Al-Qadisiyah Journal of Pure Science

Volume 26 | Number 2 Article 15

4-7-2021

Listento Me: An Audio Documentation for Open Source Software

Wurood Hassan Albayati
Al-Mustansiriyah University, Collage of Science, Computer Dept., Baghdad, Iraq,
w.albayati@uomustansiriyah.edu.iq

Mohammad Azeez Abdulhassan Alshomalic
Al-Mustansiriyah University, Collage of Science, Computer Dept., Baghdad, Iraq,
mohammada.abdulhassan@uomustansiriyah.edu.iq

Jason Holdsworth
James Cook University, information Technology Academic, QLD, Australia, jason.holdsworth@jcu.edu.au

Recommended Citation

Albayati, Wurood Hassan; Alshomalic, Mohammad Azeez Abdulhassan; and Holdsworth, Jason (2021) "Listento Me: An Audio Documentation for Open Source Software," *Al-Qadisiyah Journal of Pure Science*: Vol. 26: No. 2, Article 15.
DOI: 10.29350/qjps.2021.26.2.1264
Available at: https://qjps.researchcommons.org/home/vol26/iss2/15

Al-Qadisiyah Journal of Pure Science
ISSN(Printed): 1997-2490 ISSN(Online): 2411-3514
DOI: /10.29350/jops.

ListentoMe: An Audio Documentation for Open Source Software

Authors Names
a. Wurood Hassan Albayati
b. Mohammad Azeez Abdulhassan Alshomalic
c. Jason Holdsworth

Article History
Received on: 27/1/2021
Revised on: 14/2/2021
Accepted on: 23/2/2021

Keywords:
OSS, Open source software, Readme, Documentation, Documentation structure

DOI: https://doi.org/10.29350/jops.2021.26. 2.1264

ABSTRACT

Readme file is a short document associated with most open-source software (OSS) there were many attempts to uniform the structure of Readme file, this file is the main file most user read, it helps the user to decide upon using the software. In this paper we suggest using audio format for Readme file (ListentoMe), to do so we study the state of the art for Readme file in most popular GitHub repositories(or repos for short). 100 GitHub repositories documentation have been collected as a dataset. Extracting commits from this dataset results in 326793 messages. Next, commits regarding documentation have been isolated for further analysis. Finally, we observed the structure of readme file from the extracted commits and the number of key attributes used. The result shows that Github developers did not commit to the general Readme file structure suggested by literature [17] most developers will not save time to write documentation. Thus we suggest an audio version of Readme file named as ListentoMe, The structures for the ListentoMe file have been proposed to bridge the gap in documentation structures. To evaluate our suggestion, we conduct questionnaire targeting software developers in GitHub. The questionnaire results showed that more than 70% of developers prefer ListentoMe (audio version of Readme), they mostly agreed about the suggested ListentoMe structure.

1. Introduction

Open source software (OSS) is a new trend of built software. The main challenges that face OSS are the lack of documentation or un-updated documentation [26]. Documentation plays an important role, especially in complex software. It helps overcome problems that may occur through the software life cycle [14, 10]. In 2008 Loggem argued that quality of software documentation (followed ISO standards) could be produced by verifying documentation and following guidelines for good software documentation involved taking the nature of software into account as well as the specific problems encountered. Loggem indicates the lack of academic research in this area [24]. Capilla [2] showed the importance of architectural document

[a] : Al-Mustansiriyah University, Collage of Science, Computer Dept., Baghdad, Iraq, E-Mail: w.albayati@uomustansiriyah.edu.iq
[b] :Al-Mustansiriyah University, Collage of Science, Computer Dept., Baghdad, Iraq, E-Mail: mohammada.abdulhassan@uomustansiriyah.edu.iq
[c] : James Cook University, information Technology Academic, QLD, Australia, E-Mail: jason.holdsworth@jcu.edu.au

sustainability and from their research, they conclude the need for more effort towards standards and tools for documentations. OSS tends to lack accessible and complete documentation that keeps users [13]. The production of software documentation faces many issues the main issues are tight schedules and budget constraints, also developers are good in writing code but not documents [17, 6]. The developers spends their time on building software system while considered documentation as low-priority. For that in 2015 Nicoletti developed a tool that can help speed up the production of software architecture documents [16].

There are more than two documents that software may have, the name may differ but contents still the same. Although the literatures used different names for those documents, most researchers used documents specifies in Fig.1. From the literatures, we summarized important software documents that developers produce.

