Virtual Embeddedness of Platform Companies on Social Media

Student Name: Tran Vu Thanh Huyen

Student Number: 480869

Supervisor: Dr. Ju-Sung (Jay) Lee

Master's Thesis

June 2019

ABSTRACT

Embeddedness refers to the fact that social relationships under an industrial context affect organizational activities and outcomes. The emergence of the Internet and the advance of communication technologies have promoted the concept of virtual embeddedness, which combines the ideas of structural embeddedness and virtuality. In the era of social media, the connection between organizations and the formation of virtual communities in social media platforms, such as Facebook, have been facilitated by social media affordances. More specifically, companies can be embedded in their online networks through different kinds of affiliation linkages, for example, liking one another's pages on Facebook. The advances of the Internet and technologies are not only changing the way in which actors in the industry interact, but also enable the emergence of a new market, the (digital) platform market

The thesis aims to examine the pressure of embeddedness and isomorphism to activities of organizations on social media by analyzing the Facebook liking networks and topic modeling of platform companies in their population. A qualitative mixed-methods approach that includes some digital methods – network analysis and topic modeling – is applied. Overall, the finding reveals that the Facebook liking networks of the focal companies are sparse and fragmented, indicating the organizations in the networks are less connected. However, in the company-company network via topics the companies are linked with one another, indicating that they are embedded by the overlapping topics. The expectation that the companies become similar to each other via online content (i.e. topics) due to the pressure of embeddedness and isomorphism happens. In addition, the findings reveal that the companies in the network also manage to resist the isomorphism by topic misalignments, which make them distinctive in relation to others in their networks.

KEYWORDS: Social Media, Virtual Embeddedness, Population Ecology, Platform Market, Network Analysis

Table of Contents

1. Introduction

1.1 The impact of social media on organizational performance and the industry

Identifying industry boundaries and industry competitiveness is central to organizational study. An organization is a social unit comprising individuals that endeavor to achieve collective goals (Bílková et al., 2018). In the study of organizational performance, the resource-based perspective is a fundamental approach as resources and the transformation of resources within an organization determine their success (Bílková et al., 2018). When it comes to the features of organizational change, Bettis and Hitt (1995) describe that a new technological landscape driven by "technology is rapidly altering the nature of competition and strategy in the late twentieth century" (p.7). More specifically, the Internet and social media play a pivotal role in the new competitive playing field. Since the 1990s, the Internet has emerged as "a revolutionary communications technology" (Bettis & Hitt, 1995). By applying communication technologies to enhance their performances and spread their information on social media, companies have improved their operational effectiveness and approached wider markets (Porter, 2001).

In the Internet era, social media are dramatically changing the way we communicate, coordinate and consume information (Goktepe, 2018). Recently, the use of social media in organizations has seen tremendous growth. Corporations now develop and maintain relationships with customers or stakeholders through social media platforms such as Facebook, Twitter or LinkedIn (Parveen, Jaafar, & Ainin, 2013). Thus, these platforms' affordances offer avenues of communication for organizations to interact with the public that constitute the modern digital relation. In the supply chain, modern social networks have been developing on the idea of disintermediation as the Internet and modern telecommunications allow individuals to easily and directly connect to one another. Companies and customers can "move quickly and be more adaptive as they do not need to go through layers of a hierarchical chain" (Ressler, 2006, p. 2).

The advances of the Internet and technologies are not only changing the way in which actors in the industry interact, but also enable the emergence of a new market, the (digital) platform market. In general, this is a market where clients can purchase platform-based services. A distinct characteristic of digital platforms is that they are built on the aggregation and sharing massive streams of data (Strategic Policy Forum on Digital Entrepreneurship, 2016). There are

multiple types of platform-based companies. They can be marketplaces like Airbnb or Amazon, which offers digital platforms to connect to two or more distinct group of customers. In addition, there are companies providing platforms that make digital tools available online, such as customer data platform (CDP) companies that offer platforms as a "type of packaged software", which "creates a unified customer database that is accessible to other systems" (Customer Data Platform Institute, 2019, p. 2). With the customer database, marketers are able to develop marketing strategies and evaluate their effectiveness, as well as predict customer behavior (Earley, 2018).

As for the B2B platform market, it is prosperous and predicted to "represent $10 trillion in socio-economic value creation from 2016 to 2025" (World Economic Forum, 2017, p. 5). The emergence of this new industry market opens room for academic research. In the marketplace, a business operates through relations with other players, such as suppliers, consumers, and competitors. The inter-organizational networks provide channels for exchange useful information and needed resources that a firm can exploit to gain the competitive advantage (Eschol & Tsai, 2005). The thesis aims to examine the inter-organizational linkages of the population of B2B platform companies to understand their performance.

1.2 Embeddedness of organizations in population ecology

The population ecology of organizations pays attention to the population level for studying the interaction and competition between organizations in the market (Hannan, 2005). More specifically, the organizational ecology theory describes growth and decline processes within organizational populations (Hannan, 2005). "No business is an island" (Håkansson & Snehota, 1989), as companies are involved in constant interactions and relationships with other actors in their networks. Thus, the network approach emphasizes on analyzing companies in the network they are embedded to understand their performances (Ratajczak-Mrozek, 2017). In his article, Uzzi (1996) highlights the influence of social relations on organization actions by advancing the concept of embeddedness. According to him, embeddedness and the structure of the network – in which market entities interact – influence their behaviors (Uzzi, 1996).

Embeddedness refers to the fact that social relationships under an industrial context affect organizations' economic actions and outcomes (Uzzi, 1996; Montanari & Mizzau, 2007). Companies that operate in embedded ties are likely to be influenced by its network partners. On

the one hand, embedded ties increase the survival chances of firms operating in networks. On the other hand, over-embeddedness can hinder a firm ability to reach new information or opportunities that exist outside the network (Uzzi, 1996). The concept of embeddedness is important in studying organizational performance. However, this approach to embeddedness is outdated in dealing with the technological conditions in today's economy, especially in the platform market. Thus, the concept of virtual embeddedness, which was introduced by Fowler, Morse, and Lawrence (2004) is more suitable when it comes to the investigation of the inter-actor ties in the platform market. Virtual embeddedness refers to the establishment of inter-organizational connections by electronic technologies (Fowler, Morse, & Lawrence, 2004). In addition, it "represents the image of traditional physical/social embeddedness on the Internet" (Vasileiadou & Missler-Behr, 2011, p. 192). Virtual embeddedness is realized through virtual groups, communities, and markets (Vasileiadou & Missler-Behr, 2011).

The embeddedness of organizations affects their behaviors and market relations (Gulati, Dania, & Lihua, 2002). In the era of social media, the connection between users (e.g. individuals or organizations) and the formation of virtual communities in such platforms, such as Facebook, have been facilitated by social media affordances. More specifically, companies can be embedded in their online networks through different kinds of affiliation linkages, for example, liking one another's pages on Facebook. Online affiliation can be considered as resources for competitive advantage and affects organizational performance (Palacios-Marqués, Merigó, & Soto-Acosta, 2015). Thus, the purpose of the thesis is to examine the way in which social media reflects the virtual embeddedness of platform organizations. In addition, the influence of embeddedness on activities of organizations on social media is explored in this thesis. The research question and sub-questions that the thesis endeavors to answer are:

Research question:

How do social media reflect virtual embeddedness of platform companies in their population ecology?

Sub-question 1*: In what manner are platform companies embedded on social media?*

Sub-question 2: *How does virtual embeddedness influence activities of platform companies on social media?*

In order to answer these research questions, a qualitative mixed-methods approach that includes some digital research methods – network analysis and topic modeling – is applied.

Network analysis allows the researcher to discover how organizations are embedded in their market through various network visualizations and measurements. In addition, posting information and updates is the most popular and regular activity of corporations on social media. Topic modeling is employed to reveal how companies distinguish themselves in their market through the topics they post, and the impact of embeddedness on organizational activities on social media.

1.3 Academic and societal relevance

For the academic side, the thesis contributes to the studies of population ecology and the interaction between organizations and stakeholders in the era of the Internet and social media. Through social media and social networking sites, digital technologies are changing the structure of social relationships in both the consumer and the organization (Susarla, Oh, & Tan, 2012). Due to the rising adoption of social media, research conducted in this field has grown among scholars. However, academic research on the actual influence of social media to organizational performance is overlooked (Schultz, Schwepker, & Good, 2012). Moreover, few studies have investigated the network structure of online activities that have occurred on social networking sites (Penni, 2017). This thesis fills these research gaps by examining the inter-organization networks of companies on social media and the influence of these networks to corporations' online activities.

In addition, this study contributes to our understanding of embeddedness in contemporary society by using the virtual embeddedness framework. Although the concept of virtual embeddedness was introduced in 2004 by Fowler, Morse, and Lawrence, the research in this field is underestimated. Furthermore, researchers have focused on investigating one organization as a case study (Fowler, Morse, & Lawrence, 2004; Morse, Fowler, & Lawrence, 2007; Vasileiadou & Missler-Behr, 2011), rather than examining the inter-organizational ties in the population of organizations. Additionally, while recent studies have paid attention to individuals and consumers when examining the following or liking behavior on a business page, this study considers firms and the liking network created between them to understand their business behaviors and activities on social media. In addition, the thesis paves the way for redefining economic relationships. While traditional economic relationships are built on face-to-face communications and interactions (Morse, Fowler, & Lawrence, 2007; Uzzi, 1996, 1997), modern

businesses relationships and transactions have taken place in Internet-based platforms using algorithms and are facilitated by digital advices.

For the societal perspective, new industries – the platform market – is the target of this thesis. Platforms have become important elements in our lives. While the negative effects of platforms on information security and consumer privacy have received considerable attention from the public and policy makers, their positive values to the economy and society seem to be unnoticed (Dutch Transformation Forum, 2018). Research in this field can not only benefit recent works in the platform economy but also explore new features and trends of contemporary organizational life. Thus, the findings of this thesis are valuable to both scholars and practitioners by providing insights into the communication strategies that platform firms can implement to benefit their business, especially on social media. Furthermore, the concept of virtual embeddedness is also significant to managers of a variety of companies. As the virtually embedded ties are built on the argument that "communication technologies have established a new form of inter-actor connections", which provide considerable benefits and reduce the cost of managing inter-firm relationships (Fowler, Morse, and Lawrence, 2004, p. 663).

As for economists and policy makers, they could apply the virtual embeddedness perspective when examining changes and growth in industrial fields, since they have actual evidence on the emergence of inter-firm relations in the virtual world and the way in which these relationships enhance or exert pressures on organizational performance. These constraints and the interconnectedness in the environment entail changes and innovations that affect business performance and management strategies of the companies in the industry (Capioto et al., 2019; Ratajczak-Mrozek, 2017). In addition, this thesis provides deep analysis of the characteristics of the platform market, which can be valuable for policy makers who want to investigate the effects of this market in relation to the economy (Dutch Transformation Forum, 2018).

The rest of the thesis is organized as follows. Chapter 2 will provide an overview of the relevant theories on population ecology, social media and social networking sites, organizational networks and embeddedness. In Chapter 3, the researcher provides arguments in the selection of research methods and describes the research design and operationalization. The findings of the study will be elaborated in Chapter 4 by reflecting the theories discussed in the previous chapters. Capter 5 will discuss further the results and the theoretical and practical implications of

the study. The thesis concludes by summarizing remarks, discussing the limitations of the study, together with suggesting avenues for future research.

2. Theoretical framework

The theoretical framework provides definitions and discussions of the important concepts used throughout this thesis and several related issues. The theories include organizational ecology and its theoretical components, embeddedness together with its influence on organizational activities. Moreover, the theories on social media and social networking sites are discussed. In addition to those theories, the chapter also provides an overview of the characteristics of the platform market.

2.1 Organizational ecology and its elements

2.1.1 Organizational ecology

In the theory of population (or organizational) ecology, populations of organizations are the core of the analysis. According to Haveman and Kluttz (2015), a population includes organizations "that produce similar goods or services, use similar resources, and have similar identities" (p. 3). The organizational ecology refers to a group of organizations that interacts and competes with each other in the same market (Hannan & Freeman, 1977). More specifically, organizational ecology theory examines the effects of characteristics of organizations, demographic processes (i.e., the birth, aging, and death), and the population ecology in which organizations either prosper or fail (Hannan, 2005). Gathering population information of a set of organizations can discover changes in the structure and forms as well as the broader population of organizations (Hannan, 2005). In addition, by observing organizations within an industry, researchers investigate how the environment and demographic variables of size and age influence organizational performance and outcomes (MacMillan & Komar, 2018).

Moreover, organizational ecology aims to identify patterns of how organizations are born, change, and die by analyzing how industries change over time (Hannan, Pólos, & Carroll, 2012; MacMillan & Komar, 2018). Research into organizational ecology also provides information on the inertia and diversity of organizations by looking at each organization's structure and identity (Hannan, 2005). MacMillan and Komar (2018) explain that as the environment changes, "structural inertia hampers the ability of organizations to adapt to new conditions" (p. 378). In the following part, the relevant concepts used for analyzing

10

organizational performance and the competition between organizations in the population will be discussed.

2.1.2 Elements of organizational ecology

Organizational ecology contains a variety of distinct theoretical components that explain organizational performance and survivability. These are identified as follows:

Niche width

A niche illustrates a company's distinctiveness in relation to its competitors in a market (Enchols & Tsai, 2005). Therefore, the concept of a population's niche provides a framework for understanding the shaping and distribution of populations under environmental variations and competition (Enchols & Tsai, 2005; Hanna, Pólos, & Carroll, 2012). As a niche consists of resources required for the population to survive and expand, Hannan and Freeman (1977) propose that each firm has to find a unique domain of resources for its survival. When it comes to organizational resources, this concept includes all tangible and intangible assets that a firm uses to implement business strategies and improve its efficiency and effectiveness (Mwai, Namada, & Katuse, 2018). Thus, firm resources are diverse and are categorized into four main groups: financial, human, informational and relational resources. Organizations can exploit these kinds of resources to achieve their business goals (Zhang et al., 2018).

In a market, companies will compete with others if their fundamental sets of resources overlap (Hannan & Freeman, 1989). In other to make themselves distinctive, firms can offer a product niche or process niche (Enchols & Tai, 2005). Regarding product niche, corporations in a niche (or niche-firms) create value by providing products or services that significantly differ from those of rivals (Enchols & Tai, 2005). In addition to product niche, a niche firm can provide particular operational processes, or "process niche" (Enchols & Tai, 2005, p. 220). Process niche expresses how a company operates and practices business in an innovative way to create value, which makes it different from its rivals (Enchols & Tai, 2005). Moreover, Podolny, Stuart, and Hannan (1996) highlight the niche position's effect on competitive dynamics when they argue that niche-firms exhibit enhanced survival rates.

An examination of an organization's niche involves looking into the adaptive capacity of organizations over the various states of the environment (Hannan, Pólos, & Carroll, 2007). Niche width theory refers to organizational fitness over a range of environmental resources (Hannan,

Pólos, & Carroll, 2007). More specifically, an organization whose resource use is limited in a narrow range of the resources spectrum (narrow niche) is defined as a specialist, whereas a generalist has a wide niche due to its high fitness in a broader range of environment (MacMillan & Komar, 2018). As a result, specialists offer a small range of products, while generalists provide products "with a larger range of variation on the dimension of interest" (Dobrev, Kim and Hanna, 2001, p.1300). Thus, 'width' represents a range of not only environmental resources but also of the product offerings of organizations (Hannan, Pólos, & Carroll, 2007). In addition, environmental characteristics affect specialist and generalist organizations differently. When an environment changes or transforms, specialists can outperform generalists as they respond quickly to these changes (MacMillan & Komar, 2018).

Resource partitioning

Resource partitioning is an important part of organizational ecology theory (Carroll & Hannan, 2004). The concept of resource partitioning originated from the field of ecology, which is the process of dividing the scarce resources in an ecosystem of similar species in a way that helps them to coexist (Hannan, Pólos, & Carroll, 2007). The organizational ecology perspective explains that internal partitioning of markets is an outcome of competition between generalists and specialists (Hannan, Pólos, & Carroll, 2007). More specifically, when the level of concentration in a market is high (i.e. a few generalists dominate the core of the resource space), the survival rate of generalists decreases whereas the mortality rate of specialists increases (Hannan, Pólos, & Carroll, 2007). Together with the well-being of specialists, the theory claims there is less competition between specialists and generalists. The reason is that in a concentrated market, generalists are likely to compete with each other to control the central position in resource base. As a result, their survival rates decrease due to the increased number of competitors at this center (Swaminathan, 2001).

Due to the competition among generalists, resources located on the periphery of the resource base are open for specialists to exploit. Therefore, specialists can take advantage of peripheral market segments without having to compete directly with generalists. It is easier for them to thrive on the periphery and outperform generalists (Swaminathan, 2001). In addition, organizational identity can affect competition levels because of consumer perceptions. For instance, it is difficult for generalists to compete with specialists as they produce particular products associated with their identity, which are in demand (MacMillan & Komar, 2018).

In the platform market, B2B digital platform survival is dependent on the contribution and relations of a multitude of actors such as owner, partners, and end users. Each actor plays relevant roles in the ecology and offers complementary resources that the platform needs in order to serve the end-users (Blaschke et al., 2018). Therefore, the ecosystem of a digital platform is a complex network of actor-to-actor interactions. In other to sustain the viability and growth of the digital platform, the platform owner or company has to facilitate effective processes of co-creation value between different actors in the ecosystem (Blaschke et al., 2018).

Isomorphism and organizational change

The concept of isomorphism describes the process that organizations within the same field resemble one another (Findik & Bedük, 2014). According to Hannan and Freeman (1977), organizations in the same environmental conditions often have similar structures and characteristics due to the environment exerting the same pressures on them. While the niche theory argues that organizations try to differentiate themselves in the market (Hannan & Freeman, 1977, 1989), the isomorphism theory explains the tendency for organizations to become similar to one another due to the pressures of isomorphism (DiMaggio & Powell, 1983). DiMaggio and Powell (1983) identify three types of isomorphism, which can intertwine and influence organizational characteristics in different ways, namely coercive, mimetic, and normative isomorphism (DiMaggio & Powell, 1983; Findik & Bedük, 2014; Frumkin & Galaskiewicz, 2004).

