Volume 15 • Issue 1 • 2023

sciendo

ISMSME International Society for Manufacturing, Service and Management Engineering

received: 1 October 2022
accepted: 20 February 2023

pages: 86-105

Exploring the usability and user experience of social media apps through a text mining approach

Anna Baj-Rogowska
Marcin Sikorski

ABSTRACT
This study aims to evaluate the applicability of a text mining approach for extracting UUX-related issues from a dataset of user comments and not to evaluate the Instagram (IG) app. This study analyses textual data mined from reviews in English written by IG mobile application users. The article's authors used text mining (based on the LDA algorithm) to identify the main UUX-related topics. Next, they mapped the identified topics with known theoretical constructs to place them in their nomological network relevant to the usability (the 5Es framework by Quesenbery) and UX (the Honeycomb model by Morville). Finally, to expand the study with an emotional diagnosis, sentiment analysis was performed on two levels: (i) for each recognised topic, and (ii) for the full dataset to uncover general insights into users' emotions within all reviews. The case study of the IG app confirms the usefulness of user feedback data for software development and points out that the review data have the potential for the early detection of frustration and negative feelings introduced during the use of the application. Conducting conventional UUX evaluations with users is problematic since they are remotely located, and the user-generated content of a social app undergoes continuous and frequent changes. Thus, the consecutive stages of the proposed methodology, based on text mining algorithms, constitute a proposed framework for examining the user-perceived quality projection of applications from user feedback, and they are the main contribution of this article. The used approach can be valuable for helping developers, designers and researchers to reveal user problems and fulfil user satisfaction regarding UUX aspects for specific software features.

KEY WORDS
usability, user experience, 5Es model, honeycomb model, text mining, mobile apps

10.2478/emj-2023-0007

Anna Baj-Rogowska

Gdańsk University
of Technology, Poland
ORCID 0000-0002-8694-4461

Corresponding author:
e-mail: anna.baj-rogowska@pg.edu.pl

Marcin Sikorski

Gdańsk University
of Technology, Poland
ORCID 0000-0001-7930-7790

Introduction

Social media, such as Instagram, Facebook, Twitter and many others, have become increasingly popular for communicating with other users and sharing user-generated content, mostly text, photos and videos. Social media users are exposed to targeted advertisements; thus, in fact, they are online customers who make choices driven by individually anticipated value.

For all software-based online services, usability and user experience (henceforth — UUX) have long been considered important factors that describe user-

Baj-Rogowska, A., & Sikorski, M. (2023). Exploring the usability and user experience of social media apps through a text mining approach. *Engineering Management in Production and Services, 15*(1), 86-105. doi: 10.2478/emj-2023-0007

perceived quality, eventually determining the users' loyalty to a specific service. From the IT project management viewpoint, it was assumed that quality requirements should always be defined at the start of software product development (Berenbach et al., 2009). Furthermore, it was taken for granted that designers and developers have almost complete knowledge of the factors shaping software product quality. Most importantly for quality assurance, from the very early start of a project, frequent UUX evaluations are recommended, preferably with a sample of prospective users (Hartson & Pyla, 2012).

However, many online services (including social media) continuously evolve after deployment, while users and customers use them daily. To improve UUX and make a service or app more attractive, online service owners frequently add new functionalities, modify existing ones, and make various changes in data visualisations or other aspects of the user interface.

Contrary to classical software projects, users of online services are remotely located, and they are often eager to publish online their reactions to recently implemented changes. Users frequently post critical comments, especially if they find that specific modifications are too radical, make common tasks more complex, or just worsen their UX for any other reason.

As a result, in developing online services, despite conducting small-scale user-based pilot studies during design and development, it is never certain whether a specific community of users will widely accept specific user interface modifications.

There may also be many intangible factors that may attract or discourage users of a specific online service, such as the attractiveness of content published by other users, innovative functionalities to enhance social communication or service vendors' efforts to protect their users' privacy and security.

Gaining the trust of potential users of mobile applications is a continuous task that consists of many elements. These include, among others: attractive functionality and graphic design, the usability of the application, resilience to errors and the ability to produce instant value for the user, such as solving a specific problem. To achieve success, user comments should be constantly monitored, and the application should be adjusted to the users' expectations. A good application does not make users tired while using it or frustrated while browsing the functionalities of interest. It is this area, in a nutshell, UUX focuses on the basic principle of customer value orientation (Park et al., 2013; Adikari et al., 2011).

The validation of whether the application is useful and effective is carried out mainly through user-based research aimed at checking the level of user satisfaction. A group of representative users is usually employed as usability testers and interviewees. Obtaining this kind of information and analysing user behaviour are key elements for the entire UUX assurance process. This work aimed to propose a novel method of such an assessment based on the use of the text mining (TM) method.

Nakamura et al. (2022) revealed that so far, no method has been found yet specifically designed to analyse app store user reviews. Across their systematic literature review from 2012 to 2019, only five papers were found to apply a topic modelling method. However, a drawback of these studies is that the results of topic modelling are lists of terms which require a high cognitive load to interpret them. This study aims to fill this gap with a novel approach by introducing modifications, such as (i) labelling topics, (ii) mapping them to theoretical constructs from the field of UUX and (iii) including sentiment analysis to gain insight into the emotional dimensions of UUX. Therefore, focusing on filling this research gap, this paper aims to explore whether user opinions published online can be used:

- to disclose the main factors composing UUX for a pilot sample of Instagram users;
- as a projection of user-perceived quality, described especially as UUX;
- to reveal how users rate the IG mobile application and what emotions it evokes in them.

The classical text mining methods and algorithms were considered for this analysis. First, topic modelling based on Latent Dirichlet Allocation (LDA) was performed on user reviews to recognise the main topics in the dataset. Then, the identified topics were mapped with known theoretical constructs to place them in their nomological network and to identify factors composing usability and user experience. Finally, to reveal the emotional tone hidden in the reviews, a method of sentiment analysis was used.

The approach adopted of this study's authors is in line with the CRoss Industry Standard Process for Data Mining (CRISP-DM) framework, which comprises six phases: business understanding (consisting of research questions), data preparation and understanding, modelling, evaluation, and deployment (interpreted by the study authors). This framework is widely considered the most suitable and comprehensive set of guiding rules for performing analytics projects (Abbasi et al., 2016).

The paper is organised as follows. A brief overview of UUX-related research is presented in Section 2. Section 3 is dedicated to the process of data collection and the methodology used in this study. In Section 4, the results are presented and discussed. Section 5 concludes the paper by presenting relevant limitations and the possibilities for further studies.

1. Related research

Usability and user experience constitute important quality attributes for software, websites and mobile applications. This section starts with a discussion on the concepts and components of usability, user experience and social user experience, respectively. Then, it discusses related studies concerning the application of the text mining approach for exploring UUX issues.

1.1. Usability

For software developers to facilitate the creation of high-quality software, the standard ISO/IEC 9241-11: 1998 specified usability as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use". Furthermore, the three principal usability components were defined (Bevan, 2008):

- effectiveness — the accuracy and completeness with which users achieve specified goals;
- efficiency — the resources expended to the accuracy and completeness with which users achieve goals;
- satisfaction — the comfort and acceptability of use, including "freedom from discomfort, and positive attitudes towards the use of the (specific) product, system, service or environment".

Soon after, the ISO/IEC 9126-1:2001 was published, specifying the following general quality characteristics for software products: functionality, reliability, efficiency, maintainability and usability. In this standard, usability was described by a different set of attributes, which causes some confusion, namely:

- learnability — the capability of the software product to enable the user to learn its application;
- understandability — the capability of the software product to enable the user to understand whether the software is suitable for a specific purpose and how it can be used for particular tasks and conditions of use;
- operability — the capability of the software product to enable the user to operate and control it;
- attractiveness — the capability of the software product to be attractive to the user;
- compliance — the capability of the software product to adhere to standards and conventions.

These attributes are surely significant for assuring high usability, but (except compliance) they are subject to individual susceptibility and far-from-objective assessment. For this reason, the terms specified in the earlier standard ISO/IEC 9241-11: 1998 were commonly operationalised for evaluation. Effectiveness was specified as user task completion rates and task completion times, respectively. Efficiency includes the relationship between an achieved task outcome and expended resources, such as time, mental effort and a number of steps or attempts to find support in the case of problems.

User satisfaction remains a purely subjective measure, so it is usually rated on a numerical scale, employing usability questionnaires or surveys. The identification of satisfaction components is usually performed using expert knowledge, user interviews or observations of the user activities in an actual context of use.

ISO-defined concepts of usability, unfortunately, require that the interpretation of usability components in each case (whether an office software, e-commerce website or social media mobile app) be made by choosing an adequate set of measurable characteristics suitable to a specific context of use, usage scenarios and user needs.

Deficiencies in ISO definitions for usability have been tackled by many researchers who attempted to adapt the understanding of usability to newly emerging systems and to rely mainly on the user's subjective assessment rather than on strict measurements.

For instance, Quesenbery (2004) proposed the "5Es" usability model, stating that a high-usability interactive product should be:

- effective — the software product helps to achieve the users' goals accurately;
- efficient — the software product provides the speed and accuracy with which tasks are successfully completed;
- engaging — the software product influences the emotions of the users, involving them in the pleasant, rewarding and interesting operation of a product;

- error-tolerant — the software product has the ability to help the user who commits errors or encounters other difficulties during product operation;
- easy-to-learn — the software product provides the user with both the initial orientation and, as the user's skill gradually advances, deeper understanding and guidance.