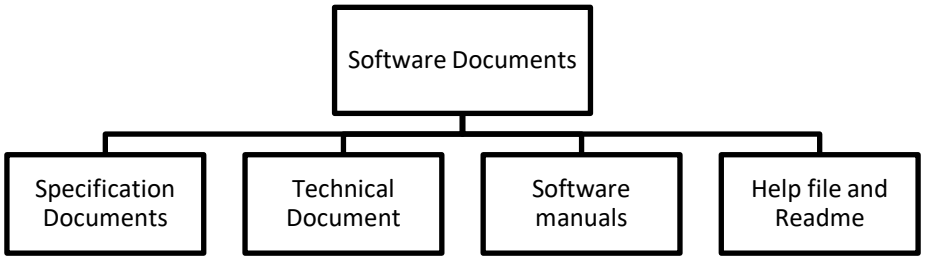

Fig. 1 software documents

Fig.1 shows four types of documents: specification documents this document mainly used by programmers and testers (both programmers and testers are types of developers), technical documents used by internal users, while software manuals targets all type of users, finally helps and Readme file also for users , they help users install use and troubleshoots main errors or conflicts may occur..

This paper focuses on the last block presented in Fig.1 which is Readmefile. The paper proposes using a structured audio version of the readme file instead of writing it, to move the developer's burden from writing a piece of short or too much information in the file. We keep the name close to the original well known readme file. Most developers feel bored or they lack experience required to write understandable and accurate documentation for their software [22]. We think the recording is more fun and less time consuming.

2. Literature Review

In OSS the responsibility of writing the documentation lay on developer's shoulders, developers may lack the skill of writing clear documents for their software [26], as could be seen in Fig.1, there are many types of documents need to be associated with the software some of the documents need to be updated frequently when the software change. Following up with all these tasks add more responsibilities. The most important type of documentation according to Sommerville [19] is process document (which involves specification and technical documents in Fig.1). Process document records maintenance and the process of development. Product documentation (corresponds to software manuals in fig.1) used after the production of the system [24].

Writing short documentation important to gain more users because users do not read long manuals [18] thus most OSS comes with Readme files and detailed documentation in WiKi's. Readme file is a short document associated with most OSS provides useful information about the software, there are a specific information expected to be included in each readme-files [23]. Readme file is not an exhaustive list of everything the software does, and often targets new users or those who are still deciding if they want to use the package, it created by a project's developers and hosted on the project's webpage/repository [3].

Documentation structure not always facilitate developer work, there is a gap between information needs by developers and the structure of documentation [22]. In 2013 Demeyer et al proposed a guidelines structure for software documentation [5] Demeyer agreed with John et.al. in their emphases on quality issues of software documentations [12].

John et al In 2006 [12] have mined data dump from Source Forge to extract information about open-source developer social media, John proved that developers relay heavily on social media when they work on a projects. Open source platforms facilitate developer communication and this increases productivity. Storey et.al. in 2012 confirms the impact of social media on OSS developers and investigated the social media influence in software engineering research and practice [11].

The widespread of OSS results in the development of many tools and environment to facilitate developer's collaboration and help in documentations. Wiki is one of those tools that provide a more friendly environment for documentation, most GitHub developers nowadays, (GitHub is the biggest OSS platform that hosts repository and provides collaboration tools for free), use wikis to include videos, images, and text to the documentation wiki help in associate different kinds of contents in a single document[1].

Developers increasingly using social media during and after developing a project [20]. Researches shows trends for developers to communicate using different collaborative technologies [15]. Menten et al in [15] proposed using audio for requirement documentation which is close to our proposal but we used audio documentation for recording basic information (same information found in readme file) in audio format.

3. The proposed audio document (ListentoMe)

Readme file sound to be easy and agreed upon in OSS environments (e.g. GitHub). It is a short document (or project description) attached to source code that gives a snapshot about the software [4,9]. In this manuscript, we observed the most popular GitHub repositories since 2018, (the more star the repo have, more popular it is). Next, commits (that represents developers and users communication messages regarding the repos) have been processed and analysed, then we find the number of occurrence of the key points (attributes) in Readmefile. According to [12, 9] there are ten attributes recommended by GitHub founder to be included in GitHub readme files.

