

Design method of Open Source Software

Hiyani Goyal 1[†] and Jasbir Kour 2^{††},

BMS College of Engineering, Punjab, Indian

Summary

Free/Open Source software is a kind of software whose source code is available for comprehension, modification and re-distribution. This kind of software has increased in popularity in recent years and becoming an interesting topic for research. Most Free/Open Source software is produced through the facilitation of Free/Open Source Hosting (FOSPHost) sites and investigations into these sites may yield results that have theoretical and practical significance. Open source software is becoming the most interesting 'new' phenomenon of the entire information technology landscape, generating a level of interest similar to that of the first moments of the Internet. The study of the possibilities and limits of open source software at the enterprise is the main concern. The research work analyzes the history of the open source movement, describes the open source community and collaboration model, analyzes the open source development process, describes business models based on open source software, analyzes possible cost savings and presents case studies of popular open source projects. Recommendations are presented, how companies and organizations might benefit from open source software and in which cases it should be avoided, because the hidden costs will not pay off the license costs savings. The impact of open source technology is expected to be quite noticeable in the software industry, and in society as a whole. It allows for novel development models, which have already been demonstrated to be especially well suited to efficiently take advantage of the work of developers spread across all corners of the planet.

Key Words:

Community, Descriptive, Explanatory, Exploratory and Open Source.

1. Introduction

The Free/Open Source phenomenon is a surprise with a mystery. The market share of a popular Free/Open Source web server, Apache, was 69% comparing to 23% for Microsoft servers in January 2004[1]. In the operating system market at the end of 2001, Linux server, a Free/Open Source system, had 26% while Microsoft had 49% of the market share. Microsoft was still the leader of the market, but 45% of all new servers shipped were predicted to be Linux in the year of 2006 or 2007[2]. Another survey undertaken by a magazine for IT managers using Microsoft servers showed that two out of five enterprises also employed Linux. More than 800 enterprises were surveyed with an average number of servers running in these companies of 400 [3]. Though a number of companies such as IBM and HP now support Linux

development as a strategy to combat Microsoft, the idea of Linux is owned or controlled by neither of these companies defies common business logic. To explain simply, Free/Open Source software is a piece of software whose source code is made freely available. Source code is the original form of a computer program as written by the programmer [4].

A number of Free/Open Source communities participate and shape political movement online. Some people also have been trying to apply the idea of Free/Open Source in other areas such as education and even forestry management. Therefore, in order not to lead readers to focus only on software or software development, the author will use a broader term 'the Free/Open Source phenomenon' to refer to what has happened so far in a broader context. The reader may wonder why the term 'Free/Open Source' is used to qualify software that the source code is made freely available in this study, rather than the more commonly used term, 'Open Source'. 'Free/Open Source' is a combination of the terms 'Free Software' and 'Open Source'. The term 'Free Software' is promoted by the Free Software Foundation, which advocates Free Software as a social movement that non-Free Software is morally wrong[5]. On the other hand, the term 'Open Source' is promoted by the Open Source Initiative, which advocates the practical benefit of Open Source software development to the commercial world. These two views are both relevant and thus the term 'Free/Open Source' is used. The author here maintains a political view that is neutral to both movements. The Free/Open Source phenomenon has the potential to attract the attention of the academic circle, as there are a number of issues that require explanations. First, it is hard to reconcile that the cost of the development of some highly complex Free/Open Source projects can be so low. For example, Red Hat Linux 7.1 was estimated to cost more than one billion US dollars to develop using conventional software development approach. Significant monetary investment towards Linux is only a recent phenomenon and thus the estimation above was huge discrepancy with the reality[6]. In Raymond's it was explained that the development of Linux showed us how a collective effort of co-developers over the Internet (the Bazaar) could possibly produce quality software with better reliability and more useful features in a shorter time[6]. A FOSPHost site is the infrastructure that supports and co-ordinates the development of Free/Open Source software projects on the

Internet. In short, Free/Open Source developers collaborate through the FOSPHost sites to produce Free/Open Source software [7].