The 5Es usability model is more comprehensive than early ISO-based usability concepts, and it better highlights the need for the correct understanding of the local context of use by designers of a specific interactive product.

Furthermore, other researchers, such as Bevan (2008), Shneiderman and Plaisant (2016) or Sharp et al. (2023), among many others, have long promoted extending the notion of usability to encompass additional satisfaction factors typical for newly emerging systems (like games or online services), including flexibility, learnability, likability, pleasure, comfort and trust.

Subsequently, further revisions of ISO-based definitions of usability resulted in specifying a separate term, "quality in use" (ISO 925010:2011), which covers an expanded set of product quality aspects that are user-perceived during product operation.

Although the relevant ISO standards contain very detailed user interface guidelines and are an excellent source of reference, they are very time-consuming to employ in usability evaluations. Both expert-based and user-based evaluations have to be manually planned, performed and supervised by a researcher, which leaves room for subjectivity and unintended skipping of an important aspect shaping user satisfaction in a particular context (Hartson & Pyla, 2012).

This aspect proved especially important in usability evaluations of newly emerging systems, such as games, social media, websites, online services or mobile apps. In such systems, users are consumers and no longer system operators. Their satisfaction is shaped mostly by factors that are no longer task- or performance-related but largely emotional, like the pleasure of shopping in e-commerce, enjoyment and curiosity on social media, excitement and competition in games, or the convenience of instant booking in a mobile travel app.

Therefore, the term "user experience" was predominantly used. With its rapidly growing popularity, usability and all its task-related aspects were gradually absorbed into a novel notion of UX, discussed in the next section.

1.2. User experience

Building upon predetermined usability design goals, user experience (UX) has recently become the popular term covering additional quality criteria against which a digital product or service should be evaluated from the user's or consumer's viewpoint beyond task-focused usability.

Gradually, the term "user experience", initially introduced by Norman (1998), has emerged to cover the components of user interaction that go beyond conventional, task-related interpretations of satisfaction. Subsequently, the new ISO International Standard 9241-210 (2019) introduced the term meaning "a person's perceptions and responses that result from the use or anticipated use of a product, system or service". The users' perceptions and responses may include their emotions, beliefs, preferences, perceptions, comfort, behaviours and accomplishments, which occur in three phases: before, during and after using a specific digital product or service (Adikari et al., 2011).

In particular, with new technologies, such as the Web and mobile apps, users are not necessarily seeking to achieve a task but also to amuse and entertain themselves (Sánchez et al., 2012; Adikari et al., 2011; Blythe et al., 2003), for pleasure and hedonic qualities (Hassenzahl, 2004), which can also be considered a legitimate objective.

In and beyond the context of mobile apps, user experience is a consequence of the brand image, presentation, functionality, system performance, interactive behaviour and assistive capabilities of a system, product or service (Park et al., 2013). It also results from a user's internal and physical state resulting from prior experiences, attitude, skills, abilities and personality relevant to a specific use context (Hassenzahl & Tractinsky, 2006).

Beyond ISO documents, different writers (Sutcliffe, 2009; McNamara & Kirakowski, 2006; McCarthy & Wright, 2004) have explored different aspects of and perspectives on the complexity of user experience in various contexts. Sadly, no common framework for measuring and assessing UX components has been established in IT-related development and research. Alternative models also leave a selection of meaningful UX components to a researcher's expertise and subjectivity.

Two of the most distinctive user experience concepts were proposed by Hassenzahl (2004) and Morville (2004). The hedonic/pragmatic model of user experience (Hassenzahl, 2004) assumes that people perceive interactive products along two different dimensions. The pragmatic dimension refers to the product's perceived ability to support the achievement of specific objectives, like the outcome of a specific task (i.e., booking a hotel). On the contrary, the hedonic dimension refers to the product's perceived ability to perform a process in a smooth, confident and pleasant manner. For instance, it validates whether the process of hotel booking was visually attractive, transparent, and free from any frustrations, just reinforcing the user with positive emotions and memories.

According to Forlizzi and Battarbee (2004), user experience is always tied to actual use situations: to be valid, it cannot be separated from an actual experience of use during or after system operation by a specific user(s). Hassenzahl and Tractinsky (2006) strongly emphasised that user experience is always influenced by the user's internal state (predispositions, expectations, needs, motivation, mood etc.), the characteristics of the designed system (complexity, purpose, usability, functionality etc.) and the context (or environment) within which the interaction occurs (e.g., organisational/social setting, the meaningfulness of the activity, voluntariness of use, etc.). This view of user experience marks it as a definitely contextual construct, hard to formalise and unify; hence, it is rather more suitable for qualitative studies than for quantitative assessments.

Morville (2004) proposed the Honeycomb user experience model (Fig. 1), addressed mainly to the e-commerce domain, where the popularity of commercial websites and online services is very much shaped by the emotional load transferred by e-marketing tools and techniques. The Honeycomb model nowadays also applies to mobile apps, which are just another access channel for digital services, which used to be delivered primarily by websites when the model was created.

Fig. 1. Honeycomb user experience model

Source: Morville, 2004.

According to the Honeycomb model, an outstanding user experience is created by the consolidated outcome of the following interrelated components (Meuthia et al., 2021; Vivakaran & Neelamalar, 2015):

- useful — the app or service should deliver an outcome that is useful for the user;
- usable — ease of use is essential for usability (the accomplishment of the user's goal) but not sufficient to attract a user (consumer) for repetitive use;
- desirable — the combination of image, identity, brand and other elements of emotional design should make the service highly desirable for a specific segment of users;
- findable — the app or service (website) should be easy to navigate, and it should be easy to locate objects needed by users, i.e., by a local search engine or recommending tips;
- accessible — full access to the service functionality for impaired users is essential to avoid user frustration and disappointment beyond the legal requirements that are mandatory for public websites;
- credible — the app or service (website) should be trusted, instantly communicating credibility, competence and care, essential for hassle-free customer service;
- valuable — the app or service (website) must deliver value not only to project sponsors but primarily to users (customers), who would keep returning and remain loyal, ready to develop even stronger relationships in the future.

The Honeycomb model does not specify specific guidelines for delivering its user experience components, leaving particular means and solutions to be designed by qualified user experience designers and digital marketing specialists. Because user experience components in each case have to be individually identified, user experience evaluations must be based on extensive user research studies, interviews, surveys and user experience mapping workshops. The Honeycomb model addresses the product or service quality from the user's perspective, including primarily emotional responses and factors shaping further online customer behaviour.

Many studies from the e-commerce domain (Bilgihan, 2016; Eid, 2011; Chang & Chen, 2009; Chen et al., 2015) have shown that systematic, cumulative positive user experience episodes are essential for building online trust, customer loyalty and valuable relationships between online service providers and specific segments of customers. Now, the popular concept of online Customer Relationship Management (e-CRM) is based on perfecting instances of user experience during single transactions and converting them into stimuli for strengthening customer relationships with a specific online service provider (Sikorski, 2008).

Attempts to discuss the complicated relationship between UX and the concepts of usability in the scientific literature were undertaken by Følstad and Rolfsen (2006) and Sauer et al. (2020). They conclude that usability is more objectively measurable than UX. According to Sauer et al. (2020), usability can be assessed by implementing more diverse methods, while there are fewer methods and instruments dedicated to UX.

To sum up, user experience appears as a complex concept which is difficult to measure objectively. Moreover, it extends the usability concept to the broad spectrum of users' perceptions during the usage of a product and/or service. Most importantly, the concept of user experience has taken the user's subjective emotional perspective, going far beyond the former perspective of task-related usability.

1.3. Social user experience

With the development of social media in the recent decade, the majority of e-commerce marketing communication with retail consumers was transferred to fan pages embedded in service providers' websites. As a result, the user experience of social media users nowadays is composed of "official" product-related content (if available) and user-generated content, like a relevant stream of photos, videos and texts, such as posts and comments.

With the further development of social media, the user experience concept needed to be extended to social user experience (social UX). Again, as in the case of "regular" user experience, no established, commonly accepted models could be used for social UX design and evaluation in social media or similar online services. Much of the relevant development work remains experimental (Saavedra et al., 2019), testing innovations by the trial-and-error scheme, or stuck in already accepted user interface design patterns, apparently reducing the risk of existing users rejecting novel solutions.

Väänänen-Vainio-Mattila et al. (2010) revealed that the main drivers of social UX include self-expression, reciprocity, learning and curiosity, whereas unsuitability of content and functionality, incompleteness of user networks and the lack of trust and privacy are often experienced as hindrances for social UX. Thus, the main guidelines for designing social UX are based on the general concept of sociability, interpreted as a feeling of togetherness and sheer pleasure in the company of others. As such, in other words, sociability exemplifies one's sociable conduct, not the personality characteristics of sociable conduct.

General guidelines for designing social UX were presented, for instance, by Pereira et al. (2010), who revisited the honeycomb pattern and defined key functional blocks recommended for a market-successful social software: presence, conversation, sharing, reputation, groups, relationships and identity. Building valuable social relationships by escalating user engagement can be pointed out as the culminating element, essential for retaining a loyal base of users/consumers.

Later, with further developments in social media and the growing anxieties of users as to possible abuses, their attitude changed to the increased role of perceived trust and the required activities of a service provider regarding adequate moderation of tone and emotions and user privacy protection and data security (Souza & Maciel, 2015).

Guidelines provided by the Norman Nielsen Group (2021) also provide an adequate set of tools for social networking and enable users to select specific types of communication or collaboration, depending on their needs.

