and captures the whole organization offering not just the attributes of service delivery (the customer experience with the internet banking will have influence also the image of the bank as whole).

4 Semantic analysis of unstructured data

Tremendous strides were made in recent years to automate the analysis of unstructured text data. Current approaches adopt a single-variable approach, focusing on individual metrics as word length, the presence of keywords, or the overall semantic orientation of terms within the data (Robinson, 2012). That means decomposing corpora with varying latent structural dimensionality into smaller collections based on dimensions, factors, topics, sentiment, or concepts. On the issue of sentiment analysis has been in research identified several approaches that some of them can be found in research article Opinion and Sentiment Analysis by (Pang and Lee, 2008). To this approaches we can contain sentiment classification methods, feature-based opinion mining methods or comparison-based opinion mining methods. Sentiment analysis is usually carried out at three levels: document level, sentence level, feature level (character or entity). Most classifications is now implemented by the identification of words and phrases (Mishra & Jha, 2012; No & Kaur, 2013).

The problem of semantic analyses is that their results should be quantifiable. Complexities in the analysis of unstructured textual data often results in only minimal use of the data (Ashton et al., 2014). So it is necessary to find a way how to generate outputs consumable to service providers. Ashton et al. (2014) suggested the application of control charts to the solution generated by a text mining algorithm and latent semantic analysis to extract concept factors related to service quality categories which will quantitatively describe service quality, characteristics the customer prefers, and desired characteristics. Hadegaard and Simonen (2013) uses a machine-learning-based classifier that tags sentences in reviews according to whether they contain usability or user experience-related information and according to the dimensions of usability or user experience they pertain to.
Limitations in semantic analysis of WoM

Internet reviews and discussion are generally open and frank. Main advantage is that they are the direct feedback, can be detailed and easily tracked their progress over time. Prerequisite is that users of the product must be sufficiently active and comment the product in an online environment. Experience (e.g. Han & Niu, 2012) is also that negative information spreads faster than positive.

In the open discussion the problem is also the unknown identity of the contributor as it could be in the case of social networks like Facebook where we can connect the identity with the concrete person and her details like age, gender, occupancy, preferences. Some reviews can be also fake. Then it is important to look at the debate only relevant unaffected information.

Online communication text data is written informally and often contain noise in form of typos and spelling mistakes, grammatical errors, improper punctuation and irrational capitalization, oversimplified expressions, newly coined phrases and slang expressions.

We have to also take in count, that result of Hadegaard and Simonen (2013) shows that internet reviews does not contain much detailed information about specific situations of use or of measurements. They conclude their paper with the claim, that 'it seems highly unlikely that mining Internet reviews can supplant traditional usability testing or user experience studies. We can suppose that the reason can be the choice of the eWoM source or a volume of gained data. Small amount of data can lead to contradictory results.

Methodological Framework

Due to the service intangibility, the quality cannot be accessed directly. Based on questions identified by Seath et al. (2005), we have decided to find more complex service quality model which would take into account IT service components, customer perspective and remain applicable to different types of services. We have the opinion that IT services and non-IT services (the consumed business service which is supported by IT) appear always together and it is difficult to distinguish between them. Hence we adopted the service quality model outline (Vencovský, 2014) that decomposes quality characteristics into layers. Concept is inspired by the SPSPR (Voříšek & Basl, 2008). First dimension of the model takes in account the distinction of the IT quality service components. Different frameworks can be applied in each layer of the model. The lowest level represents information system’s individual resources, namely IT artefacts like hardware, software and IT related service personnel. Next level includes resources to a system view where they act as a whole. The highest layer includes service dynamic in terms of business process design. The second dimension of the outlined model uses intrinsic and extrinsic quality aspects. These layers and aspects form a production quality of service.

<table>
<thead>
<tr>
<th>Distinction of intrinsic and extrinsic aspects of IT service quality (type of quality requirements)</th>
<th>Intrinsic</th>
<th>Extrinsic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distinction of stakeholder of IT service quality (source of quality requirements)</td>
<td>-</td>
<td>Principles</td>
</tr>
<tr>
<td>Distinctions of components of IT service quality (object of quality requirements)</td>
<td>IT resources</td>
<td>Information system</td>
</tr>
</tbody>
</table>

Tab. 2: IT service quality model by (Vencovský, 2014)
The model separates production quality and the Quality in Use (QinU) is in accordance with the SQuaRE standard (ISO/IEC 25000, 2014). Similar characterisation can be found in service quality literature (Grönsoos, 1984; Praeg & Spath, 2010). Standard claims that intrinsic quality aspects are prerequisite of extrinsic aspects and the production quality is prerequisite of Quality in Use.