2. Problem Definition

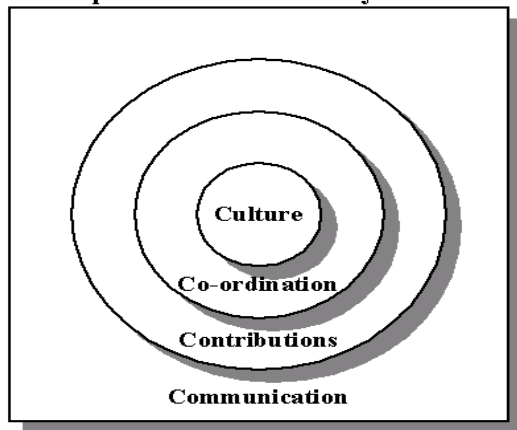
This was a survey of about the practice in open source development, with particular emphasis on the modular extensibility interfaces within several of the most successful projects, including Apache, Eclipse, Mozilla Firefox, Linux kernel, and the World Wide Web and many other scientific research related software.

The problem was to study the design methodologies of the open source development. This problem was studied by comparing of four models of development of open source software given Nakakoji; Gacek, Lawrie & Arief; Sharma, Sugumaran & Rajagopalan; Feller & Fitzgerald. Exploratory study was conducted to explain the Free/Open Source Software Development phenomenon.

Free/Open Source Community's 4C Model :

Based on the four important aspects identified in a Free/Open Source community, a model of a Free/Open Source community is built and shown in Figure 1.1.

Free/Open Source Community



The model is presented in a four-layer (4C) model.

Figure 1.1: Free/Open Source Community's 4C Model

The four layers represented in the model in Figure 1.1 are communication, contributions, co-ordination and culture respectively. The communication medium is the basic infrastructure for any interaction. Contributions referred to the different pieces of assistance given by individual developers via the communication media. Co-ordination is the process of organizing fragments of contributions into

usable products and the culture of the community in turn governs the rules in co-ordination.

A Model of Individual Participation to a Free/Open Source Community:

After introducing a model to a Free/Open Source community, one can consider to represent the relationship of individual participants to the community by a model. Individual participants, who are probably one of the most influential groups on the assessment of FOSPHost, is chosen. Other stakeholders such as user communities, commercial organizations, and the non-commercial organizations that managed Free/Open Source projects are excluded to limit the scope of investigation.

The model built to explain this relationship is shown in Figure 1.2. The model includes the mentioned 4C model, the motivations and barriers when a developer decides to join a Free/Open Source community together with the positive and negative results after interaction with a Free/Open Source community. The motivations and barriers are analogous to the "variables which affect individuals' decision to join virtual communities" and the results analogous to the effects from the three phase model on virtual communities. Since the group of individual participants is chosen, all these four factors are related just to them and a feedback loop is included as well.

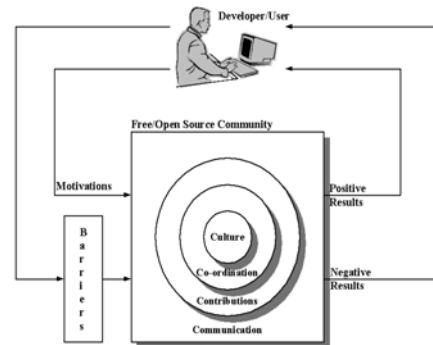


Figure 1.2: A Model on individual participation in an Open Source/Free Software Community

There are a number of motivations for a user or a developer to join a Free/Open Source Software community. An oftenly regarded motivation was stated in Raymond's 'the Cathedral and the Bazaar' - 'Every good work of software starts by scratching a developer's personal itch.' This essentially means that a developer needs a computer program to do a task for him or her. However, this need does not necessarily lead to joining a Free/Open Source community. The most common example is a developer needs a new PC to work so this person installs a copy of Microsoft Windows. Alternatively, a developer may write a

piece of software to meet his or her need but the source code of the software may never be shared. Therefore, when a developer joins a Free/Open Source community, he or she may be motivated by other factors also, such as reciprocal behavior, reputation and attraction to community.

