

Relevant Aspects for Sustainable Open Source Pandemic Apps and Platform Deployment with Focus on Community Building

Daniel THOLE,^{a,1} Till Ole DIESTERHÖFT^b, Stefan VOGEL^a, Maike GREVE^b,
Elena BAUER^b, Stefan STRATHMANN^{a,c}, Christian ELSNER^d, Lutz KOLBE^b, and
Dagmar KREFTING^a

^a*Department of Medical Informatics, Univ. Medical Center Göttingen, Germany*

^b*Georg August Universität Göttingen, Germany*

Chair of Information Management, University of Göttingen, Germany

^c*Göttingen State and University Library, Germany*

^d*University Medical Center Mainz, Germany*

Abstract. The global COVID-19 pandemic revealed the necessity for mobile and web-based solutions for a variety of medical processes, e.g., individual risk calculation, communication of health information and contact tracing. Many such solutions are provided in form of open source software. However, there are major obstacles to the sustainable long-term continuation of such projects. As the topic of sustainability strategies is complex, a classification would be useful to help new projects to identify relevant sustainability factors. Based on a literature review a classification for long-term success of open source software was created. This paper presents a classification focusing on five unique categories: (1) structural decision, (2) revenue generation, (3) user focus, (4) openness and (5) community building. It was developed within the NUM-COMPASS project, focusing content-wise on pandemic apps and structure-wise on open-source provision. We provide some insights into the community building dimension by discussing factors that go into building sustainable communities.

Keywords. COVID-19, Pandemic, Open Source, Sustainability

1. Introduction

The global COVID-19 crisis that erupted in 2020 is an ongoing challenge for health systems. Many countries established social distancing regulations for their population. As a consequence many mobile and web-based applications have emerged to address different aspects of handling the pandemics. These aspects encompass contact tracing, telemedicine, personal risk assessment, and surveys and studies on impact of the pandemics on different disorders [1,2]. A couple of these apps have been developed in scientific projects and are provided under an open source license. In order to fight the

¹ Corresponding Author, Department of Medical Informatics, Univ. Medical Center Göttingen, Germany, Robert-Koch Straße 40, 37073 Göttingen, Germany; E-mail: daniel.thole@med.uni-goettingen.de.

pandemic the German government established the Netzwerk Universitätsmedizin (NUM). The aim of the network is to employ joint forces and expertise of Germany’s university medical centers for a better understanding and handling of the pandemic [3]. One of in total 13 projects of NUM is NUM-COMPASS [4]. The project is established to develop best practices for pandemic apps and to develop an open source technical framework to ease implementation of these best practices including a reference implementation with specific focus on pandemic preparedness.

As pandemic preparedness would require the sustainable maintenance of this framework during periods of low awareness on pandemic risks, the question how the project results can be sustainably preserved, has been explicitly included into the workplan. This is especially important since it is known, that the vast majority of open source projects fail to reach maturity [5].

This paper presents a classification of relevant aspects going into the development of pandemic apps. Since the project is developing an open source application and the WHO recommends using open source technology [6] in such a case, the focus has been on open source only. The advantages of open source for the development of pandemic apps have been addressed by Tom-Aba et al. [7], who identified that tools based on open source are more sustainable through a dynamic community.

2. Methodology

The term classification describes a way of organizing knowledge [8], in this case the knowledge of aspects for sustainable open source projects.

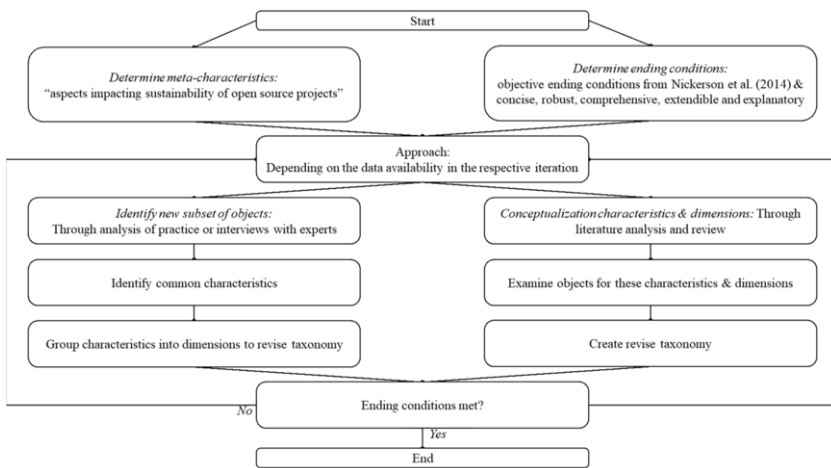


Figure 1: Taxonomy Development Process based on Nickerson et al. [9]

Development of the classification follows the process presented by Nickerson et al. [9]. We selected this method due to its interdisciplinary applicability [9]. Furthermore, the approach is iterative and formalized[10] and thus offers advantages over the phenetic[11] or cladistic[12] development techniques. The process starts by defining a meta-characteristic, the central aspect of the classification, and some ending conditions

that need to be met. During each step the classification is altered, e.g. by newly identified characteristics. The objects correspond to actual companies and projects (e.g., COMPASS) which were discovered through practice analysis, interviews and literature research. The process terminates once the ending conditions are met.

