



## **Participation of male and female developers in the project development processes of selective Open Source Library Software: An evaluative study.**

Sanghamitra Jana Chatterjee  
PhD Scholar in IGNOU and In-Charge-Library,  
Centre for Women's Development Studies  
New Delhi – 110001  
Email: jsanghamitra@gmail.com; sanghamitra@cwds.ac.in

Supratim Chatterjee  
PhD Scholar in IGNOU  
Manager: Programs, American Center  
New Delhi – 110001  
Email: supratim.chatterjee@gmail.com

Uma Kanjilal  
Professor of Library and Information Science  
and Director, IUC; IGNOU  
New Delhi – 110068  
Email: ukanjilal@ignou.ac.in

### Abstract

While many researchers attempt to understand the processes and structure of the available Open Source Library Software (OSLS), their development, impact and management, few found a gender composition in participation of the development processes worth studying. This study observed, highlighted and compared the patterns of women's and men's participation in the OSLS project development process with select OSLS viz. DSpace, E Prints, Evergreen, Greenstone Digital Library Software, Koha, NewGenLib and OpenBiblio. To find out gender participation in development of OSLS projects, the methodology involved analytical as well as theoretical study. The dataset involved both secondary and primary level of data. An online questionnaire was developed for primary data collection. Analysis of the survey has confirmed that women do not participate to a great extent in the development of OSLS. There are differences between male and female participation with respect to their role and nature of involvement in the development process.

*Keywords:* Open Source Library Software, Gender Dimension, Project development

### **1 Introduction**

Information and communication technologies (ICT), like any other technologies, has an impact on men and women in different ways. Software is at the heart of the development of ICTs. In an



ICT-based society, it is increasingly important that software is designed to meet the requirements of diverse users. The development mode of Free/Libre Open Source Software (FLOSS) is another most common and successful examples towards this aim. FLOSS or FOSS has had a major impact on the computer industry since the late 1990s and has changed the way software is perceived, developed and deployed in many areas. Libraries are also part of it. Libraries have been using ICT based resources and services to satisfy the diverse information needs of their users. Libraries have been transforming to ICT based resources and services and gradually becoming more of infobraries. And open source software in library and information science is proving to be the major driving force in this regard.

## 2 Research background

Geleijnse (2002) looked into the effects that information technology has had, and could have, on the goals, objectives and organisation of libraries. He identified three phases in the development of digital libraries. In the first phase, most of the universities and libraries move to electronic resources, enter into license agreements with vendors and publishers and start to digitise important collections. In this phase, many institutions invest large sums of money in local developments. This amounts to repeatedly reinventing the wheel, since the same things would be done at many different locations. However, in the second phase discovery of important new opportunities to support teaching, learning and research, and new roles for the library are being identified. Developments in the area of E-learning and E-publishing are key elements in this phase. The third phase is a phase of more cooperation on common tools and infrastructure and of more customer orientation. It is the second or third phase where the Library Open Software plays a very important part. From the above observation it can be inferred that with the development of ICT, there is a change in the pattern of services provided by the library and also there is a change in the software being used in the Library. Different authors talked about open source software as becoming a preferred solution, owing to the freedom to copy, modify and distribute it and the freedom from contracts, as well as for greater opportunities for interoperability with other applications. But that compliance of moving towards more technology dependency is shown not equally by both the genders.

Michlmayr (2007) researched and established in his thesis that free and open source software, or FOSS, is typically developed in a collaborative fashion and the majority of contributors are volunteers. Even though this collaborative form of development has produced a significant body of software, the development process is often described as unstructured and unorganised. His dissertation studied the FOSS phenomenon from a quality perspective and investigates where improvements to the development process are possible.

Gilbertson (2011) found out that Free/Open Source Software (FOSS) is a powerful development paradigm for creating software. Claire compared several surveys and have commented that existing FOSS communities are very homogenous populations and made up of mostly men. Open Source Library Software (OSLS) being an integral subset of FOSS are also a party to these studies.



