

Investigating Relationships Between FLOSS Foundations and FLOSS Projects

Juho Lindman¹ and Imed Hammouda^{2(✉)}

¹ Applied IT, University of Gothenburg, Gothenburg, Sweden
juho.lindman@ait.gu.se

² Department of Computer Science and Engineering, Chalmers and University of Gothenburg, Gothenburg, Sweden
hammouda@chalmers.se

Abstract. Foundations function as vital institutional support infrastructures for many of the most successful open source projects, but the role of these support entities remains an understudied phenomenon in FLOSS research. Drawing on Open Hub (formerly known as Ohloh) data, this paper empirically investigates the different ways these entities support projects and interact with different projects and with each other.

Keywords: Open source · Open source foundations

1 Introduction

Continuing Free/Libre Open Source Software’s (FLOSS) success is based on the evolution of FLOSS projects and contributors [1–5], but there is a research gap concerning the entities¹ that support FLOSS, such as foundations [6]. These entities support individual FLOSS projects in different ways, but the dynamics they are engaged in remains an understudied phenomenon. In addition, the cooperation of these entities and between developers poses several questions for further study.

We address these gaps in our empirical investigation of how these entities support and interact based on Open Hub data. In detail, our research question: *How FLOSS entities support FLOSS projects?* Our findings reveal traces of a complex interplay in this ecosystem when describing this dynamic.

The paper is organized as follows: Section 2 gives background on earlier related research. Section 3 presents the methodological details of data collection and analysis. Findings are then reported and discussed in Sect. 4, followed by Sect. 5, Discussion. Finally, future research directions and concluding remarks are presented in Sect. 6.

¹ In this work we call these “support entities” noting that in many cases “foundation” would also be applicable. However, we note that (1) not all of these entities are foundations and (2) even if they were, there are subtle differences regarding their legal and tax status in different jurisdictions. Thus, we limit these legal considerations outside the scope of this study and call these support entities.

2 Background

FLOSS foundations support their community projects in different ways. We explore FLOSS support entities (“Foundations,” “Organizations”) and their relationships to the projects they are affiliated with. Riehle [6] demonstrates how FLOSS support entities manage and ensure the long-term survival of their projects.

These entities are linked to projects and help by providing financial support and legal assurance. This makes the FLOSS projects a bit less dependent on the volunteer efforts. In addition, FLOSS support entities have other responsibilities related to the hosting and management of the FLOSS projects. Responsibilities include (i) organizing community projects (ii) marketing, (iii) managing intellectual property (IP) rights and (iv) setting strategic directions. Support entities may also provide means to protect community-generated content using IP legislation [6].

In this paper, we investigate empirically the FLOSS environment, the role of the supporting entities and the relationships between support entities.

3 Methodology

This study was conducted by using **Theoretical Saturation Grounded Theory** approach which is a form of a qualitative data collection and data analysis methodology. According to [7], theoretical saturation is associated with theoretical sampling for grounded theory. A grounded theory is a scientific research approach used by the researchers for the collection and analysis of qualitative data. The main purpose of choosing this research approach is to develop a theory (or) a model through a continuous comparative analysis of qualitative data collected by theoretical sampling process. This flexible research approach is required to collect huge volume of data because, data collection will be done simultaneously along with the data analysis process. A theory (or) a model can be formulated from the collected data. This research approach is also used to assess any sort of patterns (or) variations out of an investigated research area. The selection of cases during this research process will most likely produce the most relevant data that will evaluate emerging theories. However, each new case might offer a slightly different outcome. The researcher will be having a continued sampling of data and he/she will analyze the data until no new data emerges. The end point of theoretical saturation indicates that, the approach has reached a point where no new data were identified and it shows the researcher that the enough data were collected for data analysis purposes.

3.1 Data Collection

The **Open Hub data repository** (formerly known as Ohloh) is used as a primary data source for this study. This source holds key information about the support organizations concerning their sectors, development focuses, licensing policies, membership types and structure. The data repository also holds other information, such as projects and a committer’s list, which can be used to determine the relationships between support

entities and projects. Open Hub can be accessed using their **API keys** [8]. We use this repository to identify the relationship a support entity can have with another entity and a support entity’s portfolio projects. We used Open Hub data from all FLOSS support entities that host at least one project.

