

Results from survey among active contributors in the openSUSE community



Survey start: 22.10.2007
Survey end: 19.11.2007
Number of participants (N): 281
Topic: Firm-sponsored open source communities
By: Jan Fredrik Stoveland (janfst@student.matnat.uio.no),
University of Oslo
Date: 29.11.2007

Background

This report presents the results of a questionnaire survey that was conducted among the openSUSE community in the fall of 2007. The survey is part of a larger study of Novell and the openSUSE community. The full study will be published June 2008. The purpose of this report is to present the results for various stakeholders (Novell, the community, etc). This document is not a proper academic submission, and so it should not be cited in any academic publication.

This report includes the distribution of frequencies for all questions in the survey. The answers for the two questions with text fields are not included (3.9 and 5.2). All results are commented, and in some cases I will present statistical correlations that are found in the data. Some results and interpretations will be discussed, but note that any interpretations in this report are only temporary. Since the data from this survey will be analyzed with other data at a later stage, the comments in this report are not final. Nor have I chosen to conclude on any of the topics.

Note on the validity of the results in this survey

The survey targeted active contributors in the openSUSE community, and 281 people participated in the survey. Due to statistical methods, this number of participants in a survey would normally be enough to render the findings in the survey applicable to the entire community. However, this requires that the selection of participants for the survey is also done according to statistical methods (i.e. randomly selected among the entire community). Unfortunately the selection of participants for this survey has not been done according to such methods, for two reasons: (1) I was not given the liberty to use individual contact information to target randomly selected participants directly by Novell, due to reasons of confidentiality and use of contact details for purposes not authorized by

the individuals beforehand. (2) In addition, the boundaries of the community are unclear (see the discussion of this topic in the thesis), meaning that it is difficult to determine exactly who is within the target group beforehand. For this reason it would be hard to determine exactly which individuals to include in the selection of participants.

Rather, this survey has been conducted with an open invitation, and is based on voluntary participation. The invitation was sent directly to the following mailing lists: *project*, *factory*, *buildservice*, *translation*, *wiki* and the *main* list (opensuse@opensuse.org). A link to the survey was also posted on *news.opensuse.org*. As it is very likely that active contributors in the community follow one or more of these communication channels, it is also likely that the invitation reached a majority of the target group.

Although the results of this survey cannot be statistically generalized, they are definitely true for the group of people who answered the questionnaire (meaning the 281 participants whom voluntarily participated). Furthermore, these participants are representatives of the target group. *We can therefore say that it is likely that the findings in the survey apply to the larger group of contributors in the community, but the results may not be statistically generalized to the entire community of active contributors.*

Statistical measures

Besides the distributions of frequencies shown below, I will also present some correlations in the data where this has been found significant through statistical analysis. Correlation between two variables means that they are related somehow (for example, the more X increases, the higher value of Y). I will mainly use the statistical measure Gamma to determine the strength of the correlation, since this measure is suitable for variables on the ordinal level. It ranges between -1 and 1, where 0 means no correlation at all, and -1 or 1 is a perfect correlation (linear relationship). Negative and positive numbers indicate the direction of the correlation.

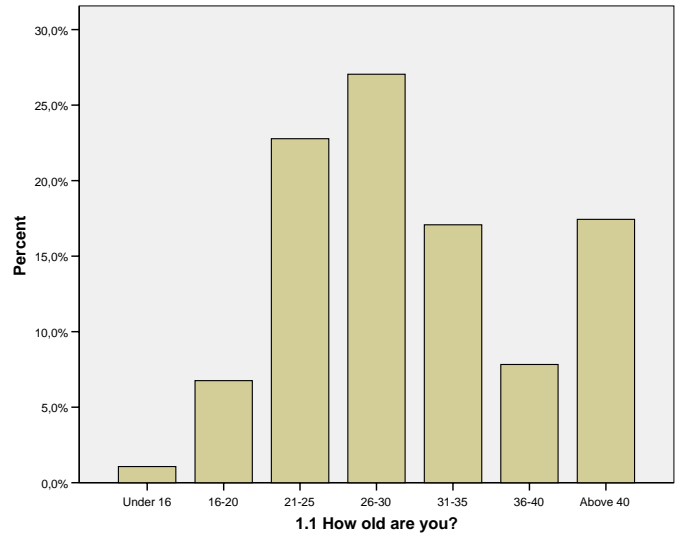
There are several common pitfalls in using correlation. Correlation is symmetrical, not providing evidence of which way causation flows. If other variables also cause the dependent variable, then any covariance they share with the given independent variable in a correlation may be falsely attributed to that independent. Also, to the extent that there is a nonlinear relationship between the two variables being correlated, correlation will understate the relationship.

