Code of Conduct in Open Source Projects

Parastou Tourani   Bram Adams
Polytechnique Montréal, Canada
{parastou.tourani, bram.adams}@polymtl.ca

Alexander Serebrenik
Eindhoven University of Technology, The Netherlands
a.serebrenik@tue.nl

Abstract—Open source projects rely on collaboration of members from all around the world using web technologies like GitHub and Gerrit. This mixture of people with a wide range of backgrounds including minorities like women, ethnic minorities, and people with disabilities may increase the risk of offensive and destroying behaviours in the community, potentially leading affected project members to leave towards a more welcoming and friendly environment. To counter these effects, open source projects increasingly are turning to codes of conduct, in an attempt to promote their expectations and standards of ethical behaviour. In this first of its kind empirical study of codes of conduct in open source software projects, we investigated the role, scope and influence of codes of conduct through a mixture of quantitative and qualitative analysis, supported by interviews with practitioners. We found that the top codes of conduct are adopted by hundreds to thousands of projects, while all of them share 5 common dimensions.

I. INTRODUCTION

Since modern software engineering technologies like version control systems, review environments, issue repositories, mailing lists, chat systems and wikis have enabled new forms of online collaboration, a large crowd of project members with diverse experiences and interests are able to collaborate in open source projects to produce large, complex and successful systems [14]. Indeed, according to the Open Source Initiative, "In order to get the maximum benefit from the [open source] process, the maximum diversity of persons and groups should be equally eligible to contribute to open sources" [3]. Furthermore, Vasilescu *et al.* have shown that gender diversity is beneficial for productivity of GitHub teams [26].

However, a mixture of people with different cultures, personalities and interests may increase the risk of offensive behaviours happening. For example, on June 18 2015, one of the core maintainers of OpalRB, a Ruby-to-Javascript transpiler, left transphobic comments on Twitter[1]. The resulting pile-on of responses ranged widely from those discussing that publicly stated opinions of a member have no bearing on the community, to those who expected consequences for the offensive person (*e.g.,* excluding him from the community) or those thinking that as long as the contributed code was good, working with the offensive member was acceptable[2]. Similarly, Vasilescu *et al.* report gender-related incidents as cause for leaving a project [25].

Therefore, it seems essential for open source communities to protect their members from these kinds of unacceptable, destroying behaviours and provide a welcoming, safe, friendly, and inclusive environment in which people can collaborate effectively towards presenting successful products. Thus far, the most common means for projects such as OpalRB[3] to achieve such an environment, is the concept of "code of conduct". Such a code of conduct basically establishes ground rules for communications between participants, outlines enforcement mechanisms for violations and tries to codify the spirit of a community, such that anyone can contribute comfortably regardless of *e.g.,* gender, ethnicity, or sexual orientation.

However, some communities find codes of conduct repressive and a threat for open source communities. One common argument is that participants in open source communities are mature enough to deal with debates and differences, and hence it should be obvious for members how to behave. Some strong opponents explicitly picked a "No Code of Conduct" for their communities[4]. Furthermore, even in projects that do adopt a code of conduct, the adoption process suscitates substantial discussion because of doubts whether codes of conducts work and disagreement about what should go into them and the exact wording to use[5].

Since there is no empirical evidence regarding the status, nature of, and procedure for establishing codes of conduct in open source projects, the primary purpose of this paper is to empirically examine codes of conduct in open source projects, identifying the procedures followed in their implementation and monitoring, as well as understanding its scope and impact in open source communities. To this end, we address the following research questions:

RQ1) How common are codes of conduct in open source projects?

Eleven codes of conduct are commonly used in open source projects, with seven of them ranging from 500 up to several thousands of adopting projects.

RQ2) What do major codes of conduct stipulate?

Codes of conduct outline the expectations and values of an OSS community against its members' behaviours to create a friendly and inclusive community, while violations have consequences. These codes of conduct target all collaboration spaces of the community, either online or offline, but the scope can be broader as well.

---

[1] https://twitter.com/krainboltgreene/status/611569515315507200
[2] https://github.com/opal/opal/issues/941

[3] https://github.com/opal/opal/issues/942
[4] https://github.com/domgetter/NCoC
[5] https://github.com/opal/opal/issues/942

24

RQ3) How are codes of conduct used in open source projects?

The concerns, needs and history of a project's community play an important role in the design of its code of conduct. However, similarities among communities may lead to reuse of existing codes of conduct.

In the remainder of this paper, we first describe the necessary background notions and related work (section II). Then, we investigate the three research questions, present its approach and results (section III) and discuss our findings (section VI). After mentioning the threats to validity (section VII), we present our conclusions (section VIII).

II. Background and Related Work

A. Open Source Projects and Diversity

Eric Raymond [22] summarized the differences and properties of open source software (OSS) in comparison to other types of software development. More specifically, the open access model of OSS development encourages participants with maximum ability and skill, specific expertise and minimum restrictions in geographical issues to participate. This leads to a high degree of diversity amongst project members in terms of gender, ethnicity, religion and age, which is likely to be an influential factor impacting OSS project success [7]. Bazile-Jones et al. [5] stated that managing and valuing diversity in workplaces, as one intellectual asset, can bring long-term wealth. Valuing diversity refers to recognizing individual differences and dissimilarities, and respecting them by considering everyone's needs and expectations.

Sherae et al. expanded on the theoretical understanding of diversity and its implications in OSS projects [9]. They defined three types of diversity, i.e., disparity (based on contribution reputation), separation (based on culture) and variety (reputation) diversity. They discussed how each of these types plays a role in the project success from a community engagement point of view or market success. Vasilescu et al. in [26], discussed about the various aspects of diversity in open source projects. Using GitHub, the largest public code repository for OSS projects for data extraction, they showed that more diversity in gender and tenure is associated with higher productivity and turnover. In other words, diverse teams consisting of both men and women, with varying degrees of experience, show better performance.