Possible objects of the analysis in our supposed method are IT service artefacts as software, service operation logs, input and output data. Other objects are IT service stakeholders as customers and personnel. Instead of interviewing them their communication and feedback is analysed. The Vencovský model outline will be taken as a framework for the extension of the quality service measurement of computer-aided content analysis as it can distinguish between different objects. He also suggests measuring different aspects of quality in the matrix between components of quality dimensions overlap.

Generally, the measuring method should follow basic managerial steps. For this we offer well-known Deming cycle. It should comprise steps of measurement planning, executing, evaluating and future modification. For this reason we have chosen CRISP-DM methodology (Cross Industry Standard Process for Data Mining). Even the content analysis and text mining is different from data mining on the base of the collections of documents composed of unstructured textual data instead formalized database records, they have many similarities (Feldman & Sanger, 2006). CRISP-DM methodology consists of six basic steps: business understanding, data understanding, data preparation, modelling, evaluation, deployment.

7 General method design

At this section we outline a method design based on the previous research we made and described in this paper. This method will be modified in greater detail as our research will continue in the future and validated at real data.

According to Vencovský (2014) customer satisfaction poll permeates through all components of the service quality model within the relational aspect. We also have to take in account dimension overlap. First of all we have to model the relationships between all the notions mentioned in the section Customer Perspective in Service Quality. This relationship is designed in Figure 1.

![Fig. 1: Relationship between Customer perspective components entering the Service Quality](authors)

The important part of this view is the fact, that analysed Customer Voice can bring a knowledge about new customer needs which should be linked to redesign or design new attributes of the service which are therefore accessed by customer again and the cycle is closed. That leads to Continual Service Improvement, which is one of the conceptual phase described in ITIL v3 (ITIL3 CSI, 2007).

We detected the sources of Customer Voice in the section Customer voice sources. As a repository and analytics tool for content analysis we have chosen open source Elasticsearch software based on Apache Lucene library. Elasticsearch is a distributed scalable system for real-time search and analysis.
tool whose main function is the full-text search. It also supports structured search, geolocation and recording the relationships between data. In Elasticsearch, the data from all sources are collectively analysed. Data are loaded to Elasticsearch through connectors. In case of the Internet resources as forums or social networks web crawlers for automatic browsing website content have to be programmed. This part of the model presents the Quality of Use which is involved by the results from the Product Quality measurement. That can be done by the designation of attributes or defined metrics from the Product Quality results, which will be analysed by the Elasticsearch in the content.

Fig. 2: Framework of measurement of the Service Quality using unstructured data (authors)

We defined WoM earlier as a claim carrying some sentiment about the product or service. Simply stated, it includes the subject (service, product, some attribute of the product) and related appraisal words (usually adjective or adverb like bad, nice, tasty) which is evaluating the discussed subject and carries the sentiment. With this assumption we are working in our model, when we are detecting the single subjects - according to manually created database by the users with the domain knowledge, which can be improved by the time, or as a result of the clustering method running above the content repository. For the appraisal words there is another repository, their meaning and sentiment are evaluated by Natural language Library commonly available by different research institution (e.g. Stanford1) or open source Apache OpenNLP2.

The problem is that eWOM content could include one or more subjects. Although computers are able to find subjects and appraisal words in the eWOM content, they cannot distinguish which appraisal words are used for which subjects, and this can cause problems with regard to further analysis of the content. For the appraisal words there is another repository, their meaning and sentiment are evaluated by Natural language Library commonly available by different research institution (e.g. Stanford1) or open source Apache OpenNLP2. The last thing which should be treated is the negation of the adjectives (phrases like “not good”), where the negation word turns the sentiment of the appraisal word.

Next issue which is problem to solve is that some appraisal words are neutral if they stay alone, but they get sentiment in the context of the sentence. For example adjective “slow” is carrying neutral sentiment, but if we put it to the context when “some service transfer is slow”, then it is getting negative sentiment. This is the reason why is necessary to create appraisal words repository where those words should be define in relation to service purpose. The last thing which should be treated is the negation of the adjectives (phrases like “not good”), where the negation word turns the sentiment of the appraisal word.

