

Henri Nieminen

**THE PROCESS BY WHICH USERS BECOME ACTIVE
KNOWLEDGE CONTRIBUTORS IN PROGRAMMING-
FOCUSED ONLINE COMMUNITIES**



UNIVERSITY OF JYVÄSKYLÄ
FACULTY OF INFORMATION TECHNOLOGY
2018

ABSTRACT

Nieminen, Henri

The process by which users become active knowledge contributors in programming-focused online communities

Jyväskylä: University of Jyväskylä, 2018, 74 pp.

Information Systems, Masters thesis

Supervisor(s): Soliman, Wael

The recent popularity of social knowledge sharing platforms has encouraged people to seek and to share knowledge on them. Motivating users to contribute knowledge has been a major challenge for online communities, in which only a relatively few online knowledge sharing communities have been able to succeed. The goal of the thesis was to find out why users share knowledge, and how users become active knowledge contributors in programming focused online communities. With this information the communities could better identify the users that are likely to become knowledge contributors, and find ways to encourage them to begin contributing knowledge. A qualitative study was carried out to find answers the research problem. Semi-structured interviews were used to collect stories from 18 knowledge contributors on Internet Relay Chat (IRC) network Freenodes programming focused communities. The interviewees were selected based on their activity in knowledge sharing during a selected period of time. Content analysis and narrative analysis were used to analyze the interview data in order to build a process theory. The process theory explains how users become knowledge contributors after initially joining a community. The process theory indicates that users initially join a community to learn more about the topic of the community, to get help with a problem, and to keep up to date with the topic of the community. After joining, the users begin familiarizing themselves with the community by observing or by socializing, and when familiar with the community they start seeking knowledge. Only after the users have gained enough knowledge and confidence in their knowledge do they begin sharing knowledge. The majority of the users were motivated to participate in the online knowledge sharing communities by the possibility to learn. Sharing knowledge was also found to be motivated by learning opportunities. Some users share knowledge in order to validate their knowledge, if they make mistakes someone in the community might correct them.

Keywords: knowledge sharing, information sharing, online community, virtual community, process theory, process model, internet relay chat, freenode.

FIGURES

FIGURE 1: SECI-model (adapted from Nonaka & Konno, 1998).....	11
FIGURE 2: Degrees of community participation (adapted from Wenger et al., 2002, 57).....	20
FIGURE 3: Channel counts per category.....	29
FIGURE 4: Interviewee age groups.....	31
FIGURE 5: Native countries of interviewees.....	32
FIGURE 6: Work experience.....	33
FIGURE 7: Educational background of the interviewees.....	34
FIGURE 8: Levels of structure in narrative (adapted from Pentland, 1999).....	39
FIGURE 9: Nerd Sniping (Munroe, 2007).....	48
FIGURE 10: Pathway #1.....	55
FIGURE 11: Pathway #2.....	56
FIGURE 12: Pathway #3.....	57
FIGURE 13: Process model: how users become knowledge contributors.....	59

TABLES

TABLE 1: Summary of motivations for participation per interviewee.....	43
--	----

TABLE OF CONTENTS

ABSTRACT.....	2
FIGURES.....	3
TABLE OF CONTENTS.....	4
1 INTRODUCTION.....	6
2 LITERATURE REVIEW.....	8
2.1 Concepts of knowledge sharing.....	9
2.1.1 Knowledge creation.....	10
2.1.2 Knowledge sharing.....	11
2.1.3 Communities of practice.....	12
2.2 Online knowledge sharing communities.....	13
2.2.1 Online communities.....	14
2.2.2 Knowledge sharing platforms.....	15
2.3 User participation.....	18
2.3.1 Types of participation.....	18
2.3.2 Why users participate.....	20
3 METHODOLOGY.....	23
3.1 Research goals.....	23
3.2 Research approach.....	25
3.3 The case of the study.....	26
3.4 Data collection.....	27
3.4.1 Selection of interviewees.....	27
3.4.2 Interviewees.....	30
3.4.3 Interviews.....	34
3.5 Research ethics.....	36
3.6 Analysis.....	38
4 RESULTS.....	41
4.1 Why users participate in programming focused online communities.....	42
4.1.1 Reasons for participation in online communities.....	42
4.1.2 Advantages of IRC to knowledge sharing.....	49
4.1.3 Reasons for not sharing knowledge.....	51
4.2 How users become active knowledge contributors.....	52
4.2.1 Pathways to becoming active knowledge contributor.....	53
4.2.2 Process model: how users become active knowledge contributors in online communities.....	58

5 DISCUSSION.....	61
5.1 Discussion of the empirical results.....	61
5.2 Strengths and weaknesses of the research.....	64
5.3 Implications for theory and practice.....	65
5.4 Topics for further study.....	65
6 CONCLUSION.....	67
REFERENCES.....	68
APPENDIX 1 INVITATION LETTER.....	72
APPENDIX 2 BACKGROUND QUESTIONS FOR INTERVIEWEES.....	73
APPENDIX 3 SOME OF THE RECURRING QUESTIONS IN INTERVIEWS....	74

1 INTRODUCTION

Knowledge sharing in online communities has been thriving after the rise of social media. Social networking technologies have changed the way people share knowledge, which has led companies to encourage their employees to share knowledge on the company initiated knowledge sharing platforms. (Charband & Navimipour, 2016) Motivating users to share knowledge has been the biggest challenge online communities face (Chiu, Hsu, & Wang, 2006), and only relatively few online knowledge sharing communities have been able to succeed in motivating their users to become active knowledge contributors (Lai & Chen, 2014).

The goal of the thesis is to find out how users become active knowledge contributors in online communities. Plenty of research exists on why users of online communities share knowledge, but a much less studied topic has been how they become active knowledge contributors. It has been widely accepted that without knowledge contributors, these communities might cease to exist (e.g. Joyce & Kraut, 2006; Ridings & Wasko, 2010; Nov, Naaman, & Ye, 2010). Finding out how users become active knowledge contributors in online communities makes it possible for the communities to identify the users that are likely to become knowledge contributors, and then find ways to encourage them to start sharing.

This study answers to the following two research questions:

- **RQ1:** *Why do users share knowledge on Freenode's programming focused communities?*
- **RQ2:** *How do users of Freenode's programming focused communities become active knowledge contributors?*

The empirical study attempts to answer the research questions by building a process theory that will cover the time from initial use until a user becomes a knowledge contributor. The case of the study is Internet Relay Chat (IRC) network Freenode and its programming focused communities. Semi-structured interviews are used to collect stories from knowledge contributors, who were selected based on their activity in knowledge sharing during a selected period of time. Qualitative methods, namely content analysis and narrative analysis, are

used to analyze the data of 18 interviews, with the aim of building a process theory covering events from the initial use of the community until a user becomes a knowledge contributor. The levels of structure in narrative (Pentland, 1999) are used in the analysis to aid in building the process theory so that the theory is able to show the different levels of the narrative, including stories, fabula, and the generating mechanisms for each stage in the process.

The resulting process theory indicates that users initially join a community to learn more about the topic of the community, to get help with a problem, and to keep up to date with the topic of the community. After joining the users begin familiarizing themselves with the community, which is done by observing the discussions or by socializing with community participants. Once the users are familiar with the community, they start seeking knowledge by presenting problems, asking questions, and following the discussions. Once the users gain enough knowledge, their confidence in their knowledge raises, and when the users have gained enough knowledge they begin sharing their knowledge. First they begin by helping others that have easy problems, and then they gradually move to helping with harder problems when they gain further knowledge and confidence.

The results indicate that the majority of the users are motivated to participate in the online knowledge sharing communities by the possibility to learn. Learning as a motivation applies to the users in most of the stages in the process theory. Even sharing knowledge is found to be motivated by learning opportunities, which is a finding that has not been widely discussed in previous literature. The results of this study also indicate a new motivation for sharing knowledge, which is learning by validating ones knowledge. Some users are motivated to help others without the fear of providing false answers, they feel that even if they provide false information, sharing this knowledge can be beneficial to them when other users correct this false information. When this happens the users learn from their mistakes.

2 LITERATURE REVIEW

In this chapter you can find the literature review, which will showcase the literature relevant to this study. The goal of the literature review was to explore the key concepts of knowledge sharing, online communities, and user participation. The literature review is a narrative review of the literature, as the aim is to summarize previous research on the topic (Templier & Paré, 2015). Templier and Paré's (2015) general procedure for conducting literature reviews was followed by first figuring out a research problem. This was done with the help of the supervisor of the thesis and by reading journal articles around the area of knowledge sharing. After the research problem was identified, the selection of the articles for the literature review began. The articles selected for the literature review were found by using multiple different methods. First of all Google Scholar was used to get started, and later on also Scopus was used for searching articles. Various keywords were used in combinations when trying to find relevant articles. Some of the keywords used were "knowledge sharing", "online community", "user participation", and different combinations of these. When some relevant articles were found, backwards and forwards reference searching was also used; which articles were referenced in each article, and which articles had referenced said article. Finding forwards references was mainly done by using Google Scholar. Some of the most popular information systems journals' fairly recent issues were also manually checked for relevant articles, as they might not show up in searches due to having only a few citations. After gathering the articles, it was verified that the articles were relevant to this study. Only relevant articles were chosen with the aim being on choosing articles from respected journals and conferences. Then data was acquired and analyzed from the articles, and the data acquired was synthesized by comparing and summarizing it.

2.1 Concepts of knowledge sharing

Knowledge is the human ability to use information, and it is constructed in the moment when thinking through a problem (McDermott, 1999). It is often defined as a "justified true belief" (Nonaka, Toyama, & Konno, 2000). Knowledge is an understanding that can be gained through experience. (Alavi & Leidner, 2001). Wilson (2010) sees knowledge as a set of mental processes that involve understanding and learning. Two types of knowledge exist: explicit knowledge, which can be shared by communication as words and numbers, and tacit knowledge that is considered to be hard to formalize and therefore hard to share. Tacit knowledge is highly personal as it is related to persons actions and experience. (Nonaka & Konno, 1998) Wenger, McDermott, and Snyder (2002, 9) note the following about sharing tacit knowledge:

Sharing tacit knowledge requires interaction and informal learning processes such as storytelling, conversation, coaching, and apprenticeship (...)

McDermott (1999) lists six reasons why knowledge is different to information: knowledge is a human act, the residue of thinking, created in the present moment, belongs to communities, circulates through communities, and new knowledge is created at the boundaries of old. Information on the other hand is something that can be stored or saved while knowledge can't be (McDermott, 1999). McDermott (1999) also explains the difference of knowledge and information in common man's terms:

Knowledge always involves a person who knows. My bookcase contains a lot of information on organizational change, but we would not say that it is knowledgeable about the subject. (McDermott, 1999, 105)

Knowledge and information have not been distinguished in many knowledge related studies (Wang & Noe, 2010). For example Kogut and Zander (1992) define knowledge as information and know-how, which can be replicated in in the context of a firm, but cannot be easily transferred to other firms. This view contradicts with the idea that information is not the same thing as knowledge (e.g. Nonaka & Konno, 1998; McDermott, 1999; Wilson, 2010). According to Wang and Noe (2010) the terms knowledge and information might have been used interchangeably because distinguishing them has not served any practical purposes. The same might also apply for information sharing and knowledge sharing (Savolainen, 2017).

2.1.1 Knowledge creation

According to Wilson (2010) people construct knowledge from information received, a process that is not fully understood. In principle knowledge is personal (Wilson, 2010) and each individual can construct different kind of knowledge out of the same information. In order for knowledge to be created information has to be exchanged between individuals and groups (Cabrera & Cabrera, 2002).

Nonaka and Konno (1998) have introduced a knowledge creation concept known as *ba*, which they consider to be shared space in which knowledge creation happens. The space can be either physical, virtual or a combination of the two. According to the concept, knowledge is intangible and can be acquired through experience only, and when knowledge is shared outside of this shared space, it becomes information. (Nonaka & Konno, 1998.) This is because knowledge is dynamic, it is created in social interactions within some context by people through action and interaction (Nonaka et al., 2000). When knowledge is shared outside of this space, it is missing the context.

Nonaka and Konno (1998) have demonstrated knowledge creation in the form of a model known as SECI. The SECI model describes four different phases that spiral endlessly, the outcome being organizational knowledge. Firstly, individuals create knowledge from peer to peer, then it is externalized within a group, which leads to combination of knowledge of the groups inside an organization, and then lastly an individual acquires this knowledge from the organization. After this the individual shares the knowledge to another individual and the spiral loops again, as can be seen in Figure 1. (Nonaka & Konno, 1998; Nonaka et al., 2000)

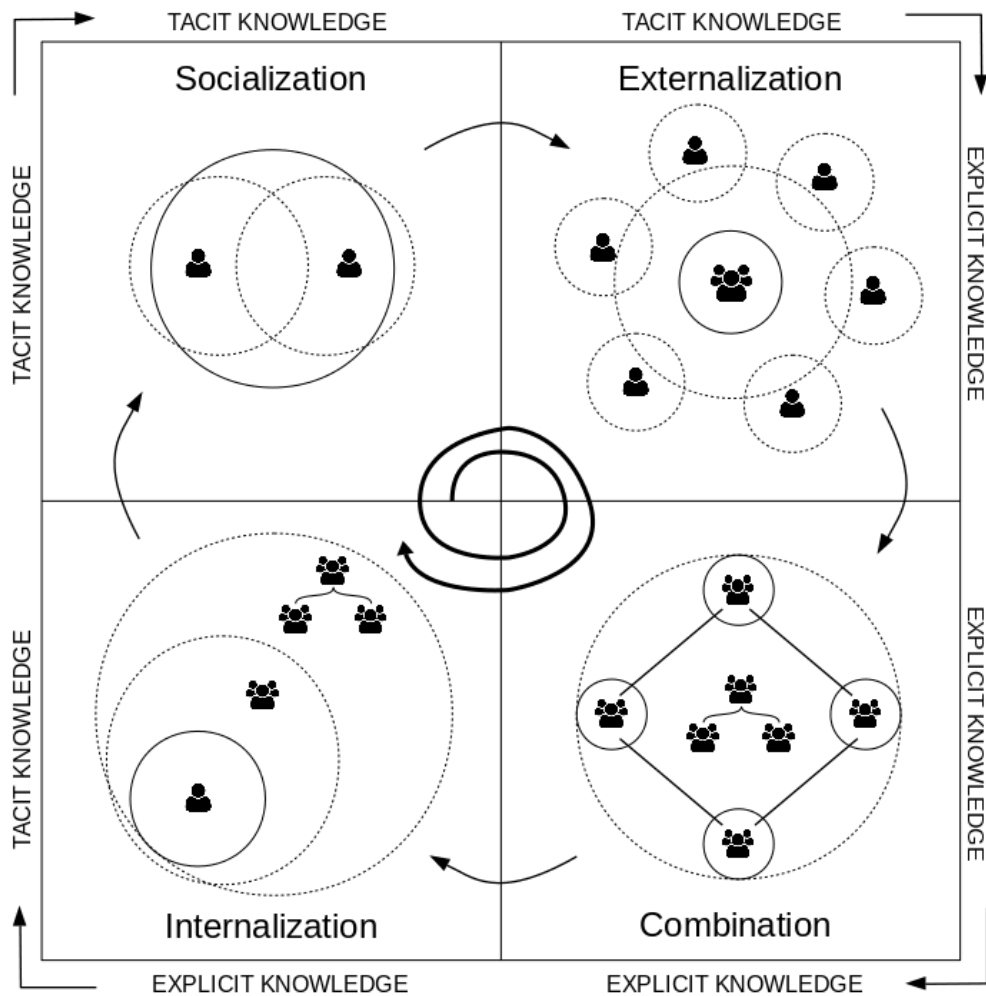


FIGURE 1: SECI-model (adapted from Nonaka & Konno, 1998).

2.1.2 Knowledge sharing

Wang and Noe (2010) define knowledge sharing as an act in which knowledge is provided by individuals to others. According to Hendriks (1999) knowledge sharing is related to communication and distribution of information, but it is not the same as either of those. For knowledge to be able to be shared, first it has to be reconstructed in the minds of individuals. The act of sharing knowledge doesn't necessarily have to be a conscious or an intentional act, as people can learn by observing and the "sharer" might not be aware of being watched. To make it easier for others to be able to reconstruct the knowledge, it should be presented in an understandable way. (Hendriks, 1999)

According to Ardichvili, Page and Wentling (2003) people often feel that knowledge doesn't belong to them, it rather is a public good and sharing it is a moral obligation and a community interest. Much of knowledge sharing depends on the individuals behavior, knowledge sharing is not something that can be forced to happen (Bock, Zmud, Kim, & Lee, 2005). The individuals need motivation to share knowledge, otherwise it is not likely to happen (Chiu et al., 2006). For knowledge sharing to happen in an organizational level, employers must take effort in creating environments that motivate individuals to share knowledge (Bock et al., 2005). Bock et al. (2005) suggest that feedback should be given to those individuals that take part in knowledge sharing to make them notice that their actions are appreciated.

Knowledge collaboration happens in online communities when individuals get together and collectively take part in knowledge creation and sharing. Faraj, Jarvenpaa and Majchrzak (2011) define knowledge collaboration as: "the sharing, transfer, accumulation, transformation, and co-creation of knowledge." (Faraj et al., 2011, 1224).

