amat_2022_quality_4_0_big_data_analytics_to_explore_service_quality_attributes_and_their_relation_to_user_sentiment_in_airbnb_reviews

Year

2022

Author(s)

Natalia Amat-Lefort and Federico Barravecchia and Luca Mastrogiacomo

Title

Quality 4.0: big data analytics to explore service quality attributes and their relation to user sentiment in Airbnb reviews

Venue

International Journal of Quality & Reliability Management

Topic labeling

Manual

Focus

Secondary

Type of contribution

Established approach

Underlying technique

Label from previous studies or Manual labeling assisted by associated documents

Topic labeling parameters

Label generation

In order to provide a label for each attribute identified through the topic modelling process, the reviews with the highest topic weight for each attribute were used as a reference. Additionally, the most relevant keywords associated with each attribute were also used as a reference for the labelling process. Table 2 shows the top 5 keywords for each attribute, which are the words with the highest probability within each topic.

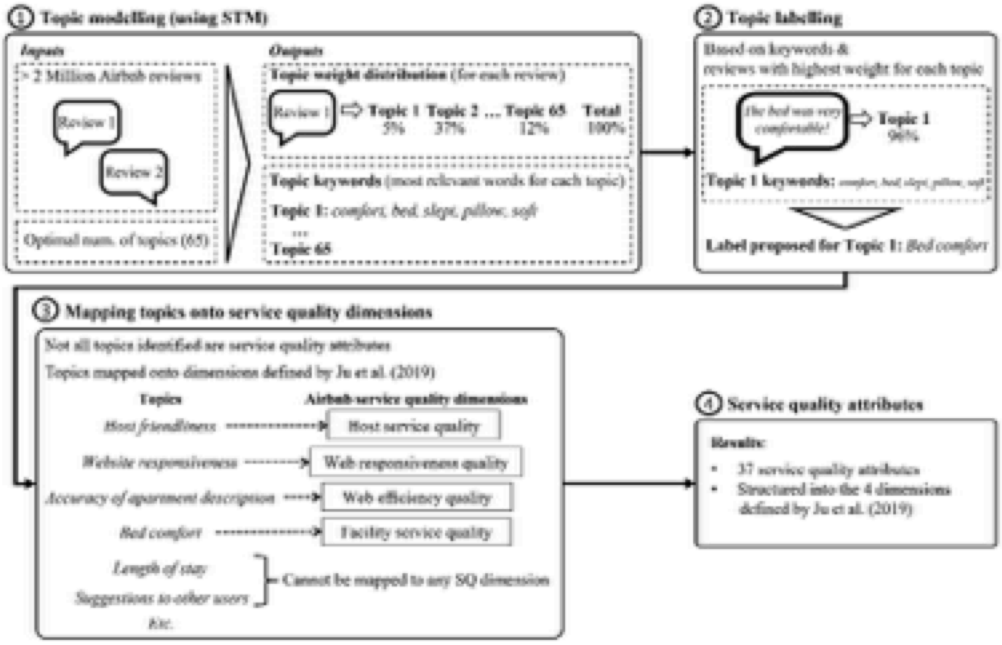
Dimensions (Ju *et al.*, 2019)	Id	Service attributes identified	Topic label source	Relevant keywords (stemmed for text pre-processing purposes)
Host service quality	1	Host tips and advice	Adapted from Ju *et al.* (2019)	Visit, place, see, tip and short
	2	Host welcome	Adapted from Tussyadiah and Zach (2017)	Welcom, inform, warm, met and gave
	3	Host friendliness	Adapted from Ranjbari *et al.* (2020)	Help, friend, kind, pleasant and thought
	4	Host kindness	Adapted from Ju *et al.* (2019)	Host, autom, accommod, offer and wonder
	5	Feeling at home	Adapted from Zhang (2019)	Home, made, make, feel and trip
	6	Communication with host	Adapted from Kiatkawsin *et al.* (2020)	Easi, communic, get, check and checkin
	7	Host responsiveness	Zhang (2019)	Quick, response, always, question and respond
	8	Late check-in	Adapted from Zhang (2019)	Late, wait, reserv, flight and delay
	9	Check-in/check-out flexibility	Ranjbari *et al.* (2020)	Check, even, earli, leav and accommod
Web responsiveness quality	10	Website responsiveness	Parasuraman *et al.* (1988)	Person, guest, ask, detail and care
Web efficiency quality	11	Accuracy of apartment description	Adapted from Ranjbari *et al.* (2020)	Expect, photo, descript, accur and smaller
	12	Accuracy of apartment pictures	Labelled by authors	Just, look, pictur, exact and describ
Facility service quality		*Apartment features*		
	13	Apartment safety	Ranjbari *et al.* (2020)	Build, door, safe, felt and outsid
	14	Room aesthetics	Cheng and Jin (2019)	Room, bathroom, bedroom, kitchen and live
	15	Cleanliness	Ranjbari *et al.* (2020)	Clean, good, flat, hous and near
	16	Bed comfort	Adapted from Zhang (2019)	Comfort, bed, slept, pillow and soft
	17	Apartment dimensions	Labelled by authors	Small, size, unit, shower and cloth
	18	Sleeping capacity	Kiatkawsin *et al.* (2020)	Space, access, plenti, spot and coupl
	19	Views	Ranjbari *et al.* (2020)	Beauti, amaz, stay, view and absolut
	20	Relax/unwind areas in apartment	Adapted from Kiatkawsin *et al.* (2020)	Day, long, relax, sightse and end
	21	Ease of access (stairs)	Labelled by authors	Floor, stair, top, elev and luggag
	22	Food and drinks	Adapted from Zhang (2019)	Breakfast, food, coffe, even and wine
	23	Thermal management (water/heating issues)	Kiatkawsin *et al.* (2020)	Work, problem, shower, water and hot
	24	Hygiene issues (uncleanliness)	Labelled by authors	Toilet, bad, smell, unfortun and seem
		Equipment		
	25	Room equipment	Adapted from Zhang (2019)	Provid, use, kitchen, includ and cook
	26	Internet connection	Adapted from Lee *et al.* (2019)	Wifi, connect, Internet, watch and view
	27	Apartment furniture	Ranjbari *et al.* (2020)	Well, quiet, spacious, equip and decor
		Location characteristics		
	28	Location (overall)	Tussyadiah and Zach (2017)	Locat, apart, stay, recommend and love
	29	Neighbourhood features	Lawani *et al.* (2019)	Neighbourhood, huge, plus, far and favorit
	30	Accessibility	Kiatkawsin *et al.* (2020)	Also, apart, around, near and easy
	31	Proximity to attractions	Ranjbari *et al.* (2020)	Street, away, step, market and block
	32	Distance from city centre	Labelled by authors	Citi, can, center, centr and park
	33	Proximity to public transports	Adapted from Ranjbari *et al.* (2020)	Close, metro, station, conveni and nearbi
	34	Proximity to restaurants	Adapted from Zhang (2019)	Restaur, area, lot, shop and bar
	35	Reachability from the airport	Labelled by authors	Airport, bus, train, take and stop
	36	Sleep disturbance (night noise)	Kiatkawsin *et al.* (2020)	Night, bit, window, sleep and nois
		Other attributes		
	37	Family friendliness	Kiatkawsin *et al.* (2020)	Famil, kid, adult, children and apart

