Chapter 5
Emotion Analysis in Software Ecosystems

Nicole Novielli and Alexander Serebrenik

Abstract Software developers are known to experience a wide range of emotions while performing development tasks. Emotions expressed in developer communication might reflect openness of the ecosystem to newcomers, presence of conflicts, problems in the software development process, or source code itself. In this chapter, we present an overview of the state-of-the-art research on analysis of emotions in software engineering focusing on the studies of emotion in context of software ecosystems. To encourage further applications of emotion analysis in the industry and research, we also include a table summarizing currently available emotion analysis tools and datasets as well as outline directions for future research.

5.1 What Is a Software Ecosystem?

Several definitions of software ecosystems can be found in the literature [14, 64, 77, 78, 82, 83]. Rather than selecting one of these definitions a priori, we have decided to start with adopting a nominalistic approach, i.e., state that an ecosystem is whatever is being called an ecosystem. Following this approach, we conduct a literature review of sentiment and emotion in software ecosystems recording the definitions of the ecosystems used in the primary studies, the ecosystems considered, and the insights obtained in the primary studies.

N. Novielli (✉)
University of Bari, Bari, Italy
e-mail: nicole.novielli@uniba.it

A. Serebrenik
Eindhoven University of Technology, Eindhoven, The Netherlands
e-mail: a.serebrenik@tue.nl

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

T. Mens et al. (eds.), *Software Ecosystems*,
https://doi.org/10.1007/978-3-031-36060-2_5

To conduct the literature study, we reuse the collection of 186 articles collected by Lin et al. as part of their systematic literature review of opinion mining[1] for software development [75]. The authors used ACM Digital Library, IEEE Xplore Digital Library, Springer Link Online Library, Wiley Online Library, Elsevier ScienceDirect, and Scopus. The following search query was used to locate primary studies in these online databases:

("**opinion mining**" OR "**sentiment analysis**" OR "**emotion**") AND ("**software**") AND ("**developer**" OR "**development**")

We perform a full-text search for the term **ecosystem** in the 186 articles. While additional articles in the collection of Lin et al. [75] might have studied ecosystems without using the term "ecosystem," the goal of this section is not to provide a comprehensive overview of emotion analysis in software ecosystems but rather to identify what kind of artefacts are usually being called "ecosystems" in the literature on opinion mining for software development. After excluding the articles that mention the term "ecosystem" only in the bibliography, we obtain 28 primary studies. None of them has provided a formal definition of an ecosystem.

Ten articles refer to the "software development ecosystem" or "social programmer ecosystem" as an entire collection of different channels and communication means available to a contemporary software developer, e.g., the Software Engineering Arousal lexicon (SEA) has been specifically designed to address the problem of detecting emotional arousal in the software developer ecosystem [80], and Novielli et al. applied sentiment analysis to such components of the ecosystem as GitHub and Stack Overflow [96]. Two articles explicitly talk about "a rich ecosystem of communication channels" [42, 97]. Four articles refer to the ecosystem of mobile apps: online reviews from an unnamed store [57] or the iTunes and Google Play marketplaces [41, 51, 79, 87] and StackOverflow questions about Android, iOS, and Windows phone [74]. Similarly to the studies of the "software development ecosystem," this line of research seems to implicitly focus on the presence of a shared communication platform (e.g., app store, GitHub, or Stack Overflow) akin to the definition of Bosch and Bosch-Sijtsema [14]: "A software ecosystem consists of a software platform, a set of internal and external developers and a community of domain experts in service to a community of users that compose relevant solution elements to satisfy their needs." Differently from this definition, "a software platform" in this line of research is also conceptualized as a collection of interrelated technical platforms or communication channels, e.g., GitHub and Stack Overflow.

Seven studies have focused on open-source software communities: Tourani et al. [130] and Ortu et al. [104] considered Apache; in a different paper Ortu et al. [107] further extended the data to include Spring, JBoss, and CodeHaus. In addition, Ferreira et al. studied the Linux kernel [38], Umer et al. [133] the

[1] "Opinion mining" is a broader area than sentiment and emotion, but lions' share of the opinion mining studies in the software engineering context have been dedicated to sentiment and emotion.

reports from the Mozilla issue tracker collected by Nizamani et al. [92], Tourani and Adams [129] Eclipse and Open Stack, and finally Boudeffa et al. [15] OW2. These studies tend to focus on several projects within the ecosystem chosen, e.g., on ten Open Stack and five Eclipse projects [129] or on XWIKI, Sat4j, and asm from OW2 [15]. The focus on projects within the ecosystems is shared with the way Lungu has approached ecosystems as collections of jointly developed projects [77].

Finally, several papers have used the word "ecosystem" in a very generic sense not necessarily disclosing a particular meaning [4, 39].

> **Definition**: In the context of the sentiment and emotion studies in software engineering, "ecosystems" are often seen as:
>
> - either platforms or collections of interrelated communication platforms supporting software development (e.g., GitHub, Stack Overflow, Google Play app store)
> - or as collections of interrelated software projects (e.g., Apache, Eclipse, OW2).

5.2 What Is Emotion?

Emotions have always been in the center of human inquiry with pre-Socratic philosophers being among the first to think about this topic [122]. Emotions have been studied by numerous philosophers [122], historians [109], sociologists [71], psychologists [91], biologists and neurophysiologists [1], economists [136], musicologists [68], literature scientists [53], and computing researchers [23]. Despite this, or maybe due to this, a definition of emotion proved to be elusive [40, 122]: as aptly stated by Fehr and Russel, "everyone knows what an emotion is, until asked to give a definition. Then, it seems, no one knows" [37]. It should come as no surprise then that multiple theories of emotion have been proposed in the literature. Gross and Barrett [47] and Meiselman [81] propose to arrange emotion theories along a continuum ranging from theories of basic emotion through theories of appraisal to psychological construction theories. Among these theories, those on the extremes have found their way into studies of emotion in software engineering: theories of basic emotions such as the one of Ekman [36] consider emotions to be universal and distinct from each other, and while those psychological construction theories tend to see emotions as continuous space organized along several dimensions, e.g., Russel's circumplex model of affect [114].

Russel [114] has observed that the distinction between such emotions as *sadness* and *anger* present in English is absent from some African languages, while English misses words for Bengali *obhiman*, which refers to sorrow caused by the insensitivity of a loved one or German *Schadenfreude*, which refers to pleasure derived from another's displeasure. Based on these and similar arguments, he argued that emotions can be operationalized along several dimensions: valence,[2] arousal, and dominance. Valence expresses the degree of pleasantness of the emotion and typically can be characterized on the scale from negative to positive. Arousal corresponds to degree of activation in the emotion and can be scale from low to high. Dominance is related to feeling in control or feeling controlled. Using these dimensions, Russel states that both *excitement* and *calmness* can be characterized by positive valence and feeling in control, with arousal being high for *excitement* and low for *calmness*. *Anger* shares with *excitement* high arousal and dominance but differs from it by negative valence.

Ekman [36] believes emotions to be separate and distinct from each other. He associated emotions with facial expressions [35] and distinctive patterns of activation of the autonomic nervous system [34], as well as connected emotions in humans to comparable expressions observed in other primates [34]. Ekman further argued that there are more emotional words than actual emotions and that only emotions satisfying specific criteria can be seen as basic emotions. These emotions are *anger*, *surprise*, *disgust*, *enjoyment*, *fear*, and *sadness*; later research suggests that *contempt* should be seen as a basic emotion too. Starting from a similar list of basic emotions, Shaver et al. [119] have proposed a tree-like structure gradually refining these emotions to emotion names. This hierarchy of emotion labels includes basic (primary) emotions, which are further refined into secondary and tertiary ones, e.g., *anger* is refined to such secondary emotions as *envy*, *rage*, and *exasperation*, with such tertiary subspecies of *rage* as *outrage*, *hatred*, or *dislike*.

Plutchik's wheel[3] of emotions [111] combines discrete and dimensional elements: while Plutchik argues that only a small number of basic emotions exist (and other emotions can be synthesized by combining the basic ones), he also recognizes that each emotion can exist at different levels of arousal, distinguishing between, e.g., *"blues"*, *sadness*, and *grief*. Moreover, emotions on the opposite sides of the wheel are opposing: e.g., *joy* and *sadness* and *expectancy* and *surprise*.

The aforementioned models have been used when studying emotions in the context of software engineering: for example, Murgia et al. [89] have used the model by Shaver et al. [119], as presented in Parrott [108], Khan and Saleh [72] chose the Plutchik's wheel; Girardi et al. [44] opted for Russel's circumplex model of emotion. Similarly to the latter work, many studies of emotion in software engineering implicitly adopt a dimensional model; however, as opposed to it, these studies focus only on the valence of emotion. Such studies tend to call valence "sentiment" and consider it to be negative, neutral, or positive [12, 110, 123].