1 http://nlp.stanford.edu/software/
2 https://opennlp.apache.org/
The method keeps phases according to CRISP-DM methodology which are the main pillars of the model life-cycle. In general the description of the method is described in the following paragraphs. The Business Understanding phase is more explained on the particular example of the online internet banking service (the grey italics paragraphs), where we are planning to implement our method first in our future research. The quoted comments are real commentaries from the customers of the service.

A. Business Understanding (identification of service quality objectives)
   a. definition of the service from the business perspective:
      i. in which context and which role is presenting (support, facilitation, expansion, core-service)
         The core of the service is the bundle of the IT services focused on money transfer and account management. Online banking IT service plays the facilitation role only. Customer has still the possibility to visit their local branch. The online banking service consumes sources that may be defined as the low level services. Stakeholders are service user (client), service provider (IT department), service customer (bank).
      ii. purpose of the service
         The purpose of the service is to enable online activities like account management, a money transfer and a bill payment. The main value of the service is the spare of time and convenience. “I don't have to rush to my local branch for my transactions”, “I haven't been to a branch in years, thanks to banking machines and online banking.”
      iii. determination of the single elements (components) of the service
         Core functionality
         ➔ account management
            • basic information settings
            • limit management
         ➔ bank card management
         ➔ checking account balance and transaction history
            • searching
            “I can't see my balance after each transaction”
            • results filtering
            • account analytics (aggregation of transaction categories, building charts)
         ➔ money transferring
            “move money around”, “transfer between accounts”, “payee online”
         ➔ bill paying
         ➔ cheque deposit
         Extended functionality
         ➔ budgeting tool
            “The budgeting tool is great”
         ➔ financial and banking product shopping
         ➔ information service
         Support functionality
         ➔ help / advice system
            • online chat
            • helpline
            • textual help / Q&A
         Components
         ➔ Application software and its integration mechanism to the core IS
            • technical difficulties
         ➔ Application platform
         ➔ Hardware
            • Application server
            • Database server
            • Network devices
Personnel
“The employees were awesome very efficient and friendly”, “The staff
is knowledgeable, polite, educated in their product”

b. definition of the dimensions which will be evaluated

Service quality dimensions
➔ Tangibles

➔ Reliability

➔ Responsiveness

➔ Guarantee/Assurance

➔ Empathy

Software quality in use dimensions
➔ Effectiveness

➔ Efficiency

➔ Freedom From Risk (see Guarantee/Assurance)

➔ Satisfaction

➔ Context coverage
c. **creation of the subject database**: database of the notions (words, phrases) related to the service

*Definition of the service synonyms*: app, application, web, system, banking.

For the next analysis must be defined the sentiment words. In generally, predefined sentiment dictionary or Natural Language Library can be used. For example, we identified during the pre-reading phase these positive and negative words.

- **positive words**: like, love, thanks, easy, superior, good work, happy, capable, usable, simple, intuitive, powerful, excellent, “positive experience”, pleased, appreciate
- **negative words**: inconvenient, annoying, glitch, terrible, frustrated, hard to understand

The most complicated is to distinguish which data are really related to the service. Even if we have a data source focused on the specific service, we must filter out texts that are oriented to related services in a way to measure the quality evaluation more precisely.

For example, during the pre-reading phase we found textual reviews that assessed the quality of the service at the branch, although the reviews was located on the online banking message board. The words that determine whether the comment is not focused on the online service may be: branch, door. In any case filtered texts need to be reviewed manually.

“whenever I walk in the door!”; “My branch is the ABC.... They are awesome..... always helpful and friendly.... they make me feel special as a customer.... I may not go in for months and they still know who I am..... They call me when I can get better products... I will never change my bank!”