Availability of funding also enables members of Free/Open Source community to work on project devotedly such as support in BSD by DARPA and Linux by University of Helsinki. Lastly, altruism or idealism may also motivate developers to contribute. Although there are a number of motivations for developers to join a Free/Open Source community, barriers also exist to deter them, as in any virtual communities. Technically, Free/Open Source communities only accept developers who attain a high degree of competence. The complexity of source code also created a barrier for contribution. On the other hand, software with poor design and inadequate documentation may deter contribution. Another barrier is that a developer may not be willing to share his or her own code. Cultural barriers may also exist. Firstly, language can be a barrier because people from certain backgrounds in some part of the world may find it hard to join a Free/Open Source community using English as the common language of communication. Cultural mysteries also exist and they have to be solved before a member could be accepted by certain Free/Open Source communities. The last but obvious reason is that a developer cannot afford the time for one's involvement in a Free/Open Source community. There are several positive outcomes as a result of joining a Free/Open Source community. A developer may have one's own itch scratched and found that he or she enjoyed programming in collaboration. He or she may learn more skills and build up one's own reputation in the community as well.

Negative results from participation in a Free/Open Source community may include a lack of interest on one's project, rejection from others, hurts in management issues and burn-out. An example of the model can be that a computer literate required a certain application to fulfil her needs. She found a piece of Free/Open Source software (positive result) and added some modifications to fulfil her needs more comprehensively. She then tried to contribute the code back to the community but she found the code had to conform to the coding standard (barrier) and the core members of the project were not too friendly (negative result). Later on, a new version of the software was released with new features but not compatible with her modifications.

It was a nuisance that she would need to adjust the modifications for each release. Then, she finally got her code to conform to the standard (motivation). Also, she was no longer new to the community and knew the core members better. Her modification was eventually accepted and it stayed in the code for the versions to come (positive result). The burden of maintenance was therefore shared

(positive result). The Model of Individual Participation to a Free/Open Source Community and FOSPHost Design and Deployment.

After the development of the analytical frameworks, namely the 4C model and the model of individual participation in a Free/Open Source community. Recalling that 4C model of a Free/Open Source community consisted of communication, contributions, co-ordination and culture, a FOSPHost site is the communication tool that holds the contributions of the community. A FOSPHost site indeed creates a basis for the existence of a community. Moreover, the model of individual participation in a Free/Open Source community suggests that the important issues in improving a FOSPHost site are how well does a FOSPHost site support collection of contributions, co-ordinations of project(s) and cultivate a constructive culture for community. Other important issues include how the design of FOSPHost motivates users to participate and maximises positive results. On the other hand, barriers of participation should be lowered and negative results should also be minimised. From the derivation above, the models thus suggested distinct focuses on how the study should proceed. The issues obtained above will be the starting point for the data collection stage

3. Design Method of Oss

This investigation in external hosting sites and finally the construction of an evaluation model for FOSPHost. An exploratory approach was taken in this research. Moreover, the conclusion of this research will be constructed from the empirical data collected, and thus an inductive approach was also adopted. One way to classify social research is by the purpose of study. There are mainly three types of purposes, namely exploration, description and explanation. Exploratory studies are conducted to learn more about topics that are little known to construct mental pictures based on basic facts and stakeholders. Descriptive studies are conducted to observe and describe details of social phenomena. Explanation studies are conducted to verify certain theory on the relationships of different variables in a system.

When a new topic is studied, the sequence for three types of study to be executed would be exploration, description and finally explanation. This concept is illustrated. The knowledge that is unknown is denoted as a cube in red, the area where it is known is denoted by white. For exploratory research, it is like increasing the white area of known knowledge on the surface of the cube. In descriptive research, it is to increase the depth of known knowledge, based on the results from previous exploratory research. Explanation studies are done last as substantial understanding of the topic was required before formulating theories about the topic.