Readme file targets the user, the more clear and precise information in it, the more likely it that the project (software) will be forked (download). In addition to Readme file in GitHub, a detailed documentation about the project will be available in wikis.

In this paper, dataset consists of 100 most popular Github projects has been collected and processed to extract commits (using keyword search- filter 1) which results in obtaining 326793 commits related to documentation of the projects. Then another filter applied (filter2) to identify commits regarding Readme file only which was 5953 commits. Fig. 2 below depicted the framework used to extract Readme file structure used in GitHub.

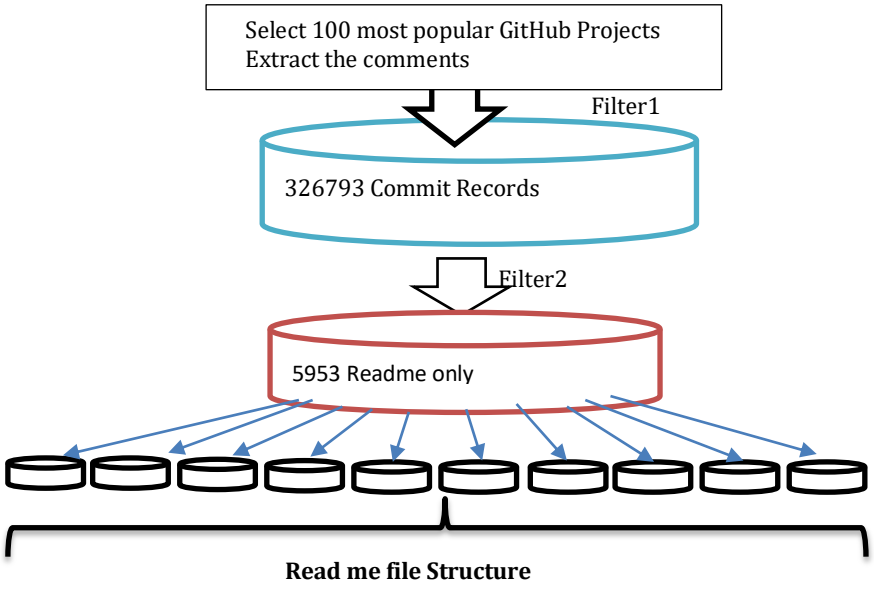

Fig. 2. The proposal framework

Analysing the commits of readme files in our selected dataset (100 repos) shows that more than 25% of developer's correspondence spend on editing and fixing issues related to Readme file, most of the commits were about clarifying some key points in the file, fixing some spelling errors and adding a key point's declaration.

As an open-source users, we judge on the project according to readme file contents [3], unfortunately, this file may miss a lot of basic information and sometimes developers have to read the whole documentation section to decide upon using the software. In this paper, we suggest using audio version of Readme file called ListentoMe, because many developers find it hard to dedicate time to write a text documents, as not all developers are as good in writing as they are in coding [3].

ListentoMe will reduce time and effort as a developer could record while amidst other coding tasks, or driving their way home. Based on the literature, there were key points (attributes) that should be declared in ReadMe files [9]. Analysing the state of the art for using these key points (or attributes) in the studied repos shows that not all developers considered declaring these attributes in their Readme file. The percentage of the occurrences of each attributes on our datasets depicted in Fig 3.

