

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

Guilherme Avelino\*, Eleni Constantinou†, Marco Tulio Valente‡ and Alexander Serebrenik†

\*Federal University of Piauí, Brazil, gaa@ufpi.edu.br,

†Eindhoven University of Technology, Netherlands, {e.constantinou,a.serebrenik}@tue.nl,

‡Federal University of Minas Gerais, Brazil, mtov@dcc.ufmg.br

**Abstract**—The loss of core developers might be detrimental for projects. However, it is possible that new core developers assume the project maintenance and allow the project to survive. The objective of this paper is to provide empirical evidence on project abandonment and survival and the motivation and difficulties faced when assuming an abandoned project. We adopt a mixed-methods approach to investigate project abandonment and survival on 1,932 popular GitHub projects and conduct a survey with developers that have been instrumental in the survival of the projects.

## I. REPORT

An easy way to communicate and understand the dependency of an open source software (OSS) project on key developers is the notion of Truck Factor (TF), i.e., the minimal number of developers that the project depends on for its maintenance and evolution [1]; if TF developers abandon the project the project maintenance will be heavily affected. We refer to the situation of TF developers abandoning a project as *TF developers detachment* (TFDD). Existing studies investigated contributor abandonment and retention [2]–[4], and proposed heuristics to compute TFs by mining software repositories [5], [6]. However, there are no studies that allow a more profound understanding of what happens when influential TF developers leave a project.

In our paper [7], we investigate TFDD by conducting a mixed-methods study. We start by collecting, curating, and analyzing a dataset of 1,932 popular GitHub projects to answer RQ1–RQ3 concerning the frequency of TFDDs. Next, we utilize our survey results to answer RQ4–RQ6 concerning the motivation and challenges faced by the developers who assumed the maintenance of abandoned projects.

*(RQ1) How common are TFDDs in open source projects?* TFDD is not merely a theoretical concept: 16% of the projects faced at least one TFDD; 66% of these TFDDs happened in systems with TF=1, which are 55% of the projects. 59% of the TFDDs happened in the first two years of development; but 71% of the projects with TFDDs have now between 4 and 7 years of development.

*(RQ2) How often open source projects survive TFDDs?* 41% of the projects survived their last observed TFDD, usually by attracting a single new TF developer (86%). Newcomers are crucial to recover from TFDDs. They contributed to recovery of 48% of the surviving projects.

*(RQ3) What are the distinguishing characteristics of the surviving projects?* At the moment of the TFDDs, we found

no major difference between surviving and non-surviving projects, in terms of number of developers, commits, and files. On the contrary, we found that surviving projects are younger at TFDD time compared to the non-surviving ones.

*(RQ4) Do new TF developers perceive risks of project discontinuation?* 77% of the new TF developers were (partially) aware of the risks faced by the surviving systems, before making the contributions responsible for the project recovery.

*(RQ5) What motivates a developer to assume an open source project after a TFDD situation?* The developers responsible to reactivate the maintenance of the surviving projects were motivated by their own usage of the projects (17 developers, 53%). They also intended to contribute back to an open source community (34%) or avoid the project discontinuation (16%).

*(RQ6) What project characteristics most facilitate or hamper the work of recently arrived TF developers?* The characteristics that helped on the attraction of new TF developers have a social, technical or external nature. Friendly and active maintainers is the most mentioned facilitator, indicated by 12 developers (41%). Lack of time is the most common barrier faced by new TF developers.

Our work shows that TFDDs indeed happen in open source projects, but projects can survive such situations by attracting new core contributors. Also, the motivations, principal enablers and barriers faced by developers taking over abandoned projects have been investigated by this work. During the workshop, we will present our work, and discuss how our results can become actionable for researchers and OSS.

## REFERENCES

- [1] L. Williams and R. Kessler, *Pair Programming Illuminated*. Addison Wesley, 2003.
- [2] B. Lin, G. Robles, and A. Serebrenik, “Developer turnover in global, industrial open source projects: Insights from applying survival analysis,” in *ICGSE*, 2017, pp. 66–75.
- [3] E. Constantinou and T. Mens, “Socio-technical evolution of the Ruby ecosystem in GitHub,” in *SANER*, 2017, pp. 34–44.
- [4] —, “An empirical comparison of developer retention in the rubygems and npm software ecosystems,” *Innovations in Systems and Software Engineering*, vol. 13, no. 2-3, pp. 101–115, 2017.
- [5] V. Cosentino, J. L. C. Izquierdo, and J. Cabot, “Assessing the bus factor of Git repositories,” in *SANER*, 2015, pp. 499–503.
- [6] G. Avelino, L. Passos, A. C. Hora, and M. T. Valente, “A novel approach for estimating truck factors,” in *ICPC*, 2016, pp. 1–10.
- [7] G. Avelino, E. Constantinou, M. T. Valente, and A. Serebrenik, “On the abandonment and survival of open source projects: An empirical investigation,” *2019 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM)*, pp. 1–12, 2019.