Support entity websites are another main source of data. These websites hold key information about support and services, incubation processes, project governance, maintenance, project development practices, IP management, license agreement policies, hosting services and so on. This information is used to map out how the entities provide support for projects.

3.2 Data Analyses

We used a Java program to parse the API data from the XML data format to plain text and stored it in a database. We collected data from 88 FLOSS support entities (“Organizations”). We have set a criterion to analyze our sampling cases (i.e. data) that we collected from the support entities. Our criterion for data analysis is that, if we go through 20 sampling cases without no new data/findings, then it is our saturation point.

In the generated database, we identified whether entities with unique organization IDs have (1) connections with projects affiliated with other entities and (2) whether an organization’s affiliated developers contribute to the projects of other entities. We also noted when developers contributed to some other projects (e.g., to their own projects). These different scenarios are described in Fig. 1 below.

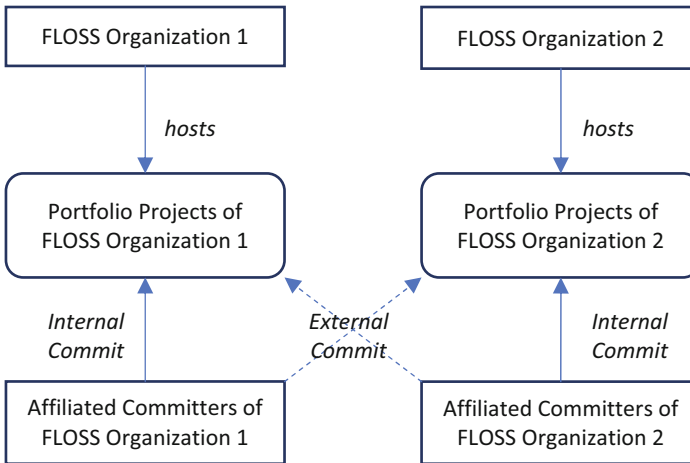


Fig. 1. The relationships between different FLOSS organizations.

We then used a manual approach to search for the appropriate information through relevant online sources—such as foundation websites and forums—to describe the identified relationships between the FLOSS organizations. We investigated each relationship between any two support entities within the FLOSS relationship network.

Finally, we qualitatively analyzed the identified details of the relationships and grouped them.

4 Findings

We defined a FLOSS support entity taxonomy (Fig. 2) that describes some of support entities' characteristics. We then outlined the data of the different characteristics available.