All correlations I present here are statistically significant at the 0,05 level, meaning that the risk of reaching the results by coincidence is less than 5% (most findings presented here are in fact significant at the 0,01 level).

Results

1. Personal background

The purpose of the questions in this section is to establish some knowledge about the demography of the community. These variables will also be used to see if there are any systematic differences in answers to other topics in the survey.



1.1. How old are you?

Under 16	16 - 20	21-25	26-30	31-35	36-40	Above 40
1 %	7 %	23 %	27 %	17 %	8 %	17 %

N = 281

1.2. Male or female?

Male	Female
98 %	2 %

N = 281

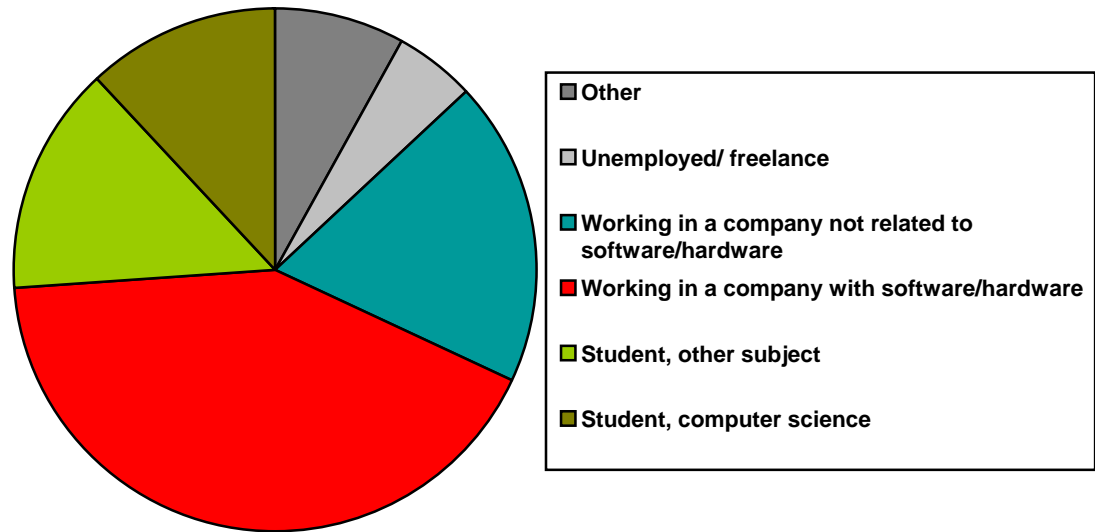
1.3. What is your occupation?

Student, computer science	Student, other subject	Working in a company with software/hardware	Working in a company not related to software/hardware	Unemployed/freelance	Other
12 %	14 %	42 %	19 %	5 %	8 %

N = 280

Comments

It is clear that the main body of the contributor-community consists of men between 20 and 35 years of age, although a quarter of the respondents are older than this. The main occupational group consists of people who work in a software or hardware company, which accounts for close to half of the respondents (42%). In comparison, all the students put together account for only 1/4 of the contributors. The high amount of IT professionals within the community might be related to openSUSE's and Novell's business-profile. As some of the results in the survey show, many contributors participate as part of their daily work at a software/hardware company. But, without any comparative data from other communities it is hard to say whether this is a stronger trend within openSUSE.



Not many females participated in the survey. There is little reason to believe that men have a greater preference for answering surveys in general, and so we could assume that it is highly likely that the group of females within the community of active contributors is also fairly small. The low amount of participation by females means that it is impossible to do any comparative analysis between genders with the results of the survey. Otherwise, gender could have been used as an independent variable to find out if there are any systematic differences between men and women on the other issues in the survey.