Coercive isomorphism is driven by external pressures, such as political influence or cultural expectations from the society or other organizations upon which the organization is dependent (DiMaggio & Powell, 1983). More specifically, political influence is conveyed through institutionalized rules and regulations of the government or trade associations that companies are obliged to follow (Katopol, 2016). Thus, companies in the same market become homogenized. Similar to coercive isomorphism, normative isomorphism is pressures from professionalization (DiMaggio & Powell, 1983). Katopol (2016) explains that employees from similar educational and professional backgrounds will perceive and attempt to solve problems in a similar way. In other words, professional norms or standardizing practices established and developed during education are adopted similarly in organizations by their staffs. For example, lawyers or doctors must comply strictly with their professional codes of ethics.

In contrast to coercive and normative isomorphism, mimetic isomorphism can be described as one company's imitation of a successful organization due to uncertainty in the course of their own actions (Findik & Bedük, 2014). For instance, when a new technology has been not yet understood, organizations can copy or mimic others, whom they perceive as a successful model, because of the uncertainty the environment has created on how they should behave. Moreover, Findik and Bedük (2014) argue that organizations tend to mimic others in their own population rather than the ones in other populations, given that organizations in the population consume similar resources, have similar structures, and experience the same structural pressures. In the case of a new organization, imitating the success of others could result from "an attempt to obtain legitimacy" in order to look like other organizations in the market (Katopol, 2016, p. 2). As a result, isomorphism could enhance organizational legitimacy and survivability (DiMaggio & Powell, 1983; Katopol, 2016).

Density dependence

This segment discusses the relationships between density (i.e. the number of organizations in a population) and competition among players. The growth in the size of the industry can initially benefit organizations as they can acquire greater access to the resources they need (MacMillan & Komar, 2018; Singh & Lumsden, 1990). However, as the number of organizations continues to rise, the competitive pressure between organizations will rise, which make the death rates increase correspondingly (Singh & Lumsden, 1990).

In general, population ecology offers diversified approaches to explore how new organizations and populations rise, change, and fall. Organizational ecology is a popular framework that helps to explain the nature of competitive landscapes and examines interactions between populations of organizations (Hannan & Freeman, 1989; Hannan, Pólos, & Carroll, 2012). This approach offers a good point of understanding the way in which organizations in the platform market interact, as well as the business competition it generates.

2.2 Social media and social networking sites

2.2.1 Social media

Social media is defined as a group of Internet-based applications and interactive platforms that "build on the ideological and technological foundations of Web 2.0, and allow the creation and exchange of user generated content" (Kaplan & Haenlein, 2010, p. 61). In other

words, these applications permit users to create, discuss and exchange information in virtual or online communities and groups. This behavior confers the emergence of a social network; hence, the term 'social network' is often used to refer to social media platforms.

Social media has drastically disrupted the manner in which individuals, communities and organizations communicate. Given that social media encourage users to generate content, social media are seen to be more effective than traditional media such as television and radio in influencing other users' attitudes and behaviors (Thackeray et al., 2008). In addition, while communication flow in traditional media platforms is "monologue" (one-to-many), social media enables a "dialogue" model (many-to-many) through conversations between authors, peers and communities on new media platforms (Hensen, Shneiderman, & Smith, 2010). With these characteristics, various types of social media can be identified in practice. Kaplan and Haenlein (2010) identify six different types of social media, including (1) collaborative projects (e.g., Wikipedia), (2) social networking sites (e.g., Facebook), (3) content communities (e.g., YouTube), (4) virtual social worlds (e.g., Second Life), and (5) virtual game worlds (e.g., World of Warcraft) and (6) blogs. Smith et al., (2017) argue that understanding the utilization of social media among companies is important, as their organizational behaviors are embedded in social networks. Among these social media applications, the most popular type is social networking sites (e.g., Facebook, LinkedIn, and Twitter) (Go & You, 2016). This thesis focuses on SNSs as these platforms have the highest number of users and reflect a fundamental change in the interactions among individuals and organizations (Smith et al., 2017).

2.2.2 Social networking sites (SNSs) and Facebook

SNSs are web-based platforms on which members can create personal profiles, facilitate friendship connections and socially interact with others by uploading, liking, and commenting on digital content such as photos, messages, and videos shared on newsfeeds (Ellison et al., 2015). SNSs with their relational affordances (i.e., unique technical capabilities) provide a specific context for understanding networks and the relationships between users (Boyd & Ellison, 2007). Hundreds of SNSs have been created, but Facebook is the largest social networking site worldwide in terms of the number of users (Statista, 2019). Facebook was the first social networking site to achieve the milestone of 1 billion registered accounts as of April 2019, it had 2.32 billion active users in a month (Statista, 2019). Facebook is chosen for this study since it is

not only the most popular SNS site but also is regarded as a relationship-oriented social media platform (Go & You, 2016). The site emphasizes on the maintenance and development of the social relationships between users in virtual communities (Kietzmnn et al., 2011), since online social networks are likely to break up into smaller communities (De Meo et al., 2012). Social networking sites such as Facebook enable two-way symmetrical communications (i.e., dialogical model), given that they allow high levels of interactions between an organization and the online public in a virtual platform (Go & You, 2016).

2.2.3 Social media as a competitive tool of organizations in population ecology

Realizing the important benefits of social media and those of SNSs in particular, organizations are now developing their social media public sites in order to build relationships with a multitude of online communities (e.g. other relevant organizations within its segment or organizations outside, even firms in its niche, stakeholders, etc.). Thus, social media has grown to become an integral strategic tool among organizations. Every organization has a presence on a social networking site, particularly the popular ones such as Facebook, LinkedIn and Twitter (Icha & Edwin, 2016). Curtis et al., (2010) point out that social networking sites (54.5%) are the most popular among users, followed by video sharing services (51.1%) and blogs (48.4%). The 2018 social media research of Fortune 500 from UMass Dartmouth reveals that less than 1% of Fortune 500 companies have no social media presence (Barnes, Kane & Maloney, 2018). Moreover, LinkedIn (98%) remains a must-have professional social media platform. It is closely followed by Twitter (91%) and then Facebook (89%) (Barnes, Kane & Maloney, 2018)

According to Parveen, Jaafar, and Ainin (2015), organizations have used social media for three main purposes, including (1) branding and promoting their products, (2) sharing and collecting information, and (3) audience engagement. These sites become a channel that external groups of users (e.g., other corporates, customers, stakeholders, and community) can learn about their business. Companies use social media for sharing the latest information about company activities (e.g. upcoming events) and products or services with the public in a real time basis (Icha & Edwin, 2016). For example, recently there has been a tendency for firms to live stream an event they are taking part in, or organize a webinar to share information about their services. Moreover, firms use social media for conducting marketing research by gathering customer information, as social networking sites such as Facebook provide useful analytic tools about their

customers (e.g. the information about the demographics of their customers) (Parveen, Jaafar, & Ainin, 2015). In addition, social media have been exploited for collecting information about competitors, for example which channels are used, the number of fans/followers and the communication strategies on their social media pages. The information about customers and competitors enables companies to construct an effective marketing strategy on social media. Additionally, the usage of social media for fostering relationships with loyal fans and reaching new customers has also highlighted (Parveen, Jaafar, & Ainin, 2015).

Besides these three main purposes, social media are useful for organizations in developing business networks and partnerships (Gloor et al., 2016). Icha and Edwin (2016) support this argument when saying that "the more time marketers invest in social media networking, the more they gain business partnerships" (p. 4). While most previous research has emphasized social media as a powerful marketing tool, Kuhn et al., (2016) stress that social media can be utilized for accessing advice. According to them, small businesses, especially those spearheaded by woman and young entrepreneurs, are more likely to use social media for seeking advices and supports from peers and experts whom they never meet in person. Thus, the authors proclaim that communications on social media and online forums are associated with business growth (Kuhn et al., 2016).

Social media as a competitive tool for organizations to interact with their environment can be considered in the following two ways (Goktepe, 2018). First, organizations can understand their environment and competitors more easily in virtual platforms. The second is that the more valuable and complete knowledge and information can be obtained through social media. In order to understand rivals, it is important for companies to examine the content strategies that the rivals deploy on their social media platforms. In doing so, companies can produce competitive communication strategies. In addition, social media are used to understand the realities of an existing economy, as social media platforms provide "the ability to analyze the topics that are popular in the industry" (Goktepe, 2018). For example, researchers and practitioners can examine the high-interaction posts to discover the popular topics and trends in a field, or investigate different types of posts to see the prominent features in the content produced among organizations (Goktepe, 2018).

In conclusion, the discussion shows that social media has a greater impact on the performance of organizations in terms of enhancing customer relations and improving

information accessibility. Therefore, social media is seen as an important tool for organizations to achieve a competitive advantage under changing environmental conditions. Additionally, by analyzing the online topics of companies operating in the same market, the researcher can discover the characteristics of that specific market and communication strategies deployed among companies.

2.3 Embeddedness conceptualization

2.3.1 Social embeddedness

The concept of embeddedness, which originated in sociology, is first introduced by Polanyi (1944) then resurrected and popularized by Granovetter (1985) (Ratajczak-Mrozek, 2017). Embeddedness refers to the process that non-economic institutions affect economic activities and outcomes, as economic activity "is closely embedded in networks of interpersonal relationships" (Granovetter, 1985, p. 496). This concept underlines the fact that economic behaviors are embedded in a context that includes economic as well as social structures (Granovetter, 1985). In other words, the economic activities of organizations are influenced by the social context in which organizations exist; therefore, the economic actions "are embedded in and influenced by that particular system of social relationships" (Montanari & Mizzau, 2007, p. 34). Briefly, embeddedness implies that a particular pattern of relationships enables the shaping of actions of actors in the network. Firms, which are involved in a dense group or are part of an inter-organizational network, encounter distinctive sets of resources as well as restraints in comparison with those who are not integrated in these networks (Moody & White, 2003).

Not only does Granovetter develop the concept of embeddedness, but he also proposes a basic division of embeddedness into two types: relational and structural embeddedness (Ratajczak-Mrozek, 2017). In general, while structural embeddedness reflects the network structure, relational embeddedness refers to the nature and quality of relationships (Granovetter, 1992). Relational embeddedness refers to the quality and the depth of relationships (i.e. strong ties or weak ties) between organizations in the network. Granovetter (1992) defined weak ties as acquaintances and strong ties as close relationships. In the inter-firm network, the strong ties are characterized by the high frequency of interaction and dissemination of information and resources between companies, while the weak ties refer to the low frequency of communication inside the relationship (Choi & Kim, 2008). Moreover, the frequency of the interaction between

companies can be inferred from the cohesion of the relations, which indicates the intensity of the connection (Capioto et al., 2019). Weak ties are beneficial because they bridge the unconnected groups and spread information, while strong ties are necessary for building networks of trust (Burt, 2000; Granovetter, 1992). When it comes to social networking sites like Facebook, as these platforms are fragmented into distinct communities, corporations may belong to several communities and generate different types of interactions suitable to each specific group.

De Meo et al., (2012) alternately define weak ties as the edges that connect users in different communities. Conversely, strong ties refer to the edges links individuals in the same groups. In the liking network on Facebook, it can be inferred that the organizations in a detected community or group have a strong tie with their group partners. However, most ties on Facebook are thus considered weak (De Meo et al., 2012), indicating that the size of communities on Facebook are relatively small, compared to the vast number of users in any given genre or field. Gloore et al., (2016) weigh in on the characteristics of the online social networking ties, by characterizing them as "more typical of low density, low identity, and low clustering coefficient weak ties" (p. 11).

In contrast to relational embeddedness, structural embeddedness emphasizes the architecture of the network, as behaviors and actions of firms are affected by the structure of the overall network (Gloor et al., 2016; Granovetter, 1992). The structure entails the patterns and characteristics of linkages between companies as a whole and the position of an organization in spreading information in the general network structure (Capioto et al., 2019). Granovetter (1992) specifies that the concept of structural embeddedness is related to the degree to which organizations being embedded in cohesive groups in the following way:

> [T]o the extent that a dyad's mutual contacts are connected to one another, there is more efficient information spread about what members of the pair are doing, and thus better ability to shape behavior. Such cohesive groups are better not only at spreading information, but also at generating normative, symbolic, and cultural structures that affect our behavior" (p. 35).

Moreover, Van Eck et al. (2011) argue that the centrality and reach of the network of firms determines the extent to which they can get benefit from the embedded ties. Yang et al. (2011) support this argument by stressing that strategic position eliminates redundant information and ties, and helps companies in gaining an advantage in relation to resources as

compared to its partners. In addition, structural embeddedness perspective has paid attention to the impact of network density and position of organizations on collaborative activities (Capioto et al., 2019; Gonzalez et al., 2014). The higher network density is, the greater interaction and collaborations happen between companies (Gonzalez et al., 2014). In a dense network, organizations are closely linked with each other and have many common connections, resulting in intense interactions and communications, redundant information, and efficient flows of information and other resources. In contrast, organizations in dispersed networks are less interconnected and seek to improve the effectiveness of the information flow (Burt, 2000). Moreover, due to the cohesive network and shared behavioral patterns and norms, firms in dense networks tend to generate less competitive actions against each other, compared to their counterparts in diffuse networks (Gnyawali & Madhavan, 2001).

The notion of embeddedness has been perceived and conceptualized in different ways. Zukin and DiMaggio (1990) develop the original notion by introducing four basic dimensions of embeddedness, including (1) structural (inter-actor ties and direct relationships); (2) cognitive (frameworks and shared social cognitions); (3) cultural (shared systems of meaning and values); and (4) political (hierarchical distributions of power and resources) (Zukin & DiMaggio, 1990; Dacin, Beal, & Ventresca, 1999). However, the classification of Zukin and DiMaggio is considered "not really consistent categories" (Hess, 2004, p.172). While their definition of structural embeddedness is similar to Granovetter as it concerns the network structure, the last three mechanisms can merge into relational embeddedness since they are related to the qualitative aspects and social context of relationships (Hess, 2004; Uzzi, 1997). In addition to Zukin and DiMaggio (1999), Gulati and Gargiulo (1999) introduce a different division into relational, structural and positional embeddedness. Although different types of embeddedness are proposed, structural and relational embeddedness appear to be the two main mechanisms, which represent the network structure and the quality of the relationships (Ratajczak-Mrozek, 2017).

In the scope of this thesis, the researcher focuses on structural embeddedness, which shows the position of each company and the way in which they interact with other companies in their network on social media. Thus, it increases the ability of each company in knowing all the players and their constraints (Montanari & Mizzau, 2007). Additionally, the structural embeddedness perspective permits the researcher to examine how external resources (e.g., networks, the flow of resources and information) influence a firm's behavior. For example, a

company that has a superior position in a network could obtain valuable resources, which increases its competitive advantage.

2.3.2 Virtual embeddedness

The emergence of the Internet and the advance of computer and communication technologies have promoted the concept of virtual embeddedness, which combines the ideas of structural embeddedness and virtuality (Morse, Fowler, and Lawrence, 2007). While the social embeddedness is achieved through "face-to-face" interaction, virtual embeddedness is obtained through "electronic linkages" and maintained by electronic communication technologies (Vasileiadou & Missler-Behr, 2011, p. 192). Thanks to the efficiency technology devices, the formation of new virtual ties is quick and low-cost, which are typical of virtual embeddedness (Morse, Fowler, and Lawrence, 2007). In other word, virtual embeddedness increases access to a wider range of potential customers and stakeholders, as organizations can develop new connections as necessary. However, as the cost of forming virtual ties is low, leaving such a tie does not affect other actors in the network (Fowler, Morse, and Lawrence, 2004). Hence, these characteristics indicate that organizations are involved in a relatively sparse network structure and less contingent on any particular actor in the network (Fowler, Morse, and Lawrence, 2004).

In addition, Fowler, Morse, and Lawrence (2004) argue that what makes virtual embeddedness different from social embeddedness is that the solutions for exchange-related problems (e.g., opportunism, uncertainty sharing information) that they offer are dependent on technological usage. For example, the risk of uncertainty on sharing information is the problem that both types of embeddedness face (Fowler, Morse, & Lawrence, 2004). In order to solve this problem, while social embeddedness focuses on sharing fine-grained and proprietary information (Uzzi, 1996, 1997), virtual embeddedness encourages widespread dissemination of both private and public information (Fowler, Morse, & Lawrence, 2004). The information on virtually embedded ties is detailed and widespread because electronic devices enable an effective transmission and storage of information (Fowler, Morse, & Lawrence, 2004). Moreover, the information on virtual ties can be individual experiences (i.e. private information) and general knowledge (i.e. public information) that are available on unlimited digital channels (Fowler, Morse, & Lawrence, 2004).

Another distinctive characteristic of virtual embeddedness is that it is a more global and inter-organizational phenomenon because of the efficiency of virtual ties (Morse, Fowler, and Lawrence, 2007). In fact, the inter-connectivity and networking construction of virtual communities is not limited by physical space, time and organizational boundaries (Vasileiadou & Missler-Behr, 2011). As virtual embeddedness's characteristics "draw on the speed, efficiency, and global reach of contemporary technologies" (Fowler, Morse, & Lawrence, 2004, p. 648), the application of virtual embeddedness is helpful for understanding the economic activities and the building of virtual networks of actors in new and dynamic industries (Morse, Fowler, and Lawrence, 2007). In the case of this thesis, the focus network is the population of organizations in the platform industry.

2.4 The impact of embeddedness on organizational activities

2.4.1 The role of social embeddedness to business performance

The network of relationships in which firms are embedded provides a more sufficient understanding of the performance of firms (Gulati, Nohria, and Zaheer, 2000). Eschol and Tsai (2005) consider embeddedness as a mediator of the relationship between niche and organizational performance. High embeddedness means that a firm operates in a dense network in which they tend to know each other well through interconnected ties. In contrast, low embeddedness shows a sparse network in which firms are not or less connected to each other (Echols & Tsai, 2005). When deciding about "what to offer and how to operate", corporations should consider the social context, in which they interact with other firms (Echols & Tsai, 2005).