While many researchers attempt to understand the structure, domain, functionalities of the available OSLS, even their development, impact and management, few studied the gender participation and tried to check whether there is any specific motivational factor responsible for participation or non-participation towards any particular gender. OSLS is a way of working within a library organisation to derive benefit for the organisation from technology choices that promote freedom, collaboration, co-operation and openness. To maximise benefit from an OSLS strategy, an organisation must develop its OSLS ecosystem, whether this consists of internal capacity or partnerships with strategic business partners who have developed their own OSLS ecosystem. There are a number of areas within libraries where OSLS tools are often used, and can be deployed or adapted to improve the quality of experience by librarians and patrons alike. These include the operating systems that run our computers, the desktop productivity applications that we use, our institutional repositories, library information systems, and various web and distributed applications. Its success has also attracted researchers from different disciplines to analyse its unconventional innovation approach. But according to a FLOSS survey on FLOSS developers in 2002, “women do not play a role in the [FLOSS] development; only 1.1% of the FLOSS sample is female.” (Ghosh et al. 2002).

It is claimed that open source development produces more bug-free code, faster, than closed proprietary developed code, although this has yet to be conclusively demonstrated (Jorgensen, 2001; Koch, 2002; O'Reilly, 1999; Raymond, 1999; Sharma, 2002; Stamelos, 2002; Wang, 2001; Wu, 2001). Open source software development teams, are generally comprised of volunteers working not for monetary return, but for the enjoyment and pride of being part of a successful virtual software development project. Team members often come from around the world and rarely meet one another face-to-face. The open source projects are self-organised, employ extremely rapid code evolution, massive peer code review, and rapid releases of prototype code. Many of these practices are counter intuitive and the opposite of what conventional software engineering holds as the correct processes for the production of high quality code (Bollinger, 1999; Charles, 1998; Edwards, 1998; Fielding, 1999; Hecker, 1999; Lawrence, 1998; O'Reilly, 1999; Ousterhout, 1999; Payne, 2002; Raymond, 1999; Sanders, 1998; Torvalds, 1999).

The above literature studies intend to point out that while women have a chance of upgrading their ICT skills and knowledge through participating in the development of the Open Source Library Software, but the opportunity is less available due to their latent participation in the process.

The importance of FOSS/ OSLS for advancing human development, especially in poor countries cannot be overemphasised. OSLS is about sharing knowledge. The Information Meeting for Permanent Delegations of the Information Society Division Communication and Information Sector held at UNESCO Headquarters, 13 January 2011 it has pointed out that there might be gender exclusion in the FOSS world. Even though there is less participation by women in FOSS/ OSLS projects, the FOSS community is one of constant innovation, sharing and spreading of knowledge. There is extensive work on the Library Open Source Software (e.g. Gairin et al. 2008 and Crow, R. 2004 etc.) but all of them speak about the features and other details not about the gender exclusion in the development process. WSIS (World Summit on the Information Society) Knowledge community has created a community of Gender Equality in Free and Open Source Software. The Community aims at creating a network of different nature in situations and



projects that deal with the issue of the under-representation of Women in Technology and more deeply with the Gender Gap in FOSS.

The status of women participation in the OSLS need to be studied and analysed in terms of the current obstacles against bringing more women, who are technically equal competent, to participate in the OSLS development.

### **3 Objective**

The objective of this research paper is to find out the aspect of male and female participation in development processes of seven major OSLS projects. Here 'project' is defined as the all public activities performed for the intentional benefit of the respective OSLS. The seven prominent OSLSs taken for the present study are DSpace, E Prints, Evergreen, Greenstone Digital Library Software, Koha, NewGenLib and OpenBiblio

### **4 Research methodology**

For Primary data –An online questionnaire was developed by using PHP code. PHP code has been embedded into HTML code so that it can be used in combination with various web template systems, web content management systems and web frameworks. Primary level data have been collected through stratified sampling method. Strata include different categories of individuals based on their involvement in the open source software for libraries at different stages of development. The sample have been worked out by taking representatives from Developers at various levels. Questionnaire was formulated and disseminated to collect the required data. Developers playing the following roles were considered for the study:

Coding, Debugging, Testing, Algorithm design, User/Reader Communication, Social task making, Providing user support, Project maintainer, Public relations, Marketing, or promoting or organizing of event, Document Creation, Translating, Providing feedback on bugs, feature needed and enhancement, Peripheral developing, Active Developing, Core Developing, Project Leading and Pioneering.

