How do Developers Promote Open Source Projects?

Hudson Borges, Marco Tulio Valente

Department of Computer Science, UFMG, Brazil

{hsborges,mtov}@dcc.ufmg.br

Abstract

Open source projects have an increasing importance on modern software development. For this reason, these projects, as usual with commercial software projects, should make use of promotion channels to communicate and establish contact with users and contributors. In this article, we study the channels used to promote a set of 100 popular GitHub projects. First, we reveal that Twitter, user meetings, and blogs are the most common promotion channels used by the studied projects. Second, we report a major difference between the studied projects and a random sample of projects, regarding the use of the investigated promotion channels. Third, we show the importance of a popular news aggregation site (Hacker News) on the promotion of open source. We conclude by presenting a set of practical recommendation to open source project managers and leaders, regarding the promotion of their projects.

1 Introduction

Open source projects have an increasing importance in modern software development. For example, several open source projects are daily used by millions of users. However, it is very important to continually attract more participants and contributors to these projects, in order to increase the chances of long-term success [1]. Particularly, several channels can be used to promote open source software, helping to keep the interest of the community and also to attract new members.

In this article, we investigate the most common channels used by developers to promote open source projects. We manually inspected a large set of popular projects on GitHub, which is the world's largest collection of open source software, with around 27 million users and 77 million repositories [2]. Our contributions include: (i) data about the promotion channels frequently used by popular open source projects; (ii) a comparison on the use of

promotion channels by popular projects and by random ones; and (iii) an analysis of the impact of promotion on Hacker News, a popular news aggregation site, in the popularity of the studied projects. Our findings help practitioners to understand the importance of using promotion channels in the open source development context.

2   Study Design

To reveal the most common promotion channels used by developers, we manually inspected the documentation of the top-100 projects with most stars on GitHub (stars is a popular feature to manifest interest or satisfaction with GitHub projects [3]). We restricted our analysis to popular projects because they have a large number of users and therefore need better and efficient ways to communicate with users and also to attract new contributors.

Figure 1 shows the distribution of the number of stars of the projects considered in this study. This number ranges from 291,138 stars (FREECODECAMP/FREECODECAMP) to 23,322 stars (TIIMGREEN/GITHUB-CHEAT-SHEET). The considered projects are primarily developed on 17 programming languages; JavaScript is the most common one (40 projects), followed by Python (9 projects) and Go (5 projects). Furthermore, 14 projects only include markdown files with documentation purposes (e.g., projects with tutorials, books, awesome lists, etc). Finally, regarding the project owners, 69 are organizational accounts and 31 are user accounts.
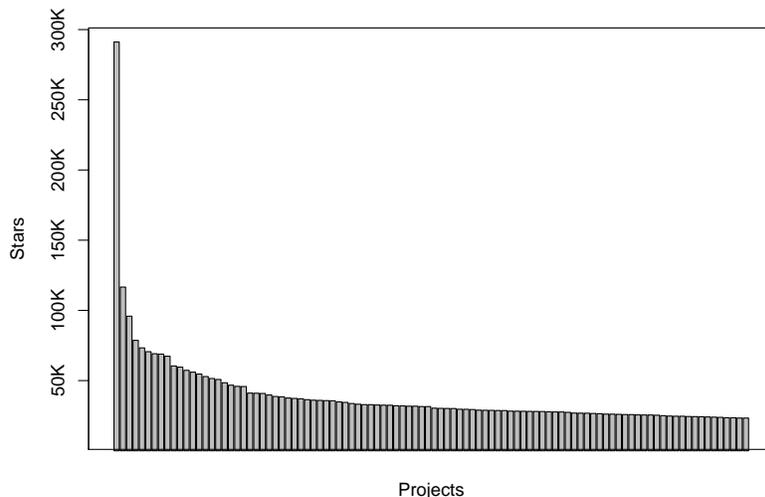