The latter resembles the early concept of social UX, presented by Battarbee (2003), who proposed that the members of a specific community should themselves shape the scope and realisation of social UX. This means that a service provider should conduct systematic studies on how available communication tools are used, and look for improvements, also by watching what the competition is doing. Furthermore, while classical models viewed user experience as the subjective response in the individual's mind, its designers and developers had very limited opportunities to provide a satisfying user experience. By contrast, social UX in social media can be seen as an individual's reaction but also as something constructed in social interaction. Most importantly, Battarbee (2003) concluded that social UX is the experience that users themselves create together in social interaction. Subsequently, it is a service provider's responsibility to contribute to the design of positive social UX by creating a safe environment for users to engage in reinforcing activities and select

Volume 15 • Issue 1 • 2023

suitable personal communication artefacts, digital tools, environments and systems to facilitate this kind of use.

Contrary to software and websites, the formal evaluation of social UX, with the intention of making improvements, faces many problems. Social communication is an ongoing process, and any studies of this kind may resemble sampling water from a fast-flowing river. Furthermore, user experience workshops with social media users should be performed remotely, and even if skilfully moderated, they are unlikely to produce results directly applicable to user experience development teams.

The available research in the three areas presented above highlights important distinctions among the three concepts:

- usability — is task-related, relevant mostly to software products, and is assumed to be created/delivered by software product designers; usability can be evaluated partly by objective measurements and observations and supplemented by subjective user opinions;
- user experience — addresses emotional components, present mostly in e-commerce and other online services, where users are aware as consumers; relevant usability factors are incorporated into a specific UX, and they should be evaluated primarily by collecting and analysing subjective user opinions; positive UX is assumed to be largely created by UX designers during the design process;
- social UX — is co-created by users during social interactions; contrary to usability and UX, it cannot be created in advance in the design process by social media developers; they can only provide the appropriate tools and features for users to generate stunning user experiences, and evaluate user feedback from actual activities, posts and comments published online.

To conclude, as discussed in prior paragraphs of this section, existing UUX models have very limited applicability to social media. They are inconsistent, context-dependent, inadequate to user experience co-created by users on the go, and incapable of streamlining a typical, predefined user research process, as happens in UX design for most IT projects.

As social UX is co-created on the fly by a specific community of users, the following reflections come to mind:

- There is no "generic" social UX for social media; it is a local development limited to a specific community of users who collaborate for some time and tend to trust each other in online communication, sharing content and collaboratively co-creating a shared user experience, specific for this group of users.
- Social UX cannot be entirely designed in advance because its bricks are dynamically generated during social interactions taking place among members of a specific community.
- Because social UX is local, it is unlikely that an external researcher would be able to identify its components unless joining the group as a "mysterious member", which could clearly raise ethical concerns.
- Hence, why not let the users of a specific community reveal which factors are important for building a positive social UX in a specific online context? It could be done by encouraging them to use a dedicated forum to post opinions on problems experienced regarding UUX, or other complaints, but also by collecting UX-related innovative ideas or improvement suggestions (possibly rewarded within one of the gamification schemes).

In this study, resulting from the above literature review, the authors selected to use (i) the 5Es framework by Quesenbery (2004, 2014) to identify attributes related to usability and (ii) the Honeycomb model by Morville (2004, 2016) to identify the user experience dimensions. Both models have established precise definitions of the dimensions and clear demarcations of the differences between them.

The authors of this article decided to choose the 5Es framework because: (i) the model is simple, and its flexibility offers the opportunity to customise each app based on the needs of the users; (ii) its dimensions well encompass usability aspects; (iii) it is relatively rarely tested in scientific studies, and the authors aimed to bridge this gap; (iv) it balances function requirements with usability requirements; (v) its dimensions, taken together, are a tool to create a more precise description of both the goals for and engagement of users, and their experience of using applications. According to Quesenbery (2014), the usefulness of this framework "does not end with understanding users", but it suggests design approaches and may then be used to evaluate why an app's interface is failing or succeeding.

On the other hand, the choice of the Honeycomb model was motivated by: (i) its applicability and popularity in domains of e-commerce and other digital services; (ii) treating users as consumers who, aware of multiple options, choose the preferred service provider; (iii) specifying complementary facets of UX, essential to convert user satisfaction into customer loyalty and valuable relationships online. Because social media apps are just interactive tools for building relationships based on users' emotions, the Honeycomb model was found suitable as a basis for user experience evaluation in a social media context, with opportunities for its prospective extensions.

1.4. Application of text mining methods for exploring UUX

The mining of user reviews has received enormous attention in recent years in many areas. Paradoxically, there are quite a few research studies focused on using TM methods for mining UUX issues/strengths based on textual data, mainly user reviews. Below, the discussion concerns those related to artefacts similar to the topic addressed in this article.

Bakiu and Guzman (2017) used a collocation algorithm to extract the features, lexical sentiment analysis to uncover user satisfaction with a particular feature, and machine learning to detect the specific UUX issues affecting the software application. Maalej and Pagano (2011) reported how user feedback could be considered in software development. In a subsequent paper, Pagano and Maalej (2013) analysed the text content and rating characteristics of user feedback from mobile application distribution platforms (i.e., app stores), which allow for the development of requirements from distributed users. Carreño and Winbladh (2013) processed user opinions to extract the main mentioned topics and some sentences representative of those topics. They revealed that this information could be useful for requirements engineers to look through the requirements for subsequent releases.

In turn, Jacob and Harrison (2013) considered mobile app reviews as a valuable source of ideas coming directly from app users. Thus, they analysed mobile app feature requests from online reviews. They designed the MARA system (Mobile App Review Analyzer), which is a prototype developed to mine for and retrieve feature requests from online reviews of mobile apps. The results of the evaluation were further analysed using Latent Dirichlet Allocation for identifying common topics across feature requests. Their study uncovered that most of the user requests refer to improved support for apps, more frequent updates, new levels for game apps, and more customisation options. In turn, Guzman et al. (2015) presented DIVERSE, a feature and sentiment-centric retrieval approach which automatically provides developers with a diverse sample of user reviews. They compared the reviews retrieved by DIVERSE with a feature-based retrieval approach and found that DIVERSE outperforms the baseline approach. Whereas Hedegaard and Simonsen (2013) examined the content of online reviews with the aim of discovering the distribution of information in reviews among different dimensions of UUX, and extracting associated keywords for each dimension using techniques from natural language processing and machine learning.

The scientific literature on the examined topic also provides works covering other systems and not related to mobile applications. For instance, Jiménes et al. (2018) reported on an approach based on text mining techniques to quickly identify usability and functionality drawbacks in a learning management system. By using these techniques, they identified more than ten usability issues and the need for seven new functionalities to be implemented in the system.

Another strand of research includes studies using sentiment analysis. Thus, Portugal and Leite (2018) studied the use of sentiment analysis to help find relationships among usability-related quality requirements in GitHub's projects. They aimed to find a list of sentiment expressions that would characterise significant relationships between usability and other qualities through text mining. Their approach yielded early positive signs.

On the other hand, Weichbroth and Baj-Rogowska (2018) used the sentiment analysis method intending to extract positive and negative keywords from user opinions of the WhatsApp mobile app. Finally, the reported problems were thematically mapped into seven attributes of usability and eight dimensions of user experience. Thus, the authors proved that online reviews reveal genuine usability and user experience issues.

This article's authors propose a different approach based on topic modelling, initially proposed by Debortoli et al. (2016) and expanded with a sentiment analysis to gain insight into the emotional dimensions of UUX.

The availability of user review datasets for mobile apps opens an opportunity to test whether text mining could help identify UUX-related factors (extracted as topics) that seem important to specific users. Thus, this paper aims to find answers to the following research questions (RQs) related to the dataset of reviews about the IG mobile application:

RQ1. What are the user-perceived components of UUX for the IG mobile app based on the selected models (presented in Section 3)? In other words, how do IG users — in the light of reviews — fit into the UUX components and what is the percentage share of each facet in the applied models?

RQ2. Are all components of both models covered/present in the textual dataset?

RQ3. What is the engagement of IG users in the topics discussed in various facets of UUX?

RQ4. Does the user review dataset have the potential for the early detection of frustration, negative feelings (e.g., brand image deterioration, weakening brand relationship), and changes (e.g., user interface, functionality, layout, etc.) introduced during the use of the application/service?

RQ5. How do users rate the IG mobile application (in the light of sentiment analysis), and what emotions does it evoke in them? Which topics trigger negative emotions?

2. Research Methodology

To answer all the RQs, this section presents the methodology applied in the study. The research methodology part is divided into two main sub-sections. Section 2.1 aims to briefly discuss data collection to create a textual dataset ready for further analysis. Then, Section 2.2 describes in detail the data analysis steps performed in the current study.

2.1. Data

The IG mobile app is available on the App Store and Google Play. This very popular app is used by its users to easily express themselves and connect with friends by (i) adding photos and videos to their Stories; (ii) messaging their friends, sharing and connecting over what is seen on their feed and Stories; (iii) creating and discovering short, entertaining videos on Instagram with Reels; (iv) posting photos and videos to the feed that one wants to show on a profile.

After downloading the app, users can write online reviews about their experience using it. Mobile app store reviews are not usually very long, and they are written and submitted from mobile devices, on which typing is not as easy as on a desktop (Fu et al., 2013).

However, such feedback constitutes an important source of information for the app developer (Instagram, Inc.) and allows developers to interact directly with potential consumers, which may help in the software development process (Cuadrado & Dueñas, 2012). Thus, data for this study cover Android and iOS mobile app reviews available respectively on Google Play and the App Store.