B. Code of Conduct

According to Wikipedia[6], a code of conduct generally is a set of rules articulating standard behaviour and responsibilities for an individual, party or group. It is commonly written for employees of a company in order to 1) protect the business and 2) inform the employees of the company's expectations. The International Federation of Accountants [1], IFAC, provided a more precise definition. According to them, a code of conduct comprises principles, values, standards and rules that act as

guidelines that have the overall benefit of the stakeholders in mind and at the same time respect the rights of members.

Recently, codes of conduct have been adopted in OSS communities, as open and welcoming communities, to deal with diversity problems. One of the oldest codes of conduct in OSS communities is the one designed for the Ubuntu community more than 10 years ago, with other codes popping up over the years and existing codes seeing updated versions. Some OSS codes of conduct such as the Ubuntu one are even being used by other OSS projects. We discuss popular OSS codes of conducts and their prevalence in section III.

It is worth mentioning that a code of conduct is different from a code of ethics. A code of ethics is adopted to clarify for members of a company the meaning of "right" and "wrong" based on the business of a company, and therefore is applied to make decisions [1] about the members' actions and manners. A code of conduct on the other hand is confined to actions or behaviours of members and is usually intended only for them (instead of for stakeholders).

C. Workplace Harassment

Most codes of conduct aim to protect members from harassment, thus it seems important to understand what workplace harassment means. According to the Oxford dictionary, harassment is "Aggressive pressure or intimidation". As such, workplace harassment is any offensive, belittling or threatening behaviour toward an individual worker or group of workers. It results in an unpleasant, humiliating or intimidating environment employees feel uncomfortable in and consequently damages effective work and productivity of employees[7].

In OSS communities, just like any other workforce, workplace harassment may include, but is not limited to, online or offline harassing behaviours such as verbal comments, sexual jokes or insults, sexual images in public spaces, intimidation, stalking, inappropriate physical contact, bullying. Evidence for such behaviour in OSS communities has been reported, *e.g.,* by Powell et al. [21] and Vasilescu et al. [25].

III. RQ1. How common are codes of conduct in open source projects?

Motivation This first research question is concerned with finding the frequently used codes of conduct as well as with determining their prevalence in open source projects. In addition to showing the relevance of this paper and motivating the need for further research on codes of conduct, a list of popular codes of conduct, with the high number of usages, will be used to answer RQ2, *i.e.*, to understand the common ingredients of codes of conduct.

Approach

We used a two-pronged approach. First, we used a list of seven codes of conduct, consisting of the Open Code of Conduct of the ToDo group[8], a well-known organization in the field, as well as six other codes claimed by the ToDo

---

[6]https://en.wikipedia.org/wiki/Code_of_conduct

[7]https://web.archive.org/web/20120328034350/http://apsc.gov.au/publications01/harassment.htm

[8]http://todogroup.org/opencodeofconduct/

group to be the giants on whose shoulders Open Code of Conduct stands. We performed a brute-force search on GitHub with the names of these codes of conduct to determine the order of magnitude of their popularity. GitHub is a popular code repository site used by millions of popular and active open source projects [8]. Since this approach only aims to provide an order of magnitude for the prevalence of the seven suggested codes of conduct, and given the number of search results, we did not eliminate duplicate search results or false positives. Furthermore, searching by name ignores projects that just mention the URL of the code of conduct used.

While the above limitations are acceptable to obtain an order of magnitude, our second approach instead uses a second group of data sets and involves manual analysis to get an accurate set of mappings between projects and the codes of conduct that they use. To do this, we used the principles of systematic reviews proposed by Kitchenham [17]: our population is "Open Source communities", our intervention is "Code of Conduct", and two electronic databases are used, *i.e.* Github[9] and Google. The latter allows to find projects not hosted on GitHub or whose code of conduct is not stored in their version control system. We used the same queries for both GitHub and Google: 1) *"code of conduct" "open source"* and 2) *"code of conduct" software community*. Quotation marks were explicitly added to reduce the number of false positive hits. Furthermore, we added "open source" and software community to filter out codes of conduct for conferences and other events as opposed to codes of conduct for software projects, which are the focus of this paper.

Although the Google search engine initially claimed that 57,000 and 2,110,000 hits were found, respectively, for the two search queries mentioned above, in reality, after Google's own duplicate filtering, the number of hits turned out to be around 500 and 700. After further manual removal of additional duplicates and incomplete matches of the query, the results yielded only 395 and 324 actual hits, respectively. Finally, after filtering out search results related to schools and conferences, we obtained 306 and 241 hits, respectively. Manual analysis of the entire collection of $547 = 306 + 241$ links yielded 108 unique open source communities using a code of conduct, including well-known ecosystems such as Apache, Eclipse, Openstack, Debian, and Scala.

For GitHub, the two queries yielded 17,498 and 2,417 textual files in markdown format (mostly README files), respectively. The top 200 hits of the first query (based on GitHub's "best match"), and top 40 hits for the second query (idem) were manually checked to see if they belong to codes of conduct of open source projects. Out of these 240 links, 184 corresponded to actual open source projects using a code of conduct. The false positive links include duplicate links across different versions of a project, or irrelevant links of non-software communities (e.g., Software Carpentry[10]).

Finally, as an additional data source, we also studied the top

TABLE II: Popularity of Codes of Conduct Obtained via Google (second approach). Codes of Conduct from Table I have been emphasized.