Embeddedness in social networks contributes to the reduction of risk and death rate, as argued by Uzzi (1996). According to him, companies embedded in social networks have "higher survival chances than do firms which maintain arm's-length market relationships" (p. 674). In addition, an important consequence of embeddedness is access to valuable intangible resource and knowledge (Ratajczak-Mrozek, 2017; Uzzi, 1997). More specifically, Uzzi (1996) points out the characteristics of embedded ties including "trust, fine-grained information transfer, and joint problem-solving arrangements" (p. 677). A high degree of embeddedness ensures the exchange of knowledge between actors and facilitates organizational joint problem-solving (Uzzi, 1997). Moreover, Echols and Tsai (2005) emphasize that a firm can use valuable information and key resources shared in networks of inter-firm relationships to cope with competitive challenges.

22

Thus, the structure of the network is crucial for the success of any firm in a niche (Echols & Tsai, 2005).

Embeddedness is not only a source of information and knowledge but also entrepreneurial opportunities. Uzzi (1996) argues that "the type of network in which an organization is embedded defines the opportunities potentially available" (p. 675). Jack and Andersson (2002) support this argument when saying "both recognition and realization of opportunity are conditioned by the entrepreneurs' role in the social structure" (p. 467). Hence, being embedded in a social context is considered an important determinant of recognizing and exploiting business opportunities (Jack & Anderson, 2002; Ratajczak-Mrozek, 2017; Uzzi, 1996).

Embeddedness can bring many positive outcomes and enhance organizational performance. However, over-embeddedness can undermine organizational performance. The negative impact of over-embeddedness is emphasized by researchers (Granovetter, 1985; Ratajczak-Mrozek, 2017; Uzzi, 1996). According to Uzzi (1997), over-embeddedness reflects the overly strong reliance of an organization in a network. He asserts that "positive effects of embeddedness rise up to a threshold, however, after which embeddedness can derail economic performance" (p. 35). In an overly embedded network, access to information outside the network of actors is constrained. That can lead to a decrease in the amount and quality of obtained information, thus having a detrimental effect on knowledge and information accumulation (Uzzi, 1997). In addition, due to over-embeddedness, firms are vulnerable to external agents, for example the development of innovative technology; consequently, their performance is diminished (Ratajczak-Mrozek, 2017).

2.4.2 The role of virtual embeddedness to relational capital management
Relational capital

Relational capital (RC), which is a component of social capital, is built on the interaction and networking of organizations with their external agents (Khavandkar et al., 2016). Thus, RC constitutes corporations' intangible assets, including relationships to stakeholders and communities, corporate image, and corporate reputation that can change over time (Sveiby, 2001). Not only do such RC factors strengthen the reputation and the competitiveness of the company in the market (Wu, Wang, Chen, & Pan, 2008), but they also attract qualified members

and employees thus contributing to the development of human capital (Vasileiadou & MisslerBehr, 2010).

Castro, Sáez and López (2004) divided these agents into two basic groups based on the influence of them to activities of a company. The first group is related to firm activities in a direct and quick way, which forms "an industrial or closer environment" of the company (Castro, Sáez, & López, 2004, p. 579). Customers, suppliers, shareholders, and allies belong to this group. Among them, customers are involved in day-to-day relations with a firm by buying goods or services (Castro, Sáez, & López, 2004). The second group of external agents is the community groups including government agencies and market regulators, trade unions, community-based agents, and mass media. The relationships with them are important for organizations but in less direct ways. This group constructs "a general or wider environment" around the focal company (Castro, Sáez, & López, 2004, p. 580). Corporations are embedded in these environments, in which they maintain a set of relations with these external agents. Moreover, firms depend partly on their embedded ties to gain profits and success (Vasileiadou & Missler-Behr, 2011).

Relational capital management through virtual embeddedness and social media

The organization can achieve the sustainable development of relational capital by means of relational capital management (RCM). According to Vasileiadou and Missler-Behr (2011), the utilization of virtual embeddedness plays as a pivotal factor to the successful RCM. Through social media, the communication and formation of virtual communities are created in digital environments. RC is built and sustained through virtually embedded ties (Vasileiadou & Missler-Behr, 2011). Social media affects relational capital in two perspectives: organization and public as well as stakeholders. On the one hand, it enables firms to "self-publish", for example blogging or Facebook posting, to attract the attention of the wider public. On the other hand, stakeholders are empowered to share experiences and their views of products as well as express their identity in the virtual community and virtual world (Vasileiadou & Missler-Behr, 2011).

In general, the concept of embeddedness is a popular tool for the analysis of the impact of social context and inter-organizational relationships on the behavior of companies (Ratajczak-Mrozek, 2017). Embeddedness benefits organizations as embedded ties provide access to sharing knowledge and network opportunities. Organizations increase their survival rate when operating in a network formed around embedded ties. In addition, in the era of social media, virtual embeddedness sustains and promotes relational capital. Thus, understanding the effect of

embeddedness on organizational performance is important for business practices and managerial implications (Echols & Tsai, 2005; Ratajczak-Mrozek, 2017).

2.5 B2B platform market

In general, the B2B platform market comprises companies providing platform as products or services. The concept of "platform" can be defined in two distinct approaches. From the technical perspective (e.g., software, production, and development), the term means "a set of shared techniques, technologies, and interfaces that are open to a broad set of users who can build what they want on a stable substrate" (Kenney & Zysman, 2016, p. 64). From the non-technical perspective, platform refers to a marketplace facilitating interactions between different group of customers (Asadullah, Faik, & Kankanhalli, 2018; Kenney & Zysman, 2016). Nowadays, technology companies have moved their products and services to the cloud so that customers easily and directly gain access to the services they want to consume. The movement of technology-based products to the cloud has created an open digital infrastructure on which the platform market operates (Kenney & Zysman, 2016).

Today, platforms reshape industries into interconnected systems. More platform companies are being established, and platform-based markets are significantly dynamic (World Economic Forum, 2017). In the market, some of the most noticeable types of digital platforms companies can be seen. For example, some companies provide platforms for platforms (i.e. the foundational platforms that facilitate the construction of cloud services, on which other platforms are built, such as Google Cloud, Amazon Web Services) (Kenney & Zysman, 2016). Other companies provide platforms that make digital tools available online and support the creation of other platforms and marketplaces (Kenney & Zysman, 2016). Such companies operate as retails platform such as Amazon or service-providing platform such as Airbnb. In addition, there are companies offering technology platform as a type of packaged software for collecting data from many sources to help users (i.e., companies) in making decisions (Customer Data Platform Institute, 2019).

It was reported that companies used business intelligent platforms for three top priorities: reporting and analysis, predictive analysis, and performance management (Robert Walkers, 2018). In the next five years, B2B platforms would continuously emerge, focusing in "six main

industries: automotive, consumer goods, electricity, healthcare, logistics and media" (World Economic Forum, 2017, p. 3). Currently, the US and China have dominated the market. 46% of the $1 bn+ platform companies are headquartered in America, 35% in Asia (mostly China), 18% in the EU and 1% in Latin America (Dutch Transformation Forum, 2018). The total platform market value is significantly different between these regions. The US is still the leader in this market with 72% of the total value, followed by China (25%). The EU has a humble number of 2% (Dutch Transformation Forum, 2018).

As for the platform economy, an increasing number of business transactions and data aggregation continue to occur on platforms. Therefore, trust is a prerequisite element that the players in the markets must consider when building and governing platform-powered ecosystems, particularly in how data and information is shared and used (World Economic Forum, 2017). Among six principles seen as being critical to building a trusted platform ecology that are developed by World Economic Forum, 'security' is the core (see Figure 1) (World Economic Forum, 2017). As the platform market is broad, this thesis focuses on the B2B platform companies for the investigation.

Figure 1: Six core principles of trusted digital B2B platforms (World Economic Forum, 2017)

3. Methodology

This chapter presents a clarification of the methods implemented to answer the reach question and sub-question. In the choice of methods, the researcher explains the reason why network analysis and topic modeling are suitable for the thesis. Then, details of the research design as well as the process of collecting data will be discussed. In addition, the thesis will describe how the important concepts mentioned in the research questions and the theoretical frameworks to be operationalized. The operationalization section also provides the relevant explanations for a set of specific tools used to analyze the data and to visualize the results.

3.1 Choice of methods

The thesis aims to answer the research question and sub-questions as follows:

Research question: *How do social media reflect virtual embeddedness of platform companies in their population ecology?*

Sub-question 1: *In what manner are platform companies embedded on social media?*

Sub-question 2: *How does virtual embeddedness influence activities of platform companies on social media?*

In order to answer the posed research questions, a qualitative mixed methods approach including digital methods of network analysis and topic modeling is employed. The application of social media in examining the network of actors in organizational ecology is a complex phenomenon because of different interpretations, contexts and company's strategies. A qualitative approach is effective because it allows researchers to describe and explain the phenomenon within their contexts (Baxter & Jack, 2008). In a qualitative study, researchers take an involved role that enables them to have in-depth data analysis and become an active member in the network when interpreting the data (Neuman, 2014; McKenna, Myers, and Newman, 2017). In addition, thanks to mixed-method approaches, data can be gathered from a variety of sources, which strengthen the credibility of the data as well as results in a comprehensive understanding of the phenomenon (Baxter & Jack, 2008).

3.1.1 Social network analysis

Social network analysis (SNA) is used to answer the Sub-question 1: *In what manner are platform companies embedded on social media?* SNA is an increasingly popular method in the social and political sciences and relies on the network as its central construct (Ahrens, 2018). Researchers use SNA to investigate social phenomenon through relationship patterns between actors. These patterns are expressed in network structures, which are a set of nodes (e.g., organizations) and edges (e.g., relationships or interactions of organizations) (Wasserman & Faust, 1994). More specifically, set of edges that connect organizations will link up to form paths, which provides a system for organizations in the network to directly or indirectly influence each other. A central principle of SNA is that the position of an organization in a network structure helps determine the chances and constraints it will face (Kane, Alavi, Labianca & Borgatti, 2014). The qualitative interpretation in SNA enables a deep understanding of the mechanisms generating these patterns (Schipper & Spekkink, 2015). Moreover, SNA is suitable for capturing the structural embeddedness of organizations. Because when investigating their embeddedness, the researcher has to look at relationships of all neighbors of the focal organizations and the density in the organization's immediate, egocentric network (Wasserman & Faust, 1994). Thus, the network analysis allows the researcher to discover how organizations are embedded in their market through various network measurements that can translate to the extent of their embeddedness.

3.1.2 Topic modeling

After identifying the connection of companies in their liking networks through the network analysis, the thesis conducts a topic modeling to find the topics mentioned in Facebook of organizations in this specified network. This examination is to answer the sub-question 2: *How does virtual embeddedness influence activities of platform companies on social media?* Topic modeling is a text analysis method that reveals hidden themes through a collection of documents (Blei, 2012). Accordingly, topic models are "algorithms for discovering the main themes that pervade a large and otherwise unstructured collection of documents" (Blei, 2012, p. 77). Especially, this method is used for the analysis of massive document collections in finding patterns in genetic data and social networks (Blei, 2012). Applying this method allows the researcher to access the alignment of topics between companies that are linked and those that are not linked in the network, and how companies distinguish themselves from each other in their

market through the topics they have posted on social media. This method is useful for answering the research question as it can lead to the expectation that companies become similar to each other because of embeddedness and isomorphism pressure.

3.2 Data sample and collection

In this thesis, a competitor database in the platform market was accessed to identify a set of competitors. As this thesis is part of a larger collaborative project by the companies Ephita and Barona Strategy Consultants, there is a special interest in the competition surrounding a particular CDP company, *NGData*. To conduct the social network analysis and topic modeling, the researcher collected data from the public Facebook pages of *NGData* and its two top competitor companies as identified from the competitor databases, namely *CB Insights* and *G2 Crowd* (see Table 1 for details of the chosen companies). Netvizz was the tool used for scraping the data on Facebook (Rieder, 2013).

Table 1: Details of three focal companies

Company	Founded	Headquarters	Products/ Services
NGData	2012	Gent (Belgium - EU)	Customer data platform to capture data of customer across all channels, provide data-rich brands with real-time insights and drive connected customer experiences (www.ngdata.com)
G2 Crowd	2012	Chicago (US)	Platform for users to share business software reviews in real time (www.g2.com)
CB Insights	2008	New York (US)	Platform to aggregate and analyze massive amounts of data, helping corporations to answer strategic questions (www.cbinsights.com)

Netvizz is a Facebook app that permits scraping of post activities and comments, and retrieving of networks from public pages (Rieder, 2013; Thelwall, 2017). Researchers can analyze data quantitatively and qualitatively from these kinds of social media content (Rieder,

2013). In the scope of the thesis, the Facebook liking networks were extracted from the two selected companies as well as *NGData*. These all indicate a variety of online organizational affiliations (Mislove et al., 2007). Then, the online content (i.e., posts) on Facebook of eleven companies across three core networks was gathered for topic modeling. Among eleven companies, one is from *NGData* network, five companies are from *CB Insights*'s network and five are from *G2 Crowd*'s network. The database has one company from *NGData* as it liked only one page on Facebook. The ten pages from *CB Insights* and *G2 Crowd* ecosystems were selected randomly as their networks are rather big (see Table 2 below). The chosen companies are platform companies as well as the other relevant ones. For being selected, the pages must have high number of post activity, which indicates that they are active on Facebook. From the Facebook page of each organization, the last 500 posts were gathered for the process of topic modeling. Ultimately, approximately 7000 posts were collected and analyzed.

3.3 Operationalization and Analysis

3.3.1 Network analysis

The network analysis was implemented in the beginning to examine the relationships and interactions between three selected companies (*NGData, CB Insights,* and *G2 Crowd*) and other companies on Facebook. Netvizz tool was used to retrieve all of their public Facebook pages (organizations) and their ego-networks (Rieder, 2013). An ego-network is a local network comprising a focal company (i.e. the ego) and its online affiliations – other companies to whom the focal one directly connects via the Facebook liking relationship (aka neighbors) – and the relationships among those neighbors (Wasserman & Faust, 1994). Crawl depth of each network can extend to a 2nd layer of affiliations from the selected firm (Rieder, 2013). However, this thesis focused on the depth 1 as it includes all of the organizations that have direct liking networks with the three focal organizations, resulting in a more accurate result in the analysis of the embeddedness of firms in these networks. The network analysis program Gephi (Version 0.9.1) was employed to visualize these network graphs. Gephi is an open-source software, which enables the researcher to explore and analyze qualitatively networks (Bastian, Heymann, & Jacomy, 2009). The tool helps the researcher to discover patterns in the network based on its structure and components not just visually but also quantitatively through various network measurements.

Shaping the graph by applying layout algorithms allows for a precise interpretation of the graph in SNA with the fewest possible biases ("Gephi Tutorial Quick Start," n.d.). Force Atlas, a force-directed algorithm was used as it arranges nodes and edges based on structural proximity as well as planarity, both leading to more accurate visual interpretation ("Gephi Tutorial layouts," n.d.). The type of layout and its properties were applied similarly for three companies and the details of their generated networks are as follows:

Table 2: Nodes and edges in the generated network of the focal companies

Company	Nodes	Edges
NGData	2	1
CB Insights	38	71
G2 Crowd	67	198

In order to understand the key characteristics of the network, measures offered by Gephi and examined in the following analysis are modularity, the clustering coefficient, degree centrality, and betweenness centrality; these are elaborated upon further below. Based on the visualization of the graph and the calculated parameters, niche width, resource partitioning, and embeddedness of entities in the networks were also examined. In addition, Gephi permits the examination of all networks as a single unified network to see if there exist overlaps or bridges between each of the egocentric networks, thus revealing some sense of their embeddedness relative to one another.

Measures for niche width and resource partitioning

From the network, generalists (wide niche) and specialists (narrow niche) in the platform market can be identified through the number of companies they connect as well as the diversity in categories of these firms. The category of the connected firms is included with the network metadata. A generalist is likely to link with many more companies than a specialist is and with companies in a variety of different areas, as they provide a wide range of products or services (Dobrew, Kim and Hanna, 2001).

Detection of structural embeddedness

Structural embeddedness is reflected in the overall connection pattern. The basic metrics of analyzing the extent of embeddedness of a firm in its network include:

(1) Community detection (Modularity): Modularity is used to measure the extent to which a network clusters into different groups or communities. Groups in a network are created by sets of nodes within which there are many edges, with fewer edges between nodes of different groups (Newman, 2004). Modularity detects and quantifies the different communities that exist within a network. In other words, it determines the community structures in the network and provides insights of the relationships and attributes of the users (Newman, 2004; Lee, 2019b). The computed modularity score lies in the range 0.0 – 1.0 (Blondel, Guillaume, Lambiotte, & Lefebvre, 2008), and a value above 0.3 is a good indicator of meaningful community structure inside a network (Newman, 2004). The division of communities indicates partitions in the network and hence either the homogeneous or fragmented manner in which the focal company is embedded.

(2) The degree centrality: The centrality of an organization is predictable from characteristics of the network in which it was embedded (Freeman, 1978). This popular measure counts the number of direct connections among nodes and shows the level of a node's involvement in their local network and the influence of the node on its neighbors (Freeman, 1978; Lee, 2019b). Organizations that are in the center of the network may have a considerable influence. Another measure, betweenness centrality, can capture another aspect of influence. This measure expresses the proportion of shortest paths – among all pairs of nodes – that a node lies on. Nodes with high betweenness play critical roles in the network structure, as they are able to be the gatekeeper/moderator who control information flows within the network; other companies will reach each other through them (Easley & Kleinberg, 2010).

(3) The clustering coefficient: this measure captures the density of an ego-network and demonstrates the degree of the direct local embeddedness of an organization in a network (Cerqueti, Ferraro, & Iovanella, 2018). The clustering coefficient of a node with its neighbors measures the number of directed links among a node' existing neighbors divided by all possible links (Mislove et al., 2007). The clustering coefficient indicator also ranges from 0.0 – 1.0 (Watts & Strogatz, 1998). A high score indicates that actors are embedded in a tightly linked local network (Cerqueti, Ferraro, & Iovanella, 2018). Then, the tightly-linked group incurs not only beneficial embedded interactions (such as cohesiveness), but also less beneficial ones (such as constraint); however, the latter issue is more applicable to interpersonal social networks than inter-organizational networks (Easley & Kleinberg, 2010).

(4) The graph density: This measures shows the number of ties relative to the number of possible ties and ranges 0.0 – 1.0. Higher density means that the firms in the overall network are more interconnected. In such a dense network, firms tend to know each other well through interconnected ties and vice versa (Drieger, 2013). This number indicates the extent of embeddedness of the entire network (Echols & Tsai, 2005). High embeddedness means that a firm operates in a dense network. In contrast, low embeddedness is associated with a sparse network in which firms are less connected to each other (Echols & Tsai, 2005).

