UNIVERSITY OF TWENTE.

Fostering Engagement in an Online Community of Practice

Author   Wouter Bosch
Email address   w.h.t.bosch@student.utwente.nl
Education   Industrial Engineering and Management
                 Bachelor

Institute   University of Twente
                 Faculty of Behavioural, Management and Social sciences
                 Enschede, the Netherlands

Internal supervisors
1st supervisor   Dr. C. Amrit
2nd supervisor   Dr. R. Effing

External supervisor   T. Potze

Date   July 2018

Management summary

In modern society, the number of online communities keeps expanding. Because of the online development, a lot of organisations see their opportunity to connect people with the same interest. For an organisation or company, there is an added value in joining these people. At the moment the knowledge on how to manage a community is available. Pillars are stated and roadmaps are created. Of course, every community differs which ensures difficulties. However, as a manager of a community, there is no visualisation of how the community develops.

In this research, the focus lies on the engagement in a community of practice. With as main goal to determine what the most important factors for the measurement and presentation of engagement in an online community of practice are. It is important for community of practice to have a common goal, a place to practice and people willing to share information. The focus can shift over time if it serves the purpose of the community.

Before knowing how a community is doing, it is important to understand in which phase they are. The most important phase is the "growth" phase. This is the phase which determines if a community evolves into a formal organization. The growth of the number of members, the interaction between members and the transparency to each other are the most important factors in this phase. If a community does well it ends up in the maturity phase. Which means that the community is self-regulating and self-sustainable.

By testing the literature on an exciting community, the relevance is determined. To keep track of the development it is important to look at the ratio in the community. Even though some trends go down it does not necessarily mean the community is not doing well. Knowing the targets and goals of a community is needed to interpret the data.

Table of contents

1. Introduction

The trends that cause society and our daily lives to become digitized, do not leave organisations untouched. In particular, communication within companies and how knowledge is shared is greatly impacted by new digital communication media. One example of developments in this field are the formation of online communities of practice (**CoP**): simply put, a group of people who share a common interest or goal. Organizations like Greenpeace, the United Nations and non-profits like Pachamama rely increasingly for their day-to-day operation on online CoP's. However, many organisations struggle with getting their members active. It follows that these kinds of organizations are invested in maintaining the health of these communities and promoting their effectiveness.

The literature suggests several factors which are important to a community. A Key Performance Indicator for CoP's is the level of engagement that the members of the community feel. It is therefore of crucial importance to be able to measure the level of engagement in online communities of practice. However, there is currently a painful lack of a tool which facilitates research into the engagement within an online CoP.

The aim of this research is to solve this problem and develop a tool that enables measurement and analysis of the (development of) engagement within a CoP. For such a tool to add value to a community, the factors that are important for that community need to be presented in such a way that the development of the community can be monitored. This would allow insight into the effects of interventions and ideally also the prediction of trends.

Since knowledge about the status of a community is crucial for community managers, this research is especially relevant for them. Insight and understanding into community health will enable community managers to achieve progress within the community. Gaining insight into the development of CoP's adds value to the community and makes it possible to set clear targets for its further growth.

A goal for every community is to have active members, which takes a lot of steps. Because every community is different it is hard to determine which exact steps should be taken for a particular community. However, by monitoring the development in the community the influence of a decision can be detected. Analysing this information can be a basis for insights into how a particular community should be managed and what a more efficient way of establishing engagement in a community could be.

Several steps are conducted in this research. First, the research concentrates on what discerns a CoP from the different other possible communities. Secondly, the different phases of maturity of a CoP will be discussed. To better explain and visualize the concept of a CoP, a case study has been conducted into Pachamama, a CoP. After gaining insight into the literature on CoP's, it will be benchmarked against a "real" CoP, to see if the literature connects with the practice. An available database is analysed by structuring and organising the data into a data warehouse. Several factors of engagement of a CoP will be discussed and based on analysis of a case study it is concluded that qualitative understanding (e.g. the goals of a community) is needed to supplement quantitative measurement of engagement.

In the context of this research, there are several key stakeholders. First, there is the company facilitating the researched online community: GoalGorilla. GoalGorilla offers its customers an online community software package called Open Social. Open Social allows organisations to evolve their community from the offline world into the online world with standard functions.

Next to the theoretical contribution through literature review and validation, this research offers a strong practical contribution for GoalGorilla. A dashboard which indicates the engagement of an online community would function as an extra tool to convince potential clients. The overview of the data about a community that the proposed tool will provide allows GoalGorilla to explain and demonstrate the added value of the Open Social platform more easily. Therefore, for GoalGorilla it is important to know which attributes contribute to increasing the quality of an online social community.

Since analysing the quality of engagement points out which communities are underperforming, the developments in a community can be acted on even before they happen.

For a potential client, it is important to know if the purchase of the product Open Social has added value for their company. A dashboard can show the level of commitment in a community, how the community is adding value for the company and developments can be monitored and acted upon. The impact of those actions can then directly be measured. This short feedback loop leads to more effective management and utilization of the online community.

The research will be based on a case study of an online community of practice called Pachamama. For Pachamama, it is interesting to know how they are doing and if they are reaching their goals.[1]

The research is guided by the research question:

What are important factors for the measurement and presentation of engagement to improve the effectiveness and value of an online community of practice?

To get an answer to this main question, sub-questions are formulated:

1. What defines an online community of practice?
2. What are the phases of an online community of practice?
3. What are important factors according to a community manager of a CoP

These answers to these sub-questions will be tested against the data available for this research to test and verify the research.

---

[1] More information about Pachamama will be given further on from page 18 onwards.

2. Research methodology

To answer the research question, a goal-reasoning approach is used where a domain-driven and a data-driven approach are combined. This chapter explains the methodology of the domain and data-driven approach. After these approaches, the results will be combined in an ESc (Engagement Scorecard). Based on this ESc the analysis will be conducted.
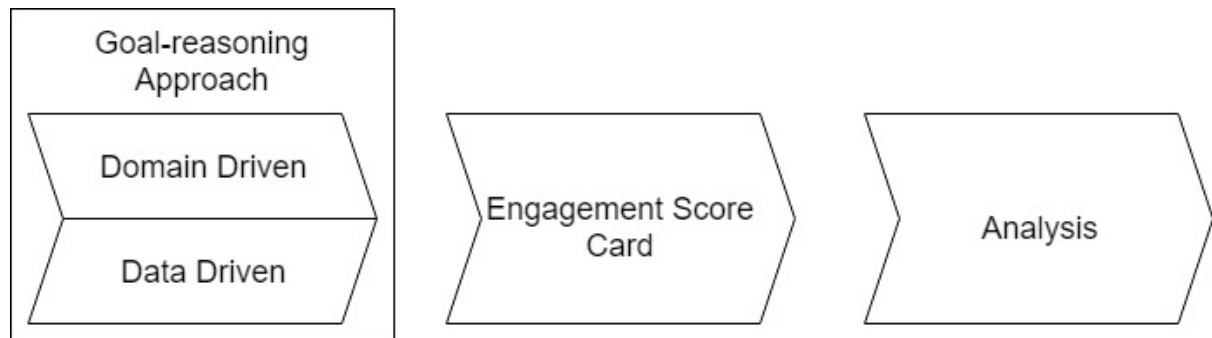

Figure 1: Research Method

2.1 Domain-driven

A domain driven approach is used to create an understanding of the domain which will be analysed. To comprehend the whole domain which needs to be studied the domain is divided into the following subdomains:

- CoP in literature
    - Through literature study the defining characteristics and factors for performance and quality of a CoP are determined. A focus lies on what metrics determine the amount of engagement in an online CoP.

- Phases of an online CoP in literature.
    - Through a literature review, the phases of the life-cycle of an online CoP will be established and the differences in what factors are important in each phase are drawn out.

- CoP's in practice
    - Through an expert interview with the community manager of Pachamama, the literature is compared and contrasted with experiences from practice.

2.2 Data Driven

For the case-study, a data-driven approach is used. The case-study makes use of data from the Pachamama online community of practice. After conducting the domain driven approach, the feasibility of the results needs to be determined. This will be done through the data-driven approach. Based on analysis of database provided by GoalGorilla the following tools are used during the research:

- Data warehouse (MySQL Workbench)
    - Since every action on a website can be saved and the designer of the database decides which actions are saved and which are not, the data needs to be organized into a data

warehouse to focus the search. The data warehouse will make sure the data is structured in such way it is easily interpretable later.

- o The biggest reason to use MySQL Workbench is that it is easy to use and it has all the abilities needed for making a data warehouse.

- Tableau Prep
  - o Tableau prep is a relatively new kind of ETL (Extract, Transform and Load) programme which is used for pulling the data out of the database and placing it into the data warehouse. In this program, the data will be structured in the way determined by the data warehouse.
  - o There are several other ETL programs which could have been used. The reason Tableau prep was chosen is because it works smooth and easy. Furthermore, the visualisation will be done in Tableau and the connection between these programs is perfect. The only disadvantage is that it is not open source software.

- Tableau
  - o After structuring and organising the data determined by the data warehouse it is visualised. Adding the "cleaned" tables and connecting relevant data can be done in a visualisation program like Tableau. Tableau is used because of its wide range of options compared to other open source visualisation programs.

2.3 Engagement Scorecard

After the domain and data-driven approaches are conducted the results are combined. To keep track of the executions of activities a balanced scorecard is used. However, the normal Business Scorecard (BSc) is not applicable in this instance. Therefore, the BSc is transformed for this research into an Engagement Scorecard (ESc). The justification of using this tool is to give an overview of the used strategy and gives an understanding of the chosen metrics. As the normal BSc also focuses on finance and the CoP researched is a non-profit organisation, an ESc is used.

2.4 Analysis

After defining a CoP, which stages it goes through, the goals of Pachamama and an overview created by the ESc, the analysis can be performed. The structured data can be presented in graphs whereby trends can be noticed and explained. All the research areas can be analysed and the state of the engagement in the community of practice of Pachamama can be concluded.

3 Theory

To get an understanding of the topic, a literature review on a CoP is conducted and other types of online communities are elaborated upon. To get knowledge on the development of a community a literature review on the phases of a community is conducted. Furthermore, the essentials and biggest risks are looked at.

3.1 Community of practice

This chapter explains what a community of practice is. Which facets are important to a community of practice and which other types of communities are possible?

3.1.1 Characteristics of a community of practice

In short, a community of practice is a group of people who share a common concern, a set of problems, or interest in a topic and who come together to fulfil both individual and group goals (Wenger 2002). However, a CoP can evolve naturally because of the member's common interest in a domain or area, or it can be created deliberately with the goal of gaining knowledge related to a specific field. It is through the process of sharing information and experiences with the group that members learn from each other, and have an opportunity to develop personally and professionally (Lave 1991).

To create a better understanding of the meaning of a CoP a further research on practice is needed.

3.1.2 Practice

As (Edelman 1993) and (Clancy 1997) argue to engage in practice, we must be alive in a world in which we can act and interact. We must have ways to communicate with one another. But the focus on practice is not merely a functional perspective on human activities, even activities involving multiple individuals. It does not address simply the mechanics of getting something done, individually or in groups; it is not a mechanical perspective. It includes not just bodies (or even coordinated bodies) and not just brains (even coordinated ones), but moreover that which gives *meaning* to the motions of bodies and the working of brains. Practice is thus about meaning as an experience of everyday life.

To associate practice and community three dimensions of the relation by which practice is the source of coherence of a community are summarized:

1. Mutual engagement
   - Practice does not exist in the abstract. It exists because people are engaged in actions whose meanings they negotiate with one another. Membership in a community of practice is, therefore, a matter of mutual engagement.

2. A joint enterprise
   - The result of a collective process of negotiation of a joint enterprise is that it reflects the full complexity of mutual engagement. Therefore, it is defined by the participants in the process of pursuing it, despite all the forces and influences that are beyond their control. It is not just a stated goal but creates relations of mutual accountability among participants.

3. A shared repertoire
   - The coherence is not gained in and of themselves as specific activities, symbols or artefacts but from the fact that they belong to the practice of a community pursuing an enterprise. The repertoire of a community of practice includes routines, words, tools, ways of doing things, stories, gestures, symbols, genres, actions or concepts

that the community had produced or adopted in the course of its existence, and which have become part of its practice. (Cook 1996)
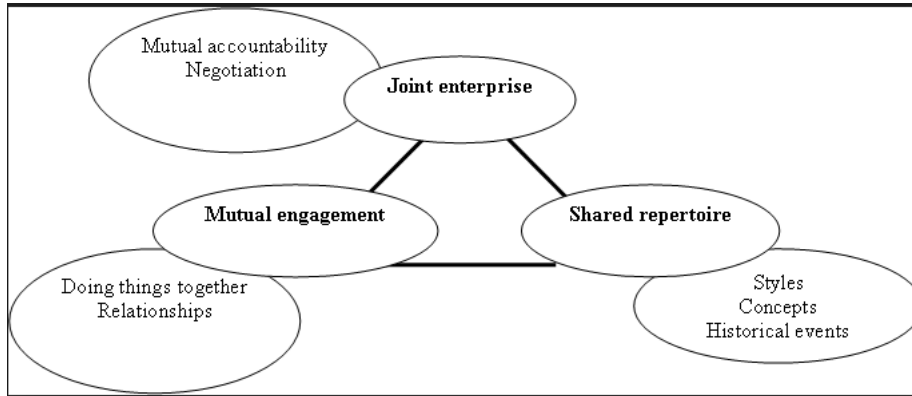

Figure 2: Dimensions of practice as the property of a community (Wennger 1998)

After conducting some further research (Wenger 2002) concluded that a community of practice is a unique combination of three fundamental elements:

- **Domain**
  - A domain of knowledge creates common ground, inspires members to participate, guides their learning and gives meaning to their actions.

