Challenges in Using Open Source Software in Product Development: A Review of the Literature

Klaas-Jan Stol
Lero, the Irish Software Engineering Research Centre
University of Limerick
Limerick, Ireland
+353 61 23 3737
klaas-jan.stol@lero.ie

Muhammad Ali Babar
IT University of Copenhagen
Rued Langgaards Vej 7
DK-2300 Copenhagen S, Denmark
+ 45 7218 5107
malibaba@itu.dk

ABSTRACT
Component-Based Software Development has become a popular approach to building software-intensive systems. Besides using Commercial Off-The-Shelf components, an organization may choose to use Open Source Software components. Using OSS has been reported to have many benefits, but there are also challenges involved. Understanding the potential challenges of using OSS in developing products is important for practitioners, so they become aware of them and can anticipate them and take appropriate measures to address these challenges. We have performed a thorough review of the literature to identify challenges that may arise, as reported in the literature. This paper presents and discusses these findings. Researchers can discuss potential causes and solutions of our synthesized findings as well as benefit from provided references to literature on OSS challenges as input for future research.

Categories and Subject Descriptors
D.2.13 [Software Engineering]: Reusable Software – Reuse model

General Terms: Management, Theory

Keywords: Open Source Software, Challenges, Component-based Development, Literature Review

1. INTRODUCTION
In the last decade or so, it has become quite common to build software-intensive systems following a Component-Based Software Development (CBSD) approach using Off-The-Shelf (OTS) components [38, 44]. Components may be built in-house, for instance as part of product line development or through an “inner source” approach [45], or acquired from third parties. When acquiring third-party components, an organization may purchase Commercial OTS (COTS) components, or decide to use Open Source Software (OSS) products. OSS products are becoming more commercially viable [16]. An increasing number of available OSS products make the use of OSS an attractive alternative to COTS components. Using OSS is often reported to have many benefits, such as significantly lower (purchasing) costs, availability of high quality products, adherence to open standards and no vendor dependency. Software development industry has taken note of these benefits, and has been increasingly using OSS components (OSCs) in combination with, or as an alternative to COTS components [27]. For small independent software vendors (ISVs) that operate in a niche market, using OSCs instead of proprietary (i.e. not free of charge) software allows them to use advanced technologies without the need to purchase expensive licenses that may minimize any margins on the final product [36, 42].

However, besides these benefits, several studies have also reported different challenges involved in using OSCs in software development. However, the studies reporting such challenges tend to focus on specific aspects. Merilinna and Matinlassi reviewed the literature on OSC integration and compare this with real-world practices [31]. Ven and Verelst studied challenges and strategies for ISVs when dealing with modifications and contributions to OSS [42]. Morgan and Finnegan [32] provide an overview of benefits and drawbacks of adopting OSS. However, their review is mostly based on online articles, such as TechSoup (www.techsoup.org), reports by commercial research institutions (e.g. Forrester) and other types of reports, rather than findings from scientific research literature. Goode reported a literature review and industrial survey on management barriers to OSS adoption [21]. Our study focuses on challenges involved in using OSS for product development (hereafter referred to as OSS challenges).

There has been no systematic synthesis of the OSS challenges reported in the literature. We assert that a synthesis of the reported OSS challenges can help practitioners to fully understand the potential OSS challenges and enable them to take appropriate measures to deal with them. Researchers can use the findings for deliberating and debating the possible causes and appropriate strategies for the identified challenges. This assertion motivated us to undertake a thorough review of the literature to systematically identify and synthesize the reported OSS challenges. This paper reports the research method used and findings from our study. The remainder of this paper proceeds as follows. Section 2 outlines our research method. Section 3 presents and discusses the results of our review. Section 4 concludes and presents an outlook to future work.

2. RESEARCH METHOD
In order to identify all the relevant research papers, we conducted a thorough search of the literature based on the guidelines for conducting a Systematic Literature Review
(SLR), as presented by Kitchenham in [25]. We do not claim that our study is an SLR. Our intention is to identify and synthesize all challenges of using OSS in product development, rather than providing a comparative overview or synthesis of any empirical evidence for these challenges.

To identify all relevant papers, we relied on two different sources. Firstly, we selected a number of papers from our search results from our ongoing extension of the SLR reported in [37]. For this extension, we searched a number of digital libraries for all studies that were related to open source software research, disregarding the type of the papers. The resulting repository contains approximately 550 papers. Secondly, we manually selected papers reported in the proceedings of the five editions of the International Conference on Open Source Systems (2005 to 2009).