2. Contributions and involvement

2.1. How many hours in average do you put in contributions each week?

('contributions' includes time used to prepare input to discussions, testing of software, or other time needed to prepare a statement, user support, code contributions or reports)

Less than an hour a week	1-3 hours a week	4- 10 hours a week	More than 10 hours a week
25 %	33 %	23 %	19 %

N = 273

2.2. In how many areas do you participate?

('involvement' means actively pursuing the development of a project by systematically contributing with testing, bug reports, code OR input to discussions. A 'project' also includes non-software work, for example wiki, translations or documentation. If you contribute for the most part with user support, you can answer alternative 2 and provide the amount of time used in the previous question)

I am only a user, I do not submit contributions	18 %
I only contribute arbitrarily where I see fit - I don't have a specific interest in any project/package	41 %
I am involved with one specific project/package	12 %
I am currently involved with one specific project/package, but have been involved with others earlier	7 %
I am currently involved with 2-3 projects/packages	8
I am currently involved with more than 3 projects/packages	14 %

N = 277

Comments

Close to half of the respondents spend more than 3 hours a week on contributions to openSUSE. 19% contribute with more than 10 hours a week. It is common for open source communities to have a core group that contributes with the largest amount of work. Also in this case, it would seem that the Pareto principle of distribution is a fitting pattern (20% of the community contribute with 80% of the work, and vice versa).

These results may give us some idea about how much effort is contributed from outside the company. Here is a thought experiment: If we try to calculate how much work the respondents contribute with - using averages for each category (½, 2, 7 and 12 hours) - we end up with 1320 hours a week. This is equivalent with the time spent by 33 full-time employees (working 40 hours a week). This is not to say that the community is doing the errand of what 33 extra Novell employees would have been doing, but certainly emphasizes the amount of work that is voluntarily contributed by the community – and this is only among the respondents in the survey¹.

The second question in this section shows that contributors are quite spread in the amount of specific projects they are involved in. Members representing the largest category are not tied to any specific project(s). However, statistical analysis shows that the two questions in this section are positively correlated (Gamma = 0,57). This means that the amount of hours contributed increases with the amount of specific projects/packages a member is involved in, and vice versa. It also means that people who contribute less are not tied to any specific projects. This poses an interesting question: Do they contribute less *because* they are not committed to any specific project, or is the lack of specific involvement due to the low amount of engagement? Unfortunately this data can not explain the causal direction between these two aspects. Between these two, I would expect that the latter explanation is more plausible. But, if the former were to be true, it would mean that a larger amount of the community might contribute more if they were given ties to specific projects. Of course, there is most certainly a large amount of completely different reasons that explain the amount of time a member is willing to contribute, so this analysis should not be taken too far.

¹ Note that some of the respondents in this survey are in fact employed by Novell (this itself is an interesting finding, as they identify themselves as contributors in the community). There are four answers in the open question on “other reasons for participating in the community” that confirm this. How many employees that have replied in total is uncertain, but looking at the remainder of answers I do not believe this is a large group. But it should be taken into account when calculating hours voluntarily contributed.

3. For what reasons do you participate in the openSUSE community?

3.1. Because I believe that openSUSE is the best Linux distribution

not at all	To a little degree	to some degree	To a large degree
3 %	4 %	35 %	58 %

N = 275

3.2. Because I respect the other community members

not at all	To a little degree	to some degree	To a large degree
5 %	10 %	42 %	43 %

N = 265

3.3. Because I respect the employees working with openSUSE at Novell

not at all	To a little degree	to some degree	To a large degree
10 %	11 %	36 %	43 %

N = 266

3.4. Because I am friends with several people in the community

not at all	to a little degree	to some degree	To a large degree
47 %	22 %	24 %	6 %

N = 268

3.5. Because it is easy to get my footprint in openSUSE compared to other Linux distributions

not at all	to a little degree	to some degree	To a large degree
34 %	25 %	30 %	11 %

N = 259

3.6. Because my work with openSUSE is acknowledged by the community and employees

not at all	to a little degree	to some degree	To a large degree
32 %	21 %	29 %	18 %

N = 261

3.7. Because I learn a lot and get support from the community

not at all	to a little degree	to some degree	To a large degree
6 %	11 %	35 %	48 %

N = 268

3.8. Because my work for openSUSE may be used by many people, including enterprise level

not at all	to a little degree	to some degree	To a large degree
14 %	17 %	32 %	38 %