Figure 1: Number of GitHub stars of the analyzed projects

For each of these 100 projects, the first author of this paper initially inspected their

READMEs on GitHub to identify the channels used to promote the projects and to keep the users up-to-date with important information about them. For example, the following sentence is available on the README of ADOBE/BRACKETS: "*You can see some screenshots of Brackets on the wiki, intro videos on YouTube, and news on the Brackets blog*". In this case, wiki and YouTube are used to support users whereas blog is a channel used to disseminate news about BRACKETS. Thus, only blog is considered a promotion channel in our study. Next, we inspected the projects' website, for those projects having one. We navigated through the site pages, searching for more channels used to promote the projects.

After this manual inspection, the following promotion channels emerged:

- **Blogs**, which are used, for example, to publish announcements of new software versions, upcoming events, and improvements.

- **Events and Users Meetings:** Organizing events and supporting users meetings are other strategies commonly followed to promote projects. On events the initiative usually comes from the development team or from the organization that supports the project, whereas on user meeting the initiative comes from the users, usually from a specific region or country. We rely on Meetup (`https://meetup.com`) to discover users meetings.

- **Twitter, Facebook, and Google+**, which are also used to connect the projects to users. We considered only official accounts, which are explicitly advertised on the project documentation or are verified by the social network (e.g., `https://support.twitter.com/articles/20174631`).

- **Newsletter and RSS feeds**, which refer to e-mails with the most relevant news about the projects and RSS feeds.

In addition, we found that developers use Q&A forums (e.g., StackOverflow), discussion groups (e.g., Google Groups), and messaging tools (e.g., IRC and Slack) to promote their projects. However, these channels are mostly used to discuss the projects and to provide answers to common questions raised by users. For example, from the 155 topics opened in 2017 in the ADOBE/BRACKETS discussion group at Google Groups, only eight (5.1%) are related to announcements of new versions, mostly pre-releases for community testing. Moreover, from almost 500 topics on FACEBOOK/REACT official forum, we could not identify any announcement related to the project development. Thus, in this study, we do not consider forums, discussion groups, and messaging tools as promotion channels.

3   Results

3.1   What are the most common promotion channels?

Figure 2 presents the most common promotion channels used by the top-100 projects on GitHub. The most common channel is Twitter, which is used by 56 projects. The second one is Users Meetings (41 projects), followed by Blogs (38 projects), Events (33 projects), and RSS feeds (33 projects). The least common channels are Facebook and Google+, which are used by 18 and 7 projects, respectively.
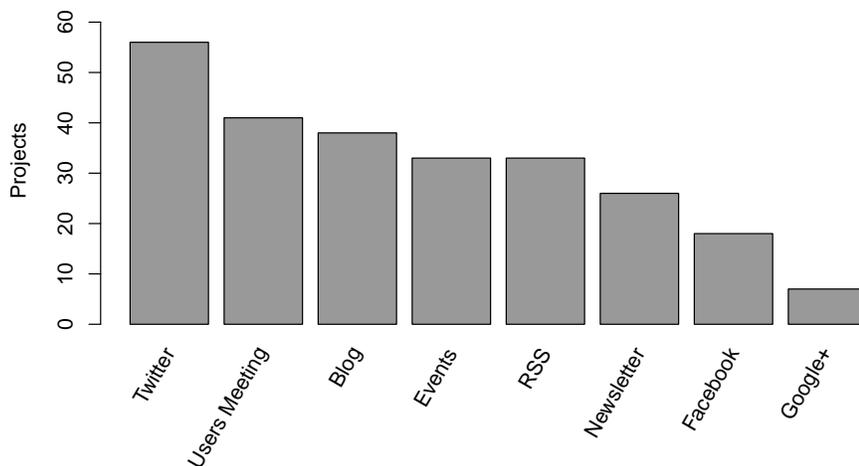