Code of Conduct	#Google Results (out of 108)
Ubuntu	20
Contributor Covenant	20
Django	13
Python	9
Citizen	9
Open Code of Conduct	7
Geek Feminism	7
Mozilla	6
Twitter	5
Rust	5
Ada Initiative	4
KDE	4
SpeakUP	3
Apache	2
Thoughtbot	3
Openstack	3
Debian	2
Puppet	2
PyCon	2

GitHub projects to analyze whether they use a code of conduct, and if so, which one. We used the number of watchers of a GitHub project as a measure of project popularity, similar to other work [27]. We studied the 150 most watched projects and found that 52 of them have a code of conduct, including such projects as Linux, JQuery, Angular and Swift.

In the obtained search results in Google and GitHub, we then compared the number of occurrences of each identified code of conduct to the order of magnitude numbers for the initial seven codes of conduct, and also tried to identify any missing major codes of conduct. However, since some projects may focus more on GitHub, while others do not, we considered each data source separately. In addition, we noticed that many Google results correspond to ecosystems instead of to individual projects, while the GitHub results mostly correspond to individual projects. As such, the Google ecosystem results actually imply that a larger number of existing projects are using the same code of conduct (Table II).

Results

The codes of conduct from the ToDo group's list, i.e., *Ubuntu*, *Contributor Covenant*, *Django*, *Python*, *Citizen*, *Open Code of Conduct*, and *Geek Feminism*) have been frequently used by more than 500 OSS projects. Table I shows the number of markdown files obtained when searching for the names of the seven codes of conduct proposed by the ToDo Group organization. Even though these numbers still contain false positives, we found that these coarse estimations are confirmed by the Google query results of table II, in which these codes of conduct seem to be the top seven most common codes used by open source communities.

Contributor Covenant [12] is a well-known code of conduct template designed to be used by open source projects. It has four versions so far and, based on its site[11], more than 40,000

TABLE I: ORDER OF MAGNITUDE FOR CODES OF CONDUCT OBTAINED FROM GITHUB (FIRST APPROACH).

Code of Conduct	Number of Hits on GitHub	Examples
Contributor Covenant	43,681	Molajo/Molajo, trevororeilly/dotfiles, SecComm/Crayon, yuluyi/Isomorphic-React-Seed, tweetstockr
Open Code of Conduct	2,167	wildland/cli-tools, KineticCafe/code-of-conduct, Tacklr/CacheManager, spotify/ios-ci, PearlCast/PearlCast
Python	2,025	PyDiff/PyDiff.github.io, 18F, brettcannon/oplop, link39/205-pi, roadcap/homebrew,sfdevs/sdcodecamp
Citizen	1,253	npr/npr-one-api-js-sdk, cworth-gh/stony, gulpjs/gulp, lkodai/Design-LK, ojs/ojs, ctdk/goiardi
Ubuntu	1,180	goodeggs/format-location, Star2Billing/cdr-stats-docs, garyjs/Newfiesautodialer,Alamofire/Foundation, Trustroots/trustroots
Django	1,054	jrief/django-angular, DBCboots, Pythonkc, Calagator, ordergroove/check_mariadb_slaves
Geek Feminism	544	nzruby, brave/chromium, crosswalk-project/chromium-crosswalk, javascripthers/javascripthers.github.io, openSNP/snpr

TABLE III: POPULARITY OF CODES OF CONDUCT OBTAINED VIA GITHUB (SECOND APPROACH). CODES OF CONDUCT FROM TABLE I HAVE BEEN EMPHASIZED.

Code of Conduct	#GitHub Results (out of 184)	#top GitHub Results (out of 52)
Ubuntu	42	3
Twitter	42	3
Django	38	2
SpeakUP	36	0
Apache	35	0
Python	32	0
Contributor Covenant	15	19
Mozilla	11	3
Open Code of Conduct	7	2
Thoughtbot	5	0
Citizen	4	2
Rust	3	4
Ada Initiative	3	0
Geek Feminism	0	1

open source projects already have adopted it, which confirms our result in table I. The *Open Code of Conduct* is developed by the ToDo group, which is a collection of open source company leaders[12] with a common interest, such as Facebook, Github, Google, and Yahoo. They collaborate to share their experiences and knowledge including practices to help Open Source projects improve and evolve. *Citizen* was created by the Stumptown Syndicate[13] for a conference, however it has been applied in open source communities as well. *Geek Feminism* refers to women who are working in geeky activities related to high technology. Their designed code of conduct is used by broad communities including FLOSS communities (Free Libre and Open Source Software). The *Django* Software foundation[14] is an open source foundation developing the

Django Web Framework. The *Ada* Initiative [15], a charitable organization supporting women in open technology and culture, is in favor of Django code of conduct, since it is very detailed and specific. Finally, *Ubuntu* and *Python* are highly popular open source ecosystems.

During our systematic review of codes of conduct, we noticed that some codes of conduct, such as *Contributor Covenant* [12] and *Open code of conduct* [2] are even used by major commercial software companies like Microsoft, Facebook and Yahoo for their open source projects.

Apart from the seven initial codes of conduct, the *Twitter*, *SpeakUP*, *Apache* and *Mozilla* codes of conduct are popular on GitHub. Speak Up![16] is a technical community to empower and educate speakers by connecting them with mentors and proper resources. Diversity is their strength and people should follow their code of conduct[17] when participating in the community. Twitter[18], Mozilla[19] and Apache[20] are popular open source projects with their own code of conduct. The Apache and Twitter codes are inspired by other codes of conduct like *Open Code of Conduct*, *Django* and *Python*, and in turn have been adopted by other projects as well (e.g., busApp[21] and stepmaniasite[22]).

> Eleven codes of conduct are commonly used in open source projects, with seven of them ranging from 500 up to several thousands of adopting projects.