For the validity or authenticity and reliability of data collection, the users and developers listed in the official website and official wiki of the respective OSLS site are only contacted. Thus there is representation from majority of the countries across the globe. This study is based on 99 validated replies.

For the purpose of better sample coverage of data collection, 'ALA Women and Gender Studies Section', 'IFLA Women and Information and Libraries Special Interest Group Mailing List (will-l)' were approached.

Proper categorization, coding, tabulation etc. has been developed to analyses the objectives against collected data. Exploratory Data Analysis (EDA) has been used to provide visual summaries of data to have quick and complete description of essential information. A measure of central tendency has been applied to questionnaire specific collected data.

For Secondary data – The notion that there is a Gender Inequality is evident in the secondary data sources like European Commission reports, postings in WSIS Knowledge communities, reports of Information Meeting for Permanent Delegations of the Information Society Division Communication and Information Sector and also reports of Community Leadership Summit and various books and journal articles.

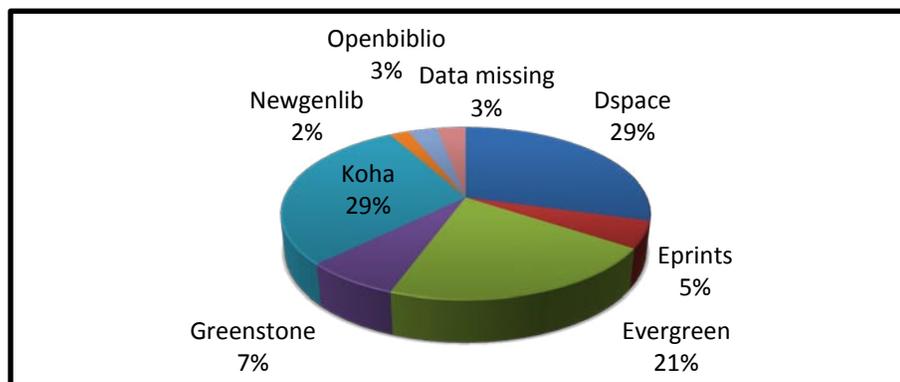
## 5 Research analysis

Amongst twenty six questions inquired in the questionnaire for the OSLS developers, four questions pertaining to their participation in the projects and role in different project have been discussed. A number of sets of answers for each question were set up and coded the responses of each individual. There is one ID code generated automatically against the response of each individual. The coding may be considered a method of summarizing the responses of each individual

## 6 Data on the name of the project to which currently (or most recently) contributed

The data regarding the name the project to which currently (or most recently) contributed has been captured by a question in the questionnaire. As per the obtained responses it is found that majority of the respondent developers have contributed in both DSpace and Koha. This is followed by those contributing Eprints. The lowest number of respondents has contributed in NewGenLib.

The data is represented graphically below with the help of a pie chart.



*Figure: 1 Data on the name of the OSLS project currently contributed*

In the given pie diagram it is getting reflected that the percentage of all the responded developers are rounded off. Almost 58% of them have contributed in DSpace and Koha respectively.

The data collected on the name of the OSLS project currently contributed by the respondent developers is further segregated between male and female respondent. The data is being captured and represented in the given table below.

**Table 1**  
**Data on the name of the OSLS project currently contributed for Male and Female Respondent Developers**



No. of Projects	No. of Female Respondent Developers	Percentage	No. of Male Respondent Developers	Percentage
DSpace	3	27.27	26	29.55
Eprints	1	9.09	4	4.55
Evergreen	1	9.09	20	22.73
Greenstone			7	7.95
Koha	5	45.45	24	27.27
NewGenLib			2	2.27
Openbiblio	1	9.09	2	2.27
Data missing			3	3.41

In the given table most of the female respondents have contributed in Koha which is followed by DSpace. Majority of the male respondents have contributed in DSpace. The percentage represented for male and female developers are calculated for their respective aggregates.

