The role of microblogging in a distributed software development

MASTER THESIS

Ilona Kuzmickaja
2012

Supervisor: Prof. Pekka Abrahamsson
Co-supervisor: Dr. Wang Xiaofeng
ABSTRACT

Communication is one of the key elements of successful distributed software development. Exchange of knowledge among remotely located software development teams must overcome barriers in order to provide interactions comparable to co-located operations. Previous studies showed that synchronous and rich media (e.g., telephone) is used, when immediate feedback is needed. Asynchronous and lean media (e.g., email) is preferred, when timing is not essential. Additionally, frequent communication and trust between team members are highlighted as a key to success. As communities grow more connected through their use of social media, the distribution of knowledge within a community can become complete and misunderstandings can be minimized. Previous studies presented six types of social media, among which microblogging has drawn increasing academic attention recently. Flexibility of access and lightweight architecture of the medium has the potential to be used in informal, collaborative contexts. In many projects, users share their experiences, but it is not clear how this information could be used to benefit development activities. In recent years, several microblogging studies have been conducted as genre analysis and investigated, user intentions, design and characteristics in different fields such as marketing, academic, and enterprise. Existing microblogging studies in software development are mainly focused on two aspects. One is microbloggers’ motivations; the other focuses on the usage of media. Obtained results showed that the development team members use Twitter (one of the most popular microblogging services) for gathering useful information for their work and life interests. Microblogging provides a new informal communication channel complementary to other media like Instant Messaging, email, face-to-face, and other social tools. Also, it helps to generate common ground to support future interactions that form connectedness among development team members. However, these studies usually used case studies based on co-located companies or individuals from different organizations. They did not analyze the use of microblogging as a communication channel in distributed software development.

This research fills the observed knowledge gap and explores how microblogging is utilized by a distributed software development community for its communication needs. The open source software development model is distributed by nature, and various open source software development projects are regarded as success stories. In this thesis, Drupal community was selected as the case study subject. It provides open source content management system created by 17,000 developers in more than 200 countries. Twitter was utilized extensively in the Drupal community as a microblogging service, therefore, the focus of the study. This study analyzes media aspects and how they can affect communication performance. We demonstrate what data types are transferred between community members, what topics appear in their discussions, etc..

Based on the literature review, a theoretical framework of microblogging layers utilized in distributed software development communication is presented. The framework consists of six layers: information processing, communication process, media capabilities, technology features, interaction types and social aspects. Each layer presents a set of corresponding communication media parameters. To validate the framework semi-structured interviews were conducted and unobtrusive measures were collected. Six community contributors from different fields (such as core developer, developer, tester, document writer) were interviewed. Tweets were collected from the interviewees, two most influential core developers and community group accounts. Quantitative and qualitative data analysis methods were applied to perform data analysis. As a result, the framework was updated based on Drupal community practices. The framework introduces the utilization of microblogging in a distributed software development environment and offers both theoretical and empirical implications.

Based on the sixteen primary empirical conclusions, the theoretical framework provides a sound understanding that suggests microblogging be an effective communication media tool in conjunction with richer and synchronous tools such as video conferencing or IM. Microblogging increases users’ availability, mutual understanding and facilitates broad information sharing. Even thought in large organizations, microblogging can cause information misusing, there are facilitated ways such as a group and list creation to manage data flow. The conducted tweet genre analysis revealed three new types of interaction types. The findings of this study propose to evaluate collected spontaneous tweets, which reflect current statuses, code tips, feature description as a new way of documentation in software development. Secondly, they promote feature to find people and start conversation using microblogging. The third finding is the media ability to collect help requests, issue solutions, questions and answers, which stimulates the creation of user support channel. Finally, the presence of social communication in parallel with work related topics in microblogging build connection between unfamiliar or weakly familiar people.

Keywords: distributed software development, microblogging, communication, media
ACKNOWLEDGEMENTS

I would like to thank to my supervisors, Prof. Pekka Abrahamsson and Dr. Wang Xiaofeng, for their engagement, advice, and support. I am grateful to Prof. Pekka Abrahamsson for his scientific advice, knowledge, many insightful discussions, and suggestions. He was instrumental in helping me learn a lot of useful information needed to perform data collection and analysis of this thesis. I am also extremely grateful to Dr. Wang Xiaofeng, she has been supportive, provided insightful discussions about the research, helped to work on the thesis structure and grammar. Her advice and suggestions helped to perform coherent data analysis.

I thank all the Drupal contributors who have accepted to be interviewed: Bert Boerland, Gábor Hojtsy, Iztok Smolič, Josef Dabernig, Jeffrey A. McGuire, Ken Rickard, and Lin Clark. Thank you for your time, insightful interviewees, which helped me to perform this study, and all the information I obtained from you during the afterwards collaboration. I especially want to thank Gábor Hojtsy and Jeffrey McGuire for their additional help provided during interview analysis and data retrieve. I would also like to thank Greg Knaddisson for his impact to provide me various metrics and interpretations of Drupal-related information. Thank you all for providing a reliable data for my research.

The most remarkable thanks go to my parents, sister and my partner. You all gave me your unconditional support through all this process. Thank you for your patient, understanding and love.
# CONTENTS

1. **INTRODUCTION** .......................................................................................................................... 7  
   1.1. **BACKGROUND** .................................................................................................................. 7  
   1.2. **RESEARCH QUESTION** ..................................................................................................... 8  
   1.3. **SCOPE OF RESEARCH** ...................................................................................................... 8  
   1.4. **STRUCTURE OF THESIS** .................................................................................................... 9  

2. **RELATED WORK** ......................................................................................................................... 11  
   2.1. **CHALLENGES IN DISTRIBUTED SOFTWARE DEVELOPMENT** ........................................ 11  
   2.2. **COMMUNICATION MEDIA THEORIES** ............................................................................. 12  
   2.3. **COMMUNICATION IN DISTRIBUTED SOFTWARE DEVELOPMENT** ............................... 14  
   2.4. **SOCIAL MEDIA AS A COMMUNICATION TOOL** ............................................................... 16  
   2.5. **THE MICROBLOGGING USE IN SOFTWARE DEVELOPMENT** ........................................ 19  

3. **RESEARCH FRAMEWORK** ......................................................................................................... 25  

4. **RESEARCH APPROACH** ............................................................................................................ 29  
   4.1. **CASE CONTEXT** ................................................................................................................ 29  
   4.2. **DATA COLLECTION METHODS** .......................................................................................... 30  
   4.2.1. **SEMI-STRUCTURED INTERVIEWS** .................................................................................. 31  
   4.2.2. **UNOBTRUSIVE MEASURES** ........................................................................................... 32  
   4.3. **DATA ANALYSIS** .............................................................................................................. 34  
   4.3.1. **QUANTITATIVE ANALYSIS** ............................................................................................. 34  
   4.3.2. **QUALITATIVE ANALYSIS** ............................................................................................... 35  

5. **EMPIRICAL FINDINGS** .............................................................................................................. 36  
   5.1. **COMMUNICATION PROCESS IN CASE PROJECT** ............................................................ 36  
      5.1.1. **COMMUNICATION TOOLS USED TO PERFORM SOFTWARE DEVELOPMENT TASKS** 36  
      5.1.2. **MICROBLOGGING USE IN DRUPAL COMMUNITY** .................................................. 38  
   5.2. **CHARACTERISTICS OF MICROBLOGGING IN COMMUNICATION CONTEXT** .............. 42  
      5.2.1. **INFORMATION PROCESSING** ....................................................................................... 42  
      5.2.2. **COMMUNICATION PROCESS** ...................................................................................... 47  
      5.2.3. **MEDIA CAPABILITIES** .................................................................................................. 50  
      5.2.4. **TECHNOLOGY FEATURE** .............................................................................................. 52  
      5.2.5. **INTERACTION TYPES** ................................................................................................ 55  
      5.2.6. **SOCIAL ASPECTS** ........................................................................................................ 59  
   5.3. **SUMMARY OF COMMUNICATION PRACTICES IN CASE PROJECT** ............................... 61  
   5.4. **PRIMARY EMPIRICAL CONCLUSIONS** ............................................................................. 65  

6. **DISCUSSION AND IMPLICATION** ............................................................................................ 67  
   6.1. **THEORETICAL IMPLICATIONS** ......................................................................................... 67  
   6.2. **PRACTICAL IMPLICATIONS** ............................................................................................... 69  

7. **CONCLUSIONS** .......................................................................................................................... 71  
   7.1. **ANSWER TO RESEARCH QUESTION** .................................................................................. 71  
   7.2. **RESEARCH LIMITATIONS** .................................................................................................. 72
LIST OF TABLES

Table 1: Classification of Social Media by social presence/media richness and self-presentation/self-disclosure (Kaplan and Haenlein, 2010) .......................................................................................................................... 17
Table 2: General information about interviewed Drupal community contributors ................................................................. 32
Table 3: Drupal community popular Twitter groups (in mid May 2012) .................................................................................. 33
Table 4: General information about Drupal core developers .................................................................................................. 33
Table 5: Drupal community members (in mid May 2012) ........................................................................................................... 34
Table 6: Communication framework practice in Drupal ............................................................................................................. 62
Table 7: Use of microblogging in Drupal community ................................................................................................................ 64

LIST OF FIGURES

Figure 1: Communication tools used by distributed teams (Thissen et al, 2007) ........................................................................ 16
Figure 2: Interaction types in enterprise (Riemer and Richter, 2010) ........................................................................................ 23
Figure 3: Tools used for communication in distributed software development ......................................................................... 25
Figure 4: Theoretical framework of microblogging layers utilized in DSD communication .......................................................... 26
Figure 5: Project usage .................................................................................................................................................................. 29
Figure 6: Developers’ increase ......................................................................................................................................................... 29
Figure 7: Number of issues ............................................................................................................................................................ 30
Figure 8: Number of commits ........................................................................................................................................................ 30
Figure 9: Research approach ........................................................................................................................................................ 31
Figure 10: Leavitt’s diamond (Leavitt, 1965) .............................................................................................................................. 36
Figure 11: Interviewee tweeting frequency .................................................................................................................................. 38
Figure 12: Core developers tweeting frequency .......................................................................................................................... 39
Figure 13: Groups tweeting frequency ........................................................................................................................................ 40
Figure 14: Groups tweeting frequency (con’t) ............................................................................................................................ 41
Figure 15: Tweets with #drupal frequency .................................................................................................................................... 41
Figure 16: The conversation containing interviewee username network ................................................................ .................. 43
Figure 17: The conversation containing “#drupal” hashtag network ............................................................................................ 43
Figure 18: Tweet contains Drupal 7.14 and 6.26 releases announcement ...................................................................................... 45
Figure 19: Contributors retweets’ number for the period of one week (June 01-June 07 2012) ...................................................... 46
Figure 20: Tweet with #drupal retweeting tendency .................................................................................................................... 47
Figure 21: Conversation representation in Twitter ..................................................................................................................... 50
Figure 22: Standard deviation of Drupal contributors’ followees’ ................................................................................................. 54
Figure 23: Tweet presents individual reputation scale .................................................................................................................. 60
Figure 24: Tools used for communication in case study .............................................................................................................. 67
Figure 25: Theoretical framework of microblogging layers utilized in case study ........................................................................ 69
1. INTRODUCTION

This chapter first discusses the background of this research and then briefly describes the objectives and the scope of the study. Finally, the structure of the thesis is presented.

1.1. BACKGROUND

Software development must adapt to the tendency where team members work together but are geographically distributed. Many factors have contributed to this increase in distributed software development (DSD). The open source software (OSS) development model is by its nature globally distributed. Distributed software development enables software production to take place independently of the geographical location of the organizations concerned, however, DSD despite offered advantages is troubled with challenges. The core challenge of any DSD development lies in the complexity of maintaining communication when teams are scattered in time (e.g. time zones) and have differences in socio-cultural mentality. The lack of communication and trust will lead to ineffective collaborative projects and the ability to communicate may be directly related to project success (Werner et al., 2005). The previous research study has shown that the better development team members are connected to each other and to the external key parties, the most successful the project is going to be (Tushman and Katz, 1980). In distributed projects, managing implicit knowledge and maintaining awareness (Gutwin et al., 2004) impact the success of a project. Communication is certainly one of the key elements of successful software development. It is the critical factor in collecting and forming relevant information, share knowledge and create functioning products.

Understanding communication process in software development and explore effective communication tools are active topics in software engineering research. In DSD communities, geographical distance decreases the possibilities for the team members to meet face-to-face. There increases the need to find such media, which will substitute face-to-face interaction. Electronic media such as audio conferencing, voice over Internet Protocol (VoIP), email, instant messaging (IM) started to be used for this purpose. However, necessary communication does not appear often enough or includes too few or too many stakeholders (Cataldo et al., 2006) or confuse the priorities and intentions of the team members who do not share common knowledge base (Hinds and McGrath, 2006). Global networking technologies had prospective to improve the communication process in the teams. Every generation of communication provides easy to use and facilitated tools. The current technology known as social media provides an easy way for co-workers to share information in the social network. Social media and Web 2.0 technologies can improve software engineering practices as team forming improvement, agreeing on tasks and goals, successful communication, share their failures and successes at work (Begel et al., 2010).

This study is based on analyzing microblogging, one of the social media technology form, use as a communication tool in distributed software development. Microblogging is a facilitated way of blogs that provide functionality to share short (140 characters) messages. Previous researches showed that engineers write messages to share technical information links, announce spontaneous meetings, and inform team members about task status and progress (Treude and Storey, 2010). Microblogging provides a new informal communication channel complementary to other media like IM, email, face-to-face, and other social tools. Also, these media help to generate common ground that can be used to support future interactions that form connectedness. Existing microblogging studies in software development mainly focused on two aspects. One is about the microbloggers’ motivations and the other focuses on the usage of media. Obtained results outlined that the development team members use Twitter for gathering useful information for their work and life interests. However, researchers usually used the case study based on co-located companies or individuals from different organizations (Riemer and Richter, 2010; Bougie et al., 2011).

This thesis faces the challenge to understand better microblogging capabilities and how they can be exploited in distributed software development. The aim is to deliver both theoretical and practical instructions for researchers and team members. First a theoretical framework is created to explore what communication media layers exists in microblogging, in order to build rich collaboration process and
develop trust and connectedness between collocated community members. Then case study is analyzed, in which contributors are using microblogging for several years. As a result, it will be seen how analyzed social media is used as a one of the communication tools during software development process, what communication needs it covers and how communication process could be improved using microblogging.

1.2. RESEARCH QUESTION

The main objective of this research was to determine and explore successful and large DSD organization and analyze what role microblogging takes in this community daily life. Research findings would be useful for organizations that are planning or executing DSD projects. Accordingly, community with a vast number of members faces communication challenges, when there is a need to share information, announce events, or communicate with co-workers. Thus, it should be presented a set of tools including microblogging, which would help to solve all these issues.

The main research question of study is:

- RQ1: How is microblogging utilized by the DSD community for its communication needs?

In addition to the main research question, we have three supporting research questions:

- RQ1.1: What part of the communication process need does microblogging constitute in the DSD community?
- RQ1.2: How does microblogging affects communication performance in the DSD community?
- RQ1.3: What types of information are transmitted through microblogging in the DSD community?

From the methodological point of view, studying daily communication in DSD is challenging. Several data collection methods where used, such as semi-structured interviews and unobtrusive measures collection. Langley’s (1999) process based research strategy for analyzing quantitative and qualitative process data was used. Quantitative research study produce results that can be used to describe numerical changes in measurable characteristics of a population interest, generalize to other, similar situations (Salkind, 2010). Obtained tweets from particular timestamp were analyzed using descriptive and inferential statistics. Qualitative research investigates answers to a question, systematically uses a predefined set of procedures to answer the question, collects evidence, produces findings that were not determined in advance (Miles and Huberman, 1994). Coding and categorizing strategy (Flick, 2007) was used to analyze transcripts of the semi-structured interviewees. Kautz et al. (2001) conclude that the combination of several perspectives provides for researcher a better understanding, which would not have been possible if only one perspective has used. In this study, we have analyzed microblogging use as a communication tool from two angles of MST and MRT communication theories. To have rich tweet analysis results open coding strategy from grounded theory is used. Open coding proceed through each line examining and then define actions within coding. In this study, we have performed coding strategy analysis to develop interaction types layer from the data as it is collected and analyzed.

1.3. SCOPE OF RESEARCH

Since there are many different kinds of OSS projects, we had to choose only one mostly successful, in the sense of a number of users, and large, in the sense of the number of collaborators, for this study. There were decided to choose Drupal, because it is large in case of contributors’ and users’ community, with frequent product releases. Drupal system is an open source content management system, which functionalities are publishing, managing, and organizing variety content on a website. The built-in functionality of Drupal can be extended by existing number of modules. Drupal community has over
800,000 members and over 17,000 developers in more than 200 countries. Also, Drupal won several Packt Open Source CMS Awards and the Webware 100 three times in a row 1.

Communication science is a broad and an old area of research. The scope of this study is limited to look at the two communication media theories: Media Synchronicity Theory and Media Richness Theory. Thus, study selected narrow field of communication media analysis as communication in DSD. There is set of tools used in DSD, research selected most recent and unexplored media as social media phenomenon. Social media contain a list of tools, but we proceeded research over the microblogging – as its popularity and variety of use increases. Accordingly, the literature review in Chapter 2 concentrates mainly on communication media characteristics provided by microblogging which can affect communication process in DSD.

Microblogging is provided by various services – Twitter, Jaiku and Pownce. For this study, we will concentrate on Twitter, because it is widely used in Drupal community. Launched in 2006, Twitter is a microblogging service where users send updates (tweets) to a network of associates (followers) from a variety of devices. Tweets are text-based posts of up to 140 characters in length. The default setting for tweets is public, which permits people to follow others and read each other’s tweets without giving mutual permission. Each user has a Twitter page where all their updates are aggregated into a separate list called user timeline.

In this study, we concentrated on what role the microblogging plays in the DSD community communication process. DSD product development, bug fixing, and documentation is done in parallel, i.e., different parties perform work at the same time. Thus, constant communication between the developers and community members was needed.

As explained earlier, there were used quantitative and qualitative data analysis methods to process collected data. We have interviewed six Drupal community members. To have a deep understanding of communication we also have collected tweets containing Drupal hashtag (#drupal), tweets from six Drupal account groups, tweets from interviewees, and two core developers, because all these data demonstrate reasonable representation of existing communication process performed during software development.

1.4. STRUCTURE OF THESIS

This work has seven chapters: Introduction, Related Work, Research Framework, Research Approach, Empirical Findings, Discussion and Implication, and Conclusions.

Chapter 2 presents a literature review of communication media research in software development. The literature review reveals challenges in distributed software development, then communication media theories, communication issues in distributed software development, role of social media and more specifically microblogging role in software development process.

Chapter 3, based on literature review that will be used as a basis of empirical analysis, presents a theoretical framework of microblogging role in DSD.

Chapter 4 concentrates on the methodological issues. The research approach and the research material and methods are described. There two types of data collection are presented: semi-structured interviews and unobtrusive measures known as tweets. Finally, collected data are analyzed using quantitative and qualitative data analysis methods.

Chapter 5 discusses the empirical results. Firstly general communication process with all used tools is analyzed. Secondly, more detailed exploration of each theoretical framework layer with its parameters is evaluated using community practices. Finally, summary of practices and primary empirical conclusions are summarized.

1 http://en.wikipedia.org/wiki/Drupal
Chapter 6 suggests theoretical and practical implications based on obtained primary empirical conclusions.

Chapter 7 concludes the thesis by answering the research question, explaining research limitation and presenting possible future research.
2. RELATED WORK

This chapter is divided into five parts – first it discusses challenges in DSD. The second part is exploring communication media theories. The third part is about communication in distributed software development. The fourth part of the chapter presents the literature about social media use as a communication tool. Finally, research of microblogging use as a communication tool for software development is presented.

2.1. CHALLENGES IN DISTRIBUTED SOFTWARE DEVELOPMENT

Distributed software development with the increased favor in the industry constantly is analyzed in the available literature. Distributed software development enables software production to take place independently of the geographical location of the organizations concerned. This phenomenon began only about ten years ago and already has its strategic importance been recognized (Davison, 2007), and related studies reviews (Prikладникii et al, 2008). The distance between teams and one team members can vary from a few meters to different countries or even continents. The situation in which teams are distributed beyond the limits of a state is called Global Software Development (GSD) (Herbsleb and Moitra, 2001). This development model is intriguing for several reasons, mainly because it enables organizations to distract themselves from geographical distance, while having human resources and minimizing cost. Thus, produce software for remote clients and obtaining a continuous workday by taking advantage of time differences (Ebert and Neve, 2001). However, a number of problems caused mainly by distance, time, and cultural differences (Krishna et al, 2004), must be faced. This depends largely on the characteristics of each organization.

Thus, distribution makes harder to handle projects quality, schedule and other project issues than in centralized projects. In addition, distribution itself may cause time problems. A recent study shows that the physical distance between development sites alone is expected to create delays in work. Problems in project management, task coordination, and communication have also been introduced (Herbsleb and Moitra, 2001).

Komi-Sirvio and Tiinen (2005) conducted a survey to gather and share lessons learned from 21 different organizations in order to understand the nature of the software development process when performing in a DSD environment and the problems that may be associated with such distributed processes. In the survey, respondents were given the option of identifying eight different problem areas. These problem areas were identified on the basis of the descriptions found in the literature (Herbsleb et al.,2001; Mockus and Herbsleb, 2001). Firstly, respondents were asked to check all the problems they had experienced in their projects and then they described each identified problem in details. In addition to problem descriptions, respondents were also asked to describe the solutions they had developed to overcome the problems encountered and to evaluate the success of these solutions.

The area “communication and contacts” was selected as a source of problems for distributed projects by the majority (74%) of the respondents. Respondents used different strategies intended to substitute for the missing face-to-face communication. The most common solutions applied are tele or videoconferences and e-mail. However, even videoconference is regarded as problematic due to the connection difficulties that are often experienced. Despite this, problems were not particularly often linked to the tools used; only 30% of respondents regarded communication tools as a problem. It may be speculated that if communication tools are experienced as strong enough and communication is still a problem, better communication tools are not likely to solve the underlying communication problem. In summary, the following strategies are considered successful in improving communication in practice: 1) informal team building sessions and face-to-face meetings, especially at the beginning of the project; 2) decreasing the need for contacting other team members by splitting projects into small, independent and manageable units; 3) appointing a contact person from each site.
Communication problems seem to be extremely common within all organizations; in all, this problem area was ranked as the second toughest (Komi-Sirvio and Tihinen, 2005).

2.2. COMMUNICATION MEDIA THEORIES

Communication is one of the critical success factors for a software development projects (Kraut and Streeter, 1995). Increasingly, as geographically distributed team members communicate via computer, understanding the effectiveness of media and its role in everyday communication, in general, becomes paramount. Kraut and Streeter established communication in the context of software development when different people are working on a common project agree to a common definition of what they are building, share information and engage their activities (Kraut and Streeter, 1995). Additionally, researchers (Kraut and Streeter, 1995; Herbsleb and Mockus, 2003) divided communication into formal and informal. Informal communication is considered explicit communication via such channels as telephone, IM, emails or verbal conversations. Formal conversations refer to written specification documents, reports, protocols or source codes. Contemporary communication media facilitate communication and data exchange among a large number of individuals across long distances. Researchers have long studied effects of media theories on media selection and effects on media use. There exist two basic communication media theories which characterize communication medium performance by its ability to convey and gather information.

Media Richness Theory (MRT) (Daft and Lengel, 1986) claims that communication media differs in the ability to facilitate understanding depending on the amount and richness of information. MRT is based on Social Presence Theory (SPT) which claims that communication channel is effective if it contains social presence elements. According to SPT, face-to-face mechanism supposes to have the highest social presence, consequently text-based communication - the lowest (Salmins et al, 2000). Since many virtual teams never get to meet face-to-face during the project, substituting mechanisms may be required to create a sense of social presence. Past studies did not find a relationship between social presence and task performance (Yoo and Alavi, 2001), while social presence is indispensable to set and support social relationships (Urry, 2002). The lack of evidence can be explained that social presence is a media characteristic. Therefore, like other communication media characteristics social presence must be designated and used to supplement communication processes.