According to Savolainen (2017) researchers have no common consensus on what is actually being shared, is it information or knowledge. Wilson (2010) argues that knowledge can't be shared by communication, and what is actually being shared is information about what one knows, knowledge is then constructed by each individual out of this information. Savolainen (2017) found that in literature the activities in information sharing and knowledge sharing are similar, and therefore the terms could be used interchangeably.

2.1.3 Communities of practice

According to Ardichvili et al. (2003) Lave and Wenger came up with the phrase "Community of Practice" in 1991. Wenger et al. (2002, 4) define it in the following way:

Communities of practice are groups of people who share a concern, a set of problems, or a passion about a topic, and who deepen their knowledge and expertise in this area by interacting on an ongoing basis.

Wenger et al. (2002, 3) mention in the story of how tech clubs emerged at Chrysler after 1988, when they tried to reduce the time it took to design a new car, that spending time together saved time and increased the confidence of engineers. Tech clubs also made it possible for the engineers to get help and stay up to date with new technologies. Chrysler's tech clubs were communities of practice that were initially formed informally by the staff, to solve problems that arose after a change in the organizational structure, and later supported by the management. (Wenger et al., 2002, 2-4)

Even though the participants in the Communities of Practice don't always meet often, they find the communities valuable because they can get and give insight into problems, and create best practices together that are tightly coupled to what they learn together (Wenger et al., 2002, 4-5). The knowledge in these communities is dynamic in nature, it cannot be simply saved as it lives organically within the participants of the community and in their actions and interactions (Wenger et al., 2002, 9). According to Wenger et al. (2002, 10) in the current fast paced environment one's own perspective isn't enough to solve complex problems anymore, collaboration with others is often needed.

In order for communities or practice to achieve the full potential, they should be managed in ways that encourage participation, the community should be given a chance to influence the decision making, and any barriers in the way of participation should be removed. The communities develop naturally, and when trying to influence their development, it should be done in an encouraging way. (Wenger et al., 2002, 13) After all, communities of practice require active participation from the majority of its members in order to operate effectively (Ardichvili, 2003).

Communities of practice bring value to its members. In the short-term they get help from other experts, they gain confidence in their abilities, it is fun to be with colleagues and feel belonging. The long-term value the members get from the community of practice include: gaining skills and expertise, keeping up to date with the field, improve reputation, and get a professional identity. (Wenger et al., 2002, 14-18)

2.2 Online knowledge sharing communities

Online communities make it possible for people to use existing knowledge that is not available locally (Wasko & Faraj, 2000). A difference between traditional communities and online communities is that users don't meet in person and most likely don't even know each other (Yan & Davison, 2013). Knowledge collaboration in online communities involves: "individual acts of offering knowledge to others as well as adding to, recombining, modifying, and integrating knowledge that others have contributed." (Faraj et al., 2011, 1224). Users share their knowledge and also combine their knowledge on top of existing knowledge in content created by the community (Faraj et al., 2011). Therefore the community itself is responsible for the generation and maintaining of the content.

It is important to understand that technology alone doesn't guarantee that knowledge would be exchanged, technologies merely offer tools that allow knowledge exchange to happen online (Cabrera & Cabrera, 2002). According to Jin et al. (2015) before technologies that made it possible for online communities

to exist were available, people tended to be passive knowledge consumers. There weren't that many opportunities for people to share their knowledge by using technologies because the technologies back in the day didn't allow back and forth communication, it rather was a one way street in which it was possible to communicate knowledge to the masses, but the masses could not take part in knowledge creation. Technologies have later made it possible and encouraged more people to take part in knowledge creation (Jin et al., 2015) by allowing back and forth communication.

Malinen (2015) argues that researchers have not found a consensus on the definition of online community. In a systematic literature review of 83 articles in the area of online communities Malinen (2015) found that online community was generally defined by the majority of the papers as: "software that allows people to interact and share content in the same online environment" (Malinen, 2015, 236). This study adopts the definition of online communities that Malinen (2015) came up with.

2.2.1 Online communities

Online communities have been referred to with several different terms, one of which is "virtual community" (e.g. Ardichvili et al., 2003; Porter, 2004; Chiu et al., 2006), another one being "online community" (e.g. Lampe et al., 2010; Faraj et al., 2011; Lai & Chen, 2014; Jin et al., 2015; Malinen, 2015), and an even older one is "electronic community" (e.g. Wasko & Faraj, 2000), one also used is "electronic network" (e.g. Wasko & Faraj, 2005). There could be even more, but these seem to be the most used terms. Probably because of this distinction the terms all of these terms have been also used together with community of practice as well. There is no need to get confused about these as they all essentially mean the same thing. Porter (2004) defines virtual community as a group of people that communicate about a shared interest via a technology.

It is not well understood how collaborative online communities structure and organize themselves, although some research has recognized that these communities are loosely coordinated and self-organizing (Faraj, Kudaravalli & Wasko, 2015). The majority of online communities don't seem to have formal role structures and clear hierarchy that would dictate who manages and which activities within a community. Rather online communities seem to mostly function on their own without clear leadership. (Faraj et al., 2015) In online communities ideas can be separated from their authors and the development of ideas does not even require the original authors to be present. Anyone can contribute in a parallel manner, making it possible for the usage of others' ideas, and further development of those ideas in different contexts simultaneously. (Faraj et al., 2011) One unique aspect to online communities is that communication can be one-to-many, meaning that for one message there can be multiple recipients. Also due to the asynchronous nature of the communication, the recipients don't

have to be present to receive the messages. (Wellman et al., 2001) Online communities can help connect different professional groups together, allowing input from a wider perspective of views (Ardichvili et al., 2003). In communities focusing on specific topics or areas, it might be typical to share information relevant to the topic in the form that might only be understood by professionals of that field. For example in programming related communities it is common to share small portions of code when it is related to the discussion (Faraj et al., 2015).

Online communities of practice will probably fail if there are no contributors. Knowledge contributions take place mainly when the users are motivated to use the community. (Wasko & Faraj, 2005) Motivating users to share knowledge is the biggest challenge virtual communities face (Chiu et al., 2006), and relatively only a few online knowledge sharing communities have been able to motivate users to become active participants in knowledge sharing (Lai & Chen, 2014). Without knowledge sharing happening, there is no supply of (knowledge) content (Chiu et al., 2006). According to Ardichvili et al. (2003) one should not try to manage communities of practice, but instead try to create conditions for the community that would encourage knowledge sharing. One such condition could be an encouraging social environment (Cabrera & Cabrera, 2002). According to Wang and Noe (2010) interpersonal ties individuals have with other users can encourage knowledge transfer and influence the quality of the information shared. The feeling of belonging to a community is yet another motivating factor for users to continue using communities for a long time (Jin et al., 2015).

According to Jin et al. (2015) people who share knowledge in online communities voluntarily without rewards, do so to gain attention. One way to show knowledge contributors that they get attention is to visualize it in some way. Users who seek for information from online communities often either don't find it, or find too much information. It is therefore important to try to make relevant information available to users as efficiently as possible. (Jin et al., 2015.) These kinds of issues should be taken into account when designing software for online communities to encourage users to contribute. It is widely accepted that online communities can only keep existing if the users are contributing into them (eg. Joyce & Kraut, 2006; Ridings & Wasko, 2010; Nov et al., 2010).

2.2.2 Knowledge sharing platforms

The age of social media has provided the internet with multiple different platforms that allow knowledge to be shared online. The popularity of these social knowledge sharing platforms has been growing in the last decade (Nov, Naaman & Ye, 2010), and it seems that they still continue growing in popularity. Social media did not initiate knowledge sharing platforms, but has rather popularized the concept. Among the oldest platforms for knowledge sharing are

news groups, such as Usenet newsgroups, and bulletin board systems, widely known as BBS's. Especially news groups have been a popular target for studies on knowledge sharing (e.g. Wasko & Faraj, 2000; Joyce & Kraut, 2006; Faraj, Kudaravalli & Wasko, 2015).

There are many collaborative online knowledge sharing platforms available for different kinds of communities, organizations, topics and needs. Wikipedia and other wikis are some of the most well known communities and platforms that aim to collaboratively produce open information in the form of articles, on a vast amount of different subjects. Wikipedia for instance tries to make the collaboratively created information openly available and searchable for free, and allow anyone to contribute.

However, not all knowledge sharing platforms are open. Many companies and organizations have their own closed knowledge sharing platforms, which in literature have often been referred to as knowledge management systems (KMS) (e.g. Alavi & Leidner, 2001). The intention of KMS's is to support creating, transferring and storing organizational knowledge (Alavi & Leidner, 2001). Examples of some modern organization only social knowledge sharing platforms include for example Jive and Yammer, while many companies might also have organization only wikis, chat-based collaboration tools and so on. The reason for keeping these platform closed from the public probably has to do with the fact that the organizations think that the knowledge they possess is valuable and gives the companies a competitive edge.

Questions and answers (Q&A) platforms are a specific type of online knowledge sharing platforms which often become knowledge repositories, be them open to anyone or just to be used within an organization. These platforms have become quite popular recently. Open Q&A platforms are often collaboratively community driven when it comes to the content: users ask questions that other users attempt to answer. Online Q&A communities often add a social dimension by linking users, questions, and topics (Jin et al., 2015). Some of the most widely known general purpose Q&A platforms include for example Yahoo! Answers and Quora. These kinds of platforms don't only aid in the immediate information hunger of the original asker: the questions and answers are saved for later use, so they will also serve the future users who search answers for the same or similar questions later on. (Anderson, Huttenlocher, Kleinberg & Leskovec, 2012; Treude, Barzilay & Storey, 2011) On Q&A platforms the community controls which content is important and which isn't (Anderson et al., 2012). The members can indicate the quality of contributions by voting on the questions and also on the answers (Jin et al., 2015). The vote count, be it positive or negative, not only indicates the quality but also the relevance of the content, for example whether or not the provided answer is really answering the question that was asked.

Stack Overflow is a open Q&A platform intended for topics related to computer programming. It has become a very popular platform to help in

learning and problem solving of computer programming related issues. (Anderson et al., 2012; Treude et al., 2011.) Due to its popularity, Stack Overflow has even become a replacement for some official documentation of computer programming related products (Treude et al., 2012). One of the reasons for its rise to popularity could be that a question and answer form is easier to understand than complicated technical documentation, at least when it comes to finding specific information about a specific issue.

Internet Relay Chat (IRC) is a protocol that enables implementing distributed chat networks (Oikarinen & Reed, 1993). Multiple different IRC networks host many chat based knowledge sharing communities that are open to anyone to participate in. For example a network named Freenode hosts official chat-based communities of several open-source projects. In these communities, due to the nature of real time chat, knowledge exchange can occur with less delay and in a more personal manner. The IRC communities are often used as support channels for open-source software, and people come there for advice when facing problems with the specific open-source software. An unique feature of IRC is that the messages are ephemeral in nature and are not stored, as mentioned by Rheingold (2000, 154):

Chat systems lack the community memory of a BBS or conferencing system or MUD, where there is some record of what was said or done in your absence. Although words are written and broadcast (and thus can be electronically captured, duplicated, and redistributed by others), they aren't formally stored by the chat system. The discourse is ephemeral.

The ephemeral nature of IRC could be seen as a benefit for knowledge sharing. The users might feel more free to express their ideas and make mistakes without having to worry that someone is monitoring all their activities. The possibility for surveillance has been found to hinder knowledge sharing, as it could be possible for someone to find out others failures (Young, Kuo, & Myers, 2012).

While the social media platforms in general, and platforms like StackOverflow and Slack, have grown tremendously in recent years, the popularity of IRC has severely declined. Pingdom (2002) interviewed the creator of IRC protocol, Jarkko Oikarinen, who thought the commercialization of the internet to be a reason for popularity decline of IRC, companies want to build walled gardens preventing users leaving their services, and open distributed solutions don't fit that scheme. Between 2003 and 2012 IRC networks had generally lost 60% of their users. This had happened to all the major networks except Freenode, which is the only major network that experienced a steady growth during this period. The reason for the growth of Freenode could be due to its focus on free and open source software, which has led to many growing open source projects to have their official chat channels on the network. (Pingdom, 2002) According to statistics presented by Netsplit.de (2018) the growth of Freenode stopped at

the end of 2013, and has since remained stable until 2018, even though experiencing a slight decrease in the number of users during this time.

2.3 User participation

Whether an individual is considered to be a participating user of an information system may depend on the point of view. From the users viewpoint, being a user of some information system might not be the same as for example the view point of the owner of the information system. The user can exist to the system owner as a user profile stored in the database when the user registered an account and used it once. However, from the users point of view the user might not consider oneself a user of the system, the user might have even forgotten about it. (Suhonen, Lampinen, Cheshire & Antin, 2010) Despite of this, according to Malinen (2015) researches often define users as participants after they have used the system in some way.

In information system research user participation has been traditionally used to refer to participation in the development process of information systems (Hartwick & Barki, 1994). In online community research however, user participation has been widely used to refer to the participants of online communities, and to how the participants use the online communities, and how much do they use them. User participation has often been conceptualized with an active-passive dichotomy (Malinen, 2015). This thesis will follow the notion of the online community research when it comes to using the term user participation, it will be used to refer to participation in online communities, rather than participation in the development process of information systems.

2.3.1 Types of participation

Users are often categorised into two participation types: active and passive. These two groups are often described as posters and lurkers. (Lai & Chen, 2014; Malinen, 2015) Whether an user is considered to be active or passive, is often defined by the quantity of actions as metrics. Passive participation has not always been considered to be a participatory action, which has caused some debate around it. (Malinen, 2015) Malinen (2015) calls for more complexity in the analysis of participation, as only two types of participation (passive and active) might be too much of a simplification of the phenomenon. Even though users might not directly contribute content, they might be participating for example by voting on posts which can be a crucial activity for Q&A platforms to indicate the quality of the posts, and for media sharing platforms indicating the quality of the posted content. Another activity type could be users reporting violating

content, which many platforms rely on to identify the violating content on their platforms.

It is known that all users don't contribute equally, and usually a smaller portion of users make the majority of the contributions (Lampe, Wash, Velasquez, & Ozkaya, 2010; Carron-Arthur, Ali, Cunningham & Griffiths, 2015). Research on online participation has not typically taken into account the quality of the activities of the users, the focus has been more on the quantity of the users activity (Malinen, 2015). According to Carron-Arthur et al. (2015) it is not possible to find out how a user contributes just by looking at the frequency of posting activity. Users can contribute in different ways for benefit the online community, but not all user activities are beneficial to the communities, trolling and other ways of disturbing online communities can even be harmful to the communities (Malinen, 2015).

For an online community to be sustainable and for it to be able to stay lively, user participation is required in the form of conversations and responses (Joyce & Kraut, 2006; Ridings & Wasko, 2010; Nov et al., 2010). Because of this, transitioning passive users to become active has been of an interest to researchers (Malinen, 2015). Users can however be active in many different ways, for example in Q&A communities users can give votes to the content created by other users. Jin et al. (2015) gave advice to visualize attention in order to encourage people to share knowledge. Voting on Q&A communities can certainly be considered as a visual cue showing the knowledge sharers that their contributions get attention and are appreciated, further motivating them to keep contributing. So depending on the features of the platform, there can be many kinds of activities that can support and keep the communities lively, and most importantly encourage future knowledge sharing contributions.