Since most websites do not offer APIs to access user reviews, text data were extracted from Internet sources via web crawling. In total, 567 reviews in English were downloaded, covering March 2020.

2.2. Data analysis

Fig. 2 presents the main steps of the research methodology adopted in this study. These issues are discussed in more detail later in this section. The R software was applied for the data analysis. According to Debortoli et al. (2016), there is no simple recipe for choosing the adequate combination of natural language preprocessing (NLP) steps, and a study's objective and its underlying dataset determine many of them. Due to this, after data collection, the corpus, representing a collection of text reviews about the IG app, was subjected to several preprocessing steps according to standard text mining procedures. First, punctuation marks (periods, commas, hyphens, etc.), numbers and white spaces were removed. Second, the characters in the entire corpus were converted to lowercase. Then, stopwords (extremely common words such as "and", "or", "not", "in", "is", etc.) were removed. Finally, to ensure the terms in the corpus were uniquely identified, Porter's stemming algorithm was used to perform stemming (Žižka et al., 2020).

This study focuses on discovering the hidden sub-themes in the corpus based on the Latent Dirichlet Allocation (LDA) modelling approach. The prepared dataset was first converted into a document-term matrix, which was then subjected to LDA using Gibbs sampling. The LDA algorithm calculates latent distributions of topics and words, given the observed occurrences of words in individual documents. The LDA algorithm was used because it gives better results in terms of generating semantically significant topics and assigning texts to identified topics (Omotosho, 2021). Moreover, LDA enables not only hidden topics to be identified but also the share of each topic in the corpus to be estimated.

First, preparation–modelling–evaluation cycles were performed to determine an appropriate number of topics to extract from the dataset. Five and ten topics were tested, but these less-grained topic models failed to suitably distinguish between topics.

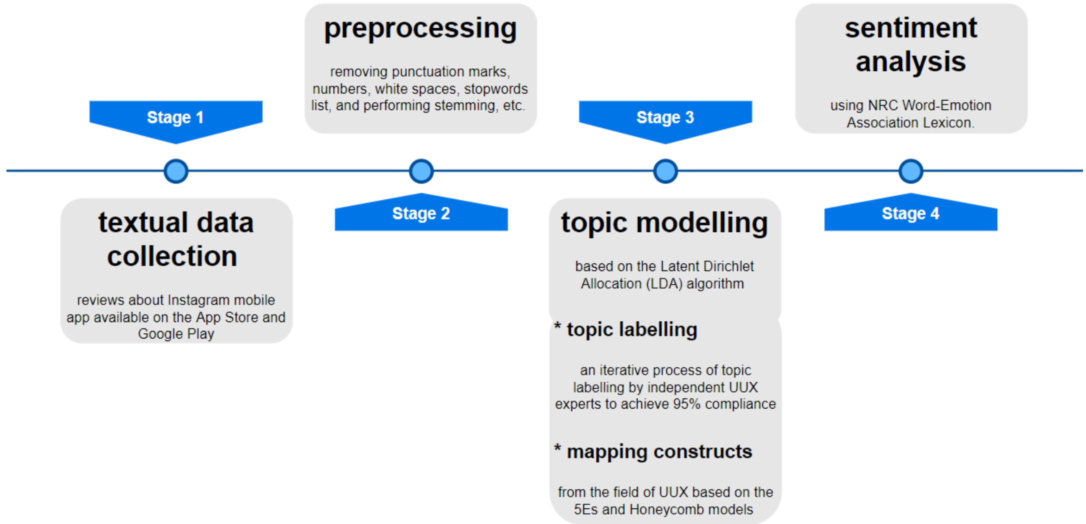

Fig. 2. Overview of the main steps of the approach

Finally, based on Rajkumar (2010) and Griffiths and Steyvers (2004), a 16-topic model was found to be optimal in terms of the average semantic coherence of the model. To validate topic models and evaluate their semantic qualities, the following questions were used following Boyd-Graber et al. (2014): (i) are individual topics meaningful, interpretable, coherent and useful?; (ii) are assignments of topics to documents meaningful, appropriate and useful?

Since each topic is actually a distribution over all words found in the corpus, topics described only by top-weighted keywords were obtained, and a labelling process was needed. According to Debortoli et al.

Tab. 1. Definitions of constructs related to the usability (the 5Es framework)

Construct	Definition	Source
Usability	"The capability of the software product to be understood, learned, used and attractive to the user, when used under specified conditions"	ISO/IEC 9126-1, 2001
	"The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use"	ISO9241-11, 1998
	"The ease with which a user can learn to operate, prepare inputs for, and interpret outputs of a system or component"	IEEE Std.610.12-1990
Sub-construct: facets of the usability	Definition *(Users think that…)*	Identification questions *(Hints posted in user opinions: …)*
Effective	An effective product helps users achieve their goals (like completing a specific task or solving a specific problem) in a specific context	Were the user goals achieved as expected regarding completeness, accuracy and suitability to the specific context?
Efficient	An efficient product helps users achieve their goals with minimal use of resources (such as time, mental effort, manual operations, number of steps, etc.)	Were the user goals achieved with acceptable time and low mental and manual effort?
Engaging	An engaging product motivates users to further activities by presenting attractive visual cues, interesting information and stimulating guidance	Was using the product interesting, attractive and motivating for the users?
Error-tolerant	An error-tolerant product lets users continue their tasks by eliminating opportunities for errors and by providing easily reversible actions for errors that occurred	Was correcting user errors straightforward and effective?
Easy-to-learn	An easy-to-learn product supports users in initial orientation and deeper learning by reusing users' prior knowledge and skills	Was it easy to learn how to operate the product? Was the acquired skill easy to recover after a long break from using the product?

Tab. 2. Definitions of constructs related to user experience (the Honeycomb model)

Construct	Definition	Source
User experience	"Momentary, primarily evaluative feeling (...) while interacting with a product or service"	Hassenzahl, 2008
	"All aspects of the user experience when interacting with the product, service, environment or facility. (…) It includes all aspects of usability and desirability of a product, system or service from the user's perspective"	ISO 9241-210:2019
Sub-construct: facets of the UX	Definition *(Users think that…)*	Identification questions *(Hints posted in users' opinions: …)*
Useful	A useful product satisfies user needs and enables producing an expected outcome (like accomplishing a task or solving a specific problem)	Does the product satisfy specific user needs and produce an expected outcome?
Usable	The expected outcome is accomplished at a reasonable expense of a user's resources (time, mental effort, manual operations, number of steps, amount of data input, etc.)	Is the outcome achieved in an easy and straightforward manner?
Desirable	The outcome of product operation is highly desirable, for instance, solving a problem, providing pleasure, convenience or a valuable relationship	Is the outcome desirable and motivating for repetitive use of the product?
Findable	Finding data and objects needed for accomplishing user goals is easy and straightforward	Are data and objects required to complete the task (data, buttons, menus, paths, search window, etc.) easy to find?
Accessible	A fully accessible product provides its complete functionality and contents for users with visual impairments and other disabilities	Are all functions and contents fully accessible for users with impairments or disabilities? Were any complaints collected regarding this issue?
Credible	A credible product (service) behaves in a predictable manner, exactly as expected by users. The same applies to its provider as to effective support in solving user problems	Is the product (service) trustworthy, supportive and protecting user privacy from harm and abuse?
Valuable	The product (service) is beneficial and attractive for users in building their positive attitudes and valuable relationships	Do users find the product (service, app) important, beneficial and valuable? Do they express positive emotions towards the service provider's brand? Would users recommend the IG app to friends and relatives?

(2016), at least two independent researchers should interpret and label the topics. Due to this, an iterative process of topic labelling was performed by two independent experts who deeply understand the UUX domain and its theoretical foundation. This stage was done in two iterations until consistent results were achieved.

To make sense of the recognised topics in relevance to existing theories, efforts were made to map the discovered topics to theoretical constructs of theories from the field of user experience and usability. First, constructs were mapped in relation to usability based on the 5Es framework by Quesenbery (2004; 2014). Then, Morville's (2004; 2016) user experience Honeycomb model was used to diagnose the UX value during mapping topics. Both models — 5Es and Honeycomb — provide a set of characteristics that can be used to organise and analyse information from IG users.

Both authors mapped topics with constructs independently based on a list of theoretical definitions and identification questions (Tables 1 and 2) to achieve at least 95 % compliance. Finally, all constructs from the two models were joined with the identified topics. The result of this stage is presented in Table A (available on the GitHub repository – https://github.com/Anna-TM-projects/-Exploring-Usability-and-User-Experience/blob/main/Appendix_Tab%20A.pdf).

While the facets of usability present its sub-constructs as the expected features of a product (software, app or website), the identification questions presented in Table 1 provide validation hints about whether a specific facet of usability was satisfied in the users' view. The information required for this validation can be sourced from the observations and measurements; however, as it was in this study, it was also collected from the text mining of user feedback published online.