[2] Russel [114] used the term "pleasure" for "valence," but "valence" is more commonly used in subsequent publications.

[3] "Wheel of emotions" is a latter term; the original paper by Plutchik referred to a "three-dimensional emotion solid" with degree of arousal providing the third dimension.

Theories: While multiple theories of emotion can be found in the literature, studies of emotion in the context of software engineering either use theories with a small number of distinct emotions or those positioning emotions in a continuous one- or multidimensional space.

5.3 Why Would One Study Emotions in Software Engineering?

Software development has been often stereotyped as a job with few interpersonal requirements [31], and, hence, one might doubt the importance of studying emotions experienced and expressed by software developers.

Our answer to the question in the section title is twofold. First, software development has long been recognized as a problem-solving activity [112], and emotions are known to impact problem-solving skills and creativity [5, 46]. Second, software development is a collaborative process [10], and such sites as GitHub and Stack Overflow further require communication to facilitate knowledge sharing and co-creation of software [124]. Previous research in organizational psychology investigated the relation between emotions and knowledge sharing in a global IT organization. The study found that pride and empathy positively impact the willingness to share knowledge and are influenced by knowledge-sharing intentions in their turn [54]. Wurzel Gonçalves et al. [141] investigated interpersonal conflicts in code review and found that they are common and often perceived as an opportunity to learn from disagreement, thus highlighting the need for developing strategies for constructive resolution of conflicts. On the other hand, Murphy-Hill et al. [48] demonstrated the potential negative impact on motivation to continue working with colleagues after receiving destructive criticism in code reviews. Indeed, the presence of negative emotions is one of the dimensions of an interpersonal conflict [7]. To illustrate the latter point, Wurzel Gonçalves et al. report the following comment made in a Linux mailing list in October 2015: "*Christ people. This is just sh*t. The conflict I get is due to stupid new gcc header file crap. But what makes me upset is that the crap is for completely bogus reasons.*" Along the same line, anger in the statements such as "*Is there any progress on this issue??*" has been considered by Gachechiladze et al. [42] for identification of actionable insights in issue handling.

This is why substantial research effort has been dedicated to understanding emotions experienced and expressed by developers, their triggers, and consequences. However, in order to answer these questions, one has to measure emotions first.

Emotions in Software Engineering: Emotions influence both cognitive processes such as problem-solving and interpersonal interaction, both important elements of software engineering.

5.4 How to Measure Emotion?

Emotion measurement is an important topic in emotion research [27, 81]. In the following, we summarize the recent advancements in emotion recognition in software engineering. Specifically, we report about available tools and dataset, specifically designed to support emotion recognition in the context of software development. For further discussion of the ways emotions of software developers are measured, we refer the reader to the recent article by Sánchez-Gordón and Colomo-Palacios [115] and for discussion of software-engineering-specific sentiment analysis tools and datasets to Lin et al. [75] and Obaidi and Klünder [102].

5.4.1 Tools

Scientific literature on emotion measurement covers a broad spectrum of techniques including psychophysiological signals (e.g., electrodermal skin response, neuroendocrine factors, or heart rate) [93], observation of behavior (e.g., vocal and verbal characteristics or body expressions and postures) [63], measurement of facial expressions [56], self-reporting questionnaires [24], and text analysis [84]. Many of these techniques have been applied in context of software engineering as well: e.g., Girardi et al. have analyzed psychophysiological signals to recognize emotions of developers in the lab [45, 86] and in the field [44, 135]; Novielli et al. [100] have argued that facial expressions should be used as a gold standard; self-reporting questionnaires such as SAM [16] and PANAS [137] have been used by Çalikli et al. [22] and Schneider et al. [117], respectively. As the communication between software developers is to a large extent text-based, the use of vocal information for emotion detection has not been explored in the software engineering context. The work of Herrmann and Klünder [52] takes the first step in this direction. The authors advocate usage of audio to analyze emotion present in software project meetings. However, their approach starts with converting audio to text and hence ignoring the tone and intonation that can reflect emotion experienced by the meeting participants.

The very same dominance of text-based communication has led to lions' share of emotion measurement techniques to focus on textual artefacts produced by software developers, e.g., code review comments, Stack Overflow questions, commit messages, or bug descriptions. Early studies of sentiment and emotion in software engineering used text analysis tools developed for very different kinds of text: e.g., a number of studies [43, 49, 104] have used SentiStrength [127], a tool originally designed for and evaluated on social Web datasets (Myspace, Twitter, YouTube, Digg, Runner's World, BBC Forums) [126]. However, as observed by Novielli et al. [96], when such tools are applied in the context of software engineering, they produce unreliable results, threatening validity of the previously published conclusions [66]. This observation has led to emergence of a series of software-engineering-specific sentiment analysis [3, 11, 18, 25, 32, 62, 118] and emotion detection tools [20, 59, 61]. As most of these tools are based on machine learning, retraining is recommended when applying them to a different kind of text than the one they have been designed for [94], as indeed different software-engineering-specific sentiment analysis tools might lead to contradictory results at a fine-grain level, when used off the shelf [99]. Further empirically driven recommendations include carefully choosing the emotion model in line with the research goals, as the operationalization of emotions adopted by the designer of an emotion detection tool might not necessarily match the focus and goal of a given empirical study. Furthermore, when a manually annotated gold standard is not available for retraining, lexicon-based tools such as SentiStrengthSE might represent a viable option [98].

To encourage further applications of emotion analysis in the industry and research, we also include Table 5.1 summarizing currently available emotion analysis tools.

5.4.2 Datasets

As an output of recent empirical research in this field, several annotated datasets have been released by SE researchers to further encourage training and fine-tuning of software engineering-specific sentiment analysis tools. Murgia et al. [89] release a dataset of Jira comments labeled according to the Shaver's primary emotions, namely, joy, love, surprise, anger, fear, and sadness. The dataset, initially composed of 400 text items, was further extended using the same annotation schema by Ortu et al. [107]. Using the same taxonomy, Calefato et al. annotated more than 4000 questions, answers, and comments from Stack Overflow [18]. Other than releasing the emotion labels, the authors also provide a mapping to the valence dimension, thus labeling each text item as either positive (joy, love), negative (anger, fear, sadness), or neutral (absence of emotion label). Surprise was mapped to either positive or negative valence depending on the context. Novielli et al. [94] adopted the same coding guidelines for labeling emotions and mapping them to positive, negative, and neutral in annotating about 7000 comments from GitHub projects. Stack Overflow posts (1500 overall) were also annotated by Lin et al. in the scope of an empirical study on mining positive and negative opinion of developers about software libraries [76]. Jira comments (500 overall) were also annotated by Kaur et al. [70] according to polarity labels, as well as by Islam and Zibran [61], who labeled 1800 Jira comments to identify the presence of excitement, stress, depression, and relaxation. Beyond Jira, Stack Overflow, and GitHub, other data sources were used such as code review comments [3] and tweets [140]. In their survey, Lin et al. [75] report a detailed list of available datasets for sentiment polarity/emotion/politeness detection, which can be used as a gold standard for training supervised classifiers. They also include consideration of dataset annotate including a broader set of emotion-related mental states, such as confusion [33].

Table 5.1 Tools for sentiment/emotion detection in software engineering. Based on previous literature reviews [75, 102, 103] and updated

Tool	Methodology	Based on	Theoretical model
		Sentiment detection	
SentiStrength-SE [62]	Lexicon-based	issue reports	positive, neutral, negative
Senti4SD [18]	Traditional machine learning	Stack Overflow posts	positive, neutral, negative
SEntiMoji [25]	Deep learning	issue reports, Stack Overflow posts, code reviews	positive, neutral, negative
SentiSW [32]	Deep learning	issue reports	positive, neutral, negative
SentiCR [3]	Traditional machine learning	code reviews	non-negative, negative
SentiSE [118]	Traditional machine learning	code reviews	positive, neutral, negative
Unnamed classifier [143]	Transformer models	issue reports, Stack Overflow posts, code reviews, app reviews, GitHub pull-request and commit comments	positive, neutral, negative
Unnamed classifier [11]	BERT-based	Stack Overflow posts	positive, neutral, negative
EASTER [125]	Deep learning	Stack Overflow posts, app reviews, JIRA issues	positive, neutral, negative
Sentisead [131]	Ensemble	Stack Overflow posts, issue reports, app reviews	positive, neutral, negative
		Emotion detection	
SO BERT emotion classifier [13]	BERT-based	Stack Overflow posts	Distinct emotions: love, joy, surprise, anger, sadness, fear
DEVA [61]	Lexicon-based	issue reports	discretization of the two-dimensional valence/arousal model: excitement, stress, depression, relaxation, neutral
MarValous [59]	Traditional machine learning	Stack Overflow posts, issue reports	discretization of the two-dimensional valence/arousal model: excitement, stress, depression, relaxation, neutral
Unnamed classifier [88]	Traditional machine learning	issue reports	Distinct emotions: joy, love, sadness, neutral
EmoTxT [20]	Traditional machine learning	StackOverflow posts, issue reports	Distinct emotions: joy, love, sadness, anger, surprise, fear. Neutral is assigned in absence of other emotions
Unnamed classifier [17]	Traditional machine learning	Stack Overflow posts	Distinct emotions: joy, love, sadness, anger, surprise, fear, objective

Measurement: While a broad specter of emotion measurement techniques can be found in the psychological literature, and many of them have been applied in the software engineering context, text-based techniques (sentiment analysis) remain dominant. Multiple sentiment analysis techniques have been designed especially for software engineering.