- **Practice**
  - While the domain provides the general area of interest for the community, the practice is the specific focus around which the community develops, share and maintains its core of knowledge.

- **The community**
  - Within their domain of interest, members engage in joint activities and discussions, help each other and share information. They build relationships that enable to learn from each other.

The three main aspects needed are, first, common ground. People in a CoP should have a common ground as a basis for their conversations. There also needs to be a focus around the community which ensures the development and maintenance of the core knowledge. Last but not least willingness and interaction between members in the community is needed. A good common ground and (for example) fifteen community members cannot do the work alone. The members need to be willing to share experiences and expertise. If this is not the case the community will not survive.

3.1.3   Other types of online communities

There is not a lot of literature which determines different kinds of online communities. However, the ones that do exist, agree on the basics.

With providing customers with the opportunity to interact with each other as well as with the company, organizations can foster deeper buyer relationships by customizing products and services

to meet consumers demands and interests (Armstrong and Hagel Iii 2000). Armstrong and Hagel focus their attention on the following four types of online communities:

- Communities of Transaction
    - Facilitating the buying and selling of goods and services

- Communities of Interest (also known as communities of practice)
    - Communities who interact extensively about specific topics or interests.

- Communities of Fantasy (also known as gaming platforms)
    - These communities allow participants to create new personalities, environments and stories.

- Communities of Relationship
    - These are the centre of intensely personal experiences and generally, adhere to masking identities and anonymity.

The first one to argue this are (Hummel and Lechner 2002). After conducting an analysis of 50 online communities they identified five genres of communities. These genres are **games**, **interest or knowledge** (community of practice) and three other mixed genres also oriented to transactions, **business-to-business** (knowledge and transaction), **business-to-consumer** (interest, commerce, and transactions) and **consumers-to-consumers** (interest, trade and transaction).

These five genres are constructed in such a broad way that others might define online communities as different or more specific, however, they always fit into this framework.

3.2 The phases of a CoP

In this chapter it will be determined which phases a community of practice can go through, what the success factors are per stage are, but also where a community of practice should look out for death traps.

3.2.1 Possible stages of a CoP

Every living thing grows. Things are not born in their final state. They will have to evolve and will go through lots of transformations. Every stage the community goes through presents distinct characteristics and needs. Community building efforts must take into consideration the needs of members and of the whole community in each stage (!!! INVALID CITATION !!! (Malhotra, Gosain et al. 1997, Preece 2000, Kling and Courtright 2003)). There are some different views on which stages an online community of practice goes through.

(Wennger 1998) identified five stages of creating an online community of practice:

- Potential
  - People face the same situations but did not form a shared practice yet.
- Coalescing
  - A common emerging point is established.
- Maturing
  - Relationships develop, standards and agenda are defined.
- Active
  - The most productive phase. Members develop shared practices.
- Dispersed
  - The CoP is no longer active and functions as a repository of knowledge.

(Malhotra, Gosain et al. 1997) illustrated four stages of evolution and design:

- Inception
- The beginning of user involvement
- Interactivity
- Growth and experimentation activities.

A newer concept, combining these studies is made in 2009 (Iriberri and Leroy 2009). They argue that the life cycle of every community contains five stages as shown below. (Iriberri and Leroy 2009) were the first to review thousands of academic articles on community development and lay out a clear set of stages and success factors.
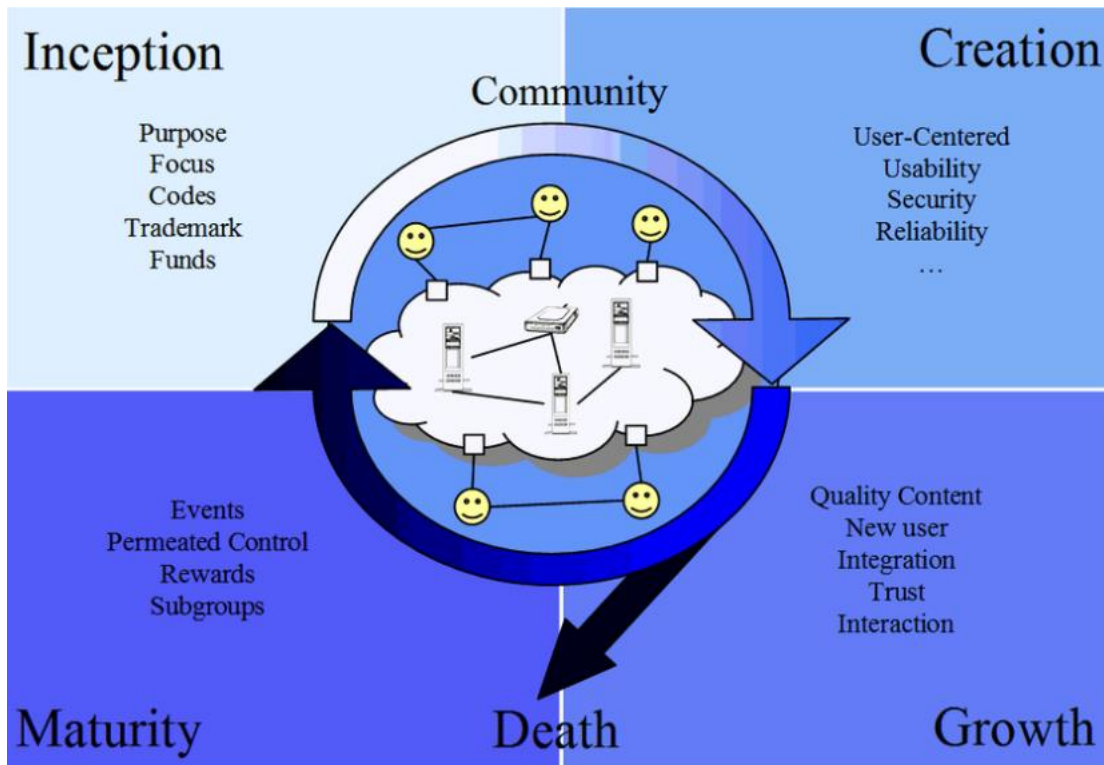
Figure 3 Lifecycle of a community of practice

Stage 1: Inception

During inception, the idea for an online community emerges to satisfy a need for information, support, recreation, or relationship. The following success factors are necessary for all types of communities in this stage:

- Purpose
  - Before an online community is created the creators must have a clear purpose for the community. These purpose needs to be visible to potential members, so they can decide to participate or not. Also, Maloney-Krichmar and Preece (2005) found evidence of the added value when there is a clear purpose.

- Focus
  - Creators must decide on the need they will address and identify the characteristics of the target audience. Wegner et al. (2002) recommend also specifying clearly the community's area of interest to its members.

- A-code of conduct
  - Principles, values, standards, or rules of behaviour that guide the decisions, procedures and systems of an organization in a way that (a) contributes to the welfare of its key stakeholders, and (b) respects the rights of all constituents affected by its operations (IFAC). This needs to be established to give the users clear guidelines on how to behave. Proof of the importance of the code of conduct is given by the study done by Leimeister and Krcmar (2003).

- Trademark
    - Kim (2000) emphasizes the need for a tagline that differentiates the community and expresses its nature. She suggests that an appealing tag would trigger the desire to participate.

- Source of revenue/ funding
    - To fund the online community, sources of funding need to be secured. The goal of the creator can influence what type of funding will be established.

Life Cycle Stage: Inception	

Success Factor	*Author*
Purpose Purpose Transparency of goals	(Maloney-Krichmar and Preece 2005) (J. 2000) (Leimester 2005)
Focus Target Audience Focusing on one target group	(Andrews, Preece et al. 2001)
Codes of Conduct Establishing codes of behaviour. Facilitator to monitor and control behaviour	(J. 2000) (Leimeister 2003) (Preece 2000)
Trademark Building a strong trademark Tag line	(J. 2000)
Source of revenue/funding Defining sources of revenue	(Leimeister 2004)

Stage 2: Creation

In the creation stage, creators select the technological components that will support the online community based on the needs of potential members and the purpose of the community. The focus of the creators must be on the needs of the users and must ensure that tools are usable.

Life Cycle Stage: Creation	

Success Factor	*Author*
User-centred Design and Evolution Evolution of the community according to the ideas of its members Knowing member preferences can maximize benefits to members Specific target groups Design with users in mind Focus on the needs of members	(Leimeister 2004) (Andrews, Preece et al. 2001) (Kollock 1996) (Williams 2000) (Cothrel and Williams 1999)
Interface Usability Intuitive user guidance/usability Sophisticated user interface Ease of use	(Ginsburg and Weisband 2004) (Tedjamulia, Olsen et al. 2005) (Maloney-Krichmar and Preece 2005) (Preece 2000)

Simple and easy to use interface	(Nonnecke and Preece 2000) (Andrews, Preece et al. 2001)
Security and Privacy Handling member data sensitively Access-rights structure Security	(Leimester 2005) (Leimeister 2003) (Leimeister 2004) (Andrews, Preece et al. 2001) (Williams 2000) (Hummel and Lechner 2002)
Identity Persistence Ability to identify other members Ability to learn the history of other members	(Hummel and Lechner 2002) (Kollock 1996)
Reliability Stability of the website Reliable interface	(Andrews, Preece et al. 2001) (Maloney-Krichmar and Preece 2005)
Performance The fast reaction time of the website Performance	(Andrews, Preece et al. 2001) (Leimeister 2004)

Stage 3: Growth

In the growth stage attracting new members is one of the most important objectives. The platform needs to grow, and new members need to join the community. It is also important to directly let the new members contribute. This can be done by, for example, starting conversations or mailing them. At the same time, the community must make sure that the quality of the content is up to speed. A first impression is important to convince new members to contribute. Finally, the community needs to build trust among their members. Members are more likely to trust other members if they have a filled out profile. This creates transparency and the fact that they are there to contribute.

Life Cycle Stage: Growth	
Success Factor	*Author*
Attracting Members The existence of an offline customer club as a starting advantage Real life status symbols Actively encourage new members to join Offering privileges or bonus programs to members	(Ginsburg and Weisband 2004)
Growth Management Continuous community-controlling regarding the growth of the number of members Sending reminders to contribute Setting numeric goals for contributions Framing similarities of opinion and uniqueness of contributions	(Beenen, Ling et al. 2004) (Ludford, Cosley et al. 2004)
Integration of New Members Assistance for new members by experienced members Room for long-term users and newcomers	(Maloney-Krichmar and Preece 2005)
Content Quality	(Brazelton and Gorry 2003) (Sangwan 2005) (Tedjamulia, Olsen et al. 2005)

Offering high-quality content Knowledge stewards to organize, upgrade, distribute knowledge Content generation by the host Interesting content Competent content management Quality of content	(Leimeister 2003) (Leimeister 2004) (Andrews, Preece et al. 2001) (Zhang 2003) (Leimester 2005)
Interaction Support Encouraging interaction between members Member directory, photographs and video clips, commenting features and recommender systems	(Zhang 2003)
Trust Building trust among the members Member directory, photographs and video clips, commenting features and recommender systems and matching profiles Clear identification of operators Member profiles Transparency of providers	(Zhang 2003) (Andrews, Preece et al. 2001) (Donath 1999) (Leimeister 2003) (Kollock 1996) (Kapoor, Konstan et al. 2005) (Leimester 2005)
Transparency The increase of market transparency for community members Trustworthy operators Affiliation to established, reputable organizations	(Leimester 2005) (Andrews, Preece et al. 2001)

Stage 4: Maturity

Provided that the other phases are past successful, online communities will mature into formal organisations. As (Andrews, Preece et al. 2001) and (Ginsburg and Weisband 2004) argue that creators and managers should facilitate the formation of subgroups, delegate control to volunteer subgroup managers, organize online events and reward and acknowledge members participation and contributions. The most important factor in the maturity phase is the recognition of members contributions. To stay a healthy community it is important that the active members are rewarded for their contributions. In this way, the community does not lean on the community management but on its members.

