

Women in Debian and Free Software

Magni Onsøien

16th June 2005

1 Introduction

The first part of this paper is a short overview of some relevant literature. The list is by no means complete and a lot of relevant literature has been deliberately skipped. Then there are some statistics of men's and women's participation in different mailing lists, compared to numbers from bigger surveys among developers of free and open source software, before the paper concludes with some remarks about why women are an important part of the developer force.

2 Literature

Quite a few articles have been written about Open Source development and developers in general (see e.g. [2, 8, 4, 10]). They are about development methods, developers, philosophy, technology, specific software, and many other issues connected to free and open source software. Rather few has dealt with gender and open source development.

Some studies that has given important statistics about women and software development are the Free/Libre/Open Source Software (FLOSS) Reports from Europe [14], USA [1] and Asia [13] (accompanied by [11]), and a study from TU-Berlin [9]. These studies were conducted as surveys among Open Source developers, and among the participants there were only 1.5% female developers.

Numbers on female developers in the Industry are hard to come up with, mainly because the term "developer" is not very well defined for statistical purposes, especially not for comparison between different surveys. Some surveys and statistics show there is between 20 and 35% women in the industry (e.g. [3, 5]).

There has however been quite a few surveys and studies on girls and women in technology and ICT. Many of them have been internal studies for the admissions department of universities, but some are more extended and serve a general purpose, e.g. [7]. Even though they deal with students in schools and universities they are interesting for our purposes because there are probably similar problems and challenges also in the Open Source community.

3 The numbers

3.1 Background

As part of my master thesis work, I have analysed open source project developers' mailing lists with regard to gender of the participants. The analysis is done

by hand, since there are no good tools for a more automatic way of doing it: had the lists had participants from only one country (or language) it would have been much easier to do an automatic categorization. This is however not the case with most open source projects: many names bear no gendered references to me, while others belong to different genders according to what country or language they are used in.

I plan to make a database with all the names I have collected and categorized so far, to assist the categorization, but this is not yet done. Overall I managed to analyse about 70% of the names.

3.2 Analysis

I chose to analyse a few developers' mailing lists during three months. The raw data is divided in male, female, unknown (uncertainty about the gender of a person), and uncategorized (not enough data to decide gender, e.g. only email address, nick, or initials). In *this* table only the distribution of gender and the total number of participants during the period is included. See table 1.

Debian has 965 developers¹. Of those only 3 are women - or 0.31% female developers. Gentoo Linux has 3 female developers among ca. 375 in total, which yields 0.80% women. The Gnome Foundation, which in some ways is comparable to Debian Developers, has 18 female members², which is about 4% of the members. See table 2.

Much of the difference between the Gnome Foundation, and Debian and Gentoo may be due to the higher emphasis on translation and documentation as a base for membership in the Foundation. In Debian translation is a typical example of an activity being done much by people who are not DDs.

During the period between November 2004 and January 2005 only 7 of 844 authors - 0.82% - on the debian-devel list were women. At the moment 9 women are in the New Maintainer queue. If all those are getting through the process (and the number of male candidates doesn't increase equally), there may be around 1% female DDs - which is still not a very high number.

For other large free software projects - KDE, Gnome, Gentoo, Apache, the Linux kernel - the numbers are comparable to the mail Debian developer list: between 0.1% and 1.6% women.

It should not come as much of a surprise that the female participation on the debian-women mailing list is significantly higher than on the other lists, albeit men are in majority also there.

According to the FLOSS studies, the percentage of female free software developers is about 1.5%. The participation in most of the bigger samples I have looked at is below this, while in some of the smaller samples, like httpd-dev, kde-core-devel, and gnome-devel, the participation of women is around the expected average.

3.3 Other surveys

According to several international studies [1, 9, 14, 13] there are 1.1 – 1.6% female free/libre software developers, and I consider this to be an accurate number. Compared to the number of women entering Computer Science studies

¹As of 2005-06-15, according to http://www.debian.org/vote/2005/vote_001

²As of 2005-05-03

	Female	Male	Total
debian-devel	0.83%	73.82%	844
debian-women	40.78%	51.46%	103 ^a
debian-vote	2.32%	90.69%	129 ^b
httpd-dev	1.32%	69.73%	152
gentoo-devel	0.63%	77.99%	318
gnome-desktop	0.74%	75.30%	270
gnome-devel	1.61%	75.81%	62
kde-core-devel	1.26%	79.08%	239
kde-devel	0.52%	70.80%	387
linux-kernel	0.13%	68.88%	797

Table 1: Gender in mailing lists (percentage analysed)

^aMarch – May 2005

^bMarch – May 2005

	Female	Male	Total
Debian	0.31%	99.69%	965
Gnome Foundation	4.04%	91.03%	446
Gentoo	0.80%	99.20%	375

Table 2: Gender in developers teams

(about 20%, but varies a lot between universities and countries), this is a very low number, especially since education is probably one of the most important entry-points into free software development. In the “IT industry” there are about 20 and 35% female workers [3, 5], depending on what jobs are counted.³

These numbers may be inaccurate. Different research include different groups in “computer technology” so it is not easy to compare them accurately and their results will differ quite much. A qualified guess is that about 20% of programmers and developers in the *industry* are female. The number of *free software* developers come from a few surveys where the participants defined themselves as developers. There may have been a tendency to over- or under-definition of oneself as a developer, but the numbers seem steady in all the three surveys and are probably rather accurate. It is commonly believed that men tend to overrate themselves and women tend to underrate themselves. In this context it means that a woman is more likely to not consider herself a developer, even when she is, and vice versa for men. Since females are probably more likely to underestimate their own value and activities, and may also be more reluctant to participate in a survey of a rather male dominated field, the proportion of female developers may be a bit too low.