Previous studies showed that media considered being “rich”, when it had the ability to convey natural language rather than numeric information in different ways of conveyance (e.g., voice tone) and to personalize a message. MRT is based on the idea that successful choose of media is influenced by situations whether the information will be processed in case of uncertainty or equivocality (Daft et al, 1987). Uncertainty is obtained when there is an absence of information. Equivocality means the existence of distinct and contradicting interpretations. The main difference between those two types of tasks is that uncertainty leads to the acquisition of data and equivocality leads to the exchange of views to obtain unambiguous result. Media richness theory was developed to theorize prove that, in some cases, lean media is more effective than rich media. As an example face-to-face is considered to be richest communication medium in MRT, when written forms as emails are low in richness (Daft et al, 1987). Daft et al stated that group of four criteria can affect communication: language variety, feedback, personal focus and multiple cues.

Since the Daft et al studies of empirical tests of media richness started and divided into two categories: those supporting media richness (Zack, 1994) and those finding mixed or no support of media richness (Dennis and Kinney, 1998). One of those studies analyzed equilibrate between the tool and the message and get two points (Lionel and Dennis, 2005). First, Lionel and Dennis stated that the use of rich media (e.g., face-to-face) with high social presence stimulates increase of motivation, but decreases the ability to process information, while use of lean media (e.g., emails) with low social presence stimulates decrease motivation, but increases the ability to process information. Second, they stated about media’s ability to effect a change in understanding. In this case, receiver is a person; channel is a part of the medium through which he communicates with sender. Channel of expansion theory (Carlson and Zmud, 1999) explains the inconsistencies in the above studies. The theory identifies experiences, which influence persons “richness perception” for media channels. Those factors are: message topic experience, channel experience,
organizational context understanding and communication co-participants experience. Carlson and Zmud (1999) obtained empirical results, stated that individual’s work experience within certain media tool reflects his ability to perceive richness of communication channel and avail different features from that media.

Media Synchronicity Theory (Dennis and Valacich, 1999) looks beyond Daft and Lengel media richness and formulates new theory. The theory states that group communication process are composed of two fundamental communication processes: conveyance and convergence. Conveyance process is the transfer of a variety of new information to enable the receiver to create and revise a perception of the situation. Individuals participating in conveyance processes engage in substantial information processing activities so that a potentially large, diverse set of information can be exchanged in a variety of information formats. Individuals participating in conveyance processes will often require time to perform information processing—the cognitive processes is necessary to analyze the information, make sense of it, and build their mental models. Otherwise, convergence is the development of shared meaning for information. The team members have to understand each other’s view and agree upon one common view. In convergence case, high synchronicity is preferred. The theory also proposes set of characteristics to affect communication process: immediacy of feedback, symbol variety, parallelism, reprocessability, and reprocessability(Dennis and Valacich, 1999). When mixed types of tasks are proceeded, media switching is proposed.

“Task” has been a key element in the development and testing of media theories (Daft and Lengel 1986; Dennis and Kinney 1998). However, studies comparing task performance between individuals working on different tasks with different media have not convincingly shown that a better match of media to the task will yield better task performance(Dennis and Kinney 1998). Task performance differences attributed to fit between media and task have not been consistent. “Task” is presented as a terms of the fundamental communication processes that must be performed. This is analogous to the concept of steps, which are the underlying acts required to accomplish a task. To better understand task outcomes, we must understand how individuals perform these fundamental steps in terms of which steps they choose to perform, in what order, and when. Every task involving more than one person requires a mix of different communication processes to perform these steps. To understand how media can influence communication performance, it is necessary to more carefully examine underlying communication processes, which include both the transmission of information and the individual cognitive processes to make sense of the information (Robert and Dennis 2005).

This study is mainly concentrated towards tasks of uncertainty (or low equivocality tasks) which represents situation, when task is analyzable, but there is a lack of information. Depending on development stage, different communication process is required. For example, in project inception phase technical problem solving may be specified or easily deduced, so that members can move quickly to execution, whose focus is on the exchange of information. Still convergence is required before the group can move to the execution phase. Group well-being and member support is less clear for tasks of uncertainty than equivocality. For a well established group, the group may quickly proceed from inception to execution using established routines. However, if the group is newly formed, or new members have been added to the group, more time may need to be spent in inception, technical problem solving, and conflict resolution (Hollingshead et al, 1993). However, the group will still use the fundamental processes of information conveyance and convergence, but the information may require a different symbol set.

Dennis at al. (2008) expanded MST theory they argue that successful completion of tasks by some participants requires convergence and conveyance processes. MST outlines media’s capability to support synchronicity. Moreover, depending on how familiar individuals are with the task, the chosen media, and with a communication participant conveyance and convergence processes will be performed in a different way and intensity. When is observed higher understanding of the task, the media, and responder, the communication will require more convergence processes. As a result, higher media synchronicity is more appropriate for the communication. Conversely, with less understanding of the task, the media, and resolver, the communication will involve more conveyance processes. As a result, lower media synchronicity is more appropriate for the communication (Dennis et al, 2008).Expanded theory claims that communication performance will increase, only, when different form of media will be used.
Communication has been defined as a process where participants develop and share information between each other to obtain mutual understanding. Sharing information is inherently an exchange process, in which developing meaning requires dissemination of information (information transmission) and individual processing of that information (information processing). Developing shared meaning requires that individuals not only understand the information they have, but also understand how others interpret it. Thus, an important outcome of successful communication is the development of shared understanding about the information and the meaning that each participant attaches to it (Daft and Lengel 1986).

Common Ground (CG) theory (Clark and Brennan, 1991) explicitly presents common ground that means mutual understanding practices and characteristics. CG theory presents eight properties that can affect grounding process: co presence, visibility, audibility, co temporality, simultaneity, sequentiality, reviewability, revisability. Any medium does not obtain all the attributes at the same time. Theory claims that face-to-face conversation easier than other tools establishes common ground, because of access to cues like facial expression. Communication over media with less number of possible cues forms harder construction of CG. As a consequence, distributed teams without mutual knowledge use audio or video channels for collaborative tasks, and teams with existing common ground can communicate effectively using lean media (Clark and Brennan, 1991).

To sum up reviewed communication theories and techniques the genre analysis is presented. Genres capture meanings and reflect practices of the communities in which they exist. As such, genre analysis can serve as an instrument to understand the communication practices of a social group, because, by labeling, genres various components of communication can be revealed (Kwasnik and Crowston 2005). Practices, in turn, can be defined, as a routine formed, of several elements: body activities, mental activities, things and their use, background knowledge in the form of understanding, know how, states of emotion and motivational knowledge. A practice outlines the routine of communication; with regards to technology, a practice view directs researchers’ attention to the technology in use (i.e., ways in which different people use technology in a time and place) (Orlikowski and Iacono 2000). Communication mediums are appropriated over time and embedded in emerging and situated collaboration practices, which are reflected in the communication genres, which structure this communication.

2.3. COMMUNICATION IN DISTRIBUTED SOFTWARE DEVELOPMENT

Earlier research has shown that distance strongly influence communication between team members (Sosa et al., 2002). Distance leads to plentiful amount of problems in communication and coordination, which impacts the performance of globally distributed software development teams (Herbsleb and Mockus, 2003). There is a large amount of communication tools that can be used by a virtual team. Global virtual teams are facing a significant challenge by understanding how and when to use different types of communication media.

The open source software development model is distributed by nature, and various OSS developments are held success stories. Distributed development has an impact on communication, control and coordination within OSS projects. Number of researchers has focused on these three fundamental processes (Evaristo et al., 2004). Communication is considered as information exchange. For communication to be successful, exchanged information should be complete and unequivocal such that the sender and receiver will obtain a common understanding. The communication process consists of transfer of knowledge and information between participants, and the media used to facilitate such interaction.

Everyday communication associated with development activities in OSS projects is usually asynchronous and Internet-based. Teams use a collaborative development environment (CDE), which propose issue tracking, mail lists, version control, and other services. Developers also use synchronous communication channels for intense technical discussions. An OSS community usually consists of motivated and self-selected contributors with a high level responsiveness. Fast feedback supports openness and responsiveness. Still, problems can occur like undervaluing activities (Lundell et al, 2006)

Exchange of knowledge among remotely located software development teams must overcome barriers in order to provide interactions and communication comparable to the original co-located operations.
Distributed teams confront different communication issues: time zone differences, cultural and language barriers. Analysis of used communication tools in three distributed successful projects was performed (Tissen et al, 2007). Study showed that synchronous media (e.g., telephone, IM) where used in case of immediate feedback need. Asynchronous media (e.g., email, document exchange system) were used when timing was not critical. Additionally, there were outlined frequent communication as a key to success.

Heeren and Lewis (1997) presented framework to understand and recommend media types of tasks by integrating MRT. The framework suggests that comparatively rich media should be chosen to achieve effective equivocality reduction, while comparatively lean media should be chosen for team activities at the operational level. Distributed teams require multiple media to intermediate communication and collaboration in different phases of work. Examples from the case studies show that using a rich medium such as face-to-face meetings is paramount in the initial phase of a project, so that shared understandings can be developed effectively (Heeren and Lewis, 1997). Experiment to use only rich media in GSD did not provide the best performance, however, switching between rich and lean media suggest progress of team performance even for complex activities (Lanubile, 2009).

Controversially empirical study to support MST as a theoretical framework was performed to analyze what communication tools are used in 12 GSD projects (Niinimaki, 2010). The study claims that it is necessary to have a set of communication tools to perform effective information sharing and communication practicing. There was agreed, that, as all technology, communication media evolve over time, so new communication media appears. It is vital to follow and introduce new communication media to the set of selected communication tools in GSD.

Communications in the OSS development environment rely heavily on electronic media (e.g., forum, mailing lists) rather than face-to-face contact. Electronic communication significantly affect the performance of virtual teams through trust, organizational identification, and communication network structure (Yang et al, 2010). Teams with high levels of trust, faced with technical or task uncertainty, are able to solve problems when they are limited to computer-mediated communication. Trust is also highly personalized parameter, because an individual who is used to being seen in person by a co-worker may avoid at appearing on camera for a video conference. If an individual can easily speak to another person in the same room, he may unwillingly interact on telephone. Help request to debug a piece of software may require scheduling in common agreeable time, setting up remote access or organizing video services. Distributed software teams development is carried out simultaneously by facing these multiple obstacles (Thissen et al, 2007). Also, studies showed, that teams that did not achieve task-focus in their communication reported low levels of trust (Yang et al, 2010). In addition, found that electronic communication is particularly important in the creation and maintenance of a common identity among decoupled virtual teams. From the network perspective, found that the fit between virtual group/organization task and communication network structure is associated with members’ perceived group performance. In general, these studies have established strong links between communication and virtual team performance. However, there is little research on the impacts of communication on virtual OSS project groups (Yang et al., 2010).

For successful team creation and maintenance, the virtual group must use unit of tools which will allow easy communication flow (Figure 1). The choice of tools is extremely valuable. It is widely believed that people differ in preferred technique of communication (Thissen et al, 2007). For some, visual presentation proves most effective. For others may work better speaking, listening and discussing. Most people combine these approaches or use one or more selectively, depending on the factors (Thissen et al, 2007). Accommodating the desire for multichannel communication and increasing frequency of contact can lead to improved remote collaboration, depending upon the choice of tools and mechanisms for exchanging information. Individuals’ needs for oral and visual communication are supported by document, image and shared operational software. Processes and systems depend on time and needs. Synchronous systems provide for participants dynamic “real time” send and receive of information know as a simultaneous. Asynchronous systems provide a technique to submit and retrieve information transferred over an interval of time which is determined by the availability of the receiver.
Frequently used synchronous tool is chat or instant messaging. IM is a balance between a phone call and email with respect to formality and immediacy. In system user obtains a dynamic contact list which represent if any individual on the list is online and/or available. Two and more people participating in IM at once, the technique is perceived as chat, usually real time. Also, a message can be sent to another user like short form of an email, even though the response may be delayed. IM and chat clients usually provide a logging option, for conversations which require a stored record (Lanubile, 2009).

Groupware is another communication tool type, which enables multiple users to use a single software resource. Systems provide shared calendars, group meeting scheduling and shared contact lists (Thissen et al., 2007). Each person maintains his own information in a private calendar; arrange meetings, telephone conferences or any kinds of this interaction; other forms of groupware are document managers. Email is still one of the most popular ways to pass information to people in different time zones or across the room. Its asynchronous nature allows the sender to transmit information regardless of whether or not the recipient is available at the moment (Niinimaki, 2010). Storage of the message is at least temporarily on a central server, which maintains a persistent copy of the mail until it is retrieved or removed by the recipient. Email may include optional attachments of audio, video, document or other electronic content.

Thissen et al draw general conclusions to form guidelines for future teams. Researchers firstly suggested allowing teams to choose their own communication tools from a variety of options. Secondly, they insist on frequent communication among all members, including some synchronous interaction (telephone, chat, web conferencing). Finally, they outline use of shared file storage to facilitate team interaction (Thissen et al., 2007).

Additional, there is necessary to outline that distributed communication performance is influenced by the users experience with a medium, participants can learn to use and favor it (DeRosa et al, 2004). DeRosa et al state that, without experience, it will be more difficult to achieve trust, in non face-to-face interaction. Research also showed that effective teams develop cognitive and affective based trust. Hence, task focus must be accompanied by social communication for trust to develop and teams to be effective.

### 2.4. SOCIAL MEDIA AS A COMMUNICATION TOOL

This section explores relatively new study of social media uses in global software systems development. Social media involve web and mobile based technologies to develop interactional dialogue between participants. Kaplan and Haenlein describe social media as set of web-based application based on Web 2.0 used to develop and exchange content. Researchers applied a set of theories from the social processes (e.g., self-presentation) and communication theories research (e.g., media richness) to present the concept of social media. Concept is based on the classification scheme consisted of six types of social media: blogs.
and microblogs, collaborative projects, social networking sites, content communities, virtual social worlds, and virtual game worlds (Kaplan and Haenlein, 2010) (Table 1).

<table>
<thead>
<tr>
<th>Self-presentation/ Self-disclosure</th>
<th>Low</th>
<th>Medium</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blogs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social networking sites (e.g., Facebook)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Virtual social worlds (e.g., Second Life)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Collaborative projects (e.g., Wikipedia)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Content communities (e.g., YouTube)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Virtual game worlds (e.g., World of Warcraft)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Classification of Social Media by social presence/media richness and self-presentation/self-disclosure (Kaplan and Haenlein, 2010)

Empirical pilot study shows that social media can enable better communication through the software system development process (Black et al, 2010). Study outlines facts that social media tools where used for several hours per day mainly to discuss work related topics and personal information. Participants used several social media tools, however, most popular where Twitter and instant messaging (Black et al, 2010). Necessary communication between developers do not occur often enough or at all (Begel et al, 2010); issues arise because it include too many or too few stakeholders (Cataldo et al, 2006) or can distort the intentions of the team members who do not share a common knowledge base (Hinds and McGrath, 2006). Considering that software engineers creates and first adopts new communication technologies, that means that they are already using modern social media tools for software development work.

However, there is other opinion that there is little advice on how, or indeed if, social media features should be integrated within modern software development environments. Through social media, users can be involved in the software development process, now they are not comprehended as “passive recipients of goods, but as co-creators of value” (Kazman and Chen, 2009).

Social media use is a moving target, and inventions are adopted and adapted faster than researcher can study them. In many projects, users share their experiences and opinions (e.g., in blogs), but it is not clear how these kinds of information could be used to benefit development activities (Storey et al, 2010). Other studies analyzed how software companies use social media methods to collect feedback from users collectively. Company size, transparency, software deployment, and number of social media tools in use and the way they influence the use of social media is analyzed in depth (Bajic and Lyons, 2011). Based on media synchronicity, researchers claim that social media can promote work performance by stimulating trust among employees and offering a communication channel. In which explicit and implicit knowledge can be effectively transferred (Cao et al, 2012).

There is also beneficial to outline that social media have a strong influence over the reputations of the participants, which has career consequences within a community (Begel et al, 2010). Social media tools can be characterized by underlying “architecture of participation” that supports crowd sourcing, as well as, many-to-many broadcast mechanism (O’Reilly, 2005).

Different traditional tools, such as instant messaging, email or social media provide effective ways to communicate, collaborate, and participate on synchronous (e.g., chats) or asynchronous (e.g., blogs, forums) communication options, they are normally viewed as asynchronous media (Farkas, 2007). Although social media offer a variety social networking services (SNS), they are normally viewed as asynchronous media (Cao et al., 2012). DiMicco et al propose for teams to use social network services to “build stronger bonds with their weak ties and to reach out to employees they do not know” (DiMicco et al, 2008). Other asynchronous social media type is microblogging. “Lightweight” version of a blog, to share short messages with a set of followers (people who “follow” the person) who have signed to receive them. Messages can be read and written from any device connected to Internet enabling constant access to information. Engineers write messages to publicize technical information links, announce events or spontaneous
meetings, and keep team members aware of status and progress on current tasks (Treu and Storey, 2010). Other study claims that engineers use microblogs to spread knowledge, ideas, and suggestions to colleagues. Usually, OSS communities grow more united regularly using social media. The distribution of knowledge through social media increases completeness and speed, thus, decreases misunderstanding between colleagues even if they do not meet face-to-face very often, or at all (Beigel et al, 2010). As the popularity of social media increases, arises number of companies integrating social media into organizations communication process. Thus, empirical research on social media in the workplace is still infrequent.

In performed studies, there were found that individuals who use social media are not just searching and sharing information. They also regard them as a platform to meet friends, obtain a sense of belongingness and develop relationships with other people. Normally, individuals accumulate their social capital as a consequence of daily social interactions, but it is also feasible to make intentional investments in social interaction. Prior research has provided abundant evidence that social media use helps build social capital.

Social capital is the resources embedded within an individual’s or an organization’s network of relationships, including both interpersonal relationships and the resources rooted in the relationships (McFadyen and Cannella, 2004). Social capital is a multidimensional concept, which can be divided into a structural, relational and cognitive dimension. The structural dimension means overall pattern of connections between people, i.e., individual connects with team members’ and how. The relational dimension describes resources embedded in the social relationship such as trust, commitment, and reciprocity. The cognitive dimension refers to a common context which increases understanding among people represented by shared language, codes, and goals. For the members of a network, the social capital benefits contain broader sources of information and opportunities that are otherwise unavailable. In this study, we mostly focus on the relational dimension of social capital, which is perceived as a trust. The most evident motivation for people to use social media is the need for social interaction. Employees in the workplace communicate to create, maintain and strengthen their interpersonal relations.

Previous studies have demonstrated that interpersonal trust plays a positive role in knowledge transfer, trust has been considered as a fundamental factor for the success of knowledge transfer. Empirically, trust has been proved to result in enhanced knowledge exchange, makes information sharing and knowledge exchange more effectively and increases the possibility that knowledge is absorbed by the recipient successfully (Levin and Cross, 2004). Trust is particularly beneficial for knowledge receivers, since it increases the credibility of knowledge and, therefore, enhances the likelihood that receivers will consider using that knowledge in their deliberations. In virtual communities, trust is also positively associated with the quantity and quality of knowledge sharing (Chiu et al, 2006). Trust has been proved as a key antecedent in shaping cooperative interaction, reducing conflicts within the organization, developing successful solutions, and increasing the effectiveness of teams (Jarvenpaa et al, 2004). Trust is expected to improve the work performance of individuals both in terms of effectiveness and efficiency.

Although social communication and interactions are not focused on the team task, they have been found to be related to positive results, such as higher satisfaction and better decision quality (Cao et al., 2012). Trust is the most crucial element in the shape of excellent communication among team members. Therefore, Cao et al proposes that social communication has an indirect effect on work performance via the impact on trust. Researchers contend that social media can stimulate trust between employees, offering a communication channel where explicit and implicit knowledge can be effectively transferred. Trust can facilitate knowledge transfer. Both trust and knowledge transfer help promote work performance (Cao et al., 2012).

However, information technology alone is typically not the main driver for any knowledge management initiatives. It is the way people use it that shapes the role of IT in supporting knowledge sharing (Huysman and Wulf, 2005). As an effective social networking platform, social media are widely employed to maintain external professional networks and strengthen ties with colleges. In a collaborative environment, individuals often develop and rely on their own ego centered networks in deciding with whom to collaborate and how to collaborate. A characteristic of social media is that the existence of the social network is visible and more accessible, enabling individuals to locate expertise in one’s social network
more effectively in processing online collaborative knowledge transfer. These social network and ties created through social media can link organizational members to various sources of knowledge and help an explanation of the knowledge, fostering a new form of information exchange. Because of informal networks, weak ties, boundary spanners and social capital, the utilization of social media is expected to facilitate knowledge transfer.

Dron (2007) developed a hypothesis, based on MST, stating that conveyance processes performance using medium with lower synchronicity will lead to better communication performance. In the networked environment, individuals face a considerable challenge in managing the following modes of communication simultaneously: one-to-one, one-to-many, and many-to-many (Dron, 2007). Among the five media capabilities, transmission velocity and parallelism are essential to media synchronicity. MST defines lower synchronicity as media that provide low transmission velocity and high parallelism (Dennis and Valacich, 1999). Social media’s low transmission velocity leaves abundant time for people to deliberate information carefully and express their opinions clearly.

Jackson et al. (Jackson et al, 2007) also found out that online social networks created by social media are beneficial supplements to offline networks, permitting employees to know detailed information about their colleagues, such as personal background, character traits, hobbies and interests. Deepening mutual understanding can reduce uncertainty about other people’s behaviors and intentions as a prerequisite of trust (Valenzuela et al, 2009). The more we recognize others, the more we may trust or distrust them. Individuals who trust each other may promote communication through various media, including social media. For instance, Ellison et al (2007) found that, among American undergraduate students, use of Facebook2 had an intimate association with maintaining and strengthen existing offline relationships. Similarly, it is unlikely that employers will use social media to communicate with people who they distrust. In fact, social media use and mutual trust have a reciprocal relationship.

As the popularity of social media rises, increasing numbers of companies are incorporating social media inside their organizations. However, empirical research on social media in the workplace is rare. By integrating media synchronicity and social capital theory, Cao et al have developed a model to investigate the influence of social media in the work context. Also, Cao et al model explored underlying mechanism for how they create value at work. Specifically, empirical study indicated that social media enhance knowledge transfer by fostering trust among employees, thus leading to better work performance (Cao et al., 2012). Although social media’s benefits for organizations are difficult to quantify, researchers have demonstrated that social media are valuable computer-mediated communication (CMC) tools in the workplace.

2.5. THE MICROBLOGGING USE IN SOFTWARE DEVELOPMENT

Microblogging is a form of blogging where message length is up to 140 characters. Templeton (2008) defines microblogging as a small scale form of blogging, made up of short, compact messages, to share news, post statuses, updates, and have conversations. Microblogging is used for a faster mode of communication. Short posts decrease users’ spent time effort and thought investment for content generation. Twitter is the most well-known microblogging platform compared with other microblogging tools, such as Jaiku3 or Identica4 (Java et al., 2007). It is considered as one of the fastest growing applications in the history of the Internet (Scoble, 2007). In fact, Twitter has received considerable attention from online communities, media companies, business companies, bloggers, researchers, etc. For example, an official study has been conducted on Twitter about user intentions in using microblogging, geographical distribution of users and the relationships between different user networks (Java et al., 2007). There are also particularly useful ideas about the use of Twitter in the workplace, such as notification of current progress of work, task assignment done outside the workplace, notification of change of meeting schedules (Riemer and Richter, 2010).

---

2 Social networking service and website launched in February 2004 (http://www.facebook.com/)
3 Was social networking, microblogging and life streaming service (shut downed on January 15, 2012)
4 Is an open source social networking and microblogging service (http://identi.ca/)
This social medium does not restrict the developer to send formal status messages (e.g. log messages) or entirely informal (e.g. chats like in IM). Developer can talk about successful implementations, problems or just cry for help (Reinhardt, 2008). Each microblogging service provider saves all the sent messages. Each message (not exceeding 3000) stays accessible until the sender decides to delete it, and thus can be accessed from any person in the sender’s network. If we assume that every developer is mutually connected to each other developer in an organization, each of the messages can durable be searched for hints, solutions or experts (Reinhardt, 2008).