Life Cycle Stage: Maturity	
Success factor	*Author*
Regular Online Events Arranging regular events	(Andrews, Preece et al. 2001) (Williams 2000)
Permeated Management and Control Integration of the members into the administration of the community Volunteers are critical to provide 24/7 service Distributed delegation to group operators Support for volunteerism Membership roles	(Ginsburg and Weisband 2004) (Maloney-Krichmar and Preece 2005) (Leimeister 2003) (Andrews, Preece et al. 2001) (Williams 2000) (J. 2000) (Cothrel and Williams 1999)

Facilitators to monitor and control behaviour Invitation only-subgroups	
Recognition of Contributions Appreciation of the contribution of the members by the operators Recognize existing volunteers with explicit reward model Real life status symbols (identity of contributors) Recognition of participation: by name, identity, positive feedback Recognizing the uniqueness of contribution and benefits to the group Extrinsic Rewards: gift, social recognition, feedback Visibility of contribution Incentives must match user values	(Ginsburg and Weisband 2004) (Chan, Bhandar et al. 2004) (Andrews, Preece et al. 2001) (Beenen, Ling et al. 2004) (Hars and Ou 2002) (Butler 2005) (Hall and Graham 2004) (Tedjamulia, Olsen et al. 2005)
Member Satisfaction Management Continuous community-controlling with regards to member satisfaction Focus on user needs	(Leimeister, Sidiras et al. 2004) (Cothrel and Williams 1999)

Stage 5: Death.

As (Wenger 2002) explains that people may begin to leave the community when it is not longer useful for them. It is important that the community watches out for death traps and even should protect the values they have. This can be done by keeping track of the following factors of death.

Life Cycle Stage: Death	
Factors of death	*Author*
Undersupply of content Poor participation Unorganized contribution Transient membership Members with weak ties Willingness to share information Lack of anonymity Concerns about privacy and safety Shyness about public posting Time limitations	(Jarvenpaa, Knoll et al. 1998) (Nonnecke and Preece 2000) (Iriberri 2005) (Zhang 2003) (Constant, Kiesler et al. 1994)

3.2.2 Essentials to a community of practice

Investment in community management is imperative to an online community's success (Blanchard and Markus 2004). Moreover, an organisation needs to:

- Establish and understand the domain of the proposed community (Iriberri and Leroy 2009).
- Develop and sustain a community management strategy according to the community's life cycle stage (Iriberri and Leroy 2009).

- Foster a sense of community (Blanchard and Markus 2004).

3.2.3 The biggest risks

The biggest risks for a CoP are a lack of a group of core members. Core members ensure fresh ideas and content. Normally such a group emerges at an early stage of the CoP and should remain stable (Borzillo 2007).

Another big risk for a CoP is a lack of one-to-one interaction (Borzillo 2007). Helping each other solve common problems adds value to the engagement and community.

Furthermore, the transition from the growth phase to the maturity phase will take a long time. The most important thing during this transition is to keep track of the developments. That is why the death traps summed up in the literature review about the phases of a CoP are important.

4 Important factors from the perspective of a community manager

To find out what important factors for a CoP are according to a community manager an interview was conducted. Among communities, the goals and targets may differ because the approach does. To give a little understanding about the Pachamama community some background information is given. Pachamama is the community that serves as the case study for this research.

To also understand the view of the community manager and the current state of the community an interview was conducted[2]. Due to the fact that the community manager of Pachamama lives in the United States, the interview was held online. To understand their ways of approaching the community and understanding what they would like to achieve was the main motivation for this interview, besides to comparing and contrasting the practical experience form the community manager by what is suggested by the literature.

Knowing where they started, how they developed and where they would like to go gives the view of engagement in their community other perspectives. It creates a clearer few on the data and more understanding about why things happened like they did. If the background or the reasoning of the CoP was not known the conclusions cannot be grounded. If you would like to analyse data, you need to understand the data!

4.1 The Pachamama Alliance

More than two decades ago, the elders and shamans of the Achuar people of the Ecuadorian Amazon began having dreams that foretold a threat coming to their territory. They saw how oil development was destroying the land and culture of their indigenous neighbours, and that threat was on the move, coming ever closer to their pristine rainforest home. This was the start of the Pachamama Alliance.

In 1995 they reached out to the modern world. The Pachamama Alliance was born. Today the Pachamama Alliance is a global community that offers people everywhere the inspiration, education and tools to become "pro-activist" leaders in the growing worldwide movement to create and foster a mutually enhancing human-Earth relationship.

In the beginning, the mission of Pachamama was to empower indigenous people of the Amazon rainforest to preserve their lands and culture and, using insights gained from that work, to educate and inspire individuals everywhere to bring forth a thriving, just and sustainable world.

After some years the vision shifted a little bit. Of course, the origins still are at the Amazon, however, the focus lies on a world that works for everyone. An environmentally sustainable, spiritually fulfilling, socially just human presence on this planet.

Since 2014 the Pachamama Alliance started using an online platform to create events where volunteers could learn more about the Pachamama mission. The platform supports organizing offline events and gives access to symposium materials. After using this platform for several years, the community became aware of a need for more functionalities. For example, the opportunity to create and collaborate in a space of partnership and possibility with people all over the world who are passionate about creating a new future for humanity. This new online platform would fit better with

---

[2]The interview with the community manager of Pachamama was conducted on the 30th of June 2018.

the vision of the Pachamama Alliance. Thus, the partnership with GoalGorilla (Open Social) started in March 2017.

4.2 Goals of Pachamama

After conducting a literature review on online communities of practice, a link with the real community must be made. Knowing what Pachamama would like to achieve with their community and understanding where they would like to be in several years is important for the development of the dashboard. The dashboard will be made for the specific community. To interpret the data and trends of the community knowing what they would like to achieve is essential. To understand their goals, an interview with one of the community managers was held.

At the beginning of the partnership Open Social and Pachamama Alliance, the administrators only asked small groups of people to join them because they knew they were interested in the subject. From September 2017 all the users from the old platform transferred to the new one. As mentioned, Pachamama already had offline events before using Open Social. However, because of the use of Open Social the range can be enlarged. They would like to reach people near local events so the offline events can grow all over the world. Also, they want more discussions (in closed groups) to be created. These will result in people with the same interests and expertise discussing with each other all over the world about the same problems. Connecting these people is the main goal. Also, they would like an increase in membership and an increase in participation. In 5 years they would like to have achieved that people know where to go if they share the same interests and that they are excited to log in because they know that more people think the same way. Eventually, inspiring each other and therefore work together in a self-sustainable world is the vision.

Summarizing, the goals of the Pachamama community manager would be described as:

- Increase of memberships
  - More members registered in the community
- Enlargement of local events
  - More enrolments at events
- More discussion groups
  - More closed groups in the community
- More participation
  - More engaging members

4.3 Why is Pachamama a CoP?

To find out if the Pachamama community has the characteristics of a CoP the key characteristics need to be researched. For Pachamama to be a CoP it needs to have a domain which creates common ground. *"The Pachamama community is an online community where you can meet a collaborate with people from all over the world who are passionate about creating a new future for humanity*."[3] Everyone in the community strives for a better environmentally sustainable future.

It also must have a practice, a specific focus around which the community develops, shares and maintains. This started in 1995 when the founders tried to save the rainforest of the Amazon. Connecting to people and doing things together proves that Pachamama has a worthy practice.

Finally, a CoP should have individuals who have the feeling of belonging to the community. The personal goals are corresponding to those of the whole and member feels they matter in the

---

[3] Connect.pachamama.org

community. In the community of Pachamama, there are a lot of volunteers which organize events for others to learn more. That proves that people are willing to put in extra effort for the community without getting something concrete back for it.

4.4 The current phase of Pachamama

According to the literature study, there are four different phases Pachamama could be in. In the inception phase purpose is created and in the creation phase, a platform is created. Pachamama is past these phases. It has a clear purpose and an online platform for several years.

In the interview, the community manager gave some pointers to they would like to achieve, for instance: getting more members and more activity per member. This indicates that they are in the growth phase. However, they already have a lot of events which suggests a maturity phase. The biggest reason why Pachamama is still in their growth phase has nothing to do with how big they are or the events they organize but the fact that the community is not (yet) self-sustainable. Currently, there are content managers working to keep the community up and running. A community in the maturity phase should self-regulate. The community should decide the way they want to go. Pachamama is getting there since people are creating events on their own and a lot of content is created by members. However, they are not there yet.

5 What data defines a CoP?

The previous domain-driven section concluded that a CoP can be divided into 3 concepts. These concepts are explained below. Besides these concepts, the metrics which indicate the concepts are noted. These metrics will determine the amount of engagement or quality of the CoP.

Platform engagement - (Butler 2001, Li, Kankanhalli et al. 2016) – After being launched a healthy community continues to grow. In the beginning, this will be fast and after the community gets older this growth will experience a slower rate. As can be seen in the study of Butler, size has a significant impact on the ability of online social structures to attract and retain members. Being continuously engaged in attracting and retaining members ensures more interaction on the platform, positively contributing to the "health" of an online community.

Content quality – (Brown 1991, Preece 2001, Soroka and Rafaeli 2006, Borzillo 2007) - Everybody from a community can post something. However, the content of the post can differ. That is why the quality of the shared knowledge is important for the health of the online community. Brown argues that the member of a CoP should work together on a regular basis to find solutions to common problems, and then evaluate the achieved the results together.

Like every other site or even store, popularity contributes to the reach of the community. If a lot of people visited the site or even saw the content of the community it will extend the awareness of the community. Also, if a lot of people saw a post, comment or even the site it indicates that the quality is high. The higher the quality, the "healthier" a community of practice. Also, people are not only attracted but are also motivated to return and contribute to communities that feel animated and vibrant.

Customer Engagement (Preece 2001)- Interactivity might be the biggest reason why online communities exist. For example, getting a question or asking if people agree with a plan you made. The more people reacting to a post the higher the interaction level is. The higher the interaction level is, the better the community functions.

Likewise, the speed whereby members of the community will respond to each other confirms the involvement of the group. The faster somebody reacts the better this is for the community.

Concepts	KPI
Platform engagement **(Domain)**	#members #new members #engaging members #new engaging members #first-time visitors #of posts #of comments
Content quality **(Practice)**	#shares Average time spend on content Average #links in a post #pageviews #different pageviews #logins #lurkers Length of a post
Customer engagement **(Sense of Community)**	The average speed of respond Participation ratio % Event attendance #events Average filled out profiles Average #comments Average #content posts #likes

5.1 The available data

The scope of this research is limited to the data concerning the customer- and platform engagement as this is the most relevant to determine the engagement of the community. To see which data is available the database is checked. The following framework is available in the database from the Pachamama community.

Platform engagement	Customer Engagement
#members	#of comments
#new members	Average filled out profiles
#engaging members	#likes
#new engaging members	Participation ratio %
#of posts	Event attendance
Average #posts	#events
	Average #comments

The content quality is important for a community as well. Delivering high-end content will add value for the members and will lead to new members. However, the content quality is not taken into account because this research is conducted on the engagement of an online community. Furthermore, the KPI's formulated above are not measurable by the given database. So even if the content quality would matter for the research it would not have been possible due to the design of the database.

5.2 Death Traps

To make sure the biggest risks in the transition from a growth phase to the maturity phase are accounted for the following table is made. This table shows which metrics could indicate a death trap. So, for example, if in the graphs made in Tableau is visible that the number of comments is dropping significantly it would indicate that the ties between the members are getting weak. This is one of the death traps according to literature. So, if this is noticed actions should be taken.

Metrics	Death traps
#of comments	Members with weak ties
#likes	Members with weak ties
Participation ratio %	Poor participation
Event attendance	Poor participation
#events	Poor participation
Average filled out profiles	Willingness to share information
#new active members	Willingness to share information
#posts	Willingness to share information
Average number of contributions per active member	Undersupply of content
Average #comments	Undersupply of content
#active members	Undersupply of content
Average #posts	Undersupply of content

Figure 4: Metrics linked to death traps

6 Engagement Scorecard (ESc)

A balanced scorecard is a performance metric used to identify and improve the internal functions of a business and their external outcomes. The balanced scorecard is also used as a strategy performance management tool. This is used by managers to keep track of the execution of activities

by the staff within their control and monitor the consequences arising from these actions. (Limited 2017)

The four balanced scorecard perspectives are (Amrit 2014):

1. Financial
   a. How do we look to our Shareholders?
2. Customer
   a. How do our customers see us?
3. Internal Business Process
   a. What should we do that is excellent?
4. Employee and Organization Innovation and Learning
   a. Can we continue to improve an add value?

The focus of the research lies on the engagement part of the community. To make use of the framework of the BSc but focus on the engagement it is transformed into an Engagement Scorecard.

To use the function of the BSc but transform it into an Engagement Scorecard the intention of the four perspectives stayed the same however they are renamed and formed.