---

[12]http://todogroup.org/members/
[13]http://citizencodeofconduct.org
[14]https://www.djangoproject.com/foundation/

[15]https://adainitiative.org
[16]http://web.archive.org/web/20141105204955/http://speakup.io/
[17]http://speakup.io/coc.html
[18]https://twitter.com/?lang=en
[19]https://www.mozilla.org/en\discretionary{-}{}{}US/
[20]https://www.apache.org/foundation/
[21]https://github.com/kalimbamgb/busApp
[22]https://github.com/stepmania

Motivation In the previous research question, we identified eleven common codes of conduct used either directly or as the basis for a custom code of conduct based on the Google search. Here we are interested in understanding the content of such codes of conduct, *i.e.*, the basic elements, measures or other guidelines, as well as the way in which these are written up. In particular, we are interested in understanding the kinds of behaviour addressed by codes of conduct, potential measures taken, but also whether codes of conduct are written up as suggestions or as stringent rules (i.e., the style of writing).

Approach To answer this research question, we manually studied the seven initial codes of conduct of RQ1, as these obtained high popularity scores in table I and table II. Studying the other four codes of conduct is left for future work. For each of them, we read and identified the underlying components using an *open coding process* [10], then looked for similarities and differences. Whereas most of the seven codes are independent, we found considerable overlap in terms of their major components and ingredients. In particular, each code of conduct included the following five components in one form or the other:

- Purpose: the rationale for the code of conduct, typically the desire to obtain a certain kind of environment for project members to work and collaborate in.
- Honorable behaviour: behaviour that is valuable for and accepted by the community.
- Unacceptable Behaviour: negative behaviour that should be avoided.
- Enforcement: mechanisms for reporting and punishing violations of the code of conduct.
- Scope: the online and offline spaces where the code of conduct applies, for example only in the mailing list versus in any online discussion forum.

Below, we detail each of these components in more detail, while Table IV summarizes how the above components and other dimensions apply to each of the seven codes of conduct. Note that for codes of conduct with more than one version, we referred to the latest one.

Results

Purpose: All codes stress the desire of diversity and of a welcoming community, while some explicitly list the desired diversity attributes (*e.g.*, gender, sexual orientation, and disability). Since OSS project communities consist of professionals and volunteers from around the world, the seven codes of conduct all promote an inclusive and safe environment to everyone, for the sake of sustainability of the community. For example, the *Contributor Covenant* refers especially to personal characteristics like gender, age, size, body, religion, ethnicity, and sexual orientation. *Python* and *Ubuntu* just generally refer to diverse groups, without explicitly naming them, while *Citizen* also mentions socioeconomic status and *Django* adds political belief and family status to the list of known diversities. The *Open Code of Conduct* contains a separate section, i.e., "diversity statement", which explicitly deals with diversity and encourages members towards certain expected behaviours in particular situations (*e.g.,* when a participant has made a mistake). Furthermore, in addition to the aforementioned personal characteristics they also refer to language and technical abilities as diversity axes.

Honorable behaviour: codes of conduct tend to pinpoint general positive behaviours like being respectful, patient, kind, focusing on the best for the community, being considerate and collaborative. The *Python* code of conduct is the least specific about positive behaviours, only mentioning being open, considerate and respectful, while *Geek Feminism* does not mention any accepted or encouraged behaviour. *Ubuntu* listed more detailed positive actions such as encouraging community members to ask questions in case of doubt, stressing everyone's responsibility to address such questions. They also encourage any members who want to leave the community to do so with minimal interruption for the project. The *Ubuntu* code of conduct also has a special section about leadership and responsibility, for instance pointing out that leadership can be taken up by anyone competent in the community, basically declaring the community to be a meritocracy in its code of conduct. Expected behaviours for leaders are also named, such as highlighting and rewarding great work of others, or being courageous to take bold decisions. Such leadership-related clauses seem specific to ecosystem-related codes of conduct.

Unacceptable behaviour: most codes of conduct denounce sexist/racist language, contempt and jokes that harass marginalized people, as well as violence and threats. The *Contributor Covenant* explicitly mentions sexualized language or imagery, trolling, insulting and publishing of private information of others as unexpected behaviors. *Django* adds discriminatory jokes and violent threats to this list. *Geek Feminism* and the *Open code of conduct* provide a more detailed list, pointing out additional issues such as deliberate misgendering, physical contact, stalking, following, harassing photography and recording, and threats of violence. *Python* and *Ubuntu* do not refer to unexpected or negative behaviors in their codes of conduct at all.

The studied codes of conduct differ in the style used in their description, *i.e.*, some like *Geek Feminism* just mention negative behaviours to discourage community members from exhibiting those, while on the opposite side of the spectrum *Python*, for instance, just states desired and valued behaviours to reinforce such positive behaviours. Finally, some codes like *Citizen* and *Ubuntu* refer to both unacceptable and honorable behaviours. The styles of the analyzed codes of conduct, either positive or negative, are shown in table IV.

Furthermore, while expressing positive and negative behaviours, some codes of conduct like *Citizen* are phrased in the form of rules like "Refrain from demeaning, discriminatory, ..." (policy/rule-based [1]), while others like *Ubuntu* instead state their intent by listing the expected values of the community, such as "Be respectful" (values/principles-based [1]). The former are very detailed and provide rules and policies in the form of dos and don'ts, while the latter are

expressed by examples and principles rather than exhaustive policies and rules. In between these two extremes of codes of conduct, some codes of conducts can mix both approaches. For instance, *Django* is rule- and values-based at the same time, since it discusses both the value of "Being respectful", but also states "Do not insult or put down other participants".

Enforcement: In communities with a code of conduct, unacceptable behaviors typically are reported to a specific group of team members with the power to decide about the appropriate actions to take. In the *Contributor Covenant*, violations can be reported via a specified email address. The reporter should be treated confidentially, while the report must be investigated and followed up appropriately. In addition, if the code of conduct is not enforced correctly, the project leadership can repress the corresponding maintainers. The *Open Code of Conduct* specifies a more detailed process, such as the information required in a report and who will be responding in special cases (*e.g.,*, when the respondent himself did the harassment). *Django* promises to answer reports within a specific period (one week) but they prioritize ongoing situations and threats to physical safety as distinguished incidents to be addressed immediately.