To implement facilitate message search microblogging uses the concept of tagging. Tags are used to classify items in an informal way, and they stand in contrast to formal top-down classification mechanisms. Golder and Humberman (2006) provide overview of tagging systems and classify the main reasons for user tagging. A common finding across these studies is that user tag provides artifact information and categorization. Part of the success of tagging comes from allowing users to create their own library. Treude and Storey analyzed the use of tags in the context of software development and described as “tags are used to describe resource such as source files or test cases in order to support the process of finding these resources” (Treude and Storey, 2009). Their findings showed that tag support may play a vital role in improving team-based software development practices.

In recent years, microblogging as an internet communication technology has drawn academic attention. Existing microblogging studies mainly focus on two aspects. One is about the microbloggers’ motivations (Java et al. 2007; Zhao and Rosson 2009). Zhao and Rosson (2009) argue that, comparing with other digital communication technologies like IM, email, or blog microblogging can promote collaborations in professional relationships through strengthened common grounds, a feeling of connectedness, and exchanging acknowledgements through follower networks. The other types of research mainly focused on the design and usage of microblogging technologies (Honeycutt and Herring, 2009).

One of the key roles microblogging provides for person-to-person interaction is collaboration. Flexibility of access and lightweight architecture of the medium has the potential to be used for sharing ideas and coordinating activities, similar to IM (Quan-Haase et al, 2005). Study suggests that if Twitter is used for collaboration, communication in a pair or small groups would be more effective than large, open discussions (Honeycutt and Herring, 2009). Honeycutt and Herring states that Twitter can be used in formal collaborative contexts, for example, distributed teams like instant messaging have been used before. Authors mention issues according Twitter functionality. Firstly, it is a limited number of tweets the user receives per manual and automatic (intervals of not less than one minute) refresh, which cause time intervals for responses. Secondly, a bounded history of tweets from followers, which is not stored. Thirdly, a lack of an interface view that displays tweets directed to and received from the same users (i.e., all parts of an exchange in one place). Finally, the absence of a search function (Honeycutt and Herring, 2009). After three years, almost all Honeycutt and Herring requests were implemented: search function, history of send and received direct messages, user friendly interface, plenty of third party applications available to organize Twitter interface by user preferences. Tweets are not only displayed on a user’s profile page, but they can be delivered directly to followers via instant messaging. Short Message Service (SMS), Really Simple Syndication (RSS), email, or other social networking platforms, such as Facebook.

In recent years, there were made various amount of studies to understand second aspect of microblogging - why and how people uses Twitter for communication purposes. One of the studies, reported using Twitter for a variety of social purposes, including keep in touch with friends and colleagues; increasing visibility of fascinating things from social networks; collecting useful work-related or other personal interest information; looking for help and opinions, and expressing emotional stress (Zhao and Rosson, 2009). These motivations seemed to be quite similar to purposes reported in studies of other social media for (e.g., IM, Blogs, and RSS), and as such poorly explained why people use Twitter, so researchers decided to analyze characteristics and technology features to understand how these users have appropriated microblogging into their social lives.

---

5 Is one of many kinds of tangible by-product produced during the development of software. Some artifacts (e.g., use cases) help describe the function, architecture, and design of software. Other artifacts are concerned with the process of development itself—such as project plans.
Zhao and Rosson states, that people use Twitter for gathering useful information for their personal work and life interests. Usually, content is shared in real time for people having similar interests, and it is trustable information from own personal network. Authors found that following technology characteristics help explain why people use Twitter for informal communication: brevity - helps make information more tight, thus, possible benefit from the reader’s perspective; mobility - pervasive access; broadcast nature – the easy way to access masses (Zhao and Rosson, 2009). Work-relevant information sharing and knowledge seeking and interchanging has been a critical issue in organizations. Practitioners and researchers always looked for a way of motivation that would lead employees to share knowledge and valuable information; motivation is needed because information sharing often involves supplementary effort by a worker. The technology characteristics of microblogging (e.g., brevity, mobility) may offer methods to reduce users’ cost of sharing, and thus make it easy for other employees to obtain useful and trust worthy information.

Therefore, microblogging may contribute colorful, content useful, in constructing person online representation, as well as, building background receptions of others to reduce social, cognitive cost in conversation. This suggests that microblogging may supplement other social software (e.g. SNSs, blogs) in organizations, as a way of getting to know a colleague as a person and learn about his interests, as well as work responsibilities. Study showed that microblogging was useful to increase awareness of what is on individuals’ mind. This, in turn, implies that it may help to generate more common ground that can be used to support future interactions that form connectedness. The real time personal updates found in Twitter may help maintain a virtual feeling of closeness (i.e., being there) (Zhao and Rosson, 2009). People use Twitter at work to keep up with what is new, what is happening with co-workers, share and exchange information need. It provides a new informal communication channel complementary to other media like IM, email, phone, face-to-face, and other social tools. Frequent small updates of personal life events, one of the unique characteristics of content shared on Twitter. Because microblogging trends to occur in real time, Twitter posts were regarded more valuable than other media for linking information to personal goals, for knowing what is on individuals’ minds at this moment, and for motivating opportunistic interactions (Zhao and Rosson, 2009). The principle of a “people-based RSS feed” was another characterization described by study participants. They were able to get truthful and useful information from people who they know individually and select to follow. Often these individuals are selected because they share similar interests with the subscriber. These technology characteristics were found useful for reducing cost of sharing and promoting frequent updates in real-time, as well as making easy browsing and monitoring of a large amount of information updates. Study suggests that microblogging may help colleagues to know each other better as an individuality, in addition, to professional relationships. This benefit is achieved by staying aware of small details about colleagues’ personal live, interests, and current moods. Which, in turn, creates more opportunities for exchanging acknowledgements and social assistance, generating new common ground, and creating and sustaining a sense of connectedness. All of these can improve colleagues’ efforts toward future collaboration at work (Zhao and Rosson, 2009).

In order to identify usual communication patterns in Twitter, authors have used genre analysis. What can be observed in context is the communication events people engage in during their daily life, such as a post in a microblogging stream. Conceptually, a genre is a form of communicative events; communication events in turn are instantiations of a genre. Some shared set of communicative purpose transforms a collection of communicative events into a genre (Askehave and Swales, 2001). The actual genre analysis can then be carried out using several techniques, e.g., document/text analysis, interviews.

Among different Twitter genre analysis studies, we have outlined some of them. Firstly, one study (Naaman et al., 2010) shows that 80% of Twitter users can be classified as so-called “me-formers”, who mainly make themselves the object of their communications. The authors identified that 20% are seen to be true “informers”, who post content that is targeted towards other users’ interests. Most frequent communication types:

- 41% of all posts contain information about oneself (e.g., “I'm tired”)
- 25% are random thoughts (e.g. “Blue sky in Winter”)
- 24% announce personal opinions (e.g., “Great game yesterday”)
- 21% exchange of information (e.g., “New Study on Enterprise 2.0: http://...”)

At the same time, a lively discussion on the benefits and risks associated with corporate social media has emerged online commentary outlets. Among the concerns is a statement that microblogging is only a
temporary fashion and that most organizations will not be ready culturally to implement social media (Marchionda 2009).

Guzzi et al studies presented generic types of the daily tweets:

- 33% of messages about future intentions (e.g., “Now I am going to”)
- 23% covers an ongoing activity (e.g., “I am”).
- 21% on a past activity (e.g., “I just did.”)
- 16% comments (e.g., “This is like so”)
- 6% to do’s (e.g., “Later, I will need to”)

Authors inspected posts remarkably close to each other (within 30 seconds) and noticed that they are directly correlated, with the second message acting as “annotation” for the previous post. Also, it is the case that the first message states the end of the previous activity. The study also indicates that the limit suggested of 140 character per message is sufficient to express what they are doing, because during the analysis message length average was about 55 characters (Guzzi et al, 2010).

Genre analysis performed one year later showed qualitative analysis of the messages posted particularly by software engineers(Bougie et al, 2011). The following four categories of tweets emerged:

- Software engineering and work-related topics (e.g., projects being worked on, seeking or providing technical help)
- Gadgets and technological topics (e.g., product news from up-to-date technologies)
- Current events outside of technical topics (e.g., politics, sports, travels)
- Daily chatter (e.g., family, weekend activities)

Software engineering-related tweets accounted about 23% of the tweets that were qualitatively analyzed. These cases included discussion of the current tasks developers are working on, and in several cases, attempts to find solutions to appropriate issues they encounter. The subtopics that authors discovered, within this category are: problem solving, self-promotion, complaints, and the use of tools for work(Bougie et al., 2011).

At the same time, others have performed analysis of using social media inside the corporations. Riemer and Richter, by applying genre analysis, explored communication patterns, in a team, that has adopted Enterprise Microblogging (EMB). For the case study, researchers used co-located large company, which was rapidly growing, and difficulties emerged with sharing information and ideas. Authors find that microblogging in this corporate context is vastly different to its public equivalent. The genre analysis identified above (Figure 2) indicates that Riemer and Richter case team has appropriated EMB to support two key team practices like awareness creation and task/task coordination. Firstly, team members, through providing updates on various issues, actively engage in the creation of awareness, which is defined as understanding of the activities of others, which provides a context for your own activity. While awareness is present naturally in face-to-face communication, in technologically intermediated environments, information about team members and their activities, is not straightly present. Such genres as update task, event, or links posting to information on the Internet thus reflect communication practices of awareness information. Secondly, EMB practice evolves teamwork on shared tasks, for example, users post lists of to do items, delegate tasks, make others aware of task progress and ask/answer task-related questions. Communication is also concerned with coordinating team matters, for example, providing updates on team events. Riemer and Richter conclude that EMB serves the role of an awareness creation and task or team coordination medium. At the same time, the above genre representation (Figure 2) also reflects that EMB is not used to support other team practices such as discussions or more in depth cooperation, as reflected in the lack of a significant number of posts in the “Discuss and clarify” category. Very rarely team members performed voice opinions and had few discussions(Riemer and Richter, 2010).
Figure 2: Interaction types in enterprise (Riemer and Richter, 2010)

Secondly, EMB practice evolves teamwork on shared tasks, for example, users post lists of to do items, delegate tasks, make others aware of task progress and ask/answer task-related questions. Communication is also concerned with coordinating team matters, for example, providing updates on team events. Riemer and Richter conclude that EMB serves the role of an awareness creation and task or team coordination medium. At the same time, the above genre representation (Figure 2) also reflects that EMB is not used to support other team practices such as discussions or more in depth cooperation, as reflected in the lack of a significant number of posts in the “Discuss and clarify” category. Very rarely team members performed voice opinions and had few discussions (Riemer and Richter, 2010).

However, there is little research about the impacts of microblogging and its follower network on team performance, especially in OSS projects. Communications in the OSS development environment rely heavily on electronic media (e.g., forum, mailing lists) rather than face-to-face contact (Yutaka et al., 2000). Electronic communication has been found to, significantly, affect the performance of virtual teams through trust (Jarvenpaa and Leidner, 1999), organizational identification (Wiesenfeld et al. 1999), and communication network structure (Ahuja and Carley, 1999). Jarvenpaa and Leidner (1999) found that when faced with technical or task uncertainty, teams with high levels of trust are able to solve problems when they are limited to computer-mediated communication. On the contrary, teams that did not achieve task-focus in their communication reported low levels of trust. In addition, Wiesenfeld et al (1999) found that electronic communication is particularly noteworthy in the creation and maintenance of a common identity among decoupled virtual teams. From the network perspective, Ahuja and Carley (1999) found that the fit between virtual group/organization task and communication network structure is associated with members’ perceived group performance. In general, these studies have established strong links between communication and virtual team performance. However, there is little research on the impacts of communication on virtual OSS project groups. Huberman et al define a person’s “friend” as someone they have sent at least two direct posts. By this definition, even though it is unidirectional, a much clearer relationship is seen between the number of friends a person has and their level of activity on Twitter. 98% of users involved in the study had fewer friends than followees6. By these findings, the authors conclude that though an individual may have a seemingly large network of followers and followees(Huberman et al, 2009). Thus, Yang et al explore that the follower network affects project performance (Yang et al, 2010).

Literature review section is enclosed by exploring microblogging use is a corporate context. Riemer and Richter compare results with similar studies on Public Microblogging (PMB) in Twitter with EMB. Results showed that EMB as a phenomenon is noticeably different from PMB. While, in both cases, the underlying technology is quite similar, assignments and use in an enterprise context are structured by the needs of the tasks at hand, by a shared group context. Communication in PMB is less egocentric than in EMB. Instead,

---

6People followed by the person, the declared friends
it is oriented towards the needs of others. Study suggests that common conviction that social media importing to the workplace will bring employees distraction is not real based. To the contrary, decision makers should vest trust in their employees to appropriate technologies in a useful and fitting way. Messages can be created directly on the web or with a wide range of mobile clients. While simplicity is often seen as a key success factor (Zhao and Rosson, 2009), microblogging is not without controversy. Critics argue that Twitter stimulates creation of a huge mass of meaningless information (Pear Analytics, 2009). Specifically, many company executives are agitated by fearing that the application of microblogging might lead to importing to their intranets “the great chatter” associated with Twitter, which would cause time wasting and productivity losses to the company. In this context, Günther et al. (2009) pointed out that, for many executives, the fear that EMB will ultimately lead to as they call “signal-to-noise” ratio “You will get flooded with information” (Günther et al., 2009), which is a main reason not to implement it in a corporate intranet. Notwithstanding, there are also many positive voices (Lynch, 2009). In 2008, Gartner, technology research specialists, has added microblogging to his hype cycle. He forecasted a sharp rise in popularity and stated that, by 2011, microblogging will be featured in 80% of enterprise social media platforms (Gartner, 2008).

Riemer and Richter results have implications for decision makers. Authors contradict opinions that microblogging apply in a corporate context can lead to the importing of unwanted behaviors. Theirs findings show that microblogging as a practice is highly context dependent. Furthermore, conclusions accent that communication and collaboration systems such as microblogging platforms are open technologies, which do not limit use, i.e., the channel does determines a form of usage (Richter and Riemer, 2009). Rather, such communication medium need to be appropriated by users in a context, in this way, becoming part of rather different practices when compared across contexts. From a practical point of view, EMB will not import the chaotic and self oriented behaviors, like concentration towards personal chatter typically associated with Twitter. Managers should trust their employees to allocate the channels in useful and focused ways. However, given the nature of the Twitter technology and comparable research on similar technologies like Skype (Riemer et.al., 2007) or SNS (Richter and Riemer, 2009) there seems that social media, in general, and microblogging, in particular, can make a useful contribution to facilitating group work, especially communication, in OSS or corporate context.
3. RESEARCH FRAMEWORK

This section presents a theoretical framework (Figure 4) that will be used as a basis of empirical analysis. The theoretical framework is a theory-based answer to the research question: How microblogging is utilized by the DSD community for its communication needs? Separate parts of framework answer to the three supporting research questions: What part of the communication process need microblogging constitute in the DSD community? How microblogging affects communication performance in the DSD community? What types of information are transmitted throw microblogging in the DSD community? The framework goal is to analyze what characteristics describe communication tool, which will perform successful communication process. As this study analyzes role of microblogging in distributed software development, we deeply explore each layer of this relatively new social medium.

Framework represents microblogging analysis as a communication segment in DSD communication tool set and detail exploration of medium elements. The fundamental view of framework visualizes separate medias’, which have been used to perform successful communication process in DSD (Figure 3). From previous studies, we obtain that telephone, email, groupware, and IM where used in successful distributed projects (Tissen et al, 2007; Niinimaki, 2010; Lanubile, 2009). Earlier research claims that case studies showed face-to-face communication effects shared understandings (Heeren and Lewis, 1997). To substitute face-to-face interaction in distributed development practices electronic media such as audio or video conferencing started to be used for this purpose (Cataldo et al, 2006).

Based on the literature review, we obtain that social media is started to be used in DSD (O’Reilly, 2005). The main aim of this thesis is to select microblogging as one of the social media type. Then we zoom in and focus on the microblogging breakdown to six fundamental layers. Every medium used in DSD communication framework (Figure 3) obtains those six frameworks only with own parameters. Framework was built in the bottom-up approach. Every layer is enumerated like L1, L2, etc. First layer, located in the base, characterizes what information processing need, regarding MRT, microblogging satisfies. Theory describes two basic needs: information uncertainty and equivocality (Daft and Lengel, 1986). Microblogging is a lean media (Kaplan and Haenlein, 2010), thus, based on literature review we can state that this media reduces information uncertainty, in the sense of absence of information.
Regarding MST characteristics, communication process is composed of two fundamental communication processes: conveyance and convergence (Dennis and Valacich, 1999). The second framework layer (from the bottom) describes, what of one of two, communication processes medium performs. Microblogging is considered to obtain low media synchronicity, so second layer obtains conveyance parameter. Based on medium characteristics, it is applied to exchange information without having focus on the same time or without agrees.

The third layer describes media capabilities and is based on the MST, MRT and CG factors. Three parameters were derived from mentioned theories. The first parameter is transmission capability. Based on literature review, parallel is obtained by asynchronous data exchange capability and led to achieve successful communication process (Dennis and Valacich, 1999). The second parameter in this layer is data processing capability. If data is asynchronous, this means that before sending data medium should provide the possibility to review and modify it that is why we present rehearsability parameter. After data has been sent, there should exist possibility to review messages, known as a reprocessability factor, which increases communication process performance (Dennis and Valacich, 1999). In turn, regarding transmission capability, if there exist several parallel conversation, there should also exist sequential flow of messages on the same topic. Based on CG factor we introduce sequentiality factor, which is included into data processing parameter (Clark and Brennan, 1991). The third parameter shows media ability to present and manage various types of symbols. Based on literature review, (Daft and Lengel, 1986; Dennis and
Valacich, 1999) there have to be substituted rich symbol sets like tone of voice or eye contact with symbols provided by parameters which support reach communication in distributed teams using microblogging.

The fourth framework layer presents five microblogging features and strengths to perform high level communication between distributed team members. Microblogging by its nature is a form of “mini” blogging where message length is up to 140 characters. With this feature, medium obtains ability to process data as it provides ability easily to review and understand short messages (Zhao and Rosson, 2009). Messages can be read and written from any device connected to Internet enabling constant access to information and mobility for users. Consequently, mobility invokes spontaneity (Treude and Storey, 2010) of messages and informal style of text as tweeting person do not spend much time for rethinking message content. The most specific microblogging technical feature is its broadcast nature – ability to sent content to dispersed audience (Zhao and Rosson, 2009).

The fifth framework layer presents communication genres. All interaction types are divided to work-related and personal life information exchange. The daily charter information is less relevant for this study; however previous studies distinguished that users mostly talks about family (Java et al 2007), or more usually personal emotions, interests. (Naaman et al., 2010). Research about work-related topics has various outcomes (Guzzi et al, 2010; Bougie et al, 2011), we have selected most popular, and success oriented genres which best describes microblogging as a communication tool. Accordingly, researchers outlined help requests (Zhao and Rosson, 2009), to do’s (Guzzi et al, 2010), daily updates (Java et al 2007; Bougie et al, 2011), and assorted types of information shares (Bougie et al., 2011). Also, these genres where highlighted by the EMB studies (Riemer and Richter, 2010). Help request is based on the team members seeking to ask experienced colleagues for suggestions or solutions of faced issues, feel the gap of missing information (Zhao and Rosson, 2009; Riemer and Richter, 2010).To do’s are based on coordination, where specified or general tasks are delegated (Guzzi et al, 2010; Riemer and Richter, 2010). Daily updates and information share are based on the last events, up-to-date product varies and structural reorganizations, which are usually accompanied by links to information source (Java et al 2007; Riemer and Richter, 2010; Bougie et al, 2011).

Finally, the upper level of the framework is dedicated for social aspects provided by microblogging. In distributed team members without meeting each other have to build relationships to share knowledge and create software. Trust, connectedness, reputation, and self-presentation were selected from literature, as a most relevant for the analyzed medium and communication needs, to strengthen social ties. Feelings of mutual trust are likely to form connected group. Thus, an increased level of informal communication throughout social network at work may increase the chances of sharing and collecting valuable information (Zhao and Rosson, 2009). Feelings of connectedness might be facilitated by proximity, similarity, and exposure to personal events. The real time personal updates found in microblogging may help support a feeling of connectedness (i.e., being there), and provide possibilities to explore similar experiences and relationships with each other (Zhao and Rosson, 2009). Social media has a strong influence over the reputations of the participants, which in turn has career consequences within an organization or community (Begelet al 2010b). Self-presentation (Kaplan and Haenlein, 2010) describes the process of public image formation of each team member or organization and increases the level of reputation.

Besides, each framework’s lower layer explains the following layer characteristics’. Thus, it strongly supports the idea of bottom-up approach model. The lowest layer based on MRT presents microblogging as a lean medium; consequently, we can claim that this medium can solve uncertainty problem in a communication process. Accordingly, based on MST we know that it is asynchronous medium. As it is lean medium, which solves uncertainty problem, conveyance process has to be performed to fill the information gap. At this point, we have a connection between first and second framework layers.

The further connection appears between two previously described layers and media capability layer. As lean and asynchronous media is explored, thus, such characteristics as message parallelism, feasibility of message reprocessability and various set of symbols appear. For example if participants perform asynchronous communication, which substitute’s face-to-face approach and they need to share some knowledge, thus they will use several representations of symbols (e.g. text and numbers). Before sending information sender will be able to review message to check for errors. Receiver will be able to review
message even after some time. Additionally several topics can be run in parallel, but messages for each topic would be viewed and saved sequentially.

Consequentially to support previously presented layers technical features are required. Microblogging possesses unique technical characteristics to maintain lean and asynchronous data. Those characteristics are - broadcasting short, informal and spontaneous messages from any kinds of devices connected to the Internet to appropriately process data.

Interaction type layer is based on previous four layers. Based on media capabilities, technological features and communication process abilities microblogging users employ medium to their daily personal and work needs. Their decision on how to use medium and what information to exchange depends on what functionalities are presented and how easily they can be adopted. Of course, this layer adapts to each organization need, however, regarding the fundamental distributed software development process tasks (Ebert and Neve, 2001) we obtain typical layer for any DSD organization.

Finally, those message categories influence strength of social aspects between team members. More personal information creates connection and ties between each team member as they can self-present. Based on those connections work-related topics contributors obtain trust and build personal reputation. In concert, team members build organization image in public and that is how top layer is connected with all rest framework layers.

Next the theoretical framework is validated against empirical data. This is done in Section5. However, before that the empirical research design and the methodology used in it are described in Section 4.
4. RESEARCH APPROACH

The purpose of the empirical part of the thesis is to explore phenomenon through the theoretical framework. This section describes the study context and the research methodologies.

4.1. CASE CONTEXT

Case study research is appropriate for the study of software development and use within organizations (Darke et al., 1998). This study applies a single-case study approach to investigate microblogging use in DSD phenomena in depth to provide a rich description and understanding. The unit of analysis identifies what constitutes a “case”. The unit of analysis can be individual, group or organization. This study explores Drupal community as a single-case study. Drupal is an open source content management system, which performs content publishing and managing, on a website. The built-in functionality of Drupal can be extended by existing amount of modules. Drupal community has over 800,000 members and over 17,000 developers in more than 200 countries. The fact is that community is still growing in the sense of active members (Figure 6) and that the product is in demand (Figure 5). Drupal has had 7 releases up to now and more than 16, 000 modules.