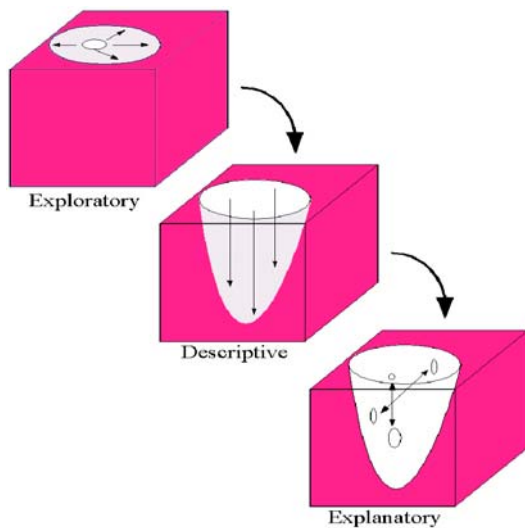


Figure 1.3 Exploratory, Descriptive and Explanatory
Research

This illustration in one sense is not totally accurate as the results from each type of study are probably not mutually exclusive. For example, during an exploratory research, the result probably will have some depth. Casual relationships of elements within the topic may already be partly confirmed. Moreover, there is always more to discover even on a well-known topic, so using the idea of using white to denote known knowledge in a certain area can be misleading. Nevertheless, it may help to understand the underlying principle of purposes of research. As mentioned above the amount of literature on the Free/Open Source phenomenon was not sufficient to form a comprehensive explanation. Therefore, an exploratory approach was adopted. In exploratory research, suggested that the researcher 'must be creative, open minded, and flexible; adopt an investigative stance; and explore all sources of information.' The disadvantages of exploration studies are the conclusion yielded may not be definitive and the representative ness of result may be weaker.

Another choice of the research strategy in this study was between deductive and inductive approach. In a deductive approach, a hypothesis is formulated from pre-existing theoretical framework and empirical data is collected to prove or disprove this hypothesis. In an inductive approach, empirical data is collected to build a theoretical framework based on a few initial concepts. Inductive approach was the obvious choice because the amount of pre-existing theoretical framework was not sufficient. The overall research strategy included the relationships between literature and data collections. The initial survey employed the model of individual participation to a Free/Open Source

community as the theoretical basis for the initial questions in the first round of the survey. After the survey, a detailed investigation was done to further collect data on different FOSPHost sites. The literature of methodology and evaluation was referred to in each of the three steps of sub-projects to ensure consistence.

4. Selection of Research Methodologies and Methods

The rationale for choosing an appropriate methodology will be presented and then the choice of research method for each phase of the research will be explained. The word 'method' of research is defined as 'the actual techniques or procedures used to gather and analyse data related to some research question or hypothesis. In contrast, methodology is a more philosophical 'analysis of how research should or does proceed.

All these three methodologies can be relevant to research in the Free/Open Source phenomenon. As the Free/Open Source phenomenon is related to software engineering, and software engineering has its roots in mathematics and science and thus positivism is actually a pre-dominant methodology in Free/Open Source research.

Examples of researches employing positivism are surveys on source code and analysis of statistics from FOSPHost sites. On the other hand, some researchers hoped that interpretive approach could provide a more meaningful description to the chaotic Free/Open Source phenomenon. A discussion in the workshop on 'Advancing the Research Agenda on Free/Open Source Software suggested that one of the methodological directions could be anthropological or even ethnographic in order to gain more insight in the organization of Free/Open Source software development. Ethnographic studies exist but the numbers are few. Lastly, though there are very few researches employing the critical social science approach, it will be interesting to see what insight can a theory of classifying society by the degree of software freedom each class possesses and how the oppressed can be empowered by Free Software. As argued above, all three major methodologies could probably yield interesting results. In this research, however, positivism is chosen, as it is the methodology that the majority of the audience is familiar with. The ontology (or world view) of positivism is that general laws, which are the fundamental operating principles of the world, exist and they are objectively observable. Though each observation of the world is atomic, they are discrete and independent of each other. Conclusions can be drawn from them to discover the basic principle of the world. Truth can thus be found on observations, not unexamined belief or metaphysics. This methodology is also consistent with the inductive strategy that is employed in this research.