3.3.2 Topic modeling

After network analysis, topic modeling was conducted to investigate the manner of content-based embeddedness of platform companies through the posts on their Facebook pages. As for the topic modeling, topics were discovered based on the words of each document; each document is a Facebook post (Diesner, 2013; Mohr & Bogdanov, 2013). Each of the 7000 posts extracted by Netvizz tool constituted a distinct document. Topic modeling was achieved through the text analysis software ConText (Version 1.2.0). This program was developed by The University of Illinois, which assigns topics based on latent Dirichlet allocation (LDA) (Diesner et al., 2015). LDA is a statistical model of detecting themes or topics from document collections (i.e. corpus), that is developed on the co-occurrence of key terms that can indicate a topic (Blei, 2012). Each document (here, a distinct social media post) can exhibit one or more topics. Thus, a topic list with numeric weights of each topic's prominence was obtained from the algorithm. This list per company's corpus characterizes the topical nature of each company examined and can be qualitatively labeled by the researcher. When labeling the topics, the researcher referred to the document-to-topic output (i.e. one of the outputs of topic modeling) to identify the presence of topics in specific documents (see Appendix A for an example).

Before the topic modeling function, pre-processing steps (delivered by the program ConText) typical in computational text analysis were conducted. Noise/stop words, such as 'the', 'a', 'an' in the data were eliminated from the text, and verbs and plurals were stemmed or lemmatized to their original forms (Diesner et al., 2015). Subsequently, the topic modeling function was applied to the cleaned data. The most time-consuming part is labeling of the topic. Topic labels will require qualitative analysis of each of the topic member words and phrases as well as examining some of the source documents (Bartian, 2015). Seven topics per each

33

company were labeled, and the results show how the companies have presented themselves on social media. Then, in order to display the comparison of the topics across selected companies, topic models were visualized again through the Gephi tool (Bastian et al., 2016).

Examining niche width, resource partitioning and virtual embeddedness

After labeling the topics (i.e. group of words) of *NGData, CB Insights* and *G2 Crowd*, as well as eleven connected firms, a total of 98 topics were inferred and compared. While the differences between topics indicate how a company differs from their competitors in the niche, the topic alignment and misalignment can examine against companies' diversity of resources and stakeholders, which are represented by the neighboring or affiliated companies in their online networks.

Measures for virtual embeddedness

Based on the comparison between topics, the researcher is able to assess the extent and manner of embedding. There might be overlaps, partitioning, alignment or diversity of resources. In addition, the detected topics from three focal companies and those companies linked to them in their liking networks may present the details of the platform market. Furthermore, the company-to-company network created via folding of the topic model as a network can reveal the partitioning or alignment of the topics between companies. In such a network, companies have weighted edges based on the extent they share similar topics that emerged from their social media posts.

3.4 Validity and reliability

In term of validity, the thesis uses Facebook liking networks and posts of platform companies and the relevant organizations to analyze their embeddedness on social media. Takac Hinz, and Spann (2011) argues that the online world can reflect the real world. Moreover, when it comes to the business performance, social media content can truly demonstrate the opinion and position of a company in the virtual world (De Meo et al., 2014). Especially, in the B2B context, the content and conversation on social media such as Facebook or Twitter can be seen as endorsements from trusted and known sources (Brennan & Croft, 2012). In addition, companies are mindful of communication strategies when expressing their presence on social media, such as choosing suitable online channels (e.g., Facebook, LinkedIn or Twitter), as well as the way they present online content, how to develop customer engagement, or following other companies

(e.g., competitors or relevant ones) (Go & You, 2016; Goktepe, 2018). Furthermore, relational capital management on social media is also central to a company's online presence (Vasileiadou & Missler-Behr, 2011). Thus, their online network relationships are likely to be made judiciously, so they would reflect instrumental linkages (Vasileiadou & Missler-Behr, 2011). That is, these companies are not going to connect to businesses are totally different from them, e.g., McDonald's, without rational reasons.

As for reliability, the information of the type and characteristics of the products that the companies offer are not identified on the Facebook liking network; therefore, niche width and resource partitioning are not explicitly exhibited. To enhance the reliability of the various findings, the researcher conducted offline investigations (e.g., collecting relevant information on their webpages and other relevant websites) to support the online data and the analysis. Moreover, labeling the topics is subjective, which might influence the reliability of the results. In order to reduce the personal bias in interpreting topics, data triangulation – by which a qualitative mixed-methods – was applied (Noble & Smith, 2015). Thanks to the mixture of research methods and the combination of qualitative analysis and quantitative numbers produced from digital tools such as ConText and Gephi, the findings are considerably more robust. Regarding the research method, LDA is a probabilistic statistical algorithm; therefore, its outputs are not always the same. Yet, the top most prominent topics tend to be quite stable (Blei, 2012).

4. Results

4.1 Virtual embeddedness reflected through the Facebook liking network

The first part of the result chapter is to answer the sub-question: *In what manner are platform companies embedded on social media?* Social network analysis of the Facebook liking network was conducted to investigate how *NGData, CB Insights,* and *G2 Crowd* are embedded in their market. Various network measurements such as modularity, clustering coefficient and centrality were calculated to translate the manner and extent of embeddedness. Netvizz tool was used to scrape the liking network of these three firms, and then their collected networks were visualized by Gephi. The results show the basic information of their liked pages, such as name (Id), category of those company according to their product or services, and post activity (i.e., posts per hour, based on the last 50 posts), the number of fan etc.

4.1.1 Detection of ecosystem and relational capital of three focal companies

The digital business ecosystem of three focal companies consisting of the pages that they like is identified. As *NGData* likes just one page, *seopus+*, its ecosystem as well as liking network is simple. In contrast, *CB Insights* and *G2 Crowd* have more complicated liking networks, and their linked firms are more diversified in terms of business categories (see Figure 2).
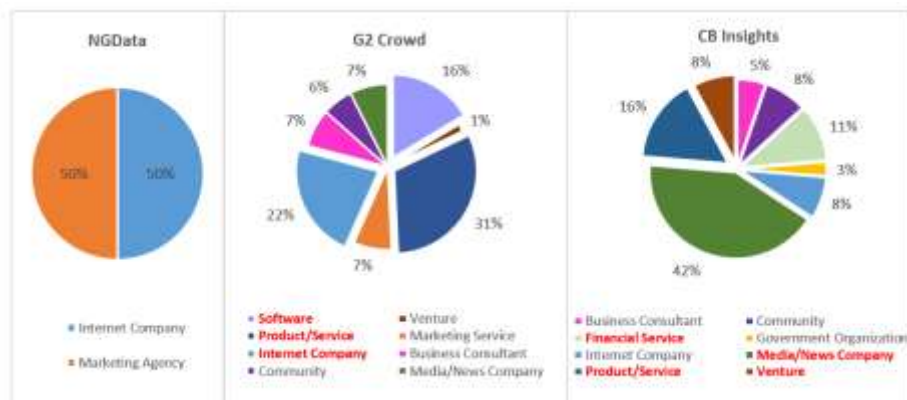

Figure 2: Companies linked to NGData, G2 Crowd and CB Insights on their Facebook pages per business category

In general, three focal companies affiliate with companies who share common interests or are doing business in relevant fields. While *NGData* is a customer platform company providing customer relationship management (CRM) solutions (www.ngdata.com), *seoplus+*, the only company connected to it is a Canadian digital marketing agency delivering solutions in relation to online marketing needs (www.seoplus.ca). Investigating further in the *NGData* website, the researcher discovered the relation between them, as Amanda Murray, a content marketing specialist of *seoplus+*, was featured on the articles detailing tips about real-time marketing on NGData's website (Stringfellow, 2016). This highlighted the argument that elements from the real world (offline) can be reflected in the online social networks (Takac, Hinz, & Spann, 2011).

The ecosystems of *G2 Crowd and CB Insights* are more dynamic than *NGData* in terms of quantity and the type of companies. The dynamic of this ecosystem and the links in the liking networks reflect the diversity of resources that these companies can exploit to achieve their business targets (Zhang et al., 2018). As for *G2 Crowd*, the company owns a review platform headquartered in Chicago, Illinois. The firm offers ratings and reviews on CRM and related, e-commerce, IT, software, etc. from real users. Reviewers, software users, and technology buyers are their target customers and stakeholders ("Company overview of G2 Crowd, Inc.", 2019). These groups can be observed in G2 Crowd's liking network. Most of the business pages G2 Crowd liked are technology-based ones. Those providing technology products or services represented 31% of the total companies. Internet companies including platform companies, e-commerce, and digital companies accounted for 22%, and software companies made of 16%. These connected companies play different roles in the ecosystem of *G2 Crowd*, creating a complex network of actors as described by Blascheke et al., (2018).

In the ecosystem of *CB Insights*, it is obvious that the company is interested in media/news companies, venture capital, and startups. Indeed, the company's products (and services) is a market focused intelligent platform analyzing data on venture capital, startup and news mentions to predict the next technology trends. In more details, more than 40% of the total companies that *CB Insights* connects to on Facebook are media/news and publishers. Companies that operate in the financial section represented nearly 19%, including financial services (11%) and venture capitals (8%). The linked companies on Facebook can contribute to their resources, for instance informational or relational resource, which *CB Insights* can employ to benefit their

business. Similar to *G2 Crowd, CB Insights* ecosystem presents a network of diverse resources including users, suppliers, allies and stakeholders (Blascheke et al., 2018).

It is important to notice that two online groups of external agents, which can contribute to relational capital of each core company, have been identified. On the one hand, the industrial or closer group that influences them directly includes customers, suppliers, shareholders (Castro, Sáez and López, 2004). In the ecosystems of these three core companies, they are other relevant companies and share the same interests with the focal company: *Seoplus+* in NGData network; media/news companies, venture capital, and startups companies in CB Insights's network; as well as Internet and software companies, product and service, business consultants in G2 Crowd's network belong to the industrial group. On the other hand, the core companies also connect to community-based agents. For example, *CB Insights* liked *The National Venture Capital Association*, a trade association that represents the U.S. venture capital industry. *G2 Crowd* connected to *1871 Chicago,* an entrepreneurial hub for digital startups. The community group is important for these companies but in less direct ways (Castro, Sáez & López, 2004). The group accounted for a small number of total connected companies in the both networks, *CB Insights* (8%) and *G2 Crowd* (6%) accordingly.

4.1.2 Virtual embeddedness reflected of each core company

As mentioned in the research design, community detection reveals the different clusters in the network structure and allows for assessing a single aspect of embedding. To recall, community detection produces a measurement indicating the extent to which a network is partitioned into distinct communities. As edges to the ego (focal company for each ego-network) bias the community detection, all of them are removed before the communities are detected. This provides a set of communities based on the affiliated companies' affiliations (likings) to one another. The isolated nodes (i.e. single and disconnected node) of such a network indicate those affiliates that have only a connection to the ego. As the network of *NGData* has only two nodes and one edge, its modularity is 0; therefore only one community was detected. In contrast to *NGDat*a, the results of CB Insights's and G2 Crowd's liking networks are more interesting.

The modularity parameter of *G2 Crowd* is 0.433, which indicates a rather meaningful division of the network into distinct groups (Newman, 2004). The result reveals 22 groups. The five connected groups were colored by green, red, blue, purple and orange. The nodes (i.e.,

companies) belonging to common communities were also assigned the common community ids and colors as seen in Figure 3.
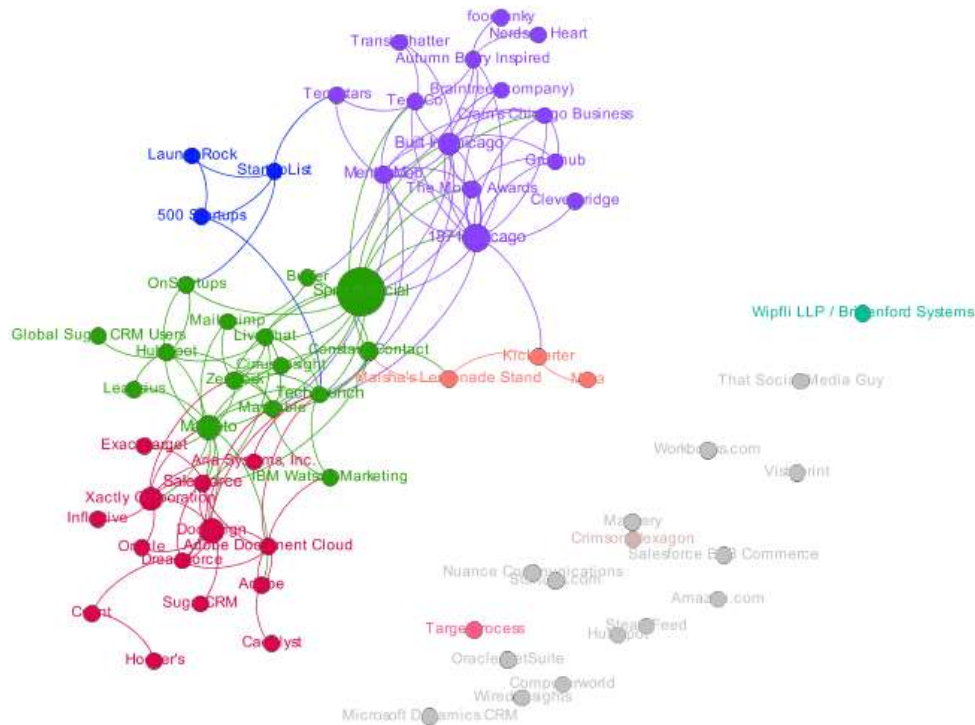
Figure 3: Community detection of G2 Crowd

The communities identified can also be somewhat distinguished visually (i.e., more edges within a community than externally). These communities formed through affiliations among G2 Crowd's neighbors characterize its embedding. Many of the nodes in this community, which are not connected to any other nodes, are called "pendants" to the ego. They consist of a diverse range of companies including software, technology, and product/service companies, and are located in the right corner of the network. The blue group is the smallest group with a modest number of three members. However, this community is a distinct group as all of its members are companies related to startups, including *Startuplist, 500 Startups,* and *LaunchRock* (a startup launch platform). Moreover, all three startups are linked together and created a triadic linkage. This type of small network structure can indicate embeddedness of a tightly-linked group that incurs beneficial embedded interactions for the actors (Easley & Kleinberg, 2010). Indeed, the mutual actors will push the interactions among the three actors, potentially raising trust and

confidence in the integrity of the transactions (social, economic, or otherwise) between them (Easley & Kleinberg, 2010).

In addition, there are overlaps between the green and red groups, as both comprise marketing-related companies as well as software companies, implying a close connection between marketing-related companies and software. One can argue that marketing-related businesses are consumers of software companies. However, two groups are somehow different and identifiable. While the green group comprises mostly of marketing-related companies, such as *Mailchimp, Marketo, IBM Watson Marketing, etc.,* the red group consists of mostly of software companies, such as *Adobe, Oracle, Xactly Corporation, etc.* Regarding the purple group, one of the outstanding characteristics is that it has three companies based in Chicago, namely *1871 Chicago, Crain's Chicago Business and Built in Chicago*. Three of them connect to one another and span the group. Among them, *1871 Chicago,* a community company, is located in the center and connects to most of the nodes in the cluster.

Additionally, some companies play the role of bridging nodes, which connect the disconnected groups together (Lee, 2019b). For example, *Buffer* (an internet marketing service company) and *Sprout Social* (a social media software company) bridge the green (marketing-related) community and the purple community, which seems characterized by miscellaneous companies in the Chicago area. *Techstars* (a venture capital) bridges the blue (startup) community with the local (Chicago) group. The partitions in the networks also highlight the diversity of resources, which is explored further below.

We can now characterize the embedding of *G2 Crowd* to be modestly diverse with communities that are homogeneous along certain traits (i.e., the marketing-related, technology-related companies, startup and local area companies) and others that are more diverse, and inclusive of businesses that appear to be (or characterize) their customers, namely product/service companies. Similarly, there is some partitioning within these communities that have a mixture of resources: customers and technology affiliates. Considering this partial mixture and partial partitioning along with the high modularity score Q of 0.433, one might then characterize G2 Crowd's embedding to be partly-partitioned and modestly resource diverse.

Regarding *CB Insights* network, the community detection is also moderate (0.405), indicating a meaningful group-division (Newman, 2004). The modularity algorithm revealed 18 groups. Similar to *G2 Crowd*, the community partitions are clearly observable. The nodes that do

not link with any other nodes (called pedants) are separated and located in the left side of the network (see Figure 4). The connected nodes were grouped and colored in pink, orange, green, and red. In the network, the green group is distinguished by having news and media corporations. The big names are *Harvard Business Review, Bloomberg* and *The WSJ.* As most of the nodes in this group are linked to one another, the group is more close-knit and has common interests (Easley & Kleinberg, 2010). Another outstanding group is the orange group. Four of six companies in this group are venture capital firms and banks. Among them, *Silicon Valley Bank* is at the center and connects to most of the nodes. The two remaining companies in this group are business magazines, namely *DC Inno* and *Silicon Valley Business Journal*. In addition, bridging nodes revealed are *Venture Beat* and *Fortune*. *Venture Beat* (a venture company) links the orange (financial-related) and the pink group, which comprises a mix of startups, venture capitals, software and news. *Fortune* bridges the green (news and media), the red and the orange (financial-related) communities.