A thorough literature review on sustainability in open source projects and platform technologies was conducted through the databases: ABI/INFORM [13], Econbiz [14], and Business Source Complete [15]. Main search terms were: "sustainability", "open source", "platform technologies", "financing", "community" and "business model". Only those texts were examined that were available as full-texts, were written in German or English and had gone through a peer-review process. Because of the vast amount of publications, only the first 15 pages of the search results have been examined. The publications were screened first, based on title and abstract before going through a full text analysis. In the end, 39 papers went into the final consideration. The literature review was conducted by three independent reviewers from October 2020 to February 2021.

In the second step of our method we examined a few real-world open source projects more closely. The projects included have been established for at least 10 years and trusted interview partners were available. Based on these criteria, the i2b2 tranSmart [15] and Kitodo/Goobi [16,17] projects have been selected. We then conducted interviews with contributors from the respective communities. Interview partners/contributors were from a senior role with many years of experience in the project. The interviews lasted for about 90 minutes and were conducted at the end of 2020/beginning of 2021. Each interview with the contributors was conducted in a semi-structured manner. We also examined 16 case studies on open source projects [18,20]. As in these projects and studies the topic of community building has been identified as essential for a project's success. By using the databases mentioned above we searched for literature which addressed the topic of community building in open source projects, by using the same process. We identified 12 additional papers and books relevant for our analysis and the results.

3. Results

Based on the research process described above we achieved the results presented in the table 1. The classification can be represented in five categories (Table 1) each having a number of criteria to describe their characteristics. These categories are:

- Structural decisions include those related to the value promised (utilitarian, emotional or social) or the type of platform being used (web-based and/or mobile-based) [19]. Other decisions in this dimension concern the organizational structure, the platform autonomy and how innovation is generated [20].
- Revenue generation is possible through different means, ranging from offering parts of the software or support in exchange for money to external financing via grants or donations. A model where monetary revenue is no concern and the platform exists only as a means of self-service is also possible [21].
- User-focus is concerned with criteria such as regionality (global, regional, local) [19] or target market (e.g., new markets, c2c, b2c, b2b) [22,23].
- Openness is described by criteria such as requirements for participation [23,31], transparency of information provision [24] as well as leadership structure (organizational vs. community-driven) [27].

- Community building is influenced by criteria such as project culture [30, 31], transparency within and accessibility to the community [24] as well as factors influencing member motivation [29].

Table 1: Classification with identified aspects impacting sustainability of open source projects.

Category	Characteristics
Structural decision	Value Proposition [19] Platform type [19] Organizational Structure [20] Platform Autonomy [20] Innovation Generation [20]
Revenue generation	Several Models [21]
User focus	Regionality [19] Focus [19] Target Market [22,23]
Openness	Requirements for Participation [23,27] Transparency [24] Leadership Structure [27]
Community building	Culture [30,31] Transparency [24] Accessibility [24] Member Motivation [29]

Given the importance of community building and openness for the success of open source projects outlined above, this paper will discuss these categories in more detail drawing on a well-known example. The chosen example is the i2b2 tranSmart project [15]. The program started out as a research project to build a system that allows access to the Data Warehouse of Harvard Medical. Initially conceptualized as a local solution, i2b2 tranSmart is now being used and sponsored by companies and institutions in Europe and the United States. The software is distributed under an open source license and maintained by an open source community. The experience from this project (as well as from the Kitodo/Goobi project), is that establishing an active community is essential for project success.

3.1. Openness in Open Source Communities

Even though communities develop in an organic manner, there are some ways to influence the community building process. First of all, an attractive community has low barriers to entry and allows people of any sorts to participate [23,27]. This is often referred to as *openness*. Openness in the reference sheet encompasses several factors, only one being the access to the community. Additional factors include the transparency of information distribution and the replicability of the results [24]. Having an open community that gives access to people from different backgrounds, allows for a more sustainable and effective innovation process to take place [20,21]. As Shah [26] further indicates, there might be users without the technical background necessary to do the coding themselves, but the need to use the software in some context. These members

provide valuable input for the future development of the software and therefore participate in the innovation process.