Communities generally have three kinds of participation groups as can be seen in Figure 2. There is a core group which is the most active group of people whose goal is to advance the community. The second group is the active group, people in this group participate occasionally. The third group includes the people who are at the peripheral, which is the majority of the people in the community. These people rarely participate and mostly just lurk. Lastly there are outsiders who are interested in the community but don't participate in any way. The people in the community move between these groups organically, as the community changes or their own motivations change. Successful communities encourage people into active participation by making the center of the community interesting and allowing people to fluidly move between groups. (Wenger et al., 2002, 56-58)

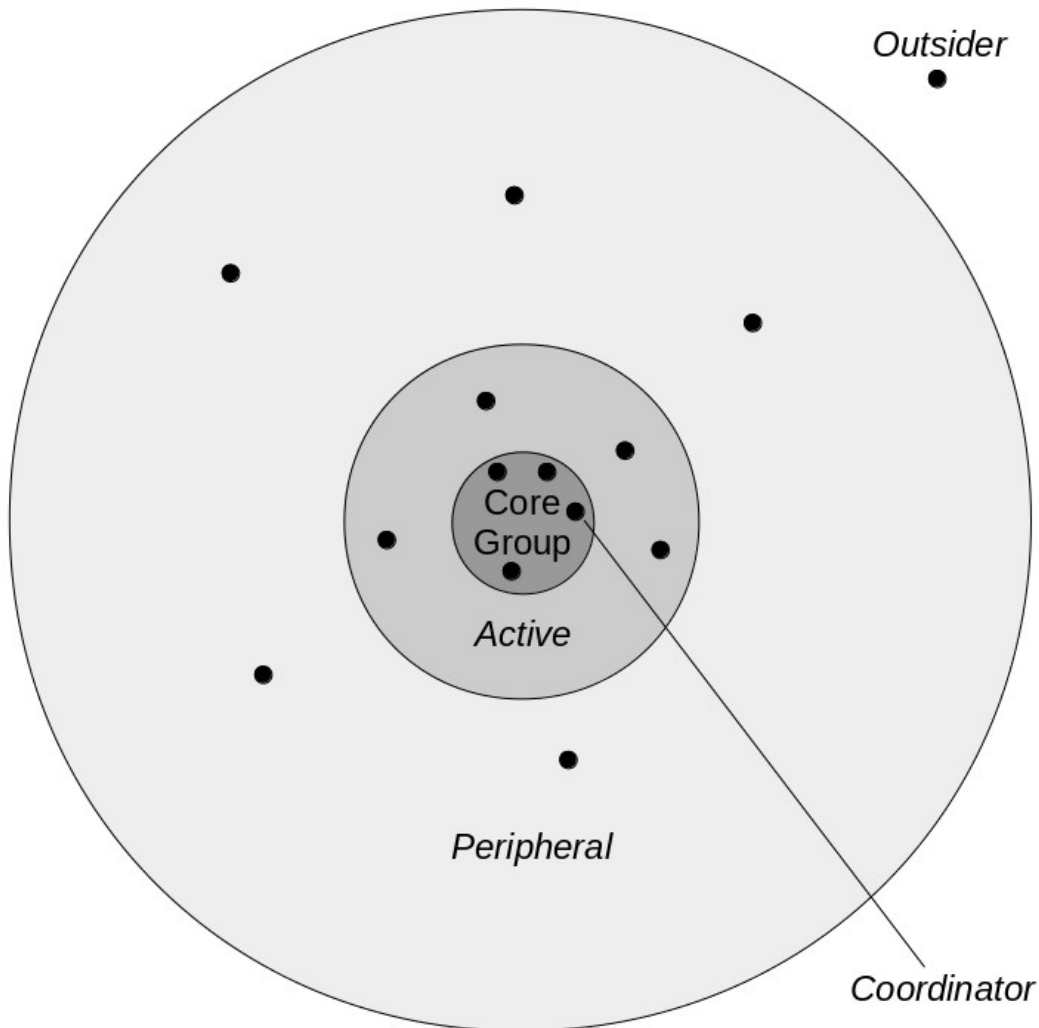


FIGURE 2: Degrees of community participation (adapted from Wenger et al., 2002, 57)

2.3.2 Why users participate

Participating in a community can be a learning experience that is both enjoyable and social (Wasko & Faraj, 2000). According to Wenger et al. (2002, 55) people participate in communities for at least three reasons: direct value, personal connections, and improving skills. Malinen (2015) identified five topics from literature that can explain user participation in online communities: motivations, personality traits, values, group processes, and technology and policy. Motivations have been the factor of interest in many studies analyzing user participation in communities. Gray (2004) conducted a study in the context of an online com-

munity of practice and found three motivations for users' participation: learning new skills, socially connecting with colleagues, and reducing geographic isolation. Learning may not be a motivation only for the passive users, it can also be a motivation for those that share knowledge, as they are looking to improve their own understanding in order to be able to explain something well (Wang & Noe, 2010). Socially connecting with colleagues, and geographic isolation might result from working in virtual teams of diverse professions, one might not get to engage in many face-to-face social interactions with others from the same profession, and the sense of belonging to a profession could be diminished due to not being able to communicate with professionals of ones own field.

Wasko and Faraj (2000) identified generalized reciprocity to be the major reason why users shared knowledge in a newsgroup. The users might participate in knowledge sharing because they expect their participation to encourage others to also participate in knowledge sharing - they don't expect the exact individual they helped to help them back, but perhaps to encourage them to help someone else. Bandura (1986) defines self-efficacy as what one believes to be able to accomplish with the abilities one has. Kankanhalli, Tan, & Wei (2005) found that knowledge self-efficacy and enjoyment in helping others had a significant impact on knowledge contributors use of electronic knowledge repositories. According to Cabrera and Cabrera (2002) the thought of belonging to a group has been shown to encourage contributions for public good. Lampe et al. (2010) confirm this idea that belonging to a community is an important reason for participation. Users also appreciate the social aspect of a community, and think that participation brings them satisfaction and that helping others is fun. Communities have also been found to be useful in keeping up to date with the topic of interest, therefore knowledge should be available so that the users would find the community to be useful in this sense. (Wasko & Faraj, 2000)

The two participation types of online knowledge sharing communities, active and passive, have slightly different motivational factors driving them (Lai & Chen, 2014). In online community literature, motivations for participation have generally been split into two categories: intrinsic and extrinsic (Malinen, 2015). According to Lai and Chen (2014) knowledge sharing intentions of active users are motivated more by intrinsic factors and passive users tend to be motivated more by extrinsic factors.

Jin, Li, Zhong, & Zhai (2015) studied why users share knowledge in a Chinese online Q&A community Zhihu. They collected data from the users actions for a time series of 15 weeks. The study is one of the few on the field of online communities that have collected panel data from user actions. They made three major findings; users who disclose more personal information (in their profiles for example) contribute more in Q&A communities by sharing knowledge; the feedback received from other users has a positive effect on contributions; and that social learning opportunities encourage more contributions. (Jin et al.,

2015.) Soliman and Tuunainen (2015) also identified feedback to be a major reason for continued participation in a crowdsourcing photo market platform. Giving feedback can make the users' feel that their contributions matter (Bock et al., 2005). The likelihood of continued participation in online knowledge sharing communities can be greatly influenced by whether the person receives responses to his or her first participation (Joyce & Kraut, 2006). For an online discussion group to be sustainable, the individuals taking part must collectively participate in creating conversations (Ridings & Wasko, 2010). Lampe et al. (2010) conducted a study on an online encyclopedia platform similar to wikipedia, and found that continued participation might happen for different reasons than initial participation. The motivations of the user for continuing participation can change over time (Lampe et al., 2010). Soliman and Tuunainen (2015) also found that motivations can change from initial to continued use, initial use being driven by selfish motivations, and continued use being driven by selfish and social motivations.

Wohn, Velasquez, Bjornrud & Lampe (2012) argue that habit might be a better defining factor for continued use than motivation. Users behave differently and engage in different kinds of actions within online communities, and might conduct these activities out of a routine without consciously choosing why and how they use the online community. They found that passive use can be explained with habit, while active participation cannot. (Wohn et al., 2012)

Lampe et al. (2010) make an important comment about the research done on user contributions; research has rarely been done from the viewpoint of why users do not contribute. By only examining the reasons why users contribute, it could potentially lead to a one sided view of the issue, and perhaps more research should be done from the viewpoint of why people do not contribute to online communities. This viewpoint has not been totally ignored though, as for example Ardichvili et al. (2003) found that some users might not want to contribute because they are afraid of criticism, or that they fear they might not be able to give accurate and relevant answers. According to Chiu, Hsu and Wang (2006) users' confidence plays a role when contributions are voluntary, when they doubt their knowledge sharing abilities they will not be likely to contribute.

3 METHODOLOGY

In this chapter the research methodology used in the empirical study will be discussed. This chapter consists of research goals and the motivations for the study, the research methods and the reasons for their selection, how the data was collected and how it was analyzed, a description of the study case, and the research ethics that concern this study. The goal was to find out how users of online communities become active knowledge contributors. Qualitative approach was the selected research method for this study. The case of the study was programming focused communities on Freenode IRC network. 18 semi-structured interviews were undertaken, lasting between 30 minutes and 2 hours.

3.1 Research goals

The aim of the empirical research was to find out why users share knowledge, and how do users become active participants in knowledge sharing on programming focused online communities. In this study the focus was on those users that were currently in the active group and the core group in the degrees of community participation (Figure 2), meaning that the users were actively participating in knowledge sharing. Users move between the groups organically, but finding those users that were once active but were no longer active would have been a troublesome task, therefore the focus was on the currently active users, as it would be possible to identify them. Freenode IRC networks programming focused communities fit well with the research goals, as programming focused knowledge is actively shared on them, and identifying the active users was possible with little effort.

This lead to the following research questions:

- RQ1: *Why do users share knowledge on Freenode's programming focused communities?*
- RQ2: *How do users of Freenode's programming focused communities become active knowledge contributors?*

The study concentrates specifically to what happens in-between the first visit to the community, and the point when the users can be considered to be active knowledge contributors. The point of interest of the study were the users in online communities that were active during a specified duration of time, and could be considered to be participating in knowledge sharing within a certain community. Multiple studies have researched why users participate in knowledge sharing in online communities (e.g. Wasko & Faraj, 2000; Ardichvili et al., 2003; Gray, 2004; Jin et al., 2015), but only a few have studied how users become active knowledge contributors in online communities of practice. Joyce & Kraut (2006) did study how responses to first posts affected continuing participation for newcomers in newsgroups, and found that the users whose first posts received replies were more likely to continue posting.

It has been noted in literature that active participation is necessary for the sole existence of online communities (e.g. Joyce & Kraut, 2006; Ridings & Wasko, 2010; Lampe et al., 2010; Nov et al., 2010) and that it is often a small portion of users that make most of the contributions (Lampe et al., 2010; Carron-Arthur et al., 2015). It can be said that this relatively small group of users is the reason why these communities keep existing. Therefore it can be important from the viewpoint of online communities focused on knowledge sharing to understand how the users become active knowledge contributors and what motivates them in each step of their journey.

By finding out the different paths the users go through before becoming active knowledge contributors, it might be possible to identify people that could be likely to become active knowledge contributors. With this information it might be possible to find ways to encourage more users to become active knowledge contributors, thus helping online communities to become more sustainable, and perhaps even more useful for newcomers. If a community starts to appeal to new users, it might bring on more users, making the community more lively. This information could also interest those who are launching new services focused on knowledge sharing, as bootstrapping is a known problem for example to new social networks. Bootstrapping in this context means that the appeal of a social network is in its user base, without many users the social network isn't attractive to new users. It has been widely accepted that without user contributions, communities are not very likely to offer their members anything worth staying for, and therefore might cease to exist in the future (Joyce & Kraut, 2006; Ridings & Wasko, 2010; Nov et al., 2010). Motivating users to contribute knowledge is challenging (Chiu et al, 2006), and only some communities have succeeded in this (Lai & Chen, 2014). It is therefore crucial for communi-

ties to find ways to help them gain more contributors, especially in the case if they are an emerging community or if they lose existing contributors. The results of the study could provide new insights on the subject, which could then later be verified by quantitative studies.

3.2 Research approach

The qualitative approach was the method of choice for this study. In a fairly recent systematic literature review of user participation in online communities, Malinen (2015) reported that out of 83 selected articles from 2002 to 2014, only 16% were qualitative in nature. This indicates that there is room for more qualitative studies in this research area, thus making the choice of research method relevant. Qualitative methods have been used in the research of user participation in online communities to gain a deeper understanding of the experiences of the users (Malinen, 2015). As this thesis aims to gain a deeper understanding of the users' experiences, and specifically how they become active knowledge contributors after their initial use of the online community, the choice of qualitative method seems justified.

This thesis uses the inductive research strategy, meaning that the aim is to build theory from observations, the observations in this study's context being stories. Inductive approach was chosen because it was decided to do narrative interviews, which basically means collecting the stories of the interviewees, and when analyzing narratives, it can be useful to have an open mind and not to be directed by some theory. This can allow the researchers to make findings that are beyond the scope of some theory, and doesn't therefore restrain the thinking of the researcher.

In order to answer the research question, this study aimed to build a process theory that would show the different stages and the users go through before they become active knowledge contributors in online knowledge sharing communities. So in essence the target of the study is the time in-between the first use of the knowledge sharing community and the time when a user became an active knowledge contributor. Stories (or narratives) were collected in order to build the process-theories.

Because narrative embodies sequence and time, it is naturally suited to the development of process theories and explanations. (Pentland, 1999, 717)

The stories were collected by interviewing users who were identified to be actively contributing to knowledge sharing in programming focused online communities. The process theory was built based on the analysis of the interview data.

3.3 The case of the study

This section will describe the chosen case and explain why it was chosen. The chosen case was Freenode's programming focused communities. Freenode is an internet relay chat (IRC) network focused on open source software. The programming focused communities on freenode were chosen as the case because they represent a specific kind of knowledge sharing community that is mature, but has not been studied before from this point of view. IRC is unique for example in the way that the messages are not saved and that there is no reputation point system present. Freenode is also representative for programming focused communities as many open-source projects have their official community (support) chats on it. What clearly differentiates IRC from almost anything, is that generally you don't need to register anywhere to use it, although some communities on IRC require that. IRC is free to use for anyone.

IRC is a protocol that people can use to communicate with each other in real-time. It was implemented in 1988 to replace BBS (Bulletin board system) chats. In 1993 a Request for Comments (RFC) 1459 was submitted to the Internet Engineering Task Force (IETF) by Jarkko Oikarinen. (Oikarinen & Reed, 1993) RFCs are publications or documentation of IETF standards (Hoffman & Harris, 2006, 29-30). The IRC protocol is a ruleset for implementing IRC servers and clients. In the context of IRC, IRC network is a collection of servers connected to each other as a network, and can be considered being a platform in the context of this study. Users can discuss with other users connected to the same network by joining channels of certain topics, or by privately sending messages directed to a desired user name. Users are identified by a user name, which can't be reserved according to the protocol, but some IRC networks have implemented features that allow users to register and reserve user names in order to protect one's identity.

IRC has a special characteristic not present in other online platforms used for knowledge sharing: the messages to a channel are generally not logged by the IRC network to be later shown to users who were not present at the time. This means that a user has no direct access to messages sent during the user's time offline, making the user possibly unable to take part in discussions that started before the user was online. When thinking of an IRC channel as a place for knowledge sharing, the knowledge available to the user is usually only available in that space during a relatively short period of time, because as more messages come in, the earlier messages kind of get lost in the list of messages. IRC channels can be considered to be virtual spaces for knowledge sharing in which the knowledge available at a certain time is dependent on the users that are online and actually present at that time period, and the knowledge that these users hold. This distinguishes IRC as a knowledge sharing platform from many other social knowledge sharing platforms: it is potentially possible to get help

faster, but also if you don't get help quickly, it is likely that you won't get any help at all, especially if the channel is very active and unless you repeat your question. Because of this, the effort the user has in forming questions and providing information about them could have a direct effect on whether the user receives any responses to them, especially during busy times when there are a lot of questions coming in rapidly.

IRC networks (as platforms) could be compared to other chat-based collaborative communication platforms e.g. Slack and Gitter, which have been gathering some popularity as of lately. One of the major differences between these newer platforms and IRC is that the newer platforms often save the logs of the discussions, and possibly show different kind of (push) notifications to their users. It is also possible to log discussions on IRC with the client program, but the users are generally required to set up logging themselves, and during offline time logs cannot be produced. When comparing IRC to Q&A communities, one major difference is that the questions can't be searched for on IRC, and that there are no voting mechanisms present to indicate the quality and the relevance of the answers. The knowledge seeker therefore has to evaluate the quality of the answers themselves. Users could perhaps also evaluate the answers depending on who gives them, by observing the channel dynamics and figuring out who seem to be knowledgeable or considered to be experts. Some users might have gained a reputation for providing quality answers or advice in the eyes of the knowledge seeker.

3.4 Data collection

The chosen method for data collection was interviews. The case of the study was IRC network Freenode's users on programming focused communities (also known as channels). A total of 18 Freenode IRC network users on programming related communities were interviewed. The communities from which the interviewees were selected from are not listed in order to keep the interviewees identities as anonymous as possible. By naming the communities, the possibility of identifying interviewees would increase, for example based on their writing style due to the usage of direct quotations.

3.4.1 Selection of interviewees

In this section the procedure of interviewee selection will be explained in detail. The goal was to find about 20 individuals to interview using purposive sampling. Purposive sampling, which is also known as purposeful sampling, is a type of nonprobability sampling (Daniel, 2012, 87-88). In purposive sampling the subjects are selected based on how well they match the goals of the study

and how well they fit the study's inclusion and exclusion criteria (Daniel, 2012, 87-88).

The case of the study was chosen to be Freenode IRC networks' users on programming and software development related chat communities. The communities were selected based on several criteria that were thought to be fitting with the scope of the study. One criteria was that the channels had to have over 300 users at the moment of inquiry for user counts per channels. The reason for excluding channels that were smaller was to try to make sure that the communities were somewhat active. The user count criteria for the communities was decided after an initial look at the general user counts of IRC channels on Freenode.

The data collection of the user counts on the channels happened on November 24th at 13:57 (UTC+2) in 2017. A list of channels with more than 300 users was collected by utilizing a bot known as "alis", which Freenode IRC network offers as a way for gathering information about the channels on the network (Freenode, 2018a). Freenode also suggests an alternative method for gathering user counts, which would have been using a website netsplit.de (Freenode, 2018a). The "alis" bot offered easier access to the data, which could be easily acquired by running a few commands on the bot, which then would return the results as plain text. The other option, netsplit.de, only shows 20 results per a web page, which would have required building a web scraper, and then the results would have needed to be parsed from the HTML (Hypertext Markup Language) of the web pages. In the end the "alis" bot ended up being less cumbersome for performing this task, and was therefore chosen for the job.

The search resulted in 163 channels which had over 300 user at the time of the query. For some reason at least one channel that was later found to be within the query parameters was left out. It might be that the channels user count happened to change during the time the queries were made to the bot. It is possible that some other channels were left out too, although it is not very likely. The missing channel was noticed because it was known that it would fit in the scope of the study and it had clearly more than 300 users on it.

The next step was to categorize these 163 + 1 channels. It was done in two steps, first the channels were categorized into rather specific groups, and after that to more general groups. The categorizing of the channels resulted in 11 categories in total. The general categories and their channel counts can be seen in Figure 3. The missing channel from the dataset belongs to the category programming language, boosting that category's channel count to 24 in total. Out of the 11 categories, 2 were chosen because they are directly programming related. The two chosen categories were "programming language" and "framework/library". The amount of channels was 37 and when adding the one that was missing, there were 38 channels in total from where to look for interviewees.