N = 263

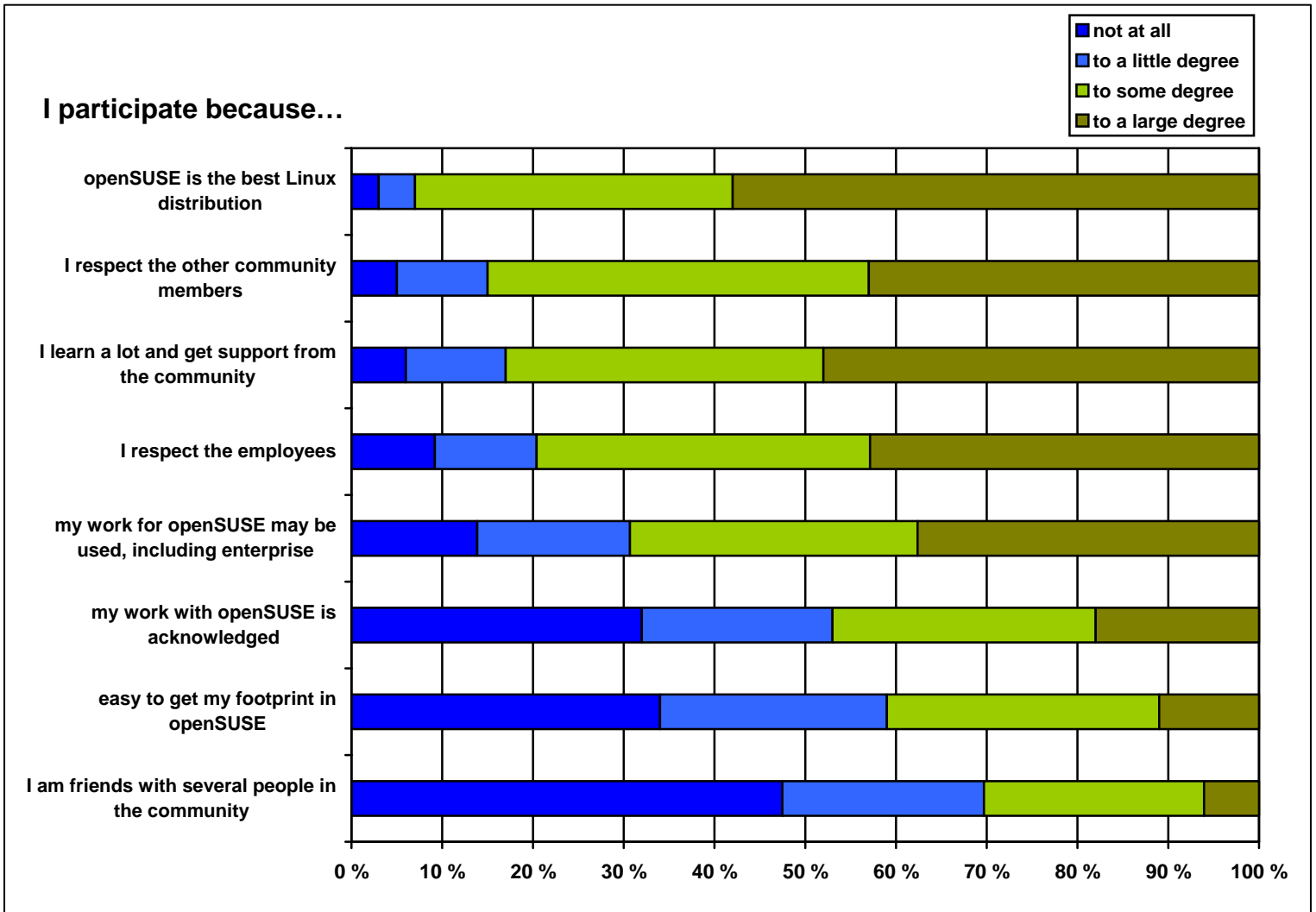
3.9. Are there any other reasons for why you participate in the openSUSE community?