Capra and Wassermann [27] discuss the topic of governance in open source communities. They argue that the more open a community is, the more decentralized decisions are made. Whereas in rather closed communities decisions and leadership falls to one or few organizations, truly open communities decentralize decision, maybe up to a point where decisions are made in an informal process.

In the context of i2b2 tranSmart this means, that everyone interested in the project can participate. The project organizes its members in working groups, each dedicated to a certain task. Every group decides about its workflow and goals set and even though the board checks in on the progress regularly there is no set deadlines other than those set by the working group itself. Hence, decision-making is rather decentralized.

It is important to note, that openness is essential for open source communities to work out in the long-run. If the community feels that its needs for openness are not met, its members start to search for other communities to participate in or even start their own projects [27].

3.2. Motivating and Integrating Community Members

A transparent and accessible community provides the basis for a sustainable community but in order to create a community we need to look at the motivation for participants to become active in a project. Research by Bonaccorsi and Rossi [29] indicates that the motivations to participate in open source communities are related to economic, social, and technological reasons. Shah [26] further indicates that economic reasons are less important for the majority of developers. These people are more motivated by the enjoyment of coding and the challenge it provides. In order to keep a vibrant community it is therefore important to make sure members feel welcome, challenged and have the feeling that their contributions matter. As Bacon [30] points out mentoring can play a huge role to integrate new members in the community.

One final important aspect we want to mention in that regard is that open source communities rely on regular member events [30,31]. These are normally held in the form of some kind of conference where the current state of the project is discussed and future plans are examined. Additionally it gives community members the opportunity to build more personal relationships with each other. The experiences of i2b2 tranSmart and Kitodo/Goobi support these claims.

Following this research, we conclude that building welcoming communities is the most efficient way to build sustainable open source projects.

4. Discussion

So far, we have given a rough outline of the classification developed for successful open source projects in the context of pandemic apps. The classification described in this paper was based on an iterative method outlined by Nickerson et al. [9]. As pointed out in their paper Nickerson et al. describe their process as a hybrid of a taxonomy and a topology. The results can therefore also be described by other terms. In our case, we opted for classification. One advantage of this approach is that the different categories and characteristics derived are likely to be more useful in different contexts than traditional development methods that provide a rather tight corset for the development procedure.

Our classification can therefore be seen as a flexible construct that is easily extendable by other factors. During our development process, we could indeed experience this, as the first version did not include the community building characteristic. This has only been added after experience from real-world open source projects strongly hinted at the importance of such a category.

The results suggest that having an open community is crucial for the success of open source projects, a generalization that applies to pandemic apps in particular. Since pandemics tend to be quite chaotic and subject to several changes during their duration, a community developing and maintaining a pandemic app needs to be dynamic in order to adjust to changes quickly. This is best achieved through an open community, as discussed by Tom-Aba[7] in the context of pandemic apps for the Ebola in mid-2010.

The results discussed in this paper provide only a limited overview over the actual work that is necessary to build sustainable open source projects, mainly because we focus on community building only. There are, however, many different factors that play a role in making these projects sustainable in the long-run. It will be challenging to see how these factors interact and to what extent they influence one another. We intend to discuss additional aspects in further publications.

5. Conclusion

In this paper, we presented a classification for open source projects and platform technologies. The reference sheet developed is characterized by an easy and flexible approach can be applied to all kinds of open source projects. By following it, one can be assured that all important aspects for successful open source projects are covered. It does, however, require some familiarity with the open source mindset or the flexibility to adjust certain characteristics based on experience. Therefore, it remains to be seen to what extent it can be applied in a real-world context. Further Research into the matter is suggested, since we believe the classification has potential benefits for open source projects and can help to make projects in medical informatics more sustainable.

Declarations

Conflict of Interest: The authors declare that there is no conflict of interest.

Author contributions: DK, LK and CE designed and directed the project; DT, EB and TD developed the theoretical framework; DT wrote the paper with input from all authors; SV, SS, MG assisted by concretizing the methodology MG, TD, SS, SV and DK substantially revised the manuscript. All authors discussed the results and commented on the manuscript. All authors approved the manuscript in the submitted version and take responsibility for the scientific integrity of the work.

Acknowledgements: This work has been funded by the German Ministry of Education and Research (NaFoUniMedCovid19, grant number 01KX2021) COMPASS and B-FAST.