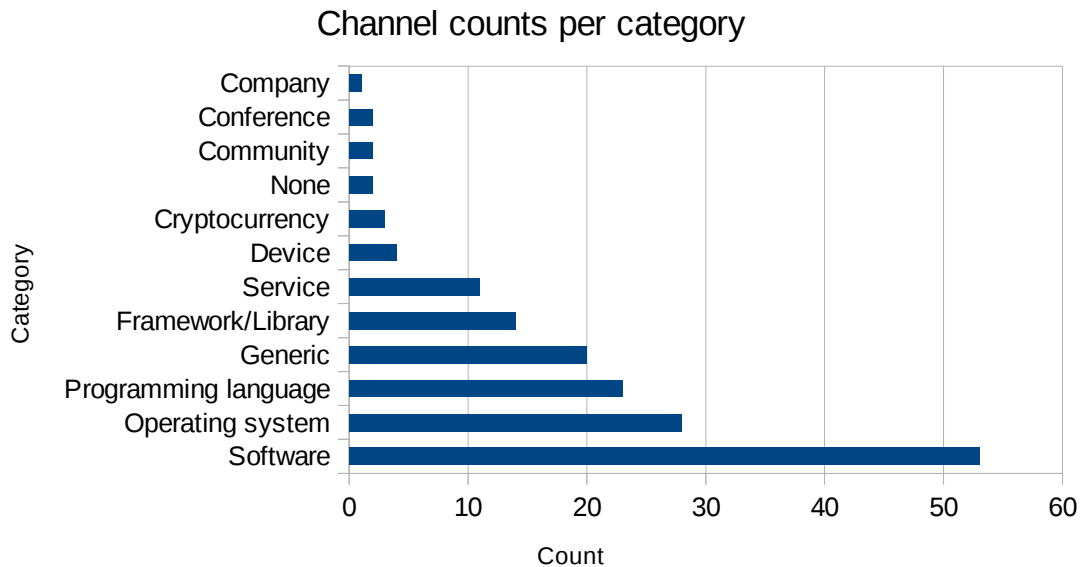


FIGURE 3: Channel counts per category

To help picking interviewees, chat logs were acquired for as many of the target channels as possible. In the end logs were acquired for 22 channels in total, with the addition of acquiring end result data for 2 channels which were processed by a third party according to given instructions, because the helping users didn't want to share the chat logs directly. Freenode's channel guidelines mention that sharing logs is not encouraged, at least if the channel does not explicitly state that it is being logged (Freenode, 2018b). Because of this, getting logs for some channels was problematic. Due to problems acquiring log data for all of the channels, 14 channels and their users were left out of the list of potential interviewees.

In order to produce a list of potential interviewees that could be considered being active in knowledge sharing, the logs were processed to count messages sent by each user during the time period of September 1st of 2017 until the end of 31st December of 2017. After this was done, data was available for 24 channels total, listing their most active users by messages sent to the channel. Users with less than 100 messages sent during the time period were removed from the data set, so that the list would not include user names clearly not relevant to the study. It should be noted though, that some users might have used several different user names during this period. It was not attempted to calculate total message counts for users, as it would have been close to impossible to identify this. It could have also violated the privacy of the users using different user names, as they might have their own privacy based reasons for choosing to do that.

The interviewees were then chosen from the lists of users that had written the most lines per channel during the selected time period. It was also manually confirmed from the chat logs that the user was active in knowledge sharing related activities, and that the user was not just a off-topic talker with a writing style that would account to many lines, thus making the line count of the user higher than those contributing knowledge. This manual verification was done to users of the 22 channels that logs were acquired for. After this it was possible to find out which users to ask to interview. The goal was to ask users to be interviewed from all of the 24 channels in an attempt to cover multiple channels instead of focusing on just a few.

3.4.2 Interviewees

In this section the backgrounds of the interviewees will be examined in detail, so that it is more transparent what kind of a group of people were interviewed. Potential interviewees were contacted by sending them a private message on the Freenode IRC network asking whether they would be willing to participate in the study by being interviewed. The people that agreed to be interviewed were from multiple different communities on Freenode. They were selected based on the processed chat log data, from which they were identified to be active in knowledge sharing, meaning that they helped other users who asked for help in the community. The total amount of interviews that were conducted ended up being 18. After making more inquiries for interviews it was decided that it was not worth the effort to push for more interviews, as it started to become increasingly more difficult to find interviewees.

The group of interviewees that agreed to be interviewed had very diverse backgrounds. The ages of the interviewees ranged from under 18 to under 50, with the biggest group being 18 to 30 year olds, as can be seen in Figure 4.

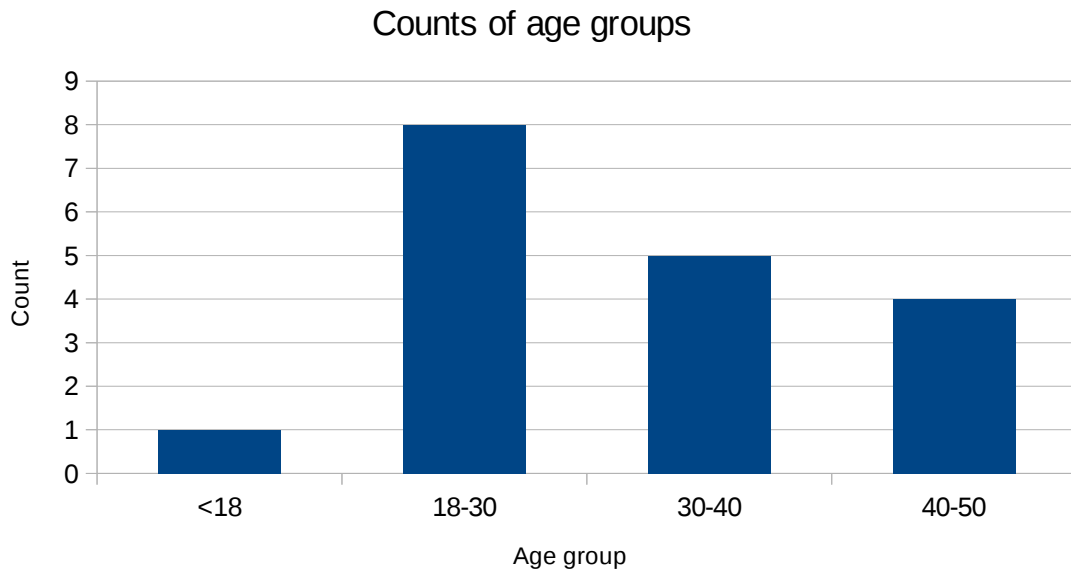


FIGURE 4: Interviewee age groups

Perhaps surprisingly the majority of the interviewees originated from the United States, as can be seen in Figure 5. All of the interviewees were from North America or Europe, except one who was from Australia,. It was to be expected that the majority of the users would come from these continents, as they include mostly developed nations in which the level of technology usage is high. However, it was rather surprising that there were no interviewees from continents like Africa, Asia and South America. There can be many reasons for this, including language barriers and cultural differences. Even time zones could have an effect, if the majority of Freenode users come from Europe and North America, it could be possible that users from other continents wouldn't show up in logs with high line counts, because there simply aren't that many discussions going on during suitable times for people from different time zones. It can also be that people from other continents simply don't use IRC, and possibly use something else in its place. In the interviews there were mentions of Chinese and Indian users, so there probably are some, but none of them happened to fit the criteria, or they could have declined the interview offer. As evident, there can be many reasons for getting interviewees mainly from Europe and North America, but the reasons can only be wondered about.

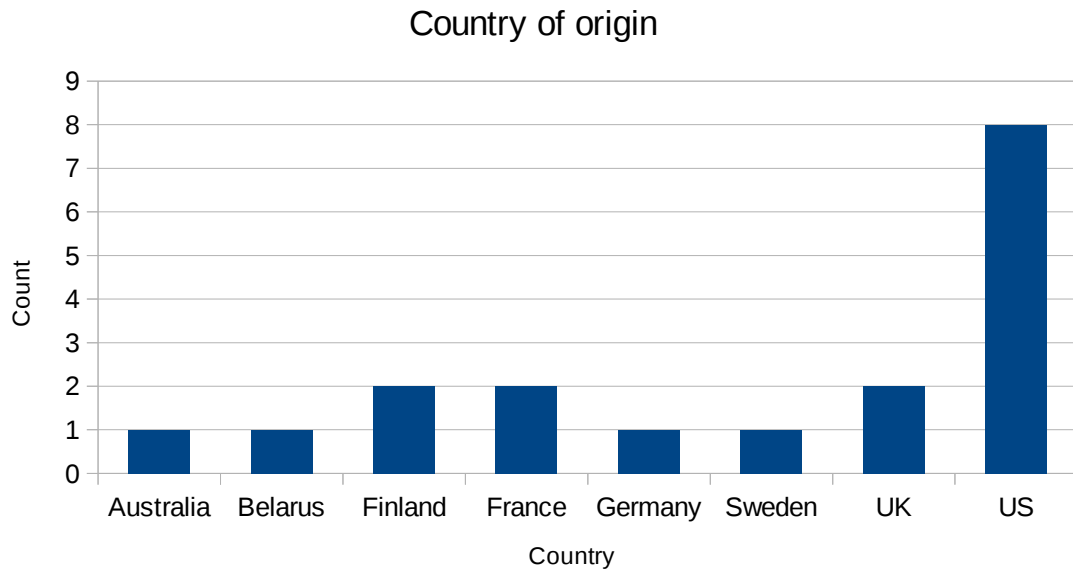


FIGURE 5: *Native countries of interviewees*

The majority of the interviewees had plenty of work experience in the field of software development, as can be seen in Figure 6. This perhaps isn't that surprising, as one might expect that helping people with programming related questions probably requires some experience and knowledge of programming. It is however entirely possible that someone knows a lot about programming but isn't doing it professionally, but this wasn't typically the case within the interviewees of this study. One interviewee was under 18 and didn't have any work experience because of this, and another interviewee had work experience in completely different field than software development. Most of interviewees had work experience in the field of software development.

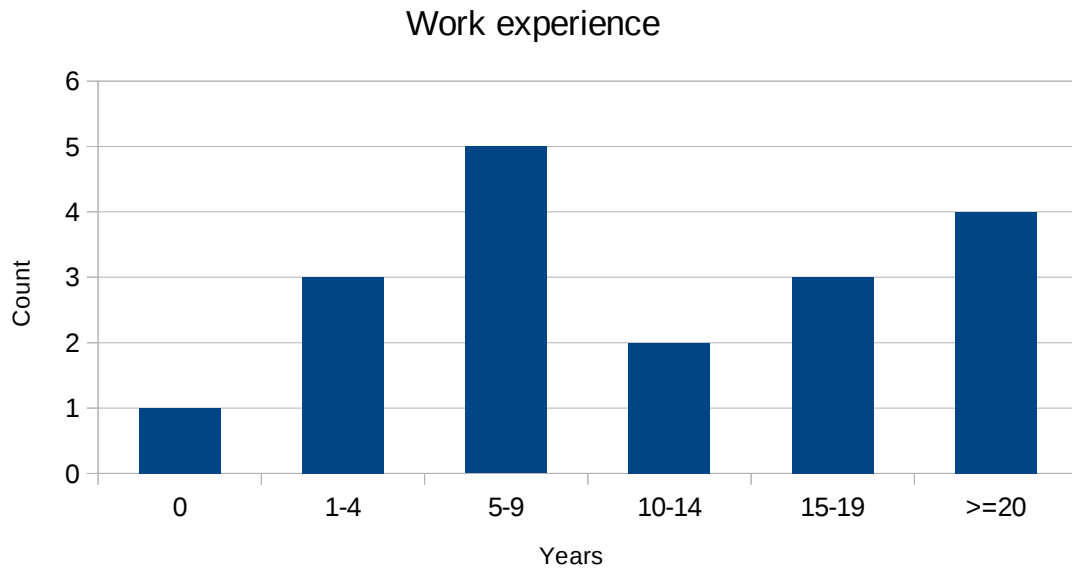


FIGURE 6: Work experience

Educational background varied a lot, and perhaps surprisingly about half of the interviewees' educational level was high school or below. The categories for educational levels try to illustrate the Finnish school system. The results might not be entirely accurate as the interviewees come from many different countries that have different school systems. The categories only give a very general view of the educational levels of the interviewees. PhD means that the person has done a dissertation, MSc refers to a Masters level of university degree, and BSc to Bachelors level of university degree. High school refers to the Finnish high school, which can be equivalent to upper secondary school, sixth form college (in Britain), or senior high school (in the USA).

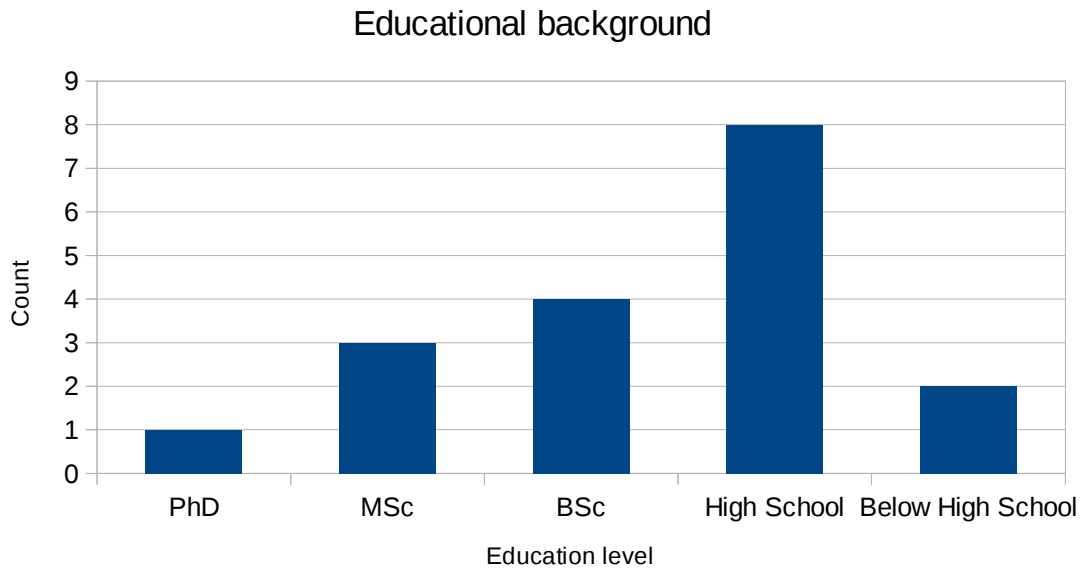


FIGURE 7: Educational background of the interviewees

It is important to note that the users interviewed were all identified to be active in knowledge sharing, and they all still used IRC. This means that any users that might have stopped using Freenode's programming focused channels were not interviewed. It could be that these users might have totally different stories to tell, but that is beyond the scope of this study. It is good to keep in mind, that the interviewees were all active current users of Freenode's programming focused chat channels.

Getting interviewees on Freenode proved to be both hard and easy. It seemed to be much easier to find interviewees on programming communities that covered higher level programming, while on the other hand communities focused on lower level programming proved to be harder to get anyone to interview. High level programming has more abstractions from the details of the computer, while low level programming deals with less abstractions and work on a closer level to the processor instructions. One possible reason for this observed difference might be that the users that were asked to be interviewed on the low level programming language themed channels were much more privacy focused, and did not want to share anything about themselves, especially to someone they did not know beforehand.

3.4.3 Interviews

The aim of the interviews was to find out how the interviewees became active knowledge contributors in Freenode's programming focused communities. The interviews were conducted online, since it would have been hard to find users

for face to face interviews, as Freenode users come from all over the world. The interviewing was performed on the Freenode IRC network. It was the chosen medium because of several advantages: the users were already comfortable with IRC and there was likely to be very few technical problems, the interview logs could be automatically be saved by both parties reducing the risk of losing the data, the interviews could be conducted online so that users from all over the world could be interviewed, and the interviews didn't have to be transcribed because the interviews would already be in text form. Another possibility would have been to conduct the interviews by phone or via video conference calls, which would have made the interviewing process much more complicated, and the risk of something going wrong with the recording would have been much higher. Video conference calls might have had some benefits though, for example it could have allowed picking of verbal cues, and some users might have preferred it over IRC. On the other hand there might have been difficulties with communication due to different pronunciations of English, and some interviewees might have even refused to do video conferences, but were comfortable doing the interview on IRC.

The choice of IRC as the interviewing medium proved to be suitable for this case. Only one hiccup was noticed that could have been removed by other means, sometimes it was hard to tell whether the interviewee was still thinking and was about to write more about something. That might have also caused the interviewing to take more time due to possibly unnecessary waiting done by the interviewer, but on the other hand the interviewees had more time to form their thoughts before they had to submit them.

The chosen interview type was semi-structured interview. According to Myers and Newman (2007) semi-structured interviews typically have an incomplete script and they require improvisation. In semi-structured interviewing the researcher has to give room for the story to develop and try to avoid directing the interview too much (Myers & Newman, 2007). According to Schwarz, Chin, Hirschheim and Schwarz (2014) individuals must be guided to look into their minds and to examine their behavior so that they are able tell their story. The best case happens when the interviewee can tell his or her story and analyze the reasons of his or her actions without the need for the interviewer to intervene much. The interviews weren't intended to be structured because the aim was to collect the stories from the users, and with a clear structure the interview might have lost some insight about certain things, as the interviewees could not ponder freely. Corbin and Morse (2003) note that unstructured interviews don't usually follow the same pattern and thus differ quality and length. The same is true for semi-structured interviews as well. This became evident in practice, as the interview types ranged between unstructured and semi-structured, depending on each individual interview. Some questions were prepared beforehand to aid in driving the interviewing forward in case the interview didn't flow naturally, but they were used in an improvised manner. Many of the questions

asked were formed during the interview in order to guide the interviewee to tell more about some topic or to concentrate on a specific viewpoint for example. It was attempted to intervene as little as possible, perhaps naturally some interviews required more intervening than others.