The definition and identification questions of sub-constructs from both models were not derived solely from expert interpretations but arose from aggregated expertise from literature sources. Thus, 5Es were supported by Quesenbery (2001; 2003;

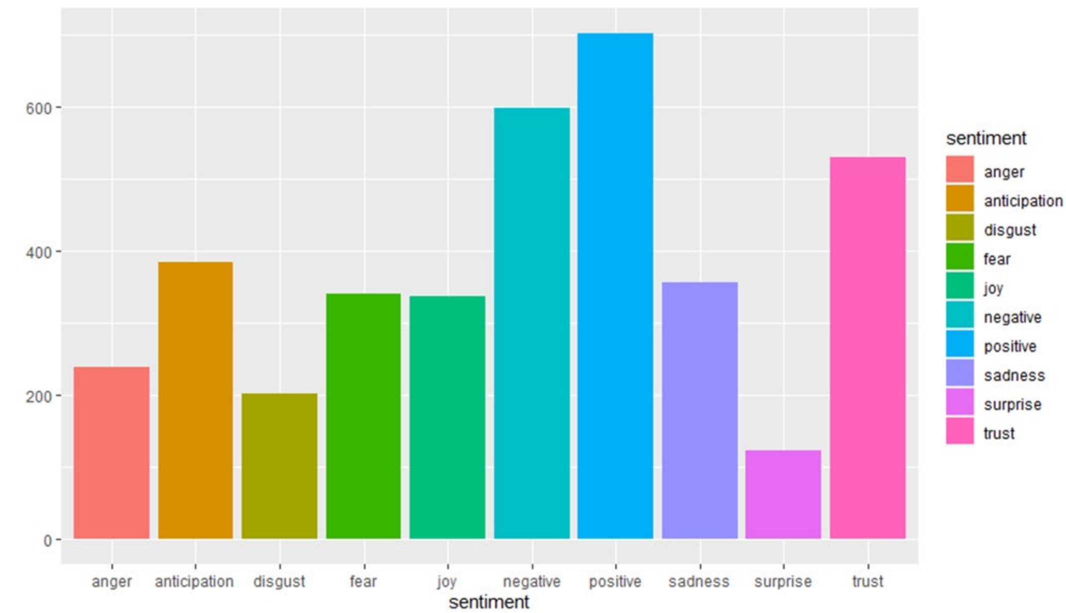

Fig. 3. Sentiment scores distribution in the IG dataset

Source: elaborated by the author in R software.

2004; 2014); Korhan and Ersoy (2016), Häkkilä et al. (2005), Kumar and Mohite (2018), Liu et al. (2018) and Wang and Brennan (2017). Whereas Table 2 presents the "user experience" definition for the Honeycomb model from two sources and its facets derived from aggregated expertise from literature sources (Morville, 2004; 2016; Hartson & Pyla, 2012; Meuthia et al., 2021; Vivakaran & Neelamalar, 2015).

The identification questions presented in Table 2 provide validation regarding whether a specific facet of user experience was satisfied in the users' opinion and relevant to a particular product or service (predominantly, a website or mobile app). The information required for this validation can be sourced from direct observations or emotional responses, but primarily, as it was in this study, for mobile apps, it should be collected directly from users and by the text mining of user feedback published online.

Each facet of the Honeycomb model can provide help as a singular-looking glass, enabling the exploration beyond the standard way. Similarly, insights from the 5Es model can be extracted in terms of the usability of the IG app. This way, each user need statement can be turned into UUX goals or requirements. Thus, calculations were made to find the share of topics (in per cents) included in the facets of each model. This way, information was obtained about how IG users with their reviews fit into the UUX components and what was the percentage share of each facet (Figs. A and B in the Appendix).

A sentiment analysis (SA) was performed to expand the context of this study with emotional information. First, SA was performed for the full dataset to detect general insights into user emotions within all reviews. NRC Word-Emotion Association Lexicon was used to perform this kind of SA. Due to this choice, it was possible to use pairs of opposite emotions from Plutchik's Wheel (Cardone et al., 2021) as a base, i.e., joy as the opposite of sadness, fear — anger, anticipation — surprise, disgust — trust. Plutchik's Model of Emotions (PME) organises feelings and makes sense of them logically, and for this reason, it is used very often. Fig. 3 visualises counts of emotions detected within the full set of IG reviews.

Subsequently, sentiment analysis was also performed for each of the 16 topics at a binary level (positive or negative sentiment). The results are shown in Fig. C in the Appendix.

The analysis of the sentiment for each topic was based on two categories of opinions: positive and negative. Then, the Net Sentiment Rate (NSR) was calculated, according to Baj-Rogowska (2020), by adding a standardised interpretation of the index result as well. The following algorithm was used for the computation of the Net Sentiment Rate:

$$NSR = \frac{(PO-NO)}{(PO+NO)} \tag{1}$$

where: PO — positive opinions, NO — negative opinions. NSR ∈ (-1; 1), where: -1 — opinions are totally negative, 1 — opinions are totally positive.

Whereas to define the strength of the net sentiment, the following classification was applied (Baj-Rogowska, 2020):

- 0.0 < | NSR | ≤ 0.1 — weak positive/negative sentiment,
- 0.1 < | NSR | ≤ 0.3 — average positive/negative sentiment,
- 0.3 < | NSR | ≤ 0.5 — high positive/negative sentiment,
- 0.5 < | NSR | ≤ 0.8 — very high positive/negative sentiment,
- 0.8 < | NSR | < 1.0 — almost complete positive/negative sentiment.

The outcomes provided by the algorithms provided information not only on the user perception of an examined issue in a binary way (negative or positive) but also provided insights into more sophisticated emotions, such as disgust or trust. This provided us with quite comprehensive insights into the issue regarding the content of Instagram reviews.

3. Results and Discussion

Many experiments were performed within the textual dataset analysis, providing the following results. First, the topic modelling study showed how, based on naturally occurring text data in the form of user reviews, it is possible to attain user experience and point out usability attributes (RQ1). The relationship between the topics and the recalled theories from Section 3 of this paper is presented in Table A (available on the GitHub repository – https://github.com/Anna-TM-projects/-Exploring-Usability-and-User-Experience/blob/main/Appendix_Tab%20A.pdf). Sixteen inductively identified topics were mapped to the existing theoretical constructs from the UUX area, thus forming a nomological network. The results showed that the aspects Usable (50 %) and Useful (31 %) had the strongest presence in the set of reviews.

According to the definitions (Tables 1 and 2) and based on SA, the IG app is useful or/and meaningful and does satisfy specific needs or solve specific user problems. Users perceived it as easy to-use and trouble-free, in general. According to Porat and Tractinsky (2012), more useful products evoke a higher level of pleasure, which may, in turn, generate the desire of users to repeat the experience and continue using the application (Cockburn et al., 2017).

On the other hand, user engagement in the topics attributed to the usability facets (the 5Es model) points out that Engaging and Effective were covered the most often in user feedback. Thus, based on the construct definitions (Table 1) and SA outcomes, it can be very generally maintained that users perceive the IG app as engaging, pleasant and satisfying to use. The app's functionalities are engaging and appropriate to the tasks, users and context, and the users appreciate using this app. In addition, during the use of the IG app, the user goals are met successfully, in general, although there have been negative experiences connected with the subjects presented in Table 3. According to Quesenbery (2004), the relationship between the 5Es facets recognised in the study can set the direction for the interface design and help determine the techniques for user research and usability evaluation during a project. This may suggest design approaches and recognise places where changes are necessary to meet user needs.

Furthermore, the analyses did not identify additional explanatory factors, which go beyond the existing UUX theories, and all facets from the two models were present in the textual dataset (RQ2).

To answer RQ3, the obtained results were visualised and presented in Figs. A and B, in the Appendix. They show that the highest user engagement of IG users in the topics discussed in the Honeycomb model concerns two facets, namely, Usable (50 %) and Useful (31 %). In the 5Es model, facets concerning usability are the highest in the area of Engaging (30 %) and Effective (23 %). In turn, the lowest includes the area Efficient (9 %). This confirms the earlier remark that the IG app, to a small extent, may cause some of the difficulties that its users are exposed to.

Jacob and Harrison (2013) claimed that mobile app reviews were valuable repositories of remarks and ideas coming directly from app users. According to Guzman et al. (2015), it is important to consider user feedback when creating and maintaining useful and usable software. Our study also provides evidence in this scope. In answer to RQ4, we can say that we gain knowledge about general problems reported by users by identifying topics with negative sentiment (Table 3), as well as discovering knowledge about positive user remarks from positive sentiment. This information might be helpful for the early detection of frustration and negative feelings, identifying missing features and changes in terms of, e.g., user interface, functionality, layout, etc., and could finally bring an improvement in software quality. This has also

Tab. 3. Topics with negative sentiment

Topic No	Topic Label	NSR value	The strength of the net sentiment
#1	The app's changes (unsuccessful/frustrating)	-0.1724	average negative sentiment
#2	Technical problems	-0.3005	high negative sentiment
#8	Quality issues	-0.0136	weak negative sentiment
#9	Functionalities to be brought back	-0.1000	weak negative sentiment
#10	Issues of hiding story/viewer highlights	-0.4545	high negative sentiment
#11	Posts publishing	-0.1200	average negative sentiment
#13	Adding some options to switch on and off	-0.7000	very high negative sentiment
#15	User suggestions on what to fix	-0.6667	very high negative sentiment

Tab. 4. Sample reviews on the hot topic "Viewer lists are not available after 24 hours"

Review No	Rate (1–5)	Sample Reviews (original spelling)
# 191	Rated 1 star out of 5	"wait, what? we can't watch the viewers on our own stories (the ones with more than 24 hours in archives)? this is bad. really very bad. what are you guys doing!! bring our privacy back! what was wrong earlier if we can see who viewed our own stories after 24 hours???!!!"
#304	Rated 4 stars out of 5	"I HATE this newest update for instastory: „Viewer lists arent available after 24 hours." ...it's INCONVENIENT and LESS USEFUL. I like to be able to see instastory-viewer-list whenever I want. I think it's way better if you change it back like it was before"
#330	Rated 1 star out of 5	"I want to see the viewers list after 24 hours. if people can check my highlights after 24 hours, then I should be able to see them. they go hand in hand! you shouldn't have tried to fix what isn't broken. #bringbacktheviewerslist"
#399	Rated 1 star out of 5	"Hate that we can no longer see who has viewed stories after 24 hours. Not everyone has the time to constantly check their viewers while their story is live. Also, what's the point of telling us how many viewers but not who? (Especially for close friends)"
#505	Rated 3 stars out of 5	"Please explain why you remove the feature who viewed my stories?!!! why only 24 hours?! ill rate back on 5 when that feature got back"
#555	Rated 2 stars out of 5	"*viewers list not available after 24hours* this new update, trust me is of no use. To be very honest it's horribleeee!? and this is needed to be changed :)"

been confirmed by Pagano and Bruegge (2013) and Pagano and Maalej (2013), who reported that users very often shared their needs, ideas and experiences in their reviews, thus users provide useful feedback to the application vendors and developers for improving software quality.