Figure 2: Most common promotion channels

Figure 3 shows the distribution of the number of promotion channels per project. Almost one third of the projects (32 projects) do not use any channel. By contrast, more than half of the projects (55 projects) use at least two promotion channels. The highest number of promotion channels is seven, which is the case of FACEBOOK/REACT, FACEBOOK/REACT-NATIVE, METEOR/METEOR, GOLANG/GO, IONIC-TEAM/IONIC, ANGULAR/ANGULAR, and ADOBE/BRACKETS. We also found that Blog and Twitter is the most frequent combination of channels (35 projects). Other frequent combinations include, for example, Blog and RSS (31 projects), Events and Users Meetings (31 projects), and Twitter, Events and User Meetings (31 projects).

3.2   How often do developers promote their projects?

In this second question, we investigate how often developers promote their projects on blogs and social networks. For blogs, we calculate the promotion frequency as the number of posts
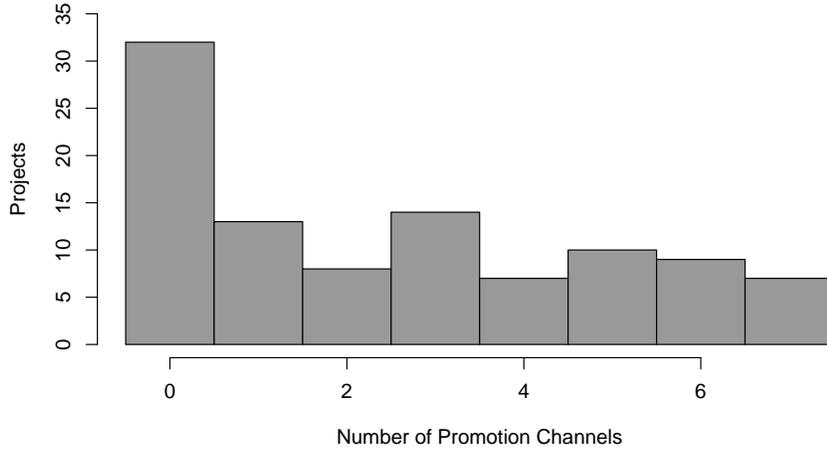
Figure 3: Number of promotion channels per project

on the last 12 months. For social networks, we could not retrieve all posts for all projects because their APIs restrict the search to a recent period (e.g., last seven days for Twitter and last 100 posts for Facebook). Thus, in this case, we only classified each social network account in two distinct groups: active and inactive. An *active* account has at least three posts on the last three months; otherwise, it is considered an *inactive* account. This classification was performed by manually counting the number of posts on the social network pages.

Figure 4 presents the distribution of the number of blog posts on the last 12 months. The number ranges from 1 (NYLAS/NYLAS-MAIL) to 1,300 (FREECODECAMP/FREECODECAMP); the first, second, and third quartile values are 7, 19, and 54 posts, respectively.
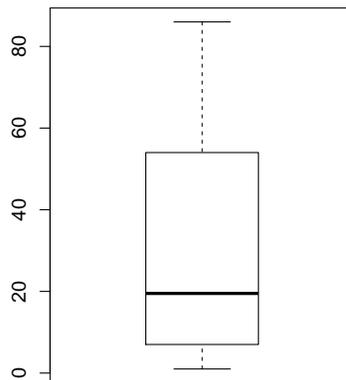

Figure 4: Distribution of the number of posts on the last 12 months (outliers are omitted)

Table 1 lists the activity status of the Twitter, Facebook, and Google+ accounts. We found that 83.9% of the projects that use Twitter have an active account; 55.6% of the projects have an active Facebook account and only 28.6% have an active Google+ account.