We included papers that report on product development with OSS, including CBSD with OSS. Papers that report on adoption of, or migration to OSS were excluded, since they do not address issues in product development but rather focus on usage (by end-users) of the software. We acknowledge that the difference between product development and adoption can be a very thin line; we required that a paper had to report on some sort of development activity, rather than just replacing an existing proprietary or commercially purchased system with an OSS alternative.

Based on these criteria, we identified 44 studies. We inspected each of these papers to identify issues, concerns and challenges related to product development with OSS. We did not require that challenges in the studies were empirically grounded. In other words, a paper could be an experience report, workshop report, or otherwise. A number of papers that we initially selected did not mention any challenges. Others discussed the use of OTS in general, not differentiating between OSC and COTS components. After our selection procedure, we were left with 17 papers.

We thoroughly read all identified papers to identify any reported challenges related to using OSS in product development. Some papers explicitly listed challenges, whereas others implicitly reported the issues. We extracted both challenges that were reported as authors’ experiences or as cited from other literature. While reading the papers, we recorded all challenges in a spreadsheet. After extracting the challenges, we annotated each one with one or more keywords. In particular, we found a number of challenges that could be classified as ‘community’, ‘support’ and ‘maintenance’, since maintenance (e.g. bug fixes) is a type of support, which is provided by the community. Based on these keywords, a number of categories emerged, which we used to cluster related challenges. During analysis of the challenges, we merged similar challenges that essentially stated the same issue.

3. RESULTS AND DISCUSSION
The results of our review are presented in Table 1. We identified 21 challenges that have been reported in literature. The table indicates the number of times each challenge has been reported (column “#”) as well as references to the reporting studies. The bracketed number after the category names indicates the total number of reports of challenges in that category.

3.1 Product selection
3.1.1 Too much choice
Various studies reported that identifying quality products among the many available OSS products is difficult due to an uncertainty about the quality (C1). The quality is typically referred to in terms of quality attributes such as usability, reliability and performance. Sourceforge, the largest repository for OSS projects, alone hosts more than 230,000 projects. This challenge has long been recognized, and has resulted in a variety of OSS evaluation methods and frameworks, such as Capgemini’s Open Source Maturity Model (OSMM) [14], Navica’s OSMM [20], OpenBQR [39] and QSOS [4]. However, despite these efforts, research has shown that practitioners typically do not use these evaluation methods and frameworks [28, 31]. Rather, they use ad-hoc approaches and information sources to select components, such as experiences of colleagues. Hauge et al. found that a “first fit” rather than a “best fit” principle is applied [22].

3.1.2 Lack of time to evaluate
A related challenge to C1 is a lack of time to evaluate components (C2). Though this is of course a direct consequence of having a large number of OSS products to evaluate, we decided it should be listed separately, since it was separately identified by a study ([5]) that also reported challenge C1. Furthermore, possible measures that practitioners could take to address C2 are different from measures for C1 and distinguish C1 more clearly from C2. Such measures include: 1) limiting the number of components to evaluate to a small number, and 2) a decision on management level to allocate more resources (time, manpower) for evaluating OSCs.

3.1.3 Choosing a fork
Another related challenge to C1 is deciding what “fork” of the OSS project should be used (C3). If a fork occurs, a new project is spun off from the original project, and can occur if a project’s core developers have fundamental disagreements about the future of the project [36]. This challenge is different from C1, since C1 refers to making the decision on what product to select, whereas C3 refers to what fork of that product should be selected. Forking projects rarely happens, and there is a strong social pressure against forking [34]. Nevertheless, if it happens, developers need to decide which fork of the project to select. One experience paper reported that this decision caused a temporary delay in development [7].

3.2 Documentation
A lack of good quality documentation remains to be a challenge that is difficult to overcome (C4). Well-documented software is easier to understand by others, which makes it easier to modify the software. However, OSS contributors are typically more interested in coding, and some consider adding comments in the source code is sufficient [17]. Researchers have proposed various architecture recovery methods to overcome the lack of design and architecture documentation [13].

The availability of different descriptions of the same components is problematic as well (C5). OSS products may have documentation, but due to the active evolution of many OSS products, this documentation may quickly go out of date.
3.3 Community, support and maintenance
A wide variety of challenges have been reported with respect to the interaction with the community. This interaction can be related to (future) support for the product as well as contributions to the project. Maintenance can be done by both the OSS product’s community as well as through contributions from the product’s users. In fact, the boundary between “community members” and “users” may not be that clear according to the onion model [12], which states that the social structure of an OSS community is layered, and users are just another layer. As Gacek and Arief state: “all OSS developers are users, but not all users are developers” [17].