Table 2. Mapping and labelling of identified topics onto service quality attributes

In the case of attributes identified in previous Airbnb studies, the label source indicates the study from which the name of the attribute was extracted. Six of the attributes were labelled manually by the authors through group discussion and evaluation, given that they could not be found in previous literature. Lastly, the reviews with the highest weight for each topic were analysed to confirm the adequacy of their respective labels.

Figure 3 illustrates (with examples) the process of extracting topics, labelling them and mapping them

onto service quality dimensions in order to identify service quality attributes.



Big data analytics to explore service quality

Figure 3. Process to identify service quality attributes from topic modelling results

Motivation

Topic modelling can capture the fact that the same review could mention several different topics, and it can also reflect the relevance of each topic in a particular review. However, topic labels are not provided by the STM.

Labels are used to identify specific service quality attributes.

Topic modeling

STM

Topic modeling parameters

Nr of topics: 7 to 100

Nr. of topics

65 (out of the 65 topics identified in the corpus, 37 were mapped onto service quality attributes.)

Label

service quality attribute identified in previous Airbnb studies or manually assigned single or multi word label

Label selection

\

Label quality evaluation

\

Assessors

\

Domain

Paper: Quality management
Dataset: Tourism and Hospitality

Problem statement

PQuality 4.0 is a new paradigm of quality management, which emphasises the need to adapt to recent technological innovations by updating traditional quality approaches. Amongst the most important factors for adopting Quality 4.0 is the leveraging of big data to collect insights and quality perceptions from clients. Therefore, user reviews have emerged as a valuable source of information, which can be analysed through machine learning procedures to uncover latent quality dimensions. This study applies a combination of text mining techniques to analyse Airbnb reviews, identifying service quality attributes and assessing their relation to the users' sentiment.
More than two million reviews written by guests in four European cities are analysed. First, topic modelling is applied to find the quality attributes mentioned by reviewers. Then, sentiment analysis is

used to assess the positiveness/negativeness of the users' feedback.

Corpus

Origin: Airbnb

Nr. of documents: 2,735,437

Details:

- reviews collected from 100,454 listings in Paris, Rome, Barcelona and Lisbon as of July 2020 (review dates ranged from 2010 to 2020).

Document

Text of an individual review

Pre-processing

- Removal of non-english reviews
- Text normalisation (conversion of all text into lowercase letters to reduce ambiguity)
- tokenization
- stemming
- Removal of the punctuation, numbers and common English stop words
- Removal of the words shorter than 2 or longer than 15 characters, as well as the reviews with less than 10 words
- Removal of words with an extremely low frequency (appearing in less than 20 documents)
- Removal of the words generally not significant to identify topical content (such as: "another", "review", "made", "did", "done", etc.)
- N-gram analysis to find frequently co-occurring words and transform them into unigrams (e.g. "customer service" to "customerservice").

```
@article{amat_2022_quality_4_0_big_data_analytics_to_explore_service_quality_attributes_and_their_relation_to_user_sentiment_in_airbnb_reviews,
    author = {Natalia Amat-Lefort and Federico Barravecchia and Luca Mastrogiacomo},
    date-added = {2023-04-12 23:12:01 +0200},
    date-modified = {2023-04-12 23:12:01 +0200},
    doi = {10.1108/ijqrm-01-2022-0024},
```

```
    journal = {International Journal of Quality \& Reliability Management},
    month = {sep},
    number = {4},
    pages = {990--1008},
    publisher = {Emerald},
    title = {Quality 4.0: big data analytics to explore service quality
attributes and their relation to user sentiment in Airbnb reviews},
    url = {https://doi.org/10.1108%2Fijqrm-01-2022-0024},
    volume = {40},
    year = 2022}
```

#Thesis/Papers/FS