5.5 What Do We Know About Emotions and Software Ecosystems?

Following our observations in Sect. 5.1, we organize this section in two subsections according to the two interpretations of the concept of an "ecosystem," as a (collection of interrelated) communication platform(s) or as a collection of interrelated projects. A word of caution is in place though: due to use of different datasets and tools, conclusions derived by similar studies might appear contradictory. Moreover, validity conclusions about texts created by software engineers but derived using general-purpose sentiment analysis tools that have not been adjusted to the software engineering domain, should be reassessed as those tools are known to be unreliable in the software engineering context [66, 96]. In particular, this is the case for all results published prior to 2017 as the first software engineering specific sentiment analysis tool has been published in 2017. For example, such insights of Guzman et al. [49] as "Java ⟨GitHub⟩ projects tend to have a slightly more negative score than projects implemented in other languages" or that comments on Monday were more negative than comments on the other days could not have been confirmed when a different sentiment analysis tool has been used [66].

5.5.1 Ecosystems as Communication Platforms

In this section, we discuss two popular developer communication platforms, Stack Overflow and GitHub.

5.5.1.1 Stack Overflow

Stack Overflow is a major Q&A platform that has been frequently considered in the research literature in general [2, 6, 9, 90] and through the lens of emotion analysis in particular [17–19, 21, 65, 66, 76, 85, 95, 97, 131, 132]. However, many papers have merely used the data from Stack Overflow to evaluate the sentiment analysis rather than to obtain insights in the development practices on Stack Overflow. We exclude these papers from the subsequent discussion.

Several studies have tried to relate sentiment expressed in Stack Overflow posts to success (e.g., ability to receive an answer) or quality (e.g., as expressed in terms of upvotes and downvotes). Mondal et al. [85] have observed that the upvoted questions tend to be more positive than the downvoted ones. Jiarpakdee et al. [65] have shown that inclusion of sentiment-related variables improves prediction of whether a Stack Overflow question will get an accepted answer. Refining this insight, Calefato et al. [21] recommend the users to write questions using a neutral emotional style as expressing emotions is associated with lower probability of success, i.e., receiving an answer that is accepted as a solution. Finally, Calefato et al. [19] observed that comments rather than questions and answers tend to express emotions and that this can be attributed to the fact that comments do not influence the reputation scores and hence can be seen as a kind of “lawless region” where anything goes.

Focusing on the Stack Overflow discussions about API, Uddin and Khomh [132] have observed that certain aspects of APIs such as *performance* triggered more opinionated statements than other aspects of APIs such as security. Among these opinionated statements, *security*-related ones are predominantly positive, while the opinions related to *performance* and *portability* are much more mixed. Zooming in on specific domains, Uddin and Khomh observed that the distribution of opinions for a given aspect varies, e.g., the opinions about performance are mostly positive for API features related to serialization but mostly negative with regard to the debugging of the performance issues of the APIs.

Finally, Cagnoni et al. [17] have used sentiment information to complement indicators of the popularity of programming languages such as the one by TIOBE.[4] They have observed that programming languages associated with the highest share of positive posts on Stack Overflow are not necessarily the same as those developers indicate as the most loved languages in a survey: while MATLAB and R trigger the highest share positive emotions in Stack Overflow posts, Rust and Kotlin are the languages indicated as being “loved” by the highest percentage of the Stack Overflow survey. In fact, Python is only one programming language shared by the top-10 of the most loved languages and the top-10 of the programming languages that have triggered the highest share of positive emotions. One might wonder what makes the language “loved”: the insights of Cagnoni et al. [17] suggest that there is more to this than positive atmosphere in the support community.

5.5.1.2 GitHub

Similarly to Stack Overflow, GitHub has been extensively studied in the research literature [28, 128, 134], even triggering a methodological research on how GitHub-based studies should be conducted [69]. Emotion analysis has been also repeatedly conducted on GitHub data [29, 32, 49, 55, 58, 60, 67, 105, 106, 110, 116, 120, 121, 123, 138, 139, 142]. Also similarly to Stack Overflow, manually labeled datasets

[4] https://www.tiobe.com/tiobe-index/.

derived from GitHub have been used to design and evaluate sentiment analysis tools; we do not discuss those papers below.

The first group of studies has considered the impact of negative or positive GitHub-related artefacts such as issues and commits on the software development process. Souza and Silva [123] have observed that commits with negative sentiment are slightly more likely to result in broken builds. In a similar vein, Huq et al. [55] have observed negative emotions in contributor commits to indicate that fix-inducing change might be needed, while the statistical model of Ortu et al. [106] suggests that issues expressing dominance and *sadness* are less likely to be merged. Taken together, these studies [55, 106, 123] suggest that commits expressed more negatively deserve a more careful review as they might have undesirable consequences. When it comes to positive emotions, Huq et al. [55] also claimed that "too much positive emotions in discussion may lead to buggy code" as positivity "can turn developers overconfident and careless, reducing their ability to scrutinize their own code" and potentially biasing other reviewers. The latter point might also be related to the observation that positive valence and specifically emotion of *joy* is linked with a higher probability of merge [106]. Preference of neutral emotional style stemming from these studies is reminiscent of a similar recommendation of Calefato et al. [21] to write Stack Overflow questions using a neutral emotional style.

The second group of studies has investigated the software engineering context where specific emotions can be observed. For example, Souza and Silva [123] have shown that commits following a build breakage tend to be more negative. Singh and Singh [120] found that developers express more negative sentiments than positive sentiments when performing refactorings. This finding contradicts the observation of Islam and Zibran [60] that positive emotions are significantly higher than negative emotions for refactoring tasks, despite the fact that both studies [60, 120] have considered the general-purpose sentiment analysis tool SentiStrength [127] and in both cases the tool has been adjusted to the software engineering context. However, the adjustment has not necessarily been carried out in the same way, and the datasets considered have been different, which might explain the difference between the results. Furthermore, differences between the conclusions might be related to different kinds of refactoring activities carried out by the developers: e.g., Singh and Singh [120] observed that high negativity could be particularly attributed to move-refactorings, rename classes, or attributes being pulled up, while no such information is available for the study of Islam and Zibran [60]. Rather than distinguishing between specific types of software development activities such as refactoring or bug fixing, Pletea et al. [110] have focused on the application domain these activities take place in and compared security-related GitHub entities and non-security-related ones. The authors have concluded that security-related entities are more negative than the rest of the entities and that this does not depend on whether one considers as "entities" commits or pull requests and individual comments or entire discussions. This conclusion has been confirmed by subsequent replication studies [66, 99].

Several studies have focused on the influence of the day of the week on the sentiment. Guzman et al. [49] have stated that comments on Monday were more negative than comments on the other days, but this finding could not be confirmed through replication [66]. Sinha et al. [121] reported that overall, the most negative day was Tuesday, while for projects with the highest number of commits in their dataset, the most negative days were Wednesday and Thursday, suggesting that the differences in the distribution of the sentiment over the week might be project-specific. Islam and Zibran [60] observed negative emotions to be slightly higher in commit messages posted during the weekends than those posted in weekdays and not much differences are visible in the emotional scores for commit messages posted in the five weekdays. Ultimately, evidence on presence of day-related differences in developers' sentiment is inconclusive at the very least.

Ortu et al. [105] and Destefanis et al. [29] compared sentiment in communication of developers and users. According to Ortu et al. [105], when commenting users express more *love*, *sadness*, *joy*, and *anger* than developers; for replies, however, the situation is partially reversed: developers tend to have expressed more positive emotions (*love* and *joy*) and less negative ones (*sadness* and *anger*). Using a complementary perspective on the theory of emotion, Destefanis et al. [29] observed that commenters expressed fewer emotions than users, while they communicated with higher levels of arousal, valence, and dominance.