The interviews were thought mostly successful, even though interviewing was much harder than assumed beforehand. The interviewees were informed that the interview would take from 30 minutes to an hour. Almost all the interviews took longer than expected. The least time used was 30 minutes, but the majority of the interviews took at least an hour or more. The longest interview might have taken about two hours. Thankfully in the interviews that needed more time, the interviewees agreed to continue even though it took longer than what they agreed to. It seemed that majority of the interviewees were interested in the topic, and wanted to help by participating by being interviewed. This might have something to do with why they share knowledge with others to begin with. After the interviews had been performed and while doing the analysis, one interviewee was presented a question after the interviews in order to clarify the order of events in interviewees story.

3.5 Research ethics

National Advisory Board on Research Ethics (2009) mentions three ethical principles research should follow: the autonomy of research subjects, avoiding harm, and privacy and data protection. Regarding the autonomy of the research subjects, the participation should be voluntarily and it should be based on a consent. The participants should also be informed about the research topic, data collection method and the estimate of required time, the purpose of the data and how will it be archived for secondary use, and that participating in the study is voluntary. In case the participant is a minor, the researchers have to evaluate if it is necessary to ask for the permissions from the parents, as in some cases it is justified not to ask for their consent. (National Advisory Board on Research Ethics, 2009) The interviewees were informed about this before the interview, and they were asked for consent. It was mentioned to the interviewees that the interview data shall only be used for this study, and that it shall not be used by anyone else but the researcher and the supervisor, with the exception for the case if the interview data has to be verified in some way by a third party. It was decided that if the interviewee would be under 18 years old, the parents of the interviewee should be asked for consent. This was done because the goal of the research was to get participants not only from Finland, but from all around the world, and the laws and rules might be different to Finnish ones. The National Advisory Board on Research Ethics is Finnish and it was not certain whether the guidelines would be fitting to every possible country.

According to National Advisory Board on Research Ethics (2009) harm caused to participants can stem from the collection and storage of data, and from the publication of research. There are at least three kinds of harm that should be avoided: mental, financial, and social. In order to avoid these kinds of harm, the participants should be treated with respect, and privacy and data of the participants should be protected. (National Advisory Board on Research Ethics, 2009)

Protecting privacy and data is considered to be important in research ethics, while also being part of the Constitution of Finland. The participants' privacy may be at risk if the research data is not handled with care and according to plans prepared beforehand (National Advisory Board on Research Ethics, 2009) When the participants of a study share personal and intimate information about their lives, they put their anonymity at risk (Corbin & Morse, 2003). This research intends to keep the interviewees identities as anonymous as possible, and extra effort has been placed in to making sure that the interviewees' identities would stay anonymous. During the interviewee acquiring period it was made sure that no information was shared to anyone concerning whom, if any, in a community had been approached. Unfortunately, in one instance this information was leaked, while asking one person to be interviewed: the person decided to publicly announce that the person had been approached to be interviewed.

Serious consideration has been placed to make sure that no information regarding who have been interviewed would be leaked at any time, even after the research would be done with. A serious emphasis was also put on data storage, and it was decided that none of the interview data should be stored on any cloud storage platforms, even after removing identities from the data. This was done in order to make sure no data was leaked due to possible mistakes by any 3rd party. The data was only stored locally, and backed up to a local file storage server not accessible from the internet. Before the interview data was handled any further, it was anonymized by removing nicknames and names from the data. When directly quoting any interviewee in this thesis, it has been considered could the identity of the interviewee be at risk. According to Corbin and Morse (2003) it is always possible that someone might recognize the interviewee from a quote. It has been acknowledged that total anonymization of interview data is probably impossible, which is a reason for this careful handling of the data. In order to better protect the identities of the interviewees, the IRC channels where interviewees were looked for is not listed in this thesis. Before interviewing, the interviewees were informed that the data wouldn't be shared to any 3rd parties, and that anything the interviewee says can be quoted in the thesis in a way that the interviewee shouldn't be identifiable from it. If the interviewee wouldn't consent to this, the interview couldn't begin.

3.6 Analysis

The goal of the analysis was to build a process theory from the stories gathered in the interviews. The targets of the study were currently active knowledge contributors, and the objective was to study them from their initial use of the community until they became active knowledge contributors. This was done to narrow down the scope of the study. Pentland (1999) argues that stories (narratives) are well suited for building process theory, because the sequence of events and their timing can be identified from the stories and they can be placed on a timeline. In a process model the events form a chain in which one event leads to another. The scope of this study was to study the time after the initial use of the community until the user became an active knowledge contributor. Therefore the events are placed on a timeline with a beginning and an ending, and because of this the events before the beginning and after the ending are not within the scope of the study.

This research used an inductive research approach, which means that the goal was to build theory by identifying underlying patterns from observations. Theory was constructed based on the identified patterns, and in the case of this thesis the theory was a process theory. Keeping an open mind was one of the main goals in the analysis, which was possible because of using the inductive approach. The analysis concentrated on finding patterns in the data rather than following assumptions made in previous theories. Siggelkow (2007) notes that case data rich in details can be a great inspiration for new ideas, while it can also be useful for finding and filling gaps in existing theories. Even if having an open mind is good, it can be useful for a researcher to use the information from existing research as a compass guiding the direction of the analysis. It can even be argued that completely blocking out existing assumptions is not possible for a researcher. (Siggelkow, 2007)

The methods used for the analysis of the interview data were content analysis paired with narrative analysis. According to Tuomi and Sarajärvi (2002, 105), content analysis is an analysis method in which documents are analyzed in a systematic and an objective way, in order to produce a compact description of the collected data. In narrative analysis the aim often is to produce short summaries of stories that condense the storylines of the stories. These summaries can then be used to come up with a base story that represents the typical storyline of multiple stories in a condense format. (Saaranen-Kauppinen & Pulusniikka, 2006)

Building of the process theory was guided by the levels of structure in narrative (Pentland, 1999), which can be seen in Figure 8. The interview data was processed according to levels of structure in narrative. First the text interview data of each interview was processed to form a coherent story that was in a chronological order. After this the stories were compared to each other to find

similarities, which was done multiple times until it seemed that the most common story types were gathered. After the story types were ready, the key forces for each event were analyzed. Once the story types and their key forces were identified, the stories were converted to the fabula level of narrative, which included a generic description of the events and explained the relationship between the events. Then lastly the generating mechanism were identified for each event in the fabula.

Levels of Structure in Narrative

	Level	Definition	Example	
Participants' stories	Text	Particular telling of a story by a specific narrator	Actual text of his or her story: "When I showed up at the interview..."	Researchers' stories
	Story	Version of a fabula from a specific point of view	A new employee's own version of how he or she was hired	
	Fabula	Generic description of a particular set of events and their relationships	How a particular person was hired: what happened, who did what	
	Generating mechanisms	Underlying structures that enable or constrain the fabula	Overall recruiting process: how people in general are hired	

FIGURE 8: Levels of structure in narrative (adapted from Pentland, 1999).

When the interviews were finished and there was data from 18 individuals, the interviews needed to be transformed into the story level according to the levels of structure in narrative (Figure 8), by identifying the different stages in-between the initial use and active participation in knowledge sharing. The analysis began by transforming the interview data, which is at the text level of the narrative, into chronologically ordered events from the viewpoint of each interviewee. After the overall events had been formed for each interviewee, the narratives were then transformed to the story level. It was made sure once again that the events were ordered correctly in a chronological order, one event leading to another. After each interview was transformed to the story level, the stories were fine tuned while also looking at the generating mechanisms. At this point the interview data was viewed again in order to figure out the reasons for the events in the stories, and the structures that enabled or constrained the stories were briefly looked at. Once this was done, it was possible to start grouping the narratives based on the stories and also the generating mechanisms. Then began the process of comparing the stories with each other in order to find any recurring patterns from the interviewees' stories. The goal was to group the sto-

ries based on the patterns found, so that it would be possible to come up with only a few story types that would best describe the different kinds of stories. This part of the analysis required quite a few iterations, as while grouping the stories to story types there were constantly new things coming up when looking at the data from different perspectives. The events that were identified also changed names and meanings multiple times during the process, and with each iteration there were fewer story types left, meanwhile the stories were improving as well.

Once it was decided that the story types were at a satisfactory level, there were a total of 3 story types left. Each story type showed a different pathway to becoming an active knowledge contributor in online communities. Next the pathways and their stages were analyzed for generating mechanisms, meaning that work began to identify the forces that were affecting the users decision on moving from one stage to another. Previously the generating mechanisms were looked at for each interviewees story, and now this information was used when identifying the generating mechanisms for the pathways (story types). Once again the interview data, which is at the text level of the narrative, was also looked at to confirm the generating mechanisms and to possibly find hints of generating mechanism previously not identified.

When the generating mechanisms had been identified for all of the stages in the three pathways, it was time to begin creating the process theory. The process model sits on the fabula level of the narrative (which can be seen in Figure 8). The fabula level of narrative describes the overall story that best explains the generic set of events. To transform the narratives to the fabula level, the pathways were directly used to identify the most common story. The time spent on the numerous iterations when analyzing the interviews, forming stories, and grouping the stories into story types (later pathways) made it rather easy to come up with the fabula, as at this point it was pretty clear what the overall story would be. After the fabula had been formed, the final part of the analysis began: identifying the generating mechanism for the fabula, which would then form the process theory. The generating mechanisms identified in previous stages of the analysis were looked at and added to the fitting stages in the fabula. The outcome was a process theory that describes the overall story of how users of online knowledge sharing communities become active knowledge contributors.

4 RESULTS

This chapter showcases the findings of the empirical study. First will be shown the reasons for why users participate in programming focused online communities paired with the advantages of IRC to knowledge sharing. Then the three pathways showing how users become active knowledge contributors are presented, after which the process theory based on the pathways will be shown. While reading through the results, it is good to keep in mind that case of the study was active knowledge contributors in programming focused IRC channels on Freenode IRC network. Active knowledge contributor is defined in this research as posters or active users. According to Wenger et al. (2002, 57) definition of degrees of community participation which can be seen in Figure 2, most of these users would probably belong to the core group and some to the active group. Some communities tend to be more active than others, which was not taken into account when contacting users for interviews. It is important to note that these findings concentrate on users that became active knowledge contributors and still continue to contribute knowledge actively, not on users that did not become active. There are also users that have become knowledge contributors, but for some reason stopped being active. This study does not cover these users, nor their motivations for stopping being active.

The results section will include relevant quotes directly from the interviews. The direct quotes from the interviews are shown as is, with the following exceptions: non-relevant parts are cut out by using three dots inside brackets (...), and some parts that are censored to protect anonymity have been replaced with square brackets and an explanation (like [channels]). Also some obvious but small typing errors might have been corrected.

4.1 Why users participate in programming focused online communities

In this section the motivations for participating in freenodes programming focused online communities are discussed. In the social media era there are many different kinds of communities of varied topics to choose from. IRC has been around for a long time, and it still is actively used especially by programmers. The reasons and motivations for participating in online communities concerning RQ1 are presented, and also the advantages of IRC in knowledge sharing are discussed next.

4.1.1 Reasons for participation in online communities

There are many motivations behind participation in programming focused online communities. The motivations identified for each interviewee can be seen in Table 1.

TABLE 1: Summary of motivations for participation per interviewee

Inter- viewee	Advance commu- nity	Belong to commu- nity	Enjoy helping	Get help	Influ- encing	Learn	Problem solving	Reputa- tion	Social- ize	Stay up to date	Validate knowl- edge
#1		x		x	x						
#2	x	x			x	x	x		x	x	
#3			x	x		x		x		x	
#4		x	x	x		x	x		x	x	
#5	x	x	x		x	x			x	x	x
#6	x		x			x				x	
#7						x			x	x	
#8	x	x	x		x	x	x	x	x	x	x
#9	x	x			x	x	x	x	x		
#10	x	x	x								
#11		x				x			x	x	
#12		x	x			x			x		x
#13	x					x					
#14		x	x	x		x			x		
#15		x	x			x				x	
#16				x		x		x	x		
#17		x	x	x		x		x	x	x	x
#18			x		x				x		x

Most users want to learn or to keep up to date with some technology:

(...) with my profession being [software development], I have a strong reason to be online and be interested in the regular topics and conversations in the channels I'm active in. (Interviewee #8)

Some just want to belong to a community and look for a place to chat to socialize with others. Knowledge contributions seem to be motivated by reputation and the chance to influence others by helping them learn a technology. Some users really like to solve puzzles, and are willing to drop anything they are currently doing when presented an interesting one. IRC was thought to be good for especially new users who don't necessary know the required terminology, due to the immediacy present they can have back and forth discussions explaining their problem by answering to questions made by other participants. The

ephemeral nature of messaging on IRC was also thought to be beneficial as the users might feel more free to say what they think.

It is vital for the communities that their knowledge contributors don't stop contributing. It has been acknowledged that only a small portion of users provide the majority of the contributions (e.g. Lampe et al., 2010; Carron-Arthur et al., 2015). It was thought that if the top contributors stopped contributing it would have a major negative impact on the communities:

Well, honestly each channel really only has a limited number of regular question answerers. If you stopped the top five answerers from each channel, it would severely cripple the channels. (Interviewee #8)

4.1.1.1 *Learning*

One of the most common motivations for participation in online communities was learning. People can learn from the discussions by observing others questions and answers, asking questions or presenting problems themselves, and they can learn by helping others. One can even learn by making mistakes when giving answers to others questions, if there is a mistake another community participant might correct the misinformation. One interviewee had a strong opinion about correcting misinformation:

(...) if someone else answers it incorrectly, i'm very likely to jump in and correct them
 (...) correcting misinformation is more important than providing information (Interviewee #6)

Learning was a motivator for almost all of the interviewees, whether the interviewee was seeking knowledge or helping others by sharing knowledge. One interviewee even loved learning:

(...) the desire to get help got replaced by the desire to learn more, then the desire to help people (while still learning more. I love learning) (Interviewee #4)

Learning is not limited to just observing or asking questions, several interviewees mentioned that they learn but helping others:

one of the ways I learn is by helping others; that's something I enjoy doing, but it also serves as a sort of check for me -- do I know something well enough to explain it (Interviewee #15)

One perhaps surprising way how people learn on Freenode programming communities seems to be validating ones knowledge. An interviewee mentioned that helping others is an opportunity to learn even if one makes mistakes:

it helps me learn especially when i am wrong and someone corrects me (Interviewee #12)

There were a total of 5 interviewees identified to be motivated by validating knowledge which means learning when others correct ones mistakes.