5

Table 1: Active Twitter, Facebook, and Google+ accounts

Channel	Active (%)	Inactive (%)
Twitter	47 (83.9%)	9 (16.1%)
Facebook	10 (55.6%)	8 (44.4%)
Google+	2 (28.6%)	5 (71.4%)

Finally, we investigate the characteristics of the user meeting groups promoted on Meetup (such meetings are the 3rd most common promotion channel studied in this article). A Meetup group is a local community of people that is responsible for organizing meeting events [4]. These groups are identified by topics to help members find them. Here, we rely on these topics to collect meetups about the studied open source projects, along with their locations (i.e., city and country). For example, the topic for JQUERY/JQUERY is *jquery* and a summary of the meeting groups about this topic can be found at `https://www.meetup.com/topics/jquery/all`. Figure 5 presents the distribution of the number of groups, cities, and countries of the projects with meetings registered at Meetup. For groups, the values ranges from 2 to 2,261 groups; considering the cities, the values range from 2 to 725; finally, for countries, the values range from 2 to 96. The maximum values always refer to TORVALDS/LINUX. In other words, TORVALDS/LINUX has 2,261 meetup groups, which are spread over 725 cities from 96 countries.
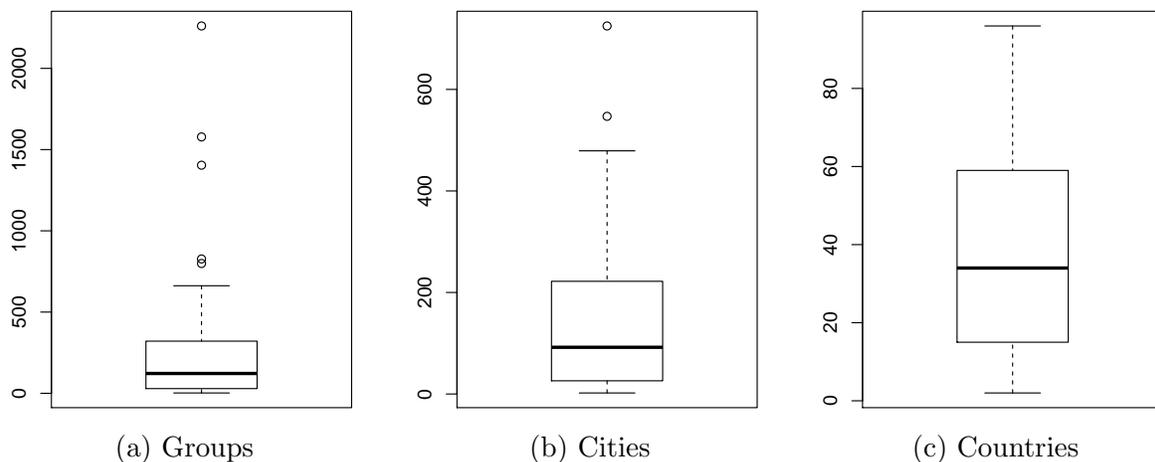

(a) Groups   (b) Cities   (c) Countries

Figure 5: Number of groups, cities, and countries of the user meetings

3.3 How popular and random projects differ on the usage of promotion channels?

In Section 3.1, we investigated the most common promotion channels used by popular GitHub projects. In this section, we contrast the usage of promotion channels by these projects and by a random sample of GitHub projects. For this purpose, we randomly selected 100 projects from the top-5,000 repositories by number of stars and manually inspected their documentation using the same methodology reported in Section 2. The number of stars of this random sample ranges from 2,297 stars (UBER-ARCHIVE/IMAGE-DIFF) to 22,558 (VSOUZA/AWESOME-IOS).
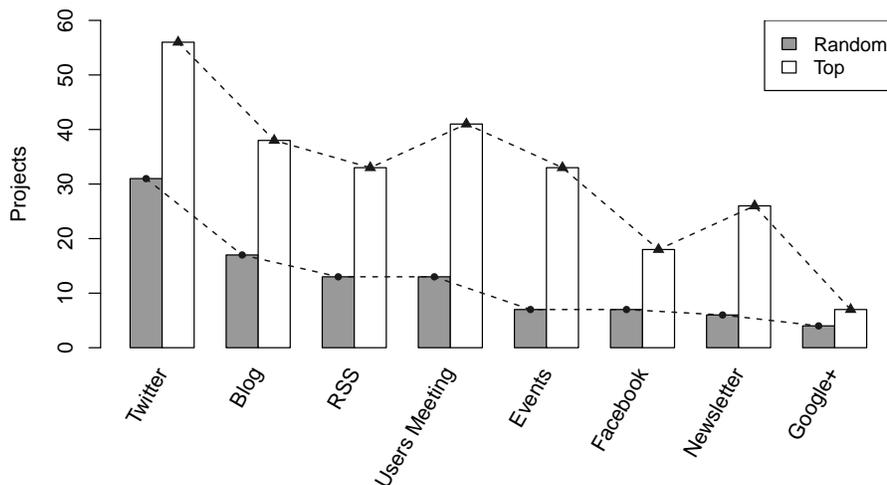
Figure 6: Most common promotion channels used by random projects