Finally, Yang et al. [142] and Jurado and Rodríguez Marín [67] have conducted studies on collections GitHub *projects* focusing on similarities and differences between these projects. We discuss these papers in Sect. 5.5.2.1.

5.5.2 Ecosystems as Interrelated Projects

5.5.2.1 GitHub

Yang et al. [142] and Jurado and Rodríguez Marín [67] have conducted studies on collections of GitHub *projects* focusing on similarities and differences between these projects, i.e., as opposed to the studies discussed in Sect. 5.5.1.2, they interpret the notion of an ecosystem as a collection of projects rather than as a shared communication space. Yang et al. [142] observed that the rate of bug-fixing speed increases with emotional values increasing for 13 projects of their dataset, while it decreases for seven projects. Unfortunately, no explanation has been provided for this phenomenon. Jurado and Rodríguez Marín [67] have studied distribution of emotions across nine projects: they have observed that at least 80% of the communication does not express emotion and that the most expressed emotion is *joy* accounting for 4.66–11.94%. *Disgust* is the least present emotion barely found in the dataset. The authors have also observed differences between the projects: e.g., *fear* is overrepresented in Raspberry Pi compared to other projects. Pandas has shown two instances of *fear*; however, one of these instances merely reflected presence of lexicon related to *fear* rather than actual *fear* experienced by developers:

"The terrible motivating example was this awful hack." These findings suggest that (a) different projects might have different project culture impacting frequency of different emotions being expressed and (b) pure lexicon-based approaches cannot capture complexity of software engineering communication.

5.5.2.2 Apache

Several studies have considered the Apache ecosystem [26, 70, 88, 89, 113, 130]. Based on the Parrott's model [108], Murgia et al. [88, 89] have observed that developers express all emotions from the model. While some emotions can refer both to software artifacts and coworkers (e.g., *joy*, *anger*, and *sadness*), others target only artifacts (e.g., *surprise* and *fear*) or coworkers (e.g., *love*). One should keep in mind though that in this study, *love* is mostly represented as gratitude ("Thanks very much! I appreciate your efforts"), *joy* as satisfaction of the development process or its results ("I'm happy with the approach and the code looks good"), and *sadness* as developers apologizing for their mistakes ("Sorry for the delay Stephen") or expressing their dissatisfaction ("Apache Harmony is no longer releasing. No need to fix this, as sad as it is"). Rigby and Hassan [113] describe "developer B," the top committer for 1999 and 2000 who has left the project later on. As developer B was preparing to leave, their language shifted from describing new insights to explaining previously taken decisions, and the number of positive emotions decreased as well. Focusing on emoticons, Claes et al. [26] have observed that in more than 90% of occurrences, they are used to express *joy*. Moreover, emoticons are used more often in Apache projects during the weekend than during weekdays; the effect size was, however, small. Tourani et al. [130] have studied the Apache mailing lists and observed that almost 70% of the communication is neutral, about 20% are positive, and slightly more than 10% are negative. For emails with the positive sentiment, user mailing lists contain substantially more *curiosity* than developer mailing lists, while for emails with negative sentiment, user mailing lists contain more *sadness* and less *aggression* than developer mailing lists.

5.5.2.3 Other Ecosystems

Next we review studies of sentiment and emotion in other ecosystems.

Two studies have targeted the Eclipse ecosystem [101, 129]. Using SentiStrength, Tourani and Adams observed that increase in the lowest sentiment score, i.e., most negative score becoming less negative, has a positive but small effect on defect proneness. By manually analyzing the initial posts of threads at the Eclipse forum and the corresponding first replies, Nughoro et al. [101] have observed that Junior contributors and Members post more positive messages than Senior contributors; moreover, Juniors both start positive interactions and receive positive responses, while Seniors post initiate positive, neutral, and negative communication. These observations seem to concur with the idea that the relative power of the actor and target affects the extent to which display rules require controlling one's expressions [30, 50].

Two further studies have focused on Mozilla. Umer et al. [133] have shown that inclusion of sentiment improves prediction of whether the enhancements proposed by Mozilla contributors will be integrated. While similarly to Apache the most popular emoticon in Mozilla represents *joy*, the share of *sad* and *surprised* emoticons in Mozilla is more than twice higher [26].

Ferreira et al. [38] have studied sentiment expressed on the Linux Kernel mailing list. While no differences in sentiment across releases, months, and weeks have been observed, Sunday, Tuesday, Wednesday, and Thursday had more positive than negative sentiment. Referring to the specific event of Linus Torvalds taking a break from the community in September 2018, the authors investigated whether it has affected sentiment within the community. While the difference in the sentiment was not immediate, positive sentiment at the level of months and weeks after his break has increased.

Lanovaz and Adams [73] compared the sentiment in two mailing lists of the R community: one targeting developers and another one helping users. Developers showed marginally more positive and negative tones than users, and while negative messages by the users did not receive response, this was not the case for developers. One might wonder whether this difference could be attributed to developers seeing users as customers and hence neutralizing their emotions [30].

Sentiment in Ecosystems: When ecosystems are treated as communication platforms, sentiment is used to predict the outcome of developer's activities on these platforms or to understand differences in the context where different sentiment is observed. When ecosystems are treated as collections of projects, the studies focus on differences between different kinds of contributors, different projects, and popularity of different emotions.

5.6 What Next?

In this chapter, we have provided an overview on the state-of-the-art resources for sentiment analysis in software engineering, with specific focus on their application to software ecosystems. The good performance achieved by the available SE-specific sentiment analysis tools provides an evidence that reliable sentiment analysis in software development is possible provided that SE-specific tools are used. Still, open challenges remain for sentiment analysis, in general, and on developers' communication traces, in particular.

Tools based on supervised machine learning might produce a different performance on different data sources due to platform-specific jargon and communication style [98]. As such, we recommend retraining supervised tools using a gold standard from the same domain and data source being targeted. Model retraining or fine-tuning might be required also in a within-platform setting, as language constantly evolve, especially in the context of online interaction. It is the case of emoji, for example, that recently emerged as a predominant way of conveying emotional content [25].

Traditionally, sentiment analysis research has predominantly focused on the English language, also due to the availability of resources for this language (e.g., sentiment lexicons and toolkit). We highlight the need for future research to focus on different languages, in order to support effective interaction in projects and communities that do not use English as a predominant language in their online communication.

Finally, while most of the approaches used so far have focused on a single type of measurement, Lisa Feldman Barrett has recently developed a constructionist approach for measurement of emotions, advocating a multimodal approach toward measurement going beyond solely facial analysis or self-reporting or psychophysiology [8]. Along the same dimension, Novielli et al. advocate in favor of the design and implementation of tools combining multiple approaches for emotion assessment, to fully support emotion awareness during software development. Specifically, they envisage the emergence of tools and practices including both self-reporting of emotions through experience sample and emotion detection through using biometrics, as they might provide complementary information on the emotional status of an individual [100].

5.7 What Have We Discussed in This Chapter?

Software engineering processes depend on the emotions experienced and expressed by software developers. To get insights in these emotions, psychological theories and automated tools have been developed. Using these theories and tools, multiple studies have investigated software ecosystems through the lens of emotion. Most such studies have considered ecosystems as interrelated communication platforms supporting software development such as GitHub or Stack Overflow, e.g., recommending developers how to ask questions on Stack Overflow or aiming at understanding the impact of emotions on software engineering or context where emotions are likely to emerge. Other studies investigate on ecosystems as collections of projects such as Apache or Eclipse focusing on experiences of developers in these communities.

References

1. Adolphs, R., Anderson, D.J.: The Neuroscience of Emotion: A New Synthesis. Princeton University Press, Princeton (2018)
2. Ahasanuzzaman, M., Asaduzzaman, M., Roy, C.K., Schneider, K.A.: CAPS: a supervised technique for classifying Stack Overflow posts concerning API issues. Empirical Softw. Eng. **25**(2), 1493–1532 (2020). https://doi.org/10.1007/s10664-019-09743-4