To check the statistical relevance of the data the central tendency has been measured in the form of Mode. Here mode is calculated as it is the value that appears most often in a set of data. As observed in the graph 1 it has a discrete probability distribution. The mode of a discrete probability distribution is the value  $x$  at which its probability mass function takes its maximum value. Here we have two names with maximum value. So the mode of the name of the OSLs project currently contributed:

$x = 29$  i.e. mode is valued at both DSpace and Koha

The mode of the name of the OSLs project for the male respondent developers currently contributed:  $x = 26$  i.e. mode is valued at DSpace for male respondent developers.

The mode of the name of the OSLs project for the female respondent developers currently contributed:  $x = 5$  i.e. mode is valued at Koha for female respondent developers.

## 7 Data on the name of the first project to whichever contributed

The data on the name of the first project to whichever contributed has been captured by a question given in the questionnaire.

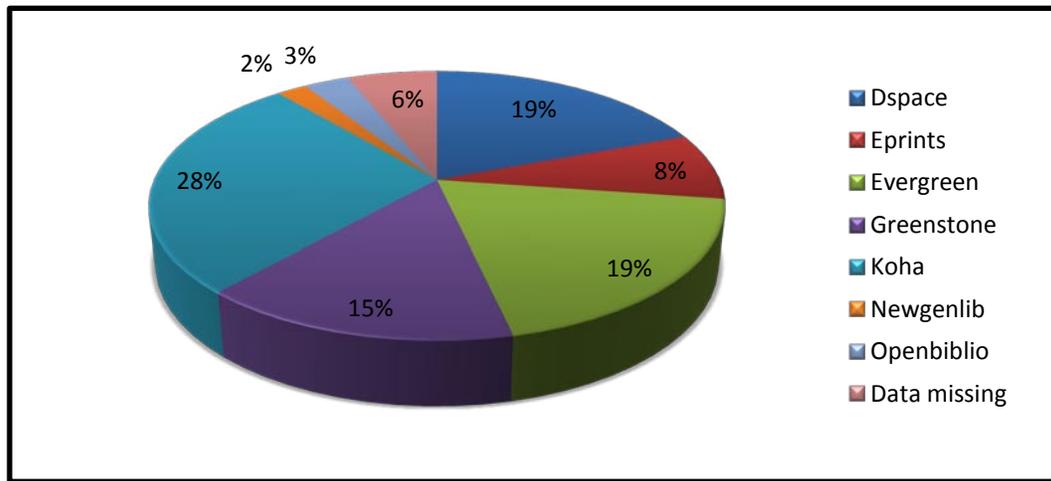


Figure: 2 Data on the name of the first OSLS project which ever contributed

The percentages of all the responded developers are rounded off for the diagram. In the given pie diagram it is getting reflected that the majority of the responded developers have contributed in Koha as their first project which is followed by DSpace and Greenstone. Almost 6 % of the data is missing. The least number of respondents have started with NewGenLib.

The data collected on the name of the first OSLS project whichever contributed by the respondent developers is further segregated between male and female respondents. The data is being captured and represented in the given table.

The segregated data of the Male and Female Respondent developers is given below.

No. of Projects	No. of Female Respondent Developers	Percentage	No. of Male Respondent Developers	Percentage
DSpace	1	9.09	18	18.18
Eprints	2	18.18	6	6.06
Evergreen	1	9.09	18	18.18
Greenstone	2	18.18	13	13.13
Koha	3	27.27	24	24.24
NewGenLib		0.00	2	2.02
Openbiblio	1	9.09	2	2.02
Data missing	1	9.09	5	5.05



In the given table most of the female respondents have contributed in Koha which is followed by DSpace. Majority of the male respondents have contributed in Koha.

As observed in the Table 2 it is a discrete probability distribution. The mode of the name of the first OSLS project whichever contributed:  $x = 27$  i.e. mode is valued at Koha

The mode of the name of the first OSLS project which ever contributed for the male respondent developers:  $x = 24$  i.e. mode is valued at Koha for male respondent developers.