B. **Data understanding (collection and review data)**
   a. identification of sources of content analysis, assessment of their quality and costs of acquiring data
   b. **WoM collection**: collection, indexation in repository and filtering WoM-related content from the Internet

C. **Data preparation (first content analysis)**
   a. pre-reading of data
   b. **clustering analysis**: to explore data, found the most discussed topics
      i. determination of frequency of words
      ii. **subject collection**: determine when the comments relate / do not relate to services - search for the occurrence of specific notions (name of the brand, service components, problems,.... entered by users
      iii. **subject filtering**: filtering the notions (words, phrases) from the content by comparing with the database of subjects created before
   c. **categorization**: of subjects to predefined dimensions set in first phase, to new dimensions recognized by clustering
   d. **setting the appraisal attributes analysis rules**: (appraisal word is word which is evaluating the subject of the estimation)
   e. **part of speech tagging** - to recognize appraisal word in sentence
   f. **appraisal word monitoring** - calculation of appraisal word weight

D. **Modeling (deep analysis)**:
   a. **subject matching**: matching the sentences according to the collected subjects
   b. **appraisal sentences evaluation**
      i. **appraisal sentence marking**: sentences that include appraisal words are singled out and marked as appraisal sentences (claims, talks)
      ii. **appraisal sentence analysis**: With the help of Natural Language Library adjectives are then examined to see whether they are positive or negative, and whether they have been negated (occurrence of the words like “not”)
      iii. **determination of sentiment** of the claims about the service components/attributes
c. descriptive frequency statistics analysis: to normalize the results for synthesis with the results gained by other components metrics - measurement of the subjects in claims, assigning weights to individual subjects and dimensions and their sentiment

E. Evaluation (evaluate model and conclusion):
   a. determination of overall sentiment of the service
   b. comparing results with the results from rating scales (if they are the part of the feedback)
   c. combination the normalized results with the results from the components quality part and assessment of the overall Service Quality.

F. Deployment (application of the conclusion into business)
   a. assign the findings to the service improvement: redesign the attributes, development of the new attributes

This method assumed to be used long term so it is learning as the service cycle continues and also transferable in the same business. We can say it has some iterative elements and is more accurate as each iteration passed. We proposed to establish the roles accountable for the each phase and the evaluation process as whole.

8 Conclusions and future research

In this paper we designed a method which assist in the overall assessment of the service quality. It includes perspective of the Customer Voice, mainly Word-of-Mouth as an input to the measurement. Theoretical foundation of the method lie in the Service Quality model literature research which is presented in a chapter 2. We have subjected analyzing Customer Perspective and its measurement in the chapter 3. Semantic analysis of unstructured data and its limitation is discussed in chapter 4 and 5.

This paper answers the two of three issues presented by Seath et al. (2005) in relation to IT and service quality - we defined a way how to listen to the voice of customer through information systems and how to collect data related to customer perceptions and his/her possible behaviour.

The frequency of the listening and collection is not as necessary as the possibility to collect full history, because when we have full data in the repository, we can slice them according to time period we want to see and from this view we can read the need of the frequency which will probably change with the type of service.

The method design is presented in a chapter 7. We described how we approach to the relation between phenomenon like Service Experience, Service Usability, Quality in Use, Customer Voice, Word-of-Mouth and others. We designed the implementation environment. We have chosen as a repository the Elasticsearch software, which is open source and can handle amount of the data from different sources. Finally we proposed the methodological process. Its phases are designed in accordance with the CRISP-DM methodology. The business understanding phase is described more in depth using the example of the internet banking service.

In the future research we are planning to implement our method for a concrete IT service. The remaining phases like data understanding, data preparation, modelling and evaluation will be presented closely. In these phases will be the critical to have enough data. More data we gain, more accurate results we get. In the standardized services as online banking or internet providing or broadcasting the method is transferrable, we can use what the system learned in the past - dictionaries, repositories, also associations and results. The aim is to teach the method to be fast, accurate and easy to use for its users which can be from marketers -for the better advertisement of the service, customer service department - to better communicate and solve problems of the customers, to managers who has the rights to enforce specific changes to improve services.

Acknowledgements

This paper was prepared thanks to the IGA grant VSE IGS IG406025 and with contribution of long term institutional support of research activities by Faculty of Informatics and Statistics, University of Economics, Prague.
References

Ashton, T., Evangelopoulos, N. & Prybutok, V. R., 2015: Quantitative quality control from qualitative data: control charts with latent semantic analysis, Quality & Quantity, 49(3), pp. 1081-1099


Divya, M. S. & Goyal, S. K., 2013: An advanced and quick search technique to handle voluminous data


**JEL Classification:** C80, M15