Geek Feminism proposes a responsible team called the "Geek Feminism Anti-Abuse Team" that promises to not name the victim publicly and respect her confidentiality. *Citizen* just introduces a specific group of members with their contact information for receiving the violation reports, claiming that respondents will help victims. Surprisingly, *Ubuntu* and *Python* do not mention any reporting or enforcement guidelines in their codes of conduct.

Since some codes of conduct, especially those intended to be customized by other open source projects, express honorable and unacceptable behaviours in a generic way, declaring concrete enforcement and punishment mechanisms is not possible for them as these would differ based on the specific project that adopts them. However, even in those cases, some codes of conduct like *Django* and *Geek Feminism* still provide boundaries for any punishments. For example, *Django* proposes a list of punishments ranging from "Nothing" to "A request for a public or private apology" that will be performed in response to violations, while *Geek Feminism* warns to exclude offenders from the community as response to a violation, or to publicly identify the harasser to the project's (or even general) community. *Citizen* briefly mentions the consequences of unaccepted behaviors like permanent expulsion from the community as worst case penalty.

Scope: codes of conduct apply to all community members, *i.e.*, both paid and volunteering contributors, in all community spaces (online and offline). The *Contributor Covenant* and *Citizen* define their scope not only as the communication spaces within the project, such as mailing lists, but also as any outside space where a community member is representing the project, for example when using an official project email address. The enforcement scopes of *Django* and *Geek Feminism* seem wider, as they cover not only the community spaces, but also any intervention of a member out-side the community (either when representing the community or not). In case of any violation, the responsible committee should be informed and this may have consequences for that member. *Python*, *Ubuntu*, and the *Open Code of Conduct* do not explicitly state the scope of their codes of conduct.

The different scopes of codes of conduct are related to the intention behind the code. For example, *Geek Feminism*, as a community supporting women who do geeky activities, has as goal to support women within these activities everywhere, with as consequence that their code of conduct ranges even outside the community spaces. However, some communities like OpalRB only cares about things that are strictly related to their projects' community[23]. This is why the Opal community, after the notorious incident discussed in section I, adopted[24] *Contributor Covenant* v1.0[25], a version of this code of conduct where the boundary of enforcement is not specified and can be restricted to the inside of the community.

Codes of conduct outline an OSS community's expectations and values against the behaviour of members to create a friendly and inclusive community, while violations have consequences. Their scope is all spaces of the community, either online or offline, however a broader scope is possible.

V. RQ3. How are codes of conduct used in open source projects?

Motivation So far, we focused on the popularity of codes of conduct in open source projects, and tried to understand the basic elements and attributes of codes of conduct. In this research question, we investigate why and how they emerged in the open source world, which problems they are dealing with, the process and thought behind them, and their influence and limitations in open source communities.

Approach

Since no quantitative data is available about the adoption of codes of conduct, nor their enforcement (*e.g.,* there is no such thing as a code of conduct complaint repository), we opted for interviews with leaders and creators of codes of conduct in open source communities. Creators of 6 of the top codes of conduct in table II were contacted and invited to a Skype call interview[26]. Of the 5 positive reactions, we were able to perform 4 interviews, two via email and two via Skype (each interviewee belonged to a different open source community). We refer to the anonymous interviewees as A, B, C, and D in the rest of the paper.

Skype interviews were held in semi-structured form and email interviews in structured form [6], both based on our list of questions in table V. In the semi-structured interviews, interviewees could also bring up new ideas to complement our

---

[23]https://github.com/opal/opal/issues/942
[24]https://github.com/opal/opal/pull/947
[25]http://contributor-covenant.org/version/1/0/0/
[26]Creation of a code of conduct usually is a collaborative process.

TABLE IV: COMPARISON OF THE CHARACTERISTICS OF THE SEVEN STUDIED CODES OF CONDUCT.

Code of Conduct	Length (pages)	Phrasing	Style	Scope	> 1 Version?	Dependencies
Citizen	3	Rule	Both	All community venues online or offline. Also outside the scope of the community when violations adversely affect the well-being of the community.	Yes	Django, Geek Feminism
Contributor Covenant	2	Rule	Both	Within project spaces and in public spaces when representing the community, for example by email.	Yes	
Django	4 (with reporting guidelines)	Both	Positive	All spaces online or offline, outside spaces may affect person's participation	Yes	Speak UP!
Geek Feminism	3	Rule	Negative	All Geek Feminism sponsored spaces, both online and offline, and also outside.	No	
Open Code of Conduct	3	Rule	Both	Not mentioned	No (stopped)	
Python	1	Value	Positive	Not mentioned	No	
Ubuntu	4	Both	Positive	Not mentioned	Yes	

TABLE V: PREPARED QUESTIONS FOR INTERVIEWS WITH OPEN SOURCE PRACTITIONERS.

1-	What motivated your community to create a code of conduct?
2-	To create your project's code of conduct, how were the needs of the community gathered?
3-	Did you reuse an existing Code of Conduct (from other industries or software industry)?
4-	How did your project's community react to the idea of having a Code of Conduct?
5-	How is your Code of Conduct enforced? Have any violations been reported already?
6-	What are the strengths and weaknesses of a Code of Conduct for open source projects?

predetermined questions, while in our email interviews they could just answer the predetermined questions. One Skype call took half an hour, the other took one hour (based on the interviewees' availability).