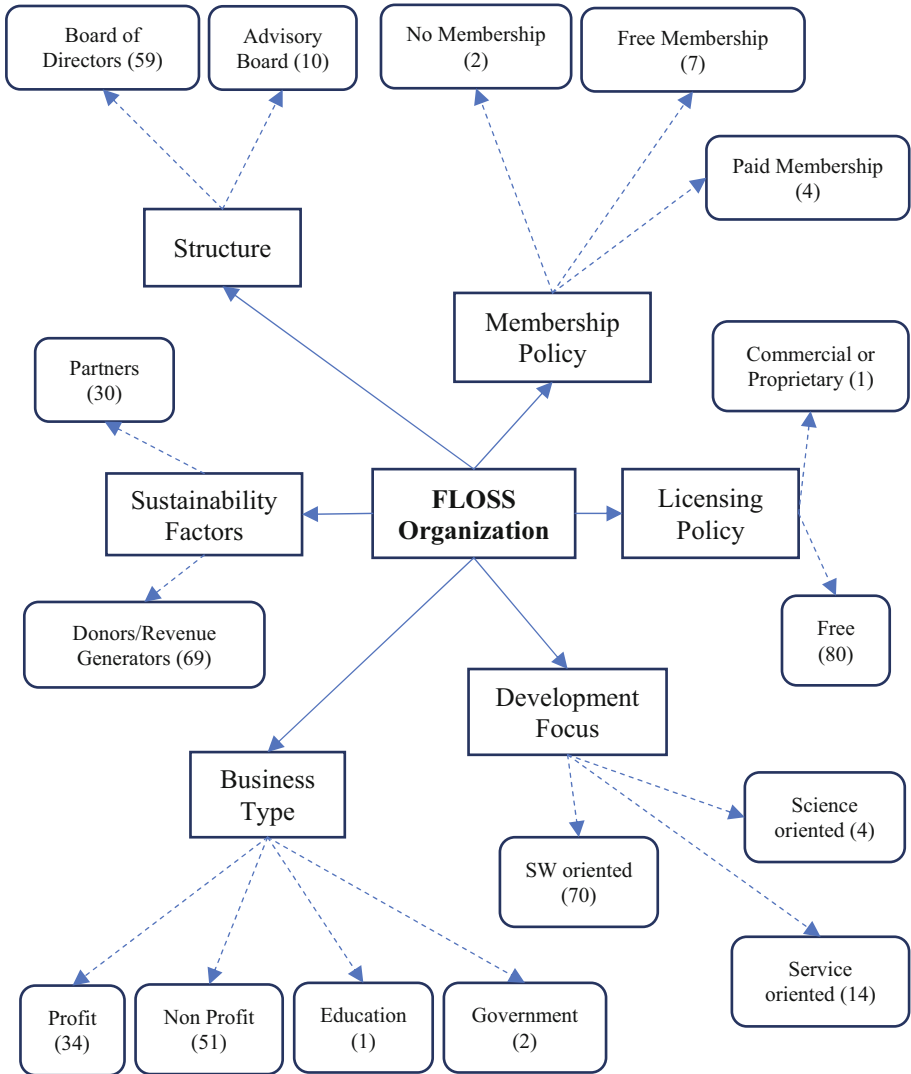


Fig. 2. FLOSS support entity taxonomy.

A **Profit (or) Commercial FLOSS** support entity generates revenue via sales of products, services and solutions. They collaborate with different corporations and technical partners. **Nonprofit** foundations are primarily sustained through volunteer donations. They collaborate with external companies, educational institutions and other stakeholders to get funds to support the projects. Most of these organizations are also primarily governed by the Board of Directors (BOD). **Government FLOSS** mostly consists of science-related projects. The funding for such projects comes mainly from public sources. **Education FLOSS** comprises primarily educational institutions. These support entities mainly focus on providing education to the general public and are sustained through donations from public sources and student fees.

Organizations list different development focuses. Options include **S/W orientated**, **service orientated** and **science orientated**. Most service-orientated support entities were of the profit type, while science-orientated support entities were often government based and educational.

FLOSS entities may support projects that use either **free software license** projects or **commercial or proprietary software license** projects. A free software license allows the user of a piece of software the extensive rights to modify and redistribute that software. A commercial or proprietary software license is produced for sale or to serve commercial purposes.

FLOSS support entities evolve through different kinds of **donors and revenue generators** and **partners**, such as volunteers, corporations, open-source organizations, software products, government agencies, educational institutes and investors.

FLOSS organizations are governed by two different governance modes: the (BOD) and the **Advisory Board (AB)**. The BOD has the decision-making authority and responsibility for governing the support entity. BOD committee roles may include Founder, Investor and Director. In contrast, the AB does not have the decision-making authority, and they are only responsible for assisting or giving advice within the organization. AB committee members can have roles like Senior Manager, Executive, Volunteer and so on.

FLOSS organizations have different types of membership schemes. The **no membership (NM)** type does not have any members within the support entity. The **free membership (FM)** type allows any members to join without any membership fee. The **paid membership (PM)** type allows only the paid members to take part.

5 Discussion

To answer our research question (*How FLOSS entities support FLOSS projects?*), we explored how entities support FLOSS projects. We grouped our findings (described in detail in Table 1 below): services, incubation process, project governance, project maintenance, IP, project acceptance and hosting services. Table 1 summarizes the key support mechanisms.