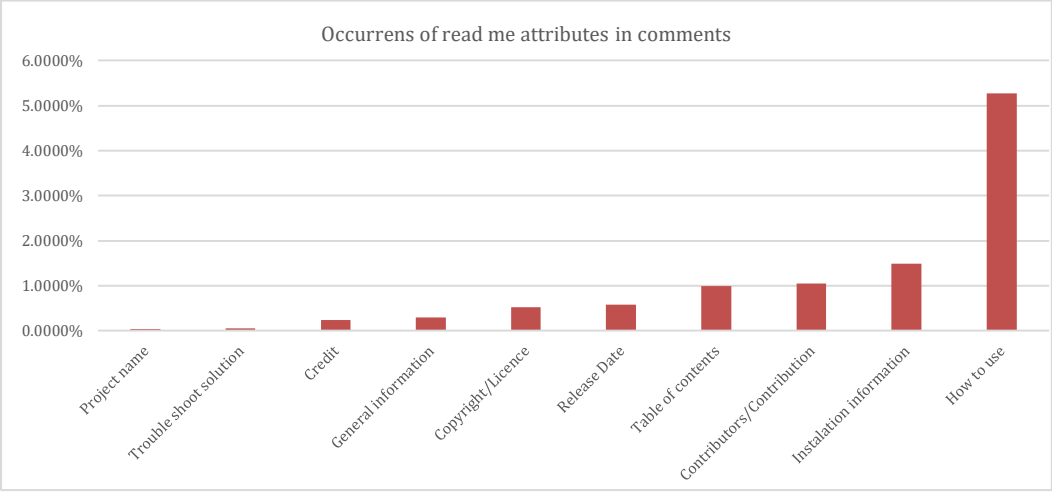

Fig. 3. The attribute occurrence in the 100 repos

The key points (attributes) are also considered when recording the ListentoMe file. We also propose a structure for the audio file –ListentoME. Fig. 4 below illustrated ListentoMe structure

prerequisites
general information
how to use
constraints
license and
contact us

Fig. 4 Listen to Me file structure

Fig. 4 shows that the basic information that the user is interested in seven only which we selected carefully from [9]:

- Software (or project) name.
- Prerequisites in this section the developer states what platform, operating system, or libraries need to be there in order for the software to work.
- General information a short definition of the software what's the input, when we need to use it, what is the goal.
- How to use, how user can install and use the software on his side.
- Constraints: is there any storage, hardware, or input constraints.
- License some software had a license and agreements user need to be agreed and aware about the license. The type of license should be specified in this section
- Contact us, it is an optional option.

A questioner have been used as an evaluation criteria to our proposal, a survey was delivered to a numbers of GitHub users asking them about audio documentation. We get only 58 responses that reflect different nationalities and opinions. Two groups of questions have been asked: one related to the audio version of readme file (if the developer will consider do so for his future application). The other group of questions rated the proposed audio file structure fig 5,6 and 7 illustrates developer response to the questioner.

4. **Results**

Readme file is a short documentation associated with most OSS. Based on the literature, software developers found it difficult to write documentation, on another hand, they use social media (audio or video conferencing) a lot to communicate, when developing a project many researchers suggest reusing developer communication history for documentation [22, 5, 8]. Developers seems to use social media a lot [25, 8] to communicate when developing a new project (software), this motivates us to start searching in audio documentation. Thus we propose an audio version of readme file with seven attribute (Prerequisites, general information, how to use, constrains, license and contact us) as a minimal requirement (see Fig.2). In order to evaluate the proposed file, we collected 58 developers' opinions using a questionnaire to evaluate audio documentation (ListentoMe).

The results for the question (Do you prefer to do a Voice recording or writing a Readme file when you do the documentation for your software?) shows that there are around 70% of developers prefer recording (audio documentation). When asking (What prevents you from developing audio version of Readme?), hesitating about audio recording, although. Also we ask developer about the selected key points (attributes- please see Fig.4) that should be included in the audio Readme file, The 48% agreed upon consider it as required, others suggest adding more information figs. 5,6 and 7 depicted these responses.

Therefore, we nominated seven attributes as depicted (in Fig.2) according to Kozlowski's recommendation [23], Kozlowski studied many Readme file structures then he concluded a readme file layout to be considered when writing such files. Hence, we think this information is enough for a user to decide upon using the software.

In the questionnaire we asked developers to give us feedback about the ListentoMe structure; more than 70% recommend and agreed to the suggested structure (positive feedbacks). In the questionnaire, we also asked many different GitHub developers from many backgrounds and nationalities and we are still expecting more responses.

5. **Conclusions and Future Work**

Although there were attempts to use audio documentation [15, 21], still in its early stages. According to the documentation most developers consider reading documentation is boring, users try products without reading its documentation [18]. ListentoMe is an attempted to bridge the gap in OSS documentation. In this paper we collected commits from 100 popular GiHub projects. Then we studied the state of art for readme in these projects . Our finding confirms the lack of key attribute information presence in those projects, thus 25% of the commits were related to key attributes that should be declared earlier in Readme files. If the developers spent more time in formatting and structuring their Readme file, they will probably save more time

and effort rather than spent it in responding to user commits. To facilitate developer efforts in producing easy and good Readme file we suggest an audio version of readme file, it is easier and more convenient for most developer[8,15, 20, 21]to use audio, Also a structure for it have been proposed .