The mode of the name of the first OSLS project which ever contributed for the female respondent developers:  $x = 3$  i.e. mode is valued at Koha for female respondent developers.

### **8 Data on the role played in the current and the first ever OSLS project**

The data regarding the role played in the current and the first ever OSLS project has been captured through a question in the questionnaire.

The given table indicates the role played by the respondent developers in the current and the first ever OSLS projects. The options given are Coding, Debugging, Testing, Algorithm design, Communication/ Social task making, Providing user support, Project maintainer, Public relations, marketing, or promoting or organising of event, Document Creator, Translating, Providing feedback on bugs, feature needed and enhancement, Peripheral developing, Active Developing, Core Developing, Project Leading and Pioneering.



Table No: 3 Data on the role played by the Male and Female Respondent Developers in the current and the first ever OSLs project										
Role	Number of Respondent Developers for -									
	Current/ Most Recent Project (CRP)	Current/ Most Recent Project & First Project (CRP & FP)	First Proje ct (FP)	Combi ned Percen tage	Not particip ated in the role	Current/ Most Recent Project (CRP)	Current/ Most Recent Project & First Project (CRP & FP)	First Proje ct (FP)	Combi ned Percen tage	Not particip ated in the role
	Female					Male				
Coding	2	1		27.27	8	17	43	3	71.59	25
Debugging	3	1		36.36	7	17	31	10	65.91	30
Testing	1	4		45.45	6	19	9	19	53.41	41
Algorithm design	3	1		36.36	7	21	12	3	40.91	52
Communication/ Social task making	3	6		81.82	2	5	2	2	10.23	79
Providing user support	4	1	1	54.55	5	11	9	11	35.23	57
Project maintainer	2	2	1	45.45	6	10	4	6	22.73	68
Public relations, marketing, or promoting or organizing of event	4	3	2	81.82	2	6	4		11.36	78
Document creator	3	2	1	54.55	5	8	2	3	14.77	75
Translating	2		1	27.27	8	4	2		6.82	82
Providing feedback on bugs, feature needed and enhancements	3	1		36.36	7	10	4	3	19.32	71
Peripheral developing	1	2	2	45.45	6	4	19	15	43.18	50
Active developing	1	3		36.36	7	22	24		52.27	42
Core Developing	1	2		27.27	8	34	13	1	54.55	40
Project leading	2	1		27.27	8	25	2		30.68	61
Pioneering					11	1			1.14	87

The above table is a representation of the male and female segregated data. The majority of the female respondent developers working on their first project and current or most recent projects



have played role either in communication/ social task making (81.82%) or in public relations, marketing, or promoting or organising of events (81.82%) for their software. It is also observed that the least number of female respondent developers are involved (27.27%) on coding, translating, core developing and project leading while working on their first project and current or most recent projects. On the contrary, the majority of the male respondent developers are involved in coding (71.59%) and debugging (65.91%) and least number played role in communication/ social task making (10.23%), translating (6.82%) and pioneering (1.14%).

## 9 Time spent for contributing in different project related activities

This question has six options to choose by the respondents i.e. whether the time spent in the current/most recent project involves contribution in Coding, Debugging/Testing, Designing algorithms or user interface, Communication with other programmers, Documenting the code, Public relation or promoting the project. Each option in turn has four alternatives to pick. Whether the time spent is all, majority, a substantial amount or a bit.