Table 1. How foundations support FLOSS projects.

Support	Explanations
Services	FLOSS support entities can provide legal, financial and consulting services to their projects. Support entities can provide tools and offer advice on how to raise funds. Support entities can also provide essential support on how to protect the IP and financial contributions, and it can limit the legal exposure of an individual contributor in portfolio projects; for example, ASF and Gentoo
Incubation process	Support entities have different guidelines on how a portfolio project can be created. Many support entities require an incubation process. Created projects enter the incubation process. Some of the processes are mandatory quality control mechanisms. In some FLOSS organizations, incubation processes are used to create new versions of the existing projects and not for creating new projects. Some FLOSS projects start with pre-existing code before they go through the incubation process. These incubation processes are useful for new projects in learning community norms and processes. Projects in incubation will be monitored by the nominated mentors There are some variations: <ul style="list-style-type: none"> • The incubation process is only used to create the new versions of an existing project and not for creating entirely new projects; for example, the Wikimedia Foundation • Individuals are responsible for the creation of projects. However, under the Eclipse Foundation, a project can be started/created with some pre-existing code • A project can be started/created by anyone with the necessary skills
Project governance	Support entities may assign a project management committee (PMC) consisting of people to govern or manage projects and subprojects. Support entity mentors usually work with PMC to help in the evolution of the project; for example, ASF and Tryton
Project maintenance	Project data are maintained either by a PMC or by projects; for example, ASF
Intellectual Property (IP)	FLOSS support entities' IP management enables the participation of software developers from different organizations to develop software. Tried-and-true practices exist to support software IP management and to foster a growing community. FLOSS organizations protect the developer's contribution to portfolio projects when the developer signs a Contributor License Agreement (CLA). The CLA is specially designed to protect the developer's contribution. Organizations usually do not protect the hosted projects managed by third parties with the CLA; for example, Outercurve Foundation, Eclipse and Gentoo <ul style="list-style-type: none"> • A project might receive organization IP clearance for contributions and third-party libraries • IP management enables and encourages the participation of organization software developers to develop software collaboratively in a FLOSS community • When a CLA is signed by the developers, the entity protects the contributions on its portfolio projects; for example, Twitter and 52 NIFGOSS • However, third parties managing the hosted projects within the entity are not protected by the CLA
Project acceptance	Projects need to be championed by a sponsor (i.e., if the sponsor is the foundation board); for example, Outercurve Foundation
Hosting services	Organizations provide project hosting services and tools to promote FLOSS development; for example, OSGeo and Genivi Alliance <ul style="list-style-type: none"> • The support entity hosts projects and a wide variety of other mailing lists for projects, committees and special interest groups

Table 1 shows that FLOSS organizations like ASF, Gentoo and SpringSource provide various support and services to portfolio projects. Organization incubation processes are used in ASF, Wikimedia Foundation, Eclipse Foundation and the MirOS project. Foundations such as ASF and Tryton assigns a PMC to govern their projects. Some foundations, such as KDE, have limited hierarchical structures. Some support entities like the Outercurve Foundation, Eclipse and Gentoo and own IP rights to protect their portfolio projects while restricting their contributors.

Based on our qualitative analyses, we list the identified reasons that describe why the support entities interact. Two FLOSS support entities can have a relationship because of the following key reasons: plugins, sponsorship, tie-ups, packages, reliance, key persons and hosting (see Table 2 for more detailed descriptions).

Table 2. The relationship between two support entities.