To evaluate our proposal, we seek developer's opinions thus questionnaires were used. The questionnaire was posted to GitHub developers (GitHub is one of the biggest Open-source software hosting platforms for version control and collaboration, GitHub has millions of developers [7]). The questionnaire has 6 questions, the first three questions were general (asking general questions regarding developer gender, nationality, and programing language they familiar with., We asked about developer preference (audio versus writing). From 58 responses, about 70% of developers prefer audio recording, 25% percent prefer writing a Readme file, others did not care about the format they have no preferences. Fig. 5 illustrated the results of the questionnaire regarding this question..

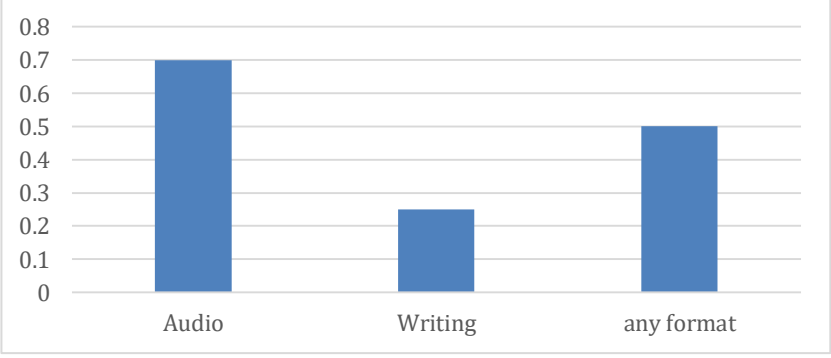

Fig. 5 Developer responses for their Readme file (Audio or writing)

The fifth question targets those who prefer not to record an audio Readme, the question was "What prevents you from developing an audio version of Readme?", 30% of the developers have their security awareness, others were more worried about the quality of recording(technology issues), some have their own personnel reasons. Fig 6 below depicted the responses for that question.

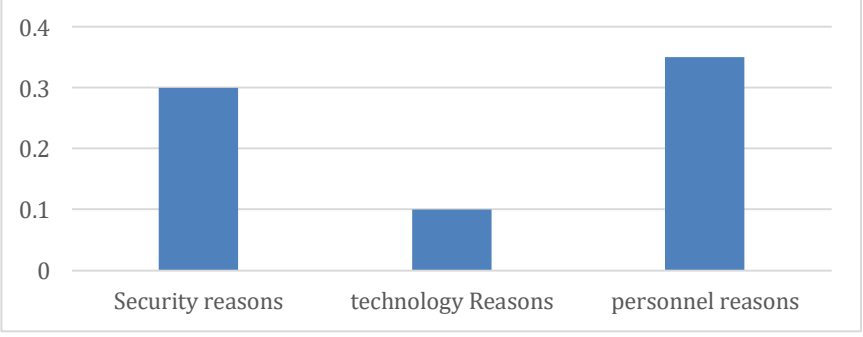

Fig. 6 Developer responses for the question " What prevents you from developing an audio version of Readme"

Finally, the developers asked to rate and suggest upon the seven key points (attributes) suggested in Fig.4. The developer agreed mostly on the suggested attributes(see fig. 7), positive feedbacks were received from the participants, this encourages us to pursue research in that field.

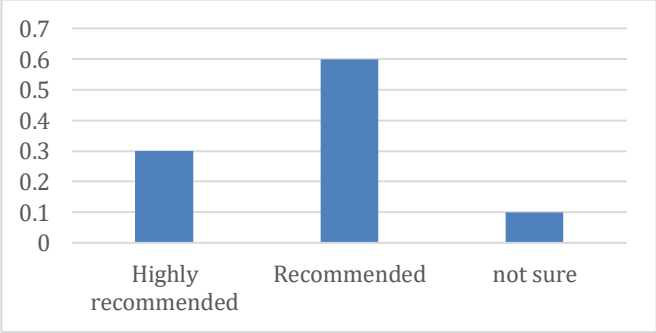

Fig. 7 Developer recommendation regarding the suggested key attributes.

We think audio documentation is a promising, easy, and fast way to do documentation; therefore, our future work will focus on: searching for more audio tools that could facilitate audio recording. Moreover, presenting a detailed study to employ the compressed audio file format for ListentoMe.

Acknowledgment

We would like to acknowledge a debt of gratitude to Mustansiriyah University, and College of Science, for the support and encouragements.