1. Community manager
   a. The stakeholder in this instance is the community manager. He/she would like to know how their goals are doing. So how do we look in the eyes of a community manager?
2. Customer engagement
   a. How do our customers interact with each other? The goal here is to understand the behaviour members have with each other.
3. Platform engagement
   a. To excel in engagement the community should have a good platform. People should feel comfortable joining the platform and should not hesitate to post something.
4. Innovation / development
   a. The innovation and development perspectives are used to keep improving

To get a clear view of the goals set in this research an Engagement Scorecard is created. The BSc is evolved for measuring the engagement in the CoP of Pachamama. The goals per perspective are explained and the metrics (KPI's) are shown.

Perspective	Goals	Metrics
Community manager	More discussion groups More participation	#closed groups #contributions per member
Customer Engagement	More interaction between members More offline activity	#of comments Average filled out profiles #likes Participation ratio % Average number of Event attendance #events Average #comments
Platform Engagement	Enlargement of the platform	#members #new members

		#engaging members
		#new engaging members
		#of posts
		Average #posts
Innovation / development	Rewarding member	#tags
	Linking member with the same interests/ expertise	#filled in interest
		#filled in expertise

The community manager's perspective has some overlap with the customer and platform engagement perspectives. For example, enlargement of members and more event attendance are goals for the community manager as well. They are not notated in that section because they are already accounted for.

Furthermore, the innovation and development perspective will ensure the transition from the Growth phase to the Maturity phase. Rewarding members for their contributions to the Pachamama community by means of giving them tags will ensure members to feel pleased. Which will motivate them to create more content or be more engaging in the community. This also applies to the interests and expertise. There is a search option on the site which will ensure that people can search for each other based on their expertise on interest. This means that people can find each other instead of getting connected by the community manager. These are steps towards a self-sustaining community.

7 MySQL Workbench (Data warehouse)

To review the status of the Pachamama community the database is received. This database is unstructured and very unclear. To organize the database a data warehouse is built as this will result in a structured overview of the necessary data. Below the data warehouse is shown.[4]
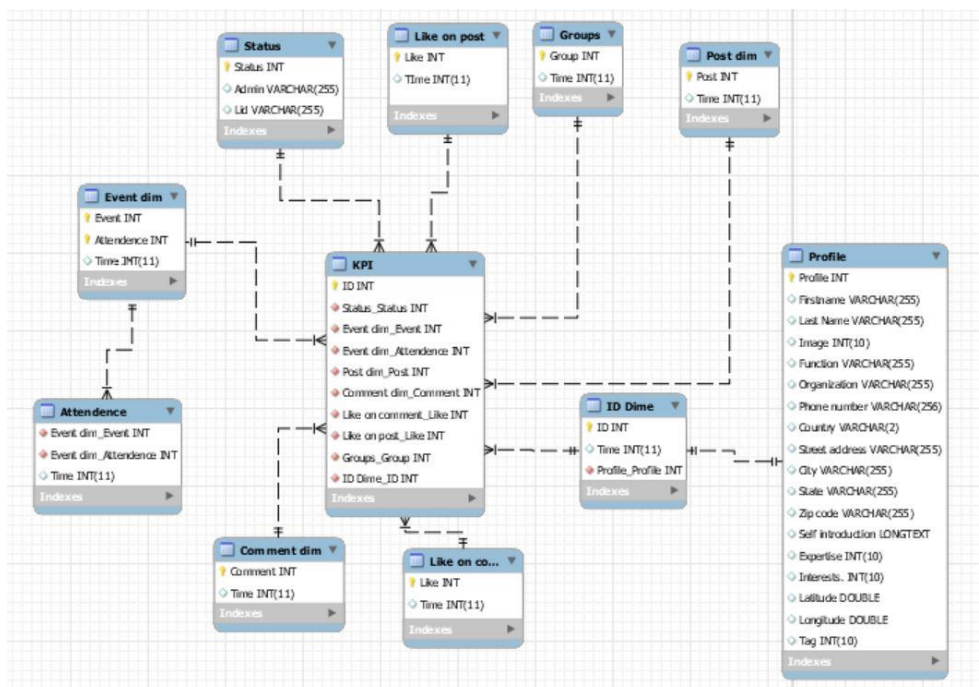

Figure 5: Data warehouse

---

[4] At the point of the study, the database from April 2018 is used.

8 Tableau Prep

Tableau prep is the ETL program used to clean and structure the data in the way determined by the data warehouse. The database of the Pachamama community is connected to the program and all the tables in the database are accessible. Each table has its own content and the right tables need to be connected to subtract the desirable data. In the Appendix, the most important steps are evaluated.

9 Tableau

Tableau is used as a visualization program. This means it transforms the data into graphs. These graphs will indicate the development of the engagement in the Pachamama community. The most important steps are shown in the Appendix.

10 Results

There are a lot of metrics which need to be handled. That is why there are several dashboards created. Each dashboard will be useful to discuss a part of the metrics. The following dashboards will be discussed.

- Members
- Engaging members
- Node
- Enrolment
- Post
- Comment
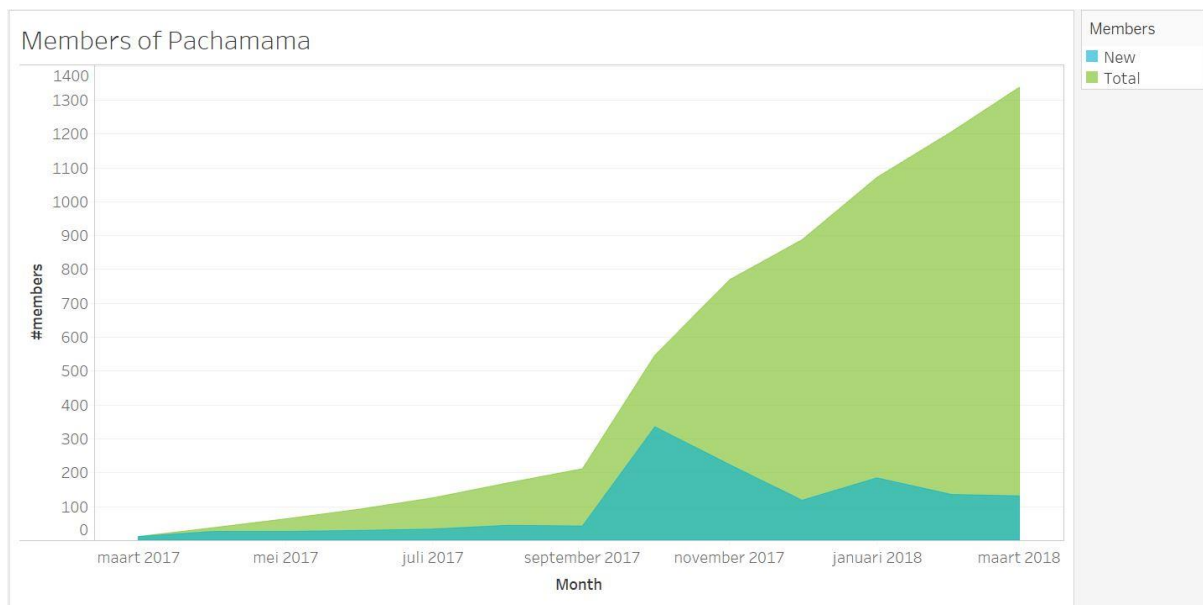- Like
- Groups
- Profile

10.1 Members

Figure 6: Members

As can be seen above is the total number of members at 27-3-2017, is 1,337. The peak of new members at October 2017 can be explained by the fact that since then the old community of

Pachamama (The Pachamama Alliance Facilitator Hub) was closed. So, all the members who were still on the old platform joined the new (Open Social).

Since October 2017 at least 115 members join the Pachamama community per month. The biggest reason why this growth more than doubled compared with before September 2017 is because every lead refers to the new community of Pachamama. Before September 2017 there were two platforms active so not everyone would join the Open Social version.

10.2   Engaging Users

Figure 7: Active VS Engaging User

To be active in a community someone, for example, would need to log in. However, the aim of the research lies in the engagement of their members: people need to contribute to the platform. This can be done in several ways. Someone can leave a like, post or comment. However, they also can enrol in an event, create a topic or event. If a member would take one of these actions they are "engaging" in the month they did the action. In figure 10 the differences between an active user and an engaging user illustrated. The engaging user has the "green" characteristics and the active user has the "green and orange".
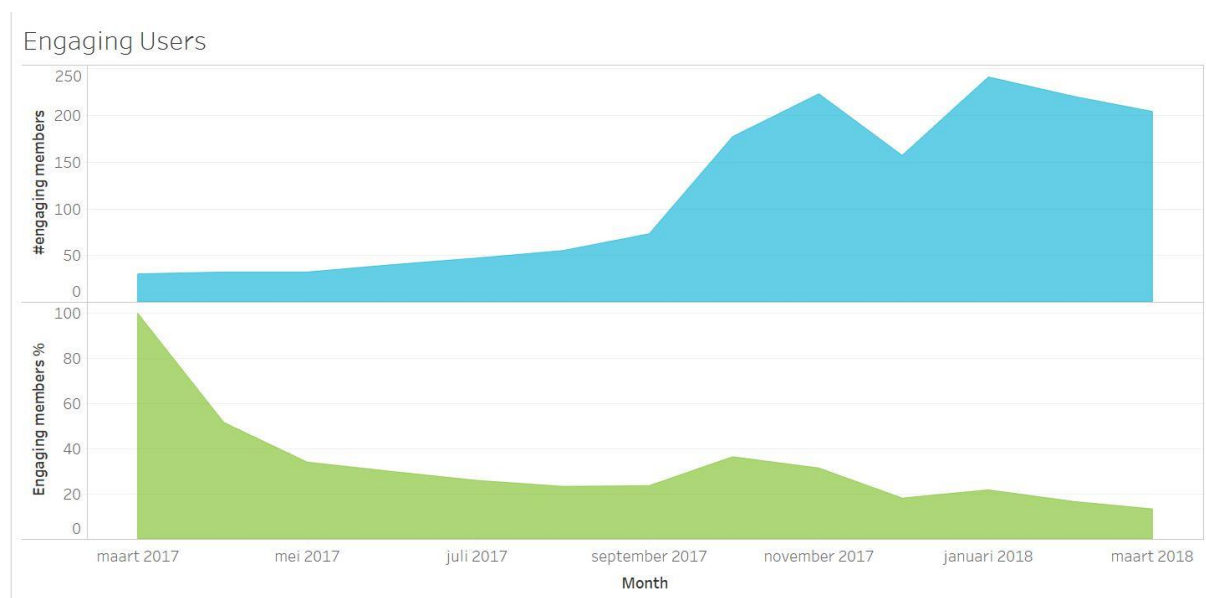

Figure 8: Engaging users

In the graphs above can be seen that the number of engaging members almost tripled when the members from the old community joint the new community. However, the percentage of people being active in the community decreased. This is not a big issue though. There will always be a big group of members lurking. Lurkers are members who do log in and read the information on the site but do not engage online. They might go to a meeting or event but will not press the attend button

online. Also, it is not realistic for the whole community to be active. At the end of March 2018, almost 14% of the community performed an engaging activity.

Of course, the higher the engagement percentage is the better the community would function. However, first new members will need to find a place they are comfortable within the community before they will contribute. The slight peak in October 2017 would contradict this, however, the new members around October were not new members. They were new to the platform but already active in the Pachamama community.
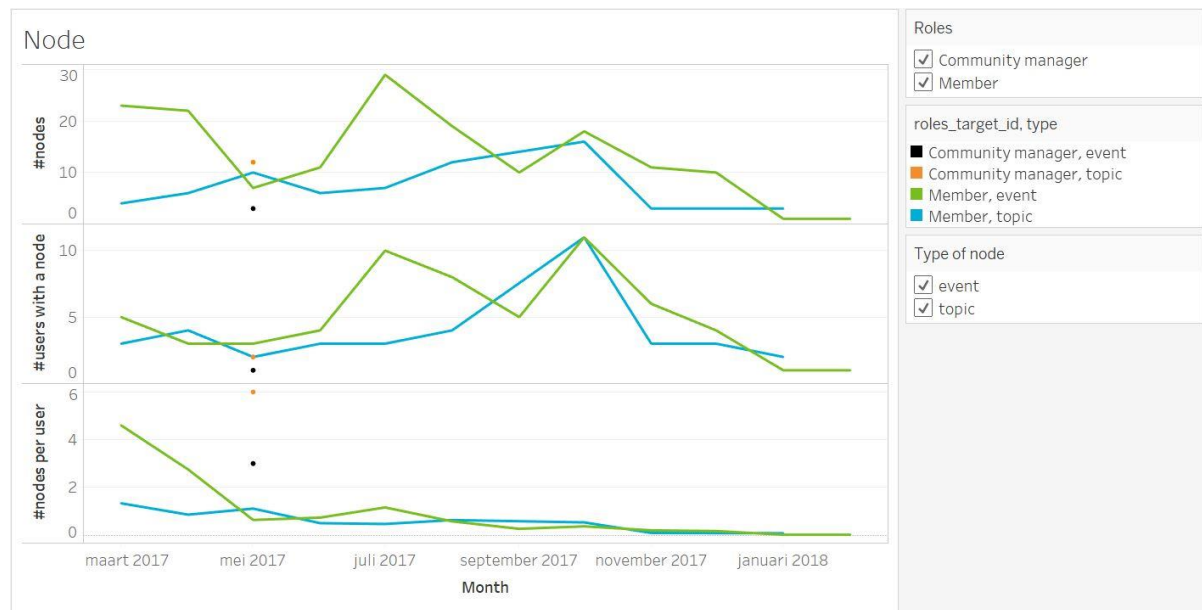
10.3   Node

Figure 9: Node

In Open Social a node is a collective name for six different types of actions. A node could be an article, book, event, page, section or a topic. However, the community of Pachamama does not make use of them all. To get a clear overview on what is happing in the community only the event and topic creations are taken into consideration. The reason why only those are chosen is that Pachamama has the objective to create more attendance during these events. So, knowing how many are created is important. The same holds for the topics. The goal is to create more closed discussion groups. Using these topics, a member tests the importance of their topic. If a lot of people react on it a closed group can be formed.