3. Ahmed, T., Bosu, A., Iqbal, A., Rahimi, S.: SentiCR: a customized sentiment analysis tool for code review interactions. In: International Conference on Automated Software Engineering (ASE), pp. 106–111. IEEE, Piscataway (2017). https://doi.org/10.1109/ASE.2017.8115623
4. Ali, N., Hong, J.E.: Value-oriented requirements: eliciting domain requirements from social network services to evolve software product lines. Appl. Sci. **9**(19), 3944 (2019)
5. Amabile, T.M., Barsade, S.G., Mueller, J.S., Staw, B.M.: Affect and creativity at work. Administrative Sci. Q. **50**(3), 367–403 (2005). https://doi.org/10.2189/asqu.2005.50.3.367
6. Anderson, A., Huttenlocher, D., Kleinberg, J., Leskovec, J.: Discovering value from community activity on focused question answering sites: a case study of Stack Overflow. In: SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), pp. 850–858. ACM, New York (2012). https://doi.org/10.1145/2339530.2339665
7. Barki, H., Hartwick, J.: Interpersonal conflict and its management in information system development. MIS Q. **25**(2), 195–228 (2001)
8. Barrett, L.F.: How Emotions Are Made: The Secret Life of the Brain. HarperCollins (2017)
9. Barua, A., Thomas, S.W., Hassan, A.E.: What are developers talking about? An analysis of topics and trends in Stack Overflow. Empirical Softw. Eng. **19**(3), 619–654 (2014). https://doi.org/10.1007/s10664-012-9231-y
10. Begel, A., Herbsleb, J.D., Storey, M.A.: The future of collaborative software development. In: Conference on Computer Supported Cooperative Work (CSCW), pp. 17–18. ACM, New York (2012). https://doi.org/10.1145/2141512.2141522
11. Biswas, E., Karabulut, M.E., Pollock, L., Vijay-Shanker, K.: Achieving reliable sentiment analysis in the software engineering domain using bert. In: International Conference on Software Maintenance and Evolution (ICSME), pp. 162–173. IEEE, Piscataway (2020). https://doi.org/10.1109/ICSME46990.2020.00025
12. Biswas, E., Vijay-Shanker, K., Pollock, L.L.: Exploring word embedding techniques to improve sentiment analysis of software engineering texts. In: International Conference on Mining Software Repositories (MSR), pp. 68–78. IEEE / ACM (2019). https://doi.org/10.1109/MSR.2019.00020
13. Bleyl, D., Buxton, E.K.: Emotion recognition on stackoverflow posts using bert. In: 2022 IEEE International Conference on Big Data (Big Data), pp. 5881–5885 (2022). https://doi.org/10.1109/BigData55660.2022.10020161
14. Bosch, J.: From software product lines to software ecosystems. In: International Software Product Line Conference (SPLC) (2009)
15. Boudeffa, A., Abherve, A., Bagnato, A., Thomas, C., Hamant, M., Montasser, A.: Application of computational linguistics techniques for improving software quality. In: International Conference on Product-Focused Software Process Improvement (PROFES). Lecture Notes in Computer Science, vol. 11915, pp. 577–582. Springer, Berlin (2019). https://doi.org/10.1007/978-3-030-35333-9%5C_41
16. Bradley, M.M., Lang, P.J.: Measuring emotion: the self-assessment manikin and the semantic differential. J. Behav. Therapy Exp. Psychiatry **25**(1), 49–59 (1994). https://doi.org/10.1016/0005-7916(94)90063-9
17. Cagnoni, S., Cozzini, L., Lombardo, G., Mordonini, M., Poggi, A., Tomaiuolo, M.: Emotion-based analysis of programming languages on Stack Overflow. ICT Express **6**(3), 238–242 (2020). https://doi.org/10.1016/j.icte.2020.07.002
18. Calefato, F., Lanubile, F., Maiorano, F., Novielli, N.: Sentiment polarity detection for software development. Empirical Softw. Eng. **23**(3), 1352–1382 (2018). https://doi.org/10.1007/s10664-017-9546-9
19. Calefato, F., Lanubile, F., Marasciulo, M.C., Novielli, N.: Mining successful answers in Stack Overflow. In: Working Conference on Mining Software Repositories (MSR), pp. 430–433. IEEE, Piscataway (2015). https://doi.org/10.1109/MSR.2015.56
20. Calefato, F., Lanubile, F., Novielli, N.: EmoTxt: a toolkit for emotion recognition from text. In: International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACII Workshops), pp. 79–80. IEEE, Piscataway (2017). https://doi.org/10.1109/ACIIW.2017.8272591

21. Calefato, F., Lanubile, F., Novielli, N.: How to ask for technical help? Evidence-based guidelines for writing questions on Stack Overflow. Inf. Softw. Technol. **94**, 186–207 (2018). https://doi.org/10.1016/j.infsof.2017.10.009
22. Çalikli, G., Al-Eryani, M., Baldebo, E., Horkoff, J., Ask, A.: Effects of automated competency evaluation on software engineers' emotions and motivation: a case study. In: International Workshop on Emotion Awareness in Software Engineering (SEmotion), pp. 44–50. ACM, New York (2018). https://doi.org/10.1145/3194932.3194939
23. Calvo, R.A., D'Mello, S., Gratch, J., Kappas, A.: The Oxford Handbook of Affective Computing. Oxford University Press, Oxford (2014)
24. Cardello, A.V., Jaeger, S.R.: Measurement of consumer product emotions using questionnaires. In: Meiselman, H.L. (ed.) Emotion Measurement, pp. 165–200. Woodhead Publishing (2016). https://doi.org/10.1016/B978-0-08-100508-8.00008-4
25. Chen, Z., Cao, Y., Lu, X., Mei, Q., Liu, X.: SEntiMoji: an emoji-powered learning approach for sentiment analysis in software engineering. In: Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE), pp. 841–852. ACM, New York (2019). https://doi.org/10.1145/3338906.3338977
26. Claes, M., Mäntylä, M., Farooq, U.: On the use of emoticons in open source software development. In: International Symposium on Empirical Software Engineering and Measurement (ESEM). ACM, New York (2018). https://doi.org/10.1145/3239235.3267434
27. Coan, J.A. (ed.): Handbook of Emotion Elicitation and Assessment. Series in Affective Science. Oxford University Press, Oxford (2007)
28. Dabbish, L., Stuart, C., Tsay, J., Herbsleb, J.: Social coding in GitHub: transparency and collaboration in an open software repository. In: International Conference on Computer Supported Cooperative Work (CSCW), pp. 1277–1286. ACM, New York (2012). https://doi.org/10.1145/2145204.2145396
29. Destefanis, G., Ortu, M., Bowes, D., Marchesi, M., Tonelli, R.: On measuring affects of github issues' commenters. In: International Workshop on Emotion Awareness in Software Engineering (SEmotion), pp. 14–19. ACM, New York (2018). https://doi.org/10.1145/3194932.3194936
30. Diefendorff, J.M., Greguras, G.J.: Contextualizing emotional display rules: examining the roles of targets and discrete emotions in shaping display rule perceptions. J. Manag. **35**, 880–898 (2009). https://doi.org/10.1177/0149206308321548
31. Diefendorff, J.M., Richard, E.M.: Antecedents and consequences of emotional display rule perceptions. J. Appl. Psychol. **88**(2), 284–294 (2003)
32. Ding, J., Sun, H., Wang, X., Liu, X.: Entity-level sentiment analysis of issue comments. In: International Workshop on Emotion Awareness in Software Engineering (SEmotion), pp. 7–13. ACM, New York (2018). https://doi.org/10.1145/3194932.3194935
33. Ebert, F., Castor, F., Novielli, N., Serebrenik, A.: Confusion detection in code reviews. In: International Conference on Software Maintenance and Evolution (ICSME), pp. 549–553. IEEE, Piscataway (2017). https://doi.org/10.1109/ICSME.2017.40
34. Ekman, P.: An argument for basic emotions. Cogn. Emotion **6**(3-4), 169–200 (1992). https://doi.org/10.1080/02699939208411068
35. Ekman, P.: Facial expression and emotion. Am. Psychol. **48**, 384–392 (1993). https://doi.org/10.1037/0003-066X.48.4.384
36. Ekman, P.: Basic emotions. In: Dalgleish, T., Powers, M.J. (eds.) Handbook of Cognition and Emotion, pp. 45–60. Wiley, London (1999)
37. Fehr, B., Russell, J.A.: Concept of emotion viewed from a prototype perspective. J. Exp. Psychol.: Gener. **113**, 464–486 (1984)
38. Ferreira, I., Stewart, K., Germán, D.M., Adams, B.: A longitudinal study on the maintainers' sentiment of a large scale open source ecosystem. In: International Workshop on Emotion Awareness in Software Engineering (SEmotion), pp. 17–22. IEEE/ACM (2019). https://doi.org/10.1109/SEmotion.2019.00011