[open text field] See *the raw data in the other document for replies to this question.*

Note: Some of the questions in this section are double-sided, and this affects the interpretation of the data. They ask if a specific phenomenon motivates to participate in the community. But in some cases, one may also use the reply of the question to determine whether the phenomenon is present/applicable for the member in the community. For example, statement 3.4 (I participate because I have friends in the community) can have the following meanings. If the informant answers negatively to the question, it can be interpreted in two ways: (a) The informant does not have friends in the community and is therefore not motivated by them, or (b) the informant does have friends in the community, but it is not a motivating factor for him. If, on the other hand, an informant answers positively, it can only mean one thing: (c) that the informant has friends (confirming that the phenomenon is present) AND is motivated by them to participate. This duality may lead to some uncertainty in the interpretation of the data, so the results should be considered wisely².

There are many factors that may motivate individuals to participate and take part of a community, ranging from egoistic to altruistic, and intrinsic to extrinsic reasons. Many of these are covered in the literature on open source. With the questions in this section I have focused on some aspects that are not so much described in the literature, and that I expected might be relevant for the openSUSE community after some interviews with community members. The thesis will include a larger discussion on this topic. Below is a ranked presentation of the results to the questions in this section.

² In order to avoid this dilemma one could have asked a double set of questions – one question to find out if the phenomenon is present, and a second to determine if it is motivating. I chose not to do so as this would increase the amount of total questions, and could affect the amount of participants willing to do the survey.



From these frequencies it is clear that the quality of the product itself is the motivating factor that unites the community. Personal learning also matters a lot. Although respect for other community members and employees seems to be important, I would be careful with putting too much weight on these two aspects. The reason is that these questions might be more sensitive and difficult to answer negatively to, as they concern respondents' view on other people. Respondents that have an opinion according to answer 'b' above, might be concerned that it will be interpreted as an 'a' (meaning that they do not respect other community members or employees).

It is interesting to observe that 'social ties' is the least motivating factor in the list. Since humans are genuinely social beings, it is difficult to not be motivated by people that we like. The replies to this question should therefore be interpreted as 'a', meaning that members are not motivated by friends due to the fact that there doesn't exist many strong social ties in the community. This is natural as open source communities are distributed networks, largely mediated through the internet. Still, 30% of the respondents *do* claim to

have friends in the community, and this corresponds well to the percentage of people who contribute with a large amount of work in the community (question 2.1). With a Gamma equal to 0,4, statistical analysis confirms that there is a correlation between amount of work contributed and motivation by friends. This seems reasonable, as more work means more interaction with other people, increasing the likelihood of establishing friendship.

I believe that the question about ease of establishing a ‘footprint’ in openSUSE should be interpreted in a similar way as the former comment. It is largely documented that ownership to a product or process motivates people. When this question receives such a low score, it should be interpreted as a lack of experience with influence on the product within the community. As with the question of social ties, one could expect that people who contribute more hours are prone to experience the ‘footprint’ as more of a motivating factor as they would more likely be able to influence the projects they work on. However, there is little correlation between these factors in the data (Gamma = 0,2). This means that there are many who contribute a lot who still do not experience ease-of-influence as a motivating factor. This is a finding that is backed up by individual interviews I have conducted with members of the community.

4. Do you agree with the following statements?

4.1. I believe Novell is an open source company

not at all	to a little degree	To some degree	to a large degree
12 %	19 %	50 %	19 %

N = 273

4.2. openSUSE and Novell are one and the same to me

not at all	to a little degree	To some degree	to a large degree
37 %	28 %	26 %	10 %

N = 275

4.3. The patent deal between Novell and Microsoft worries me

not at all	to a little degree	To some degree	to a large degree
28 %	30 %	21 %	21 %

N = 274

4.4. Novell's relation to other companies impacts my motivation for contributing in openSUSE

not at all	to a little degree	To some degree	to a large degree
39 %	28 %	25 %	9 %

N = 266

4.5. It would be fine with me if openSUSE was designed in red Novell colors

not at all	to a little degree	to some degree	to a large degree
55 %	21 %	15 %	25 %

N = 268

Comments

These questions address some topics concerning the relation between the community and Novell, the sponsor of the community. The majority of the respondents clearly separate Novell and openSUSE from one another. Not surprisingly, the respondents who view them as one and the same are largely the same people that believe Novell is an open source company (Gamma correlation = 0,55).