![Figure 5: Project usage](http://drupal.org/project/usage/drupal)

At any given time, there are two major supported Drupal releases. At this moment, these are Drupal 6 and Drupal 7. Updated versions are issued on a regular basis. For example, within the Drupal 6, several versions have been and are still released like 6.0, 6.1, etc. These versions are sometimes generally referred to as “6.x”. Drupal 8 is currently in the early stages of development, a stable version is likely to be released in late August 2013. Once that happens, Drupal 6 will no longer be supported and the user will need to upgrade Drupal 6 sites to Drupal 7 at that time. New stable release of Drupal core is scheduled for the last Wednesday of each month. A new release will only come out on this date if it is contain enough important bug fixes to warrant a release. If a release is warranted, a “code freeze” will be called a few days prior to the release date, and the impending release will be announced via the @drupal Twitter account and Planet Drupal.

![Figure 6: Developers’ increase](http://drupal.org/documentation/version-info)

---

1. [http://drupal.org/project/usage/drupal](http://drupal.org/project/usage/drupal)
3. RSS feeds about all Drupal news.
Diagram (Figure 6) presents the number of developers who made their first commit in the period from middle of July 2010 till middle of March 2012 (1 year and 9 months). The big jump of commits on March 2011 is explained by two causes. Main reason is the migration from Concurrent Versions System (CVS) to Git, this simplified commit process and increased number of developers. The second reason was a global annual Drupal conference (DrupalCon), where new commit process was presented, and participants and newcomers started to use simplified process.

The second diagram (Figure 7) presents the number of issues in the period from middle of July 2010 till middle of March 2012 (1 year and 9 months). The significant issue increase on January 2011 aroused because of preparation for Drupal 7.0 release. Issues increase on March 2011 was induced by previously explained community migration from CVS to Git.

Finally, diagram (Figure 8) presents the number of all commits in the period from middle of July 2010 till middle of March 2012 (1 year and 9 months). The significant increase on February 2011 is explained by the fix of the issues aroused in January 2011 (Figure 7) and also, when Drupal 7.0 was released developers presented a new set of modules. There is a visible dependence between issues number and commits number. If in a month the quantity of issues grows, in the next month the number of commits will increase. Commit number is large, because of the constantly developed and committed vast number of new modules.

4.2. DATA COLLECTION METHODS

This study is exploratory research in nature, because there are few studies about the use of microblogging as a communication media and no study about the use of this media in distributed software development. For the analysis of selected case study and use of Twitter, within this community, single case study (Figure

---

http://en.wikipedia.org/wiki/Git_(software)  
http://chicago2011.drupal.org/  
http://drupal.org/project/modules?page=120 and filters=bs_project_sandbox%3A0 and solrsort=created%20desc
9) was selected. It helps to develop intensive understanding of the events and practices of organization (in this study case it is Drupal OSS community). This research goal is to provide an interpretation of unique and typical experience of community members in the communication process (Cunningham, 1997).

<table>
<thead>
<tr>
<th>Single case study</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Collection</strong> methods:</td>
</tr>
<tr>
<td>• Interviews</td>
</tr>
<tr>
<td>• Unobtrusive measures - tweets</td>
</tr>
<tr>
<td><strong>Data Analysis:</strong></td>
</tr>
<tr>
<td>• Quantitative</td>
</tr>
<tr>
<td>• Qualitative</td>
</tr>
<tr>
<td><strong>Primary Empirical Conclusions:</strong></td>
</tr>
<tr>
<td>• Theoretical implication</td>
</tr>
<tr>
<td>• Practical implications</td>
</tr>
</tbody>
</table>

Figure 9: Research approach

Both quantitative and qualitative data were collected. Quantitative data was grounded on the tweets which were collected in unobtrusive measures - way of collecting data that do not require the cooperation of the subjects. Qualitative data was gathered using tweets and semi-structured interviews. Interviews are ways for participants to get involved and talk about their views. In addition, the interviewees are able to discuss their perception and interpretation in regards to a given situation. For the data analysis, quantitative and qualitative data analysis methods were selected. Based on the empirical evaluation results, both theoretical and practical primary empirical conclusions will be drawn. Data collection, analysis and implication methods are described in detail in the next subsections.

4.2.1. **SEMI-STRUCTURED INTERVIEWS**

Semi-structured interviews are nonstandardized and are frequently used in qualitative analysis. This type of interview does not apply to do the research to test a hypothesis (David and Sutton, 2004). The researcher has a list of key questions, topics and issues to be covered. In this interview type, the order of the questions can be changed depending on the direction of the interview. An interview guide is used, but additional questions can be asked. In this study, semi-structured interview data collection was selected to obtain feedback and offer the interviewer the opportunity to explore an issue. It allows the interviewee to express their opinions, concerns and feelings. Semi-structured interview allows the conversation to flow where it needs to in order to deal with issues, as opposed to cutting someone off because they stray from the topic. Additionally, questions can be asked which have not been anticipated in the beginning of the interview. This interview type gives the opportunities to probe for views and opinions of the interviewee. Probing is a way for the interview to explore new paths, which were not initially considered.

The list of Drupal organization contributors was reviewed to identify key informants. There was selected I1 as a member with one of the longest experience, core developer (I2), and evangelist with reputation (I4). Succeeding interviewees were suggested by I1, I2, and I4 as relevant and potential interviews. There were decided that 6 community contributors can cover all topics of interest in this study. Regarding the literature, OSS development teams have a hierarchical or onion-like structure (Crowston et al., 2003). At the centre are the core developers, surrounding the core are the developers, surrounding the developers are the users. Active users contribute bug reports or feature requests, write documentation or translate the system. Thus, the selected subset of interviewees cover OSS development team hierarchy, because we have interviewed core developer, developer, tester, document writer, and all of them are users at the same time. General interview questions and topics (Appendix 1) were prepared after literature review analysis, setting research questions and quantitative data collection. Questions with interview plan and conditions were sent prior to interview day. These semi-structured interviews lasted averagely from 60 to 90 minutes. The respondents are located in different countries, so to conduct interview we have used VoIP application Skype\(^\text{13}\) to obtain both video and audio conversation. To record audio CallGraph\(^\text{14}\) Skype recorder tool was used.

---

\(^{13}\)http://www.skype.com/intl/en/home

\(^{14}\)http://scribie.com/free-skype-recorder
Interviewees were asked to describe communication media tools they use to perform Drupal development tasks. In the beginning of each interview, we did not present this study interest in the microblogging analysis as it could lead the interviewees to talk only about microblogging. Beforehand quantitative data results showed that Drupal community members use Twitter fairly often (Section 4.2.2). The goal of this study was to understand what media respondents mostly use and what role microblogging plays in their chosen set of communication tools. The interview topic was presented as “Exploring communication media selection in distributed software development”. Starting from the point when interviewees themselves had mentioned Twitter as a way to communicate, we were asking more detailed questions related to microblogging use. Below (Table 2) is presented a list of interviewees, their responsibilities and experience in the community; experience with OSS and Twitter.

<table>
<thead>
<tr>
<th>Interviewee</th>
<th>Responsibilities</th>
<th>Member in community</th>
<th>Experience with OSS</th>
<th>Started use Twitter</th>
</tr>
</thead>
<tbody>
<tr>
<td>I1</td>
<td>handbook writer, tester and evangelist</td>
<td>&gt;11 years</td>
<td>&gt;11 years</td>
<td>April 2007</td>
</tr>
<tr>
<td>I2</td>
<td>core developer, module maintainer</td>
<td>&gt;8 years</td>
<td>~ 12 years</td>
<td>June 2009</td>
</tr>
<tr>
<td>I3</td>
<td>developer, module maintainer</td>
<td>&gt;7 years</td>
<td>~ 8 years</td>
<td>April 2007</td>
</tr>
<tr>
<td>I4</td>
<td>evangelist, documentation writer</td>
<td>&gt;5 years</td>
<td>&gt; 5 years</td>
<td>October 2007</td>
</tr>
<tr>
<td>I5</td>
<td>translator, support forums</td>
<td>&gt;5 years</td>
<td>&gt;5 years</td>
<td>March 2008</td>
</tr>
<tr>
<td>I6</td>
<td>developer</td>
<td>&gt;4 years</td>
<td>~ 4 years</td>
<td>October 2008</td>
</tr>
</tbody>
</table>

Table 2: General information about interviewed Drupal community contributors

4.2.2. UNOBTURSIVE MEASURES

Secondary data is the data that have been already collected by and easily accessible from other sources. Unobtrusive measures are method of gathering data that do not require the collaboration of the subjects and, in fact, may be invisible to them. Webb et al. (1966) describe these measures as “nonreactive research”, because the researcher is expected to observe or gather data without interfering in the ongoing flow of everyday routine. Data collected in this manner are categorized as archival records, documents, and physical evidence. Unobtrusive measures are particularly useful for triangulation15. As a complement to interviews, nonreactive research provides another perspective on a phenomenon, developing its complexity. These methods can be used without noticing subjects. Additionally, this data collection process is relatively easy, because it often involves using data already collected by someone else (e.g., archival records, bills).

Tweets are one of the representations of unobtrusive measures. They are not collected as e.g. sale records, but they are stored and available publicly. To collect and retrieve data from Twitter API, there were used twitteR16 package within R17 (open source statistic data mining tool). R code was used to extract tweets and create graphics. A set of definitions of the Twitter terms are provided which will be used from this point and in the next sections.

To follow someone means to have chosen to subscribe to individual’s Twitter updates. Every time an individual posts a new message, it will appear on his/her followers’ Twitter home page. If user tweets publicly, anyone who runs a search for a keyword existing in that tweet may be able to see that message. Tweets are public by default. It is also possible to protect tweets and keep updates out of search. To keep tweets public is the default account setting and to protect tweets require manual approval of each and every person who may view that account’s tweets. Public Tweets are visible to anyone, thus protected tweets may only be visible to approved followers. If a message begins with @username, meaning it was directed to

---

15 Indicates that data validation is performed from more than two methods
16 http://cran.r-project.org/web/packages/twitteR/index.html
17 http://www.r-project.org
another user, it is a @reply. By clicking the Reply button on another person's tweet user replies to it. A retweet is a reposting of someone else's tweet. Retweet feature helps to share post with all of individual’s followers. Sometimes people type "RT" at the beginning of a tweet to indicate that they are reposting someone else's content. This is not an official Twitter command or feature, but signifies that one is quoting another user’s Tweet.

There are some limits on account actions like maximum 1,000 updates/tweets per day. The daily update limit is further broken down into smaller limits for semi-hourly intervals. Retweet is counted as updates. The # symbol, called a hashtag, is used to mark keywords or topics in a tweet. Hashtag is made up of a single word that begins with the ‘#’ symbol and are placed anywhere in a 140-character post. It was created organically by Twitter users as a way to categorize messages. There is an option to post video, media, pictures directly to Twitter, with the option to add a message. A timeline is a Twitter term used to describe a collected stream of Tweets listed in real-time order. User at once can see 20 most recent statuses posted on his timeline. Using twitteR package it is possible to extract maximum user’s most recent 3,200 tweets; this maximum is provided directly by Twitter developers.

<table>
<thead>
<tr>
<th>@name</th>
<th>Focus Group</th>
<th>Creation date</th>
<th>Number of Followers</th>
<th>Number of Tweets</th>
</tr>
</thead>
<tbody>
<tr>
<td>drupal</td>
<td>Main Drupal account</td>
<td>25-Mar-07</td>
<td>30975</td>
<td>1356</td>
</tr>
<tr>
<td>drupal_modules</td>
<td>About Drupal modules</td>
<td>19-Jan-09</td>
<td>7835</td>
<td>12902</td>
</tr>
<tr>
<td>drupalplanet</td>
<td>Announce about Drupal(code, marketing, infrastructure, etc.)</td>
<td>19-Feb-09</td>
<td>6767</td>
<td>5095</td>
</tr>
<tr>
<td>drupal_security</td>
<td>Drupal security advisories</td>
<td>29-Feb-08</td>
<td>4315</td>
<td>338</td>
</tr>
<tr>
<td>drupalcore</td>
<td>Drupal core major happenings and opportunities to contribute</td>
<td>24-Aug-11</td>
<td>2764</td>
<td>94</td>
</tr>
<tr>
<td>drupaldocs</td>
<td>Drupal documentation</td>
<td>16-Jul-10</td>
<td>935</td>
<td>294</td>
</tr>
</tbody>
</table>

Table 3: Drupal community popular Twitter groups (in mid May 2012)

Above is presented information about six Drupal group accounts (Table 3), which were suggested by Drupal.org or were considered as most popular groups; popularity was evaluated by followers’ number. The total amount of extracted tweets was 6,821 in the period from middle of October 2010 till middle of March 2012 (@drupal_modules and @drupalplanet) and in the period from middle of July 2010 till middle of March 2012 (rest of the groups from Table 3). We have obtained two different periods, because tweeting frequency of each group is abundantly distributed. Tweets were also collected from six interviewees (Table 1), and two core developers (Table 4) to perform comparison of collected qualitative data with participants’ statements during the interviews.

<table>
<thead>
<tr>
<th>Core developers</th>
<th>Responsibilities</th>
<th>Member in community</th>
<th>Experience with OSS</th>
<th>Started use Twitter</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>Permanent core developer</td>
<td>&gt;11 years</td>
<td>&gt;11 years</td>
<td>March 2007</td>
</tr>
<tr>
<td>C2</td>
<td>Drupal 7 and 8 core developer</td>
<td>&gt;7 years</td>
<td>&gt;7 years</td>
<td>March 2007</td>
</tr>
</tbody>
</table>

Table 4: General information about Drupal core developers

Below is presented table (Table 5) with the number of followers and tweets of Drupal contributors. Totally we have obtained 8834 tweets from all contributors in the period from middle of December 2010 till middle of May 2012.

---

18. Tweets that start with an @ and then a username. Also, an @reply is any update posted by clicking the "Reply" button on a Tweet.

19. https://dev.twitter.com/docs/api/1/get/statues/user_timeline

4.3. DATA ANALYSIS

Qualitative research is a type of scientific research. In general terms, this research consists of an investigation that seeks answers to a question, systematically uses a predefined set of procedures to answer the question, collects evidence, produces findings that were not determined in advance, produces findings that are applicable beyond the immediate boundaries of the study (Miles and Huberman, 1994). Additionally, it seeks to understand a given research problem or topic from the perspectives of the local population it involves. Qualitative research is especially effective in obtaining culturally specific information about the values, opinions, behaviors, and social contexts of particular populations. Quantitative and qualitative data analysis methods were employed to analyze the collected data in each phase of the research process. This section deals with how the data analysis was performed in the empirical evaluation part of the thesis. Langley’s (1999) process based research strategy for analyzing quantitative and qualitative process data was used.

4.3.1. QUANTITATIVE ANALYSIS

Quantitative research study produce results that can be used to describe numerical changes in measurable characteristics of a population interest, and that can be generalized to other similar situations (Salkind, 2010). Provide explanations of predictions and explain causal relationship. Quantitative methods include the target group and the group from which data are collected. Regarding case study Drupal community is a target group and data is collected from a small target group which consists of the interviewees, two core developers and Drupal account groups. To explore collected quantitative data exploratory data analysis is used, which is conducted using graphical techniques such as plots and data summaries. Exploratory analysis provides insight into the underlying structure of the data. The existence of missing cases, outliers, unexpected or interesting patterns in the data can be checked with exploratory procedures. The relationships between events within one case were performed. The study used extracted tweets from personal accounts of Drupal contributors (Table 5) and Drupal group accounts (Table 3).

The data obtained from particular timestamps was analyzed using descriptive and inferential statistics. Numerical descriptors including mean and standard deviation were applied to evaluate information amount the contributors handle every day. For the Drupal contributors’ and Drupal group accounts, a graphical tweet frequency analysis was performed. Statistical methods such as correlation calculation were used to analyze the relationship between frequency of retweets and the number of followers (Section 5.2). Percentage was used to describe categorical interaction types data (based on Section 4.3.3). Meanwhile tweets with “#drupal”, their frequency, and amount of retweets were analyzed, and a keyword dendogram was created. Additionally, other relevant data obtained from the Drupal community (such as amount of

<table>
<thead>
<tr>
<th>Collaborator</th>
<th>Number of Followers</th>
<th>Number of Followees</th>
<th>Number of Tweets</th>
<th>Number of tweets sent per day</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>18,404</td>
<td>8</td>
<td>5,965</td>
<td>3.2</td>
</tr>
<tr>
<td>C2</td>
<td>7,624</td>
<td>1,192</td>
<td>2,849</td>
<td>1.5</td>
</tr>
<tr>
<td>I1</td>
<td>2,203</td>
<td>1,317</td>
<td>77,222</td>
<td>41</td>
</tr>
<tr>
<td>I2</td>
<td>2,177</td>
<td>163</td>
<td>4,711</td>
<td>4.5</td>
</tr>
<tr>
<td>I3</td>
<td>1,714</td>
<td>91</td>
<td>4,386</td>
<td>2.39</td>
</tr>
<tr>
<td>I4</td>
<td>1,305</td>
<td>1,801</td>
<td>1,841</td>
<td>1.15</td>
</tr>
<tr>
<td>I5</td>
<td>563</td>
<td>624</td>
<td>1,146</td>
<td>0.75</td>
</tr>
<tr>
<td>I6</td>
<td>314</td>
<td>193</td>
<td>2,283</td>
<td>1.84</td>
</tr>
</tbody>
</table>

Table 5: Drupal community members (in mid May 2012)
commits, issues, etc.) was analyzed to decide the dependency between development information and tweets during the selected period.

4.3.2. QUALITATIVE ANALYSIS

Coding and categorizing strategy (Flick, 2007) was used to analyze the transcripts of the semi-structured interviewees. The main activity is to search for relevant parts of the data and to analyze them by comparing them with other data and by naming and classifying them. The first level of coding was based on identifying themes, units of meaning such as words, sentences. Then the interviewee statements were grouped by the topics. Finally the data analysis was performed by exploring interviewee views and perceptions of situations when microblogging was used. Investigation of the theoretical framework layers and its parameters was based on interviewees’ claims and illustrated example. Interviewees’ claim comparison within the case promote to understand what does the interviewees say about different tasks performed and tools used to perform them, how consistent/contradictory the statements across several categories are.

Kautz et al. (2001) conclude that the combination of several perspectives provides for researcher a better understanding, which would not have been possible if only one perspective has used. Thus, although each of theoretical perspective would be imprecise on their own, common exactness can be provided by a certain combination of these different theoretical perspectives. In this study, we have analyzed microblogging use as a communication tool from two angles of existing communication theories. Firstly, we have explored it by looking from the Media Richness theory viewpoint. MRT claims that communication media differ in the ability to facilitate understanding depending on the amount and richness of information. A medium is considered “rich” when it has the ability to convey natural language rather than numeric information in different ways of conveyance and to personalize a message. Based on the MRT characteristics, we have defined microblogging as a lean media with all the properties describing this tool. On the other hand, MST states that group communication process is composed of two fundamental communication processes: conveyance and convergence. Conveyance process is the transfer of a variety of new information to enable the receiver to create and revise a perception of the situation. Convergence is the development of shared meaning for information. Consequently, microblogging obtains definition of asynchronous tool with a set of characteristics corresponding MST properties. By providing two different viewpoints, we proposed a new view of microblogging and presented two capability layers in the theoretical framework obtained from MRT and MST, which can improve communication performance used in the distributed software development workplace (Section 5.2).

To have rich tweet analysis results open coding strategy from grounded theory is used. Coding is a process of simultaneously decreasing the data amount by dividing it into units of analysis and coding each unit (Langley, 1999). The categories may have several subcategories and properties or other characteristics. As the categories are developed, main categories appear which integrates all the theoretical concepts into “grounded” original evidence. Coding helps to gain a new study perspective. Open coding proceeds through each line examining and then defining actions within coding. In this study, we have performed coding strategy analysis to develop interaction types layer from the data as it is collected and analyzed. In the theoretical framework (Figure 4) the interaction types layer was created based on literature review, however, we decided no to test it, but perform case genre analysis based on open coding strategy. The process of analysis took place from unobtrusive measures – tweets. In the study, coding was based on a qualitative evaluation of each tweet of every contributor (Table 5). Once data have been collected, the next stage was to build an indexing system for the data. The open coding was proceeding. Each tweet was coded based on context and appearing keywords. When all tweets were coded, the next step of creating more general categories followed. The generalization process was repeated several times until two main categories and six subcategories of interaction types were developed. The aim of this step was to build up a group of relevant genres used in the Drupal case.
5. EMPIRICAL FINDINGS

This section describes the findings of the empirical research. The section proceeds so that, the first Section 5.1, briefly presents how communication process is performed in Drupal community and in which cases community members prefer to use microblogging as a communication media. Section 5.2 identifies how the characteristics presented in the theoretical framework (Figure 4) of a microblogging use in distributed software development were applied in the project. This is done by exploring whether the Twitter provided all functionalities and aspects proposed in the theoretical framework and by identifying benefits of applying this media type. For example, if a community member had to announce global event, did they use microblogging? Or if a developer needed help from more experienced colleague, what kinds of communication tools he used to ask a question? Sections 5.3 and 5.4 summarize the empirical findings.

First, Section 5.3 compares the aspects mentioned in the theoretical framework and concludes what role microblogging plays in the Drupal community. Then Section 5.4 identifies the primary empirical conclusions.

The text contains quotes from what the interviewees said in interviews. After each quote, there is a reference telling who said the comment. The references are declared in Table 2. As it can be seen, I1 refers to first interviewee, I2 – to second, etc. In the next columns of Table 2, there is information about interviewees’ such as responsibilities, experience in Drupal organization. The same references are used inside of the quotes. All the names have been removed from the quotes. The text also contains quoted tweets, which are presented in this format - [tweet] @username (date and time) - square brackets contains tweet context, after follow author username reference (Table 3 and Table 5), tweet date and time. Additionally, starting from this point word “microblogging” become “Twitter” synonym; words “post” and “message” become “tweet” synonyms.

5.1. COMMUNICATION PROCESS IN CASE PROJECT

The Drupal project is performed by volunteers, whose daily work routine includes development, testing and documentation tasks. Additionally, the contributors perform organizational tasks like event, release announcement. As the interviewees have stated, in Drupal every task has its own procedures. There is no strict hierarchical structure, so everyone works in the sphere he prefers. However, Drupal has strict distribution of workforce into departments (teams), where each member is responsible for the function, for example, "core committers, (...) security team, (...) documentation team, etc."(I1).

5.1.1. COMMUNICATION TOOLS USED TO PERFORM SOFTWARE DEVELOPMENT TASKS

Primarily this study analysis starts from the multivariate framework proposed by Leavitt (1965), also called Leavitt’s diamond, which consist of the four variables: actors, structure, task, and technology. Actors are part of the change process. Structure reflects the system of authority and power. Task defines the job to be done by the actors. Technology represents techniques, tools, etc. by which work is done. These four elements (Figure 10) are interdependent and change in one will change the state of each of the remaining three variables.

![Figure 10: Leavitt’s diamond (Leavitt, 1965)](image-url)
To have a better understanding what is microblogging’s (as a technology) role in Drupal community (as a structure) the interviewees (as actors) presented a list of tasks they have usually performed. The interviewees presented how they have contributed to the project and how they have communicated. Every development process involves people, who are grouped by their responsibilities. Regarding interview outcomes and information found on official Drupal website (Drupal.org), we highlighted a set of tasks and tools used to perform communication.

The main task in software development is based on code contribution. In Drupal there is a small set of core developers, usually two people (C1 – is the main core developer, and each new release second contributor is changed). Additionally, there is a small team that reviews proposed changes and maintains code. Also, there are a vast amount of system contributors and module creators, because it is module based system. Additionally, developers perform new release coding, old version support, answer to the questions of their users and provide help. For example, I2 was a Drupal 6 core maintainer with C1 and now in Drupal 8 he maintains multilingual initiatives, I3 is tracking Drupal 8 issues, and he is core maintainer of node access subsystem, I6 created a patches and feature requests. Thus, everyone performs several tasks in parallel like developing for a new release and maintaining for the previous one:

“people are coming into my issue queue and posting bugs, or asking for some help” (I3)

During the development phase users usually need help from more experienced developers, so when they are seeking for help they are suggested to look for it on IRC, Drupal.org or other tools including microblogging. Any software development process cannot be done without testing. Testing phase is vital for Drupal project as in any other software development processes to find bugs, to check functionality vs. requirements and users expectations. One of the interviewees has been involved in testing. He divided testing to three phases: bug finding, usability checking and issue solving. He also stated that communication performed during the testing phase was done using emails, IRC or Drupal.org.