Table 4 Data on the time spent working in the current/most recent project for contributing to various following factors.		
<b>For Coding</b>		
Response rate is 73.74%		
Options	No of Responses	Percentage
All	4	5.48
Majority	65	89.04
A substantial amount	3	4.11
A bit	1	1.37
<b>Debugging/Testing</b>		
Response rate is 51.52%		
Options	No of Responses	Percentage
All	2	3.92
Majority	41	80.39
A substantial amount	8	15.69
A bit		
<b>Designing algorithms or user interface</b>		
Response rate is 36.36%		
Options	No of Responses	Percentage
All	2	5.56
Majority	17	47.22



A substantial amount	13	36.11
A bit	4	11.11
<b>Communication with other programmers</b>		
Response rate is 37.37%		
Options	No of Responses	Percentage
All	2	5.41
Majority	18	48.65
A substantial amount	13	35.14
A bit	4	10.81
<b>Documenting the code</b>		
Response rate is 44.44%		
Options	No of Responses	Percentage
All		
Majority	19	43.18
A substantial amount	21	47.73
A bit	4	9.09
<b>Public relation or promoting the project</b>		
Response rate is 35.35%		
Options	No of Responses	Percentage
All	3	8.57
Majority	21	60.00
A substantial amount	6	17.14
A bit	5	14.29

In the above table it is observed that working on their current or most recent project, respondent developers by far spend most of the time in coding their software which is followed by debugging and testing the software. Many respondents also spent a substantial amount of time documenting the code or communication with other programmers.



The data for female responded developers represented graphically below with the help of bar diagram.

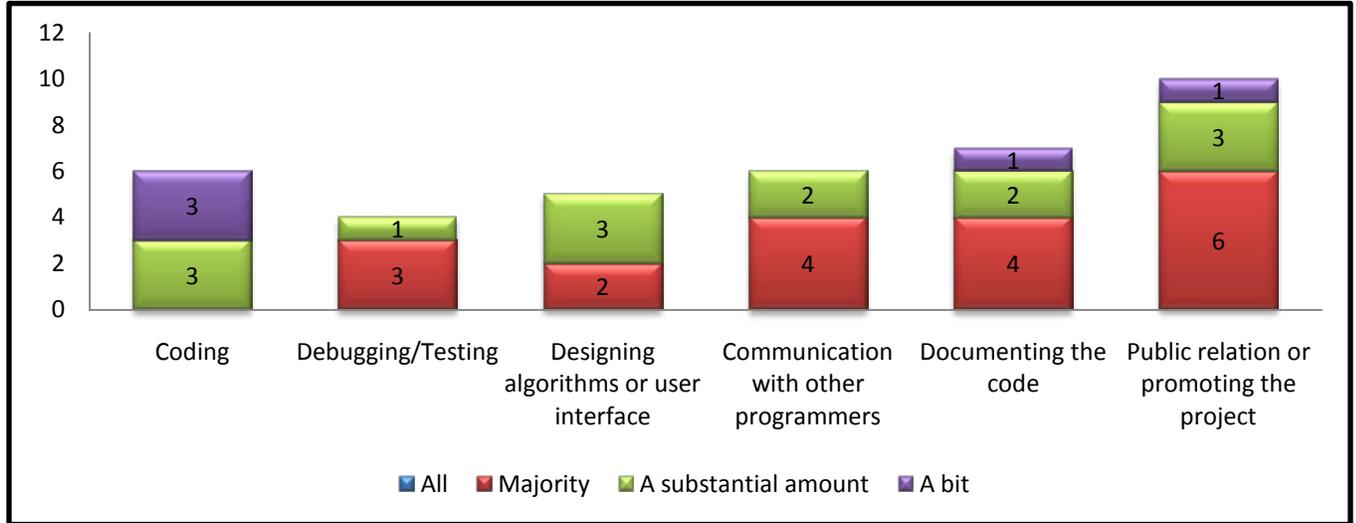
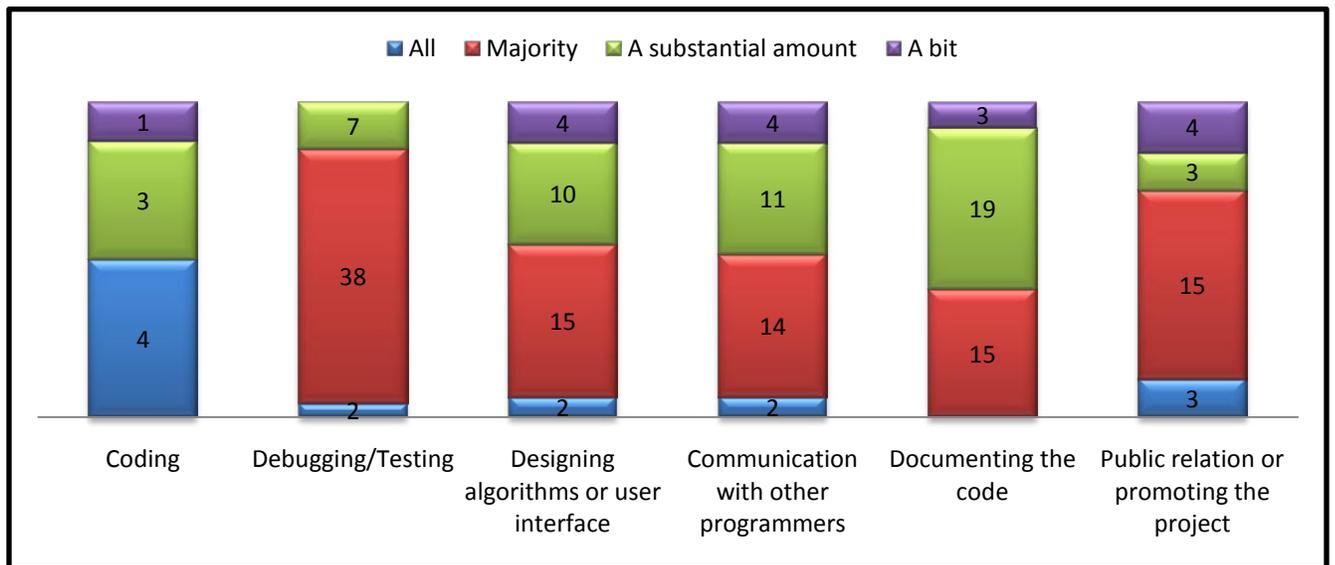


