Managing First Impressions of New Open Source Software Projects

Namjoo Choi and InduShobha Chengalur-Smith, University at Albany, SUNY
Andrew Whitmore, Simmons College

One of the most important indicators of open source software (OSS) project success is the number of developers joining the project. Some OSS projects begin with a previously networked group of experienced developers who are then able to recruit other developers. However, the vast majority of OSS projects are initiated by one or at most a few developers who are not yet part of a collaborative network and have to wait for potential developers to notice and join their projects.

An initiator’s failure to attract other developers often results in abandonment of the project. Therefore, although the OSS community has reason to celebrate the explosive growth in the number of OSS projects in recent years, we must pay more attention to how these new projects can effectively and efficiently attract more developers in light of these facts.

The OSS literature has largely ignored how new OSS projects can attract developers in order to sustain their vitality. Drawing upon first-impression-management research, we explore the potential role that the initial presentation of new OSS projects plays in attracting developers. We provide preliminary findings from a pilot study including interviews and an exploratory quantitative analysis.

Project Participation
Numerous factors can influence a potential developer’s decision to participate in an OSS project, including application type, project vitality, previous ties with the initiator(s), popularity, and license type. Even before potential developers assess project factors, the pathway through which they arrive at the project might vary by their experience level and the extent of their personal network. For example, experienced developers might depend more on their previous collaborative ties, while novice developers might search for a project to join on OSS hosting websites such as Sourceforge.net. In either case, finding and selecting a project is not a trivial task, but one that requires much time and thought.

OSS research has largely ignored one interesting aspect of OSS projects: first-impression management. The mature status of established OSS projects likely attracts developers given their greater activity and vitality. Potential developers typically join such projects by first becoming project users and then evolving into developers and stable committers. However, this pathway is unavailable for newly initiated projects.

When the initiators of a new OSS project first introduce it to the world to generate collaboration,
they present a few pieces of information such as a name, description, mission statement, license type, features and requirements, screenshots, downloadable initial work, and a project website. Because this is the very first exposure potential developers have to a new project and presumably the only information they will use to make a decision, we cannot overlook the role of first impressions in attracting developers. The initial project presentation can provide cues into the nature of the project and the initial development team. For example, if a project is poorly introduced and potential developers feel that the initiators did not invest much care in the initial presentation, they might formulate negative opinions about the prospect of project success and fail to affiliate with the project even if they have great interest in its scope.

Poor first impressions not only impact the join rate of potential developers in the short run, they can also produce negative externalities for the project in the long run. For instance, they might lose the synergies derived from large groups of developers working together, such as greater project activity and popularity. Consistent with this idea, much OSS literature stresses the importance of the size of the current developer and user base as important factors in attracting additional collaborators. Although the size of the developer and user base might be a dominant factor in driving further software adoption in the long term, it fails to explain the growth of OSS projects during project startup when these bases are small or altogether absent. Yet, despite this probable impact of the initial impressions on attracting potential developers, OSS developers haven’t paid adequate attention to this aspect of the development process. Karl Fogel indirectly captures this point by stating that programmers often value substance over form and some even exhibit an antipathy for marketing and public relations work.

**First-Impression Management**

Impression management theory refers to the process by which individuals or organizations try to control or manage the impressions others form of them. The theory dates back to Erving Goffman’s pioneering work that first brought the importance of self-presentation to light. Goffman employed the metaphor of theatrical performances as a framework to understand the way our behavior is constructed in social interactions and the way we appear to others. Consequently, a large body of literature examined self-presentation from the impression management perspective and later the research expanded to organizational impression management. Due to its intuitive importance and wide applicability, a great deal of research has applied the theory to different disciplines, ranging from multimedia to medical research.

Marketing literature has specifically examined the powerful effect a first impression has on forming a person’s long-term impression toward products. This effect is called the halo effect, which posits that a person with a positive first impression toward a product might disregard negative issues encountered later, and vice versa.

This line of research on first-impression management has been rapidly adopted as one of the main theoretical lenses in the information management literature. Due to the ever-increasing amount of information available on the Internet, the need to make quick choices among competing alternatives has risen in importance, and thus so has first-impression management. Accordingly, researchers have examined various types of websites in terms of users’ perceptions of their appeal.