The interviews were audio-recorded and afterwards we prepared a transcription of the talk. Again by applying an *open coding process*, we analyzed the interviews and decomposed them into distinct themes, which approximately mapped to the questions, such as *motivation*, *evolution*, and *strengths*, while labeling them with any important information from the transcripts. Finally, a table was built with the major identified themes as rows, and the different viewpoints of the four interviewees in the different columns.

Results

Motivation for creating a code of conduct (even early on in a project): observations of negative behaviour in previous communities. All interviewees stated that they reflected on their own personal experience of negative behaviours in the past, in different spaces of the community such as on mailing lists or IRC. Therefore, they were motivated to create a dedicated code of conduct for their community. For instance, interviewee B said they wanted to create a new community around some desired behaviours and attitudes that were uncommon in their previous community.

A and B have created their own codes of conduct from the beginning of the project when their communities were still small. They stated that the code of conduct acted as ground rules for the project helping the community to grow, without it

they would not be able to tolerate working in the community.

C stated that arguments happen all the time in open source communities, like the typical argument between people who implemented a product and pursue its stability in contrast to members who are looking to grow the product by adding new features. Hence, open source communities are contentious sometimes, and a code of conduct could be one solution to protect the community, retaining a safe and friendly environment.

Ingredients of code of conduct: concerns, expectations, priorities and even history of a community. Similar to the aforementioned motivation for creating a code of conduct, the ingredients to add to those codes of conduct also relate back to the experiences of the creators in previous communities. Two interviewees, B and C, both pointed out that these needs can also be based on the histories of certain kinds of behaviours in the community, the priorities of the community, their products and consequently situations the community might be running into. One interviewee also affirmed that during the writing phase of a code of conduct, the corresponding community should be understood well and should not be dictated by other communities on how they are supposed to behave. Interviewee C also indicated that even culture and geography play a role. For instance, they got feedback from their project members in India and Middle East stating that their code of conduct reflected issues that are too specific to the United States.

Corporate-supported open source projects[27] encountered a duality, as Interviewee C mentioned. On the one hand, they have to comply with their company's policies about campus, central buildings and so on. On the other hand, they also want to offer a welcoming and protective environment for their projects. Thus, there is a gap between existing codes of conduct and the typical issues those organizations have to deal with as corporations, which should be handled in their specific code of conduct too.

However, interviewee A believed that codes of conduct will converge to some extent, because communities already copy text from one another, and he expects this to continue. For instance interviewee A's community copied some sentences

---

[27]I.e., major commercial software companies like Microsoft that manage their own open source projects.

from another code of conduct, which were acknowledged explicitly in the text. His opinion was confirmed by interviewee C, since A realized that some communities have similarities that can lead to reuse of codes of conduct. In section III, we also observed that many projects have been inspired by or reused another code of conduct.

Evolution: Similar to software artefacts, codes of conduct evolve as well. Interviewee B explained that their code of conduct has been updated five or six times in 10 years. Starting from a short document that did not seem to be that serious, gradually several elements were changed, for example the phrasing changed from rule-based to value-based, and they added significant details on the leadership of the community, whereas leadership was not a big deal for them at the beginning. He also stated that throughout all of these versions, they only changed the textual expressions, but not the intention of the code of conduct.

According to interviewee A, every new suggestion about the code of conduct from their community members is welcomed until there are enough arguments and justifications. In interviewee C's community, the discussion for the code of conduct is started by the board, then passed around to all members via the member mailing list. This list is where community needs were addressed and implemented as necessary. Afterwards, changes were voted on by board membership.

As such, all interviewees experienced or expected some changes to their codes of conduct across time. This is because new situations come up, new members are invited, communities are expanded and new or more concrete needs are raised that need to be covered by codes of conduct. From the interviews, it followed that, similar to regular code changes, such changes to a code of conduct should be revised and investigated by a committee before applying them, although different communities might use different processes to apply changes. For instance, in community D, voting among the board membership happens.

Reaction: especially people outside a project complain about the code of conduct. Since most of the interviewed communities grew around their codes of conduct, their projects considered a code of conduct to be a good thing and codes have been adjusted as needed by the community across time. Interviewee A said that the code largely served to attract like-minded people and repel those who did not agree, which caused reinforcement of the norms embedded within the code.

B confirmed that his community rejected those who wanted to join the community but did not like their code of conduct. Based on B's opinion, a code of conduct is a means for "retention of newcomers but it does not attract any newcomers".

Interviewee C said that none of the communities he runs reacted negatively, i.e., the code of conduct was a non-issue to them. However, according to him, from outside the community, several criticisms have been voiced, mostly like "heterosexual white man can not write a code of conduct and can not understand how to do it". However, in his opinion the criticism that someone is not capable to do something because of his/her color in fact goes against the code of conduct.

Enforcement policy: varies from signing the code by all members to a responsibility for project leaders to protect the community. In B's community, members should sign their code of conduct, while in communities of other interviewees, there was no such need and the code is adopted implicitly by any contributor. However, interviewee C believed that the code of conduct is a declaration on the part of the project leader, not the community members. It should say that someone as a leader will take the responsibility to address all complaints and will not harass members for complaining.

Interviewee B stated that in their community, a code of conduct is mostly enforced in the sense that people refer to it whenever someone is being disrespectful, for example on a chat channel or mailing list. Just the act of mentioning the code suffices in most of the cases to calm down a conflict. In a select number of cases, when a difficult situation occurred, it had to be raised to a public community counselor. Interviewee A talked about actions that his community chose to enforce a code of conduct: first a polite notice, later a stern notice or warning, then a request to leave, finally a technical measure to prevent participation. During his tenure of around 3 years, only a handful of people had to be asked to leave. Interviewee D had experienced two reports during his tenure.