<table>
<thead>
<tr>
<th>Category</th>
<th>ID</th>
<th>Challenge</th>
<th>#</th>
<th>Reported in</th>
</tr>
</thead>
<tbody>
<tr>
<td>Product Selection (9)</td>
<td>C1</td>
<td>Identifying quality products among the large supply is difficult due to uncertainty about quality (e.g. usability, stability, reliability)</td>
<td>7</td>
<td>[5, 10, 11, 23, 26, 31, 41]</td>
</tr>
<tr>
<td></td>
<td>C2</td>
<td>Lack of time to evaluate components</td>
<td>1</td>
<td>[5]</td>
</tr>
<tr>
<td></td>
<td>C3</td>
<td>Decide what “fork” of the project should be chosen</td>
<td>1</td>
<td>[7]</td>
</tr>
<tr>
<td>Documentation (5)</td>
<td>C4</td>
<td>Lack of, or low quality documentation</td>
<td>4</td>
<td>[2, 5, 29, 31]</td>
</tr>
<tr>
<td></td>
<td>C5</td>
<td>Several descriptions of the same component</td>
<td>1</td>
<td>[6]</td>
</tr>
<tr>
<td>Community, support and maintenance (19)</td>
<td>C6</td>
<td>Dependency on the community for further support and upgrades; possible need to hire additional talent for maintenance; difficult to control the quality of the support; lack of helpdesk and technical support.</td>
<td>5</td>
<td>[10, 11, 26, 31, 41]</td>
</tr>
<tr>
<td></td>
<td>C7</td>
<td>Custom changes need to be maintained, which is time-consuming and may cause problems with future versions/community may take a different, incompatible approach.</td>
<td>6</td>
<td>[7, 23, 24, 30, 42, 43]</td>
</tr>
<tr>
<td></td>
<td>C8</td>
<td>Convincing OSS community to accept changes (modifications may be too specific); contributions can be difficult or costly. Difficult to control the architecture if not a core member.</td>
<td>5</td>
<td>[7, 23, 30, 31, 42]</td>
</tr>
<tr>
<td></td>
<td>C9</td>
<td>Uncertainty about product future and consequences for company product</td>
<td>1</td>
<td>[7]</td>
</tr>
<tr>
<td></td>
<td>C10</td>
<td>Community members would like to have a bigger say in features and integrating final product with company</td>
<td>1</td>
<td>[24]</td>
</tr>
<tr>
<td></td>
<td>C11</td>
<td>Contributing and investing in OSS project costs resources</td>
<td>1</td>
<td>[24]</td>
</tr>
<tr>
<td>Integration and Architecture (8)</td>
<td>C12</td>
<td>Backward compatibility concerns</td>
<td>2</td>
<td>[24, 41]</td>
</tr>
<tr>
<td></td>
<td>C13</td>
<td>Modifications needed to implement missing functionality or fit into architecture</td>
<td>2</td>
<td>[41, 42]</td>
</tr>
<tr>
<td></td>
<td>C14</td>
<td>Incompatibility between components or existing systems</td>
<td>2</td>
<td>[11, 41]</td>
</tr>
<tr>
<td></td>
<td>C15</td>
<td>Horizontal integration</td>
<td>1</td>
<td>[31]</td>
</tr>
<tr>
<td></td>
<td>C16</td>
<td>Vertical integration / Mismatch of platform/programming language</td>
<td>1</td>
<td>[31]</td>
</tr>
<tr>
<td>Migration and usage (3)</td>
<td>C17</td>
<td>Complexity of configuration</td>
<td>1</td>
<td>[41]</td>
</tr>
<tr>
<td></td>
<td>C18</td>
<td>User training/learning costs</td>
<td>2</td>
<td>[11, 41]</td>
</tr>
<tr>
<td>Legal and Business (5)</td>
<td>C19</td>
<td>Complex licensing situation</td>
<td>5</td>
<td>[1, 24, 29, 35, 41]</td>
</tr>
<tr>
<td></td>
<td>C20</td>
<td>Concerns about, or no clear strategy on Intellectual Property and Rights issues</td>
<td>3</td>
<td>[1, 35, 41]</td>
</tr>
<tr>
<td></td>
<td>C21</td>
<td>Lack of clear business models that are appealing to industry</td>
<td>2</td>
<td>[1, 11]</td>
</tr>
</tbody>
</table>

3.3.1 Dependency for future support
If an organization decides to use an OSS product, it is dependent on the community for future support and upgrades. A challenge is to acquire support for the OSS product that is of sufficient quality (C6). Support for an OSS product is provided on a voluntary basis by the community, which makes it difficult to control the level of quality of support that is needed. However, for some products, there is an option to acquire support and training from companies such as Red Hat and IBM [15]. Various studies report the support to be a challenge (see Table 1). Ven and Verelst reported a study that investigated the reliance of organizations on commercial support [43]. They found that the absence of available commercial support is not an insurmountable obstacle for adopting OSS. However, they also found that the OSS community is primarily used by organizations with a strong technical background.