For example, Heejun Kim and Daniel R. Fesenmaier recently examined the persuasiveness of travel websites by investigating information seekers’ first impressions in the context of seeking travel information. The study proposed a conceptual model that describes the online information search process of travel information seekers who do not have a predetermined travel website in mind. Such users depend on a search engine to identify a set of alternative sites from which they can choose. The model consists of three distinct stages: search (users input one or more search terms and the site returns a list of travel websites), primacy (first-impression formation is used to select a travel website), and elaboration (how search within a website can help users learn about the website and a destination and subsequently adjust their first impression). The Kim-Fesenmaier conceptual model can explain how novice developers search for projects on OSS hosting websites (see Figure 1).

---

**Figure 1. The information search process for an open source software (OSS) project (adapted from Kim and Fesenmaier).**
the first stage, potential developers type a search term into an OSS project registry and locate project listings related to their topic of interest. The second stage of the search process involves selecting a particular OSS project from the located listing of projects.

At this stage, the cues used for project selection depend heavily on the information presented in the search results. Information such as project name, short description, screenshot availability, downloadable initial work availability, rank, activity level, latest file release, and downloads plays a significant role because it functions as the primary cue for a potential developer to raise enough interest for further investigation. If a project fails to properly provide this initial information at the second stage, it might not have the opportunity to be exposed to the third stage.

Once potential developers select a project via this first-impression-formation process, they explore the project using the project page. At this third stage, the information fields preset by OSS hosting systems such as Sourceforge.net are provided—for example, intended audience, operating system, programming language, and license type. For further information, the project can also provide a direct link to its website. The project website lets the initiator describe the project in far greater detail, including perhaps a mission statement, progress to date, future plans, and so forth.

Finally, with more information available in the third stage, a potential developer adjusts the first impression and decides whether to invest more time and effort, such as sending inquiries to the project administrator(s), or whether to go back to the search results (the second stage).

A Pilot Study
To explore the potential, partial role that the initial presentation of newly initiated OSS projects plays in attracting developers, we first conducted interviews based on the conceptual model (see Figure 1). Our goal was to identify the most informative presentation cues provided by newly initiated projects to potential developers—cues such as project name, short description, screenshot availability, downloadable initial work availability, and project website availability. The participants reflected the target population of potential novice developers, and we selected Sourceforge.net as the hosting website because it is one of the largest OSS repositories.

We interviewed six computer science major college students (three senior undergraduate students and three graduate students) who have one to two years of OSS development experience and some level of familiarity with Sourceforge.net. We asked them to find a project less than one-year old that they thought was interesting enough to attract their attention as a potential project to contribute to. This served to de-emphasize the impact of other factors such as project rank, activity level, latest file release, and downloads that are absent or small during initial project stages.

After observing each participant’s search process, we interviewed them to capture their decision-making process, specifically to determine the presentation cues they used most. Each interview lasted about an hour. Our observation results indicated that the participants followed the process described in Figure 1. All the participants first typed a specific search term into the search bar (stage one), then searched through the identified project listing and clicked into several of those OSS projects (stage two). On the project pages they selected, the participants looked for further information about the projects (stage three).

During the interviews, the participants reported that project name had little impact on their decisions. All the participants used the project description as an important presentation cue, in terms of whether it provided a more detailed and specific description of the project and how well it communicated the software’s intention. Finally, all the participants pointed out that screenshot, downloadable initial work, and project website availability were important presentation cues because they provide the most information about the project.

We followed up the interview results with a quantitative analysis using project data obtained from Sourceforge.net. Based on the interview findings, we included the following presentation cues: project description, screenshot availability, downloadable initial work availability, and project website availability for the analysis. We selected projects from the security category in Sourceforge.net and then narrowed our focus to cryptography, the biggest subcategory. This let us control for differences across projects in different categories as well.

Again, to focus on the initial presentation cues while de-emphasizing the impact of other factors, we limited the sample to projects that were registered within the past 18 months (choosing only projects registered between 1 August 2006 and 1
January 2008) and collected data (the total number of page hits to the second page) for a one-year period for each project. Some projects that were not directly related to cryptography were dropped, and the final sample consisted of 94 projects.