The first graph shows the number of nodes during the past year. Until September on average there where 15 events created and around 10 topics. After September the number of nodes did not enlarge, which would be expected because a lot of new members entered the community.

As can be seen at the number of nodes per person the community grew a lot. It is hard for a community to keep the same percentage of nodes per person if the community grows. However, if the number of people joining (enrolment) the events grows as well as the goal of more attendance would still be reached.

Another striking fact is the influence of the community managers. Only in May, the community managers made some events (3) and topics (12). Most likely these topics and events were already created in the old community of Pachamama and the community managers also wanted members of the new community to join. However, the fact that only in May 2017 they created a node indicates

that the node creation is self-sustainable. Only the members create them, and they do not have to be initiated by the community managers.

A side note which must be made is the fact that not every event is mentioned in the community. There are a lot of small groups which have a weekly or bi-weekly meeting. If they would create an event every week member near the event other members might notice it and join.
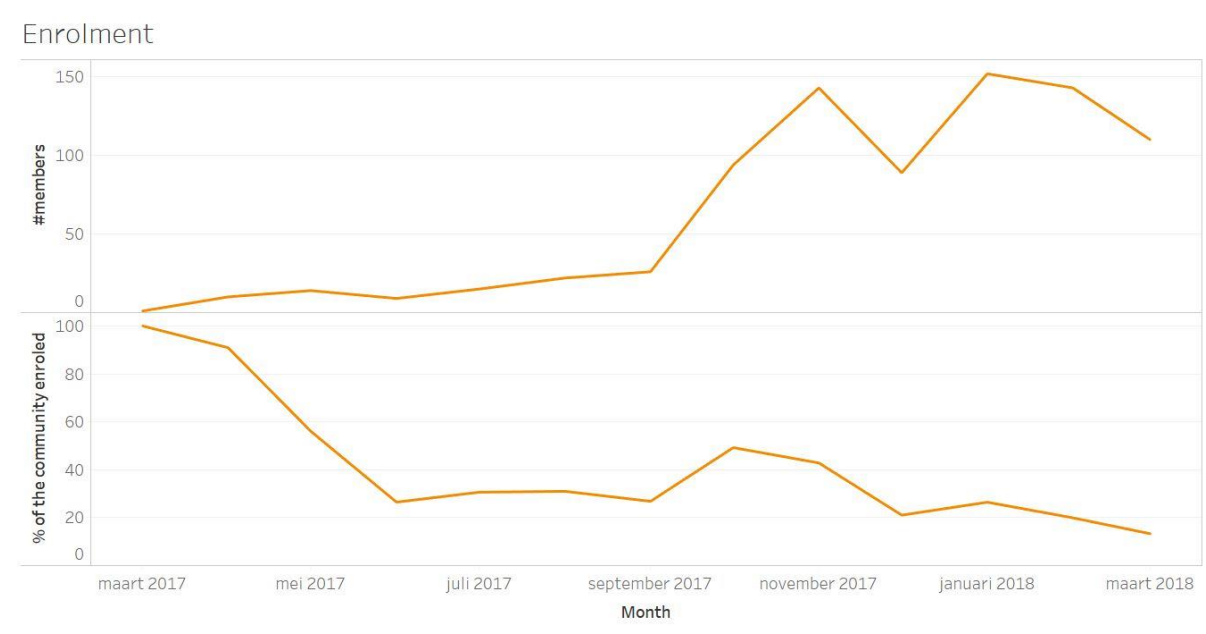
10.4   Enrolment

Figure 10: Enrolment

Before it was concluded that the number of events did not grow at the same pace as the community members which could indicate one of the community death traps (i.e. poor participation). However, the number of people joining the events (at least enrolling) increased significantly. In March the percentage of members enrolling in an event was 13,29%. Which means 110 members enrolled in an event. It is the lowest percentage since the existence of the community, however, the influence on the growth must not be forgotten.
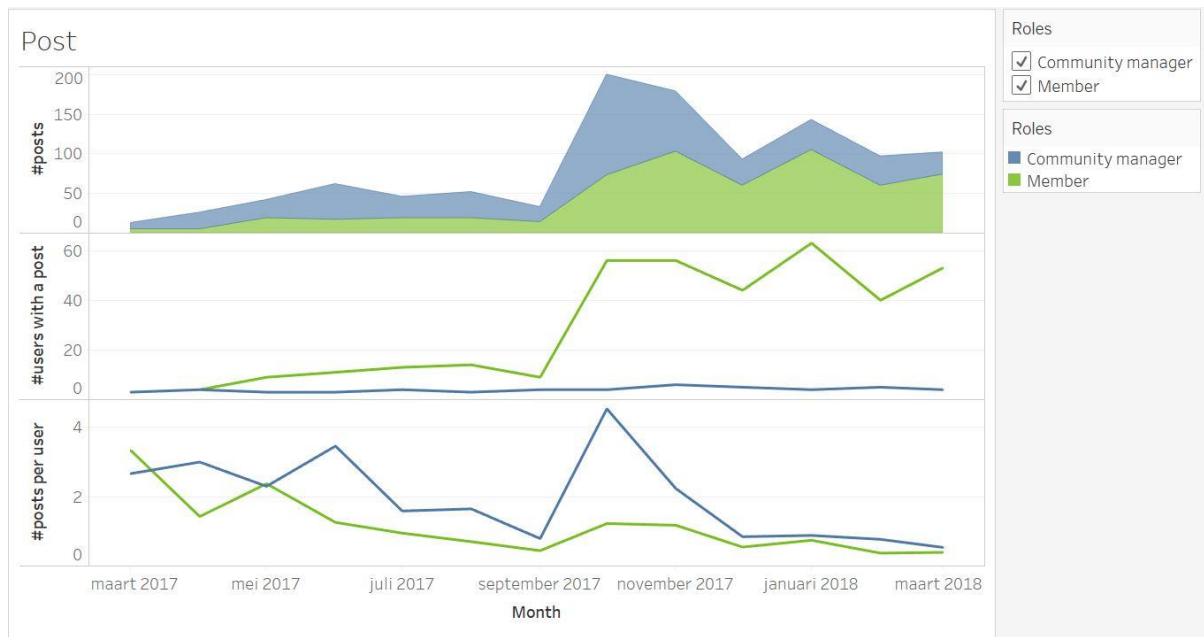
10.5 Post

Figure 11: Posts

Around a hundred posts are posted in the Pachamama community every month. The peak in October 2017 can be explained by the fact that the community managers always react to new members to check what they would like to reach in the community. Around that time the old platform joined the new community. However, the number of posts after this peak stayed high. The influence of the community manager is still there (probably to welcome the new members) but around the 50 different members post every month. With the increase of the members in mind the posts per member are at 0.4, which indicates that a lot of members act on their own.
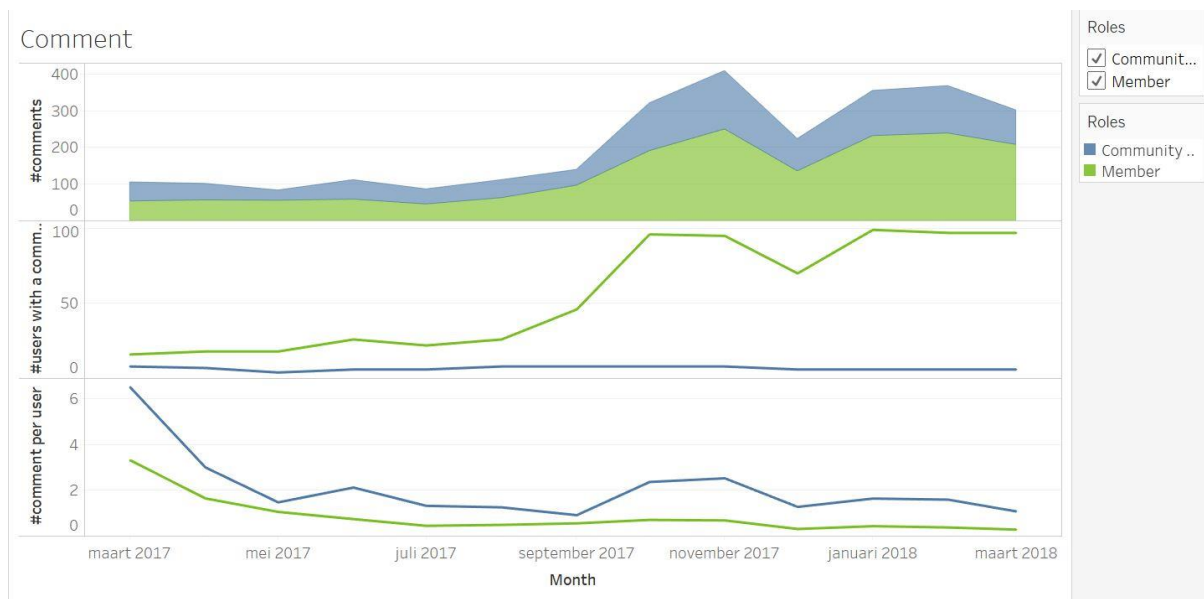
10.6 Comment

Figure 12: Comment

Before the old community joined, the average number of comments per month was around 100 comments. These were placed by about 20 different members and 4 different community members.

After they joined the number of different users increased a lot. Around 100 members are commenting per month. So, when the community tripled the number of users the number of members commenting almost quadrupled. This indicates that there were a lot of enthusiastic members still on the old platform which. The best thing from the number of different users with a comment graph is that this was not a "one-time occurrence". The 6 months after around the 100 members kept on commenting.

The peak in comments from the community managers around October/November 2017 can be explained with the same reason as done by the Post dashboard. The community managers encourage new members to tell them what they would like to reach and when/if the member reacts they will comment on it again.

10.7   Like

Figure 13: Like

The like function is a good indicator of the engagement of users. A lot of (new) members hesitate to post or comment in the community because they are not comfortable (yet). However, hitting the like button is a safe option because you do not have to say anything, but the others do know your opinion.

What applies to the Post and Comment dashboard is the same for the Like dashboard. The increase of likes given by community members around October 2017 is explainable by the approach towards new members.

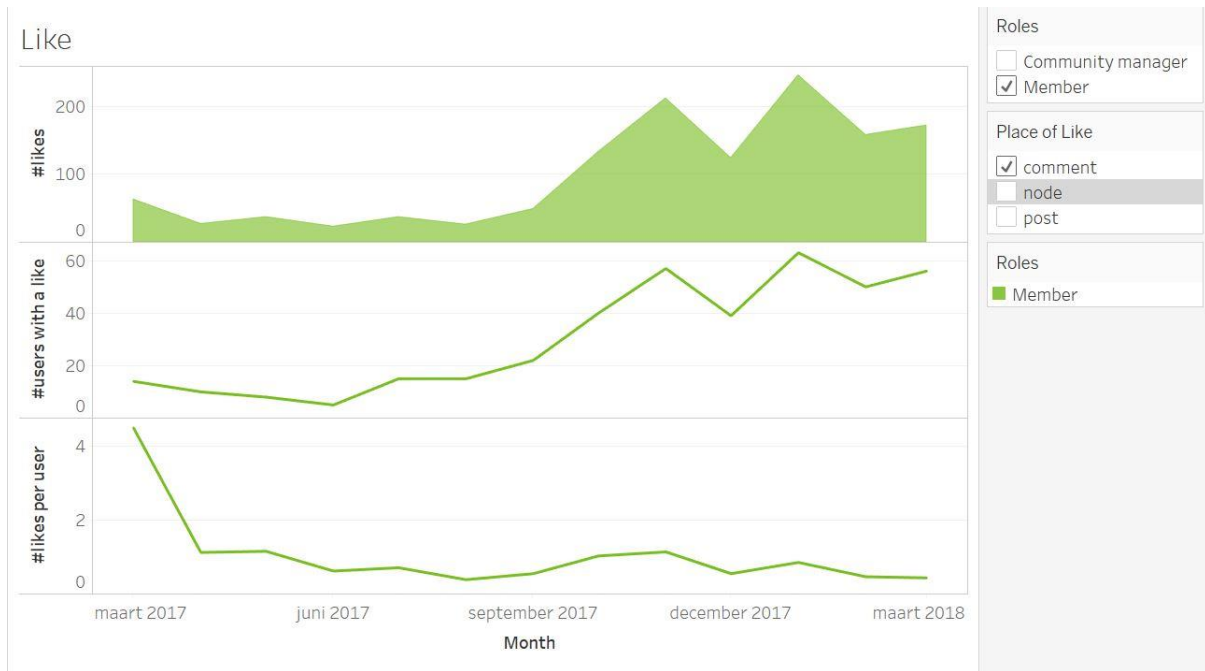*Figure 14: Like on a Comment*

As can be seen, the 150 likes are given by members on comments. The biggest source of likes is the comments, which indicates the fact that a lot of people do read the discussions on the platform.