3.3.2 Need to maintain custom changes
If changes are made to an OSS product, and these modifications are not given back to the product’s community, then the software developers that made such changes need to maintain these custom changes themselves (C7). This means that additional resources must be allocated for the maintenance
efforts. Furthermore, if an organization does not give back the modifications, it effectively ‘forks’ the project as the customized modifications define a new version of that product. This may have serious consequences for future compatibility. Modifications may have to be re-applied whenever new versions of the OSS are used. As an OSS product evolves, patches that implement modifications may no longer be easily applied. Alternatively, an OSS product’s community may tackle a certain feature or issue themselves by taking an approach that may be incompatible with the customized modifications [42].

3.3.3 Difficult to get changes accepted
An organization may decide to contribute the changes made back to an OSS component (that is integrated into a product). However, sending patches to an OSS’s community does not automatically imply that these patches get accepted (C8). A key characteristic of an OSS development process is that contributions are thoroughly scrutinized by community members that have commit access [15]. In general, any changes or proposals for change will be subject to a review process [34]. Furthermore, an OSS project may have specific practices that developers will have to get familiar with [7]. Specific extensions may be rejected to prevent that too many new features (code bloat) are introduced by one-time contributors [42].

3.3.4 Uncertainty about product future
A realistic concern that organizations may have is the future, or longevity of the OSS product (C9). Obviously, if a certain OSS product is adopted, an adopter does not like to be in a situation where the community supporting that product disappears. If that happens, it means no support or updates for that particular product. In such a case, an organization may choose to take over the maintenance of the project. However, this would result in additional maintenance efforts, and may distract the organization from its core business.

3.3.5 Community wants more influence
One study reporting experiences of product development with OSS at Nokia reported a challenge similar to C8, but in the opposite direction. OSS developers expressed their wish to be more closely involved in features of the final product. The closed way of integrating these OSS products can cause frustration among OSS developers [24] (C10). To partially address this issue, Nokia started a special distribution, to allow anybody participate more closely in the development. However, as stated in [24], product companies must have the final control over their products.

3.3.6 Contributing costs resources
An organization may choose to use OSS as-is without further development. Alternatively, using OSS becomes more effective if the organization actively participates in a community’s development process. This, however, requires additional resources (C11). The amount of resources required depends on the level of involvement. Bonaccorsi et al. [9] list three kinds of involvement: 1) project coordination; 2) code development collaboration, and 3) provision of code.

3.4 Integration and architecture
3.4.1 Backward compatibility issues
An OSS product is continuously evolving, depending on the liveliness of a community. Changes to products include new features, bug fixes and architectural changes. After an organization starts using an actively evolving product, new versions are released. As a product’s development continues, at some point newer versions are no longer backward compatible, which can become a problem if the product in which the OSS is integrated depends on certain features or APIs (C12).

An organization will have to adopt a strategy for updating any used OSCs. On the one hand an organization may choose to stay close to the latest version of the OSC. However, this has consequences for backward compatibility, as features may be deprecated and architectural changes may occur [24]. Ven and Mannaert describe four possible strategies for contributing to OSS projects [42]: 1) contributing any modifications, 2) taking regular snapshots, 3) forking and 4) initiating an OSS project as a set of patches to an existing OSS project. One solution to this problem is to use only those distributions that are provided by packaging companies [35].

3.4.2 Need for modifications
A consequence of a CBSD approach is that components must be fitted into a system. That means the components may have to be modified. Furthermore, OSCs may have to be modified if they do not have all required functionality. Such modifications require additional resources (C13). Obviously, OSCs are more flexible than COTS because the source code of OSCs is available for modifications. However, many organizations do not usually make any changes to the source code before using OSCs [28].

3.4.3 Component and architecture incompatibilities
OSC may not be compatible with each other, or with existing architectures (C14). This phenomenon is called architectural mismatch, which may have serious consequences for the development schedule and costs [18]. Another compatibility issue that may arise is that components may have dependencies on conflicting libraries [42].