4.1.1.2 *Influencing*

Trying to make others interested in the same topics was identified to be a motivation towards helping others by sharing knowledge, like one interviewee mentioned:

(...) I also want people to be interested in topics I care about, and sharing knowledge to people who are interested is the best way to do it (Interviewee #18)

If someone wants to spread the use of a technology they love, helping others to use that said technology is a concrete way to try to encourage them to continue to do so. One interviewee mentioned that if the interviewee and other volunteers would stop contributing knowledge, it would have an effect on the popularity of the technology:

I imagine the technologies I like to use will become less popular (Interviewee #1)

This could be one reason for why so many technologies come and go, once the users discover a new thing, they might stop their voluntary support to the previous one and direct it towards the newest thing. This would likely encourage less people to use the previous technology. It also might be that people are drawn towards communities that are currently popular. One interviewee had an interesting view of this phenomenon:

Communities focused around things like web frameworks are essentially ephemeral -- because those frameworks themselves tend to turn over very quickly. And once the people who were using it move on to something else, it is somewhat rare for the majority of them to all move on to the same thing. So it is more like looking at activity hives, with transient swarm aggregation happening, and each actor in a swarm following its own rule set. The focus for each actor is not the swarm, it is the interest of the hive and then incidentally the size or "feel" of the swarm collected around it. The hive (whatever the focus of a channel or forum may be) MUST have intrinsic interest for each actor. Nobody goes there looking to make friends as a primary goal. (Interviewee #5)

The interviewee thought that if the communities grow too big too fast, they start to fragment:

Some communities have a very high signal::noise ratio. Usually small ones. Once communities grow beyond a certain size, though, it turns to chaos, so you see frag-

mentation. If you go to a JavaScript channel or forum, for example, you are unlikely to find really sound advice there. Noobs and larval-state hackers dramatically outnumber the old gurus who know everything. That is uncomfortable for the gurus. They get worn out, and you get this problem with students coming in every new semester hammering the channel about beginner stuff, and then other students who are about a semester ahead telling them basically wrong answers. The blind following the blind. Because there is no inherent social ranking system outside of super old and insular communities, an old guru will get shouted down by people who don't know what they are talking about. So gurus tend to avoid such communities entirely, making the channel even MORE noisy relative to when they were sticking around. Eventually you wind up with a starvation of good ideas and the community fragments. Most of those people move on somewhere else, grow up, graduate, whatever. They sort of disappear. (Interviewee #5)

And the interviewee also thought that there might exist a sweet spot for the size of a community:

There is a sort of Goldilocks size for an online community. And I don't know if that really is a question of total size, or whether it is a Goldilocks ratio of experts to newcomers. A community that is too small can't aggregate enough useful knowledge without wearing out the one guy who knows what is going on (he'll leave, and his absence spells disaster for the community). A community that is too big is too noisy and troublesome, so once again the experts leave. A community that is a sort of medium sized and starts out heavily populated by long-time members who have all amassed a lot of knowledge can easily metabolize and take care of newcomers. And these communities are often very durable. So that is the "goldilocks size" I mean. And it may really be more of a ratio than a total size. (Interviewee #5)

4.1.1.3 *Boosting reputation*

Boosting ones reputation can be a motivation to some. As in probably any community there are visible and nonvisible social structures present:

I'd like to think I haven't messed up enough times that some might even respect me. I think IRC's like anywhere else, one came make a name for themselves (...) (Interviewee #16)

Having an ego is thought to be a negative thing among the interviewees. One interviewee mentioned that some users stroke their egos by treating new users badly when they don't conform to their standards. Users that have an ego can also be useful for the community, for example when they are motivated to try to be the first to answer questions. Gaining respect within a community can also be useful for the user, as others might be more willing to help users they respect. Two interviewees characterizes having an ego in the following ways:

It's also fun to try to answer a question before anyone else. (...) Answering questions makes you feel smart and puffs up your ego. I didn't learn to be humble early on, so I have a strong natural tendency to be pretty arrogant. It's a bit of a source of pride that I can answer questions in the IRC channels that I frequent. (...) It's bad because pride really isn't a good thing generally. But, it does help to offset "impostor syndrome", which is the tendency for a developer/programmer to feel that they aren't actually very good. It's sort of a reverse-pride in your ability. Having the small pride boosts helps keep you on track that you must be at least somewhat competent, or you wouldn't be able to answer questions. (Interviewee #8)

(...) it'd be a lie to not acknowledge that it also gives my ego a boost, which is a bad reason but still there. also, i suspect it builds my 'street cred'(ibility) such that i'm more likely to get help and explanations from other active users when i want it. (Interviewee #17)

Gaining a good reputation within the community can have many positive side effects, some might even get job offers due to their good reputation in the community:

well, I guess I do conduct business on here also, have connected with a few clients, been offered jobs, etc. it's more a side effect, I get business sometimes through my exposure on here (...) because (a) you sometimes see people asking for work, and (b) I'd meet people who needed work done (Interviewee #9)

4.1.1.4 *Belonging to a community*

Some like to belong to a community, for instance to fight the feeling of geographic or social isolation. Sometimes programmers work in environments where they have no colleagues within the same profession, and therefore are unable to discuss their work on a desired level with anyone in the work community. Some people also work alone for example from their home and they don't get to meet any colleagues, if they even have any. One interviewee expressed thoughts concerning this:

working from home and not being a really social [person]... this is where i get my socialization. (Interviewee #17)

Online communities were thought to reduce the feeling of isolation, be it social or geographical, although this wasn't the case for everyone.

4.1.1.5 *Problem solving ("Nerd sniping")*

Several interviewees mentioned "nerd sniping" as a reason for why they might participate in knowledge sharing. Nerd sniping is a term popularized by Ran-

dall Munroe, the creator of the comic xkcd. The comic can be seen in Figure 9. In short "nerd sniping" means that people who are prone to being "nerd sniped" will stop anything they are doing at the time if they are presented with an interesting problem to solve, and will instead start working on solving the problem. At least some of the interviewees might do the same if an interesting problem is presented. One interviewee told about ones motivations to help others with their problems:

(...) I like solving interesting problems. I'm particularly vulnerable to being "nerd sniped", per that old xkcd comic (<https://xkcd.com/356/>). (Interviewee #8)

Another interviewee mentioned the comic when telling that sometimes the questions include self-contained puzzles to solve:

it's enjoyable to solve some puzzles yes (...) even if they're not intentionally constructed as such (Interviewee #2)

It seems that problem solving is something that strongly motivates some of the users to help others. Problems are not always solved alone, users also collaborate together to solve problems and solving problems with others is enjoyable:

Also like collaboration on a problem with other experts (Interviewee #1)

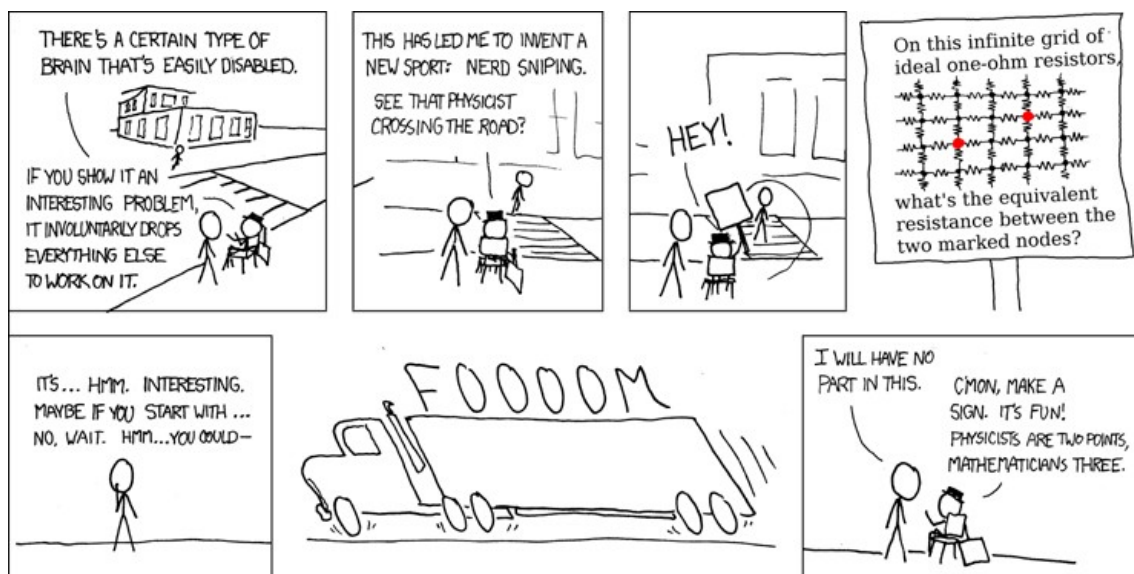


FIGURE 9: Nerd Sniping (Munroe, 2007).

4.1.2 Advantages of IRC to knowledge sharing

In the era of social media, there are many kinds of platforms available for forming online communities that encourage knowledge sharing. IRC is a chat-based medium that doesn't keep logs of the previous discussions. For instance IRC communities were thought to be a good way to keep oneself up to date with technologies and the direction they are heading to. It was mentioned that by following the discussions on IRC you stay ahead of the trends compared to other media. Being a chat-based medium has some advantages and disadvantages, which will be discussed next.

4.1.2.1 *Fast feedback cycle*

IRC was generally considered to be a suitable place for seeking and sharing knowledge. Some interviewees thought that IRC fits well to people that are beginning to learn about a certain programming related topic like a programming language or a software framework. When comparing IRC to Q&A web platforms like StackOverflow some interviewees mentioned that the faster feedback cycle helps newcomers and that asking on IRC is more informal:

IRC has faster feedback cycle and more "immediate" feel, so it's more suitable for people early in their learning process (...) (Interviewee #18)

(...) asking on IRC is usually faster and more informal than e.g. posting on stackoverflow (Interviewee #3)

On the otherhand some thought that Q&A platforms are better suited for finding solutions to specific technical problems, but only if the user is aware of the required keywords. The benefit of IRC is that one can explain the problem in ones own words:

(...) sometimes it's easier to explain your problem than try to find keywords for it (Interviewee #14)

That could be because asking good questions might also require some knowledge about the topic:

when you're starting out with something, it can be hard to get answers because you don't even know enough to ask the right questions (Interviewee #10)

This may be one of the advantages IRC and other chat based communities have, as newcomers can often get informed of the proper terms they should look into to solve their problem.

It was also mentioned that IRC works well for mentoring and teaching others, unlike Q&A web platforms like StackOverflow, which might be better suited for helping someone with only one question. IRC is a socially engaging platform, at least more so than many Q&A platforms:

IRC also allows for more back-and-forth discussion compared to SO where you have comments but it's mostly one question and then answers, ideally without having to clarify things (Interviewee #3)

It is much easier to have conversations with other users, as the response cycle is faster due to the nature of live chat. The chat format might allow people to form social relationships with each other more easily, for example due to the possibility of having in-depth discussions. Dialogue is something that many users seem to enjoy. Several interviewees found this to be an advantage of IRC:

Here you also kind of 'see' the person you are helping, their reactions and struggles, it's more interactive (Interviewee #11)

(...) IRC specifically is a very conversational medium that models real-life interactions pretty well (Interviewee #15)

4.1.2.2 The lack of a reputation point system

Many Q&A platforms use vote-based reputation systems to indicate the quality of questions and answers (Anderson et al., 2012). This can lead to a situation in which the answerer's main motivation might not be helping the user, but instead looking good by answering and gaining reputation points. Reputation point systems can motivate users to give more answers, but there is a risk that some knowledge contributors might begin to focus on the low hanging fruit, the easy questions that don't take too much effort to answer. On IRC however there is no reputation point system. At least one interviewee thought that this leads to the users encouraging others to work through their problems rather than just focusing on gaining reputation points:

People actually encourage others to work through their problems instead of trying to look good by answering (Interviewee #11)

Since there is no reputation point system, the reputation of a user is built by others observing the actions of a user. So each user who wants to gain reputation within the community has to convince each participant, as the community given collective reputation can't be observed by a quick glance from the users points.

4.1.2.3 *Ephemeral messaging*

Many platforms, like Slack and StackOverflow, save messages automatically for them to be viewed later on. This is the case especially on Q&A platforms where the questions and answers are meant to be looked up later on as they form sort of a library of information. This is not generally true on IRC as the servers don't log the messages, although the users can do that themselves if they want to. Essentially though the questions aren't archived for long-term use. This was thought to benefit IRC in multiple ways, for example because the users don't have to worry as much about what they say. This was mentioned by several interviewees like so:

a positive is that people are more free to simply speak and let the conversation flow .. any mistakes or possibly embarrassing things automatically go away after a time (...) perhaps another way to put that is, there's less of a consequence for being wrong, which I think is important for productive learning (Interviewee #15)

We tend to consider IRC as totally ephemeral. People feel more free to express any crazy idea here. Because if you change your mind later, make some controversial statement in a moment of anger, or whatever, it just disappears over time and people move on. If you make a controversial statement in Slack, a mailing list, twitter, etc. with the current level of political insanity in the West, well, you might lose your job for it 5 years later when someone digs it up. So ephemeral messaging is one of the very last places you can still be known by your real identity and say what you actually think. (Interviewee #5)

The lack of logs was also thought to be a negative thing in some respects like for example in repeated questions:

one of the common downsides to lack of records is the frequent questions, there is nowhere for someone new to find the answer because they cannot see that it was already asked yesterday .. or 10 minutes ago (Interviewee #15)

4.1.3 **Reasons for not sharing knowledge**

In programming focused communities it is often expected that the asker or the knowledge seeker puts effort into the questions. This way the knowledge seekers might solve the problems while improving the presentation of them, and at least then it is easier to figure out what is really being asked. It is often expected that at least stacktraces of errors and the relevant code are shared:

(...) IRC (and other technical resources) have a different set of rules from everyday conversation. It generally expected that you (a question asker) puts more effort into getting your question answered than the answerers. (...) (Interviewee #8)

It seems that some active knowledge contributors don't like to help people that come asking for concrete solutions to their problem without making any effort to solve it themselves. A special case seems to be students trying to solve their programming homework. Knowledge contributors might refrain themselves from handing out straight answers to the questions when they notice for example that the one seeking for help is a student:

I might hold back a bit if it were obvious that they were a student and needed to learn on their own (...) (Interviewee #16)

An interviewee mentioned that the active contributors often ask newcomers to show code or other artifacts to indicate that they have attempted to solve the problem themselves. The interviewee stated three reasons for doing this:

giving more information on the problem in question, determining the skill level of the person asking, and warding off "help vampires", who just want others to solve their problem without giving anything back (Interviewee #18)

The so called help vampires were also described by another interviewee as a reason for deciding not to help someone. Another interviewee mentioned that help vampires are users who want to be shown code that they can directly use in their own projects. The interviewee doesn't like helping that kind of users because it feels like doing someone else's work:

(...) it's the kind of user who feel entitled to get everything spoon-fed[.] i'm not there to do someone else's work, which is what these people are often asking for (Interviewee #3)

4.2 How users become active knowledge contributors

The following paragraphs will showcase the stories that were built as a result of the analysis of the interviews. Also the findings concerning RQ2 are presented. First the different stages in the pathways to becoming active knowledge contributor will be introduced, after which each pathway is looked at one by one.

4.2.1 Pathways to becoming active knowledge contributor

The analysis first came up with three different pathways to becoming active knowledge contributor in online communities. The pathways were built based on the analysis of the interviews, and each of the 18 interviewees were placed into one pathway. The process of building these pathways was multi-staged, first a story was constructed for each interviewee individually, after which the stories were intensively compared and then grouped together to form the 3 pathways.

The pathways describe the stages users go through before becoming active knowledge contributors in an online community. It is important to note that when a user moves from one stage to another, it doesn't necessarily mean that the activities of the previous stages stop, it rather means that a new activity starts. All of the three pathways begin with joining the community, which refers to the first time the user used or visited the online community. The initial use of the communities happened for various reasons, for example because the user was interested in IRC, was looking for a community to meet like-minded people, wanted to get help, or just wanted to learn about the topic of the community. One user described the reasons for initial use in the following way:

(...) I was like "Hey, IRC is often populated with technical people, maybe I can get help on there!" (Interviewee #4)

The stage familiarizing with community means that the users were not participating in knowledge sharing related activities, but paid attention to the discussions and tried to learn from them and also to understand what the community was like. The term "lurk" is often used in literature to mean passive participation in a community. Often the users familiarizing themselves with the community were lurking, although some might have socialized with others in order to getting to know them. The lurking users on IRC can be seen participating in the community in the list of user names that are online, but the users don't necessarily take part in knowledge related discussions and prefer to just read them. It could also be that the users might be observing the community in order to figure out whether it is worth it for them to continue participating in it. One interviewee mentioned this as a motivation for lurking in a community:

it was to get a sense of the channel dynamics and whether I would want to participate or not. (Interviewee #7)

The lack of knowledge about the topic of the community, and especially the confidence in ones knowledge might be the underlying factors why the users decide to lurk before participating in knowledge related discussions. Multiple interviewees felt that they learn by lurking and observing the ongoing discus-

sions in the communities, thus gaining more knowledge and confidence. It was also mentioned that lurking in an IRC community of a specific topic is a quick way to check which direction the general community of the topic is heading towards without the hype that is present on other media. Several interviewees mentioned that IRC communities are a good way to keep abreast of new trends:

So the quicker way to check the pulse and direction of the community is, of course, to hop into IRC and lurk. (Interviewee #5)

Some users seem to tend to skip lurking altogether, and immediately begin socializing with others, some users just want to get to know new people:

(...) when I got online everything was smaller and more personal, and I have always been into getting to know new people, so probably I didn't lurk at all (Interviewee #11)

The stage seeking knowledge refers to a period in which the user begins asking questions or presenting problems in the community. Some of the interviewees seem to immediately move to the seeking knowledge stage, and only after that familiarize themselves with the community. These users first come to the communities with questions or problems they hope to get some help with.

In the stage sharing knowledge users begin answering others questions or helping them to find the answers. One interviewee described the reasons for helping others in the following way:

a. I like to share knowledge with people, it helps me better encompass what I do and do not know, and more often than not a question is the starting point of a process of documentation and research about the things I help people with (Interviewee #4)

b. A developer career is based on building a technical expertise, it helps my career that I cultivate the skill of helping people, documenting myself and learning stuff. Reciprocally, getting better at knowing-learning stuff makes me want to help people more, and so on :P (Interviewee #4)

This explains some of the points on why users share knowledge. Users that are sharing knowledge are motivated by learning, just like they are in every other stage. Helping others is also seen by some as a way to validate ones knowledge, as other users have the chance to correct the advice one gives. Generally when users learn, they gain more confidence in their knowledge and thus can help with more challenging questions:

(...) as i learned about various useful features of [programming language] itself, libraries, general programming techniques etc my knowledge and confidence grew and would increase the questions i would answer (Interviewee #17)

Next we will take a closer look at each of the three pathways one by one.

4.2.1.1 Pathway #1

Pathway #1 consists of 4 stages: joining community, familiarizing with community, seeking knowledge, and sharing knowledge. These stages can also be seen in Figure 10. This pathway was constructed based on the interviews of 7 interviewees.