Finally, documentation of working system and modules is enclosing the process of development. Two interviewees (I1, I4) have performed documentation tasks. Initially, the process of how documentation should be performed was not described and any motivated individual could document technical part, but “now it is maybe a couple of hundred of people”(I1). When the community became larger, the processes came across documentation process, and documentation group was created. Document writers mostly interact with engineers and developers to make sure that they understood the technical sight and can properly describe functionalities. The interviewees have mentioned different collaboration tools used during documentation phase like face-to-face, emails, issue queues. Microblogging is described to support communication between document writer and other contributors to receive review and feedback.

“when I contribute to documentation, if it is new documentation, usually, I write it up and then either tweet about it or post an issue asking for feedback”(I2)

As the Drupal development and its use are distributed in different countries, the product requires to be translated into different languages. There are created tools for translations and code contributors have only to manage task distribution. Also, the community contributors take a part in local and global event organizations. Every interviewee outlined that personal communication is necessary to create connection between distributed team members. As one of them mentioned “enough to meet one time and have an understanding of a person” (I3). That is why Drupal community so intensively organizes global and local events to share information and develop new knowledge.

Participation in OSS development is volunteering work. Thus, it was essential to understand how much time users spent working for community and how much, in percentage of that time, they spent communicating. Another goal was to analyze what tools they usually prefer to use and how they rate them by the use. Averagely the interviewees spent about 10 hours per week working for Drupal, their tasks are decidedly different and even the same person can have extensive distribution of tasks. Also, as I6 explained it was difficult to clearly separate conversation and work activities “we are communicating plus we are working at the same time” (I6). Averagely the interviewees spent about 50% of work for Drupal by communicating. To the question what tools they prefer to use to reach distributed team members, the
The interviewees mentioned different ways, but the main set of used tools was build from issue queues, IRC, email, VoIP (usually Skype) and social media (usually Twitter and Facebook). The majority of the time was spent in issue queues (around 33% of the time spent for communication), the minority of the time in writing emails (in the range of 5-10%). Other mentioned tools were used interchangeably and were selected depending on the communication process, thus, synchronous (IRC) or asynchronous (Twitter) communication is a need to be performed.

Concluding this section, we state that microblogging is used relatively often to reach people. Thus, online availability of developers’ simplifies fast communication on demand. Based on interviews we conclude that tweets are used for announcement and information presentation for collaborators.

### 5.1.2. MICROBLOGGING USE IN DRUPAL COMMUNITY

The interviewees have stated that microblogging is part of their Drupal work related communication process. They averagely are spending about 20% (2 hours) of time for Drupal community per week by being available on Twitter and making “couple tweets per day, depending what is going on” (I5). We decided to explore intensiveness of posts and try to understand how individual’s tasks and responsibilities are reflected (if they are even reflected) on tweets. There were gathered tweets from 6 interviewees (I1 is tweeting so frequently that was impossible to draw his data with other interviewees or even find some behavioral patterns, because he has about 500 tweets in a month) and two current core developers (Table 5). Period was selected from middle of December 2010 till middle of May 2012, so we got a year and a half period. As we can see from the diagrams (Figure 11 and Figure 12) each interviewee has tweeted with different frequency but the tweeting behavior pattern can be obtained.

The interviewees have stated that they use the same Twitter account for work and personal life, so we have extracted all tweets and analyzed proportion between personal and work-related tweets. There were extracted 5805 tweets (Figure 11) for the period for one and a half year from six interviewees and 3029 tweets from C1 and C2 (Figure 12), totally we have obtained 8834 tweets. All tweets were labeled using open coding strategy. It was found that averagely 70.59% of all tweets were strictly Drupal related, because they contained #drupal hashtag. I1 has tweeted the least amount of work related tweets (21.6% out

---

21It is possible to extract only 3200 public tweets per user; if a user has protected some of messages, we could not retrieve them.
of all his tweets) compared with the other interviewees. On the other hand, C1 and C2 obtained largest work-related tweet amount (both have almost 84% out of all their tweets).

From the middle of December 2010 till middle of May 2012, there have been two significant development events: Drupal 7 release and start of Drupal 8 development. The first alpha release of Drupal 7 was on January 15 2010, followed by six more alpha releases. The first beta release happened in October 2010. The milestone release, Drupal 7.0, was on January 5 2011, welcomed by the Drupal developer parties all around the world. The current release is 7.14, Drupal 7 version is still maintained. Drupal 8 release development started on March 10 2011. The Drupal 8 release timeline was announced on February 15 2012, now it is in the development phase till February 1 2013.

![Graph of Core developers tweeting frequency](image)

**Figure 12:** Core developers tweeting frequency

Analyzing the tweet frequency of the interviewees (Figure 11) we can state that most of them tweet in average about 50 tweets per month. I3 has the most distributed amount of tweets, the standard deviation of tweets is $\sigma = 59$ and the average amount of tweets during the selected period is 131 tweets per month. I5 tweets the least amount of tweets - 15 per month, and $\sigma = 5$. Thus, we decided to explore if the increases of I3 tweets are reflecting development events (like Drupal 7 release). Based on I3’s explanations, his tweet increase is influenced by global annual conferences, which have taken place at times. This statement is proved also by the tweets statistics: 21.32% of all Drupal related tweets were about upcoming, current or past conferences and meetings. Those annual global and more frequent local events bring all developers together, and they have time not only to have a face-to-face conversation, but also make some development together. Also during these conferences people create larger follower circle. As I4 have claimed during “DrupalCon there is a whole Twitter thing going on, there is a whole IRC channel going on”, so they not only use real time communication, but also strengthen their virtual communication ties. Comparing I3 tweet increase with the rest of the interviewees we can observe evident behavioral pattern. Tweet increase starts slowly before upcoming conference. The culmination is during the conference and slow decrease happens after the conference.

By analyzing the core developers tweet frequency (Figure 12), C1 as a product creator and main core developer is tweeting in average about 100 tweets per month and usually only about the work-related topic (84 % out of all his tweets). I1 described him as “/C1 tend to be talking only about Drupal and a bit about personal stuff”. I1’s significant tweets increase in January 2011 is explained by the Drupal 7 official release. Then second increase of tweets is on March 2011, because he announced official Drupal 8 start of development. After that date, there is stable amount of tweets each month. The marginal increases are invoked because of the conferences, where he provides talks and invites new people to join development group. Comparing C1’s and I3's tendency of tweets, we can observe that C1’s tweets are more development based, and I3’s are event announcement based. Thus, a strict pattern between development related topics
and tweet frequency does not exist. But, a pattern between community related events and tweet frequency does exist. Additionally, we have analyzed data from six group accounts (Table 3), frequently tweeting groups are presented in Figure 13, in average they have from 100 till 600 tweets per month. Figure 14 presents the tendency of the rest group, accounts which does not exceed 100 tweets per month.

![Figure 13: Groups tweeting frequency](image)

@drupal_modules group is sandbox (experimental) and the full module websites\(^{22}\) RSS feeds, which announces about newly created modules, updates of old one, ask for help, etc. Most increase of tweets in this group is in January and February 2012. In February 2012 C1 announced about upcoming Drupal 8 timeline, which caused interest and motivation to create new modules and prepare old ones for upgrading. As issues have increased in January 2012 (Figure 7), then next month issues from January were fixed and number of fixes increased number of commits in February (Figure 8) supported by tweets:

\[
\text{[Drush Entity: As a result of our effort for the Drush issue \#1005480: Drush entity support we decided to make it... http://t.co/Uqtpq3]} \text{ @drupal_modules (2/2/2012 11:01:18 AM)}
\]

@drupal_planet group posts information about general changes and issues in Drupal development, Drupal.org structure, events, releases, feature changes, or retweets from other groups like @drupal_modules after filtering most valuable information. Consequently, tendency of @drupal_planet reflects tweeting frequency of @drupal_modules in a more compressed way.

By analyzing the rest group accounts (Figure 14), there is a significant @drupaldocs tweet increase till March 2011 and then a slow decrease. This culmination is because of Drupal 8 development start and that on the first week of March there was DrupalCon, so people were discussing upcoming event and work Sprints:

\[
\text{[Went to my first Drupal doc sprint today - thanks @arianesays for on boarding me! #drupalcon#drupalchi”]} \text{ @newcomer (3/11/2011 10:01:11 AM)}
\]

After the conference, newcomers actively joined and contributed documentation team in the first weeks. The same behavior pattern was observed before Drupalcon 2012 on March 2012, with a significant increase before the conference and a slow decrease afterwards. This users’ behavior proves the statement that newcomers start to look for initial information in social media groups like Twitter “generally I would suggest for people[newcomers] to search [information] in social media for Drupal activities” (I4).

\(^{22}\)http://drupal.org/project/modules and http://drupal.org/node/1011196
However, the big increase of @drupal tweets in August 2010 and then fast decrease in the next month does not have any reasonable explanations. Based on the obtained amount of data tweets sometimes can not reflect the real situation in the development or organization itself.

Finally, Drupal hashtag (#drupal) usage over some time period was explored (Figure 15). Due to the Twitter limitations\(^23\) we have extracted 1,500 tweets for two periods for March 23-26 and June 18-20. Tweets were selected from all users (even those who are not community contributors) who used hashtag in their tweets. First of all, we can see that usage of #drupal increases over time. Second outline is that a significant increase on 26 of March happened, because after DrupalCon on March 19-23, the community members started to discuss Drupal 8 progress; new commerce job vacancies appeared (more detailed topics description is presented in Section 5.2.5).

Summarizing this section, we can state that group tweets does not have rich explanatory power as they are usually links to the information source. On the other hand, personal tweet content can explain the increases of tweets, which are usually based on community events, product releases and other product changes.

\(^{23}\) It is possible to extract maximum 1,500 tweets containing a hashtag (https://dev.twitter.com/discussions/6147)
5.2. CHARACTERISTICS OF MICROBLOGGING IN COMMUNICATION CONTEXT

This section identifies how different characteristics of a microblogging layers’ were exploited by community members’ during the communication process. Each characteristic is discussed in its own subsection. The idea is to examine how different practices defined in the theoretical framework (Figure 4) were applied by the case project members.

5.2.1. INFORMATION PROCESSING

Information processing task is responsible for changing participant understanding within the time interval. Drupal community faces collaboration challenges, when members have to ask a number of questions and acquire more information to learn the answers. Information flow is enormous because of the constant increase of new users (Figure 6), who creates new modules, develop new additions (Figure 8), etc. Regarding interviewee statement “sometimes I get this feeling that I am listening to so many voices” (I6) describes existing critical situation in such scale organization. To be able to monitor all existing changes, upcoming events and be up-to-date, a member should use media that would help him/her collect all needed information and provide the possibility to retrieve it in a reliable amount of time. All these issues describe the definition of uncertainty task, because the absence of information persists in the community.

As I6 states he prefers Twitter as a reliable media source to reduce uncertainty problem, because “I can maintain, I can follow and unfollow people”. Thus, microbloggers and their followers form a unique and informal communication network where microbloggers and their followers are “nodes” and the “following” relationships between them are “links”. In this network, followers select one or more microbloggers with whom they may share similar interests, experiences, or opinions about OSS development and whom they wish to follow. Also, following process is one directional communication. That means if user A is following user B he can directly read all B’s posts, it is not required that user B will follow user A. Therefore, the follower network provides a stable infrastructure that enables developers constantly receive a stream of OSS development information from the set of microbloggers whom they selected to follow.

To present different network of participants based on a conversation, we have extracted 50 last tweets contained two specific topics. The size of the vertex is proportional to the number of followers of user, and the color of vertex represents the number of tweets found in the search results by that user (the darker color – the larger amount of tweets). Each user is marked with U1, U2 etc., each group is marked with G1, etc. First network (Figure 16) contained “@[I6 nickname]”, thus conversation was directed to I6 or tweeted by him personally. We expected to see a tightly knit community of users who all seem to know each other (at least virtually), sending each other personal or work updates. We get results as we expected, because every user was connected either by symmetric\(^{24}\) (represented by ⟷ links) or asymmetric following\(^{25}\) (represented by ⟷ links). In the network (Figure 16) we have obtained I6 followees and followers and two Drupal group accounts.

\(^{24}\)A follows B and B follows A  
\(^{25}\)A follows B, but B does not follow A
The second network was based on the conversations containing “#drupal” hashtag (Figure 17). We expected to see platoons of separate not connected users, as “#drupal” hashtag is used not only by community members, but by a wide range of users (Figure 15). Again, we obtained expected results, most of the users were not connected (thus they were users, newcomers, interested in) and those who were connected were personal accounts of Drupal contributors (including II interviewee) or Drupal group accounts.

Uncertainty task (that means a low equivocality task), describes case, when the task is analyzable, but there is a lack of information. Independently on the current development status there is a different amount of information needed. Even if tweets redirect user to more extensive source, the amount and complexity of the information can be remarkably different. It can be just a short summary of done work: [quick summary of #drupal.orgD7 porting sprint, day 1: http://t.co/VeEFyLQ] @C2 (4/24/2012 8:34:12 AM). On the other hand, it can be explanation how to upgrade module to a new version: [That is to say: #domain_access 7.x.3.0-rc1 is hot off the presses. Read http://t.co/kzaKNelibeforeproceeding! #drupal] @I3 (9/16/2011
8:15:00 PM). The members still have to convey information, deliberate, and converge on a shared set of goals, but the volume and degree of complexity will be different.

Likewise, technical problem solving is done using “rich” media type like IRC chats or video/audio conversations, so that members could move quickly to execution, whose focus is on the exchange of information. Nonetheless, convergence is required before the group can move to execution. Group well-being and member support is less clear for tasks of uncertainty. For a well established group, the members quickly pursue tasks using a mix of “lean” and “rich” media to share personal achievements and personal to do’s, personal “blocks”, cries for help, etc. For tasks coordination “rich” media (as phone conversation) is used:

“in Drupal 8 we have initiative meets that are on different topics, (..) every two weeks (..),on the phone” (I2)

However, if a new member has joined community, he can be facilitated from Twitter as the point of start, because “it is a community of people, where Drupal groups are on [different social media]” (I4). If something is not understandable or a newcomer seeks more information, he can ask questions using other channels as issue queues or using “rich” media.

Amount of developers in the community is very large and to follow most of them is impossible. Also, the information received from developers is not pure, as it mixed with personal chatter. Thus, information amount for existing members can be reduced by following official groups such as “@drupal_modules”, “@modules_sandbox”, etc. Usually, information is presented as RSS feed. Thus, user receives up-to-date, short, automatically generated posts with formal, topic related information, like tweet presented below:

[Scrollbox: Provide a tooltip box that shows up when user scrolled down the page. It can contain anything, from v... http://t.co/KUMzXOgt]@drupal_modules (12/28/2011 11:38:06 AM)

Interviewees have outlined that using microblogging and social media they can obtain answers to different kinds of questions and provide a strong base information. This is particularly useful for the newcomer, or people who have never heard about existing product functionalities or community activities. Additionally, as I1 outlined, it is a new way of free marketing “I do not think, that we have much better marketing”. Without leaving habitual social network channel users can learn about new technologies and even became a member of the community:

“a lot of Drupal groups launched on the other [social media sites]. They are really active, go there.”(I4)

Independently if new and/or incomprehensible information is sought by newcomer, or developer with long time experience Twitter seems to be valuable media choice, because it provides characteristics of “lean media”. Thus, this media increases ability to process information, because it provides short messages, which can be fast read. At the same time, information is stored, and the reader can review it whenever he wants. There are several restrictions on Twitter. One of them is that a user can at once review 20 newest feeds and it is possible to review only last 3,000 tweets. However, if a user wants to backup all tweets, he has to use third party applications. On the other hand, Twitter provides specific functionality as “favorites”, which bookmarks tweets a user would like to review later or just to store them separately. This medium uniqueness interviewee perceives as a benefit:

” [Twitter] is a little more focused than [other social media]. So, I think it [Twitter] is good for social engagements” (I3).

To receive feedback using microblogging can take a time, or user can never be answered, because tweet To receive feedback using microblogging can take time, or user can never be answered, because tweet can be lost in the stream of feeds, this is the reason to call this media “lean”. There are different solutions how to improve and get the maximum, possibly fast, feedback, however, it will never be as face-to-face. Nevertheless, to receive unknown information, Twitter provides quite various sets of symbols to process. It is possible to tweet not only text with numbers, but also put links (as was presented in previously
mentioned tweets) and pictures, and video (if it is stored in particular web hosting). Drupal contributors have adopted this medium facility to their own needs. Contributors perform visual expression of transmitted information to motivate people to read, because people prefer variation of text and graphical presentation.

The community members also are used to attach types of tags to their messages – like project (#drupal), event (#drupalcon), release (#drupal7 or #d7), etc. In projects, project leaders or teams regularly define common tags to be used. Through user tagging, content structures are being adapted to changing situations on the fly without extra effort. The project tag is used to identify the name of the OSS project mentioned in the microblogging message and is used in exceptionally wide range of context. According to I4’s opinion, hashtags are widely used: “There are a lot of people around the world, who follow, for example, the Drupal hashtag (#drupal)” (I4). The release tag classifies data by project release number, to have focused attention towards message. The event tag helps to identify the information about upcoming or held events, places, etc. Together they provide an efficient navigation process for developers to locate the OSS projects, release, events information involved in the microblogging messages. In addition, this tagging feature reflects the information structure and relationships that the microbloggers use, which is crucial for the microblog content analysis in this study.

However, tags can cause information misuse, because they can be used by anyone. Regarding analyzed participants network(Appendix 4), most of the users are not development related, thus information seeker has carefully selected information. Also, an increased amount of information should be read, which automatically increase users time abuse.

One of the major aspects in communication is to obtain message receivers personal focus, otherwise transmitted information will be simply ignored. The personal focus in microblogging can be evaluated by two existing parameters: bookmarks called “favorites” and “retweeting” functionality. By analyzing these two parameters, we can detect how many community members have read an individual’s post. If they have used “retweet” option that means that the information was paramount and beneficial; thus, they decided to share information with their personal communication network. If the followers favorite the tweet, that means they found this useful and will probably review/reread it later.

“I have seen people retweeting announcements and post what I have made, and looking on issues what I post on Twitter” (I2)

Below presented is an example (Figure 18) from @drupal account. Tweet announces two maintained and mostly used (Figure 5) Drupal 7 and 6 versions. This news is relevant to a large group of users and developers (for example, module maintainers will have to check if it is compatible with their module). That is why tweet have more that 50 retweets and several favorites (it is instant knowledge there is no need to review it later). This tweet (Figure 18) is an example of the message which reduces information uncertainty. Information describes concrete development phase as a release, with such details as a release number and date, short explanation and link to find all associated data.

Additionally, from (Figure 14) we have obtained yet another way to be aware of community news and achievements by following community accounts. There is no any strict description who and how (in the sense of the title and content) should create this type of account. However, the idea is that anyone can create it. We have analyzed only six most popular groups in the terms of the amount of followers (Table 3). For example, “@drupal” account was proposed by Drupal Community and is used to help communicate
about Drupal related topics. Another example is “@drupal_modules”, it is oriented on the current appearing modules and changes associated with existing modules “you can “follow” modules, new projects submitted to Drupal.org it is a good thing” (I6). It is enough to follow e.g., “@drupal” to have an idea of modules, monitor each channel, and the user will obtain up-to-date clean and topic-oriented data.

The number of followers has some power to state about an individual’s popularity, however, the count of retweets presents real followees interest and focus on the author’s tweets. While followers’ counts can be tricked, inflated or simply represent the number of people actively ignoring individual, retweets show people engaging with what an individual is putting out there. Having retweeted content by user network means adding exposure, additional authority, and, more importantly, that people like what user brand is doing or saying. Figure 19 presents the number of retweets’ (sum of one day tweets) of analyzed contributors’ (Table 5). Some of them are not included to the diagram, because nobody was retweeting their tweets during the selected one week period (June 01 - June 07 2012). However, we decided to calculate correlation between user followers’ number and retweets’ number. Correlation coefficient was r=0.93497, we obtained a strong positive linear relationship. Thus, in the case of analyzed contributors the number of retweets depends on the number of their followers. C1 has the biggest amount of followers, he is most popular and with the strongest reputation in the community. It is explained by the fact that he is the creator and main core developer of Drupal community.

Additionally, we have performed analysis of tweet (T) and retweet (RT) containing “#drupal” proportion during the two periods of time (Figure 19). We can observe that on 23 and 26 of March the number of retweets achieved almost 45% of all tweets, usually retweets form about 30% of all tweets with “#drupal”. These statistics present that assuredly users put personal focus on the every third message on Twitter. Additionally these retweeted posts are transferred to a wider audience. From the analyzed tweets and interviews, we clearly can state that microblogging has not been used for the equivocality tasks. Equivocality exists when there are multiple (and possibly conflicting) interpretations for information or the framework with which to interpret it. Equivocality requires negotiation among members. For this task, like document creation user uses wiki-like pages, for instant reply like issue solution – IRC is preferred, etc. Drupal contributors stated that there is a set of “rich” tools to collaborate and build new knowledge. Thus, microblogging is not used to build shared understanding and common ground - “[Twitter] is bad for community problem solving” (I3).

Figure 19: Contributors retweets’ number for the period of one week (June 01-June 07 2012)
The microblogging fulfills several characteristics of MRT as feedback and rich set of symbols. These characteristics will be analyzed in details, in media capabilities section (5.2.3).

Below presented are the primary empirical conclusions obtained from this section:

PEC1: Hashtags can be used, by anyone, to simplify information retrieval; however, they can trigger information misuse.

PEC2: “Retweet” or “favorite” functionality can be used to measure individual focus on a tweet.

### 5.2.2. COMMUNICATION PROCESS

For communication to be successful, the receiver must understand the message that sender intended to send, and both the sender and receiver must agree that the receiver has understood the message. In Drupal community, to have approval from all receivers that they have understood the message/information is impossible. Also, it is impossible to have all or even at least majority of team member focus. Community with such big and still growing number of distributed contributors’ (Figure 6) is finding facilities in microblogging functionality. Regarding the fact that information is sent with time intervals, this media is defined as asynchronous and supports communication process known as a conveyance. Thus, participants do not have to focus at the same time on the current post and do not have to agree on the meaning.

The high parallelism of social media enables people to join a conversion simultaneously, easily generating different information or idea. Rather than focusing on the same issue or topic, this multiple conversation facilitates participants to develop a shared focus and understanding. As a lower synchronicity media type, microblogging is more effective for conveyance process than for convergence process, and subsequently influences the effect of knowledge transfer.

Microblogging messages are short (140 characters), however, as analysis showed 39% of all analyzed tweets (including daily chatter) contained links. Development related post analysis presented that 56% of tweets with “#drupal” contained links. These numbers represent the community members practice of supplying posts with references to the information source.

[Drupal entity property translation storage vs. revisions keeps being a hard point. Any great ideas? http://t.co/pjxU418A #D8MI] @I2 (4/10/2012 7:44:30 AM)

In order to reply to the sender to answer or even to validate understanding of newly received data like to that presented as an example above, the followers of I2 will have to spend more time than reading 140 characters. On the other hand, the followers receive concise informative tweet, which contains label “#D8MI”. Thus, information is about Drupal 8 multilingual initiatives (department working on one of the generatrix parts of core) and with references to the detailed information.
When contributors perform exchange of information to reduce uncertainty problem, they need not only gather but also convergent tidings. That is why microblogging is presented as one of the possible tools in whole pool of existing communication tools. The interviewees stated that they needed different tools to perform information convergence. Also, they claimed that it is a personal decision to choose mediums. Community has an official website Drupal.org, where issue queues represent a main place to locate problems and solve issues. Regarding interviewees the majority of time they spent in issue queues (about 33% of their time) and other tools like IRC, VoIP’s or social medias’ have highly distributed number of use depending on individual choice (together they obtained 59% of time). The minority of time they spent by writing emails. In any case, to build knowledge or solve urgent problems respondents prefer to choose synchronous media tools (phone or VoIP’s). For example to solve problems they prefer tools providing fast feedback and low parallelism:

“If you are solving specific code situation or getting path to the issue then you go on the phone quickly just to talk it through” (I4)

Regarding previously described Drupal conventional development tasks (Section 5.1.1) we have realized that community contributors require different types of information to convey. As the group accounts table shows (Table 3), there exists information separation into organizational departments depending on people responsibilities and interests in work. Initially, did not exist hierarchical structure in the community, each member did what they need and could “I had been doing what I can do best” (I1), when community started to grow and, for example, documentation team increased from “10 or 15 people. Now it is maybe couple of hundred people” (I2) naturally raised the need to create teams and have strict workforce breakdown. Group accounts (Table 3) created in Twitter are dedicated to updating information for each team and have relevant data under one account name. Below we describe how community members take advantage of each group.