Similar to *G2 Crowd*, the network of *CB Insights* is relatively diverse, and partitioned with a moderate community detection (0.405). The pendants are more diverse, and include companies that appear to be their partners, namely startups and venture capital firms. They are situated in the bottom left corner of the graph. Meanwhile, other groups have some common attributes (i.e., the news and media group, the financing-related groups) spread throughout the remaining part of the graph. In addition, *CB Insights* network is a mixture of resources (i.e., customers, startups and news and media affiliates). Thus, it can be inferred that the embedding of *CB Insights* is partly partitioned and intensely-resource diverse.
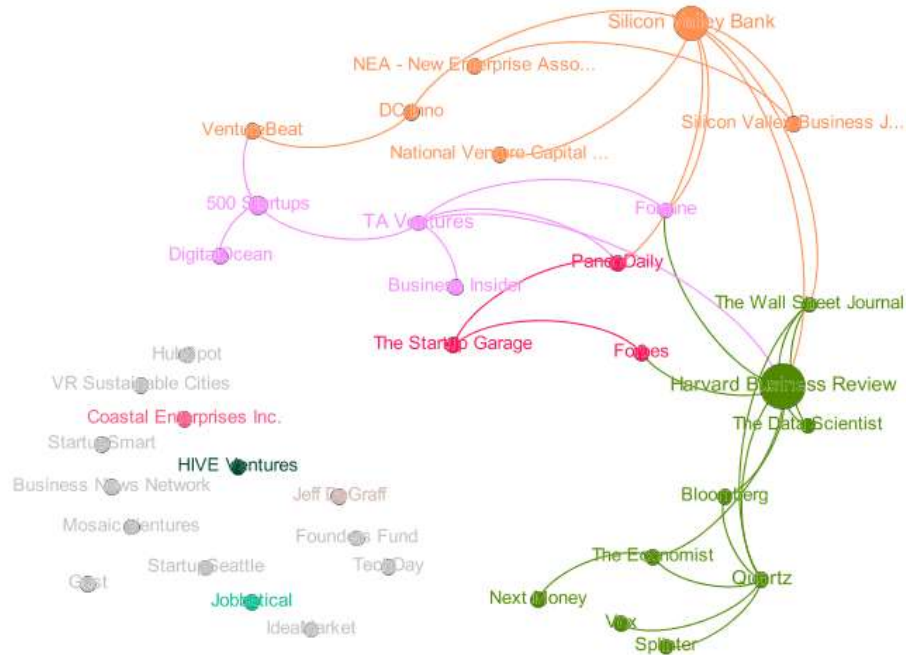
Figure 4: Community detection of CB Insights

In a network, the behavior of nodes (firms) is likely to be influenced by their neighbors (Cerqueti, Ferraro, & Iovanella, 2018). Thus, the clustering coefficient is run to understand the level of cohesion around the focal company (Cerqueti, Ferraro, & Iovanella, 2018). As discussed above, the clustering coefficient of a node ranges from 0 – when none of the neighbors of the focal node are connected to one another – to 1 – when all of its adjacent nodes are connected to each other (Easley & Cleiber, 2010). In an ego-network, the ego will have the highest degree centrality and betweenness centrality, thus they may not be worth mentioning. Alternately, the researcher investigated the corporations having the second highest of each score. The organizations identified are *1871 Chicago* and *Silicon Bank Valley* (see Table 3). As these organizations have the highest degree centrality and betweenness centrality after the focal companies, they are characterized as gatekeepers and span local structural holes (Burt, 2000). The high numbers also demonstrate that they are the most influential player in the network due to their ability to control more of the information flows (Borgatti et al., 2009, Drieger, 2013).

Harvard Business Review (HBR) is the broker in CB Insights's network. In the network, this management magazine reaches the pink group – which includes a mixure of venture, software and media – via *Fortunes,* and the red group via *Forbes.* On the one hand, through the direct connections with the central player of each group, HBR could obtain fewer redundant and

42

more valuable sources of information (Burt, 2000). On the other hand, the diversity in HBR's connections across the separate groups makes it more attractive to other organizations as the contact in its respective communities (Burt, 2000).

In G2 Crowd's network, *1871 Chicago* belongs to the purple group, which is distinct with three companies based in Chicago – the headquarters of the focal company. The organization is a Chicago center for technology and entrepreneurship (1871.com). As a community organization, *1871 Chicago* is considered having a less and indirect impact on the activities of companies (Castro, Sáez and López, 2004). However, its position of high betweenness in the network of *G2 Crowd* could raise its power. *1871 Chicago* is tied with most of the organization within its group*;* in addition its relationship with *Tech Crunch,* a conduit for information in the green (marketing-related) group and *Kickstarter,* a conduit for information in the orange group. Being as a gatekeeper gives it an advantage in terms of information access (Burt, 2000). In addition, *1871 Chicago* could be a mediator of business deals and 'a crucial hub for communication and political decision making' (Borgatti et al., 2009, p. 10); thus, *G2 Crowd* can take advantage of the relationship it has with *1871 Chicago* to gain benefits towards its business.

Table 3: Linking network characteristics of gatekeepers and the two focal companies

Company	Clustering Coefficient	Betweenness Centrality	Degree Centrality	Graph Density
G2 Crowd	0.029	769.55	71	0.05
1871 Chicago	0.230	161.83	18	
CB Insights	0.024	62.5	39	0.045
Harvard Business Review	0.188	14	10	

In addition, the graph density algorithm was calculated. The results indicate a low number in both networks, *G2 Crowd* (0.05) and *CB Insights* (0.045). This low density signifies sparse networks, where firms in these networks are either partially connected or totally disconnected from each other (Echols & Tsai, 2005). Since G2 Crowd' and CB Insights' networks are fragmented into many communities due to high community detection values (De Meo et al., 2014). Then, the clustering coefficients of two gatekeepers were calculated and the results strengthen the argument of the sparse network. These two organizations have a low value of clustering coefficient, *1871 Chicago* (0.230) and *Silicon Valley Bank* (0.160), indicating that

they are embedded in a loosely connected local network (Cerqueti, Ferraro, and Iovanella, 2018). In general, the structure of the liking networks of both focal companies, *G2 Crowd* and *CB Insights,* are characterized as sparse networks, low density and low clustering coefficient, which are described by De Meo et al. (2014) and Gloor et al. (2016).

4.1.3 Niche width detection of each core company

In terms of the niche structure in the platform market, based on the center position plus a wide range of products and services (see Figure 5), *CB Insights* is likely a generalist platform company. In contrast, *Gust,* a SaaS funding platform connecting startups and investors, is a specialist platform company as it offers services for a specific group of customer, as well as it does not connect to other companies in this network. Another specialist that has been discovered is *IdeaMarket*, a crowd-sourced marketplace that matches entrepreneurs to establish new companies. These specialists are pendants and are located in the periphery of the network without any interactions with the other nodes. Investigating further into their websites, the researcher discovered that while *Gust* is doing well since its foundation in 2007, *Idea Market* ended its business in 2016, two years since it was established (ideamarket.com). Hence, it is worth noting that their roles in *CB Insights* network cannot reveal why one succeeded and one failed, unless the ego-networks of each of these companies are examined.
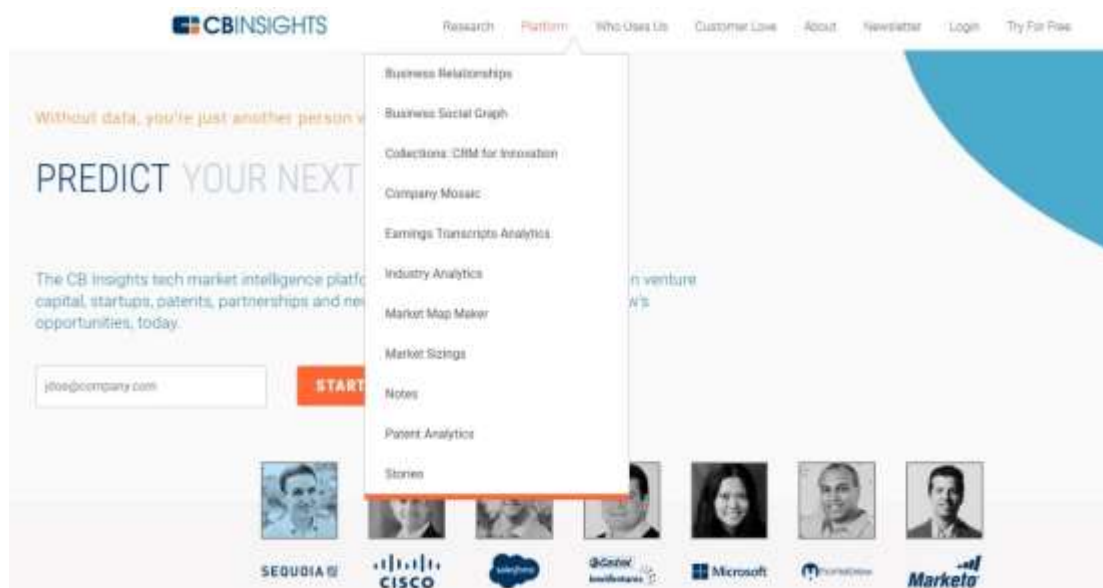
Figure 5: Platform provides by CB Insights (www.cbinsights.com)

Similar to *CB Insights, G2 Crowd* is a generalist platform company in its network acting as its central actor and provides a broad range of software and services (see Figure 6). Except for *Hubspot*, which provides diverse software in marketing, sales, and service, other companies such as *Docusign, Marketo, Sugar CRM* provide a particular product (the so-called products niche) meeting the specific demands of their target customers (Enchols & Tai, 2005). For example, *Docusign* provides an e-signature solution platform, *Marketo* offers marketing automation software, etc. However, as we can see from the products of all these companies, they are all listed in Software and Services reviewed by *G2 Crowd*. Hence, they are considered as specialist companies and are scattered in the periphery of the network.

Figure 6: Software and Services provided by G2 Crowd (www.g2.com)

4.1.4 Niche structure and virtual embeddedness of the three core companies in their population ecology

After detecting the liking network of each company, an analysis consisting of the population of three core platform companies was created to examine the niche structure and some sense of their embeddedness relative to one another. First, two metrics – graph density and average clustering coefficient – were calculated to examine the whole structure. While the graph density algorithm captures the overall interconnectivity of organizations in the network, and the average clustering coefficient gives an overall indication of the clustering in the network ("Gephi Tutorial Quick Start," n.d.). The results reveal that the population network is sparse and non-cohesive with low numbers of both of the measured metrics, graph density (0.025) and avg.

clustering coefficient (0.277). It implies that the companies in this population ecology are less connected to each other.
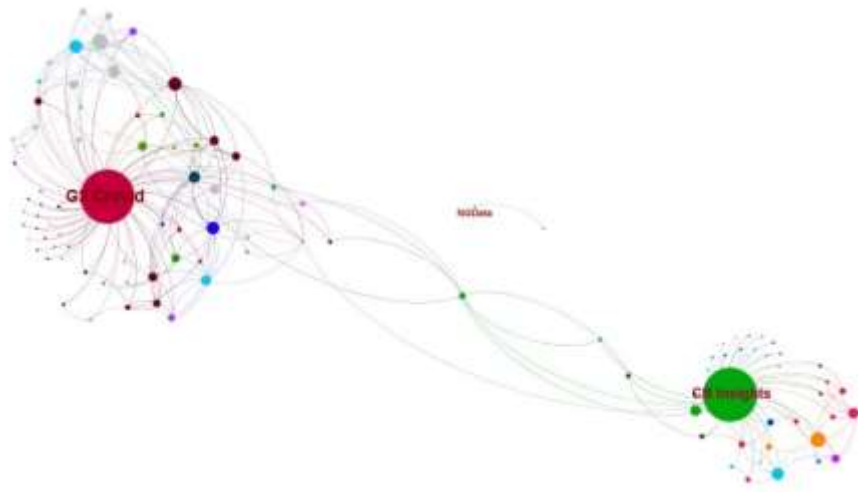
Figure 7: The liking network of the platform companies in their population ecology

In the network, *G2 Crowd* and *CB Insights* are the prominent actors. The network is divided into two hubs, in which each of them is the central actor (see Figure 7). In a stark contrast, *NGData* is a least active actor with the lowest degree centrality value of one and does connect to a single component network created out of CB Insights's and G2 Crowd's ego-networks. The partitioning of the population network is easily observable. The liking network of *NGData* is small with two companies and is separated from the overall population ecology. It can be inferred that *NGData* is a specialist company. *G2 Crowd and CB Insights* communities link together via mediators. The standout one is *500 startups*, who sits in the center of the network. The companies to whom *500 startups* connects on both sides are ventures, startups and related ones such as *TA Ventures, Ventures Beat, and TechCrunch.* Besides *500 Startups,* both *CB Insights* and *G2 Crowd* connect to *HubSpot,* a provider of a full platform of inbound marketing, sales and customer service. The overlapping of resources could lead to an increase in competition between companies (Hannan & Freeman, 1989). However, as the number of common connection is insignificant, this situation hardly happens between these focal companies.

As *NGData* network is small and simple, the researcher investigated deeper into the 2nd layer of its liking network. The results include all organizations that are indirectly linked with it through the directly linked firm, *seoplus+*. 15 organizations and 20 edges are revealed in the

expanding network. *Seoplus+* becomes the central actor as the company connects to all nodes. A new analysis comprises three networks of the core companies (i.e., the expanding liking network of *NGData,* CB Insights's and G2 Crowd's network) were established and visualized by Gephi (see Figure 8). Even when looking further into NGData's ecosystem, the company does not connect to its main competitors *G2 Crowd* and *CB Insights*. These findings remain similar to those previously discussed. More specifically, networks of *G2 Crowd* and *CB Insights* are linked together via the bridging company, *500 startups*. In the new network, *NGData* is not linked with any company, except *seoplus+*; hence, the result enhances the argument that *NGData* is a specialist in this population ecology.
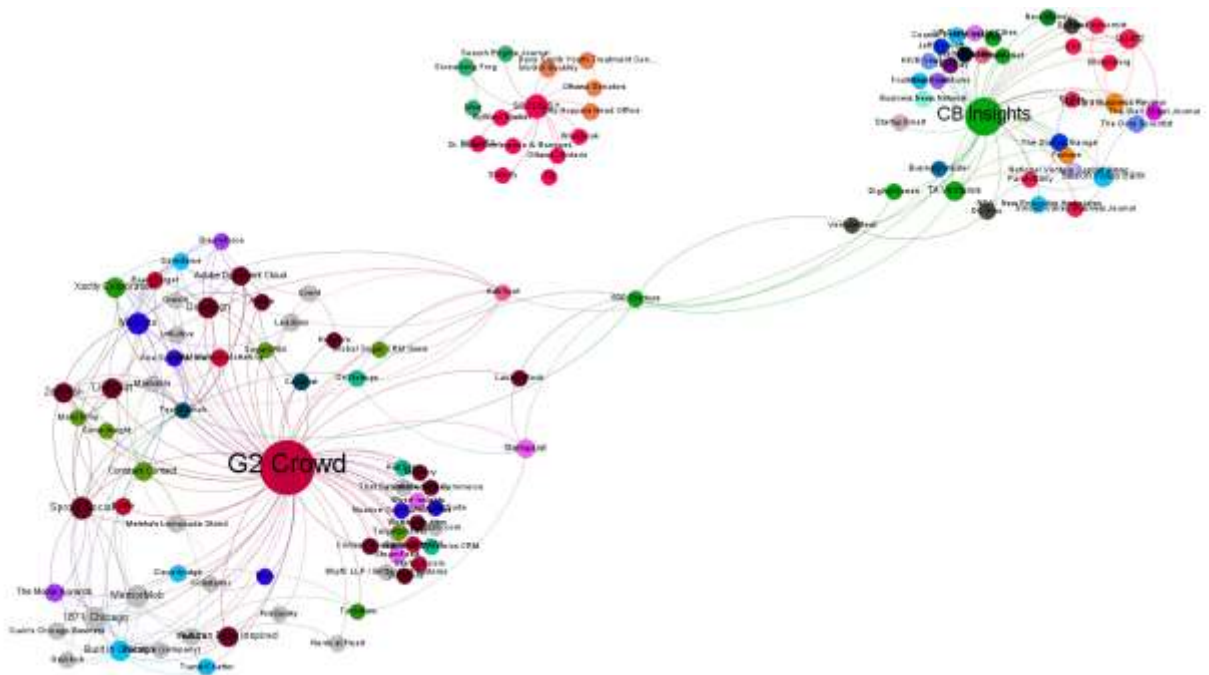
Figure 8: The new network of the platform companies with the 2nd layer of NGData liking network

4.2 Virtual embeddedness reflected through the company-company network via topics

After the network analysis, 14 companies including three core firms (*NGData, CB Insights and G2 Crowd*) and other eleven relevant ones were selected randomly for topic modeling analysis. The company database includes technology-related companies, marketing-related companies, startups and venture capital firms. The main findings will answers Sub-question 2: *How does virtual embeddedness influence activities of platform companies on social media?*

By doing topic modeling, the distinct topics in a corpus are revealed. To explore the themes uncovered in the text corpus, raw data from the recent 7000 post of 14 firms (500 posts

per each company) were scraped by Netvizz. Then the data were parsed in ConText for conducting topic modeling. A total of 98 topics (7 topics per each company) were discovered. Although topic-modeling results are quantitative, the labeling the topics is subjective and requires a qualitative analysis (Diesner et al., 2015). Thus, the researcher interpreted and labeled the topics. Subsequently, the network of companies connecting via topics was created through Gephi for investigating their virtual embeddedness.

4.2.1 Company network via topics

A folded network – which is a company-by-company network, with the edges representing the overlapping topics between companies – was created. The folding is performed through matrix multiplication form of the company-by-topic adjacency matrix and the topic-by-company adjacency matrix derived from the topic modeling outputs (Lee, 2019b). The edge weights indicate the number of topics shared between companies. Then, the interconnectedness between companies via topics was visualized by Gephi (see Figure 9).
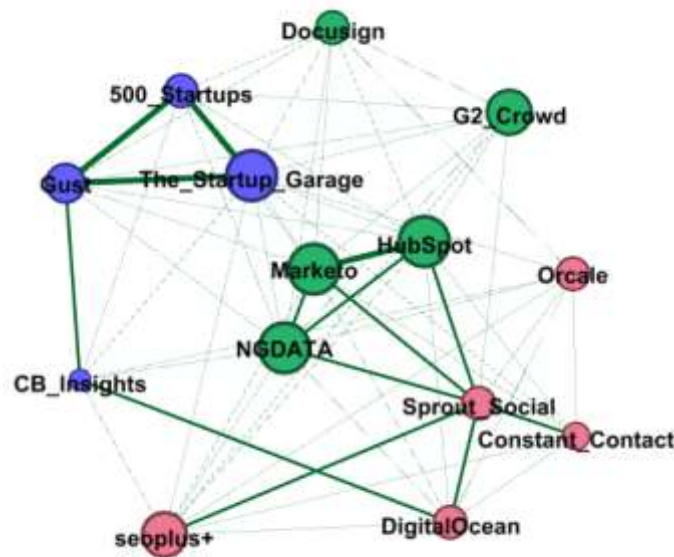
Figure 9: Company network via topics

The researcher obtained two network measures through Gephi, graph density and average clustering coefficient to investigate the overall structure of the network. The network has an extremely high graph density (0.659) and avg. clustering coefficient (0.727), which highlights a dense network indicating that the firms in this network are more interconnected via topics. This

48

is a salient finding as companies are likely to be much more embedded in the network by the overlapping topics.