One interesting aspect of positivism is that if each observation of the outside world by a subjective human being is regarded as subjective; then how can an objective conclusion be drawn? Objective results can be obtained by drawing conclusions on common patterns from subjective observations collected in a scientific manner, and this is the main methodological philosophy of this research. Obviously, the validity in seeking the truth by employing this view of objectivity can be critiqued, but an in-depth debate is beyond the scope of this research. Though it is not the author's intention to go into methodological debates, but one of the objections to positivism is important enough to be discussed here- relevancy, claimed that there could a danger that 'positivism reduces people to numbers and that its concerns with abstract laws or formulas are not relevant to the actual lives of real people.' As the author would like the final evaluation model to be relevant and useful to the general public, a more lenient approach from the orthodox positivism worldview will be taken when required. Lee suggested that though positivist and interpretive approaches were usually views as irreconcilable and incompatible approaches, it was possible to integrate them and reaped the benefits from both methodologies. As mentioned above that exploratory research required flexibility to construct a richer picture of the situation, interpretive approach will be used when needed to construct meaning to increase relevancy.

5. Result and discussion

Comparison between the Bazaar Model and the Model of Individual Participation to a Free/Open Source Community:

The model of individual participation to a Free/Open Source community presented above covered technical and socio-economical aspects of Free/Open Source as well as context of the community. On the other hand, pointed out that 'the Cathedral and Bazaar' described the process of how to run a Free/Open Source project as a replica of Linux. This focus unfortunately reduces the phenomenon of Free/Open Source into a series of technical processes. This is, however, not to say that Raymond did not know about culture. On the contrary, he was the compiler of 'The New Hacker's Dictionary' . Moreover, in the 'Homesteading the Noosphere, the next essay after 'The Cathedral and Bazaar', he mentioned various aspects of the different sub-cultures within Open Source. Unfortunately, probably in the process of marketing Free Software and by de-politicisation and renaming it to 'Open Source', the complexity of the phenomenon was reduced to technical processes. To conclude, the metaphor of the Cathedral and Bazaar is useful as an introductory, first estimate to the phenomenon

of Free/Open Source but more is needed to explain the phenomenon. The model presented above is one of the many attempts to contribute towards a more comprehensive and complex explanation, which covers contextual, technical and socio-economical aspects.

Comparison between the Other Models and the Model of Individual Participation to a Free/Open Source Community:

Other than the models presented above, researchers around the world also devised different explanations to describe and investigate the Free/Open Source phenomenon. The models to be compared are 'Evolution patterns of Open-Source software systems and communities. Open Source characteristics - common and variable', OSS (Open Source Software) Model and 'A framework analysis of the Open Source software development paradigm. The focus of the first two models was mainly on the software development process and the latter two were attempts to develop a more comprehensive explanation.

Comparison of the Models:

After presenting the four models, the result of the comparison is tabulated in Table 1.1. Most of the explanations from the four models are more elaborate than the model of individual participation to a Free/Open Source community. For example, the motivation categories proposed. There are also areas that are not included in the model of individual participation to a Free/Open Source community such as qualification by the Open Source Definition and stakeholders such as commercial organizations. Nevertheless, the model of individual participation to a Free/Open Source community is yet flexible enough to incorporate most of the materials in the four models. Moreover, less discussed areas such as contributions, barriers and positive and negative results are also included. Also, though most of the content in the four models were based on actual facts, some of the facts might only reflect particulars of certain Free/Open Source communities. In contrast, by being less prescriptive, the model of individual participation to a Free/Open Source community may have the advantage of allowing its users to discover alternatives. Recalling the aim of creating the model of individual participation to a Free/Open Source community is to identify important aspects in a FOSPHost site for further investigation. This aim can be regarded as completed since the model of individual participation to a Free/Open Source community includes most of the important aspects that the four models discussed.