@drupalcore group is announcement group, which usually tweets core news available on Drupal.org as a RSS feeds. However, also some updates from group maintainers are also presented. Tweets usually contain links (to Drupal.org) and #hashtags. The group is relatively “young” (created in August 2011) comparing with other analyzed groups (Table 3) and has constant tweet frequency. Even if, there are some small increases, they are not supported by any remarkable development events. Half of the interviewees, even those who are not contributing to the core, are also the followers of this group.

[Test and reroll patches at #Drupal core office hours 1h from now. See you then and sorry we missed last week! #DrupalCOH](http://drupal.org/core-office-hours) @drupalcore (3/6/2012 3:10:07 AM)

@drupal_modules similarly as a @drupalcore group provides the sandbox and the full module websites RSS feeds, which announce about newly created modules, updates of old one, etc. Most increase of tweets in February 2012 (Figure 13) was related to the official announcement from C1, about upcoming Drupal 8 release timeline, afterwards interest and motivation for new modules increased. Also, there has been increased number of issues (Figure 7) from January 2011 (e.g. [@thagler weigh in on the issue. We have to solve the numeric gid problem in [node_access], #domain#drupal] @I3 (1/25/2011 1:14:00 AM)), which automatically increased number of commits on February 2011 (Figure 8), because issues where fixed and contributors have tweeted about it

[@davereid I committed to cvs today and may release on Wednesday.] @I3 (2/22/2011 1:14:00 AM)

In the previous section we have mentioned that Drupal product is module based, that means that with every new release team members produce new modules to improve existing core functionality and maintain old modules, to make them compatible with the current release version. However, the existence of 16,500 modules (as in the middle of June 2012) arises a significant problem. Community is distributed in different countries with a huge amount of developers, so to be sure that developers invented new module with unique functionality can be proved only by reviewing all existing modules. When community was not such large as today it was an easy task done by manual reviewing, as I3 explained “in two-three days, you could read them [modules] all, you could test them all”. However, today “you cannot do that anymore” (I3), because, despite the large amount of existing modules, the new one emerges constantly. In this case,
Microblogging was used to stay informed and help reduce duplication. As we see from interviewee comments this is practically used in Drupal community:

“@drupal_modules Twitter feed helps me intercept that stuff occasionally, like “Hey, this sounds familiar to this other project”. (I3)

The @drupaldocs group dedicated for all types of document updates, team updates, and upcoming documentation team meetings on IRC, etc. Thus, this group followers obtained “hot” topics and had the ability to create base understanding on documentation base. As it was mentioned before, for convergence tasks, such as writing collaborated documents, reviewing and editing collaborators synchronous communication (IRC channels) has been used.

There was a significant increase of tweets on March 2011 (Figure 13), which is explained by the fact that DrupalCon (annual community event) took place. After the conference, newcomers started to join documentation team and actively contribute in first weeks, look for information, express satisfaction of joining the team. The same behavior pattern was repeated on March 2012, with a lower number of tweets, however, significant increase before Drupalcon in 2012 and slow decrease afterwards. This users’ behavior proves the statement that newcomers start to look for initial information in social media groups. Usually tweets were accompanied by links to information updates or help requests:

[ Anyone able to do a quick tech review + mentor on this issue? http://drupal.org/node/1205398 Thanks! #drupal] @drupaldocs (6/30/2011 9:49:06PM)

The other group type is @drupal_security, which provides short automatically generated references to the posts on website. The number of tweets is remarkably stable (Figure 14), without any big variances, however, only one significant increase is observed on March 2012, because the security team introduced improved security issue reporting process, thus the number of issues increased (Figure 8). This group has a circle of followers containing basic Drupal’s security team.

Such groups like @drupal and @drupal_planet post general information like changes in Drupal development, Drupal.org structure, events, releases, feature changes, or retweets from other groups like @drupal_modules or retweets of influential people like core maintainers. One of the interviewees described @drupal group as “a way of being informed about the changes” (I6). @drupal has largest circle of followers counting about 30,000 individuals (Table 3). Considering the fact, that there are 18,000 developers, half of the followers are ordinary users who seek news and updates about Drupal product. Comparing with about 900 @drupaldocs followers, we obtain concrete information distribution, because documentation is not relevant information for general user, only for contributors follow this group. To summarize we can divide all groups to public (e.g. @drupal) – which information is relevant for all types of users and strictly development group (e.g. @drupal_security) – which information usually is relevant for one working team. Information from all these groups provides kind of reference to the source of information, thus to be familiar with context participants have to spent time for reading and understanding the meaning. That is why microblogging was used, by community members, to announce news, explain when and where to find up-to-date information and ask for feedback “tweet about it (..) asking for feedback” (I2), however feedback usually requires to take some time.

The active audience, in the sense of retweeting and posting in previously described groups consist of product developers, maintainers or any kinds of contributors, who are interested in not only general information like product releases, but are more deeply involved in the development process. Thus, developers additionally follow each other (Appendix 3) in the sense of interests and common work, tasks, or working area.

“I follow a lot of Drupal developers on Twitter to see if they post interesting stuff about how they are solving certain problems” (I3).

Microblogging fulfills low media synchronicity description and provides high information conveyance. Conveyance process requires cognitive resources to read, understand and integrate this information into a mental model. Lower synchronicity allows more time between messages so that the receiver has time to
process cognitively the message. When a user has decided to follow one of the groups described above, tweets start appear in the user’s personal timeline feed by feed mixed with the followed developers’ tweets. Other possibility is to group users to the controlled lists. Use of this functionality increases conversation parallelism and information processing performance. Summarizing microblogging functionalities disclosed during Drupal community members interview and tweet analysis, we observed that Twitter fulfills several MST media characteristics as immediacy of feedback, symbol variety, parallelism, information reheasability and reprocessability; thus, these will be detailed explained in section 5.2.3.

Below presented is the primary empirical conclusion obtained from this section:

PEC3: Twitter group accounts are used to simplify information search and gather it by work fields.

5.2.3. MEDIA CAPABILITIES

This section presents media capabilities, which are formed based on communication and common ground theories, data collected from interviews and tweets.

The microblogging information transmission is based on the parallel communication, which increases communication asynchronicity. As I4 stated “every single Drupal communication, is always happening in parallel”. Developers are used to operate several tools at once like social mediums, IRC, email, VoIP and others “it never stops!” (A4). This practice improves the idea that, for example, Twitter in case of Drupal community, captures only part of communication need. Exploring only Twitter capabilities, based on evidences, we state that it obtains high parallelism, because at once there were executed several topics, and contributors could monitor all of them at once and do not lose the sequence.

The sequence is obtained by using “@reply” function, which creates visualization of conversation at the same time. In the case, if some user e.g. A have mentioned user B, and B wants to respond; thus he uses reply function. In this way, conversation is created (Figure 21). The tweets appear in user timeline based on the last tweet time. If first tweet was created several days ago, the whole conversation will appear with the last tweet time. Analyzing presented example (Figure 21) first tweet was sent at 11:56 p.m. (we obtain the time by expanding first tweet), so respond time took 12 hours delay, during the delay time in both users timelines other tweets appeared, however, conversation sequence was continued and information has not been lost. Message sequence outlines that at the same time, even if it is not real time conversation, the broadcast nature of microblogging can create a vision of real time conversation “because you can be up-to-date with everything really quickly” (I6)

As interviewees have claimed, real time communication helps a lot in solving issues fast, so when they have a problem, they “have an issue tracking system on Drupal.org” (I2), where users can post problems. But because of the large amount of users who develop, maintain modules and find bugs, there are a lot of issues. I2 explained the cruciality of the situation “just for Drupal core there are 9,000 open issues”. That means if someone posts a new issue this will get lost in whole pool of issues and community member can wait for a long time or even never receive a reply. As a core developer states, to get noticed about newly upcoming problems or to cry for the help by asking whole community support, developer has to share post
on several different communication tools. That means if a developer wants to reach masses he has used different channels including Twitter to reach thousands of people.

“I have some followers who are interested in the same multilingual area, so I tweet about my posts” (I2).

Tweets about releases, cries for help, personal work achievement goes in parallel and appears in a user’s timeline. That is why a user can be always in the center of current information and, independently on timeframe, user is able to retrieve data as soon as he can. The interviewees mentioned that they followed group accounts based on interest and for data reduction. As in the case of modules people frequently duplicate work “so using the @drupal_modules Twitter feed helps me intercept that stuff” (I3). As we have presented before the interviewees are following several groups at once (see Appendix 3), because their community daily work consists of different tasks. Each group (Figure 13 and 14) is tweeting several times per day. Thus, it is possible to stay informed about all changes by short tweets. If any topic is relevant for the current task, user can just perform further analysis by exploring attached link or by asking responsible person. It is easy to reach responsible person, because he is usually mentioned in the tweet. Consequently, there were analyzed if microblogging can provide reachability of information processing. Twitter is described as a spontaneous medium. By exploring medium abilities, it was found that there exists media rehearsability capability. Thus, a user can edit message before sending. Also, if a message is send, it will be always possible to delete it from personal timeline and followers timelines. As was mentioned in 5.2.1 section, user can always review sent or received messages, these increase information reproprocessability factor and user’s control over the data.

“I can maintain, I can follow and unfollow people (...). There [in Twitter], I have like more control about it[information flow]” (I6)

Ability to ply parallel topics almost at the real time and present them on the same time board provides the possibility for team members attend several conversation at once. In some of the conversations user is a passive listener in some - an active participant. Also, Twitter is used to extract mostly relevant topics from the big pool of updates. Because of the large amount of data from different sources, like group accounts and individual accounts, a person can easily get lost in information appearing in parallel a mix of work-related and personal tweets.

For productive information processing, community contributors are facilitated from microblogging supported ways of information reproprocessability. They have used list functionality. For data grouping, there is possible to create list - a manageable group of Twitter users. Those lists can be formed of a mix of personal and group accounts. When user preview list, he see a stream of Tweets from only the users included in that group “so I have a Drupal lists, because I am not only “following” Drupal guys” (I6). Thus, I6 has created 5 lists where two of them a dedicated for Drupal work. Also, I6 is a member of other groups, and most of those are Drupal related groups of other community contributors. By exploring who, from analyzed individuals, is adopting lists to sort information, we found out, that 5 out of 8 participants (6 interviewees and 2 core developers) have created from two to twenty lists. Lists have functionality which does not require, from user, to follow set of users, but select extraneous list and follow it. For example, I1 is following 21 lists, and he has created only one list by himself, that means he is following twenty foreign lists. Extraneous list selection is based on the support and trust of other people (more detailed explanation in Section 5.2.6). Additionally, community members are facilitated from grouping information into the “favorites” list. When they like a tweet they click on favorite and all tweets goes to separate “favorites” section. Other exploited possibility is to search information by known names of users using mentions (@) or keywords using hashtags (#). In this way, users can always receive list of the latest information on different topics and manage to build own knowledge from different sources.

“using the “hashtag” you can use it, you can “channel” your communication for a specific audience, for a specific group of interest” (I6)

Microblogging contains various sets of symbols to express thought in a richer way. Text and numbers are basic characters, text is used to explain the main thought and numbers to qualify provided information, for example, release number or the date when release should appear. 43.6% of all analyzed tweets contained
links, which were accompanied as a resource to familiarize with topic in-depth. Videos and picture are a rare phenomenon; however, they also appear as a proof to some of the statements. Usually pictures were associated with personal life, with Drupal community social communication or some moments from Drupal conferences, as presented below:

[#drupal caching by @slashrsm http://pic.twitter.com/B48xKIxV]@I5 (4/5/2012 7:55:06 PM)

Some of the analyzed contributors practices to attach different types of documentation in various formats like ppt²⁶, doc²⁷. Usually, C1 facilitates from this option. The main difference comparing personal and group accounts is that, in personal accounts, people express emotions, using emoticons²⁸, like

[It fills my heart with joy to see all of the former #GitHub projects coming back home to #Drupal.org now. :) Yay! Welcome back!] @C2 (2/24/2011 8:54:55PM)

Despite of virtual communication in microblogging, it is possible to express language variety using emoticons, a set of words like representing positive or negative sentiments. These symbols can be evaluated as an advantage and disadvantage at the same time. Spontaneously expressed emotions can provide better characteristic of person and fill out his virtual profile. However, they can arise some conflicting situation if negative emotions where expressed towards specific person (detailed explanation is presented in Section 5.2.6)

Below presented are the primary empirical conclusions obtained from this section:

| PEC4 | Lists (manageable groups of Twitter users) are used to perform data grouping. |
| PEC5 | Team members are able to execute several conversations in parallel and do not lose the meaning. |
| PEC6 | Various symbols help to express emotions. Thus, information is presented not as a set of words, but as a personalized thought. |

### 5.2.4. TECHNOLOGY FEATURE

After microblogging was explored as a communication tool for information receiving and processing, this section investigates technological aspects of used type of social media. One of the basic functionalities is broadcast nature “If you post on Twitter it could be spread” (I5). With the retweeting power and use of hashtag, not only focused people can receive useful information by viewing their timeline feeds. Even without personal knowing and agreement to follow author posts community contributers can receive and exchange flow of information provided by the source. The interviewees outline the need to use tools which would be able to present information publicly and spread to a wide audience:

“people can see your answers and more people can get in touch, I mean, more people can get into conversation and contribute”(I5)

At this point new aspects of technology feature, which have not been mentioned in literature yet, expand understanding of broadcasting definition. In Wikipedia broadcasting is described as “distribution of audio and video content to a dispersed audience via any audio visual medium”²⁹. And it is based on the idea of “getting one message out”, thus one-to-many approach. However, Twitter (as other social media tools) provides rare possibility to distribute information by many-to-many communication. This automatically

---

²⁶ An abbreviation of power point filename format for presentation file.  
²⁷ An abbreviation of “document” is a filename extension for word processing documents.  
²⁸ An emoticon is a pictorial representation of a facial expression using punctuation marks and letters, usually written to express a person's mood. Emoticons are often used to alert a responder to the temper of a statement, and can change and improve interpretation of plain text; emoticons for a smiley face :-]) and sad face :-(_
²⁹ http://en.wikipedia.org/wiki/Broadcasting
decreases time and a person’s effort to send a message to a large audience. The interviewees supported this statement by clarifying, that open communication facilitates by broadcasting answer and help for a wider audience at once:

“And it could take me more time to answer, but I know I will not be helping one person, but maybe 20 or 50” (I5)

Other broadcast facility is to share all decision-making publicly. As I3 states, in Drupal community exists the problem of covering decision-making, there exist Drupal 8 initiatives and some of the contributors who works with them are communicating in private. Other members are dissatisfied, because those decisions, which affect the direction of the project, are being made in private. That is why he suggests presenting all information in public. Initially mailing lists were actively used for information spread, for large groups, to share new information and build knowledge, however, as it was stated several times before, the community grow fast and mailing lists were discarded:

“I think the reason that we moved out of mailing lists was that we grew too big for them, to be useful media for the discussion.” (I2)

Even if Twitter uniqueness is based on its broadcasting nature, it provides the ability to send direct messages too. Those are secured from publicity and only determined person will receive a message. With direct message feature, user private communication is safe and closed for publicity. We could not analyze content of private messages; however we saw, that community contributors have facilitated from this option.

[@ry5n Rock!! I figured we could just meet up after my session, but Twitter DM logistical works too. :)] @C2 (11/05/2012 9:27:01 AM)

The Twitter’s broadcasting nature is associated with spontaneous tweeting. Positive site of spontaneity is a generation of new ideas, brainstorming, and capturing instant thoughts like [If I could change one thing about mapping in Drupal it would be...] Drupal Groups http://t.co/pHOLTlhT#drupal#mapping] @I6 (4/26/2012 8:11:11 PM). But, one of the interviewees described this characteristic as a possible disadvantage for the media; because participants tend to express not only positive excitement feeling, but often express negative emotions “I cannot believe that this idiot just did this” (I3). I3 states, that most of people do not put a lot of thought into tweet. However, tweets’ analysis provided negative emotions directed to all community, not direct person:

[I know its volunteer work, but you seriously need to put some effort into design not only code http://t.co/wnFHN6YV#ugly #drupal] @I5 (1/30/2012 8:12:16 PM)

However, we have analyzed limited number of tweets that could be the reason we did not find directed negative emotions. Of course, tweets from personal accounts usually contain emotional spirit, in group accounts there are not any emotion, usually formal information. From those emotional tweets addressed to the team members we have distinguished usually positive emotions consisted of motivational or thankful note (more details are provided in 5.2.6).

As we have mentioned before, Drupal community is relatively large and still growing (Figure 5 and Figure 6). Thus, amount of information to process is enormous and amount of people who follow community contributors is huge “I have 2,000 followers on Twitter and together with them we have about 10,000-20,000 people” (I1). In Table 5 we obtain confirming results, that user obtains more than 2,000 followers. Also it is important to outline that Drupal contributors have dispersed number of followees (Figure 22). The number of followees’ represents the amount of information contributors obtain. Received information is not only work or development related; there is constant mix of information flow.
From the diagram we obtain, that community contributors followees’ number is 674 and their standard deviation is \( \sigma = 679 \) (contributors have varied amount of followees in the range from 8 till 1800). Contributors averagely post about 2.19 tweets per day (Table 5). They follow not only Drupal related people, but anyone from their personal circle, which averagely tweet 1.85 tweets per day. Thus, C1 daily obtains at least 16 tweets, I4 about 3600 tweets and in average interviewees obtain about 1,200 tweets daily in their user timeline. Additionally they may search for supplementary information by hashtags, mentions, or in Drupal account groups. These observations present the huge number of data users retrieve everyday.

The advantage of having the possibility to ply this amount of data is that firstly huge mass of contributors can be reached at once - “you can retweet, and if you are lucky it might have 100,000 impressions” (I1). Not only community contributors can be reached, but also people from outside, who even have not considered the existence of such product or its functionality, facilities. For example, @drupal (Table 3) group account has more than 30,000 followers and about 1,300 tweets, comparing with @C1 (Table 4) has more than 18,000 followers and about 6,000 tweets. Thus, regarding the fact that the tweets from these accounts (@drupal, @C1) are retweeted, by their followers, to even wider audience, tweet is reaching masses. The disadvantage is that the bigger is the amount of data, the more increases difficulty to manage information, retrieve it and understand. However, for information reprocessability are used methods described in section (5.2.2).

During the tweet analysis and regarding interviewees “I was reading [tweet] on my iPhone”(I3), we have noticed various types of devices users used to tweet. Under each tweet there is a note from what device tweet was sent, like:

“1:28 PM - 3 Jun 12 via Twitter for iPhone”

This research aim was not to analyze what device are used for tweeting, so we had not made any deep analysis, only brief scan of possible device, we distinguished these categories: Mac, iPhones, Android, via the web, via Mobile Web, etc.

Interviewees used mention “@” to directed message for user. Mentions usually contained not full names, but nicknames that are why there is not any kind of formality in conversations between users. If the messages were sent by group account they have some kind of structure and formality, however some of them are generated automatically; thus there is some type of formulations. Personal messages contain emotions and expressions having more kind of friendly form of communication. Any restrictions for tweets format facilitate spontaneity and real expressions and personalization. Informal communication often consists of unplanned and brief “catching-up” conversations. It supports sharing of work-relevant information among co-workers and creating potential collaboration opportunities. Social functions such as transmission personal life details (e.g. “we know, when somebody broke the leg or we post pictures of our cats” (I2)) maintain common ground and a feeling of connectedness between co-workers (details are presented in Section 5.2.6)

---

As it was described previously Twitter provides the possibility to write text-based messages up to 140 characters that are pretty limited amount of text. We have analyzed 8834 personal tweets and 6821 group account tweets to understand tweet average. The results we get, stated that group message length contained averagely 107 characters and personal tweets averagely about 101 characters. Considering the fact that from all personal account tweets 44% contained a link, which averagely is about 15-20 characters. Thus, tweets are even shorter than default settings; however they provide enough information to have an idea and get information.

Averagely users tweet 2 times per day, where at least one tweet is associated with Drupal development or community related work. Only I1 tweets about 40 tweets per day and his 75% of all tweets are about his personal life. Also, interviewees who are not native English speakers tweet in both languages (English and their native language). Group average is 5 tweets per day, but interval varies between 1 to 19 tweets per day depending on group. All group tweets are written in English, only local group accounts as @drupalat (Austrian Drupal community) are written in German.

Below are presented primary empirical conclusions obtained from this section:

PEC7: Microblogging enables many-to-many paradigm.

PEC8: Spontaneous tweets can generate useful, unique ideas and solutions; however, in some cases they provoke controversy within the development team.

PEC9: Microblogging is applicable for large groups of collaborators.

### 5.2.5. INTERACTION TYPES

This layer is dedicated to investigating communication patterns and mostly communication topics - genres. We present two basic parts in this section. Firstly we compare how interviewees described their daily development related practices using microblogging and after those statements we present supporting or contradicting evidences as a tweets. Secondly we outline mostly repeating and characteristic communication genres used by Drupal community contributors. As the interviewees stated, they have one account for personal and work interaction. Thus, we have analyzed all tweets (8834 totally) from middle of December 2010 till middle of May 2012. Additionally we have analyzed several groups (Table 3) mostly followed by Drupal contributors including interviewees. As was described in Section 5.2.2, tweets in the groups usually are expressed as a type of RSS feed to the relevant posts located on the main Drupal.org website. However, we also have analyzed repeating keywords and supported interviewees’ statements using group account tweets. Below presented tweet percentage is calculated by analyzing only those tweets, which contained “#drupal” hashtag, as it is direct evidence of drupal related context.

In section 5.1.1 we have outlined common tasks performed by community members. Now we present how interviewees used microblogging media regarding those tasks. Comparing two general communication types in microblogging (personal and work-related), interviewees differently perceive Twitter assignment. As I1 have stated “on Twitter I talk a lot. So I tweet a lot (...), Twitter is for me, that is how I look on it”. On the contrary, I6 claimed “I have started my Twitter account, because of Drupal and I am still using it mainly for Drupal”. Thus, from analyzed tweets, we can see that analyzed set of people tweeted mostly about work related topics (7 out of 8 people have mentioned drupal keyword in more that 50% of their tweets) and only A1, as he has claimed, has only about 30% work-related tweets. Rest group of contributors uses Twitter accounts to perform Drupal community and development work-related tasks, because they tweet about Drupal related work and have most of the followers from Drupal community “I think 80% of people who are following me are somehow connected with Drupal” (I5).