In this company-company network, all firms of this subsample are connected such that there is only one cluster/component. The edges are sized by their weight. Thicker edges illustrate that companies have many shared topics and vice versa. Companies in the network are linked together by a maximum of three topics and a minimum of one topic. The four thickness edges characterized by the three overlapping topics are *500 Startups-Gust, The Startup Garage-Gust, 500 Startups-The Startup Garage, and HubSpot-Marketo.* In the company-company network, the startup companies have the strongest connections to one another, indicating the formation of a group with the most shared topics. Another outstanding group consists of marketing companies and the other related companies, including *HubSpot, Marketo, NGData and Sprout Social*. In their small group, members share at least two topics and a maximum of three topics.

The nodes are sized by the degree centrality; the smaller the node is, the fewer direct connections (of indirect/topic edges) the company has. While *Hubspot* is the biggest node, indicating the company connects with almost all examined companies, *CB Insights* is the smallest node and has the least number of connections. Another interesting finding is that *NGData* and *seoplus+,* who are separated in the liking network of the population, connect to most other companies in the company network via topics. More specifically, *NGData* has shared topics with 11 of 13 companies and *seoplus+* links with 10 companies via shared topics.

Moreover, the marketing-related companies group is not only a strongly connected group characterized by the thick linkages, but also the group connected with most companies in the overall network. The most outstanding company is *Hubspot,* who connect to 12 of the 13 companies in the network. Similar to *NGData*, the connection of *Marketo* comprises 11 companies. The network itself is partitioned into three distinct groups based on the thickness of the edges and the number of connections. The startup group and the marketing-related group include companies having the thick links and a large number of direct connections. The third group consists of a mix of companies (e.g., startups, software, and marketing-related companies) having many thin edges and is situated in the periphery.

In the company network via topics, there is a big difference between three core companies. While *NGData* and *G2 Crowd* connect to most companies worth noting*, CB Insights* appears to have the least number of connections with six companies and is located in the

periphery of the graph. In addition, although *G2 Crowd* spans the whole network, all its nine links are weak, indicating that its topics are quite generic and relevant to a broad set of companies. Regarding *NGData,* among the 11 links, there are three strong connections with marketing-related companies, implying that the company has concentrated in marketing and related topics. On the other hand, *CB Insights* shares topics only with six companies in the network and most of its edge are weak.

4.2.2 Overlapping topics in the company-company network

In this part, the overlapping topics will be investigated. As all 14 companies connect to each other via at least one topic, the researcher chose overlapping topics that are shared by four companies and above for a deeper examination. In addition, topic modeling shows the weight of each topic, indicating the relative prominence of the topic for each company (Lee, 2019a) (see Appendix B for the list of 14 companies in the subsample and their labeled topics). The tables of the overlapping topics include companies arranged in decreasing order of the topic weight. The document-to-topic output produced by topic modeling also reveals the extent to which each post is related to each topic. The value ranges from 0.0 to 1.0; the higher the number means the greater the fit between topic and post (Lee, 2019a). When analyzing the topics, the most related posts from these statistics were chosen as examples.

Among 98 topics, the most overlapping topic, which is shared by eight companies, is CEO/founder talks.

CEO/founder talks (eight companies)

Table 4: Companies and the topic members of the topic 'CEO/founder talks'

Company	Weight	Topic Members
HubSpot	0.324	hubspot - customer - inbound - shah - dharmesh - representation - set - hire - join - online
Gust	0.209	investor - startup - david - rose - pitch - ceo - share - founder - angel - entrepreneur
TheStartupGarage	0.159	startup - san - diego - garage - capital - entrepreneur - community - jensen - event - founder
Marketo	0.089	marketer - marketing - year - cmo - learn - market - list - steve - leader - marketober
500 Startups	0.086	ceo - founder - company - strong - share - portfolio - investment - feature - christine - tsai
NGData	0.083	ngdata - ceo - burgelman - luc - talk - digital - gdpr - noels - steven - intelligence

| G2 Crowd | 0.070 | ceo - abel - godard - news - intent - buyer - learn - datum - woman - partner |
| Docusign | 0.063 | docusign - ceo - dan - innovation - springer - release - read - announce - news - international |

In the 'CEO/founder talks' topic, the main topic words shared between all companies are 'CEO' and 'founder'. Besides these keywords, the words that have similar meanings such as 'share', 'presentation', 'talk', 'announce' are used among companies when they have publicly presented this topic. Eight companies involved in this topic are startups with three companies and five marketing-related companies. The dominance of marketing-related companies indicates that they have an increased interest in using the reputation of their CEO/founders for branding.

Given that this topic is shared by startups and marketing-related companies, the content of most of their speeches is related to providing advice on investment and marketing. For example, a post with a high topic association from *Gust* illustrates this topic: 'Are video pitches important to #investors for #startups? Gust CEO David S. Rose shares his advice for #entrepreneurs'. Regarding the marketing-related companies, the post 'Listen to the latest episode of the Marketing Today podcast featuring our very own Steve Lucas! Steve discusses the engagement economy the future of marketing and lots more' from *Marketo* is a typical example. The content also revolves around the latest legislation, which is of specific relevance to this industry, as mentioned in an example from *NGData* 'CEO Luc Burgelman discusses about how the May 25th deadline for the GDPR was just the first step in enabling companies to have a trusted dialogue with customers based on a standard framework for data privacy'. In addition, most of the talks have taken place on digital channels, for example on the company blogs, companies' websites, or webinars. Not only on these virtual channels, some of their talks have happened at actual business events. In the following post, the founder of *HubSpot* Dharmesh Shah shared his speech at one of the largest inbound marketing events 'Join HubSpot co-founder Dharmesh Shah for his live keynote at #INBOUND18.'

Generally, the posts related to this topic include a link of external media (i.e., the links to their other online channels such as their websites, blogs) for audience to learn more about the speeches. Besides this example, webinars (live, web-based video conferences) and live streams are their preferred method for information sharing, demonstrating their attempt to provide the information in real time fashion (Icha & Edwin, 2016) and focusing on building online customer

engagement (Parveen, Jaafar, & Ainin, 2015). In addition to this, the topic stresses the purpose of using social media among firms for branding by using the voice of their own CEOs/founders and sharing information and advice (Gloor et al., 2016, Icha & Edwin, 2016). Although the CEO/founder talks is the most shared topic, the extent of its prominence is different between companies. The topic seems to be more important to *Hubspot, Gust* and *The Startup Garage* than the remaining companies.

Marketing strategy (five companies)

Table 5: Companies and the topic members of the topic 'Marketing strategy'

Company	Weight	Topic Members
HubSpot	1.187	business - marketing - company - make - work - change - year - facebook - people - day
seoplus+	0.168	marketing - digital - content - ottawa - media - social - specialist - top - year - business
Sprout Social	0.142	marketing - social - strategy - brand - customer - pinterest - experience - top - business - put
NGData	0.119	marketing - video - strategy - marketer - effective - round - single - acquisition - important - campaign
Marketo	0.097	marketing - sale - learn - content - lead - team - marketo - strategy - guide - secret

This topic comprises companies providing products (e.g., software, platforms) and services in the marketing sector. Most of the organizations have concentrated on giving tips (or advice) on digital marketing and the ways to leverage social media to build a company's brand (Parveen, Jaafar, & Ainin, 2015), given that the topic contains the words such as 'Facebook', 'Pinterest', 'digital', 'content' 'social', and 'media'. The following post is from *Sprout Social*, a provider of social media management software and solutions, 'Is your social media presence strong enough to do a pull-up? Use these examples to strengthen your digital marketing strategy and build a community around your brand'. One of the top posts from the core company *NGData* also illustrates the topic 'Viral video is a lofty target and, while it shouldn't be the ultimate goal of your video marketing campaign, it is a very nice icing on the cake of great video that fulfills all your marketing objectives'.

It is worth noting that 'Marketing strategy' topic appears as the most prominent/weighted topic for the companies in the table. The high weight indicates that providing informative content

about marketing strategy is a part of the communication strategy of these platform companies on social media (Icha & Edwin, 2016; Parveen, Jaafar, & Ainin, 2015). Among them, *HubSpot* is impressive as its topic weight (1.187) is much higher than that of other companies. The topic also implies that digital and online marketing is the core to the industry, not only platform companies, as a post from *Marketo* emphasizes 'According to eMarketer, 84% of US companies will utilize digital content marketing this year [2018]. Discover why so many B2B companies rely on content marketing'.

Product introduction (five companies)

Table 6: Companies and the topic members of the topic 'Product introduction'

Company	Weight	Topic Members
Oracle	0.439	oracle - cloud - learn - data - business - autonomous - discover - join - customer - webcast -
NGData	0.195	customer - experience - video - business - journey - deliver - analytic - lily - company - brand -
CB Insights	0.084	market - today - product - insight - research - start - day - data - team - trial -
Sprout Social	0.082	sprout - social - team - data - marketer - work - index - meet - learn - attention -
DigitalOcean	0.080	droplet - plan - month - celebrate - dohq - black - happy - cpu - digitalocean - nyc

The companies in this group are marketing-related companies (*NGData, Sprout Social*) and software companies (*Oracle, Digital Ocean, CB Insights*). *CB Insights* is the only company in this group that its topic member contains the words 'product' and 'trial', making the topic is more easily labeled. The following post illustrates the topic 'Making a market map used to take days even weeks plus the blood sweat and tears of far too many analysts. With CB Insights it only takes one click. Start your 30-day free trial to build your first market map today'. In contrast, the representative topic members of other companies do not contain the direct word 'product'. Therefore, the researcher interpreted and labeled this topic by examining deeply their posts. In addition, 'Product introduction' is the most prominent topic of *Oracle* with the highest weight (0.439) compared to other companies, indicating their focus on using social media for promoting products (Parveen, Jaafar, & Ainin, 2015). Their prominent product (Oracle database) appears in their following post, 'Oracle's Autonomous Database raised a lot of questions among

database pros and business leaders. Exactly how does it improve security? Here are some of the questions we've tackled this past year [2018], with answers gleaned from articles on our blogs and other publications.' This post includes an additional link to their blogs, in which clearly explains the features of their product. Security, the biggest concern of companies in the platform market is also emphasized (World Economic Forum, 2017).

Other companies have also mentioned their products on their posts. For example, *Sprout Social* introduced Sprout Social Index by posting 'Join us for a deep dive into our 2018 Sprout Social Index Data!' *Digital Ocean* presented their new products and promotions in the post 'Introducing our latest CPU Optimized Droplet plan - an entry-level 1 vCPU for $20'. Lily tool, a product from *NGData* is also highlighted in its post 'Are you looking for a new and results-driven way to onboard customers? Lily SmartVideos are the perfect tool to drive engagement and they generate real business value with your very own customer data. https://www.ngdata.com/lily-smartvideo-call-reduction-onboarding/'.

Internal activity (four companies)

Table 7: Companies and the topic members of the topic 'Internal activity'

Company	Weight	Topic Members
The Startup Garage	0.118	startupsavvy - startuplife - tsg - sandiego - getfunded - thestartupgarage - events - day entrepreneurs - great
seoplus+	0.137	happy - day - seoplus - team - director - today - office - great - month - friend
Constant Contact	0.070	day - comment - giveaway - winner - win - december - purchase - visit - prize - unique
Oracle	0.054	sailgp - team - francisco - san - app - australia - race - oracle - day - sydney

The companies have used social media to not only promote their products but also to reflect their corporate culture as well as enhance networking (Gloor et al., 2016). The companies presenting this topic comprise startup, marketing and software companies. The words in this topic such as 'day', 'team', 'life', 'director' and 'office' infer the workplace. The audiences are attracted by the positive words, for instance 'happy', 'unique', 'prize' and 'winner', implying a healthy workplace environment at these companies. Businesses have shown that they pay attention to team building activities to improve teamwork at the office. For example, employees at *The Startup Garage* enjoyed the time after working by watching a sporting event together.

They posted on their Facebook page: 'Fun night at the San Diego Gulls game with The Startup Garage Team! #Winning'.

The team-building activities have taken place regularly, like monthly creative workshops at *seoplus+* to increase satisfaction and innovation among employees. The topic is shown in the post 'Another great CreativeMornings Ottawa shows this morning. This month talk was on Compassion by Alexander Shelley, Music Director of the National Arts Centre Orchestra! Our seoplus+ team loves attending these monthly talks for creative'. While other companies have organized events for their local teams, *Oracle* reported the activities of their international teams at the world sailing event on their posts. The internal activities will increase brand images and attract potential qualified employees, which contributes to developing human capital (Vasileiadou & MisslerBehr, 2010).

Funding (four companies)

Table 8: Companies and the topic members of the topic 'Funding'

Company	Weight	Topic Members
The Startup Garage	0.151	investor - capital - people - raise - founder - relationship - money - deal - year - cover
Gust	0.112	startup - entrepreneur - zwilling - focus - note - martin - angels - funding - market - convertible
CB Insights	0.120	report - download - funding - deal - complete - global - publish - increase - moneytree - pdf
500 Startups	0.075	million - raise - investor - round - sanfrancisco - ventures - funding - ticket - conference - fund

As all corporations involved in this topic are startup companies or related, 'Funding' is their most weighted topic. The topic is constituted of the terms related to funding such as 'investor', 'capital', 'raise', 'money' and 'ventures'. Each company has expressed the topic in different perspectives, including funding tips, funding events and funding reports. The post associated with giving funding advice is shown in the following post from *The Startup Garage* 'These days one Investor may not be enough. You probably raised just enough to reach your next milestone but not enough to achieve self-sustaining profitability...Our consulting services can help you stay on track to continually reach the milestones that investors care about.' In addition, the funding activity is presented through investor events, which is illustrated by the post of *500 Startup* 'Are you as excited as we are for PreMoney Investor Conference 2018 on October 2nd?

Join us to hear Ben Einstein of Bolt give a tear down on his hardware investing expertise. Get your tickets today…' As for *CB Insights*, they produced a report on the trends in funding 'We just published a new report: Venture Capital Funding Report 2018. Download the complete PDF here…' These posts demonstrate social media as a digital channel for companies to offer their services, which can be free or with relevant costs, such as tickets to various events (Parveen, Jaafar, & Ainin, 2015).

In general, the overlaps reveal a commonality between the companies' communication strategies as well as the characteristics of the overall market (Goktepe, 2018). In addition, the overlapping topics illustrate not only the influence of structural embeddedness as argued by Granovetter (1992), but also the pressure of isomorphism on organizational behaviors in the platform market (DiMaggio & Powell, 1983; Findik & Bedük, 2014). More specifically, the companies in the subsample focus on activities of sharing and giving advice to meet the expectations from others companies upon which they are dependent (DiMaggio & Powell, 1983). The reason likely is that information and advice seeking is the most concern of their customers and stakeholders (Gloor et al., 2016). Moreover, these companies have paid attention to building and promoting their company brand on social media. This attention is expressed through the speech or sharing of CEOs/founders as well as the internal employee-focused activities. Additionally, the finding reveals the startup group and the marketing group are the strongly embedded groups due to a high number of overlapping topics.

4.2.3 Non-overlapping topics of three core companies

Overall, all companies in the company-company network via topics are linked with one another, indicating that they are embedded in the network by the overlapping topics. However, among seven topics shared between the companies, some topics lack alignment across the three focal companies. Indeed, although *NGData, G2 Crowd* and *CB Insights* have shared topics with most companies in the company-company network, there is just one overlapping topic ('CEO/founder talks') between *NGData* and *G2 Crowd*, and one ('Product introduction') between *CB Insights* and *NGData.* Miss alignment of topics is possibly translated to the manner to which these companies are embedded, which could be due to the diversity in resources and stakeholders. This part will examine the topics of each core firm, which do not overlap with

those of any companies in the subsample. The findings uncover different ways of embedding between the three core companies and their communication strategies on social media.

NGData

Table 9: NGData's top seven topics

Topic	Weight	Topic Members
Customer experience webinar	0.203	customer - marketing - free - webinar - experience - learn - make - datum - guide - brand -
Product introduction	0.195	customer - experience - video - business - journey - deliver - analytic - lily - company - brand -
CDP market	0.130	datum - company - customer - data - cdp - platform - technology - ngdata - machine - intelligence -
Marketing strategy	0.119	marketing - video - strategy - marketer - effective - round - single - acquisition - important - campaign -
Bank industry	0.096	ngdata - report - service - market - banking - world - bank - industry - analytics - financial -
CEO/founder talks	0.083	ngdata - ceo - burgelman - luc - talk - digital - gdpr - noels - steven - intelligence -
Data management	0.079	datum - company - machine - lake - business - tool - process - leverage - analytic - intelligence -

In the topic modeling of *NGData*, there are four topics non-overlapping with other companies in the company-company networks, namely 'Customer experience webinar', Business report', 'Bank industry' and 'Data analysis'. Among them, 'Customer experience webinar' is the best-fit theme (0.203), which contains the following words: customer - marketing - free - webinar - experience - learn - make - datum - guide – brand. It is easy to interpret the topic as most of the terms are related to 'marketing' and 'webinar'. The post with the highest fit of the topic (0.93) illustrates the product niche of the company 'Looking to improve your customer experience? Learn how to make every micro-moment count to achieve true customer centricity. Save your seat for Tuesday s webinar to find out how!' The company offers customer data platform (CDP) for marketing solutions to improve customer experience, which distinguishes itself from its rivals (Enchols & Tai, 2005). The way of promoting their products and giving advice is similar to other companies in the subsample, by organizing free webinar on social media sites.