Figure 6 compares the usage of promotion channels by the random projects and by the most popular ones. In the random sample, the number of projects using the investigated promotion channels is significantly lower compared to the most popular ones. However, by applying the Spearman's rank correlation test, we found a strong correlation between the number of projects using the promotion channels on each group ($rho = 0.904$ and $p\text{-value} < 0.01$). For example, Twitter is also the most used promotion channel among the random projects (31 projects), followed by Blogs (17 projects) and RSS (13 projects). Compared to the most popular projects, Users meetings and Newsletter are less common (13 and 6 projects, respectively). Finally, Facebook and Google+ also have a very limited usage (7 and 4 projects, respectively).

3.4 What is the impact of promotion on Hacker News?

After publishing content on blogs, Twitter, etc., open source developers can also promote this content on social news aggregator sites. These sites aggregate contents from distinct sources for easing viewing by a large public. The most popular and important example is Hacker News (`https://news.ycombinator.com`), which is dedicated to Computer Science and related technologies content. Hacker News posts just include a title and the URL of the promoted content (e.g., a blog post about a new version of an open source project). Any user registered in the site can post a link on Hacker News, i.e., not necessarily the links are posted by the contributors of an open source project, for example. Other Hacker News users can discuss the posts and upvote them. An upvote is similar to a *like* in social networks; posts are listed on Hacker News according to the number of upvotes. In this research question, we use Hacker News due to its popularity; posts that reach the front page of the site receive for example 10-100K page views, in one or two days (`https://goo.gl/evyP4w`). Furthermore, Hacker News provides a public API, which allows search and metadata collection.

For each popular project considered in our study (100 projects), we searched for Hacker News posts with a URL referencing the project sites or pages, including GitHub pages (READMEs, issues, etc). As result, we found 3,019 posts on Hacker News referencing content from 96 studied projects (i.e., only four projects are never referenced on Hacker News). Figure 7 presents the distributions of the number of posts per project, upvotes, and comments. The number of posts ranges from 1 to 298 posts per project (RAILS/RAILS); the first, second, and third quartile values are 4, 10, and 43 posts, respectively. Regarding their upvotes, the most popular post is about APPPLE/SWIFT (*"Swift is Open Source"*), with 1,824 upvotes; the quartile values are 2, 3, and 12 upvotes, respectively. Finally, the highest number of comments is 760, about a GitHub issue opened for Microsoft Visual Studio (*"VS Code uses 13% CPU when idle due to blinking cursor rendering"*); the quartile values are 0, 0, and 2 comments, respectively. On the one hand, these results show that most Hacker News posts do not attract attention. By contrast, a small number of posts attract a lot of attention. For example, the top-10% posts have at least 132 upvotes. These posts are called *successful posts* in this investigation.