	Communi- cation	Co-ordi- nation	Cultur- e	Motiv- ation	Barrier- s	Positiv- e Results	Nega- tive Result- s
Nakakoji et al. 2002		√	√		√		
Gacek, Lawrie & Arief 2001	√	√	√	√			
Sharma, Sugumaran & Rajagopalan 2002	√	√	√	√	√	√	
Feller & Fitzgerald 2002	√	√	√	√	√	√	

Table 1.1

Moreover, it also includes other significant issues that the four models have less emphasis on. Furthermore, the omission of stakeholders other than developers is favourable as to narrow down the scope of investigation to the most important group of stakeholders. After the comparing the advantages and disadvantages of the four models and the model of individual participation to a Free/Open Source community and reviewing how suitable it is for the investigation, other observations can be discussed. From the analysis above, for comprehensive models, social theories are employed as a basis to derive explanations. Even for software development based models, discussions on social issues on Free/Open Source are included.

This probably suggests the importance of the social aspect the discussion of the, contributions is one of the least discussed topic within the 4Cs. The obvious reason is that many regards contributions to be coding for Free/Open Source software. The user-support as a significant type of contributions and thus conducted a study in Apache mailing-list on the responds of request for user assistance. Gabriel suggested that other contributions such as marketing and standards development were also notable. He further commented that hierarchical analysis of Free/Open Source communities based on authority on code could be misleading. The code development community is just one of the many communities within the Free/Open Source phenomenon and the boundary of a community should be defined by these different kinds of contributions or interests in order to represent their significance. It is then not surprising that the effects of Free/Open Source are less discussed in the models mentioned. In the model of individual participation to a Free/Open Source community, only the effects affecting individuals are mentioned. Also, negative factors such as barriers and negative effects of Free/Open Source are less discussed. Therefore, research in these areas will yield new knowledge. From the models

presented above, pointed out that Free/Open Source projects with different co-ordination models possessed a number of different attributes. Also showed that there were variables between different projects. Moreover, also claimed that there were different practices in different organizations and developers. Indeed, flexibility was also an important consideration in the design of the model of individual participation to a Free/Open Source community. May be this collection of differences and variables are where the chaos of Free/Open Source lies. Therefore, further research on these variables will be profitable. To conclude, after comparing the model of individual participation to a Free/Open Source community with four other models, the quality of the model is acceptable as the basis for this research.

6. Future Work

This study will help the future development of open source software. The compression of the open source software models will further help in the take over process of commercial software by the open source software. Here this work will widen the possibilities of the open source software development in future.

REFERENCES

- [1] Apple Computer Inc. 2002, Darwin - Open Source, <<http://developer.apple.com/darwin/>>.
- [2] Bezroukov, N 2000, 4.1. Linus and Linux; Linus Torvalds' Short Unauthorized Biography, <http://www.softpanorama.org/People/Torvalds/Linus_Torvalds_biography.shtml>
- [3] Bezroukov, N 2002, California Sunrise: Linux Hype Festival and IPO Gold Rush, <http://www.softpanorama.org/People/Torvalds/linus_california_sunrise.shtml>.
- [4] Crowston, K & Scozzi, B 2002, 'Open Source Software Projects as Virtual Organizations: Competency Rallying for Software Development', IEE Proceedings - Software,
- [5] DiBona, C, Ockman, S & Stone, M 1999, 'Appendix A: The Tanenbaum-Torvalds Debate', in C DiBona, S Ockman & M Stone (eds), Open Sources: Voices from the Open Source Revolution, Eunice, J 1998a, Beyond the Cathedral, Beyond the Bazaar, <<http://www.illuminata.com/public/content/cathedral/intro.htm>>.
- [6] Eunice, J 1998b, Beyond the Cathedral, Beyond the Bazaar - The Key is Community, <<http://www.illuminata.com/public/content/cathedral/cathedral5.htm>>.
- [7] Hamerly, J, Paquin, T & Walton, S 1999, 'Freeing the Source: The Story of Mozilla', in C DiBona, S Ockman & M Stone (eds), Open Sources: Voices from the Open Source Revolution, O'Reilly & Associates, Sebastopol, California, <<http://www.oreilly.com/catalog/opensources/book/netrev.html>>.