39. Ferreira, J., Dennehy, D., Babu, J., Conboy, K.: Winning of hearts and minds: integrating sentiment analytics into the analysis of contradictions. In: IFIP WG 6.11 Conference on e-Business, e-Services, and e-Society (I3E). Lecture Notes in Computer Science, vol. 11701, pp. 392–403. Springer, Berlin (2019). https://doi.org/10.1007/978-3-030-29374-1%5C_32
40. Frijda, N.H.: The Psychologists' Point of View, 3rd edn., pp. 68–87. Handbook of Emotions. The Guilford Press (2008)
41. Fu, B., Lin, J., Li, L., Faloutsos, C., Hong, J.I., Sadeh, N.M.: Why people hate your app: making sense of user feedback in a mobile app store. In: International Conference on Knowledge Discovery and Data Mining (KDD), pp. 1276–1284. ACM, New York (2013). https://doi.org/10.1145/2487575.2488202
42. Gachechiladze, D., Lanubile, F., Novielli, N., Serebrenik, A.: Anger and its direction in collaborative software development. In: International Conference on Software Engineering: New Ideas and Emerging Technologies Results Track ICSE-NIER, pp. 11–14. IEEE, Piscataway (2017). https://doi.org/10.1109/ICSE-NIER.2017.18
43. Garcia, D., Zanetti, M.S., Schweitzer, F.: The role of emotions in contributors activity: a case study on the Gentoo community. In: International Conference on Cloud and Green Computing, pp. 410–417 (2013)
44. Girardi, D., Lanubile, F., Novielli, N., Serebrenik, A.: Emotions and perceived productivity of software developers at the workplace. Trans. Softw. Eng. 1–53 (2022). https://doi.org/10.1109/TSE.2021.3087906
45. Girardi, D., Novielli, N., Fucci, D., Lanubile, F.: Recognizing developers' emotions while programming. In: International Conference on Software Engineering (ICSE), pp. 666–677. ACM, New York (2020). https://doi.org/10.1145/3377811.3380374
46. Graziotin, D., Wang, X., Abrahamsson, P.: Happy software developers solve problems better: psychological measurements in empirical software engineering. PeerJ **2**, e289 (2014). https://doi.org/10.7717/peerj.289
47. Gross, J.J., Barrett, L.F.: Emotion generation and emotion regulation: one or two depends on your point of view. Emot. Rev. **3**(1), 8–16 (2011)
48. Gunawardena, S.D., Devine, P., Beaumont, I., Garden, L.P., Murphy-Hill, E., Blincoe, K.: Destructive criticism in software code review impacts inclusion. International Conference on Human-Computer Interaction (CSCW), vol. 6 (2022). https://doi.org/10.1145/3555183
49. Guzman, E., Azócar, D., Li, Y.: Sentiment analysis of commit comments in GitHub: an empirical study. In: International Conference on Mining Software Repositories (MSR), pp. 352–355. ACM, New York (2014). https://doi.org/10.1145/2597073.2597118
50. Hall, J.A., Coats, E.J., LeBeau, L.S.: Nonverbal behavior and the vertical dimension of social relations: a meta-analysis. Psychol. Bull. **131**(6), 898–924 (2005)
51. Hatamian, M., Serna, J.M., Rannenberg, K.: Revealing the unrevealed: mining smartphone users privacy perception on app markets. Comput. Secur. **83**, 332–353 (2019). https://doi.org/10.1016/j.cose.2019.02.010
52. Herrmann, M., Klünder, J.: From textual to verbal communication: towards applying sentiment analysis to a software project meeting. In: International Requirements Engineering Conference Workshops (RE), pp. 371–376. IEEE, Piscataway (2021). https://doi.org/10.1109/REW53955.2021.00065
53. Hogan, P.C., Irish, B.J., Hogan, L.P. (eds.): The Routledge Companion to Literature and Emotion. Routledge, London (2022)
54. van den Hooff, B., Schouten, A.P., Simonovski, S.: What one feels and what one knows: the influence of emotions on attitudes and intentions towards knowledge sharing. J. Knowl. Manag. **16**(1), 148–158 (2012). https://doi.org/10.1108/13673271211198990
55. Huq, S.F., Sadiq, A.Z., Sakib, K.: Understanding the effect of developer sentiment on fix-inducing changes: an exploratory study on github pull requests. In: 26th Asia-Pacific Software Engineering Conference, APSEC 2019, Putrajaya, Malaysia, December 2-5, 2019, pp. 514–521. IEEE, Piscataway (2019). https://doi.org/10.1109/APSEC48747.2019.00075

56. Hwang, H.C., Matsumoto, D.: Measuring emotions in the face. In: Meiselman, H.L. (ed.) Emotion Measurement, pp. 125–144. Woodhead Publishing (2016). https://doi.org/10.1016/B978-0-08-100508-8.00006-0
57. Iacob, C., Faily, S., Harrison, R.: MARAM: tool support for mobile app review management. In: International Conference on Mobile Computing, Applications and Services (MobiCASE), pp. 42–50. ACM/ICST (2016). https://doi.org/10.4108/eai.30-11-2016.2266941
58. Imtiaz, N., Middleton, J., Girouard, P., Murphy-Hill, E.R.: Sentiment and politeness analysis tools on developer discussions are unreliable, but so are people. In: International Workshop on Emotion Awareness in Software Engineering (SEmotion), pp. 55–61. ACM, New York (2018). https://doi.org/10.1145/3194932.3194938
59. Islam, M.R., Ahmmed, M.K., Zibran, M.F.: Marvalous: machine learning based detection of emotions in the valence-arousal space in software engineering text. In: Symposium on Applied Computing (SAC), pp. 1786–1793. ACM (2019). https://doi.org/10.1145/3297280.3297455
60. Islam, M.R., Zibran, M.F.: Towards understanding and exploiting developers' emotional variations in software engineering. In: International Conference on Software Engineering Research, Management and Applications (SERA), pp. 185–192. IEEE Computer Society (2016). https://doi.org/10.1109/SERA.2016.7516145
61. Islam, M.R., Zibran, M.F.: DEVA: sensing emotions in the valence arousal space in software engineering text. In: Symposium on Applied Computing (SAC), pp. 1536–1543. ACM, New York (2018). https://doi.org/10.1145/3167132.3167296
62. Islam, M.R., Zibran, M.F.: SentiStrength-SE: exploiting domain specificity for improved sentiment analysis in software engineering text. J. Syst. Softw. **145**, 125–146 (2018). https://doi.org/10.1016/j.jss.2018.08.030
63. Jacob-Dazarola, R., Ortíz Nicolás, J.C., Cárdenas Bayona, L.: Behavioral measures of emotion. In: Meiselman, H.L. (ed.) Emotion Measurement, pp. 101–124. Woodhead Publishing (2016). https://doi.org/10.1016/B978-0-08-100508-8.00005-9
64. Jansen, S., Finkelstein, A., Brinkkemper, S.: A sense of community: a research agenda for software ecosystems. In: International Conference on Software Engineering, pp. 187–190 (2009). https://doi.org/10.1109/ICSE-COMPANION.2009.5070978
65. Jiarpakdee, J., Ihara, A., Matsumoto, K.i.: Understanding question quality through affective aspect in Q&A site. In: International Workshop on Emotion Awareness in Software Engineering (SEmotion), pp. 12–17. ACM, New York (2016). https://doi.org/10.1145/2897000.2897006
66. Jongeling, R., Sarkar, P., Datta, S., Serebrenik, A.: On negative results when using sentiment analysis tools for software engineering research. Empirical Softw. Eng. **22**(5), 2543–2584 (2017). https://doi.org/10.1007/s10664-016-9493-x
67. Jurado, F., Rodríguez Marín, P.: Sentiment analysis in monitoring software development processes: an exploratory case study on GitHub's project issues. J. Syst. Softw. **104**, 82–89 (2015). https://doi.org/10.1016/j.jss.2015.02.055
68. Juslin, P.N., Sloboda, J.A. (eds.): Handbook of Music and Emotion: Theory, Research, Applications. Series in Affective Science. Oxford University Press (2010)
69. Kalliamvakou, E., Gousios, G., Blincoe, K., Singer, L., German, D.M., Damian, D.: The promises and perils of mining GitHub. In: Working Conference on Mining Software Repositories (MSR), MSR 2014, pp. 92–101. ACM, New York (2014). https://doi.org/10.1145/2597073.2597074
70. Kaur, A., Singh, A.P., Dhillon, G.S., Bisht, D.: Emotion mining and sentiment analysis in software engineering domain. In: International Conference on Electronics, Communication and Aerospace Technology (ICECA), pp. 1170–1173 (2018). https://doi.org/10.1109/ICECA.2018.8474619
71. Kemper, T.D.: A Social Interactional Theory of Emotions. Wiley, New York (1978)
72. Khan, K.M., Saleh, M.: Understanding the impact of emotions on the quality of software artifacts. IEEE Access **9**, 110194–110208 (2021). https://doi.org/10.1109/ACCESS.2021.3102663