Figure 8 shows boxplots with the number of GitHub stars gained by projects covered by successful posts, in the first three days before and after the publication date on Hacker News. The intention is to investigate the impact of a successful promotion on Hacker News, by comparing the number of stars gained before and after each successful post publication. On the median, the projects covered by successful posts gained 74 stars in the first three days before their appearance on Hacker News; in the first three days after the publication, the projects
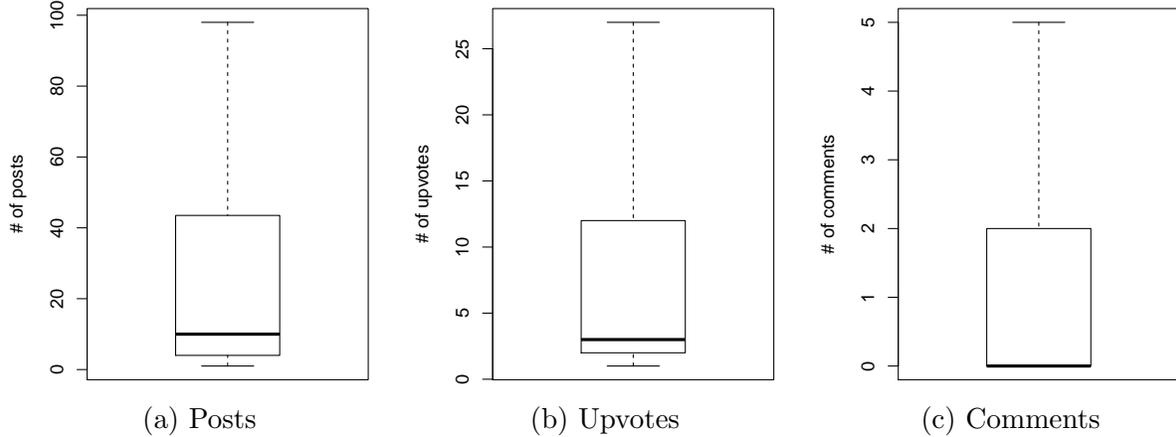
Figure 7: Number of posts, upvotes, and comments (outliers are omitted)

gained 138 stars. Therefore, Hacker News has a positive impact on the project's popularity, measured by GitHub stars. Indeed, the distributions are statistically different, according to the one-tailed variant of the Mann-Whitney U test (p-value ≤ 0.05). By computing Cliff's delta, we found a *medium* effect size ($d = -0.372$).
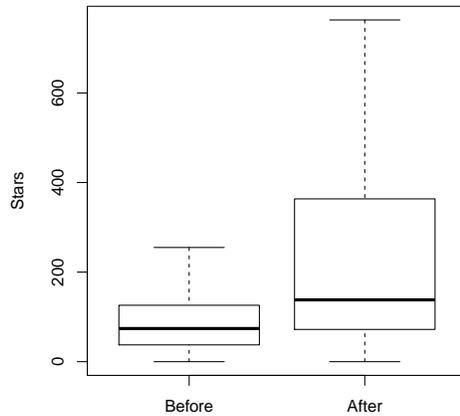

Figure 8: Number of GitHub stars received by projects covered by successful Hacker News posts in the first three days before and after the post publication

Finally, we inspected the titles of each successful post, aiming to categorize the post purpose. The most common category includes posts announcing new releases of open source projects (44.9%; e.g., *"Angular 2 Final Released"*). Other popular categories include posts promoting articles or reports about the projects (25.4%; e.g., *"Vue.js vs. React"*), announcing the first release of a project (16.5%; e.g., *"YouTube-dl: Open-source YouTube downloader"*), highlighting new project features (10.6%; e.g., *"Git and GitHub Integration Comes to Atom"*) and open sourcing products (1.6%; e.g., *"Visual Studio Code is now open source"*).

9

4 Related Work

Although open source software has been exhaustively explored recently, little is known about how developers promote these projects. The main exception is a work conducted by Bianco et al. where the authors analyze marketing and communication strategies of three companies that develop open source software [5]. By means of interviews, they found that websites and product launch events are adopted by the three organizations; however, the organizations differ considerably on the use of other communication channels, mainly when promoting the projects in open source communities and among industrial users.