Figure:3: Data on the time spent by the female respondent developers working in the current/most recent project for contributing to each of the following factors.

From figure 3, it is observed that majority of the female responded developers have spent time on public relation or promoting the project, communication with other programmers and documenting the code. The female respondent developers have spent least time on designing algorithms or user interface.

The data for male responded developers represented graphically below with the help of bar diagram.





*Figure 4: Data on the time spent by the male respondent developers working in the current/most recent project for contributing to each of the following factors.*

From figure 4 it is observed that majority of the male responded developers have spent time on debugging/testing and coding. The male respondent developers have spent least time on public relation or promoting the project.

## 10 Conclusion

In the above four questions, the major concern is made to the OSLS developers' overall experience in OSLS projects, because all these tables and figures together indicates the role of the OSLS developers over a longer period, while their current project involvement may be influenced by accidental or casual circumstances. Therefore, this study concentrates on the overall number of projects performed and on the duration of membership through their participation in the OSLS community. Occupational issues must have an impact on project involvement, too. As already mentioned in the methodology that this article is a part of a major research project which has confirmed that women do not participate in great extent in the development of OSLS; only 11% of the respondent OSLS sample is female.

The role of programmers decreases with growing project performance. Astonishing is the role of female OSLS developers for public relations, marketing, or promoting or organising of event and other occupations that are not specified here. Their shares increase considerably with growing OSLS project experience. Thus, these groups of OSLS developers may not play a major role in terms of quantity, but they are obviously very important for the community in terms of OSLS project organisation and performance. Finally while working on their current or most recent project, respondent male developers by far spend most of the time coding their software which is followed by debugging and testing the software. However, majority of the female responded developers have spent time on public relation or promoting the project, communication with other programmers and documenting the code. The female respondent developers have spent least time on designing algorithms or user interface.

The female OSLS developers are underrepresented and are amongst those who may not have led a project, but show average shares among those who do the other works in the projects. Probably, this engagement of female OSLS developers would explain largely the findings about the role of gender dimension in OSLS project.