The overall number of developers in a project is also usually small enough that even small changes in the absolute numbers gives huge changes in the relative ones in most projects. Debian, with about 1000 developers, is an exception to this. The project is big enough that small changes are hardly detectable, while a trend can be observed.

³Numbers are hard to collect and compare, since there is no international statistical definition for the kind of IT workers - programmers and developers - that we are talking about here.

4 Why women?

Do women matter? A mixed environment has influence on the behaviour of other people. Research is done on this, e.g. the influence of women in Antarctica many years ago, when women were allowed in the research bases, and violence and drinking went considerably down and male workers started to “behave”. [12, s. 17]

Sometimes the open source community may seem like Antarctica: a hostile environment, full of brave warriors there to fight against the forces. On the other hand you have the penguins, and the community, a tight community of people who share survival and adventures, who depends on themselves and each other.

In 1914 Ernest Shackleton was looking for men for his Antarctica expedition: *Men wanted for hazardous journey. Small wage, bitter cold, long months of complete darkness, constant [...] Safe return doubtful [...] and recognition in case of success.* He got 5000 applications for 56 jobs. [12, p. 17]

Maybe is the modern open and free software developer a bit like the adventurer more than 90 years ago: into it for the adventure, not too concerned about dangers and uncertainty coming up, and happy with glorification after success.

Modern research has, however, shown that a diverse range of personalities should be represented in a group going on an Antarctic expedition. By having all kinds of people represented, the team gets stronger than the sum of all individual strengths. A parallel can be drawn to open source software: by having diverse teams, the best software is made. It’s not only the programmer who should be satisfied - it’s an often diverse mass of users standing there. Recognising this could be a key to success in the future [6].

5 Conclusions

Both debian-level and the Debian Developers have about only half the expected amount of women if compared to these studies and other comparable projects. There may be many reasons for this, but it is interesting to see that so many more women are participating in the debian-women project. This shows that there are in deed women interested in Debian, and it is possible to find them. In fact, they are already here.

About the author

Magni Onsjøien is writing her M.Sc. thesis on female developers in Free Software at Norwegian University of Science and Technology. She has been using different Linux distributions for the past 10 years, and is now considering entering the New Maintainer process.

References

- [1] Paul A. David, Andrew Waterman, and Seema Arora. The Free/Libre/Open Source Software Survey for 2003. Technical report, Stan-

- ford Project on the Economics of Open Source Software, Stanford University, 2003. <http://www.stanford.edu/group/floss-us/>.
- [2] Bert J. Dempsey, Debra Weiss, Paul Jones, and Jane Greenberg. Who is an open source software developer? Profiling a community of Linux developers. *Communications of the ACM*, 45(2):67 – 72, February 2002.
 - [3] Nancy J. Hafkin. Some thoughts on gender and telecommunications/ICT statistics and indicators. June 2003.
 - [4] Gisle Hannemyr. Technology and pleasure: considering hacking constructive. *First Monday*, 4(2), February 1999.
 - [5] Indicators for the Information Society in the Baltic Region, Action line 6. Available via <http://www.riso.ee/en/nordic/nv2al6.htm>, July 2003. [Cited 2004-11-27].
 - [6] Michelle Levesque. Fundamental issues with open source software development. *First Monday*, 9(4), April 2004.
 - [7] Jane Margolis and Allan Fisher. *Unlocking the Clubhouse. Women in Computing*. The MIT Press, Cambridge, Massachusetts, 2002.
 - [8] Eric S. Raymond. The cathedral and the bazaar. *First Monday*, 3(3), March 1998.
 - [9] Gregorio Robles, Hendrik Scheider, Ingo Tretkowski, and Niels Weber. Who Is Doing It? A research on Libre Software developers. Technical report, Fachgebiet für Informatik und Gesellschaft TU-Berlin, August 2001. <http://widi.berlios.de/paper/study.html>.
 - [10] Walt Scacchi. Socio-technical interaction networks in free/open source software development processes. In Silvia T. Acuña and Natalia Juristo, editors, *Software Process Modeling*. Springer Science+Business Media Inc., 2005.
 - [11] Hiroyuki Shimizu, Jun Iio, and Kazuo Hiyane. Realities of Free/Libre/Open Source Software developers in Japan and Asia. *First Monday*, 9(11), November 2004.
 - [12] Anthony J. W. Taylor. *Antarctic psychology*, volume 244 of *DSIR bulletin / New Zealand Department of Scientific and Industrial Research*. SICP, Wellington, 1987.
 - [13] Free/Libre/Open Source Software Asian Developers Online Survey. Technical report, Mitsubishi Research Institute, Inc. (MRI), Japan, March 2004. <http://oss.mri.co.jp/floss-asia/>.
 - [14] Free/Libre and Open Source Software: Survey and Study. FLOSS Final Report. Technical report, International Institute of Infonomics, University of Maastricht, Netherlands/ Berlecon Research GmbH, Berlin, Germany, June 2002. <http://www.infonomics.nl/FLOSS/report/>.