References

[1] Aguiar, A. and G. David. Wiki weaving heterogeneous software artifacts. in Proceedings of the 2005 international symposium on Wikis. 2005. ACM.

[2] Capilla, R., et al., 10 years of software architecture knowledge management: Practice and future. Journal of Systems and Software, 2016. 116: p. 191-205.

[3] Carvalho, N., A. Simões, and J. Almeida, DMOSS: Open source software documentation assessment. Computer Science and Information Systems, 2014. 11(4): p. 1197-1207.

[4] Carvalho, N. R., Simões, A., & Almeida, J. J. (2013). Open source software documentation mining for quality assessment. In Advances in Information Systems and Technologies (pp. 785-794). Springer, Berlin, Heidelberg.

[5] Demeyer, S., S. Ducasse, and O. Nierstrasz, Object-oriented reengineering patterns. 2013: Lulu. com.

[6] Geiger, R.S., et al., The types, roles, and practices of documentation in data analytics open source software libraries. Computer Supported Cooperative Work (CSCW), 2018. 27(3-6): p. 767-802.

[7] Han, J., Deng, S., Xia, X., Wang, D., & Yin, J. (2019, July). Characterization and Prediction of Popular Projects on GitHub. In 2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC) (Vol. 1, pp. 21-26). IEEE.

[8] Karolina Muszyńska. Gameflication of communication and documentation processes in project teams, Procedia Computer Science 176 (2020) 3645–3653.

[9] Kozlowski, W. (2014). Guidelines for writing''readme" style metadata. 2014-10-12]. http://data. research. comell. edu/sites/default/files/SciMD-ReadMe — Guidelines — v4-l_0. pdf.

[10] Korir, N.J.K.a.W.P.K., Importance of Software Documentation. IJCSI International Journal of Computer Science Issues,, 2013. 10(5): p. 223-228.

[11] J.B.a.P.J.H.S., The Economics of Open Source Software Development-Chapter 11:Application of Social Network. 2006.

[12] John P "Jack" Barr; Rosenbaum, S., Documentation and Training Productivity Benchmarks. Technical Communication; Washington, 2003. 50(4): p. 471-484.

[13] Levesque, M., Fundamental issues with open source software development. First Monday, 2004. 9(4).

[14] Luqi, et al., Documentation driven development for complex real-time systems. IEEE Transactions on Software Engineering, 2004. 30(12): p. 936-952.

[15] Menten, A., S. Scheibmayr, and L. Klimpke. Using audio and and documentation. in 2010 Third International Workshop on Managing Requirements Knowledge. 2010. IEEE.

[16] Nicoletti, M., S. Schiaffino, and J.A. Diaz-Pace, An optimization-based tool to support the cost-effective production of software architecture documentation. Journal of Software: Evolution and Process, 2015. 27(9): p. 674-699.

[17] Rafiq, M. and K. Ameen, Issues and lessons learned in open source software adoption in Pakistani libraries. The Electronic Library, 2009. 27(4): p. 601-610.

[18] Rettig, M., Nobody reads documentation. Communications of the ACM, 1991. 34(7): p. 19-25.

[19] Sommerville, I., Software documentation. Software engineering. 2001(2): p. 143-154.

[20] Storey, M.-A., et al. The impact of social media on software engineering practices and tools. in Proceedings of the FSE/SDP workshop on Future of software engineering research. 2010. ACM.

[21] Suzanne V. Blackley, Valerie D. Schubert, Foster R. Goss, Wasim Al Assad, Pamela M. Garabedian, Li Zhou, Physician use of speech recognition versus typing in clinical documentation: A controlled observational study, International Journal of Medical Informatics, Volume 141, 2020, 104178, ISSN 1386-5056.

[22] Treude, C., M.P. Robillard, and B. Dagenais, Extracting Development Tasks to Navigate Software Documentation. IEEE Transactions on Software Engineering, 2015. 41(6): p. 565-581.

[23] Turner, R.C., T.A. Douglass, and A.J. Turner, Readme. 1st: SGML for Writers and Editors. 1996: Prentice Hall PTR.

[24] van Loggem, B. Software documentation: a standard for the 21 st century. 2014. ACM.

[25] Vasilescu, B. (2014, May). Human aspects, gamification, and social media in collaborative software engineering. In Companion Proceedings of the 36th International Conference on Software Engineering (pp. 646-649). ACM.

[26] Yeates, altruism S., Documentation issues in open source. OSS Watch, 2006.