It is noteworthy that topics with negative sentiment are more related to features and functionality issues (e.g., "#9: Functionalities to be brought back", "#13: Adding some options to switch on and off" etc.), whereas positive topics are more connected to general perceptions and human aspects (e.g., "#14: Assessment of the Insta app", etc.). This is in line with Nakamura et al. (2022) and shows that unsatisfied users are willing to provide details about the functionalities and aspects that evoke frustration while giving positive reviews; they are prone to describing the overall qualities and aspects of the app.

In the full review collection, the hot topic (i.e., the largest number of reviews written) with high negative sentiment was "#10: Issues of hiding story/viewer highlights". Similar assessments were given to "#2: Technical problems". Examining the comments assigned to a given topic allows for obtaining more detailed information about the reported problems. Sample reviews about the reduction in viewing time for Instagram Stories after the new update (#bringbacktheviewerslist) are given in Table 4. It clearly shows what users cannot accept and what they find very irritating. Meeting these expectations can prevent brand image deterioration and the weakening of the brand relationship.

In line with Pagano and Maalej (2013), it was also found that the spectrum of feedback quality was varied, from helpful opinions for other users and developers to useless noise to insults. Valuable and useful comments are present fairly often and include information about errors and app crashes, technical problems, feature requests etc. This helps developers uncover user needs and experiences to develop the application in line with crowdsourcing requirements.

The study results also showed that most user feedback emerged shortly after new releases (i.e., a new update), and its frequency relatively quickly fell over time. Thus, analysing the feedback over several releases would help detect problematic features or other connected issues early.

Finally, to answer RQ5, sentiment analysis was performed per whole dataset (Fig. 3) and per single topic (Fig. C in the Appendix), and the results of emotion detection were presented by the visualisations. For the full dataset, joy and sadness were the first pair of emotions for analysis. Both emotions are on a relatively similar level, which means a similar number of reviews are happy and sad in the dataset.

The next comparison between fear and its opposite — anger — points out that users are very eager to share their experiences to publish their annoyance on the Web. As we first expected, the analysis confirmed that the IG mobile app topic brought lots of anger, stress and frustration. Further analysis indicated the causes of these negative emotions. The very high negative sentiment was connected with the following topics:

- #13 Adding some options to switch on and off (-0.7000)
- #15 User suggestions on what to fix (-0.6667)
- #10 Issues of hiding story/viewer highlights (-0.4545)
- #2 Technical problems (-0.3005)

The next pair of opposite emotions is anticipation, which was quite high in the reviews, and surprise, which had a comparatively low value. Anticipation is linked with curiosity and exploration.

Users seem to be looking for help when searching for solutions to their problems and are prone to exchange their experiences. Surprise is associated with something shocking or unexpected. This may imply that users are not finding many special shocking experiences for them connected with the IG app.

The last comparison, between disgust and trust, reveals a value of two-and-a-half times higher for trust than disgust. Disgust has a negative meaning and links to rejection, distrust or being uninterested. Whereas the value of trust in the entire set of emotions is superior to others. Due to this, we can conclude that IG users are recognized as being trusting and welcoming of the IG app.

To sum up, the SA revealed that in the entire dataset (Fig. 3), the positive sentiment of the reviews exceeded the negative sentiment. Users were more eager to share more positives than negatives while presenting their experiences with the IG app.

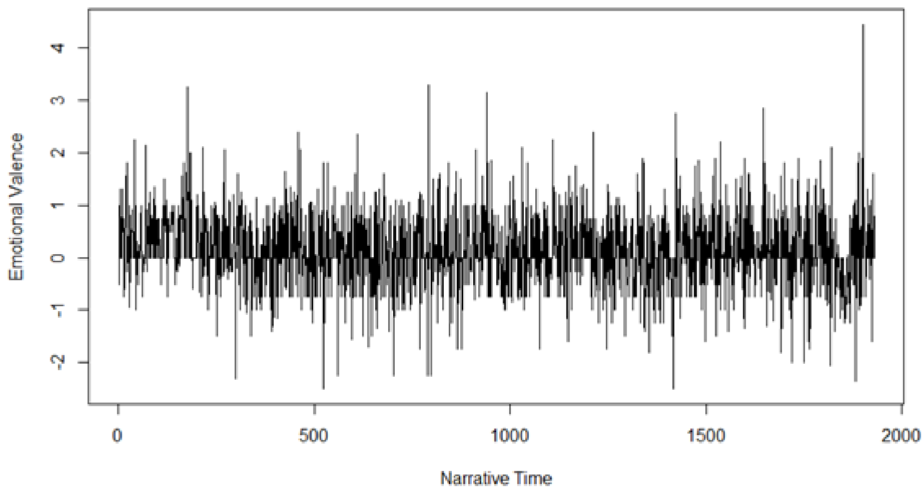

Fig. 4. Plot trajectory of emotional valence during a certain time in the full dataset
Source: elaborated by the author in R software.

Fig. 4 presents a moving average trend line for the full textual data to present how the sentiment of reviews developed over time. Data are presented in narrative time with the calculation of the mean sentiment valence for each time. This graph can be useful for getting a sense of the emotional trajectory of all reviews. From here, as well, it is clear that the distribution of positive and negative reviews is relatively even, and there are few outliers.

Because the main goal of this study was to evaluate the applicability of a text mining approach for extracting UUX-related issues from the dataset of user comments and opinions and not to evaluate the IG app, the most important conclusions and remarks are presented in the next section. This also includes potential novelties and contributions of this study to the science.

Conclusions

Starting from the IG example, this study aimed to perform a proof-of-concept of a novel approach, potentially opening new frontiers for UUX-focused exploration of available datasets of user comments and opinions regarding social media applications. The initial results seem promising, with some concluding remarks — achievements and limitations — presented below.

The present paper's contribution is fourfold. First, the study identifies, classifies and maps topics discussed in the user reviews into facets of two models from UUX. To the authors knowledge, this is the first such study in the literature. Second, the analyses pointed out that all facets of the two models were present in the user reviews dataset, which may confirm the usefulness of the methodology in software development and assessment. Third, a consolidated taxonomy chosen for this study was validated, and a clear definition of facets was created as well as identification questions for each of the two models to categorise the reviews. Fourth, the case study of the Instagram app confirmed the usefulness of user feedback data for software development and pointed out that the review data have the potential for the early detection of frustration, negative feelings (e.g., brand image deterioration, weakening brand relationship), and changes (e.g., user interface, functionality, layout, etc.) introduced during the use of the application.

Further validation of the approach should be conducted on all the proposed stages against a much larger text data set. Moreover, a good idea would be for future research to include a time perspective regarding each new app release, which could provide insights into UUX factors changing over time. In particular, modifications and changes that negatively — in the users' view — affect the functionality and user experience of mobile apps are frequently commented on online with obvious negative sentiment (Fashionunited, 2022; Digitalinformationworld, 2022). This provides an interesting opportunity for using text mining to analyse newly published user comments and to suggest responsive policies for service providers.

Definitely, the novel approach proposed in this paper requires more validation and further case studies, also aimed at reducing the limitations of this approach, briefly presented below.

First, in line with [Debortoli et al., 2016], we claim text mining methods such as topic modelling cannot replace human analysis, but should only accelerate and augment it. This is exactly the approach we recommend. However, a certain limitation of our approach is the time-consuming expert analyses.

Second, considering the assessments of experts (during labelling and mapping), the same results are not always obtained because the human factor is based on subjective feelings. To reduce the risk of researcher/expert bias, two researchers took part in the labelling and mapping processes, independently from each other, based on the prepared list of theoretical definitions and identification questions, achieving at least a 95 % compliance level.

Third, it is important to emphasise that the evaluation of UUX by analysing user feedback should not entirely replace the already existing methods but only complement them resolving limitations and provide quick, useful information for software development.

Finally, considering the findings of Hu et al. (2006) that reviews are mainly written by a group of very satisfied or very dissatisfied customers, no guarantees can be given that the set of reviews used for data processing was representative, i.e., whether it fully reflected the description of typical experiences in the user feedback base. In particular, despite the collected opinions being published only by IG mobile app users, no knowledge was available regarding the extent users' prior experiences (which surely existed in some cases) with the IG website could influence their opinions about the IG mobile app.