Strength and weaknesses: codes of conduct promote a friendly and inclusive environment, but may induce a policing environment. Finally, all interviewees believed that a community with a code of conduct has an advantage to foster clear, explicit norms that make the social environment more tolerable, friendly and welcoming to many individuals, while empowering diversity. Interviewee A mentioned that "it helps a community grow in a way that reinforces those norms, by articulating common reference points. It encourages development of social awareness, in this way reflecting on appropriate behaviour, i.e., people skills." Interviewee C believed that having a code of conduct means that leaders make a commitment that they care about their community members.

Interviewee A mentioned that it is impossible to exhaustively enumerate all situations or resolutions for conflicts in a code of conduct, which he identified as a drawback or inability of codes of conduct. He also believed that a code of conduct may embed and enforce norms more specific or larger than some community members desire. The interviewee also stated that some community members may find a code of conduct intolerable, as it may imply a degree of behavioural scrutiny or "policing". Interviewee D confirmed this issue as a weakness and mentioned that some people feel that a code of conduct prohibits them from speaking freely.

About the difference between *rule-based* and *value-based* phrasing of a code of conduct, interviewee B found that both have their benefits, such as signaling a safe and welcoming environment. Rule-based phrasing is more clear and allows to easily react when a conflict happens, whereas value-based phrasing is less likely to deter people from joining the community, which is an advantage. Since in rule-based phrasing, misbehaviours are mentioned, they may induce to others that the related community is an environment with hostile and

unfriendly issues.

> A community's concerns, needs and history play an important role in the design of its code of conduct. Yet, similarities among communities may lead communities to reuse an existing code of conduct.

VI. Discussion

Several developers, researchers and activists have proposed their vision of codes of conduct, their desired content and role. According to Hermans [15], codes of conduct should trigger discussion and change people's minds. She believes that if the code of conduct does not induce some pushback, it is not working properly. While she is referring to conference codes of conduct, we share her vision of codes of conduct as the discussion and reflection vehicle.

Similarly to Hermans, the authors of *Geek Feminism* stress the importance of effectiveness in the code of conduct[28]. They stress the need for reporting and enforcement policies as well as a "clear demarcation between unacceptable behaviour <...> and community guidelines". A recent addition to the Code of Conduct landscape are the GitHub community guidelines[29]. While this document is not finalized at the moment of writing, nor does it refer to itself as a code of conduct, it shares several common features with the codes of conduct we have studied: discussion of best practices vs. undesired behavior as well as reporting and enforcement mechanisms.

As we discussed in RQ2, open source codes of conduct have a specific section specifying their enforcement approach, in which violation penalties can range from a simple warning process to banning from the project, with many more nuanced penalties in between. Some codes like *Django* provide more specific enforcement mechanisms, for instance by introducing a time period to take proper actions in case of code violation. We do not have enough evidence to determine whether such a detailed approach is more successful. However, the Ada initiative community admired the *Django* and *Rust* codes of conduct as two open source codes with a well-defined and -documented complaint handling process [30].

Another point in line with active enforcement is the awareness of community members about their codes of conduct. In RQ3, we found that in one community members sign their code of conduct and get informed about it when joining the project, while there are communities without explicit adoption of codes. O'Dwyer et al. [19] conducted a survey to examine different properties of professional codes of conduct, like their enforcement among the top 1,000 companies based in Ireland, and the awareness amongst those companies' employees. A similar study is needed in open source communities to analyze the implementation of codes of conduct in this context, while

---

[28]http://geekfeminism.wikia.com/wiki/Code_of_conduct_evaluations

[29]https://github.com/blog/2267-introducing-github-community-guidelines

[30]https://adainitiative.org/2014/02/18/howto-design-a-code-of-conduct-for-your-community/

---

shedding light on the perceptions of individuals about their community's code of conduct.

VII. Limitations and Threats to Validity

Threats to construct validity focus on how accurately our observation measurements are done. To measure the number of open source projects with a code of conduct and also the most popular codes of conduct, we performed queries on Google and GitHub. Certainly, there were false positive hits among these results, as discussed before, while the applied search queries might not return all possible cases. To mitigate these risks, we complemented a rough name-based search approach with 3 smaller, manually analyzed data sets that were used together to evaluate the most popular codes of conduct.

Regarding external validity, for RQ2 we considered seven codes of conduct as representative codes, however these might be different from other existing codes of conduct, or not comprehensive enough to cover others. We reduced this risk by considering seven codes of conduct that seem to be among the most popular codes according to Google queries. However, further studies need to confirm our findings on other codes of conduct. Similarly, while the practitioners invited for the interviews were creators of actual codes of conduct, only 4 of the invited people could be interviewed. Given their experience with the creation and adoption of a code of conduct, and the overlap between their answers, we are confident that our findings cover a large spectrum of codes. Of course, interviews with other practitioners as well as a larger-scale survey are necessary to generalize our findings.

Threats to internal validity concern confounding factors that can influence the results. For example, the interviewees may be biased, since they were the driving force behind the creation and adoption of a code of conduct. We tried to lessen this risk by designing structured and semi-structured interviews, preparing unbiased questions that also encompass different aspects of codes of conduct. Again, more interviewees should be considered to further reduce this risk.

VIII. Conclusion

In this paper, we found that adopting codes of conduct in open source projects is an emerging phenomenon in order to deal with diversity issues and provide a safe and inclusive community. The phrasing of a code of conduct, enforcement mechanism used, scope and other properties might vary depending on the code of conduct and community. We obtained these insights through a combination of manual analysis of the codes themselves, as well as through interviews with creators and adopters of codes of conduct.

Although our study is the first step showing the role of codes of conduct in open source projects, we believe it opens up a variety of research opportunities, since it is one of the first mainstream solutions to deal with conflicts in a software project. Recently, research on detection of emotions, politeness and sentiment in software engineering interactions has taken off [13], [16], [18], [20], [23], [24]. Except for Dullemond et al. [11], most of this work focuses on measuring the presence

of some kind of conflict or negative feelings, with the aim of informing managers about these. Codes of conduct are a concrete tool to act on such information.