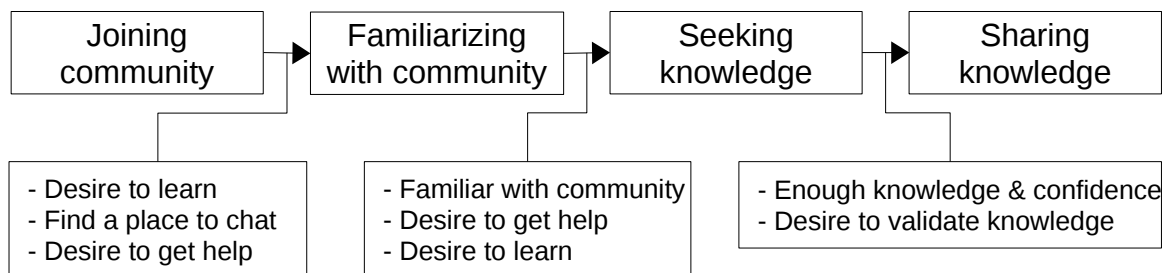


FIGURE 10: Pathway #1

The majority of the interviewees of this pathway were initially motivated to joining the community by learning. One interviewee was specifically looking for a place to chat, and two interviewees needed help. After their initial use of joining the community, they began to observe the discussions and familiarized themselves with the community. There were several reason for this, the most mentioned reason was the lack of knowledge and the lack of confidence in ones knowledge. Language barrier was also mentioned as a reason once:

initially i wasn't using [channel] that much, so was just trying to get a feel of it. i wasn't very used to english back then too. (Interviewee #13)

Other reasons mentioned were getting to know the people and figuring out the social hierarchy, or as one interviewee put it "who not to annoy". After lurking for a while, these interviewees started asking questions, and they also asked help with their problems. Asking for help was motivated by learning as well, which was specifically mentioned by four interviewees. When moving to the sharing knowledge stage, the interviewees had learned enough so that they were able to answer questions or help in other ways like showing other users where to look at for answers. At this point they had gained enough knowledge to help with what they considered to be easy questions. One interviewee started helping others to validate knowledge. The interviewee felt that when providing help or answers to others, someone would let the interviewee know if the help or answers they had provided were incorrect, and the interviewee considered that to be an important learning opportunity. Once the interviewees had learned even more they also gained even more confidence in their knowledge.

This led them to becoming active knowledge contributors within the community.

Several interviewees also thought that helping others was fun, which probably contributed to them becoming active knowledge contributors. One interviewee mentioned to be getting socialization and finding the community useful as the reason for becoming active participant. The reasons for continuing being an active participant in the community were somewhat different to each interviewee. At least three interviewees were motivated by advancing the community, for which there are several different kind of ways. For example one interviewee was motivated by encouraging other users to begin helping, which is important for the healthiness of the community. Knowledge sharing communities need knowledge contributors in order to be lively. Boosting ones ego was also mentioned by one interviewee.

4.2.1.2 Pathway #2

The pathway #2 has 4 stages: joining the community, seeking knowledge, familiarizing with the community, and sharing knowledge. See Figure 11 for a visual representation of the pathway. A total of 7 interviewees' stories were the basis for this pathway.

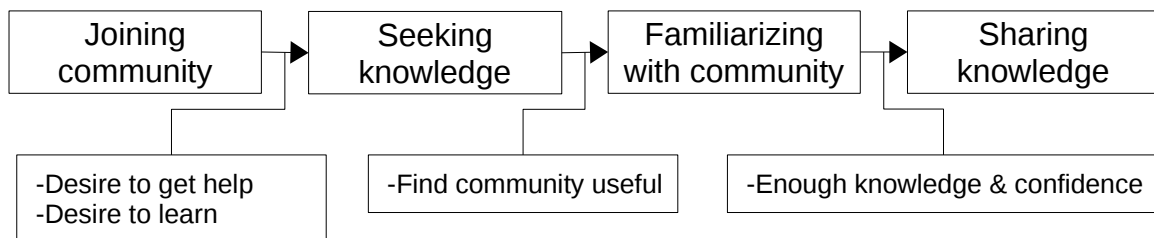


FIGURE 11: Pathway #2

The initial motivations for using the community were learning and the idea that the interviewee might get help or advice, so they were seeking knowledge. Just trying out IRC was also a reason for the initial use. After finding the community useful, these interviewees stuck around, asked questions and familiarized themselves with the community. Once they got more familiar with the community, some interviewees started to ask questions even more actively. Some were afraid to start helping others because they thought that only experts could do something like that, so there certainly exists a threshold that is blocking some of the interviewees from helping others. The interviewees began sharing their knowledge once they had learned more and had enough confidence in their abilities to do so, and one could say that perhaps some of them felt that they were becoming experts in the field. After starting to help others by sharing knowledge, some quickly found that one can also learn by helping others:

I fairly quickly realized that answering questions was as much of a learning opportunity as asking questions, so I started trying to answer any questions I could. (Interviewee #8)

Realizing that helping others is a learning opportunity made some of the interviewees even more active helpers in the community. Continuing to help others was motivated by advancing the community, the interviewees felt that they belonged to the community and were members of it and that their actions could make a difference to the community. One interviewee cared deeply about the experience of the new users, and tried to make the experience pleasant for the new users in hopes of encouraging them to continue to participate in the community. Helping other users to learn was something most of the interviewees enjoy:

I'm also very interested in specifically helping others learn. If I had a dream job in this or a similar field, it would probably be teaching this stuff and mentoring. (Interviewee #8)

Also the feedback received when helping someone motivated some of the interviewees to continue to do so.

The main difference of this model when compared to other stories is that the familiarizing with the community stage is not directly after the joining stage. This is because the interviewees initially joined the community to ask questions or to ask for help with their programming problems. They thought that they could fulfil their initial needs by visiting the community, and once they had (or had not) received what they came for, some of the parted the community, and returned after they needed help with something again. They didn't feel being part of the community initially, and viewed it as a useful knowledge resource they could utilize whenever they needed to.

4.2.1.3 Pathway #3

The pathway #3 consists of only 3 stages: joining the community, familiarizing with the community, and sharing knowledge. These steps can be seen in Figure 12. The pathway is based on the stories of 4 interviewees.

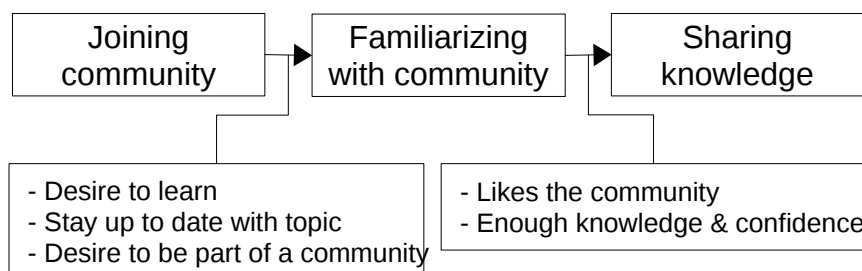


FIGURE 12: Pathway #3

The interviewees that belong to this pathway were initially motivated to visit the community to learn more about the topic of the community, but some also wanted to familiarize themselves with the community of a specific topic. Initial use of the community was also motivated by staying up to date with the topic, and the feeling of belonging to a community. After joining the community the interviewees started to familiarize themselves with the community. While familiarizing with the community, they were also learning about the topic by observing the community discussions, questions, and answers. One interviewee wanted to observe the channel dynamics in order to find out whether the interviewee liked the community or not, and if it was worth participating in. The interviewees wanted to get to know the people, social hierarchies, guidelines and so on. Some did this by observing others, while some started to socialize with the community participants. After the interviewees had familiarized themselves with the community, they began helping others. They started sharing their knowledge mainly because they felt that they had learned enough from others questions while observing and possibly by studying the topic from other resources. The interviewees started by helping other users that had easy questions or problems for which they were certain they had the answers to. Once they had gained enough knowledge and more confidence in their skills to help others, they began helping others even more actively.

One interviewee learned from the questions the interviewee was unable to answer. Continuing to share knowledge actively was motivated by learning, some interviewees thought that they learn by helping others with their problems. Overall it seems that helping others is something the interviewees in this pathway enjoy. Other motivations for continuing to share knowledge actively are to advance the community. Some interviewees even mentioned that in order to improve their teaching and communication skills they began experimenting with communication and teaching styles. This pathway differs from other stories in that the knowledge seeking stage is missing altogether. One reason for this could be that the interviewees in this pathway were all very experienced in programming or in the topic of the community.

4.2.2 Process model: how users become active knowledge contributors in online communities

Based on the three pathways it was possible to move up to the *fabula* level in the levels of structure in narrative (as shown in Figure 8). This means having one story that best describes the overall story of this study. The process model, which can be seen in Figure 13, is directly based on the *fabula*, which was the most common story among the interviewees.

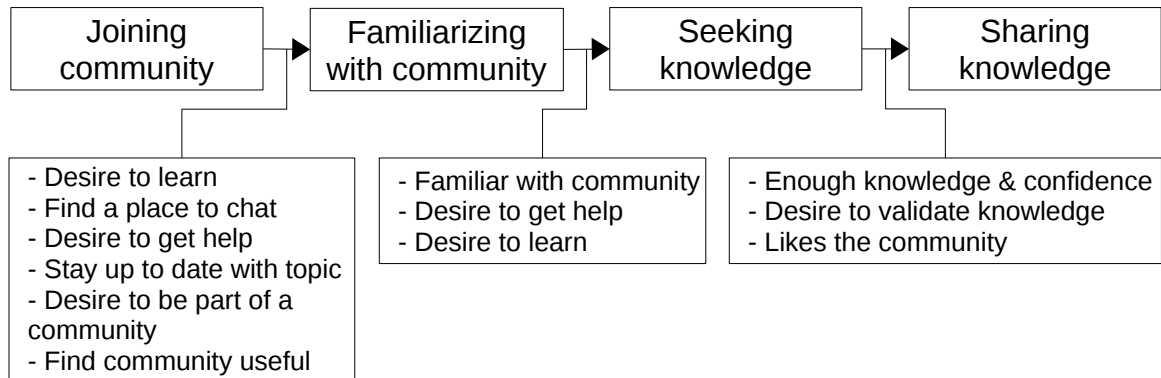


FIGURE 13: Process model: how users become knowledge contributors

The stages in the process model are exactly the same as in the pathway #3 (Figure 10), the forces affecting the moving to the next stage are quite similar too with some additions from the two other pathways. The process model shows how users become active knowledge contributors in programming focused online communities after first joining them. Common motivations for initial use of a community are learning and the desire to belong to a community. The users are interested in learning or staying up to date with a certain technology, and they find online communities to be a suitable place for that.

Users are motivated to familiarize themselves with the community because they want to learn, find a place to chat, want to be able to get help, desire belonging to a community, and when they find the community useful. Familiarizing with a community means observing the social structures and the community guidelines, norms, and acceptable behaviour, be them written on unwritten, and socializing with the users of the community in order to get to know them. It seems that familiarizing with the community is something that is a requirement for the users in the process of becoming active knowledge contributors in online communities, users don't seem to just join a community and begin sharing knowledge before being familiar with the community to some extent.

When moving to the stage of knowledge seeking, the users are generally familiar with the community, they want to get help, and most importantly they want to learn more about the topic of the community. Seeking knowledge is often done by asking questions or presenting problems that the user wants to get solved but is unable to do so alone. Sometimes the users just want to know the more experienced participants opinion about a something.

There are many obstacles in the way of users starting to share their knowledge. First of all if the user doesn't like the community or share its values, it is unlikely that the user begins sharing knowledge. The biggest obstacle is the user's own perception of not having enough knowledge to help others. One interviewee put it like this:

I viewed the channel as something on which only experts could help people, people with so much knowledge a lowly peasant like me couldn't start to grasp (I was in my 1st year of experience, mind you) (Interviewee #4)

Each individual has a different view on how much they know, and how much they would need to know in order to begin helping others. This probably also differs between communities. Anyway, the confidence in ones knowledge is a major barrier in the way of beginning to share knowledge in online communities. But once a user begins sharing knowledge, the user most likely also gains more confidence in ones knowledge:

the more I was there, the more confident I became in my answers and moved up the chain, as more newer people would help with simpler stuff. (Interviewee #9)

Learning how to explain something is thought to be a good way to become competent with it:

learning to teach something is the best way to get good at it (Interviewee #6)

Some users view helping others as a chance of validating their knowledge, as when they help others and if they make mistakes, someone might point out the mistakes. That way the user who had potentially false information can correct that and learn from it:

even in the beginning i was able to do the basics in [a programming language] having worked with it some before and just generally having lots of programming background so i could provide answers to some things. and when i got them wrong someone would jump in and correct me and i'd still learn. (Interviewee #17)

This might be something that happens naturally in some communities:

there seems to be some function of time/skill, where the higher experienced [users] sit back and let more junior people try to help first and step in when (a) wrong / poor answers are surfacing, or (b) it's turning out too difficult. this helps people on all levels to learn (...) (Interviewee #9)

5 DISCUSSION

The goal of the study was to find out why users share knowledge and how users become active knowledge contributors in online communities. The study was conducted on programming focused online communities on the Freenode IRC network. The research questions were:

- **RQ1:** *Why do users share knowledge on Freenode's programming focused communities?*
- **RQ2:** *How do users of Freenode's programming focused communities become active knowledge contributors?*

The next paragraph will discuss why users share knowledge and introduce the process theory of how users become active contributors in online communities. Also the results of the study will be compared to existing literature. Finally the strengths and weaknesses of the study will be analyzed, and the implications for future research will be discussed with suggestions of topics for further study.

5.1 Discussion of the empirical results

Most of the interviewees were initially motivated to join a community by the opportunity to learn. Throughout the stories it became evident that learning is a big motivator for the majority of the interviewees in most of the stages. Learning was a motivation even to contributing knowledge by helping others. Research has found previously that learning is a motivation for participation in online knowledge sharing communities. For example Gray (2004) found that the opportunity to learn motivated users to participate in online communities of practice. Learning is also something that motivates users to start sharing knowl-

edge. According to Jin et al. (2015) when a user learns from others it is likely that the users might begin contributing knowledge, so it seems that social learning opportunities motivate knowledge contributions. Continuing knowledge contributions seems to be motivated by learning as well, and some of the interviewees even stated that they learn more by helping others. This observation finds support from previous literature too, as Wang and Noe (2010) note that knowledge sharing may also be thought of as a learning opportunity, for example because it provides a chance to deepen ones understanding so that it would be easier to explain it to others. Several interviewees mentioned that learning how to teach and explain things is an important skill for a programmer, some even thought that teaching and mentoring might be their dream job. One way to learn while contributing knowledge was the possibility to validate ones knowledge. When sharing knowledge to help others with their problems, others users in the community can correct the possible mistakes one makes. Validating knowledge by sharing knowledge seems to be something not widely recognized in other studies.

For some interviewees the motivation for joining the community was to get help solving a technical problem. They were initially mostly interested in their own problem and didn't care too much about the community. But once they got help solving the problems they might have found the community useful, which then made them to stay around and familiarize themselves with the community while occasionally asking for help. Other motivations for initially joining the community were the desire to belong to a community, finding a place to socialize, wanting to stay up to date with a technology, and just trying out IRC. For some the feeling of social or geographic isolation was a reason for participating in the programming communities. Many professionals seek belonging to a group of people sharing the profession. When one works in an environment having no one sharing the profession, online communities provide a way to reduce the feeling of isolation. Without colleagues of similar profession there might be no one to discuss with in order to figure out solutions to difficult problems, leaving the individual to deal with the problems alone. Gray (2004) found an initial motivation for participating in online communities of practice was offsetting the feeling of isolation in a work environment, when working in an occupation not shared by anyone in the work place or working in geographic isolation.

In the familiarizing with community stage the users mostly observed the discussions and socialized with the community. Some did not participate at first by seeking knowledge directly by asking because of a language barrier, because they were not confident in their English skills at the time. In the familiarizing stage lurking was a way to get to know the community, so that the users could decide if they want to participate in the community or not. Observing the community discussions was also found to be helpful in figuring out the direction of a technology, and keeping up with the upcoming trends without much of the

hype. Some users seem to skip the familiarizing stage at first. It seems that these users just want to get help solving their problems, and then only later if they find the community useful they might move to the familiarizing with the community stage.

Users move to the seeking knowledge stage and begin asking for help because they want to learn more about the topic, or they want to solve a problem. Some users might stay longer in this stage as they are afraid to help, because they think that only experts could do that, and that they lack the knowledge to help others. Once these users realize that they can solve problems alone, they gain confidence and move to the sharing knowledge stage and start helping others.

Sharing knowledge seems to be something that requires confidence in one's knowledge, which based on this study seems to be a major threshold in the way to becoming a knowledge contributor. This finding is supported by previous research. According to Jin et al. (2015) knowledge contribution requires experience and knowledge of the domain area. Chiu et al. (2006) found that users' confidence plays a role when contributions are voluntary, and when the users doubt their knowledge sharing abilities they will not be likely to contribute. Kankanhalli et al. (2005) found that knowledge self-efficacy has a considerable effect on the use of knowledge repositories (Kankanhalli et al., 2005). Self-efficacy is what one believes to be able to accomplish with the abilities one has (Bandura, 1986). The interviewees started sharing knowledge when they felt that they have learned enough, and thus had gained confidence in their knowledge. According to Wenger et al. (2002, 14-18) communities of practice enable their members to gain confidence. The same is probably true for online communities as well, and as users learn more about the topic of the community by asking questions and observing discussions, they also gain confidence in their abilities, which at some point might lead them to becoming knowledge contributors.

Users become active knowledge contributors when they learn and gain even more confidence, some might feel that they are becoming experts on the topic. The users continue to help actively because they want to advance community in several ways, like encouraging others to start helping, and caring for the experience of new users. Many users enjoy helping and they don't like seeing others flounder, which has also been identified in previous literature (e.g. Kankanhalli et al., 2005). The feedback the knowledge contributors receive seems to further motivate them to continue doing so. Wasko and Faraj (2005) found that knowledge contributions in online communities are strongly motivated by building reputation. Jin et al. (2015) also found that other members' recognition positively impacts knowledge contribution. The findings of this study support this idea. Even though IRC doesn't have a concrete reputation system (for example a reputation point system), a reputation system still exists as a social construct within the community and in the minds of the contributors.