73. Lanovaz, M.J., Adams, B.: Comparing the communication tone and responses of users and developers in two R mailing lists: measuring positive and negative emails. IEEE Software **36**(5), 46–50 (2019). https://doi.org/10.1109/MS.2019.2922949
74. de Lima Fontão, A., Ekwoge, O.M., dos Santos, R.P., Dias-Neto, A.C.: Facing up the primary emotions in mobile software ecosystems from developer experience. In: Workshop on Social, Human, and Economic Aspects of Software (WASHES), pp. 5–11. ACM, New York (2017). https://doi.org/10.1145/3098322.3098325
75. Lin, B., Cassee, N., Serebrenik, A., Bavota, G., Novielli, N., Lanza, M.: Opinion mining for software development: a systematic literature review. Trans. Softw. Eng. Methodol. **31**(3), 1–41 (2022)
76. Lin, B., Zampetti, F., Bavota, G., Di Penta, M., Lanza, M., Oliveto, R.: Sentiment analysis for software engineering: how far can we go? In: International Conference on Software Engineering (ICSE), pp. 94–104. ACM, New York (2018). https://doi.org/10.1145/3180155.3180195
77. Lungu, M.: Towards reverse engineering software ecosystems. In: International Conference on Software Maintenance (ICSM), pp. 428–431. IEEE, Piscataway (2008). https://doi.org/10.1109/ICSM.2008.4658096
78. Manikas, K., Hansen, K.M.: Software ecosystems: a systematic literature review. J. Syst. Softw. **86**(5), 1294–1306 (2013). https://doi.org/10.1016/j.jss.2012.12.026
79. Mankad, S., Hu, S., Gopal, A.: Single stage prediction with embedded topic modeling of online reviews for mobile app management. Ann. Appl. Stat. **12**(4), 2279–2311 (2018)
80. Mäntylä, M.V., Novielli, N., Lanubile, F., Claes, M., Kuutila, M.: Bootstrapping a lexicon for emotional arousal in software engineering. In: International Conference on Mining Software Repositories (MSR), pp. 198–202. IEEE, Piscataway (2017). https://doi.org/10.1109/MSR.2017.47
81. Meiselman, H.L.: Emotion Measurement, 1st edn. Woodhead Publishing (2016)
82. Messerschmitt, D.G., Szyperski, C.: Software ecosystem: understanding an indispensable technology and industry. MIT Press, Cambridge (2003)
83. Mitleton-Kelly, E.: Ten Principles of Complexity and Enabling Infrastructures, pp. 23–50. Pergamon (2003)
84. Mohammad, S.M.: Sentiment analysis: detecting valence, emotions, and other affectual states from text. In: Meiselman, H.L. (ed.) Emotion Measurement, pp. 201–237. Woodhead Publishing (2016). https://doi.org/10.1016/B978-0-08-100508-8.00009-6
85. Mondal, A.K., Rahman, M.M., Roy, C.K.: Embedded emotion-based classification of stack overflow questions towards the question quality prediction. In: International Conference on Software Engineering and Knowledge Engineering (SEKE), pp. 521–526. KSI Research Inc. and Knowledge Systems Institute Graduate School (2016). https://doi.org/10.18293/SEKE2016-146
86. Müller, S.C., Fritz, T.: Stuck and frustrated or in flow and happy: sensing developers' emotions and progress. In: International Conference on Software Engineering, pp. 688–699 (2015)
87. Muñoz, S., Araque, O., Llamas, A.F., Iglesias, C.A.: A cognitive agent for mining bugs reports, feature suggestions and sentiment in a mobile application store. In: International Conference on Big Data Innovations and Applications, pp. 17–24. IEEE, Piscataway (2018). https://doi.org/10.1109/Innovate-Data.2018.00010
88. Murgia, A., Ortu, M., Tourani, P., Adams, B., Demeyer, S.: An exploratory qualitative and quantitative analysis of emotions in issue report comments of open source systems. Empirical Softw. Eng. **23**(1), 521–564 (2018). https://doi.org/10.1007/s10664-017-9526-0
89. Murgia, A., Tourani, P., Adams, B., Ortu, M.: Do developers feel emotions? An exploratory analysis of emotions in software artifacts. In: Working Conference on Mining Software Repositories (MSR), pp. 262–271. ACM, New York (2014)
90. Nasehi, S.M., Sillito, J., Maurer, F., Burns, C.: What makes a good code example? A study of programming Q&A in StackOverflow. In: International Conference on Software Maintenance (ICSM), pp. 25–34. IEEE, Piscataway (2012). https://doi.org/10.1109/ICSM.2012.6405249

91. Niedenthal, P.M., Ric, F.: Psychology of Emotion. Psychology Press, New York (2017)
92. Nizamani, Z.A., Liu, H., Chen, D.M., Niu, Z.: Automatic approval prediction for software enhancement requests. Autom. Softw. Eng. **25**(2), 347–381 (2018). https://doi.org/10.1007/s10515-017-0229-y
93. Norman, G.J., Necka, E., Berntson, G.G.: The psychophysiology of emotions. In: Meiselman, H.L. (ed.) Emotion Measurement, pp. 83–98. Woodhead Publishing (2016). https://doi.org/10.1016/B978-0-08-100508-8.00004-7
94. Novielli, N., Calefato, F., Dongiovanni, D., Girardi, D., Lanubile, F.: Can we use SE-specific sentiment analysis tools in a cross-platform setting? In: International Conference on Mining Software Repositories (MSR), pp. 158–168. ACM, New York (2020). https://doi.org/10.1145/3379597.3387446
95. Novielli, N., Calefato, F., Lanubile, F.: Towards discovering the role of emotions in Stack Overflow. In: International Workshop on Social Software Engineering (SSE), pp. 33–36. ACM, New York (2014). https://doi.org/10.1145/2661685.2661689
96. Novielli, N., Calefato, F., Lanubile, F.: The challenges of sentiment detection in the social programmer ecosystem. In: International Workshop on Social Software Engineering (SSE), pp. 33–40. ACM, New York (2015). https://doi.org/10.1145/2804381.2804387
97. Novielli, N., Calefato, F., Lanubile, F.: A gold standard for emotion annotation in Stack Overflow. In: International Conference on Mining Software Repositories (MSR), pp. 14–17. ACM, New York (2018). https://doi.org/10.1145/3196398.3196453
98. Novielli, N., Calefato, F., Lanubile, F.: Love, joy, anger, sadness, fear, and surprise: SE needs special kinds of AI: a case study on text mining and SE. IEEE Softw. **37**(3), 86–91 (2020). https://doi.org/10.1109/MS.2020.2968557
99. Novielli, N., Calefato, F., Lanubile, F., Serebrenik, A.: Assessment of off-the-shelf SE-specific sentiment analysis tools: an extended replication study. Empirical Softw. Eng. **26**(4), 77 (2021). https://doi.org/10.1007/s10664-021-09960-w
100. Novielli, N., Grassi, D., Lanubile, F., Serebrenik, A.: Sensor-based emotion recognition in software development: facial expressions as gold standard. In: International Conference on Affective Computing and Intelligent Interaction (ACII) (2022). https://doi.org/10.1109/ACII55700.2022.9953808
101. Nugroho, Y.S., Islam, S., Nakasai, K., Rehman, I., Hata, H., Kula, R.G., Nagappan, M., Matsumoto, K.: How are project-specific forums utilized? A study of participation, content, and sentiment in the Eclipse ecosystem. Empirical Softw. Eng. **26**(6), 132 (2021). https://doi.org/10.1007/s10664-021-10032-2
102. Obaidi, M., Klünder, J.: Development and application of sentiment analysis tools in software engineering: a systematic literature review. In: Evaluation and Assessment in Software Engineering (EASE), pp. 80—89. ACM, New York (2021). https://doi.org/10.1145/3463274.3463328
103. Obaidi, M., Nagel, L., Specht, A., Klünder, J.: Sentiment analysis tools in software engineering: a systematic mapping study. Inform. Softw. Technol. **151**, 107018 (2022). https://doi.org/10.1016/j.infsof.2022.107018
104. Ortu, M., Adams, B., Destefanis, G., Tourani, P., Marchesi, M., Tonelli, R.: Are bullies more productive? Empirical study of affectiveness vs. issue fixing time. In: Working Conference on Mining Software Repositories (MSR), pp. 303–313. IEEE, Piscataway (2015). https://doi.org/10.1109/MSR.2015.35
105. Ortu, M., Hall, T., Marchesi, M., Tonelli, R., Bowes, D., Destefanis, G.: Mining communication patterns in software development: a GitHub analysis. In: International Conference on Predictive Models and Data Analytics in Software Engineering (PROMISE), pp. 70–79. ACM, New York (2018). https://doi.org/10.1145/3273934.3273943
106. Ortu, M., Marchesi, M., Tonelli, R.: Empirical analysis of affect of merged issues on GitHub. In: International Workshop on Emotion Awareness in Software Engineering (SEmotion), pp. 46–48. IEEE / ACM (2019). https://doi.org/10.1109/SEmotion.2019.00017
107. Ortu, M., Murgia, A., Destefanis, G., Tourani, P., Tonelli, R., Marchesi, M., Adams, B.: The emotional side of software developers in JIRA. In: International Conference on Mining