'CDP market' is the second non-overlapping topic with the representative words: datum - company - customer - data - cdp - platform - technology - ngdata - machine - intelligence -. The topic demonstrates the niche status of *NGData*, a specialist in the CDP market, as most of these

terms are related to this market. The topic can be best illustrated by the post 'Get your Customer Data Platform Industry in Europe January 2019 report from the CDP Institute here'. In addition, the member words "machine", 'technology' and 'intelligence' reflects the characteristics of the CDP market, which is based on artificial intelligence/machine learning to collect customer data and deliver customer experiences (Customer Data Platform Institute, 2019). One of the best-fit posts illustrates these words is 'Business intelligence analytics is more accessible than ever before and customer-centric enterprises optimizing their customer interactions through more advanced BI systems have a critical advantage over competitors not yet taking full advantage of this technology.'

Another non-aligning topic is 'Bank industry' with the terms: ngdata - report - service - market - banking - world - bank - industry - analytics – financial. The words describing the topic revolve around bank industry. This topic reveals the target customers and potential markets that the company has tried to approach. The best-fit post (0.98) that represents this topic is 'Consumers have increasingly high expectations and demand exceptional customer service in exchange for their loyalty according to the latest consumer banking survey conducted by NGDATA'. The post also illustrates the activities of the company. By conducting their own research, *NGData* discovers the trends in the market, and thus producing their own innovative products. Another post shows how their products meet the demand of new customers 'In this video we showcase what the customer journey powered by Lily looks like in the retail banking industry'.

The topic 'Data management' has the lowest prominence (0.079) and is an unaligned topic, including the terms: datum - company - machine - lake - business - tool - process - leverage - analytic - intelligence -. The topic also illustrates the characteristics of the CDP market, which emphasizes the role of customer data management for companies in gaining competitive advantages (Customer Data Platform Institute, 2019). The following post contains most of the relevant terms 'Take a look at the benefits and challenges of using data lakes to store large datasets and how businesses can intelligently utilize them.' The post shows that sharing advice on data management and data analysis for business process is the main activity of *NGData*.

In general, the non-overlapping topics of *NGData* reveal its product niche (Enchols & Tai, 2005), as the company is a specialist on the CDP market. This niche finding is in line with

the liking network. The topics show not only their business activities but also their diversity in resources and stakeholders, which can be seen in the topics related to marketing and banking. Besides six main industries that B2B platforms can prosper through the next five years, which has been reported by the World Economic Forum (2017), the topics indicate that banking is another market that the B2B platform companies are considering for developing their business.

G2 Crowd

Table 10: G2 Crowd's top seven topics

Topic	Weight	Topic Members
Community service	0.206	review - crowd - day - work - giveback - give - software - donate - proud - leave -
Marketing event	0.192	crowd - company - marketing - review - miss - don - booth - user - check - love -
Social media usage	0.181	facebook - crowd - tool - business - blog - marketing - customer - start - video - product -
Review blog	0.138	crowd - research - blog - review - explore - datum - technology - smallbusiness - series - life -
Networking	0.132	chicago - crowd - office - join - thecrowd - today - marketing - brand - work - cmo -
Best Software	0.090	software - list - company - crowd - top - make - congratulation - recognize - teams - slack -
CEO/founder talks	0.070	ceo - abel - godard - news - intent - buyer - learn - datum - woman - partner -

Among the non-overlapping topics of *G2 Crowd*, 'Community service' is the most prominent one with the highest weight (0.206), implying that community activities is one of their largest concerns. They have presented their contributions to the community withpride 'Over the last few years G2 Crowd has focused on growing our ability to #GiveBack through our with #reviews. As our latest efforts to help the tech and business communities, we are proud to announce G2 Crowd's Entrepreneurial Scholarship program'. Through the community activities, the company can promote their product and activate customers engagement by calling on donations as illustrated in the following post 'In order to help the people and communities affected by the California wildfires, G2 Gives is immediately shifting all donations generated by our community. For every review you leave on G2 Crowd we'll donate $10 towards Habitat for Humanity s Wildfire Disaster Recovery Fund'. Hence, in addition to the speeches from the

experts like other companies, 'Community services' is distinct in enhancing the corporate image towards the public.

The topics 'Review blog', 'Networking' and 'Best software' are not overlapped with any companies, illustrating their business activities. As *G2 Crowd* provides a software review platform, the company's topics on social media highlight its products and services (Parveen, Jaafar, & Ainin, 2015). 'Review blog' (0.138) contains the words: crowd - research - blog - review - explore - datum - technology - smallbusiness - series - life – which emphasizes the role of blogs in G2 Crowd's professional activities. The review blog is the place organizations, small business in particular, can look for business advice and reviews on software and social media applications. Additionally, the 'Review blog' also highlights the trend and innovative technologies in the industry. For example, some of the top posts show that 'Are you ready for a new OASIS? The G2 Crowd Blog explores the similarities between #FacebookSpaces and the world of Ready Player One'.

The non-overlapping topic 'Networking' (0.132) contains the terms: chicago - crowd - office - join - thecrowd - today - marketing - brand - work - cmo -. It is interesting that the word 'chicago' is the first representative topic member. The word does not only emphasize the city of their headquarters but also indicates the place activities are regularly organized. *G2 Crowd* has organized events for building up the network within the tech community of the city. The post containing most of the topic members is 'Attention #Chicago #Tech community: Join #TheCrowd this Thursday afternoon at our offices for drinks #networking and a few special announcements from G2 Crowd! Hear more from our on-the-go CMO Ryan Bonnici in the video below and RSVP here to attend: https://g2.co/2zZnmil #event #specialevent'. Hence, the topic underlines the usage of social media for networking opportunities (Icha & Edwin, 2016).

The topic members: software - list - company - crowd - top - make - congratulation - recognize - teams - slack – belongs to the 'Best software' topic. The post with the highest fit (0.95) of this topic is 'Announcing G2 s Best Software Awards 2019. We analyzed 269 987 new reviews from 2018 to create the definitive ranking of the best software companies and products worldwide. Ready to reach your full potential in 2019? Find the full list at https://g2.com/best'. Moreover, it is interesting that in the post about the 2018 report, most of the software companies that *G2 Crowd* have connected to on Facebook were named 'Congrats to the winners on amazing #b2breviews especially our Top 10 stars: Slack, Shopify, Google, MailChimp, Salesforce,

WordPress, SurveyMonkey, Adobe, Zoom Video Communications, HubSpot. Do not miss your chance to congratulate this year's winners'. The post demonstrates the companies – to which G2 Crowd have connected on Facebook – are their actual resources; and they have a diverse set of resources (e.g., reviewers and software companies). This finding is similar to the one revealed in its liking network.

In general, the misalignment on topics of *G2 Crowd* expresses the company's interests in community and networking activities, especially to the local communities in the Chicago city area, which makes it distinct to the other companies. Recall back to its liking network, the community group is made up of 6% of its overall connections and three Chicago-based community pages are the 'powerful' nodes and span their small groups. Although the community groups exert less direct impact on organizations (Castro, Sáez and López, 2004), *G2 Crowd* values these relationships, given that they contribute to relational capital and strengthen the reputation of the company in the market (Wu, Wang, Chen, and Pan, 2008).

'Best software' report and 'Review blog' are other non-overlapping topics, which reflect their products and services (i.e., review software platform). These topics are general and do not show their focus on a specific market or industry. Thus, these reports can be informative and seen as referential sources for companies in any kind of industry. Thus, it is possible to infer that *G2 Crowd* is a generalist. Moreover, the topic misalignment emphasizes the diversity in resources and stakeholders (e.g., customers, community, reviewers and software affiliates). An interesting finding is that *G2 Crowd* links with most of their reported software companies on Facebook. This finding highlights social media as a tool for the company to integrate into the online tech community (Gloor et al., 2016).

CB Insights

Table 11: CB Insights's top seven topics

Topic	Weight	Topic Members
Healthcare market	0.172	trend - report - industry - healthcare - watch - wellness - biggest - top - trends - technology -
Tech trend	0.124	client - note - startups - fintech - future - blockchain - don - register - briefing - unicorns -
Funding	0.120	report - download - funding - deal - complete - global - publish - increase - moneytree - pdf -

Startup report	0.114	startup - company - tech - unicorn - year - list - investor - check - full - valuation -
Consumer advice	0.099	consumer - company - business - startup - digital - map - brand - impact - space - break -
Product introduction	0.084	market - today - product - insight - research - start - day - data - team - trial -
Tech companies	0.070	google - amazon - apple - market - world - acquire - smart - data - back - uber -

The topic 'Healthcare market' is the most prominent topic (0.172) and not overlapping with the other topics. Most of the topic members are related to the healthcare industry such as 'healthcare', 'industry', and wellness', making the topic easily identified. The post with the highest fit (0.9) of this topic is '#Healthcare is emerging as a prominent area for #AI research and applications and nearly every area across the industry will be impacted by the technology rise. We break down the top trends to keep an eye on: https://cbi.vc/2O5wGcp'. The topic reveals the emergence of healthcare in the platform market, which is highlighted in the report from the World Economic Forum (2017). A similar market, the wellness economy is also indicated as a large incoming trend. This trend is illustrated in the following post 'What's next for #wellness? We examined 18 of the biggest wellness trends to watch in 2019 covering everything from wellness #tourism to boutique #meditation. See what we found in our complete report: https://cbi.vc/2NOApsA'.

The second prominent topic 'Tech trend' (0,124) is also non-aligning, including the topic words: client - note - startups - fintech - future - blockchain - don - register - briefing - unicorns. The post is illustrated by the following posts 'Client Note: Stash's new Rewards Program - African Fintech High Momentum Startups - A Beauty Unicorn - And More' or 'Client Note: Future Supply Chain Unicorns - Uber's Autonomous Driving Business - Med Device Security - And More'. Actually, client note is a column on their webpage, providing brief reports about the new technology trends and innovation in the industry. Given that *CB Insights* is a provider of software which predicts emerging tech trends and new markets, this topic demonstrates the characteristics of the company's products and services (Parveen, Jaafar, & Ainin, 2015). In addition, the topic focusing on financing/funding with the words 'fintech' (financial technology) 'startups' and 'unicorn', indicates their clients including banks, venture capitals and startups.

The non-overlapping topic 'Consumer advice' (0,099) has the representative terms: consumer - company - business - startup - digital - map - brand - impact - space - break -. The

best-fit post (0.94) representing this topic is 'The #apocalypse is nigh: amid major bankruptcies and malls that look like ghost towns brick-and-mortar #retail is feeling the sting of unsustainable business models and new consumer preferences. We mapped out 57 major bankruptcies and why they failed'. This post is involved in sharing advice on changes in consumer tastes. That is, companies should adopt new consumer preferences in order to survive. Another example is 'On the consumer side Square is looking to challenge banking incumbents by transforming its Cash App into a digital #bank. In this report we break down how Square is leveraging its 7M active users and $200M+ in user account balances'. In this post, the company discusses how a new startup is successful by satisfying consumer needs. Generally, 'Consumer advice' demonstrates that *CB Insights* has employed social media for giving business advice through both failed and successful cases based on consumer preferences.

The last non-overlapping topic is 'Tech companies'. This is also the least weighted topic with the value of 0.070. The topic members contain the names of the biggest technology companies in the world, such as Google, Amazon, Apple and Uber, making this topic easily recognized. The following post includes most of the key words 'As #2019 gets underway take a look back at the strategic initiatives and forward-looking moves of some of the world's top companies. Download the Best of #Strategy bundle today for our analyses of Amazon Apple Google JPMorgan Uber and WeWork: https://cbi.vc/2Rlw30f'. This post mentions the competitive differences between the big technology companies, focusing on their strategies and their technology innovations. Thus, the topic highlights the use of social media for sharing information (i.e., business strategies) of *CB Insights* (Parveen, Jaafar, & Ainin, 2015).

In general, the non-overlapping topics illustrate the characteristics of the company's services, which are platforms for reporting technological change and predicting trends. *CB Insights* products are introduced directly via the topic 'Tech trend' and 'Tech companies', or indirectly by presenting the trend in the platform market 'Healthcare market' (World Economic Forum, 2017). In addition, the misalignment in topics highlights the interest of the company in using social media for giving advice and sharing information, which is similar to other companies in the subsample.

5. Conclusion

This chapter provides a further discussion of the results and the explanations of whether the expectations raised in the previous chapter are met. In addition, interesting findings are mentioned. Then, the theoretical and managerial implications as well as the limitations of the research are discussed. The thesis ends by suggesting new directions for future study.

This thesis aims at examining the virtual embeddedness of platform companies on social media by answering the research question: *How do social media reflect virtual embeddedness of platform companies in their population ecology?* In order to answer the research questions, the researcher conducted an analysis of the Facebook liking networks and the company-company network via topics of three platform companies, namely *NGData, CB Insights,* and *G2 Crowd*. In general, the influence of virtual embeddedness on the performance of platform companies on social media is reflected more obviously in the company-company network (via topics). Although some companies do not connect to others in their liking Facebook networks, they are much more embedded in the overall population through the topics of their posts. The expectation that the companies become similar to each other via online content (i.e. topics) due to embeddedness and isomorphism pressure happens.

5.1 Main findings

5.1.1 Virtual embeddedness reflected through the Facebook liking network

At first, the manner of embedding of each of the core companies was investigated. The liking network illustrates the diversity of resource that companies have taken advantage of to achieve their business goals (Zhang et al., 2018). When it comes to the B2B platform companies, the survival of these platforms is contingent on the integration resources from the owners, partners, end users and stakeholders (Blascheke et al., 2018). The liking networks of three core companies reveal the resources that are needed. As for *CB Insights,* they are news and media companies, startups, and venture capitalist companies. Marketing companies are one of the resources of *NGData*; likewise, reviewers and software companies are an important resource of *G2 Crowd*. The pages that they connect to on Facebook include all types of these resources. In addition, their networks are partitioned into two groups. The first group includes organizations (called pendants) that connect to the core companies; however, they do not connect with any other companies in the network. The second group is made of the connected organizations.

It is interesting that two of the focal companies, *CB Insights* and *G2 Crowd* have the same manner of embedding, which is partly-partitioned and modestly-resource diverse. Moreover, the pattern of their structural embeddedness aligns with the one described in previous studies. That is, the finding shows that the overall structure of the network in which they are embedded have the same characteristics. More specifically, these networks are sparse and fragmented, indicating the organizations inside are less connected (Echols & Tsai, 2005; De Meo et al., 2012). In addition, they have a low density and a low clustering coefficient (Gloor et al., 2016). In contrast, *NGData* has such a small online network with two nodes. Perhaps, Facebook is not the main channel that this company has exploited to be present on social media. In briefly investigating other SNSs such as Twitter, the researcher finds that the company is much more active on this platform (see Table 12). One can argue that the similarities in the network structure and the way of embedding of *CB Insights* and *G2 Crowd* are a result of pressure from the observed isomorphism (DiMaggio & Powell, 1983), while *NGData* probably tries to make itself distinct in order to resist isomorphism.

Table 12: Following and followers on Twitter accounts of the three focal companies

Company	Following	Followers
NGData (twitter.com/ngdata_com)	779	2168
G2 Crowd (twitter.com/G2CrowdReviews)	256	2249
CB Insights (twitter.com/CBinsights)	16	83.200

The population network established subsequently reveals the niche status of these companies. Because of their focus on a broad market plus their central actors in the network (MacMillan & Komar, 2018), *CB Insights* and *G2 Crowd* appear to be generalists. In contrast, *NGData* – a customer data platform company – is the least active actor and isolated out in the periphery, indicating it as a specialist. The resource partitioning is uncovered in the population network. The two generalists are separated into two hubs and have exploited their own resources, which are represented as their neighboring or affiliated companies.

5.1.2 Virtual embeddedness reflected through the company-company network via topics

According to the previous studies, embeddedness shapes and affects behavior of organizations (Granovetter, 1992; Gnyawali & Madhavan, 2001; Uzzi, 1996). In order to investigate the influence of embeddedness on organizational activities, the researcher conducted topic modeling of 14 companies, who were selected randomly from the liking networks of the three core companies. From the topic modeling outputs, the company-company network (via topics) was established for examining the embeddedness pressure. The result from the company network is rather contradictory to the one from the liking network. More specifically, in the company-company network, the result reveals that this is a dense network as all companies in the subsample share their topics with each other. The most outstanding cases are *NGData* and *seoplus+*. Perhaps, in other to avoid competing with its rivals on social media and to enhance organizational legitimacy towards the population, these two companies give in to isomorphism by overlapping certain topics (DiMaggio and Powell, 1983).

In addition, companies in small or particular groups tend to connect to one another strongly in both types of networks to enhance their survivability, for example the startup group. In the liking network of the *G2 Crowd,* the startup group is presented as a tightly-linked group since all companies link with one another. Likewise, in the company-company network via topics, startups including *500 Startups, Gust* and *The Startup Garage* form a group notable for having the thickest edges, showing they have many overlapping topics. Another distinctive group is the marketing-related group consisting of *HubSpot, Marketo, NGData* and *Sprout Social*. Although the liking networks do not reveal clearly that they are strongly connected, their community seen in the company topic network shows strong edges with a high number of shared topics.

Another important finding of the thesis is that the overlapping topics not only show the manner of embedding but also the behavioral patterns which organizations are inclined to adopt (Granovetter, 1992). More specifically, the companies have utilized social media for accessing advice and sharing information (Kuhn et al., 2016; Parveen, Jaafar, & Ainin, 2015; Icha & Edwin, 2016). Among these purposes, the overlapping topics emphasize giving advice as the most popular activity between platform companies. In addition, these companies also focus on promoting their corporate images through the topics of 'CEOs/founders talks' as well as internal employee-focused activities.

On the one hand, the embeddedness reflected via the overlapping topics reveals the similarities in behaviors between organizations. On the other hand, the non-overlapping topics show their niche status, as well as their distinct communication strategies on social media. As for *NGData*, the misalignment in topics indicates it as a specialist in the CDP market (Enchols & Tai, 2005). *G2 Crowd* is a distinct generalist and has paid much more attention to networking in their local (both online and offline) communities (Gloor et al., 2016). The focus on external activities such as community services rather than internal activities in order to build their image make the company different in relation to the others. Regarding *CB Insights,* the non-overlapping topics are mostly used to promote the characteristics of its products and services*,* which are platforms for predicting technology trends. In addition, the misalignment in topics also reveals the characteristics and trends of the B2B platforms market. Besides key emerging industries reported by the World Economic Forum (2017), the topic 'Bank industry' from *NGData* indicates another sector that the B2B platform industry are considering in order to develop their business.