On the one hand, most communication channels investigated in this paper are explored in other studies, but with different intentions. Singer et al. report a qualitative study focused on discovering the benefits that Twitter brings to developers [6]. They found that Twitter adopters use it to stay aware of industry changes, for learning, and for building relationships. By correlating the blogging and committing behavior of developers, Pagano and Maleej observed an intensive use of blogs, frequently detailing activities described shortly before in commit messages [7]. Bajic and Lyons analyze how software companies use social media techniques to gather feedback from users collectively [8]. Their results suggest that startups use social media mainly for competitive advantage and established organizations use it to monitor the buzz among their users. By studying a successful software development company, Hansson et al. identified that user meetings and newsletter are adopted to include and increase the participation of users in the development process [9]. Finally, Aniche et al. conduct a study to understand how developers use modern news aggregator sites (Reddit and Hacker News) [10]. According to their results, the two main reasons for posting links on these sites is to promote own work and to share relevant content.

5 Conclusion and Practical Implications

In this paper, we investigated the most common promotion channels used by popular GitHub projects. This investigation supports the following practical recommendations to open source project managers and leaders:

1. Promotion is an important aspect of open source project management, which should be emphasized by project leaders. For example, most popular GitHub projects (two thirds) use at least one promotion channel; half of the projects invest on two channels. By contrast, the use of promotion channels is less common among projects with lower popularity.

2. Open source project managers should consider the use of Twitter (47 projects among the top-100 most popular GitHub projects have active Twitter accounts), Users meetings (which are organized or supported by 41 projects), and blogs (which are used by 38 projects).

3. Open source project managers should also consider promotion on social news aggregator sites. Successful posts on Hacker News may have an important impact on the popularity of GitHub projects. However, only 10% of the Hacker News posts about the studied projects have had some success

Acknowledgments

References

[1] S. Comino, F. M. Manenti, and M. L. Parisi, "From planning to mature: On the success of open source projects," *Research Policy*, vol. 36, no. 10, pp. 1575–1586, 2007.

[2] I. GitHub. (2018, Feb.) Code search. [Online]. Available: https://github.com/search

[3] H. Borges, A. Hora, and M. T. Valente, "Understanding the factors that impact the popularity of GitHub repositories," in *32nd IEEE International Conference on Software Maintenance and Evolution (ICSME)*, 2016, pp. 334–344.

[4] Meetup. (2018, Feb.) What is a meetup group? [Online]. Available: https://www.meetup.com/help/article/902256

[5] V. D. Bianco, L. Lavazza, V. Lenarduzzi, S. Morasca, D. Taibi, and D. Tosi, "A study on OSS marketing and communication strategies," in *8th International Conference on Open Source Systems (OSS)*, 2012, pp. 338–343.

[6] L. Singer, F. F. Filho, and M.-A. Storey, "Software engineering at the speed of light: how developers stay current using Twitter," in *36th International Conference on Software Engineering (ICSE)*, 2014, pp. 211–221.

[7] D. Pagano and W. Maalej, "How do developers blog?: An exploratory study," in *8th Working Conference on Mining Software Repositories (MSR)*, 2011, pp. 123–132.

[8] D. Bajic and K. Lyons, "Leveraging social media to gather user feedback for software development," in *2nd International Workshop on Web 2.0 for Software Engineering*, 2011, pp. 1–6.

[9] C. Hansson, Y. Dittrich, and D. Randall, "How to include users in the development of off-the-shelf software: A case for complementing participatory design with agile development," in *39th Annual Hawaii International Conference on System Sciences (HICSS)*, 2006, pp. 175c–175c.

[10] M. Aniche, C. Treude, I. Steinmacher, I. Wiese, G. Pinto, M.-A. Storey, and M. Gerosa, "How modern news aggregators help development communities shape and share knowledge," in *40th International Conference on Software Engineering (ICSE)*, 2018, pp. 1–12.