10.8   Groups

Figure 15: Groups

One of the pillars of the community manager of Pachamama was the creation of closed groups. In this way, people can discuss topics with people they "trust" or know. In the graph above can be seen that again the transition from the old to the new platform had an impact.

The influence of the community manager is notable. The community manager made some closed groups and a few public groups. For the rest, the members did it themselves.
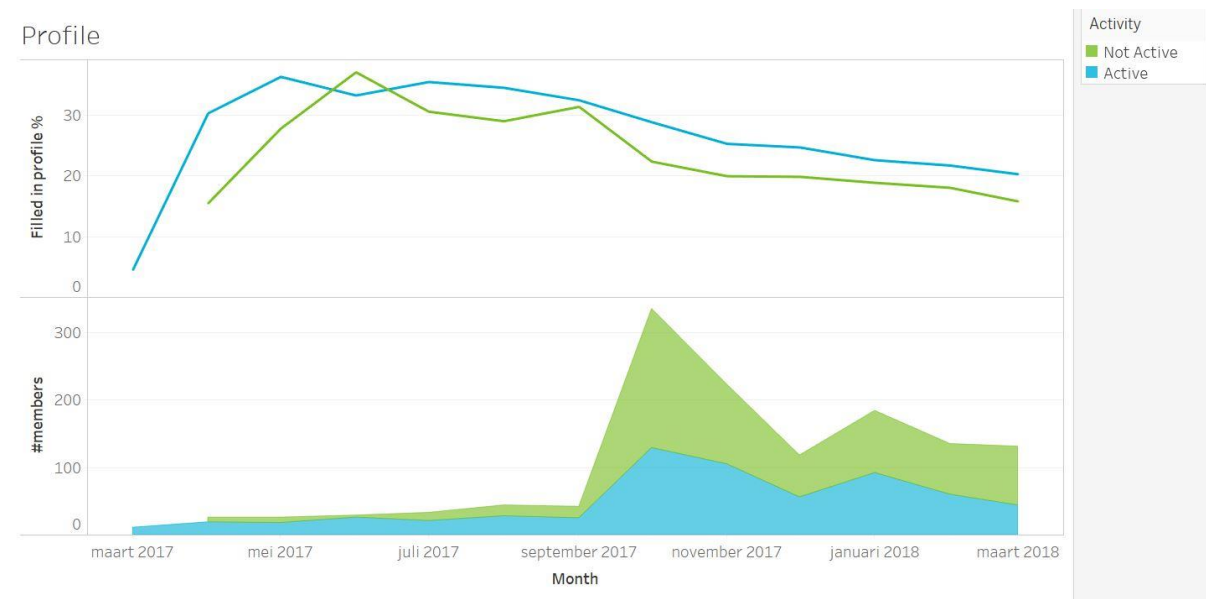
10.9   Profile

Figure 16: Profile

According to the literature, if members enter their personal information they have more credibility. Out of the profile graphs, it can be concluded that the people who are active entered more personal information in their profile. Only in June 2017, the profile was filled in more by the non-active members. However, in June 2017, only three inactive members became a member in comparison with 26 active members.

It would have been better if the change in profile information could have been monitored. A lot of new members will not directly share everything with the others. However, the longer they are in the community they might get more comfortable to do so. This change is more interesting than the initial sign up information. Unfortunately, the database was not designed that way.

10.10   Expertise, interests and Tags

Expertise and interests are fields a member can enter when making a profile. They indicate the subjects a member has knowledge on or wants to know more about. Now some members use the field to be transparent, adding 20 interests or expertise. On the other side do they not enter their first name or country of origin.

Although these are interesting KPI's to research, they are useless if they are used in another way than intended. That is why the graphs created by the expertise or interests do not say anything useful for this research.

The tags are a clever way to reward members for the effort they put into the community. These tags can only be handed out by the community managers. In this way, you can create a status for members which will be approved by the community. However, the same as said above, there are people with online tags and no information about themselves. In this way, the function of the tags does not work the way it is intended. The graphs on who has tags and if they become active when reserving one would be interesting to analyse. However, this was also not possible by database design.

11 Analysis

The analysis of the wellbeing of the engagement in the Pachamama community will be based on the created Engagement Scorecard and the literature review on the phases of a community of practice

The first perspective which should be looked at is that of the **Community Manager**. Creating more discussion groups in the Pachamama community is a goal which is going at a slow pace. Almost every month a new discussion group is formed, and people have their own discussions. To stimulate the building of these closed groups the community manager could act by combining members with the same interest or expertise. There are probably a lot of members who do not know about all the functionalities of Open Social. Before they would create their own they first have to have seen these possibilities. After that, they might do it themselves.

More participation was another goal of the **Community Manager**. While a lot of new members join the community, the average number of contribution per member went down. In contrast to the September 2017 shift, not every new member felt directly comfortable sharing their experience and expertise directly. The biggest reason the average went down is that of the growth. Every new member that enters the community needs to be compensated.

The second perspective is the **Customer Engagement**. The number of comments and likes went up. Even though the community grew a lot there is a shift in the ratio of comments. Where in the beginning the community managers reacted on everything, slowly but noticeably, the community starts to react on each other. The influence of the community manager is less obvious to the community so gets more self-sustainable.

Looking at the **Customer Engagement,** the number of events did not go up. This does not need to be a problem if the number of people attending the events does go up, which happens to be true. A lot of more people are attending events, which proves that the customer engagement is doing well. The transparency in the community could improve. Now a lot of users do not have a filled in profile or even the basics. The option for Pachamama to ask their member to fill in the basics (first name, country and introduction) gives them the opportunity to create more transparency in the community and this would add value to the customer engagement.

The **Platform Engagement** is doing well. A lot of new members are joining the community every month and a lot of them become "engaging" on the platform. At the end of March 2018, 13,3% is active. This is the lowest ever but again it needs to be considered that a lot of new members are joining the community. For every 10 members at least 2 need to become engaging to raise this number. This is probably a step too big for them, however, over time, this might change.

The number of posts and the who created the post is maybe the best indicators of self-sustainability. The community managers are no longer the main source of posts in the community. These are done by members. The function of the most posts of community members is showing people the way. The content is created by the members themselves.

Even though the incentive is good the working on the **Innovation and development** perspective is not going as it should. A lot of people use only the expertise, interest or tags as their description. However, the functionality of these attributes is not used in the right way. The reasoning behind it, so people can find each other on common grounds and the award system (tags), are good steps to be made. There needs to be a change in use will the functionality work in a proper way.

One of the important factors according to the literature on the phases of a community of practice is **Attracting members**. As can be seen from the graphs the number of members is growing. New

members are either or invited by current members because they joined an event or have a relationship with the community or they are interested because they saw the platform online. With this growth, it is important that the new members feel welcome in the community so they, one day, become engaging themselves.

Growth management is important to directly involve new members. Pachamama does that by asking every new member what they would like to achieve in the community, which is a very clever way to get in contact with your members. This reduces the step a new member must make to feel comfortable in the community.

While the growth is going well it is also important to keep track of the interaction between members. **Encouraging interaction**, even though the number of closed groups does not enlarge fast the interaction between members does rise. A good example is to look at the 'likes' given by distinct members on other members. This number indicates that the spread of the community is enlarging which suggests an interaction between members.

An aspect that community of Pachamama should focus on is the **Trust** between their members. Of course, there are a lot of engaging members which did fill out their profile and are trustworthy. However, it is a big step to interact with other people if information about them is not accessible.

12 Conclusions

It can be concluded that the engagement of a CoP can be measured if the right qualitative understanding of the community is established. Therefore, the goals of the community need to be taken into consideration and the outliers in the data need to be comprehended. From the analysis, it follows that since every community of practice differs, so do the objective of that community. However, the main metrics for engagement in a community are the same. Even though these cannot be compared, because of a different approach, the main factors are determined.

To determine what an online community of practice defines a literature study was conducted. The main aspects needed are, first, common ground. People in a CoP should have a common ground as a basis for their conversations. There also needs to be a focus around the community which ensures the development and maintenance of the core knowledge. In the beginning of a community, this could be done by community managers or other stakeholders. However, after evolving this also can be done by engaging members who did so for a long time. This will lead to self-regulation which ensures the well-being of a community. Last but not least, willingness and interaction between members of the community are needed. A good common ground and (for example) 15 community members cannot do the work alone. The members need to be willing to share experiences and expertise. If this is not the case the community will not survive.

There are 4 different phases a community of practice could be in of which the last two phases are the most interesting. The growth and maturity phase give the opportunity to set goals and work on the development of the community. Before these phases, the goals are difficult to measure. However, in these phases, the development can be monitored and each of them has its own goals and understanding but clear objectives.

Finding out what the community manager of a CoP finds important factors for its CoP became more important than first thought. Every community is different so to understand one you should get inside knowledge. However, this was specific to getting to know the community and understanding the data. It did not create insights on what factors of a CoP are important to determine the engagement.

It can also be concluded that the community of Pachamama is doing well on the engagement in their community. Even though the trends, on first sight, are going downwards. With the growth taken into account, which causes a misfit with reality, the graphs show a positive development. It would be very interesting to see the development over time and see if these occurring trends stay the same way and even if the conclusion that they are doing well is right. It could be that despite the enlargement members will be more engaging over time. However, with the acquired knowledge they are doing well.

For Pachamama, the next step should be more self-regulating. This means making the transition to the maturity phase where the main activities are done by the members. While some clues can be found in the data that the transition was already was happening, at the moment, community managers are still needed. If they could make the step to the maturity phase it would benefit the community. One step would be integrating members into the administration, in other words, enlarging the number of community managers based on activity and transparency of the members. A reward model so other members can see the achievements of a member, which also enlarges trust between members (the rewards are given by the community managers) is another. Furthermore, there could be a different kind of membership roles according to your activity on the platform. These would also create more transparency and could even create goals for new members of the CoP.

13 Recommendations

Some recommendations can be made based on this research. First, would be keeping track of the login. When does someone log in? If you would have the data which shows the login distribution some facts could be easier explained. For example, you could have a clearer view of the active members. If someone does log in but does not engage it is still valuable for the community.

Second is the ability to keep track of the changes in someone's profile. According to the literature, the more transparent the community interacts with each other the better it works (Andrews, Preece et al. 2001). Of course, new members will not directly enter all their information but after some time, when they feel comfortable, they might.

At the moment the database is one big cluster of information with a lot of fields the potential value of which is still unknown. It is a fact that when designing the database, you will need some columns which are not useable for visualization. However, determining which data you save, and which won't be saved might help to structure the chaos.

Furthermore, it would be a good step to let people be obliged to fill in some profile information. Again, of course, you should not make the bar too high to enter the community, but some standard information would benefit the trust in the community a lot.

The interests and expertise are, at the moment, in such shape that you can enter anything yourself. Structuring this in such a way that there is a list of expertise or interest, so people can choose from that list would be beneficial. If someone would like to search for another person in the community based on specific interests it would be a pity if they do not find each other because they used different spelling.

On further research, it would be interesting to see if you could predict the development. Now trends can be spotted and acted on. However, maybe it is possible by spotting small changes in the community to predict the outcome after a while. In this way, you would have a safer way of leading your community. By predicting the development, you would have a new tool to sell the product Open

Social because you could demonstrate in which way the community would work. Also, a lot of small problems can easily be adapted.

14 Appendix

14.1 Systematic Literature Review

A systematic review aims to provide a complete, exhaustive summary of current literature relevant to a research question.

- **Step 1**
  Create a structured question to guide the review
- **Step 2**
  Perform a thorough search of the literature for relevant papers
- **Step 3**
  Extraction of relevant data based on the Inclusion/Exclusion criteria
- **Step 4**
  Analyse and combine data, Concept matrix

To perform a literature review on the phases of an online community of practice the following structured question is formulated: Which phases does an online community of practice, based on knowledge sharing, go through?