Hopefully, the approach can be useful for helping developers, designers and researchers to reveal user problems and helpful in fulfilling user satisfaction regarding UUX aspects for specific software features. After further validation studies, this approach may prove useful for the early detection of UUX flaws for any online services which are already in use when their users are available only remotely and provided a dataset of their credible comments can be collected. If so, the text mining approach may open interesting opportunities for extracting information on UUX-related topics, such as sources of user discontent, frustration and annoyance. In the longer term, the input from text mining can be used in recovery procedures to identify problems, locate them in the system, prioritise and decide on further actions or mitigating policies. Last but not least, the text mining approach may offer an interesting advantage: additional possibilities for a focused view of user opinions by regions and languages according to users' locations, as far as possible, with a specific dataset.

Literature

Abbasi, A., Sarker, S., & Chiang, R. H. (2016). Big data research in information systems: Toward an inclusive research agenda. *Journal of the Association for Information Systems*, *17*(3).

Adikari, S., McDonald, C., & Campbell, J. (2011). A design science framework for designing and assessing user experience. In *International Conference on Human-Computer Interaction* (pp. 25–34). Berlin, Heidelberg: Springer.

Baj-Rogowska, A. (2020). Evaluation of a company's image on social media using the Net Sentiment Rate. In E. Lechman & M. Popowska (Eds.), *Society and Technology Opportunities and Challenges* (pp. 202–218). doi: 10.4324/9780429278945

Bakiu, E., & Guzman, E. (2017). Which feature is unusable? Detecting usability and user experience issues from user reviews. In *2017 IEEE 25th international requirements engineering conference workshops (REW)* (pp. 182-–187). IEEE.

Battarbee, K. (2003). Co-experience: the social user experience. In *CHI '03 Extended Abstracts on Human Factors in Computing Systems (CHI EA '03)* (pp. 730–731). New York, NY, USA: Association for Computing Machinery. doi: 10.1145/765891.765956

Berenbach, B., Paulish, D. J., Kazmeier, J., & Rudorfer, A. (2009). *Software & Systems Requirements Engineering: In Practice*. Retrieved from https://www.accessengineeringlibrary.com/content/book/9780071605472

Bevan, N. (2008). Classifying and selecting UX and usability measures. In *Proceedings of Meaningful Measures: Valid Useful User Experience Measurement (VUUM), 5th COST 294-MAUSE Open Workshop* (pp. 13–18). Reykjavik, Iceland.

Bilgihan, A. (2016). Gen Y customer loyalty in online shopping: An integrated model of trust, user experience and branding. *Computers in Human Behavior*, *61*, 103-113.

Blythe, M., Overbeeke, K., Monk, A., & Wright, P. (2003). *Funology: from usability to enjoyment*, Kluwer Academic Publishers.

Boyd-Graber, J., Mimno, D., & Newman, D. (2014). Care and feeding of topic models: Problems, diagnostics, and improvements. In E. M. Airoldi, D. Blei, E. A. Erosheva, & S. E. Fienberg (Eds.), *Handbook of mixed membership models and their applications* (pp. 3–34). Boca Raton: CRC Press.

Cardonen, B., Di Martino, F., & Senatore, S. (2021).Improving the emotion-based classification by exploiting the fuzzy entropy in FCM clustering. *International Journal of Intelligent Systems*, *36*(11), 6944-6967. doi: 10.1002/int.22575

Carreño, L. V., & Winbladh, K. (2013). Analysis of User Comments: An Approach for Software Requirements Evolution. In *Proceedings of the 2013 International Conference on Software Engineering ser. ICSE'13* (pp. 582–591).

Chang, H. H., & Chen, S. W. (2009). Consumer perception of interface quality, security, and loyalty in electronic commerce. *Information & Management*, *46*(7), pp. 411-417.

Chen, J. V., Yen, D. C., Pornpriphet, W., & Widjaja, A. E. (2015). E-commerce website loyalty: A cross cultural comparison. *Information Systems Frontiers*, *17*(6), 1283-1299.

Cockburn A., Quinn P., & Gutwin C. (2015). Examining the Peak-End Effects of Subjective Experience. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)* 9pp. 357–366). New York, NY, USA: Association for Computing Machinery. doi: 10.1145/2702123.2702139

Cuadrado, F., & Dueñas, J. C. (2012). Mobile application stores: success factors, existing approaches, and future developments. *IEEE Communications Magazine*, *50*(11), 160-167.

Debortoli, S., Müller, O., Junglas, I., & vom Brocke, J. (2016). Text Mining For Information Systems Researchers: An Annotated Topic Modeling Tutorial. *Communications of the Association for Information Systems*, *39*. doi: 10.17705/1CAIS.03907

Digitalinformationworld (2022). Retrieved from https://www.digitalinformationworld.com/2022/08/instagram-rumored-to-be-following-in.html

Eid, M. I. (2011). Determinants of e-commerce customer satisfaction, trust, and loyalty in Saudi Arabia. *Journal of Electronic Commerce Research*, *12*(1), 78.

Fashionunited (2022). Retrieved from https://fashionunited.uk/news/fashion/instagram-after-a-backlash-pauses-changes-to-its-algorithm/2022080864536

Følstad, A., & Rolfsen, R. K. (2006). Measuring the Effect of User Experience Design Changes in ECommerce Web Sites: A Case on Customer Guidance. In *Proceedings of the 2nd COST294-MAUSE International Open Workshop* (pp. 10–15). Oslo, Norway.

Forlizzi, J., & Battarbee, K. (2004). Understanding experience in interactive systems. In *Proceedings of the 2004 Conference on Designing Interactive Systems (DIS 04):*

Processes, Practices, Methods, and Techniques. New York: ACM.

Fu, B., Lin, J., Li, L., Faloutsos, C., Hong, J., & Sadeh, N. (2013). Why people hate your app: making sense of user feedback in a mobile app store. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '13* (p. 1276). Presented At the 19th ACM SIGKDD International Conference. Chicago, Illinois, USA: ACM Press.

Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences*, *101*(suppl 1), 5228-5235. doi: 10.1073/pnas.0307752101

Guzman, E., Aly, O., & Bruegge, B. (2015). Retrieving Diverse Opinions from App Reviews. In *Proceedings of the 2015 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement ser. ESEM'15* (pp. 1–10).

Häkkilä, J., Korpipää, P., Ronkainen, S., & Tuomela, U. (2005). Interaction and end-user programming with a context-aware mobile application. In *Proc. IFIP Conf. Hum.-Comput. Interact* (pp. 927–937). Berlin, Germany: Springer.

Hartson, R., & Pyla, P. S. (2012). *The UX Book. Process and Guidelines for Ensuring a Quality User Experience*. Elsevier.

Hassenzahl, M. (2004). The interplay of beauty, goodness, and usability in interactive products, *Human-Computer Interaction*, *19*(4), 319-349.

Hassenzahl, M. (2008). User Experience (UX): Towards an Experiential Perspective on Product Quality. In *Proceedings of the 20th Conference on l'Interaction Homme-Machine* (pp. 11–15). New York, NY, USA: Association for Computing Machinery.

Hassenzahl, M., & Tractinksy, N. (2006). User experience: a research agenda. *Behaviour and Information Technology*, *25*(2), 91-97.

Hedegaard, S., & Simonsen, J. G. (2013). Extracting usability and user experience information from online user reviews. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2089–2098).

Hu, N., Pavlou P. A., & Zhang J. (2006). Can online reviews reveal a product's true quality?: Empirical findings and analytical modeling of online word-of-mouth communication. In *Proceedings of the 7th ACM Conference on Electronic Commerce (EC)* (pp. 324–330).

IEEE Std 610.12-1990 (1990). *IEEE standard glossary of software engineering terminology*. Institute of Electrical and Electronics Engineers, Inc.

ISO 9241-11:2018 (2018). *Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts*. International Standards Organization.

ISO 9241-210: 2019 (2010). *Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems*. International Standards Organization.

ISO/IEC 25010:2011 (2011). *Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) - System and software quality models*. International Standards Organization.

ISO/IEC 9126-1:2001 (2001). *Software engineering — Product quality — Part 1: Quality model*. International Standards Organization.

ISO/IEC, 9241-11:1998 (1998). *Ergonomic requirements for office work with visual display terminals (VDT)s - Part 11 Guidance on usability*. 1998: ISO/IEC 9241-11: 1998 (E).

Jacob, C., & Harrison, R. (2013). *Retrieving and analyzing mobile apps feature requests from online reviews*. In *Proceedings of the Working Conference on Mining Software Repositories ser. MSR'13* (pp. 41–44).

Jiménes, K., Pincay, J., Villavicencio, M., & Jiménez, A. (2018). Looking for usability and functionality issues: a case study. In *International Conference on Information Technology & Systems* (pp. 948–958). Cham: Springer.

Korhan, O., & Ersoy, M. (2016). Usability and functionality factors of the social network site application users from the perspective of uses and gratification theory. *Qual Quant, 50*, 1799-1816. doi: 10.1007/s11135-015-0236-7

Kumar, B. A., & Mohite, P. (2018). Usability of mobile learning applications: A systematic literature review. *Journal of Computers in Education*, *5*(1), 1-17.

Liu, R., Cao, J., Zhang, K., Gao, W., Liang, J., & Yang, L. (2018). When privacy meets usability: Unobtrusive privacy permission recommendation system for mobile apps based on crowdsourcing. *IEEE Transactions on Services Computing*, *11*(5), pp. 864–878.

Maalej, W., & Pagano, D. (2011). On the socialness of software. In *2011 IEEE Ninth International Conference on Dependable Autonomic and Secure Computing (DASC)* (pp. 864–871).