References

- [1] Abeler J, Bäcker M, Buermeyer U, Zillessen H. COVID-19 Contact Tracing and Data Protection Can Go Together. *MIR Mhealth and UHealth*. 2020; 8(4).
- [2] Tobias G, Spanier A. Developing a Mobile App (iGAM) to Promote Gingival Health by Professional Monitoring of Dental Studies: User-Centered Design Approach. *JMIR Mhealth and UHealth*. 2020; 8(8).
- [3] Netzwerk Universitätsmedizin. Antwort auf ungeahnte Herausforderungen. last accessed: 29 March 2021. <https://www.netzwerk-universitaetsmedizin.de/aufgaben-und-ziele>.
- [4] NUM Compass. last accessed: 23 March 2021. <https://num-compass.science/de/>.
- [5] Krishnamurthy S. Cave or Community? An Empirical Examination of 100 Mature Open Source Projects. *First Monday*. 2002. Available at SSRN: <https://ssrn.com/abstract=667402>.
- [6] WHO. Ethical considerations to guide the use of digital proximity tracking technologies for COVID-19 contact tracing. World Health Organisation, 2020.
- [7] Tom-Aba D, Nguku PM, Arinze CC, Krause G. Assessing the Concepts and Designs of 58 Mobile Apps for the Management of the 2014-2015 West Africa Ebola Outbreak: Systematic Review. *JMIR Public Health Surveill*. 2018 Oct-Dec; 4(4).
- [8] Wand Y, Monarchi D, Parsons J, Woo C. Theoretical foundations for conceptual modeling in information systems development. *Decision Support Systems*. 1995; 15(4).
- [9] Nickerson R, Varshney U, Muntermann J. A method for taxonomy development and its application in information systems. *European Journal of Information Systems*. 2013; 22(3).
- [10] Schöbel SM, Janson A, Söllner M. Capturing the complexity of gamification elements: a holistic approach for analyzing existing and deriving novel gamification designs. *European Journal of Information Systems* 2020; 29(6).
- [11] Eldredge N, Cracraft JOEL. *Phylogenetic patterns and the evolutionary process*. Columbia University Press. 1980.
- [12] Sneath P, Sokal R. *Numerical taxonomy. The principles and practice of numerical classification*. W.H. Freeman and Co. 1973ABI/INFORM Collection, last accessed: 27 July 2021. <https://www.proquest.com/abicomplete/index>.
- [13] Econbiz. last accessed: 27 July 2021. <https://www.econbiz.de/>.
- [14] Business Source Complete. last accessed 27 July 2021. <https://www.ebsco.com/de-de/produkte/datenbanken/business-source-complete>.
- [15] i2b2 TranSmart Foundation. last accessed: 23 March 2021. <https://i2b2transmart.org>.
- [16] Goobi. last accessed: 23 March 2021. <https://goobi.io>.
- [17] Kitodo. last accessed: 23 March 2021. <https://www.kitodo.org>.
- [18] TODO Group. *TODO guides*. last accessed: 29 March 2021. <https://todogroup.org/guides/>.
- [19] Täuscher K, Laudien S. Understanding platform business models: A mixed methods study of marketplaces. *European Management Journal*. 2017.
- [20] Gawer A, Cusumano M. *Platform Leadership, How Intel, Microsoft and Cisco Drive Industry Innovation*. Harvard Business Review Press. 2002.
- [21] Chesbrough H, Appleyard M. *Open Innovation and Strategy*. *California Management Review*. 2007; 50(1).
- [22] Shahrivar S, Elahi S, Hassanzadeh A, Montazer, G. A business model for commercial open source software: A systematic literature review. *Information and Software Technology*. 2018.
- [23] Asadullah A, Faik I, Kankanhalli A. *Digital Platforms: A Review and Future Directions*. Twenty-Second Pacific Asia Conference on Information Systems. 2018.
- [24] Balka K, Raasch C, Herstatt C. The Effect of Selective Openness on Value Creation in User Innovation Communities. *Journal of Production Innovation Management*. 2014;31(2).
- [25] Sharma S, Sugumaran V, Rajagopalan B. A framework for creating hybrid-open source software communities. *Info Systems J*. 2002; 12.
- [26] Shah S. Motivation, Governance, and the Viability of Hybrid Forms in Open Source Software Development. *Management Science* 2006; 52(7).
- [27] Capra E, Wassermann A. A framework for Evaluating Managerial Styles in Open Source Projects. *IFIP International Federation for Information Processing*. Volume 275; *Open Source Development, Communities and Quality*, 2008; 1-14.
- [28] Zaggi M, Schweisfurth T, Herstatt C. The Dynamics of Openness and the Role of User Communities: A Case Study in the Ecosystem of Open Source Gaming Handhelds. *IEEE Transactions on Engineering Management*. 2020; 67(3).
- [29] Bonaccorsi A, Rossi C. Comparing motivations of individual programmers and firms to take part in the open source movement: from community to business. *Knowledge, Technology & Policy*. 2006; 18(4).
- [30] Bacon J. *The Art of Community*. O'Reilly Media. 2013.

- [31] Buss A, Strauss N. *Online Communities Handbook: Building your Business and Brand on the Web*. New Riders. 2009.