Building reputation was described as a motivation for knowledge sharing by several interviewees. The lack of a reputation point system was seen as an advantage of IRC, due to the lack of a reputation point system the knowledge contributors concentrate more on helping others learn, and the focus is less on gaining more reputation points.

There is a certain type of anonymity present on IRC and the messaging is ephemeral in nature, meaning that the messages aren't recorded (Rheingold, 2000, 154). This was seen as a big benefit of IRC, and IRC was thought to be one of the last places where people can freely express their ideas. Because the messages aren't stored and the users can remain anonymous if they wish to, the users might not have to worry that much about what they really think. The same might not be true to Q&A platforms for example, and some users might feel the need to be more careful there, as it could be possible that the things they say there might come back haunting them later. This could severely limit the kind of knowledge shared on those platforms, and especially the creation of new knowledge. Sometimes even the craziest ideas could potentially lead to new solutions.

Communities can vary in the pace they change, and in the case of technologies it could be related to the pace certain technologies move forward. For technologies that change with a faster pace, also the people in the community might come and go quicker pace as well. It was mentioned that mature technologies which don't evolve very quickly, have communities that don't experience rapid changes in the userbase. The userbase often stays the same for years, meaning that the community users don't easily abandon the community.

5.2 Strengths and weaknesses of the research

One of the strengths of the study is the wide range of interviewees of multiple different nationalities and backgrounds, and also from a wide variety of communities, which made the collected data diverse. With so many different backgrounds the results are more generalizable than for example in a case of studying only students from one country. However, because not every continent was not present, for example Africa, Asia, and South America, many non-western cultures were left out of the data as the interviewees were mainly from Western cultures. Because of this the results probably represent only the users that are from Western cultures, as cultures might be something that have a major effect on knowledge sharing. Also as the case of the study was freenode IRC networks programming focused communities, the results might not be generalizable to all kinds of communities. As this was a qualitative study, there might be some bias present from the researcher. The interviews were semi-structured, which made each interview a little bit of different. Some interviewees were able to pro-

vide more insight than others, and some things might have been missed by the researcher and not asked about in detail. Although this was strengthened due to some interviewees being able to provide loads of insights and make deeply detailed connections between their actions and motivations, which could then be noticed by the interviewer in later interviews.

5.3 Implications for theory and practice

The results of this study might contradict those of Nov et al. (2010) who found that those motivated by self-development (learning) were contributing less information artifacts. This study concludes that learning also motivates those that share knowledge, but it is unclear whether the learning motivations affect the amount of sharing that is happening. Nov et al. (2010) suggested that this could be because those who want to learn only share their best work, which in the context of this study could mean that those motivated by learning might only share when they are very certain they know the right answer. However it was found that some users like to validate their knowledge by contributing knowledge, so that if their knowledge was incorrect it could be corrected by other users. There is a slight contradiction here and perhaps it should be looked into in more detail in future studies.

This study identified confidence in ones knowledge as a major threshold for knowledge sharing. Community maintainers need to find ways to boost the confidence of its users in order to encourage them to begin to share knowledge. Based on the interviews it seems that some communities on Freenode seem to have unwritten rules that guide senior users to give a chance to junior level participants to help others with easy questions, and that the more senior users would jump in when it seems that the junior is struggling. Perhaps other communities should consider trying to create a similar atmosphere that would give its users a chance to help others with easy questions. It would allow the users to gain more confidence in their knowledge, which in turn might result in that these users would begin to share even more knowledge.

5.4 Topics for further study

As this study concentrated on the active knowledge contributors of Freenodes programming focused online communities, it might be interesting and worthwhile to also look into the users that once were active knowledge contributors, but for some reason stopped doing that. Finding the reasons for stopping participation, either participation in the communities altogether or just stopping

knowledge contribution, could allow finding ways encourage these users to continue their knowledge contributions. Though one obstacle in the way of such study is trying to find these users that have stopped using the communities.

Another topic that emerged from the interview data was ephemeral messaging, which means that the messages are not stored for later use by the servers. One interviewee mentioned that IRC is one of the last places where this is possible. The effects of this to knowledge sharing should be studied in further detail. This paired with the findings of Young et al. (2012), that surveillance can restrict knowledge sharing, would be an important subject to study in more detail. Could ephemeral messaging positively affect knowledge sharing intentions?

It might also be interesting to interview non-Western people and find out if their paths to active knowledge contributors in non-Western online communities are different. How might the culture affect the process, and might there be differences between cultures when it comes to knowledge sharing in online communities. Some of these topics might be worth studying further.

Research needs to be done on what happens after users become active knowledge contributors. The data collected for this study indicated that there might be mechanisms that further affect how these knowledge contributors' behavior changes even after becoming active contributor. One such thing is for example giving more room to more junior people to help and learn to help others while still observing and jumping in when needed.

6 CONCLUSION

The aim of this study was to find out why users share knowledge, and how users become active knowledge contributors in online communities. To answer these questions a qualitative study was conducted with the objective of building a process theory by identifying the stages from the initial use until the user became a knowledge contributor, and explaining the reasons for why users move from one stage to another. The case of the study was Freenode IRC networks programming focused communities, from which active knowledge contributors were identified and then interviewed using semi-structured interviews.

After initially joining a community, the users familiarize themselves with the community by observing the discussions or by socializing with others. Then the users begin seeking knowledge by asking questions and presenting problems they encounter, while also learning by observing others problems and the suggestions for solving them. When the users have gained enough knowledge and confidence in their knowledge, they begin contributing knowledge into the community by helping users first with easy problems, while gradually moving to answer harder questions once they gain even more knowledge and confidence. The findings in the study indicate that confidence in ones knowledge is a major treshold in the way to becoming a knowledge contributor.

Most of the users are motivated to participate in the online knowledge sharing communities by learning. Learning as a motivator applies to all of the stages users go through on their way to knowledge contributors. This finding has implications to theory by suggesting that knowledge sharing is motivated by learning, something that the previous studies have not widely recognized. Also validating knowledge by contributing knowledge was found to be a motivation not widely recognized in previous studies.

REFERENCES

- Alavi, M., & Leidner, D. (2001). Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues. *MIS Quarterly*, 25(1), 107-136.
- Anderson, A., Huttenlocher, D., Kleinberg, J., & Leskovec, J. (2012, August). Discovering value from community activity on focused question answering sites: a case study of stack overflow. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 850-858). ACM.
- Ardichvili, A., Page, V., & Wentling, T. (2003). Motivation and barriers to participation in virtual knowledge-sharing communities of practice. *Journal of knowledge management*, 7(1), 64-77.
- Bandura, A. (1986). Social foundations of thought and action. *Englewood Cliffs, NJ*, 1986.
- Bock, G. W., Zmud, R. W., Kim, Y. G., & Lee, J. N. (2005). Behavioral intention formation in knowledge sharing: Examining the roles of extrinsic motivators, social-psychological forces, and organizational climate. *MIS Quarterly* 29(1), 87-111.
- Cabrera, A., & Cabrera, E. F. (2002). Knowledge-sharing dilemmas. *Organization studies*, 23(5), 687-710.
- Carron-Arthur, B., Ali, K., Cunningham, J. A., & Griffiths, K. M. (2015). From help-seekers to influential users: a systematic review of participation styles in online health communities. *Journal of medical Internet research*, 17(12).
- Charband, Y., & Navimipour, N. J. (2016). Online knowledge sharing mechanisms: a systematic review of the state of the art literature and recommendations for future research. *Information Systems Frontiers*, 18(6), 1131-1151.
- Chiu, C. M., Hsu, M. H., & Wang, E. T. (2006). Understanding knowledge sharing in virtual communities: An integration of social capital and social cognitive theories. *Decision support systems*, 42(3), 1872-1888.
- Corbin, J., & Morse, J. M. (2003). The unstructured interactive interview: Issues of reciprocity and risks when dealing with sensitive topics. *Qualitative inquiry*, 9(3), 335-354.
- Daniel, J. (2012). *Sampling essentials: Practical guidelines for making sampling choices*. Thousand Oaks, CA: SAGE Publications Ltd doi: 10.4135/9781452272047
- Faraj, S., Jarvenpaa, S. L., & Majchrzak, A. (2011). Knowledge collaboration in online communities. *Organization science*, 22(5), 1224-1239.
- Faraj, S., Kudaravalli, S., & Wasko, M. (2015). Leading collaboration in online communities. *MIS Quarterly*, 39(2), 393-411.

- Freenode. (2018a, November 24th). Finding Channels. Retrieved on 18.1.2018 from <https://freenode.net/kb/answer/findingchannels>
- Freenode. (2018b, January 18th). Channel Guidelines. Retrieved on 18.1.2018 from <https://freenode.net/changuide>
- Gray, B. (2004). Informal learning in an online community of practice. *Journal of Distance Education*, 19(1), 20.
- Hartwick, J., & Barki, H. (1994). Explaining the role of user participation in information system use. *Management science*, 40(4), 440-465.
- Hendriks, P. (1999). Why share knowledge? The influence of ICT on the motivation for knowledge sharing. *Knowledge and process management*, 6(2), 91.
- Hoffman, P., & Harris, S. (2006). *The Tao of IETF-A Novice's Guide to the Internet Engineering Task Force* (No. RFC 4677)
- Jin, J., Li, Y., Zhong, X., & Zhai, L. (2015). Why users contribute knowledge to online communities: An empirical study of an online social Q&A community. *Information & management*, 52(7), 840-849.
- Joyce, E., & Kraut, R. E. (2006). Predicting continued participation in newsgroups. *Journal of Computer-Mediated Communication*, 11(3), 723-747.
- Kankanhalli, A., Tan, B. C., & Wei, K. K. (2005). Contributing knowledge to electronic knowledge repositories: an empirical investigation. *MIS quarterly*, 29(1), 113-143.
- Kogut, B., & Zander, U. (1992). Knowledge of the firm, combinative capabilities, and the replication of technology. *Organization science*, 3(3), 383-397.
- Lai, H. M., & Chen, T. T. (2014). Knowledge sharing in interest online communities: A comparison of posters and lurkers. *Computers in Human Behavior*, 35, 295-306.
- Lampe, C., Wash, R., Velasquez, A., & Ozkaya, E. (2010, April). Motivations to participate in online communities. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 1927-1936). ACM.
- McDermott, R. (1999). Why information technology inspired but cannot deliver knowledge management. *California management review*, 41(4), 103-117.
- Malinen, S. (2015). Understanding user participation in online communities: A systematic literature review of empirical studies. *Computers in Human Behavior*, 46, 228-238.
- Munroe, R. (2007, December 12th). Nerd Sniping. Retrieved on 1.5.2018 from <https://xkcd.com/356/>
- Myers, M. D., & Newman, M. (2007). The qualitative interview in IS research: Examining the craft. *Information and organization*, 17(1), 2-26.
- National Advisory Board on Research Ethics. (2009). Ethical principles of research in the humanities and social and behavioural sciences and proposals for ethical review.
- Netsplit.de. (2018, August 28th). IRC Network *freenode*. Retrieved on 28.8.2018 from <http://irc.netsplit.de/networks/statistics.php?net=freenode>

- Nonaka, I., & Konno, N. (1998). The concept of "ba": Building a foundation for knowledge creation. *California management review*, 40(3), 40-54.
- Nonaka, I., Toyama, R., & Konno, N. (2000). SECI, Ba and leadership: a unified model of dynamic knowledge creation. *Long range planning*, 33(1), 5-34.
- Nov, O., Naaman, M., & Ye, C. (2010). Analysis of participation in an online photo-sharing community: A multidimensional perspective. *Journal of the Association for Information Science and Technology*, 61(3), 555-566.
- Oikarinen, J., & Reed, D. (1993). *Internet relay chat protocol* (No. RFC 1459).
- Pentland, B. T. (1999). Building process theory with narrative: From description to explanation. *Academy of management Review*, 24(4), 711-724.
- Pingdom. (2012, April 24th). IRC is dead, long live IRC. Retrieved on 28.8.2018 from <https://royal.pingdom.com/2012/04/24/irc-is-dead-long-live-irc/>
- Porter, C. E. (2004). A typology of virtual communities: A multi-disciplinary foundation for future research. *Journal of computer-mediated communication*, 10(1), JCMC1011.
- Rheingold, H. (2000). *The virtual community: Homesteading on the electronic frontier*. MIT press.
- Ridings, C., & Wasko, M. (2010). Online discussion group sustainability: Investigating the interplay between structural dynamics and social dynamics over time. *Journal of the Association for Information Systems*, 11(2), 95.
- Saaranen-Kauppinen, A., & Puusniekka, A. 2006. KvaliMOTV - Menetelmäopetuksen tietovaranto. Tampere: Yhteiskuntatieteellinen tietoarasto. Retrieved on 14.08.2018 from <http://www.fsd.uta.fi/menetelmaopetus/>
- Savolainen, R. (2017). Information sharing and knowledge sharing as communicative activities. *Information Research*, 22(3), paper 767.
- Schwarz, A., Chin, W. W., Hirschheim, R., & Schwarz, C. (2014). Toward a process-based view of information technology acceptance. *Journal of Information Technology*, 29(1), 73-96.
- Siggelkow, N. (2007). Persuasion with case studies. *Academy of management journal*, 50(1), 20-24.
- Soliman, W., & Tuunainen, V. K. (2015). Understanding continued use of crowdsourcing systems: an interpretive study. *Journal of theoretical and applied electronic commerce research*, 10(1), 1-18.
- Suhonen, E., Lampinen, A., Cheshire, C., & Antin, J. (2010, November). Everyday favors: a case study of a local online gift exchange system. In *Proceedings of the 16th ACM international conference on Supporting group work* (pp. 11-20). ACM.
- Templier, M., & Paré, G. (2015). A Framework for Guiding and Evaluating Literature Reviews. *Communications of the Association for Information Systems*, 37.

- Treude, C., Barzilay, O., & Storey, M. A. (2011, May). How do programmers ask and answer questions on the web?: Nier track. In *Software Engineering (ICSE), 2011 33rd International Conference on* (pp. 804-807). IEEE.
- Tuomi, J., & Sarajarvi, A. (2002). *Laadullinen tutkimus ja sisällönanalyysi*. Helsinki: Tammi.
- Yan, Y., & Davison, R. M. (2013). Exploring behavioral transfer from knowledge seeking to knowledge contributing: the mediating role of intrinsic motivation. *Journal of the American Society for Information Science and Technology*, 64(6), 1144-1157.
- Young, M. L., Kuo, F. Y., & Myers, M. D. (2012). To share or not to share: a critical research perspective on knowledge management systems. *European Journal of Information Systems*, 21(5), 496-511.
- Wang, S., & Noe, R. A. (2010). Knowledge sharing: A review and directions for future research. *Human Resource Management Review*, 20(2), 115-131.
- Wasko, M. M., & Faraj, S. (2000). "It is what one does": why people participate and help others in electronic communities of practice. *The Journal of Strategic Information Systems*, 9(2), 155-173.
- Wasko, M. M., & Faraj, S. (2005). Why should I share? Examining social capital and knowledge contribution in electronic networks of practice. *MIS quarterly*, 35-57.
- Wenger, E., McDermott, R. A. & Snyder, W. M. (2002). *Cultivating communities of practice: A guide to managing knowledge*. Boston, Mass: Harvard Business School Press.
- Wellman, B., Haase, A. Q., Witte, J., & Hampton, K. (2001). Does the Internet increase, decrease, or supplement social capital? Social networks, participation, and community commitment. *American behavioral scientist*, 45(3), 436-455.
- Wilson, T. D. (2010). Information sharing: an exploration of the literature and some propositions. *Information Research* 15(4), n4.
- Wohn, D., Velasquez, A., Bjornrud, T., & Lampe, C. (2012, May). Habit as an explanation of participation in an online peer-production community. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2905-2914). ACM.

APPENDIX 1 INVITATION LETTER

Hello, sorry for bothering you via pm. I would like to interview you for my masters thesis, the topic is around knowledge sharing in online communities in the context of IRC. Would you be willing to take part? The interview can be held on IRC and would take somewhere between 45min-60min, would you have time on [week]?

APPENDIX 2 BACKGROUND QUESTIONS FOR INTERVIEWS

Before we begin, I would like you to answer to the following background questions:

age, gender, country of origin, profession, work experience (in years), education

Which freenode irc communities do you participate in (name 1-5)?

How often do you visit those channels (daily, weekly)?

How many hours per day/week do you spend on freenode?

APPENDIX 3 SOME OF THE RECURRING QUESTIONS IN INTERVIEWS

- When did you first come to freenode?
- Why did you (initially) visit freenode for the first time?
 - What motivated you to do so?
- How did you use freenode back then?
- Before starting to share, did you use freenode to get information?
- Do you ask questions yourself, how often? If not, why not?
- What motivates you to contribute knowledge on freenode?
- Does feedback received (or not received) affect your desire to continue sharing?
- Why do you use freenode for these purposes?
- Do you use any other platforms for the same purposes?
- How has your use evolved since you started using freenode to this date?
- How has your participation in freenode changed throughout time?
- Have you ever stopped participating on freenode, why was that?
 - What made you come back?
- Can you identify stages in which you used freenode for different reasons or purposes?
- What motivated you in you each stage? Why did you move from one stage to the other?