Plugins/Add-ons	A FLOSS support entity may provide or produce plugins/add-ons to other FLOSS support entity projects and their produces; for example, the Xfce desktop provides add-on to Mozilla's Thunderbird application
Sponsorship	A FLOSS support entity may provide funding or sponsorship their contributors to other FLOSS support entities and portfolio projects; for example, Twitter provides financial funding and contributes to the Apache Software Foundation. Yahoo also provides financial funding to the OpenStack Foundation
Tie-up	FLOSS project software might have a tie-up with other FLOSS organizations' software. The Xfce and KDE desktops have tie-ups with Debian operating system
Packages	A FLOSS support entity may provide packages for other FLOSS products and services. For an example, Homebrew provides the packages for KDE desktop applications to install on OS X. Homebrew also provides packages to Mozilla's add-ons on OS X
Reliance	A FLOSS support entity might be using other FLOSS organization software, services, infrastructure, tools or products for its own business operations and services; for example, Sony Mobile and Yahoo are using the OpenStack platform infrastructure for their business purposes
Key person	A key person—such as the founder, lead developer, maintainer or manager—from one FLOSS support entity might be employed by other FLOSS foundations. Both FLOSS entities might have a single person as a common manager to manage FLOSS projects; that is, a single person acts as a manager for both organizations' projects. For example, Tarent solutions Gmbh and the MirOS project have a single person managing their projects and the same person is the founder of the MirOS project and is employed by Tarent Solutions Gmbh
Hosting	A FLOSS support organization might host and distribute other FLOSS organizations' products and services; for example, BlackBerry hosts and distributes Adobe apps on BlackBerry World to BlackBerry mobiles. A FLOSS organization may provide generic modules and functions to work with other FLOSS organizations' software implementations; for example, SaltStack is providing generic modules and functions to work with the Apache Software Foundation implementation

From the collected API data, we explored the relationship between two FLOSS organizations. We could not find any projects hosted under or claimed by multiple organizations as a portfolio project. We also explored whether there are relationships among different FLOSS support entities.

Based on the collected data, different FLOSS organizations have relationships when affiliated developers from one FLOSS organization contributes to other FLOSS entities' portfolio projects. As this study mainly focuses on the support entities, we did not consider the individual project information in-depth that could give more insight regarding specific projects and their committers.

6 Conclusions and Future Avenues for Research

This research study investigated FLOSS support entities, their role in FLOSS projects and the relationships among them within the FLOSS ecosystem. Based on our findings, we claim that our proposed methodology could identify the key attributes and values of a FLOSS support entity through a developed taxonomy and the FLOSS organizations key roles in FLOSS projects.

This research opens several new areas for further research. There are interesting research opportunities related to verifying and measuring the impacts of developer contribution and entities. Our methodology focuses on chosen parts of the interplay between the support entities, so we expect future studies to shed more light on their important and understudied role in supporting and governing FLOSS.

Acknowledgements. The authors would like to thank Bharat Kumar Mendu and Joshua Smith Soundararajan for their contribution to this research.

References

1. Godfrey, M., Tu, M.: Growth, evolution, and structural change in open source software. In: Proceedings of the 4th International Workshop on Principles of Software Evolution, pp. 103–106. ACM Press (2001)
2. Robles, G., Amor, J.J., Gonzalez Barahona, J.M., Herraiz, I.: Evolution and growth in large libre software projects. In: Proceedings of the Eighth International Workshop on Principles of Software Evolution (IWPSE 2005), pp. 165–174. IEEE Computer Society (2005)
3. Roy, C.K., Cordy, J.R.: Evaluating the evolution of small scale open source software systems. In: Proceedings of the 15th International Conference on Computing, Research in Computing Science (2006). Special issue on CIC
4. Succi, G., Paulson, J., Eberlein, A.: Preliminary results from an empirical study on the growth of open source and commercial software products. In: EDSE3 Workshop, pp. 14–15 (2001)
5. Aksulu, A., Wade, M.: A comprehensive review and synthesis of open source research. *J. Assoc. Inf. Syst.* **11**(1), 576–656 (2010)

6. Riehle, D.: The economic case for open source foundations. *Computer* **43**(1), 86–90 (2010)
7. Glaser, B., Strauss, A.L.: *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Transaction Publishers (2009)
8. Gallardo Valencia, R.-E., Tantikul, P., Sim, S.-E.: Searching for Reputable Source Code on the Web. <http://www.drsusansim.org/papers/group2010-gallardo.pdf>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