The dynamic OSLS world is unquestionably attracting potential developers to contribute and enhance this domain. However, participation of women developers is very less in proportion. One reason might be due to the fact that women particularly find it unapproachable to get started with contributing to OSLS. For firmly supporting the OSLS women developers and addressing their questions outreach programs are required for throwing light on the issues that they usually



have. It should be intended to bring women in the arena of OSLS development and elaborate all comprehensive opportunities array ranging from coding to research. It is needed to inspire potential women developers and showcase women who could have made noteworthy contributions to technology. This type of outreach should target not only females but academicians, students in university and colleges, open source developers and enthusiasts, supportive men, thought leaders and other underrepresented groups for widespread inclusion. It should aim to encourage female developers and users to involve into developing aspects. They need to converse the vast range of available opportunities.

## REFERENCES

Colford, S. (2009), Explaining free and open source software. *Bulletin of. American. Society of Information Science and Tecnology*, Vol-35, Issue - 2: Pp-10–14.

Deckelmann, S.M. (2008) *Women Who Code - Where Are They?*  
(<http://www.chesnok.com/daily/2008/04/02/women-who-code-where-are-they/>)

Geleynse, J. P. J. "Information technology as an agent of change." (2002).  
(<https://www.oecd.org/edu/imhe/23274588.PDF>)

Gilbertson, Claire E. *Gender and participation in early stages of the free/open source software joining process*. Diss. 2011.

Ghosh, R.A.; Robles, G.; Glott, R. (2002). "*Free/Libre and Open Source Software: Survey and Study Part V*". Maastricht: International Institute of Infonomics.

Jana, Sanghamitra, Supratim Chatterjee and Uma Kanjilal. " Study of different factors leading to the choice of Developers and Users to participate in major Open Source Library Software (OSLS) with special emphasis on gender differences" *International Journal of Research in Library Science*, 3.1 (2017): 150-161.

Jana, Sanghamitra, Supratim Chatterjee and Uma Kanjilal. " Analyzing various choice factors leading to the selection of certain Open Source Library Software (OSLS) with emphasis on Gender Differences." National Conference on trends, challenges & future of library and informaion science profession in digital era, *April 16th, 2017*.

Lakshané, Rohini. The trouble with being a woman in FOSS  
<https://deepdives.in/the-trouble-with-being-a-woman-in-foss-75181981bfdd#.q1n8df35h>



Landry, John; Rajiv Gupta (September 2000). "*Profiting from Open Source*". Harvard Business Review. doi:10.1225/F00503.

Michlmayr, M (2007). "Quality Improvement in Volunteer Free and Open Source Software Projects Exploring the Impact of Release Management" Retrieved 2015-03-30 (<http://www.cyrius.com/publications/michlmayr-phd.pdf>)  
*Open Source Initiative*. "The Open Source Definition (Annotated)". *opensource.org*. Retrieved 22 July 2016.

Open Source Pioneers Meet in Historic Summit. *1998-04-14*. Retrieved 2014-09-20.

Open Source Software Institute. "Home". *Open Source Software Institute*. Retrieved 22 July 2016.

Open Source vs. Proprietary Software. *PCWorld Business Center*. *Pcworld.com*. 2011-11-03. Retrieved 2012-03-25.

Open Standards, Open Source Adoption in the Public Sector, and Their Relationship to Microsoft's Market Dominance by Tony Casson, Patrick S. Ryan :: SSRN". *Papers.ssrn.com*. Retrieved 2015-03-30.

Rosen, Lawrence. "Joint Works - Open Source Licensing: Software Freedom and Intellectual Property Law". *flylib.com*. Retrieved 22 July 2016.

Rothwell, Richard (2008-08-05). "Creating wealth with free software". *Free Software Magazine*. Retrieved 2008-09-08.

SCILAB License. *INRIA*. Retrieved 2008-01-04.

Vance, Ashlee (2007-07-25). "SugarCRM trades badgeware for GPL 3". *The Register*. Retrieved 2008-09-08.

Verts, William T. (2008-01-13). "Open source software". *World Book Online Reference Center*. Archived from the original on January 1, 2011.

Wallach, Hanna M., Women in Free/Open Source Software Development (<http://people.cs.umass.edu/~wallach/talks/jitp.pdf>)