3.4.4 Horizontal integration issues
Merilinna and Matinlasi [31] distinguish horizontal integration issues at the architectural level and the component level (C15). No specific architectural level issues have been identified. For the component level, four design level contracts have been discussed in [8]. These levels are: 1) syntactic interface, 2) (pre- and post-conditional) constraints, 3) synchronization and timing and 4) quality-of-service. Each of these levels can have associated challenges.

3.4.5 Vertical integration issues
A mismatch of platform and programming language is an example of vertical integration issues (C16). Platforms may be hardware (i.e., processor types) or software (operating systems and virtual machines (VM) such as the Java VM, .NET and Parrot VM). Some of the techniques to overcome the vertical integration problems are use of middleware (e.g. CORBA), virtual machines (e.g. JVM, a hardware platform-independent virtual machine) and Model-Driven Architecture (MDA) [31].

3.5 Migration and usage
Complexity of configuring or setting up a user-environment can be an issue (C17). One study reported that significant effort was required to set up an installation (approx. three weeks) [41].
Two studies reported additional cost involved in migration to an OSS alternative and staff training to be a challenge (C18). In our study, we focus on product development with OSS rather than adopting OSS in favor of proprietary solutions. Migration cost seems to imply the migration from a proprietary solution to an OSS solution, such as the migration from Microsoft Office to OpenOffice.org. However, such end-user applications may be integrated as a sub-system of a larger solution.

3.6 Legal and business

3.6.1 Complex licensing situation

Not surprisingly, several studies reported the complex OSS licensing situation to be an issue (C19). One study reports a lack of consistency between licensing agreements and little guidance on interpreting the open source licenses [41]. At the time of writing, the Open Source Initiative lists 65 licenses that comply with the “Open Source Definition” [33]. It is therefore not surprising that OSS licensing is perceived to be a complex issue. Some research efforts have been made to address this issue. Alspaugh et al. [3] present a license analysis scheme, and an approach to automatically analyze license interactions. German and González-Barahona have documented a number of strategies that developers have used to legally circumvent some restrictions of the GPL [19].

3.6.2 Concerns and issues regarding IPR

Organizations that use OSS products may have concerns about intellectual property and rights (IPR) (C20). Code may have been illegally used and propagated. For instance, in recent years there have been some claims from Microsoft saying that Linux uses their intellectual property [40]. In that particular case, a deal was made with Novell, so that customers of Novell’s SUSE Linux distribution are protected from any claims.

3.6.3 Lack of clear business models

Two studies reported a lack of clear business models for using OSS (C21). We note that both studies were published in 2005; however, this issue has not been mentioned in more recent literature. In [17], three business models have been identified that motivate organizations to get involved in OSS: 1) software for own use, 2) packaging and selling of the software and 3) a platform for commercial or research software development.

3.7 Limitations of this study

Our study has some limitations, which we discuss here. Though we performed a rigorous literature search, we may have unintentionally excluded studies due to the subjectivity of our inclusion and exclusion criteria. Our classification of challenges is necessarily subjective. However, it is not our intention to present a definitive classification; rather, we intend to present our findings in a structured way that can help practitioners to inform them of challenges in using OSS in product development that have been reported so far, and may therefore arise in their situation.

4. CONCLUSION AND FUTURE WORK

Open Source Software products are being increasingly used as an alternative to Commercial Off-The-Shelf components. Using OSS has been reported to have many benefits, but also has various challenges. Therefore, practitioners may be reluctant to integrate OSS products as part of their final product. We assert that a good understanding of OSS challenges can help practitioners to be well prepared for such challenges so that they can take appropriate measures to address those challenges. This paper presents the results from our study aimed at systematically identifying and synthesizing the challenges reported to be involved in using OSS in product development. We believe these findings can be equally useful for practitioners and researchers. Practitioners will become more aware of the potential challenges of using OSS in product development. The research community can deliberate and discuss the potential causes and solutions of the synthesized challenges. Furthermore, the references to the literature on OSS challenges can benefit researchers interested in doing future research in this area. We intend to continue this research, which will focus on the challenges regarding the software architecture and the effects on a system’s quality as a result of integrating OSS components.

5. ACKNOWLEDGMENTS

This work is partially funded by IRCSET under grant no. RS/2008/134 and by Science Foundation Ireland grant 03/CE2/1303.1 to Lero—The Irish Software Engineering Research Centre (www.lero.ie).

6. REFERENCES


[34] Raymond, E.S.: The Cathedral & the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary (O'Reilly, Revised edn, 2001).