Software Repositories (MSR), pp. 480–483. ACM, New York (2016). https://doi.org/10.1145/2901739.2903505
108. Parrott, W.G.: Emotions in Social Psychology: Essential Readings. Psychology Press (2001)
109. Plamper, J.: The History of Emotions. Oxford University Press, Oxford (2012)
110. Pletea, D., Vasilescu, B., Serebrenik, A.: Security and emotion: sentiment analysis of security discussions on GitHub. In: Working Conference on Mining Software Repositories (MSR), pp. 348–351. ACM, New York (2014). https://doi.org/10.1145/2597073.2597117
111. Plutchik, R.: Outlines of a new theory of emotion. Trans. N.Y. Acad. Sci. **20**(5), 394–403 (1958)
112. Powell, P.B.: Planning for software validation, verification, and testing. Tech. Rep. 98, US Department of Commerce, National Bureau of Standards (1982)
113. Rigby, P.C., Hassan, A.E.: What can OSS mailing lists tell us? A preliminary psychometric text analysis of the Apache developer mailing list. In: International Workshop on Mining Software Repositories (MSR), pp. 23–23 (2007). https://doi.org/10.1109/MSR.2007.35
114. Russell, J.: Culture and the categorization of emotions. Psychol. Bull. **110**(3), 426–450 (1991)
115. Sánchez-Gordón, M., Colomo-Palacios, R.: Taking the emotional pulse of software engineering: a systematic literature review of empirical studies. Inform. Softw. Technol. **115**, 23–43 (2019). https://doi.org/10.1016/j.infsof.2019.08.002
116. Santos, M.F., Caetano, J.A., Oliveira, J., Neto, H.T.M.: Analyzing the impact of feedback in GitHub on the software developer's mood. In: International Conference on Software Engineering and Knowledge Engineering (SEKE), pp. 445–444 (2018). https://doi.org/10.18293/SEKE2018-153
117. Schneider, K., Klünder, J., Kortum, F., Handke, L., Straube, J., Kauffeld, S.: Positive affect through interactions in meetings: the role of proactive and supportive statements. J. Syst. Softw. **143**, 59–70 (2018). https://doi.org/10.1016/j.jss.2018.05.001
118. SentiSE. https://github.com/amiangshu/SentiSE
119. Shaver, P., Schwartz, J., Kirson, D., O'connor, C.: Emotion knowledge: further exploration of a prototype approach. J. Pers. Soc. Psychol. **52**(6), 1061–1066 (1987)
120. Singh, N., Singh, P.: How do code refactoring activities impact software developers' sentiments? An empirical investigation into GitHub commits. In: Asia-Pacific Software Engineering Conference (APSEC), pp. 648–653. IEEE, Piscataway (2017). https://doi.org/10.1109/APSEC.2017.79
121. Sinha, V., Lazar, A., Sharif, B.: Analyzing developer sentiment in commit logs. In: International Conference on Mining Software Repositories (MSR), pp. 520–523. ACM (2016). https://doi.org/10.1145/2901739.2903501
122. Solomon, R.C.: The Philosophy of Emotions, 3rd edn., pp. 3–16. Handbook of Emotions. The Guilford Press (2008)
123. Souza, R.R.G., Silva, B.: Sentiment analysis of Travis CI builds. In: International Conference on Mining Software Repositories (MSR), pp. 459–462. IEEE, Piscataway (2017). https://doi.org/10.1109/MSR.2017.27
124. Storey, M.A.: The evolution of the social programmer. In: Working Conference on Mining Software Repositories (MSR). IEEE (2012)
125. Sun, K., Shi, X., Gao, H., Kuang, H., Ma, X., Rong, G., Shao, D., Zhao, Z., Zhang, H.: Incorporating pre-trained transformer models into TextCNN for sentiment analysis on software engineering texts. In: Asia-Pacific Symposium on Internetware, pp. 127–136. ACM, New York (2022). https://doi.org/10.1145/3545258.3545273
126. Thelwall, M., Buckley, K., Paltoglou, G.: Sentiment strength detection for the social web. J. Am. Soc. Inf. Sci. Technol. **63**(1), 163–173 (2012)
127. Thelwall, M., Buckley, K., Paltoglou, G., Cai, D., Kappas, A.: Sentiment in short strength detection informal text. J. Am. Soc. Inf. Sci. Technol. **61**(12), 2544–2558 (2010)
128. Thung, F., Bissyandé, T.F., Lo, D., Jiang, L.: Network structure of social coding in github. In: European Conference on Software Maintenance and Reengineering (CSMR), pp. 323–326 (2013). https://doi.org/10.1109/CSMR.2013.41

129. Tourani, P., Adams, B.: The impact of human discussions on just-in-time quality assurance: an empirical study on OpenStack and Eclipse. In: International Conference on Software Analysis, Evolution, and Reengineering (SANER), pp. 189–200. IEEE, Piscataway (2016). https://doi.org/10.1109/SANER.2016.113
130. Tourani, P., Jiang, Y., Adams, B.: Monitoring sentiment in open source mailing lists: exploratory study on the Apache ecosystem. In: International Conference on Computer Science and Software Engineering (CASCON), pp. 34–44. IBM / ACM (2014)
131. Uddin, G., Guéhénuc, Y.G., Khomh, F., Roy, C.K.: An empirical study of the effectiveness of an ensemble of stand-alone sentiment detection tools for software engineering datasets. Trans. Softw. Eng. Methodol. **31**(3) (2022). https://doi.org/10.1145/3491211
132. Uddin, G., Khomh, F.: Automatic mining of opinions expressed about APIs in Stack Overflow. Trans. Softw. Eng. 1 (2019). https://doi.org/10.1109/TSE.2019.2900245
133. Umer, Q., Liu, H., Sultan, Y.: Sentiment based approval prediction for enhancement reports. J. Syst. Softw. **155**, 57–69 (2019). https://doi.org/10.1016/j.jss.2019.05.026
134. Vasilescu, B., Posnett, D., Ray, B., van den Brand, M.G., Serebrenik, A., Devanbu, P., Filkov, V.: Gender and tenure diversity in GitHub teams. In: Conference on Human Factors in Computing Systems (CHI), pp. 3789–3798. ACM, New York (2015). https://doi.org/10.1145/2702123.2702549
135. Vrzakova, H., Begel, A., Mehtätalo, L., Bednarik, R.: Affect recognition in code review: an in-situ biometric study of reviewer's affect. J. Syst. Softw. **159** (2020). https://doi.org/10.1016/j.jss.2019.110434
136. Wälde, K., Moors, A.: Current emotion research in economics. Emot. Rev. **9**(3), 271–278 (2017). https://doi.org/10.1177/1754073916665470
137. Watson, D., Clark, L.A., Tellegen, A.: Development and validation of brief measures of positive and negative affect: the PANAS scales. J. Pers. Soc. Psychol. **54**(6), 1063–1070 (1988)
138. Werder, K.: The evolution of emotional displays in open source software development teams: an individual growth curve analysis. In: International Workshop on Emotion Awareness in Software Engineering (SEmotion), pp. 1–6. ACM, New York (2018). https://doi.org/10.1145/3194932.3194934
139. Werder, K., Brinkkemper, S.: MEME: toward a method for emotions extraction from GitHub. In: International Workshop on Emotion Awareness in Software Engineering (SEmotion), pp. 20–24. ACM, New York (2018). https://doi.org/10.1145/3194932.3194941
140. Williams, G., Mahmoud, A.: Analyzing, classifying, and interpreting emotions in software users' tweets. In: International Workshop on Emotion Awareness in Software Engineering (SEmotion), pp. 2–7 (2017). https://doi.org/10.1109/SEmotion.2017.1
141. Wurzel Gonçalves, P., Çalikli, G., Bacchelli, A.: Interpersonal conflicts during code review: developers' experience and practices. Proc. ACM Hum.-Comput. Interact. **6**(CSCW1) (2022). https://doi.org/10.1145/3512945
142. Yang, B., Wei, X., Liu, C.: Sentiments analysis in GitHub repositories: an empirical study. In: Asia-Pacific Software Engineering Conference Workshops (APSEC Workshops), pp. 84–89. IEEE, Piscataway (2017). https://doi.org/10.1109/APSECW.2017.13
143. Zhang, T., Xu, B., Thung, F., Haryono, S.A., Lo, D., Jiang, L.: Sentiment analysis for software engineering: how far can pre-trained transformer models go? In: International Conference on Software Maintenance and Evolution (ICSME), pp. 70–80. IEEE, Piscataway (2020). https://doi.org/10.1109/ICSME46990.2020.00017