However, since codes of conduct are relatively young, more detailed empirical evaluation of their effectiveness and of best practices is required. More qualitative and quantitative studies should be done to provide recommendations and guidelines for the design and improvement of codes of conduct in this domain. Once codes of conduct will be more mature, more data will be available for quantitative studies, especially regarding enforcement and effect of the code of conduct in different stages and processes of the software development process. When designing future quantitative studies one should, however, be aware that the unacceptable behaviour might be under-reported akin to under-reporting of rape [4], and that the introduction of a code of conduct might even result in the increase of the number of reported violations, *e.g.,* due to increased community awareness.

REFERENCES

[1] Ifac. Available at http://www.ifac.org.
[2] Open code of conduct. Available at http://todogroup.org/opencodeofconduct/.
[3] Open source initiative. Available at https://opensource.org/osd-annotated.
[4] W. D. Allen. The reporting and underreporting of rape. *Southern Economic Journal*, 73(3):623–641, 2007.
[5] R. Bazile-Jones and B. Lynn. Diversity in the workplace: Why we should care. *CMA—the Management Accounting Magazine*, 70(5):9–12, June 1996.
[6] S. Cohen and D. Bailey. What makes teams work: Group effectiveness research from the shop floor to the executive suite. *Journal of Management*, 23(3):239–290, 1997.
[7] J. A. Colazo and Y. Fang. Following the sun: Temporal dispersion and performance in open source software project teams. *Journal of the Association for Information Systems*, 11(11/12):684–707, 2010.
[8] L. Dabbish, C. Stuart, J. Tsay, and J. Herbsleb. Social coding in github: Transparency and collaboration in an open software repository. In *CSCW*, pages 1277–1286, New York, NY, USA, 2012. ACM.
[9] S. L. Daniel, R. Agarwal, and K. J. Stewart. The effects of diversity in global, distributed collectives: A study of open source project success. *Information Systems Research*, 24:312–333, 2013.
[10] M. Di Penta. Combining quantitative and qualitative methods (when mining software data). In T. Menzies, L. Williams, and T. Zimmermann, editors, *Perspectives on Data Science for Software Engineering*, pages 205–211. Morgan Kaufmann, Boston, 2016.
[11] K. Dullemond, B. v. Gameren, M.-A. Storey, and A. v. Deursen. Fixing the "out of sight out of mind" problem: One year of mood-based microblogging in a distributed software team. In *MSR*, pages 267–276, Piscataway, NJ, USA, 2013. IEEE Press.
[12] C. A. Ehmke. Contributor covenant code of conduct. Available at http://contributor-covenant.org.
[13] D. Gachechiladze, F. Lanubile, N. Novielli, and A. Serebrenik. Anger and its direction in collaborative software development. In *ICSE NIER*, 2017. accepted.
[14] C. Gutwin, R. Penner, and K. Schneider. Group awareness in distributed software development. In *CSCW*, pages 72–81, New York, NY, USA, 2004. ACM.
[15] F. Hermans. On codes of conduct, Aug. 2016. Available at http://www.felienne.com/archives/5130.
[16] R. Jongeling, P. Sarkar, S. Datta, and A. Serebrenik. On negative results when using sentiment analysis tools for software engineering research. *Empirical Software Engineering*, 2017. accepted.
[17] B. Kitchenham, O. Pearl Brereton, D. Budgen, M. Turner, J. Bailey, and S. Linkman. Systematic literature reviews in software engineering - a systematic literature review. *Inf. Softw. Technol.*, 51(1):7–15, Jan. 2009.
[18] N. Novielli, F. Calefato, and F. Lanubile. The challenges of sentiment detection in the social programmer ecosystem. In *SSE*, pages 33–40, New York, NY, USA, 2015. ACM.
[19] B. O'Dwyer and G. Madden. Ethical codes of conduct in irish companies: A survey of code content and enforcement procedures. *Journal of Business Ethics*, 63(3):217–236, 2006.
[20] M. Ortu, B. Adams, G. Destefanis, P. Tourani, M. Marchesi, and R. Tonelli. Are bullies more productive? Empirical study of affectiveness vs. issue fixing time. In *MSR*, pages 303–313, Florence, Italy, May 2015.
[21] W. E. Powell, D. S. Hunsinger, and B. D. Medlin. Gender differences within the open source community: An exploratory study. *Journal of Information Technology*, 21(4):29–37, 2010.
[22] E. S. Raymond. *The Cathedral and the Bazaar*. O'Reilly & Associates, Inc., Sebastopol, CA, USA, 1999.
[23] P. Tourani and B. Adams. The impact of human discussions on just-in-time quality assurance. In *SANER*, pages 189–200, Osaka, Japan, March 2016.
[24] P. Tourani, Y. Jiang, and B. Adams. Monitoring sentiment in open source mailing lists -– exploratory study on the apache ecosystem. In *CASCON*, pages 34–44, Toronto, ON, Canada, 2014.
[25] B. Vasilescu, V. Filkov, and A. Serebrenik. Perceptions of diversity on GitHub: A user survey. In *CHASE*, pages 50–56. IEEE, 2015.
[26] B. Vasilescu, D. Posnett, B. Ray, M. G. J. van den Brand, A. Serebrenik, P. Devanbu, and V. Filkov. Gender and tenure diversity in GitHub teams. In *CHI*, pages 3789–3798. ACM, 2015.
[27] B. Vasilescu, A. Serebrenik, and V. Filkov. A Data Set for Social Diversity Studies of GitHub Teams. In *MSR*, pages 514–517. IEEE, 2015.