

論文紹介

On the abandonment and survival of open source projects: An empirical investigation

OSSプロジェクトの放棄と存続についての実証的調査

著者

Guilherme Avelino (Federal University of Piaui, Brazil),
Eleni Constantinou (University of Mons, Belgium),
Marco Tulio Valente (Federal University of Minas Gerais, Brazil),
Alexander Serebrenik (Eindhoven University of Technology, The Netherlands)

1. 背景

OSSの利用者が増えている現在、プロジェクトの持続可能性に対する需要↑

持続可能性の障壁

- 金銭的支援がない
- 少数のコア開発者に依存している

OSSプロジェクトからコア開発者がいなくなる
＝プロジェクトが放棄される

2. この研究の着眼点

既存研究

OSSプロジェクトの持続可能性に大きな影響を持つ
コア開発者を特定するための手法が提案^{[1][2][3]}

次の段階として

コア開発者がいなくなってしまうようにする
＝プロジェクトが放棄されないようにするための戦略を考える必要

[1] N. Zazworka, K. Stapel, E. Knauss, F. Shull, V. R. Basili, and K. Schneider, “Are developers complying with the process,” in ESEM, 2010.

[2] V. Cosentino, J. L. C. Izquierdo, and J. Cabot, “Assessing the bus factor of Git repositories,” in SANER, 2015, pp. 499–503.

[3] G. Avelino, L. Passos, A. C. Hora, and M. T. Valente, “A novel approach for estimating truck factors,” in ICPC, 2016, pp. 1–10.

3. 目的とアプローチ

目的

OSSプロジェクト存続の役に立つ戦略を見つける

アプローチ

1. 人気プロジェクトを定量的分析
プロジェクトが放棄される/存続する頻度 (RQ1, 2)
放棄される/存続するプロジェクトの違い (RQ3)
2. 存続するプロジェクトのコア開発者にアンケート
動機や障壁を調査 (RQ4~6)

4. 人気プロジェクトの定量的分析 | 方法 (1/2)

使用するデータ

GitHub上でリポジトリ数が多い
上位6つのプログラミング言語で
開発されているプロジェクト

Language	Projects	Language	Projects
Ruby	398 (21%)	PHP	334 (17%)
JavaScript	342 (17%)	Python	297 (15%)
C/C++	335 (17%)	Java	226 (12%)

各言語スター数上位500のリポジトリを選択 ⇒ 3,000プロジェクト

開発履歴が欠落しているプロジェクトを除外 ⇒ 1,932プロジェクト

4．人気プロジェクトの定量的分析 | 方法（2/2）

プロジェクトが放棄されたかどうかの判定基準

コア開発者らの最終コミット日時から1年以上経過 ⇒ プロジェクトは放棄された

コア開発者の定義

プロジェクト内のファイルのうち、50%以上のファイルでDOA値が最も高い

DOA (degree of authorship) 値^[3]：そのファイルの主な編集者かどうかを示す度合い

[3] G. Avelino, L. Passos, A. C. Hora, and M. T. Valente, “A novel approach for estimating truck factors,” in ICPC, 2016, pp. 1–10.

4. 人気プロジェクトの定量的分析 | 結果 (1/2)

RQ1: GitHubプロジェクトではコア開発者の離脱はどの程度起こるのか

315のプロジェクト(16%)が1度はコア開発者がいなくなる

RQ2: プロジェクトが放棄された後、プロジェクトに新たなコア開発者が現れることはあるのか

315中128のプロジェクト(41%)に新たなコア開発者が現れる

新たなコア開発者のうち41%は
プロジェクトが放棄されてから開発に参加した

4. 人気プロジェクトの定量的分析 | 結果 (2/2)

RQ3：放棄されるプロジェクトと存続するプロジェクトの違いは何か

項目	放棄の中央値	存続の中央値	p値	効果量
開発者数	47[人]	32[人]	0.000	0.26 (小程度)
コミット数	694[件]	384[件]	0.000	0.25 (小程度)
ファイル数	85[件]	54[件]	0.047	0.13 (無視できる)
開発期間	1,460[日]	1,095[日]	0.000	0.37 (中程度)

放棄されるプロジェクトの方が値が大きい傾向
↓
放棄されるプロジェクトは機能が完成している？

5．存続するプロジェクトのコア開発者をアンケート | 方法

調査対象の開発者

プロジェクトが放棄された後，新たにコア開発者が現れた
128プロジェクトのコア開発者

アンケート方法

メールアドレスが公開されている144名の開発者に
自由回答形式のアンケートを送信 → 33人 (24%) の回答

5．存続するプロジェクトのコア開発者をアンケート | 結果

RQ4：新たなコア開発者は，プロジェクトが放棄される可能性を認識していたか

Yes	Partially	No	Unclear
18 (60%)	5 (17%)	6 (20%)	1 (3%)

77%の新たなコア開発者は，コア開発者になる前に
プロジェクトが放棄される可能性を認識

5．存続するプロジェクトのコア開発者をアンケート | 結果

RQ5：コア開発者が離脱した後，新たなコア開発者がプロジェクトを引き受ける
動機は何か

Motivations	Devs	%
Because I was using the project	17	53
To contribute to an open source project	11	34
To avoid the project discontinuation	5	16
I have interest on the project area	4	13
I get paid to contribute	4	13
To improve my own skills	3	9
I have the skills required by this project	3	9
It is a successful project	3	9
Others	7	22

自分が使っているから

プロジェクトを
存続させたいから

5．存続するプロジェクトのコア開発者をアンケート | 結果

RQ6：新たなコア開発者が作業しやすくなる，または作業の障壁となる特性は何か

Type	Characteristics	Devs	%
Human/ Social	Friendly and active owners/members	12	41
	I liked/knew the project	3	10
Technical	Programming language	4	14
	Well-known SE principles	4	14
	Pull based development	4	14
	Continuous integration	2	7
	Clean and well-designed code	2	7
	Code revision	1	3
Others	Main repository access	3	10
	Job support	2	7
	Small or simple project	2	7
	Open source license	1	3

作業しやすくなる特性

親切で積極的な
メンバーがいること

5. 存続するプロジェクトのコア開発者をアンケート | 結果

RQ6：新たなコア開発者が作業しやすくなる，または作業の障壁となる特性は何か

Type	Barriers	Devs	%
Human/ Social	Lack of time	7	26
	Lack of experience	3	11
	Unfriendly maintainers	2	7
Technical	Need to keep backward compatibility	4	15
	Lack of well-known SE principles	1	4
Others	Lack of access to the main repository	5	19
	Large number of pending issues	3	11
	No financial support	2	7
No barriers	-	4	15

作業の障壁となる特性

時間がない

メインリポジトリへの
アクセス権がない

6. まとめ

人気のオープンソースプロジェクトでも
16%ものプロジェクトが放棄される (RQ1)

しかし

新たなコア開発者がプロジェクトに
参加してくれることもある (41%) (RQ2)

新たなコア開発者を引き入れるためには

親切で積極的なメンバーがいることよい (RQ6)