Search String	Scope	Date of Search	Date Range	Number of entries
Search protocol for Scopus				
"phases" AND "community of practice"	Title, keywords and abstract	1-6-2018	1991-present	205
"phases" AND "online community of practice"	Title, keywords and abstract	1-6-2018	1991-present	11
"life-cycle" AND "an online community of practice"	Title, keywords and abstract	1-6-2018	1991-present	0
Search protocol for ScienceDirect				
"phases" AND "community of practice"	Title, keywords and abstract	1-6-2018	1991-present	2898
"phases" AND "online community of practice"	Title, keywords and abstract	1-6-2018	1991-present	46

"life-cycle" AND "an online community of practice"	Title, keywords and abstract	1-6-2018	1991-present	7
Total				**3160**
Removing Duplicates				-269
Selecting based on exclusion criteria				-2743
Selecting based on Inclusion criteria				-90
Removed after reading				-32
Total selected for review				**26**

Number	Criteria	Reason for Exclusion
1	Pre-1994-articles	The concept of an online community of practice was first coined by Wenger (1998). It is expected that papers before 1994 do not address this topic specifically
2	"Phase" OR "Life-cycle" not mentioned in abstract	A lot of papers address the topic of an online community of practice. However, the focus lies on the phases of an online community of practice. If it is not mentioned in the abstract or title, it assumed not to be treated as a relevant factor.
3	Transactions topics	This does not referrer to communities of practice wanted
4	Game topics	This does not referrer to communities of practice wanted
Number	**Criteria**	**Reason for inclusion**
1	Knowledge sharing	The CoP researched is based on knowledge sharing. There are more types of CoP, but these are not relevant

14.2   Philosophy of science

In research philosophy, there are three areas which need to be discussed: epistemology, ontology and axiology. Within these areas, there are different views on reality and nature. These can be positivism, interpretivism, realism and pragmatism. One of these views fits a research philosophy area the best. These are explained below.

14.2.1 Epistemology

The study of the basis and nature of knowledge, its justification and the rationality of belief. The philosophy addresses questions as "How do we know that we know". The most appropriate view on epistemology is realistic. The more data obtained during the process of making a dashboard the better the dashboard gets. However, because the dashboard is made for a specific target group it is important to not only look at the data but have a more pragmatic view. Subjective knowledge may be useful in this instance.

14.2.2 Ontology

The philosophical study of the nature of being, becoming, existence or reality. A very simple definition of ontology is that it is the examination of what is meant by the word thing. A positive view of ontology fits this case the best. A lot of data is generated which is wholly external. From this data objective conclusions about the reality can be made.

14.2.3 Axiology

The inquiry into human values and their foundation. The positive outlook fits the best. Because the research will be undertaken value-free. The data will be analysed objectively indifferent to outcomes. My personal values or perceptions will not affect the data, the same research could be done by someone else and the same result should be established.

14.3   Ethics

The company where I will be conducting my internal research is "GoalGorilla". GoalGorilla facilitates online social communities. They develop, provide and maintain the online communities. Because GoalGorilla also manages the online communities a lot of data is acquired. This data can be used to analyse the development of the online communities. However, to which extent can a company or organisation use the data for their own gain? Should they be able to use personal information to determine which addresses somebody lives?

Below is the ethical cycle is shown. (Miller 2018)

14.3.1 Phase 1: Formulating the ethical problem

To clearly formulate the ethical problem the problem, the parties and the moral nature must be named.

The ethical problem is which data from which person can be collected and spread. The parties that are involved are customers of the online community, the organisation of the online community (such as website owner), the authorities and GoalGorilla, the maintainer of the online community. The moral nature of the problem is that it is undefined who owns the personal data.

14.3.2 Phase 2: Analysing the problem

This phase breaks down the problem into components. Below the stakeholders and their interests and moral values are notated.

1. Customers whose data is collected
    o   The customer wants to maintain their privacy, they want to be in control of their personal information. Also, they would like an online community which works well, they like to use the functions which are based on user-analytics.
2. Companies/organisations who collect data
    o   A business wants to be an integer, which means that they are trustworthy for customers. Trustworthiness ensures new customers and generates new profit. In the end, a company wants to make a profit. However, not all companies have a

commercial aim. Companies also need to comply with the regulation, such as privacy law.

3. Companies/organisations who use the data
   o Companies/organisations would like to make a good working online community, they would like to target goal orientated. Companies also need to comply with the regulation, such as privacy law.

4. The authorities who have a supervisory role
   o The authorities want to continue ensuring the safety of the society. On the other hand, they encourage the free market which leads to innovation. This innovation depends on which data they can use. The authorities also have a legal obligation to supervise certain activities and ensure companies oblige regulation.

Questions which need to be asked

- Who owns the personal data?
- Did the customer give permission to collect the data?
- Does the customer realize fully what they gave permission to?
- Do the customers realize that they are targeted with a goal?

14.3.3 Phase 3: Laying down some options
In phase three several options for action are discussed. There are four different types of options:

1. Black and white
   - If the option was black or white, then you should be able to either use all the data or nothing

2. More colour-spectral
   - Of course, there are more options than everything or nothing. So, could you only use data which the customer filled in? Or only use the statistics based on the behaviour of the site

3. Cooperative strategies
   - A cooperative strategy would be that customers give permission for everything they dos, so the company can use everything. On the other hand, a company could only use data which is not traceable to the customer.

4. Whistleblowing
   - Whistleblowing would be exposing an organisation which abuses personal information. Organisations/companies abuse personal information if they use data somebody did not give permission for or do not have another legal ground to use the data. Also, customers could leave the online community if they know that their data is used.

14.3.4 Phase 4: Ethical evaluation
In this phase, the moral acceptability of various options will be judged. Different approaches are discussed.

According to utilitarianism, the best option is the option which is useful for the most people. In this case that is a little complicated. At first, you would say, do not use any personal data because it can only be abused. However, this information can evolve in useful insights for example for the development of the online community, e.g. extra features where users will benefit from. Good use of

personal data would be with consent from the owner of the data. Abuse would be using the data without this consent. The big difference between these two is the fact that the company does not have the choice, but the customer decides and is in control of its own data. The term useful is definite as is the wellbeing or happiness of all people. The term happiness is defined in the utilitarianism as the fun minus the pain. When would this be? If the information gathered by a company adds value to the online community and does not harm the individual. It could harm people if personal data is collected against their will. If data is anonymised this cannot occur. After data is anonymised it cannot trace back to who it belonged. If it cannot be traced back, it cannot harm people.

Deontology is the normative ethical position that judges the morality of an action based on rules. Kant formulated deontology in two ways. The first one is if an action can be universalised. If everybody would use others personal data without others consent no one would give out the consent which is a paradox. From this view, the use of personal data without explicit consent is not ethical because it is irrational. As a company, you contradict yourself because if everybody would do what you would do the society would not work anymore.
The second formulation states if the action treats people as means to an end or that it treats people as a goal on their own. This depends on the approach of the company, if the approach is earning money it is not morally right. However, there are a lot of organisations which like to set up an online community where the society benefits from. Knowledge platforms or discussion forums for example.

Virtue ethics discusses the nature and definition of virtues and other related problems. For example, how are virtues acquired? The question here would be "What is a good company?". A virtues company/organisation would be an organisation which thinks about the consequences of their actions. For example, if they use a lot of data they make sure it cannot be traced back to the origin. They make sure the data is secured so it is not accessible to the public.

The common-sense approach is a practical judgment shared by nearly all people. This is a good approach because people can realize which data is risky to use, you could harm somebody if you use it. For example, by sharing information someone does not want to share. However, the common-sense approach is very risky. Only one person needs to be inattentive and the damage is done. Once personal data is shared it cannot be taken back.
The fact is that are no real rules. On European and national level there is privacy regulation, as well as the right to privacy laid down in universal human rights law. Recently the new European Regulation entered into force on the 25th of May 2018 (The General Data Protection Regulation). This Regulation formulates the rules for data collection. Since the regulation is technology neutral and with open norms to fit all data collection it is not very pragmatic. For every data collection and use, there needs to be a legal ground. The regulation can be interpreted in multiple ways. For example, data subjects should be given notice when their data is being collected. However, will this be done by a pop up on your screen, does the customer recognize when an organisation makes use of this? To use this framework the rules should be stated more clearly. Luckily, but there are also guidelines, so information is available for companies to work with. As well as information on the website of the supervisory authority with best practices. However, companies need to spend the time to inform the data subjects and to ensure regulation is complied with.

14.3.5 Phase 5: Reflection
In this phase, the findings are brought into a wide reflective equilibrium. Companies/organisations do not want to harm customers because this leads to distrust, which will lead to fewer customers. The fewer customers the slower the desired result is achieved. If the aim is earning money or creating a knowledge platform that does not matter. They want to create the best service they can. Customers do not want to be harmed however they want the best service as well. The best way to guarantee this is probably to argue which data is risky and which is not. For this risky data, clear rules need to be

determined. The best way would be if all the parties work together. This would be the company's customers and the authorities. With the aim of creating the best service possible without harming the customers. A combination of deontology with common sense would create a good fundament.

Concluding, the hard part about ethics is that it does not bind legally. I think that almost everybody would agree with me that information about someone's financial status is confidential and should not be spread. For instance, if you are doing well and everybody knows it people will judge you on your expenses. Also, the risk of being robbed is higher. Another good example is someone's medical status. If you are sick or have an embarrassing condition that most people would like to keep that for themselves. Spreading the information can cause a lot of awkward conversations. However, if someone should do so they cannot be charged because again; it does not bind legally. On the other hand, the law does. Legalisation is also based on ethical considerations. Because the rules do not connect to the current digital world, the European Parliament provided new "General Data Protection Regulation"-rules. The regulation is based on the following principals. (Commission 2017)

Data can only be processed on the following grounds:

- personal data must be processed in a **lawful and transparent manner**, ensuring fairness towards the individuals whose personal data you're processing ('lawfulness, fairness and transparency').
- you must have **specific purposes** for processing the data and you must indicate those purposes to individuals when collecting their personal data. You can't simply collect personal data for undefined purposes ('purpose limitation').
- you must collect and process **only the personal data that is necessary to fulfil that purpose** ('data minimization').
- you must ensure the personal data is accurate and up-to-date, having regard to the purposes for which it's processed, and correct it if not ('accuracy').
- you can't further use the personal data for other purposes that aren't **compatible** with the original purpose of collection.
- you must ensure that personal data is **stored for no longer than necessary** for the purposes for which it was collected ('storage limitation').
- you must install appropriate **technical and organizational safeguards** that ensure the security of the personal data, including protection against unauthorized or unlawful processing and against accidental loss, destruction or damage, using appropriate technology ('integrity and confidentiality').

As can be seen in these principals is that there needs to be an understanding between customer and company. Also when the customer gives permission to use their personal data it needs to be treated with respect to that certain goal. It cannot be reachable for third parties and when the purpose of the data is fulfilled it needs to be destroyed. All these principals keep in mind that the data from a person will remain with the same person. It only can be used for a greater goal but can never be traced back to them.

14.4   Code of Conduct

"Codes of conduct are codes in which organizations lay down guidelines for the responsible behaviour of their members." (Poel & Royakkers 2011, 33). The code of conduct at GoalGorilla is based on corporate codes. A code of conduct that is formulated by the company. The code can be described as an aspirational code. It expresses the moral values of the company.

At GoalGorilla there is not (yet) a code of conduct. The company functions according to the following statement.

"At GoalGorilla there is an informal, relaxed working atmosphere. People treat each other respectfully. Disoblige colleagues and/or clients, such as approaching aggressively, getting angry, shouting, cursing. Is absolutely not tolerated."

14.5   Professional responsibility

Being an engineer ensures that both active and passive responsibility play an important role. An active role means taking responsibility before something has happened. It refers to a duty or task to care for certain state-of-affairs. For me as an engineer at GoalGorilla, it is important I take care of the personal information customers give us. Before using this data to optimize our product I need to consider that the data is sanitized and not traceable to the customer.

Passive responsibility means taking responsibility after something happened. Which means if some confidential information got leaked taking responsibility for the action. This will be done by investigating where it went wrong and taking action that it will not happen anymore. Also, it always needs to be mentioned to the authorities. Depending on the size of the leak there needs to be decided if the person needs to be warned. This could be done because it is the rule or because of the moral values of the company.

15 MySQL WorkBench

As can be seen in the data warehouse, the time notation in the delivered time is an int (11). Which in this case means that it a Unix notation. This is a system for describing a point in time. It starts on Thursday; 1 January 1970 and every day is treated as if it contains exactly 86400 seconds. This is an easier way to notate your time. The only downside is that when you want to show it needs to be converted to real time.

Furthermore, the main aspects of the metrics which are needed are divided in such an organized way that when these tables are created they are easy to link. When started with the ETL program it is clear which data is needed and which data is unnecessary.

16 Tableau Prep

The data is delivered in a database. To structure the connections between the right tables, need to be made. Which results in the following:

Figure 17: Tableau Prep design

The important steps made in this process are explained below.

16.1   Join

Every table has its own information. One will contain the first name of a person and another will have the last name. When combining these tables, the first and second name are combined through a join function. This function asks for a join clause. This means it compares the data based on a column of

both tables. The data is combined based on the corresponding column and in that way, the first and second name is linked.