Novell's deal with Microsoft has been thoroughly debated by now. It was nevertheless a question worth including in the survey, as it is interesting to see how big of an impact the actions of the community's sponsor have on the motivation within its community. It could be reasonable to either expect that community members who had serious concerns about this deal no longer were present in the community, as the community is completely based on voluntary participation without other incentives to participate besides individual motivation. If this were the case we could expect the (remaining) respondents to have little concerns about this deal. Another hypothesis could be that people concerned about this deal are not so motivated to contribute, and should thus have a small amount of contribution hours. Neither of these assumptions are confirmed by the data. The distribution of frequency shows that level of concern is equally distributed. There is a majority of respondents without concerns about this patent deal, but also a significant amount of people who do worry about the deal at all. But the concern is clearly not too serious, since they still contribute. There is also little correlation between level of concern and amount of hours contributed, which would have fit with the second hypothesis. As the crosstabulation below shows, 36 % of the people who are concerned by the deal contribute a large amount of hours. Even if 64% of those with concerns do not contribute that much, the difference is not significant enough to create a correlation between the two.

How many hours in average do you put in contributions each week?	The patent deal between Novell and Microsoft worries me		Total
	No	Yes	
0-3 hours	52 %	64 %	57 %
4-10+ hours	48 %	36 %	43 %
	100 %	100 %	100 %

The only question in the survey that correlates with this Microsoft-variable is the next question about Novell's general relation to other companies, and how this affects motivation (Gamma = 0,55). In sum, from these results it does not seem like Novell's relations with other companies have a large impact on individual's motivation to participate in the community, among the existing community members.

The question about the colors of the distribution was used in an effort to compare the sentiment to the Novell brand in comparison to openSUSE. But the results to this question do not give much meaning unless it correlates with other variables. The frequencies show that most contributors would avoid red colors. A quarter would not mind at all. Some of the respondents in this group are the same that believe Novell is an open source company (Gamma correlation 0,34). From some conversations with community members, I get the impression there are some individuals who identify

themselves strongly with Novell. Nevertheless, the data in this survey is not strong enough to identify any “Novell patriots” within the community, but this is a topic that might be interesting for further analysis.

5. Decisionmaking

5.1. To what degree do you experience an ability to influence and take part in decisions involving the openSUSE project?

If you would like to answer "I do not know", please leave the question unanswered.

not at all	to a little degree	to some degree	to a large degree
19 %	31 %	37 %	13 %

N = 222

5.2. Are there any aspects you would like to see improved about the organization of the community?

[open text field] *See the raw data in the other document for replies to this question.*

Comments

How democratic is the openSUSE community? Although the community has no formal representation or decision-making power, openSUSE management in Novell do try to include the interests of community members in decisions³. The results do not indicate that community members experience a large amount of influence, but they are not without a voice either. 37 % experience some degree of influence and 13% experience it to a large degree. But it is clear that many community member are uncertain about how they may participate in decision-making, as many respondents did not reply to this question (missing = 59).

Level of influence is correlated with some of the motivation variables. Level of influence is correlated to 3.5, 3.6 and 3.8 (footprint, acknowledgement and enterprise user-base), with Gamma above 0,45 on all cases. More importantly, influence also correlates with contribution hours. It is probably natural that contributors that put in more work also have more influence. Another interesting finding is that level of influence also correlates moderately with social ties (0,37), and that this correlates again with amount of hours contributed. Put together, it seems that there exists a group of core contributors, having influence on decisions and on the product, and receiving acknowledgement on their work. They have some social ties since they communicate a lot together, and they are motivated by the fact that there is a large user base because their work that is influential to the product and is “out there”.

Although these data seem to give some support to the notion of a group of core-contributors surrounding the ‘level-of-influence’-variable, the satellite variables are not tied so tightly together themselves. In other words, the data does not give strong support to an assumption that there exists an exclusive group of ‘elites’. But it shows that there is

³ This is based on interviews and observations within the company.

a group that is more 'in-the-loop' than most others, although this is probably not a homogenous group with strict boundaries, and they experience the community itself in different ways.

A final comment I would like to add, is that I find it surprising that none of the background variables correlate with any of the data in the survey. It shows that age and occupation has no significance towards how community members experience the aspects of the community that have been discussed here.