We conducted a multiple regression analysis to predict the impact of first impressions, measured by the total page hits to the second page after one year had passed since the launch of each project. The explanatory variables were project description, screenshot availability, downloadable initial work availability, and project website availability. Although project website availability can be found only on the project page in Sourceforge.net, we decided to examine its impact as one of the explanatory variables against the same dependent variable because project website availability can indicate how much care the project initiators invested in the initial presentation and is therefore a proxy for the factor.

Although we found little difference across projects in project rank, activity level, latest file release, and the number of developers, our initial observation of the data indicated differences across projects in terms of downloads because some projects were launched with some initial or more mature downloadable work. To address this issue, we included project downloads as a control variable in our analysis. For the project description measurement, two of us who have previous experience in the security area (Choi and Chengalur-Smith) coded each project’s description, based on the findings from the interviews, as to how detailed and specific it was and how well it communicated the software’s intention. Disputes in differences (≥ 3.0) in coding eight projects (9 percent) were resolved via discussion. Finally, we examined inter-coder reliability using Cronbach’s Alpha statistic and found it to be high (0.94). We manually checked screenshot, downloadable initial work, and project website availability to verify whether they existed. The dependent variable—total page hits—was transformed by taking natural logs since it showed large variation by project.

Table 1 summarizes our results. We examined the assumptions of multivariate normal distribution, independence of errors, and equality of variance. There were no violations of these assumptions and multicollinearity was also not a concern as the variance inflation factors were all below two. The factors in our model—project description, screenshot availability, downloadable initial work availability, and project website availability—explained 66 percent of the differences in page hits. Even after controlling for the number of downloads, these variables were significant predictors of interest in a project. Thus, our findings suggest that including these pieces of information would increase the likelihood of attracting other developers to a project, particularly those who might not have prior affiliations or experience with OSS projects. It was surprising that only a minority of projects in our sample provided all these presentation cues: 60 percent provided screenshots, 36 percent provided downloadable initial work, and 27 percent provided a project website.

Due to its exploratory nature, this study has some limitations. A more systematic empirical analysis that considers first-impression management with other yet unidentified factors should be undertaken for more robust results in future research. Also, as we described earlier, the pathway through which potential developers arrive at a project to contribute to might vary by other factors such as their

### Table 1

<table>
<thead>
<tr>
<th>Variable</th>
<th>Standardized coefficients</th>
<th>P value</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.000</td>
<td>.000</td>
<td>n/a</td>
</tr>
<tr>
<td>Project description</td>
<td>0.121</td>
<td>.077</td>
<td>Supported at p &lt; 0.10</td>
</tr>
<tr>
<td>Screenshot availability</td>
<td>0.212</td>
<td>.002</td>
<td>Supported at p &lt; 0.01</td>
</tr>
<tr>
<td>Download availability</td>
<td>0.372</td>
<td>.000</td>
<td>Supported at p &lt; 0.01</td>
</tr>
<tr>
<td>Project website availability</td>
<td>0.272</td>
<td>.000</td>
<td>Supported at p &lt; 0.01</td>
</tr>
<tr>
<td>Project downloads</td>
<td>0.305</td>
<td>.000</td>
<td>Control variable</td>
</tr>
</tbody>
</table>
experience level or the extent of their personal network. More research that examines other possible pathways in different settings beyond the scope of this study is required to expand our understanding of how OSS projects can more effectively and efficiently attract potential developers in order to sustain their vitality.

References


About the Authors

Namjoo Choi is a doctoral candidate and lecturer in informatics at the University at Albany, State University of New York (SUNY). His research interests include open source software, information security, decision-support systems, and e-government. Choi has an MA in telecommunication, information studies, and media from Michigan State University. Contact him at nc236879@albany.edu.

InduShobha Chengalur-Smith is the chair of the Information Technology Management Department in the School of Business at the University at Albany, SUNY. Her research interests are in open source software, technology adoption and implementation, information quality, and security. Chengalur-Smith has a PhD in statistics from Virginia Tech. Contact her at shobha@albany.edu.

Andrew Whitmore is a visiting professor of information science at Simmons College. His research interests include Internet governance, e-government, and open source software. Whitmore has a PhD in information science from the University at Albany, SUNY. Contact him at andrew.whitmore@simmons.edu.

Selected CS articles and columns are also available for free at http://ComputingNow.computer.org.