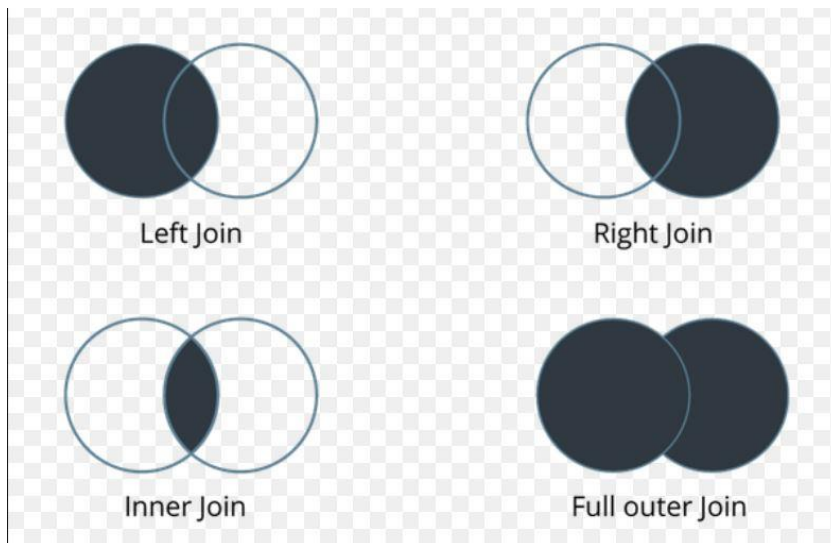There are four ways a join function can be used. The ways are shown below.

Left Join

Right Join

Inner Join

Full outer Join

Figure 18: Join types.

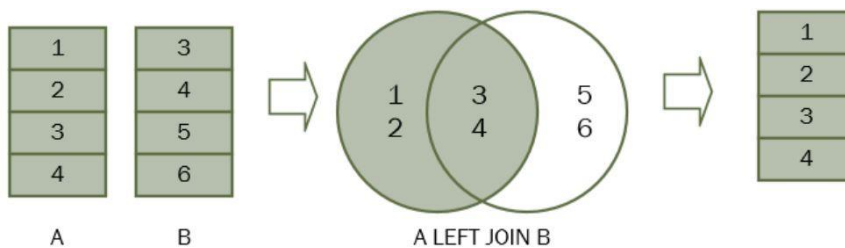A left join will keep all the rows it has and adds the corresponding ones from the second table.

Figure 19: left Join

The right join works the same but the other way around. The inner join only combines the data that is in both tables. So, in the example above the inner join would return 3 and 4. The fuller outer join combines all the data from both tables. So, in this instance, it would return 1, 2, 3, 4, 5, 6.

In the tableau prep overview can be seen that the left/right join is the most common function used. This is because, in this instance, the data is valuable if it is not possible to compare it with all the users. For example, when the number of posts is needed, this can easily be counted. However, it does not say anything about the users if it is not compared with the number of users at that moment. Let's say there are 10 posts. That would be a good score when there are only 2 users, but when there are 100 users only 1/10 of the people post (if the posts are counted distinct). So by means of using the left/right join all the users are accounted for and only the ones who posted get extra information.

16.2   Union
The union function is a simple addition. If two tables union, the data from one table is added to the other.
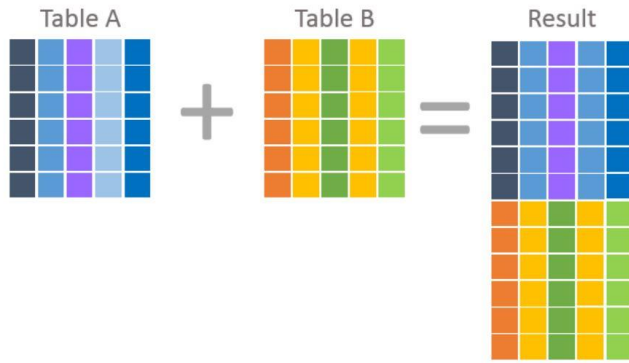
Figure 20: Union function

This is useful if, for example, the number of users with an action is needed. So, if a user gave a like, it gets added to the table. If a user posts something it gets added to the table. In that way, a list of all the users with an action is collected.

16.3 Cleaning

As can be seen in the Tableau prep overview a lot of stages are called "Cleaning". If you take a closer look it is done after every Union, Join or New Table. This is done to subtract the unnecessary data. Most of the time this is data which is used to easily navigate through the database from a designer view. However, these are not needed for the visualization. After the Union and Join most of the times the extra created table is removed.

17 Tableau Formulas

To calculate the comments/like/posts/nodes per person the following formula is used.

COUNT ([Concerned action]))/ RUNNING_SUM (COUNTD ([All users])

At first, the relevant actions are counted. After this amount is determined it is divided by the number of distinct users at that moment. This will show the number of actions per member.

This creates perspective on the development in the community. The number of posts, for example, could increase, so it would be easy to say that the community is doing well. However, it depends on the number of users. If the number of users grew a lot and the number of posts did not this would indicate that the community is not doing well.

The Unix time values gained from the database needs to be converted to a normal date notation:

DATEADD ('hour',1, (Date ("1/1/1970") + ([Timestamp]/86400)))

This done by dividing the time stamp by the number of seconds in a day (86400). In this way, the number of days can be added to the starting date 1/1/1970. In this way, the action can be determined on the second.

To find out which percentage of the profile is filled in the following formula is used:

(COUNTD([Data]))/RUNNING_SUM(COUNTD([User])

The number of entries per profile field (first name, last name, geolocation, address, state, city, zip code, country, phone number, image, function, introduction) are summed up. So, a total percentage per fields are calculated. After this is done they are summed up and divided by the number of entries to get an average of filled in profile.

Furthermore, the range of the graphs created is from 1-3-2017 until 27-3-2018. The date is because the community started at 1-3-2017. The community was active before that but only to test the possibilities. On 1-3-2017 people could join the community. There is also data available after 27-3-2018, however, the community got attacked by a DDoS-attack on 28-3-2018. Which means a lot of fake accounts where created. The reason why the data after 28-3-2018 is not included is that a lot of metrics are calculated depending on the number of members in the community. Data after 27-3-2018 is for that reason corrupted.

The graphs made in Tableau are based on the colours of the logo of Pachamama.

18 Literature list

(!!! INVALID CITATION !!! (Brown 1991, Preece 2001, Soroka and Rafaeli 2006, Borzillo 2007)).

(!!! INVALID CITATION !!! (Butler 2001, Li, Kankanhalli et al. 2016)).

(!!! INVALID CITATION !!! (Malhotra, Gosain et al. 1997, Preece 2000, Kling and Courtright 2003)).

Amrit, D. C. (2014). "Balanced Scorecard (PowerPoint)."

Andrews, D., et al. (2001). "A conceptual framework for demographic groups resistant to online community interaction." Proceedings of the Hawaii International Conference on System Sciences: 163.

Armstrong, A. and J. Hagel Iii (2000). Chapter 6 - The Real Value of Online Communities*. Knowledge and Communities. Boston, Butterworth-Heinemann: 85-95.

Beenen, G., et al. (2004). Using social psychology to motivate contributions to online communities. Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW.

Blanchard, A. L. and M. L. Markus (2004). "The experienced "sense" of a virtual community: Characteristics and processes." Data Base for Advances in Information Systems 35(1): 65-78.

Borzillo, S. (2007). Communities of Practice to Actively Manage Best Practices. Wiesbaden, Universität Genf.

Brazelton, J. and G. A. Gorry (2003). "Creating a knowledge-sharing community: If you build it, will they come?" Communications of the ACM 46(2): 23-25.

Butler, B. S., Sproul, L., Kiesler, S. and Kraut, R. E. (2005). " Community effort in online groups: Who does the work and why?" Leadership at a Distance.

Chan, C. M. L., et al. (2004). Recognition and participation in a virtual community. Proceedings of the Hawaii International Conference on System Sciences.

Clancy, W. (1997). Situated Cognition: On Human Knowledge and Computer Representations, Cambrdige UJniversity Press.

Commission, E. (2017). "What data can we process and under which conditions?".

Constant, D., et al. (1994). "What's mine is ours, or is it? A study of attitudes about information sharing." Information Systems Research 5(4): 400-421.

Cook, S. N., and Brown, J. S. (1996). "Bridging epistemologies: the generaticve dance between organizational knowledge and organizational knowing." Working paper, Xerox PARC, Palo Alto, CA.

Cothrel, J. and R. L. Williams (1999). "On-line communities: helping them form and grow." Journal of Knowledge Management **3**(1): 54-60.

Donath, J. (1999). "Identity and deception in the virtual community." Communities in Cyberspace.

Edelman, G. (1993). Bright air, Brilliant Fire: on the matter of the mind., New York: Basis Books.

Ginsburg, M. and S. Weisband (2004). A framework for virtual community business success: The case of the internet chess club. Proceedings of the Hawaii International Conference on System Sciences.

Hall, H. and D. Graham (2004). "Creation and recreation: Motivating collaboration to generate knowledge capital in online communities." International Journal of Information Management **24**(3): 235-246.

Hars, A. and S. Ou (2002). "Working for free? Motivations for participating in open-source projects." International Journal of Electronic Commerce **6**(3): 25-39.

Hummel, J. and U. Lechner (2002). Social profiles of virtual communities. Proceedings of the Annual Hawaii International Conference on System Sciences.

Iriberri, A. (2005). "Building online community: An action research project." Proceedings of the Eleventh Americas Conference on Information Systems.

Iriberri, A. and G. Leroy (2009). "A life-cycle perspective on online community success." ACM Computing Surveys **41**(2).

J., K. A. (2000). Community building on the Web. Berkeley, CA, Peachpit Press.

Jarvenpaa, S. L., et al. (1998). "Is anybody out there? Antecedents of trust in global virtual teams." Journal of Management Information Systems **14**(4): 29-64.

Kapoor, N., et al. (2005). How peer photos influence member participation in online communities. Conference on Human Factors in Computing Systems - Proceedings.

Kollock, P. (1996). "Design principles for online communities." Harvard Conference on the internet and Society.

Lave, J. W., Etienne (1991). Situated Learning: Legitimate Peripheral Participation, Cambridge: Cambridge University Press.

Leimeister, J. M., et al. (2004). Success factors of virtual communities from the perspective of members and operators: An empirical study. Proceedings of the Hawaii International Conference on System Sciences.

Leimeister, J. M. a. K., H. (2003). "Enginering virtual communities in health care." Electronic Journal of Organizational Virtualness.

Leimeister, J. M. a. K., H. (2004). "Revisiting the virtual community business model." Proceedings of the Tenth Americas Conference on Information Systems,.

Leimester, J. M., Ebner, W. and Krcmar, H (2005). "Design, implementation, and evaluation of trustsupporting

components in virtual communities for patients." Journal of Management Information Systems.

Limited, G. (2017). "2GC Balanced Scorecard Usage Survey." 2GC Active Management.

Ludford, P. J., et al. (2004). Think different: Increasing online community participation using uniqueness and group dissimilarity. Conference on Human Factors in Computing Systems - Proceedings.

Malhotra, A., et al. (1997). Evolution of a virtual community: Understanding design issues through a longitudinal study. Proceedings of the 18th International Conference on Information Systems, ICIS 1997.

Maloney-Krichmar, D. and J. Preece (2005). "A multilevel analysis of sociability, usability, and community dynamics in an online health community." ACM Transactions on Computer-Human Interaction **12**(2): 201-232.

Miller, D. P. L. F. (2018). "Reasoning of the ethics (PowerPoint."

Nonnecke, B. and J. Preece (2000). Lurker demographics: Counting the silent. Conference on Human Factors in Computing Systems - Proceedings.

Preece, J. (2000). Online Communities: Designing Usability, Supporting Sociability. Wiley, New York.

Preece, J. (2001). "Sociability and usability in online communities: Determining and measuring success." Behaviour and Information Technology **20**(5): 347-356.

Sangwan, S. (2005). Virtual community success: A uses and gratifications perspective. Proceedings of the Annual Hawaii International Conference on System Sciences.

Tedjamulia, S. J. J., et al. (2005). <u>Motivating content contributions to online communities: Toward a more comprehensive theory</u>. Proceedings of the Annual Hawaii International Conference on System Sciences.

Wenger, E. M., Richard; Snyder, William M. (2002). <u>Cultivating Communities of Practice</u> Harvard Business Press.

Wennger, E. (1998). "Communities of Practice: Learning, Meaning, and Identity." <u>Cambridge: Cambridge University Press.</u>

Williams, R. L. a. C., J. (2000). "Four smart ways to run online communities." <u>Sloan Management Review</u>.

Zhang, Y. a. H., S.R. (2003). "Factors that influence online relationship development in a knowledge sharing community. ." <u>Proceedings of the Ninth Americas Conference on Information Systems</u>.