McCarthy, J., & Wright, P. (2004). *Technology as Experience*. Cambridge, MA, USA: MIT Press.

McNamara, N., & Kirakowski, J. (2006). Functionality, Usability, and User Experience: Three Areas of Concern. *Interactions*, *13*, 26-28.

Meuthia, R. F., Ferdawati, F., & Gustati, G. (2021). Analisis User Experience Pada Aplikasi E-Kelurahan Menggunakan Model UX Honeycomb. *Journal Akuntansi Keuangan dan Bisnis*, *14*(2), 342-351.

Morville, P. (2004). User Experience Design. Semantic Studios: Interwingled. Retrieved from http://semanticstudios.com/publications/semantics/000029.php

Morville, P. (2016). User experience honeycomb. Semantic Studios: Interwingled. Retrieved from https://intertwingled.org/user-experience-honeycomb/

Nakamura, W. T., de Oliveira, E. C., de Oliveira, E. H. T., Redmiles, D., & Conte, T. (2022). What factors affect the UX in mobile apps? A systematic mapping study on the analysis of app store reviews, *Journal of Systems and Software*, *193*. doi: 10.1016/j.jss.2022.111462

Nielsen Norman Group. (2021). Social Media UX: User-Centered Social Strategies. Retrieved from https://www.nngroup.com/reports/social-media-user-experience/

Norman, D. A. (1998). *The invisible computer*. Cambridge, Mass: MIT Press.

Omotosho, B. S. (2021). Analysing user experience of mobile banking applications in Nigeria: A text mining approach. *CBN Journal of Applied Statistics*, *12*(1), 77-108.

Pagano, D., & Bruegge, B. (2013). User Involvement in Software Evolution Practice: A Case Study. In *Proceedings of the 35th International Conference on Software Engineering*. San Francisco, CA: IEEE.

Pagano, D., & Maalej, W. (2013). User Feedback in The Appstore: an Empirical Study. In *Proceedings 2013 21st IEEE International Requirements Engineering Conference ser. RE'13* (pp. 125–134).

Park, J., Han, S. H., Kim, H. K., Cho, Y., & Park, W. (2013). Developing elements of user experience for mobile phones and services: survey, interview, and observation approaches. *Human Factors and Ergonomics in Manufacturing & Service Industries*, *23*(4), 279-293.

Pereira, R., Baranauskas M. C. C., & da Silva, S. R. P. (2010). A Discussion on Social Software: Concept, Building Blocks and Challenges. *International Journal for Infonomics* (IJI), *3*(4).

Porat, T., & Tractinsky, N. (2012). It's a pleasure buying here: The effects of web-store design on consumers' emotions and attitudes. *Human–Computer Interaction*, *27*(3), 235-276.

Portugal, R. L., & Sampaio do Prado Leite, J. C. (2018). Usability Related Qualities Through Sentiment Analysis. In *1st International Workshop on Affective Computing for Requirements Engineering (AffectRE)* (pp. 20–26).

Quesenbery, W. (2004). Balancing the 5Es: Usability. *Cutter IT Journal*, *17*(2), pp. 4-11.

Quesenbery, W. (2014). The five dimensions of usability. In *Content and complexity* 9pp. 93–114). Routledge.

Rajkumar, A. V., Suresh, C. E., Madhavan, V., & Narasimha Murthy, M. N. (2010). On finding the natural number of topics with Latent Dirichlet Allocation: Some observations. In M. J. Zaki, J. X. Yu, B. Ravindran & V. Pudi (Eds.). *Advances in knowledge discovery and data mining* (pp. 391–402). Berlin Heidelberg: Springer. doi: 10.1007/978-3-642-13657-3_43

Saavedra, M. J., Rusu, C., Quiñones, D., & Roncagliolo, S. (2019). A Set of Usability and User eXperience Heuristics for Social Networks. In G. Meiselwitz (Eds.), *Social Computing and Social Media. Design, Human Behavior and Analytics*. HCII 2019. Lecture Notes in Computer Science, 11578. Springer, Cham. doi: 10.1007/978-3-030-21902-4_10.

Sánchez, J. L. G., Vela, F. L. G., Simarro, F. M., & Padilla-Zea, N. (2012). Playability: analysing user experience in video games. *Behaviour & Information Technology*, *31*(10), 1033-1054.

Sauer, J., Sonderegger, A., & Schmutz, S. (2020). Usability, User Experience and Accessibility: Towards an Integrative Model. *Ergonomics*, *63*, 1207-1220.

Sharp, H., Rogers, Y., & Preece, J. (2023). *Interaction design: beyond human-computer interaction*. 6th Edition. London: John Wiley.

Shneiderman, B., & Plaisant, C. (2016). *Designing the user interface: strategies for effective human-computer interaction.* Boston: Pearson Education.

Sikorski, M. (2008). HCI and the Economics of User Experience. In E. Law, E. Hvannberg & G. Cockton (Eds.), *Maturing Usability* (pp. 318–343. London: Springer-Verlag.

Smith, G. (2007). *Social Software Building Blocks*. Retrieved from http://nform.ca/publications/social-software-building-block

Souza, P. C. D., & Maciel, C. (2015). *Legal issues and user experience in ubiquitous systems from a privacy perspective*. In *International Conference on Human Aspects of Information Security, Privacy, and Trust* (pp. 449–460). Cham: Springer.

Sutcliffe, A. (2009). *Designing for User Engagement: Aesthetic and Attractive User Interfaces*. San Rafael, CA, USA: Morgan & Claypool Publishers.

Väänänen-Vainio-Mattila, K., Wäljas, M., Ojala, J., & Segerståhl, K. (2010). Identifying drivers and hindrances of social user experience in web services. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2499–2502).

Vivakaran, M. V., & Neelamalar, M. (2015). *From Spectators to Users: A study on the transformation caused by the Social media in the Indian Media Context*.

Wang, Y., & Brennan, R. (2017). *Building high usability consumer-oriented linked data mobile apps*. In *Proceedings Irish Conference on Artificial Intelligence and Cognitive Science (AICS)* (pp. 1–12). Dublin, Republic of Ireland.

Weichbroth, P., & Baj-Rogowska, A. (2019). Do online reviews reveal mobile application usability and user experience? The case of WhatsApp. In *Federated Conference on Computer Science and Information Systems* (pp. 747–754).

Žižka, J., Dařena, F., & Svoboda, A. (2020). *Text Mining with Machine Learning: Principles and Techniques*. Taylor & Francis Group, CRC Press.

Appendix

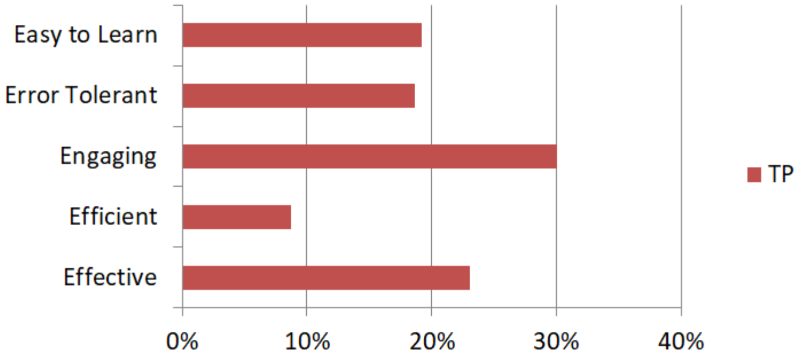

Fig. A. 5Es Model: user engagement in the discussed topics in the facets of Usability

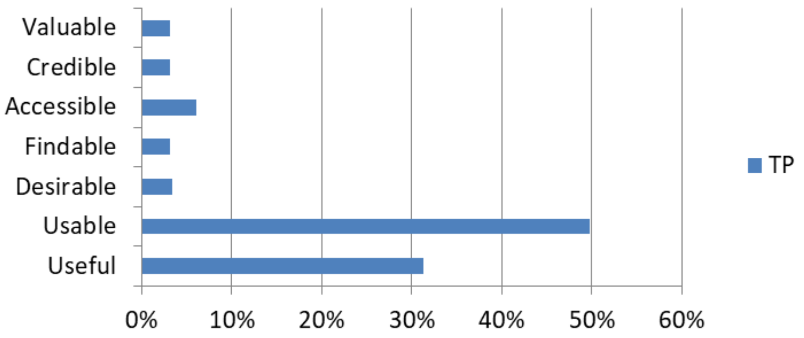

Fig. B. Honeycomb model: user engagement in the discussed topics in the facets of UX

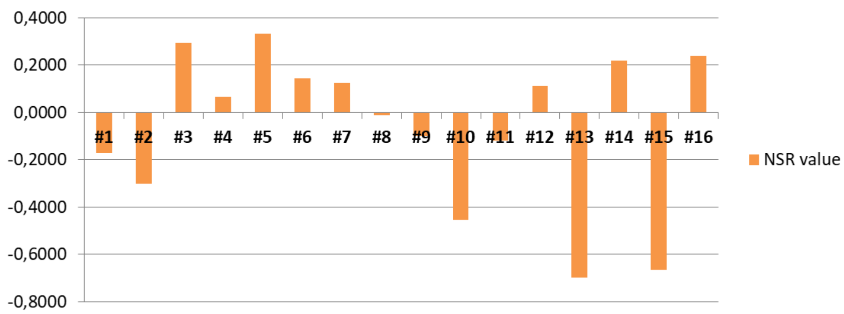

Fig. C. Sentiment distribution for each topic