---

31 There were difficulties with his data analysis as he tweets so much, that we don’t have older data as only month’s data
We did not particularly analyzed personal data (called as “daily chatter”), just distinguished several categories like type of conversation with other users, technology knowledge/news sharing, world events, personal feelings, and family, friend’s related facts. When we were analyzing team members personal focus in Twitter, we noticed, that there does not exist only work-based communication between developers. Usually developers favor to know personal life details about team members - “So we know when somebody brakes his leg or we post pictures of our cats” (I2). Also there exists clear binding of personal and work topics:

[Hm. Looks like we will need a #Drupal 8 code freeze extension. My timeline is filled with people who just got @Diablo II] @C2(05/15/2012 10:45:09 PM)

The main aim of this research was to understand how microblogging is used to communicate during software development process. Representation of all found interaction types and their subtypes with frequency in percentage is presented in Appendix 4. “I use Twitter as a tool to help me with my work” haveI4stated. Research number shows that 14.14% of tweets are about Drupal 8 version, 18.02% talks about currently released Drupal 7, and less than 1% mentions Drupal 6. Till Drupal 8 is in the development stage two previous versions of product are mainly used by users and maintained by developers (Section 4.1), tweet example:

[Updated #Drupal #localization cheat sheets (now for Drupal 6 and 7), and Drupal 7 localization tips: http://bit.ly/eQH3a7]@I2 (1/25/2011 5:58:37 PM)

The interviewees stated that to perform strong connection between team members and have a clear idea of what is happening in the community, it is suggested to attend one of the meetings or events they have organized. Majority of interviewees opinion is, that if contributor spent some time with team members not only by working on the project, but “being with someone in the same room, (...) or eating meals together”(I4), can help to know each other and “build up sympathy for them and personal connection strongly multiply our ability to cooperate virtually” (I4). Consequently, contributors participate in different global, local conferences, meetings. Also, several of interviewees mentioned that they are used to announce about upcoming events on Twitter. In this way, they let others know, that they will take a part in this event “Hey, I going to be in Denver”(I4). Often during those meetings they organize code sprints, solve some problems, set new goals.

[Having a Workbench sprint today. Come hang out in IRC#drupal-workbench at 2:30 EST. #drupal] @I1 (9/19/2011 2:23:34PM)

Drupal event updates occupy a large part of tweets about 18.99%; they are accompanied by links to the events with dates and hashtags (as presented in the above example). We also highlighted contributors’ habit to discuss events using Twitter conversations. Part of the tweets are strict announcements with forward details of upcoming events, others are expressed usually with positive (happiness) emotions, however sometimes there are negative expressions too.

I3 described microblogging as a type of message board, where people ask for others help. Tweet analysis showed that there were about 5.42% of help request’s, most of them were associated with Drupal 8, what is explained by the current pre-release status of this product. It is interesting that help requests are accompanied with links, references (@) to other people or sometimes just general cry for help to a wider audience:

[Want to get your feet wet in CMI and help #Drupal 8 Multilingual a lot? Convert the contact form settings: http://t.co/ioXzD1yJ #D8MI] @I2 (5/18/2012 11:44:11 AM)

Most of the interviewees acknowledged I3 statement that Twitter is kind of message board for expressing help requests.I2 states that he saw that people put focus on his tweets, because he saw community contributors retweeting his post and announcements, also he has received feedback from people on issue queues, because other members have looked “on issues what I have posted”. He described advantage of broadcasting to the community as it increases people focus on the development, problems and
achievements in the organization - “when I write something on the issue queues I also post it on Twitter, so my followers can be aware of it” (I6). We have noticed, that help request interaction is used as a user support service, considering the fact that developers are perceived as users in OSS.

[Looking for some help from anybody who has experience writing Views 3 display plugins that are a full page replacement. #drupal] @drupaluser (05/14/2012 7:15:19PM)

Release updates and implemented new features forms 10.66% of all analyzed tweets. Such information is accompanied by links or hashtags, this increases followers’ convenience to find further information. If there are no links, so hashtags provides specific data for readers to understand the content (e.g. number of version, subversion). If the module changed or updated the name of a developer who performed change would be mentioned. This release features usually are tweets by group accounts and community contributors just retweeting them for their personal audience. Example from group account:

[Announcing #drupal 7.6 http://t.co/8s3gG72 with important big fixes and a security update for sites using private files on comments] @drupal (7/27/2011 8:38:26 PM)

There is the occurrence of to do’s messages (3.3%), which were divided to two revealed categories: personal and coordination. Both messages are associated with Drupal current work tasks like bug fixing, module maintenance or with the upcoming event organization. Work related tweet example is “@BarisW @fschaap please do! @drupal4gov #ftw” (II). There were outlined, that there were not much coordination tweets, also interviewees have not mentioned that they coordinate work using microblogging media. Core developer outlined that for the current core initiatives they started to use phone conversations to manage work to be done. Module tasks are distributed and done either by individual or by a small team, if these teams for some reason do not use Drupal hashtag (#drupal) it is impossible to track coordination in Twitter. Also the interviewees do not perceive Twitter as a coordination tool; they have outlined other tools to carry out this task “IRC is for me real coordination tool” (I3). However, the personal coordination is more frequent case, the interviewees have not mentioned that they are used to tweet about what they plan to do, or what they have to do. Usually, they share personal to do notes with publicity. This type of tweets is also like a short report of personal activities, thus others are able to know team member’s work progress and doubts. Below is presented one of the example:

[Just started writing a blog post that looks back at Drupal in 2010. Hope to finish it by tomorrow but might be Tuesday] @C1 (1/2/2011 7:34:00PM)

These to do’s tasks smoothly transforms to the personal achievements reporting (7.95% of tweets). Each analyzed community member has posted at least several times about personal achievements. Usually such tweets are accompanied by links, or at least hashtags (e.g. [working on the @drupalat site with @the_real_fago on the train back from drupal roadshow innsbruck. local]@I6 (5/12/2012 2:08:04 PM)). These tweet types help, in the development process, to be aware of current statuses of other community members and “everybody who could be interested can come and look” (I4) to the issue queues, blog, etc. These tweets not only increase contributor’s awareness, but also increase knowledge and help to understand specific types of problems. When other product contributor faces similar issues, he will know how and who have solved similar problems. For example, during tweet analysis, we have found some tweet with code example:

[#Drupal 7 Tip of the day (when building forms using taxonomy terms). $options = taxonomy_allowed_values(field_info_field($field_name)) @I3 (5/26/2011 12:07:00 PM)

Thus, team members can share different type of information between each other to provide maximum knowledge with minimum text. We have considered it like a new type of documentation or at least knowledge board. Leading part of interaction type in Drupal community are tweets related to knowledge share (27.3% of all tweets) about “what is somehow connected to Drupal” (I5): new version, released books, videos, slides, structural changes, competitors achievements, new development practices, etc. This information quite constantly is accompanied after some technology news or changes, for example
Knowledge based information is “helpful for them, helpful for me” (I4) as one of the interviewee’s stated and can be retrieved as type of documentation. Data reliability relies on the trust of source. If it is tweeted by official Drupal group account then there are no doubts about information reliability. If it is tweeted from personal account then reliability would depend on the author’s reputation in the community (this social aspect will be discussed in section 5.2.6).

Practitioners and researchers always looked for a way of motivation that would lead employees to share knowledge and valuable information. Motivation is needed because information sharing often involves extra effort by a worker. In open source communities all contributors are volunteers and only personal gratitude for the done work is “thank you” or any motivation word. Drupal also practices motivation and “thank you” strategy. When interviewees where asked to enumerate tools where they try to motivate other members or say “thank you”, only one developer said, that part of the community located in Central America “mention each other on the tweets, they like to motivate each other” (I6). The rest of interviewees have presented other tools and methods how they motivate distributed team members. However, tweet analysis showed other statistics, that there were found that about 6.8% of tweets contained gratitude. These types of tweets were dedicated to a particular person or general part of the audience. The general gratitude where dedicated to some local communities, some group of people or for the whole community. Personal gratitude where more specified: with mentions for one or several people and usually with links, which provided further explanation for what person was gratitude, e.g.

[Wow, lots of weekend activity on #D8MI in the IRC channels (and the queue). Great people! @plach__, @Sutharsan, @ClemensTolboom, @thefubhy] @I2 (4/7/2012 3:31:51 PM)

As we have stated before, group accounts do not provide exploratory power, thus we decided to analyze if main “#drupal” hashtag can confirm or contradict examined communication genre. us, regarding the Twitter limits (only 1500 tweets per retrieve), we have extracted tweets containing “#drupal” hashtag for two periods March 23-26 and June 18-19, both periods represent beginning of the working week. From the extracted tweets using twitteR and tm packages we have mined extracted tweets. Thus we have created two keywords’ cluster dendograms for both periods and cut to 7 clusters each to perform deep genre analysis (Appendix 5 and Appendix 6). Height represent how many time keyword was mentioned in analyzed tweets. Cluster dendogram is based on hierarchical clustering. Related keyword are located nearby, farther away keywords are few or not related. Consequently, “joomla” and “wordpress” are more related than “joomla” and “modules” (Appendix 5).

32. http://cran.r-project.org/web/packages/tm/index.html
“#drupal” hashtag is publicly available to any Twitter user, thus not only community members had an ability to tweet using this particular hashtag. However, to sort data from uncommon keywords associated with “#drupal” we have removed sparse keywords, i.e., keywords occurring only in exceptionally few tweets. Based on the obtained cluster dendograms (Appendix 5 and Appendix 6). We highlight most frequent keyword “modules”, which does not have close relative keywords, however frequent occurrence of this keyword is explained by the fact, that @drupal_modules (Figure 13) group account tweets very frequently about daily module updates and innovations. Dendogram (Appendix 5) has relative keywords as “denver”, “thanks” “sprint” and “release”, this is caused by the Drupal conference in Denver, which was on March 19-23, and analyzed tweets are from March 23-26, thus cluster of keywords represent users’ share of positive emotions and information share, such as:

[The #Drupal Media module now has an official 1.0 release thanks to the #DrupalCon Denver media sprint!] @drupalcontributor (2012-03-25 19:48:02)

Relation of keywords as “dries”, “cool”, “photo”, “keynote” confirms that contributors use rich symbols (i.e., pictures) in their tweets, positive emotions to express their experience (in this case share about a past event). Also we have observed increased the number of tweets with “#drupal” hashtag after the conference, this mean that users tend to continue virtual communication after meeting each other. Cluster of “themers”, “symphony”, “twigs”, “learn” confirms one of the analyzed genres – sharing information and learning, which helps to build common ground. Previously presented keywords are associated with the themes and views developed by community contributors, thus this kind of tweets consist of several keywords as a hashtags and links for a faster search:

[Twig in #Drupal 8? http://t.co/SEpK1WZ - not only devs should learn about #symfonybutthemers too will learn new systems!] @drupalcontributor (2012-03-26 10:26:07)

Last cluster in (Appendix 5) consist of keywords as “developer”, “job”, “theme”, “app”, “css”, “looking”, etc. Thus these keywords represent, help requests [#drupal peeps - I'm having problems theming a form. Could anyone help?], questions regarding Drupal functionality [What is the most customizable #Drupal theme?], personal achievements [Has helped like 20people in drupal IRC], information and knowledge share, and even job vacancies. Second dendogram (Appendix 6) from June 18-23 period, has similar clusters of keywords, the keywords as “denver”, “sprint” were substituted by “team”, “amp”, “project”, which are related to Drupal AMP group meeting. Based on the fact, that not only contributors were able to use “#drupal” hashtag in their tweet context, both dendograms confirmed previously revealed interaction types. Main outline is that hashtag frequency and relation with other words is based to the ongoing events in the community. Secondly, there are many information share and help requests. Finally, users compound tweets using rich set of symbols (regarding to Section 5.2.3).

Below presented are the primary empirical conclusions obtained from this section:

PEC10: Microblogging can be perceived as a new way of documentations by capturing system development events and knowledge from tweets.

PEC11: Developers’ availability prompts to start a conversation or find team members on Twitter.

PEC12: Microblogging can be perceived as a new way of user (developers) support - contributors can ask for help, look for solutions, ask questions and get answers.

PEC13: Work related topics interwoven with personal data create the full profile of a person.

5.2.6. SOCIAL ASPECTS

Communication process performance increases when a sense of social connection between community members is high. The distributed development performed by people who see each other rarely or sometimes have not seen each other in person, creates issues for trust and connection. However, if proper media tool is selected, then connection between members of one community will improve productivity and
efficiency of work. Drupal community has particularly strong meeting and event practice, which provides the possibility to have local (located in the same country, region, city) community members meets and global events. Despite this, there is a need for virtual communication in-between meetings period. To create friendly connection between unfamiliar people from different countries and cultures individuals have to create personal character in a virtual world. Also contributors have to earn “name” in the community. When people select whom to follow, they prefer to receive information from reputed.

Not all interviewees have used Twitter to share personal information, however, at least those who have, are creating a kind of personal characteristics using pictures they post, personal topics they create. “I am also pretty honest about who I am” I4 starts reflecting his online personality. He states that he does not put politics, or controversial things about Drupal. He continues “I am just a real person”, however him and other interviewees sort distribution of personal and work data. When in the community exists interchange of not only work related communication, is gained vision of knowing person, learning about interests and hobbies, family member, etc.

“I never saw anyone and I could consider them my closest friends (..) If I would say I have 20 good friends, half of them are from Drupal” (I1)

When people fill connection with other community members, they also are prone to trust those people information, ask them for help, etc. “I have connections with Drupal developers” (I5) interviewee supports this study’s findings. Other strong social aspect is personal reputation. As A1 stated people go on brands, however people also can become brands, he considers himself as a brand. Interviewed core developer has the same opinion, that people fill trust, when they read “known” person tweets: “many people are following me, because I am already known name in the community” (I2)

Other interviewees confirmed that people follow “known names”, because social media is about people and it has a strong influence. However, microblogging’s spontaneous and emotional nature is able to destroy for a long time built connectedness and relation between people. As I3 states, some negative emotion like “I can not believe this person asked this question in my issue queue and [link to the issue queue]” expressed without thought can ruin close relationship and create some problems. However, he states, that he personally and community members try to avoid this type of tweets.

Use of usernames instead of real names created informal atmosphere between community members, the absence of strict work hierarchy like “fear” to talk to your manager is erased by equal contribution and everything relies on person experience and knowledge. Based on the correlation coefficient calculation in (Section 5.2.1) we received strong relation between individual followees and retweet number. Diagram (Figure 19) presents, that C1 and C2 have strongest power of word. For example, we have selected C1 tweet, which got mostly attention (109 retweets and 17 favorites):

Tweet (Figure 23) received so much attention, because it is about Drupal 8 core changes and is tweeted by main core developer in the community.

Microblogging is a relatively new technology and “it is about people” (I1), thus community have to exploit technologies that are mostly used by the current generation. It is worth to mention that interviewees describe IRC and forums as “very old fashioned” (I5), interviewees use words “geeky” and “nerdy” to describe such tools. Thus they present social media tools as a new generation communication tool, because
“IRC is a very old standard and it is “geeky” way of communication and I have discovered an IRC in the contest of Drupal community” (I4). Even if it proved that developers still use IRC, forums and state that they are efficient tools, thus interviewee claims that:

“I thought that IRC was not even used anymore and we should use communication tools that are “higher”, for example Facebook and Twitter” (I5)

Interviewees argued that people are familiar with social media tools and for newcomers it is better start with nowadays tools. Also I5 states that people are used to “getting in touch with strangers” using social media tools compared with the way to invite people to IRC which arises questions such as “in which century do you live?” (I5). However, I1 outlines that community members do not feel tied up to the particular technology, they are more tied up to innovations and people intriguing technologies “[today] it is Twitter and tomorrow it might be something else” (I1).

Below presented are the primary empirical conclusions obtained from this section:

<table>
<thead>
<tr>
<th>PEC14</th>
<th>Requirement for software development organization to survive in today’s environment stimulates to explore newest technologies such as social media.</th>
</tr>
</thead>
<tbody>
<tr>
<td>PEC15</td>
<td>Tweet openness eliminates the OSS community’s problem of person-to-person information exchanges.</td>
</tr>
<tr>
<td>PEC16</td>
<td>Social engagements increase trust and connection between unfamiliar people, and forms relationships and connections within a group of people.</td>
</tr>
</tbody>
</table>

### 5.3. SUMMARY OF COMMUNICATION PRACTICES IN CASE PROJECT

This section summarizes how different layers proposed to form microblogging utilization in distributed software development communication suggested by the theoretical framework used in the Drupal community. There were also enclosed several parameters which will be added to the framework. Firstly, we have performed comparison (Table 6) of based on literature set of tools used for DSD communication (Figure 3) and set of tools used by Drupal community members (Section 5.2).

<table>
<thead>
<tr>
<th>Communication tool</th>
<th>Theory</th>
<th>Drupal practice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Face-to-Face</td>
<td>Is considered to be richest communication medium regarding MRT (Daft et al, 1987). It has highest social presence which is vital to support social relationships (Urry, 2002). Face-to-face meetings are essential in the initial phase of a project, thus, shared understandings can be developed effectively (Heeren and Lewis, 1997).</td>
<td>Drupal community members communicate extremely rarely using face-to-face, only during annual global conference and local events. During the face-to-face meeting usually they exchange knowledge, solve issues, sometimes develop “the core developers who were in London, where in the same room working together” (I4)</td>
</tr>
<tr>
<td>Email</td>
<td>Is considered to be lean media regarding MRT (Daft et al, 1987) with low social presence, which stimulates decrease motivation, but increases the ability to process information. Email affects the performance of virtual teams through trust, organizational identification, and communication network structure (Yang et al, 2010).</td>
<td>Email is used to solve problems or exchange frequently sensitive information between two or several people “in very tide circle of emails among people who are competent to talk about subjects” (I4). Also, those people, using emails, know each other “I had to know the private contact address of the people” (I6).</td>
</tr>
<tr>
<td>Communication tool</td>
<td>Theory</td>
<td>Drupal practice</td>
</tr>
<tr>
<td>-------------------</td>
<td>--------</td>
<td>----------------</td>
</tr>
<tr>
<td>Telephone</td>
<td>Is considered to be synchronous media regarding MST (Dennis and Valacich, 1999) and is used in case of immediate feedback need (Tissen et al, 2007).</td>
<td>Generally telephone conversations are used very rare; however, recently for Drupal 8 initiatives new meeting practice was started. Every two weeks on the phone members talk about “what we are up to, what were done, what blocks us, whether there is anything we can do for each other, help to move forward” (I2).</td>
</tr>
<tr>
<td>IM</td>
<td>IM is a balance between a phone call and email with respect to formality and immediacy. Is like a short form of an email, even though the response may be delayed (Thissen et al, 2007). Is considered to be synchronous media regarding MST (Dennis and Valacich, 1999)</td>
<td>Drupal contributors use IRC as a kind of IM and comparing with previous tools exploits remarkably often. Mostly on daily-bases to ask questions, developer describes IRC as a tool “for coordination and a social interaction.” (I3). Interviewees outline two biggest advantage of using IRC - instant replay and large audience. However, it also is described as “old” and “nerdy” tool.</td>
</tr>
<tr>
<td>VoIP</td>
<td>Distributed teams without mutual knowledge use audio or video channels for collaborative tasks, and teams with existing common ground can communicate effectively using lean media (Clark and Brennan, 1991).</td>
<td>Drupal members usually use Skype or Jabber for small “organized groups” (I1) or “person to person communication” (I5). Usually it is used to perform personal communication or event organization.</td>
</tr>
<tr>
<td>Groupware</td>
<td>Groupware enables multiple users to use a single software resource. Systems provide shared calendars, group meeting scheduling and shared contact lists (Thissen et al, 2007). Each person maintains his own information in a private calendar; arrange meetings, telephone conferences or any kinds of this interaction; other forms of groupware are document managers.</td>
<td>Contributors used groupware, when multiple interpretations of information or negotiation were required. For example, for document creation users used wiki-like pages “like Wikipedia – anyone can edit it and make it better” (I1).</td>
</tr>
<tr>
<td>Social media</td>
<td>Usually, OSS communities grow more united regularly using social media. The distribution of knowledge through social media increases completeness and speed, thus, decreases misunderstanding between colleagues even if they do not meet face-to-face very often, or at all (Begel et al, 2010). In performed studies, there were found that individuals who use social media are not just searching and sharing information. They also regard them as a platform to meet friends, obtain a sense of belongingness and develop relationships with other people (Zhao and Rosson, 2009).</td>
<td>Community members facilitates from using social media ability to obtain answers to any kinds of questions and provide a strong information base. This is particularly useful for the newcomers, or people who have never heard about existing product functionalities or community activities. Also, social media are perceived as a new generation communication tools: “we should use communication tools that are “higher”, for example Facebook and Twitter” (I5).</td>
</tr>
</tbody>
</table>

Table 6: Communication framework practice in Drupal
Additionally, interviewees mentioned, that they often use issue tracking system integrated in the Drupal.org website. We also find evidences, validating the fact that contributors issue tracking systems, in the tweets, when users referred to the issue:

[“Do you have good experience with the Drupal test system? This upgrade path testing issue needs you help! Thanks! http://t.co/aawXUUV9"] @I2 (30.01.2012 14:09:55)

Social media zoom in towards microblogging tool is explored by each layer’s parameters (Table 7) presented in the theoretical framework (Figure 4) and corresponding practices in Drupal community (Section 5.2). As interaction type layer was build on open coding strategy, those obtained types which were equivalent with the theoretical framework interaction types are presented in Table 7, below are described remaining genres.

<table>
<thead>
<tr>
<th>Layer</th>
<th>Parameter</th>
<th>Drupal practice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information processing</td>
<td>Uncertainty</td>
<td>Microblogging usually was used when newcomers start building knowledge information based on documentation, existing practices. Other observed exploitation is to stay up-to-date with all development related tasks and most know issues as interviewee states “you can be up-to-date with everything really quickly” (I6). Also, used to look for help and support from more experienced developers, to write documentation or to understand how things work.</td>
</tr>
<tr>
<td>Communication process</td>
<td>Conveyance</td>
<td>Gather all references of the information, which would cover uncertainty problem solution. Various field groups (e.g. @drupalcore) are used to have better data search and sort. Also groups were used to manage several sources at once, like for example work duplication “so using the @drupal_modules Twitter feed helps me intercept [duplication] stuff”(I3)</td>
</tr>
<tr>
<td>Media Capabilities</td>
<td>Transmission</td>
<td>There is a need to communicate through several channels at once, because community contributors perform several tasks in parallel. Twitter supports the amount of conversations at the same time.</td>
</tr>
<tr>
<td></td>
<td>Rehearsability</td>
<td>As interviewee states rehearsability is needed to maintain the amount of received information, thus they “can maintain, I can follow and unfollow people (…).I have more control over [upcoming tweets ]”(I6)</td>
</tr>
<tr>
<td></td>
<td>Reprocessability</td>
<td>List functionality is used to have the possibility to analyze and recall tweets later. Contributors use the same account for personal and work communication, thus, lists enable information sort.</td>
</tr>
<tr>
<td></td>
<td>Symbols Sets</td>
<td>Using various set of symbols contributors express emotions and provides reach information for the readers.</td>
</tr>
<tr>
<td>Technology feature</td>
<td>Brevity</td>
<td>Tweet fragility increased fast information perception, did not require much time not only to read, but also to write tweets (in average contributors tweeted smaller than 100 characters posts including links).</td>
</tr>
<tr>
<td></td>
<td>Spontaneous</td>
<td>Tweet spontaneity caused idea generation and innovations like “If I could change one thing about mapping in Drupal it would be. (...)” @I6, but negative tweets disturbed communication flow and interpersonal connections (e.g., “I cannot believe that this idiot just did this” (I3)).</td>
</tr>
<tr>
<td></td>
<td>Informal</td>
<td>Personal messages contain emotions and expressions having more kind of friendly form of communication. Informal communication often consisted of unplanned and</td>
</tr>
</tbody>
</table>

63
<table>
<thead>
<tr>
<th>Layer</th>
<th>Parameter</th>
<th>Drupal practice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technology feature</td>
<td>Informal</td>
<td>brief conversations. Sharing of work-relevant information among co-workers and transmission of personal life details (e.g., “we know, when somebody broke the leg or we post pictures of our cats” (I2)) maintain common ground and a feeling of connectedness.</td>
</tr>
<tr>
<td></td>
<td>Mobility</td>
<td>Contributors tweeted from various types of devices (phones, PC, etc.)</td>
</tr>
<tr>
<td></td>
<td>Broadcast nature</td>
<td>Mass of contributors was reached at once, even people from outside “you can retweet, and if you are lucky it might have 100,000 impressions” (I1). Community collaborators facilitated from many-to-many paradigm.</td>
</tr>
<tr>
<td>Interaction types</td>
<td>Work related</td>
<td>Contributors were asking for help by directing a question to a particular person or by sending global cry for help.</td>
</tr>
<tr>
<td></td>
<td>- Help request</td>
<td>Usually contributors tweeted about personal achievements, product statuses or upcoming event announcements.</td>
</tr>
<tr>
<td></td>
<td>- Updates</td>
<td>Personal to do’s were like publicly presented work schedule; very rarely developers coordinated work and ordered to do the task for other contributors.</td>
</tr>
<tr>
<td></td>
<td>- To do’s</td>
<td>Knowledge base information about general product features or particular tips and specifications.</td>
</tr>
<tr>
<td></td>
<td>- Information share</td>
<td>Explored personal tweets contained world news, technological innovations, information about family and friends, emotions and experiences.</td>
</tr>
<tr>
<td></td>
<td>Daily chatter</td>
<td>Interviewees outlined, that followers trust information tweeted by known or respectable author “many people are following me, because I am already known name in the community” (I2).</td>
</tr>
<tr>
<td>Social aspects</td>
<td>Trust</td>
<td>Developers tend to motivate each other, send “thank you” tweets for done work, greet with work achievements or personal events (like birthday).</td>
</tr>
<tr>
<td></td>
<td>Connectedness</td>
<td>Retweet relationship with follower number showed that if a person has “name” in the community, then he has relatively large circle of followers and retweets. Trust and connectedness are dependent variables.</td>
</tr>
<tr>
<td></td>
<td>Reputation</td>
<td>Based on daily chatters and emotions expressed in tweets users create their personal virtual profile “I am also pretty honest about who I am” (I4).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Layer</th>
<th>Parameter</th>
<th>Drupal practice</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mobile</td>
<td>Contributors tweeted from various types of devices (phones, PC, etc.)</td>
</tr>
<tr>
<td></td>
<td>Broadcast nature</td>
<td>Mass of contributors was reached at once, even people from outside “you can retweet, and if you are lucky it might have 100,000 impressions” (I1). Community collaborators facilitated from many-to-many paradigm.</td>
</tr>
<tr>
<td>Interaction types</td>
<td>Work related</td>
<td>Contributors were asking for help by directing a question to a particular person or by sending global cry for help.</td>
</tr>
<tr>
<td></td>
<td>- Help request</td>
<td>Usually contributors tweeted about personal achievements, product statuses or upcoming event announcements.</td>
</tr>
<tr>
<td></td>
<td>- Updates</td>
<td>Personal to do’s were like publicly presented work schedule; very rarely developers coordinated work and ordered to do the task for other contributors.</td>
</tr>
<tr>
<td></td>
<td>- To do’s</td>
<td>Knowledge base information about general product features or particular tips and specifications.</td>
</tr>
<tr>
<td></td>
<td>- Information share</td>
<td>Explored personal tweets contained world news, technological innovations, information about family and friends, emotions and experiences.</td>
</tr>
<tr>
<td></td>
<td>Daily chatter</td>
<td>Interviewees outlined, that followers trust information tweeted by known or respectable author “many people are following me, because I am already known name in the community” (I2).</td>
</tr>
<tr>
<td>Social aspects</td>
<td>Trust</td>
<td>Developers tend to motivate each other, send “thank you” tweets for done work, greet with work achievements or personal events (like birthday).</td>
</tr>
<tr>
<td></td>
<td>Connectedness</td>
<td>Retweet relationship with follower number showed that if a person has “name” in the community, then he has relatively large circle of followers and retweets. Trust and connectedness are dependent variables.</td>
</tr>
<tr>
<td></td>
<td>Reputation</td>
<td>Based on daily chatters and emotions expressed in tweets users create their personal virtual profile “I am also pretty honest about who I am” (I4).</td>
</tr>
</tbody>
</table>

Table 7: Use of microblogging in Drupal community

When we were performing analysis of microblogging use in Drupal community based on the theoretical framework (Table 4) we obtained new parameters for some of layers, which are described further. To the “transmission capability” in the “media capabilities” layer we have added parameter, which presents microblogging as a fast almost real time media, “because you can be up-to-date with everything really quickly” (I6). In the same layer “symbol capability” obtained two new parameters: links and documents. Statistics showed that 43.6% of all analyzed tweets contained links. There also arises new practice to attach different types of documentation in various formats like ppt, doc.“Technology features” layer was supplemented, by Twitter broadcast nature, to perform many-to-many communication. Second new parameter outlines ability to send direct messages; user’s private communication is safe and closed for publicity. Finally, technology ability to retrieve large amount of data is highlighted as interviewee states. The “interaction type” layer is also updated with several new discovered genres. There were revealed genre of a new type of documentation, where users based on the context in tweets and based on the media capabilities could perceive new way for information documentation. Regarding the developers availability on Twitter, team members are able to start conversation. In the literature there was revealed genre of user
help, based on the interviewees statements and obtained tweets, we transformed this genre to a new way of user support, because it provides help, receive questions and help to provide fast answers. “Updates” genre was supplemented by a personal achievement genre, because community members often used to present report from performed work.

5.4. PRIMARY EMPIRICAL CONCLUSIONS

Based on Section 5.2 the primary empirical conclusions (PEC) can be drawn. These will be referred to in the following section when the empirical results based on case project practices are compared with the existing literature and when the practical implications are presented. Totally 16 PECs can be found. Also, the PEC’s are divided by layers (L1, L2, etc.).

PEC1: Hashtags can be used, by anyone, to simplify information retrieval; however, they can trigger information misuse.

PEC2: “Retweet” or “favorite” functionality is used to measure individual focus on a tweet.

PEC3: Twitter group accounts are used to simplify information search and gather it by work fields.

PEC4: Lists (manageable groups of Twitter users) are used to perform data grouping.

PEC5: Team members are able to execute several conversations in parallel and do not lose the meaning.

PEC6: Various symbols help to express emotions. Thus, information is presented not as a set of words, but as a personalized thought.

PEC7: Microblogging enables many-to-many paradigm.

PEC8: Spontaneous tweets can generate useful, unique ideas and solutions; however, in some cases they provoke controversy within the development team.

PEC9: Microblogging is applicable for large groups of collaborators.

PEC10: Microblogging can be perceived as a new way of documentation by capturing system development events and knowledge from tweets.

PEC11: Developers’ availability prompts to start a conversation or find team members on Twitter.

PEC12: Microblogging can be perceived as a new way of user (developers) support - contributors can ask for help, look for solutions, ask questions and get answers.

PEC13: Work related topics interwoven with personal data create the full profile of a person.
PEC14: Requirement for software development organization to survive in today’s environment stimulates to explore newest technologies such as social media.

PEC15: Tweet openness eliminates the OSS community’s problem of person-to-person information exchanges.

PEC16: Social engagements increase trust and connection between unfamiliar people, forms relationships and connections within a group of people.
6. DISCUSSION AND IMPLICATION

This section discusses the results of the thesis. The discussion concentrates mainly on the primary empirical conclusions listed in Section 5.4 and the theoretical implications (Section 6.1) and practical implications (Section 6.2) they have.

6.1. THEORETICAL IMPLICATIONS

Empirical evidences obtained from the analyzed case study provided updated list of used media to perform communication process in distributed software development (Figure 24). Empirical research of Drupal community practices confirmed the set of tools obtained from the literature review (Figure 3) and presented an additional way of communication as issue tracking systems.

![Diagram of communication media layers and parameters in distributed software development](image)

Figure 24: Tools used for communication in case study

Zoomed in microblogging media analysis developed theoretical framework of communication media layers and its parameters in distributed software development (Figure 25). Based on the primary empirical conclusions, it seems that six layers presented in microblogging layers utilized in distributed software communication framework provide a sound understanding of what advantages and disadvantages microblogging can bring to the communication process. To explore contributors needs and understand development problems or organizational issues it is better exploring personal tweets, as they present most relevant data.

Empirical evaluation brings new waves of theoretical statements and theories. Regarding MRT characteristics microblogging is a lean media which obtains low personal focus and feedback (Daft et al, 1987). A receiver’s personal focus can be measured by the number of retweets (PEC 2). The amount of retweets (Figure 19 and Figure 20) presents followers’ interest and intention to trace shared information.

Regarding MST theory (Dennis and Valacich, 1999) microblogging is categorized as an asynchronous media. By the theory descriptions, media should provide high data rehearsability and reprocessability. This statement is supported by the empirical evidences (PEC1, PEC3, and PEC 4), which present that any flow of information could be sorted, categorized and stored for the future reviewing or for the faster search. Information search can be facilitated by using hashtags (PEC1) or by following specialized group accounts.
Manageable groups called lists (PEC4) are used to store sorted data by fields or keywords.

Microblogging subsequent and parallel conversations (PEC5) verify one of the practices described by Common Ground theory (Clark and Brennan, 1991) which tends to improve community members’ mutual understanding. Microblogging capabilities serve high information sequentially similar to face-to-face communication within time intervals; however, conversation meaning is uninterrupted.

Based on MRT (Daft et al, 1987) and MST (Daft et al, 1987) theories, microblogging is considered as a media with a low set of symbols. Empirical evidence (PEC6) contradicts this statement by introducing different scenarios, where users’ facilitates from a variety of symbols like links, videos, different document types and emoticons.

Distributed software development requires a mix of communication tools to perform different types of activities (Lanubile, 2009). Tools used for virtual communication typically suggest a one-to-one paradigm (IM), or one-to-many (forums) paradigms. Microblogging raise new paradigm (PEC 7) known as many-to-many (O'Reilly, 2005). Now contributors can input and receive information dynamically within a flexibly formed scope. The definition of "many" includes entities such as organizations, products, processes, events, and other similar concepts.

Honeycutt and Herring suggested using Twitter for collaboration and communication in pairs or small groups, because, in large, open discussions, it is less effective (Honeycutt and Herring, 2009). However, study finding (PEC9) contradicts statement presented in the literature. Microblogging is applicable for large groups as it based on broadcast nature. It provides optional reply, sequential parallel conversations and control over the data amount.

Various amount of research was done to analyze the message types posted by developers and other contributors (Zhao and Rosson, 2009; Guzzi et al, 2010; Bougie et al, 2011). This research revealed three new types. Firstly (PEC10), proposes to evaluate collected short information and references to the source of information as a new way of documentation in software development. Spontaneous tweets from developers and code contributors can best reflect current statuses, code tips, feature description, etc. The second finding (PEC11) is based on the developer’s availability online. Developers spend comparatively reasonable amount of time on Twitter. Thus, their availability decreases feedback time that creates service to find people and invite them for further collaboration in any other communication channels. Finally, user (accordingly developer) support channel (PEC12) describes media ability to collect help requests, issue solutions, questions and answers.

There is a biased opinion, usually in enterprise communities, that Twitter can bring a lot of informational noise. This is the main reason of not implementation it to corporate structure (Günther et al., 2009). However, requirement for software development organization to survive in today’s environment promotes to explore newest technologies (PEC14). Microblogging is a technology which partly fulfills the communication needs.

Summarizing theoretical implications empirical data shows that microblogging is suggested to be used in distributed software development as a tool in the composition of rich communication tools (like IM or phone) to perform high-level collaboration between distributed members (Tissen et al, 2007). The empirical study did not contain comparison of several use cases including enterprise organizations. It is difficult to say how effective the communication would be in different size and organization structure communities. Nevertheless, the interviewees’ testimonials in Section 5.2 indicate their satisfactory from communication, which conforms to the results of many previous studies (Reinhardt, 2008, Riemer and Richter, 2010).
6.2. PRACTICAL IMPLICATIONS

Based on the primary empirical conclusions, it is possible to suggest relevant practical implications for open source and enterprise organizations developing software distributedly.

Information tagging is an approach, which minimizes time spent for data search (PEC1). It is necessary to create unique tags, which would be easily mapped to the labeled words. Uniqueness is required, because tags a publicly facilitated and there can be created information misuse. Tagged words would create a set of associated data and present opinion and experience of community collaborators.

Possibility to create and maintain work based groups (PEC3) not only simplifies information search, but also creates field oriented groups. Such group maintainers are responsible to filter outgoing information and present most relevant topics. Group’s main information can be the website’s RSS feeds, or group maintainers posts. Maintainers usually tweet about topic related announcements and innovations based on relevant information source.

In large communities and companies every contributor has a large set of his co-workers and familiar people. Microblogging lists are suggested to create separate people circles (PEC4). Lists can be
distinguished to two circles: work and personal. Work list can be split to several subgroups according to the following people interests. Also, there is possible to follow outright foreign list. List functionality reduces mix of information and simplifies information retrieving.

Links, videos and attached documents provide source to the information described in short messages (PEC6). Emoticons compounded in the text highlight author’s feelings. Composition of different symbols creates a feeling of real conversation in distributed software development.

Organization which decided to incorporate microblogging as a communication practice, do not have to create rules of how employees have to tweet, the practice should not restrict people thoughts and desires to select their own way to share information. Microblogging by nature is spontaneous media; the only suggestion would be to avoid direct (designated to a particular individual), negative emotions (PEC8). Public negative tweet about a certain individual or group of people provokes issues and conflicts.

Usually, in large software development communities, many valuable pieces of information are being transmitted informally from one person to another. These mainly small pieces of (micro) information then reside in personal mail boxes or are locally stored. As a result, they are neither accessible by other team members. Hence, appears need for a new means of communication that overcomes these shortcomings by publicly shared tweets (PEC15). Collection of such tweets can be admitted as a new way of documentation (PEC10).

In distributed software development, one of the biggest issues is time zone difference and problem to reach team member. The provided availability of web, mobile, desktop and other clients for microblogging enhances contribution on the go (PEC11) and makes possible to catch person online and then move to any other communication channel.

Developers in OSS development are also perceivable as users. Thus, developers and users often need help from more experienced and competent people. Further practical implication of microblogging is to create and maintain online user support (PEC12), where competent people will answer the questions and provide online help.

There are different types of people working on the same product; one of them easily can communicate with unfamiliar people, other need to have at least comprehension about team member personality. In microblogging people used to share facts, pictures and videos from their personal life and this helps for their followers have an idea of tweeting person’s characteristics (PEC13).

In the OSS communities, information is open to everyone; thus openness in development decisions should be shared with publicity (PEC15). This also encourages people from outside to start engage to the development process.

The last primary empirical conclusion (PEC 16) says that the presence of social communication in parallel with work related topics build connection between unfamiliar or weakly familiar people. The evidence is based on interviews and tweets. The evidence supports and extends statements existing in literature that claims that social engagements provide possibilities to explore similar experiences and attitudes between team members.
7. CONCLUSIONS

This thesis has explored the role of microblogging as communication media in distributed software development. Section 2 firstly analyzed what characteristics media should obtain to perform communication tasks. Subsequently, analysis of used communication tools in DSD environment was performed. Finally, role of social media and particularly microblogging as a communication tool was observed. Section 3 introduced an initial theoretical framework with six layers. Layers are formed by parameters, which characterize microblogging capabilities and aspects as a required tool in DSD communication process. The study context, research methodology, and data collection methods were explained in Section 4 and the empirical results described in Section 5. Section 6 discussed the empirical results and identified theoretical and practical implications. This section concludes the thesis. Firstly, Section 7.1 provides with an answer to the research question. Section 7.2 presents the research limitations. Finally, Section 7.3 suggests what further study is still needed.

7.1. ANSWER TO RESEARCH QUESTION

The research question of the thesis is: How is microblogging utilized by the DSD community for its communication needs? In order to answer to the question, a theoretical framework (Figure 4) was first created and then analyzed against empirical data. This section provides with an answer to the research question.

The main research question was compounded from three supporting research subquestions, so we will answer each of them to have the answer to the main question. Figure 25 shows the final model, which at the same time is the answer to the each research subquestion.

The first subquestion was: What part of the communication process need does microblogging constitute in the DSD community? This answer is based on the information processing and communication process layers. Empirical evidence showed that microblogging was used in the community, in the case of uncertain information conveyance. That means, when users needed to find knowledge sources to answer questions or to receive common understanding, they used Twitter as a tool which provided facilities to gather different types of information. This practice was suggested for newcomers, developers who needed some help, or for those who wanted to share and receive general information. Shared information usually consisted of releases or significant changes, work-related (development, testing, documentation, etc.) topics. Microblogging features provided easy functionality to gather these types of information.

The second research subquestion was: How does microblogging affects communication performance in the DSD community? The answer is compounded from three layers: media capabilities, technology features and social aspects. Media capability evidences described how uncertain data is transmitted, collected and retrieved. Microblogging is capable to transfer information fast and almost in real time. Every message (tweet) can be reviewed before being send and reprocessed after being received. Several conversations can be performed in parallel. A high symbol variety (documentation types, links as references to the source and emoticons to express feelings) is assumed to facilitate and promote communication among team members. Technological features present media uniqueness. This layer presents how amount of short and spontaneous messages can be broadcasted from any Internet-connected device using a many-to-many or one-to-one approach. The social aspect layer is based on all the previous framework layers (L1-L5), which build trust and connection between team members. That means if there exists tie connection inside the development group the higher will be a prediction for work performance achievements.

Finally, the last subquestion was: What type of information is transmitted throw microblogging in the DSD community? The work-related and personal topics are present in software development daily communication. However, the work related topics have a variety of subtopics. Based on the empirical evidence in the daily work community, contributors share different set of information to increase team members awareness. They update information about personal achievements or product, development statuses. A new way of documentation is formed from the short development announcements, references to
the documentation, tests, issues, code tips and other types of information. The community members’ presence in social media tools enabled by not only work related communication, but also sharing of personal life facts and experiences, provides high availability and fast reachability of contributors. Microblogging is also used to ask for help, which can be transformed to a new way of user support service, where a developer is perceived as a user. We have not particularly analyzed personal data (called as “daily chatter”). We distinguished several categories such as type of conversation with other users, technology knowledge/news sharing, world events, personal feelings, and family, friend’s related facts. When we were analyzing team members personal focus on Twitter, we noticed that communication between developers based only on work tasks does not exist. Usually, everyone favors to have personal communication and to know personal life details about each team member.

All the research subquestions form the answer to the main research question, stating that microblogging can be an effective communication media tool in conjunction with richer and synchronous tools such as phone, video conferencing, or IM. Because in the distributed software development not only information conveyance is required, problems solution discussions and coordination tasks also require tools with the same layers presented in the theoretical framework (Figure 25), but with the other parameters and capabilities. However, the social aspects layer and its parameters should be present in every tool, because social ties are mostly vulnerable in distributed software development.

7.2. RESEARCH LIMITATIONS

There are certain limitations that may threaten the validity of the research conducted in this thesis. Firstly, the empirical part of the thesis contained only one case study. As Darke et al. (1998) suggest, “single cases provide for in-depth investigation and rich description. Multiple case-designs allow literal or theoretical replication and cross-case comparison”. Thus, a single case study cannot provide reliable information about the broader class, but it may be useful in the preliminary stages of an investigation since it provides hypotheses, which may be tested systematically with a larger number of cases. This means that the model proposed by the thesis should be tested with organizations of different size and in different areas. The suggested framework is novel. The empirical study was a significant initial test of it. Further research need to be performed to validate the framework in a more comprehensive manner.

Another limitation is that the aim of this thesis was to explore the microblogging’s role in the distributed software development. However, selected case analysis presents one of the DSD types– OSS development. This means that the proposed framework should be analyzed, in distributed enterprise organizations, to validate framework in all types of DSD. On the other hand, OSS development is “by its very nature globally distributed” (Lundell et al, 2006). It is stated, that an OSS community usually consists of volunteers with time and effort limits. However, they are motivated and self-selected contributors with a high level responsiveness (Lundell at el, 2006). Additionally, when researchers perform DSD case analysis, they often choose OSS organizations as a relevant sample (Bird and Nagappan, 2012), because they provide a repository of software systems, without complications of dealing with software companies for permission to access repositories, legal issues, distribution of results and confidential clauses.

Finally, a limited number of community contributors were interviewed. However, the interviewed set of people included core developer, different modules developers, tester, and documentation writer. Experienced developers and community contributors provided information about different types of tasks from a longer period of time. Regarding the literature, OSS development teams have a hierarchical or onion-like structure (Crowston et al., 2003). At the centreare the core developers, surrounding the core are the developers, surrounding the developers are the users. Active users contribute bug reports or feature requests, write documentation or translate the system. Thus, the selected interviewees cover the OSS development team hierarchy. Additionally, there were collected unobtrusive measures – tweets, which covered a larger set of community population.
7.3. FURTHER RESEARCH

There are still issues that require further study. Firstly, the proposed model should be validated with additional research. It would be especially appealing to see how well it works with different open source communities in various interest areas. Also, the amount of community contributors should vary to understand how smaller groups of developers apply microblogging in their daily work routine.

Furthermore, the model suggests three new interaction types in microblogging – “documentation”, “user support”, and “find and invite” services. For that reason, it would be valuable to see how these three types are used in different organizations, which of them get more benefit, what new aspect from these interaction types can influence community contributors’ connectedness and trust.

Additionally, microblogging has just been used for a short period, thus it would be intriguing to explore the use of this media tool for a longer period. In addition, it would be worthwhile to compare companies who started to use microblogging in the initial phase of community creation, whereas in the case of Drupal community the members started to use it after Drupal 5.0 release (now Drupal 8.0 is in development).

Finally, even though the focus of the thesis has been on distributed software development, the main components of the model and many of the practices in it are not restricted to the distributed software development only. For that reason, it would be intriguing to see how the model could be adjusted for the needs of co-located enterprise using microblogging as a communication tool.
REFERENCES


APPENDIX 1. INTERVIEW QUESTIONS

1. Information about the interviewee
   a. Job title and responsibilities.
   b. Time within the organization.
   c. Why you decided to participate in Drupal community?
      i. Why this particular community?
   d. What types of tasks have you performed?
   e. In what projects do you participate now? What tasks you perform now?
   f. Do you work in groups or individually?

2. Project and task details (Media analysis through development process):
   a. How work is divided on the project?
      i. Whether developers are assigned roles?
      ii. Whether people work primarily in one area or several?
      iii. Whether there are official restrictions on work area?
   b. Provide example with detail illustration what communication tools do you use and from what media parameters you facilitate in one of these tasks (regarding your practice and where have you participated):
      i. Accident/Incident/Issue management
      ii. Planning/Task coordination
      iii. Testing/developing
      iv. Document contribution/Drupal translation
      v. Event updates in public (releases, feature updates, any changes in the organization)
   c. Types of information you gather about team members’ activities?
   d. What tools are used for work coordination
   e. What information sources exist to provide group awareness?
      i. How you use those sources to find and maintain knowledge.
   f. List advantages/disadvantages of using a particular tool.
   g. What type of information do you need? Do you usually need information conveyance or convergence?

3. Communication in general
   a. How often do you communicate?
   b. How much time you spent for communication (average per day/week)?
      i. Do you usually have (non) real time discussion? How long it takes to get a response/feedback?
   c. Describe communication in teams/between teams/individual?
      i. Do you and your colleagues use the same/different tools to communicate?
      ii. How often do you communicate with team members?

4. Experience
   a. What was your personal experience when you started to use media in this community?
   b. Do you use the same accounts in social media tools for work and personal life?

5. Problems and future plans
   a. Can you enlist existing problems in the communication process in Drupal?
   b. Your suggestions for future
      i. How would you improve communication process?
      ii. What media you would suggest using less or more to improve communication process?
APPENDIX 2. INTERACTION TYPES IN CASE STUDY

To do: 3.30%
User support: 5.42%
Updates: 26.94%
Share information: 22.84%
Documentation: 37.96%
Find and invite: 3.50%
Release updates: 10.88%
Knowledge share: 22.80%
Gratitude: 6.80%
Direct conversation: 11.49%
Idea/solution code tips: 4.85%
Event updates: 18.99%
Personal achievement: 7.93%
Coordination: 0.97%
Personal: 2.03%
APPENDIX 3. #DRUPAL CLUSTER DENDROGRAM IN the PERIOD OF MARCH 23-26
APPENDIX 4. #DRUPAL CLUSTER DENDOGRAM IN the PERIOD OF JUNE 18-23