5.2 Theoretical and managerial implications

The study provides several theoretical implications. The main theoretical implication of the thesis is that it demonstrates the relevance of population ecology and virtual embeddedness in explaining inter-firm relationships between organizations in emerging markets. This study examined the different concepts from new perspectives. More specifically, most previous researches have focused on the inter-organizational relationships of new ventures on their websites in order to examine the virtual embeddedness (Morse, Fowler, and Lawrence, 2007; Fowler, Morse, and Lawrence, 2004). The study investigates the inter-firm networks on social media via the liking networks and the behavior of platform organizations via their topics they present. As a result, the thesis extends the available academic literature of population ecology and virtual embeddedness into the era of social media.

The pressure of virtual embeddedness and isomorphism forces companies in the same market resemble one another (Capioto et al., 2019, DiMaggio & Powell, 1983; Findik & Bedük, 2014; Granovetter, 1992). However, by comparing topics (i.e., their behavior seen through their posts) and examining the two types of network (i.e., the company-company network via topics and the liking network), the study shows that companies respond to these environmental

pressures by either conforming to or resisting them. Perhaps this is a new kind of strategy of companies to differentiate themselves through the unique and non-overlapping topics. Therefore, the findings from this study can be applied in other relevant academic fields such as management theories.

In addition, the roles of gatekeepers/mediators in enhancing the survivability of specialists are revealed. More specifically, *500 Startups* and *HubSpot are* the gatekeepers in the population network. Meanwhile, *Gust* is the pendant specialist that survived in the CB Insights's network and has a strong connection with *500 Startups* in the company network via topics. Therefore, it can be inferred that being connected with the mediators could increase the survivability of this company. Due to their strong connections enough to ensure their survivability, the startup group could be a niche in this population (Podolny, Stuart, and Hannan, 1996). Moreover, the gatekeeper positions can increase the influence of community organizations, which have a less and indirect impact on organizations' activities (Castro, Sáez and López, 2004). The case of *1871 Chicago* in G2 Crowd's network is a typical example. The influential role of this communication organization is not only expressed in the liking network but also in the topics of *G2 Crowd*. These findings pave the way for developing studies on the role of gatekeepers inside a network and the benefits gleaned from relationships with them.

Additionally, this thesis highlights the influence of social media in understanding organizational performance, which is often overlooked (Schultz, Schwepker Jr, and Good, 2012). The role of social media in academic research is promoted as they provide the ability to analyze the impact of the virtual embeddedness of companies. In addition, the data collected from social media allow the researcher to identify the characteristics of a specific market, for example the platform market as well as competitive communication strategies of companies. The examination of companies' online content of companies on social media can also discover new trends of this market.

When it comes to the managerial implications, the findings of this study raise awareness among managers to the pressures from the inter-organization relationships that they can encounter (Capioto et al., 2019, DiMaggio & Powell, 1983). With regard to the new businesses, they can adopt the behavioral model or copy the execution of successful companies in the same field in order to survive (Katopol, 2016). Then, they will be capable to produce their own communication strategy. Especially, to the new tech-based companies in the platform market,

they should focus on giving advice on their social networking sites since the findings show that this is valued in this market. Meanwhile, companies can resist isomorphism by making themselves different with distinct online contents. In addition, social media is no longer a pure marketing tool. As mentioned above, it is a powerful tool in seeking advice and networking to the online community, as well as improving their relational capital (Goktepe, 2018; Vasileiadou and Missler-Behr, 2011). Thus, companies should utilize social media to enhance their business.

5.3 Limitations of the study

In the study, the social networking site Facebook was chosen for investing the embeddedness of companies in the population because the links created have some importance and are not made randomly or without reason (Brennan & Croft, 2012; Go & You, 2016; Goktepe, 2018). However, the level of resource dependence cannot be inferred through this digital data. More specifically, the researcher cannot identify whether the company they are linked with on Facebook is actually their customer or the extent to which they affiliate. In addition, the liking network on Facebook lacks of information about the strength of these ties (i.e., weak or strong) (De Meo et al. 2014), as well as the reason for the survival and success of a company (e.g., *Gust* and *Idea Market*). The technological limitations also partly effect Facebook chosen over other social networking sites. For example, it is not easy to collect the data of Twitter following or follower unless the researcher knows computer programming.

In addition, the researcher did not limit the time when scraping posts of 14 companies in the subsample, as long as they are 500 recent posts. Therefore, the collected posts of each company are different in terms of their timeline. For example, while the last posts of *NGData* were from Jun 2016, those of *CB Insights* were from Sep 2018. As a result, the researcher cannot obtain the communication strategy of each company on social media via their topics in a longer term. Moreover, as the online content is constantly changing, the changes in their strategies in response to the pressure of isomorphism and embeddedness are not revealed.

Another limitation of this study could be bias in data selection from two perspectives. First, the core companies chosen for analyzing virtual embeddedess were selected from the competitor database of *NGData*, which illustrates its own point of views and its own goals. Therefore, the results do not necessarily fully reflect the CDP market, where it operates. Second,

due to the large networks of the two focal companies, the researcher chose randomly the sample for topic modeling, which could limit the complete picture that could be made.

5.4 Suggestions for future research

Given the limitations, this paper gives the following recommendations for future research in order to add clarity to the field of this study.

Besides Facebook, LinkedIn and Twitter are other popular social networking sites that can be used to examine the network ties between companies. Especially, Twitter offers researchers the ability to gather information pertaining to following and followers of its users, indicating the relationships between corporates and their customers as well as stakeholders (Zhang, Jansen, & Chowdhury, 2011). In addition, Twitter also reflects the performance of corporations as it was reported that 84% tweets from corporate Twitter accounts are about company news (Funk, 2011). Therefore, Twitter could be a potential source for future research. The combination and comparison between those social networking platforms could provide a more comprehensive portrayal of the virtual embeddedness of companies.

As there are biases in the selection of companies for this research, the future study could focus on a specific sector of B2B platform market, for example the CDP market, which is reported as a dynamic and growing sector (Kenney & Zysman, 2016). The list of companies in this sector is reported and updated regularly by professional organizations, such as the CDP Institute. As a result, the competitor database can be made more credible, resulting in the chosen companies' being more relevant to the market.

In order to enhance our understanding on the effects of embeddedness and isomorphism on organizational activities and their response, the researcher suggests conducting topic modeling of all the companies inside the affiliation network. Potential studies could take advantage of other more advanced digital tools for creating a full company-via-topic network. The topics themselves will not be qualitatively labeled; however, only the member words could serve for the commonality of topics, when the aforementioned advanced tooling is utilized. Moreover, the dataset used for topic modeling should be collected at specific periods, and a longitudinal study is conducted to observe behavioral changes and the differences in strategies in resisting isomorphism among companies.

References

Abbott, K. W., Green, J. F., & Keohane, R. O. (2016). Organizational ecology and institutional change in global governance. *International Organization*, *70*(2), 247-277.

Ahrens, P. (2018). Qualitative network analysis: A useful tool for investigating policy networks in transnational settings? *Methodological Innovations*, *11*(1), 1-9.

Aral, S., Dellarocas, C., & Godes, D. (2013). Introduction to the special issue—social media and business transformation: A framework for research. *Information Systems Research*, *24*(1), 3-13.

Asadullah, A., Faik, I., & Kankanhalli, A. (2018). Digital Platforms: A review and future directions. *PACIS Proceedings*, 1-14.

Barnes, N.G., Kane, A., Maloney, K. (2018). The 2018 Fortune 500 target millennial and seek uncensored expression. *UMASSD*. Retrieved from https://www.umassd.edu/cmr/social-media-research/2018-fortune-500/

Barnett, W. P., & Carroll, G. R. (1995). Modeling internal organizational change. *Annual review of sociology*, *21*(1), 217-236.

Bastian, M., Heymann, S., & Jacomy, M. (2009, March). Gephi: An open source software for exploring and manipulating networks. In *Third international AAAI conference on weblogs and social media*.

Bastian, M., Ramos Ibañez, E., Jacomy, M., Bartosiak, C., Heymann, S., Bilcke, J., McSweeney, P., Panisson, A., Subtil, J., Suzuki, H., Skurla, M., & Patriarca, A. (2016). Gephi (Version 0.9.1) [Software]. Retrieved from https://gephi.org/users/download/.

Baxter, P., & Jack, S. (2008). Qualitative case study methodology: Study design and implementation for novice researchers. *The qualitative report*, *13*(4), 544-559.

Bettis, R. A., & Hitt, M. A. (1995). The new competitive landscape. *Strategic Management Journal*, *16*(S1), 7-19.

Bílková, M., Greco, G., Palmigiano, A., Tzimoulis, A., & Wijnberg, N. (2018). The logic of resources and capabilities. *The Review of Symbolic Logic*, *11*(2), 371-410.

Borgatti, S. P., Mehra, A., Brass, D. J., & Labianca, G. (2009). Network analysis in the social sciences. *Science*, *323*(5916), 892-895.

Boyd, D. M., & Ellison, N. B. (2007). Social network sites: Definition, history, and scholarship. *Journal of computer-mediated Communication*, *13*(1), 210-230.

Blaschke, M., Haki, K., Aier, S., & Winter, R. (2018). Capabilities for digital platform survival: Insights from a business-to-business digital platform. In *Thirty Ninth International Conference on Information Systems.* Retrieved from https://www.alexandria.unisg.ch/255598/1/Capabilities%20for%20Digital%20Platform%20Survival_%20Insights%20from%20a%20Busin.pdf

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM, 55*(4), 77-84.

Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and experiment*, *2008*(10), 10008.

Brennan, R. and Croft, R. (2012). The use of social media in B2B marketing and branding: an exploratory study. *Journal of Customer Behaviour*, *11*(2), 101-115.

Burt, R. S. (2000). The network structure of social capital. *Research in Organizational Behavior*, *22*, 345-423.

Capioto, G. R., Barbosa, D. H., Kurumoto, J. S., & Cotrim, S. L. (2019). Suppliers' network analysis under the perspective of structural, relational and cognitive embeddedness: An exploratory study. *Production*, *29*, 1-15.

Carroll, G. R. (1985). Concentration and specialization: Dynamics of niche width in populations of organizations. *American Journal of Sociology*, *90*, 1262-1283.

Carroll, G. R., & Hannan, M. T. (2004). *The demography of corporations and industries*. Oxford: Princeton University Press.

Company overview of G2 Crowd, Inc. (2019, June 22). Retrieved from
https://www.bloomberg.com/research/stocks/private/snapshot.asp?privcapId=228545085

Cerqueti, R., Ferraro, G., & Iovanella, A. (2018). A new measure for community structures
through indirect social connections. *Expert Systems with Applications*, *114*, 196-209.

Choi, T. Y., & Kim, Y. (2008). Structural embeddedness and supplier management: a network
perspective. *Journal of Supply Chain Management*, *44*(4), 5-13.

Castro, G. M. D., Sáez, P. L., & López, J. E. N. (2004). The role of corporate reputation in
developing relational capital. *Journal of Intellectual Capital*, *5*(4), 575-585.

Curtis, L., Edwards, C., Fraser, K. L., Gudelsky, S., Holmquist, J., Thornton, K., & Sweetser, K.
D. (2010). Adoption of social media for public relations by nonprofit
organizations. *Public Relations Review*, *36*(1), 90-92.

Customer Data Platform Institution. (2019). Customer data platform industry updated. Retrieve
from http://pearlthinks.com/files/2019/02/CDPI-Industry-Update-Jan-2019.pdf

Dacin, M. T., Beal, B. D., & Ventresca, M. J. (1999). The embeddedness of organizations:
Dialogue & directions. *Journal of Management*, *25*(3), 317-356.

De Meo, P., Ferrara, E., Fiumara, G., & Provetti, A. (2014). On Facebook, most ties are
weak. *Journal of Communications of the ACM*, *57(11),* 78-84.

Diesner, J. (2013). From texts to networks: Detecting and managing the impact of
methodological choices for extracting network data from text data. *KI-Künstliche
Intelligenz*, *27*(1), 75-78.

Diesner, J., Aleyasen, A., Chin, C., Mishra, S., Soltani, K., Tao, L. (2015). ConText: Network
construction from texts [Software]. Retrieved from http://context.ischool.illinois.edu/

DiMaggio, J.P., Powell, W.W. (1983). The iron cage revisited: Institutional isomorphism and
collective rationality in organizational fields. *American Sociological Review*, *48*(2), 147-
160.

Dobrev, S., Kim, T., & Hannan, M. (2001). Dynamics of niche width and resource partitioning. *American Journal of Sociology, 106*(5), 1299-1337.

Dutch Transformation Forum. (2018). Unlocking the value of the platform economy. Mastering the good, the bad and the ugly. Retrieved from https://dutchitchannel.nl/612528/dutch-transformation-platform-economy-paper-kpmg.pdf

Earley, S. (2018). The role of a customer data platform. *IT Professional*, *20*(1), 69-76.

Easley, D., & Kleinberg, J. (2010). *Networks, crowds, and markets* (Vol. 8). Cambridge, MA: Cambridge University Press.

Echols, A., & Tsai, W. (2005). Niche and performance: The moderating role of network embeddedness. *Strategic Management Journal*, *26*(3), 219-238.

Ellison, N. B., Gibbs, J. L., & Weber, M. S. (2015). The use of enterprise social network sites for knowledge sharing in distributed organizations: The role of organizational affordances. *American Behavioral Scientist*, *59*(1), 103-123.

Findik, M., & Bedük, A. (2014). Organizational isomorphism in the context of institutional theory. *International Journal of Social and Economic Sciences*, *4*(1), 27-33.

Fowler, S. W., Lawrence, T. B., & Morse, E. A. (2004). Virtually embedded ties. *Journal of Management*, *30*(5), 647-666.

Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social networks*, *1*(3), 215-239.

Frumkin, P., & Galaskiewicz, J. (2004). Institutional isomorphism and public sector organizations. *Journal of Public Administration Research and Theory*, *14*(3), 283-307.

Funk, T. (2011). *Social Media playbook for business: Reaching your online community with Twitter, Facebook, LinkedIn, and more.* Santa Barbara, California: ABC-CLIO.

Gephi Tutorial Layouts. (n.d.). Retrieved from https://gephi.org/users/tutorial-layouts/

Gephi Tutorial Quick Start. (n.d.). Retrieved from https://gephi.org/users/quick-start/

Gloor, P. A., Woerner, S., Schoder, D., Fischbach, K., & Colladon, A. F. (2016). Size does not matter–in the virtual world. Comparing online social networking behavior with business success of entrepreneurs. *International Journal of Entrepreneurial Venturing*, *10*(4), 435-455.

Gnyawali, D. R., & Madhavan, R. (2001). Cooperative networks and competitive dynamics: A structural embeddedness perspective. *Academy of Management review*, *26*(3), 431-445.

Go, E., & You, K. H. (2016). But not all social media are the same: Analyzing organizations' social media usage patterns. *Telematics and Informatics*, *33*(1), 176-186.

Goktepe, O. (2018). Social media *as a* competitive tool *in* organizations: The Red Queen effect. *Journal of Media Critiques*, *4*(13).

Gonzalez, G., Claro, D., & Palmatier, R. (2014). Synergistic effects of relationship managers's social networks on sales performance. *Journal of Marketing, 78*(1), 76-94.

Granovetter, M. (1985). Economic action and social structure: The problem of embeddedness. *American Journal of Sociology*, *91*(3), 481-510.

Granovetter, M. (1992). Problems of explanation in economic sociology. In Nitin Nohria & Robert G. Eccles (Eds.) *Networks and organizations: Structure, form, and action* (pp. 25-56). Boston, MA: Harvard Business School Press.

Greenberg, P. (2017, March 23). How customer data platforms can benefit your business? *ZDNet*. Retrieved from https://www.zdnet.com/article/introduction-to-customer-data-platforms/.

Gulati, R., & Gargiulo, M. (1999). Where do interorganizational networks come from? *American journal of sociology*, *104*(5), 1439-1493.

Gulati, R., Nohria, N., & Zaheer, A. (2000). Strategic networks. *Strategic management journal*, *21*(3), 203-215.

Gulati, R., Dania, D., and Lihua, W. (2002). Organizational networks. In J.A.C. Baum (Ed.) *Blackwell Companion to Organizations* (pp. 281–303). Boston, MA: Blackwell Publishers.

Guy, I., Ronen, I., Zwerdling, N., Zuyev-Grabovitch, I., & Jacovi, M. (2016). What is your organization 'Like'?: A Study of liking activity in the enterprise. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 3025-3037). ACM.

Håkansson, H., & Snehota, I. (1989). No business is an island: the network concept of business strategy. *Scandinavian journal of management*, *5*(3), 187-200.

Hannan, M. T. (2005). Ecologies of organizations: Diversity and identity. *The Journal of Economic Perspectives, 19*(1), 51-70.

Hannan, M. T., & Freeman, J. H. (1977). The population ecology of organizations. *American journal of sociology*, *82*(5), 929-964.

Hannan, M. T., & Freeman, J. H. (1989). *Organizational ecology.* Cambridge, MA: Harvard University Press.

Hannan, M. T., Pólos, L., & Carroll, G. R. (2012). *Logics of organization theory: Audiences, codes, and ecologies*. Oxford: Princeton University Press.

Hansen, D., Shneiderman, B., & Smith, M. A. (2010). *Analyzing social media networks with NodeXL: Insights from a connected world*. Burlington, MA: Morgan Kaufmann.

Haveman, H. A., & Kluttz, D. N. (2015). Organizational populations and fields. In Robert Scott and Stephen Kosslyn (Eds.,) *Emerging trends in the social and behavioral sciences: An interdisciplinary, searchable, and linkable resource* (pp. 1-15). Hoboken, NJ: John Wiley and Sons.

Heerschap, N., Pouw, N., & Atmé, Cybele (2018, December). Measuring online platform. *CBS*. Retrieved from https://www.cbs.nl/en-gb/background/2019/06/measuring-online-platforms

Heller Baird, C., & Parasnis, G. (2011). From social media to social customer relationship management. *Strategy and Leadership*, *39*(5), 30-37.

Hess, M. (2004). 'Spatial'relationships? Towards a reconceptualization of embeddedness. *Progress in Human Geography*, *28*(2), 165-186.

Hoberg, G., & Phillips, G. (2016). Text-based network industries and endogenous product differentiation. *Journal of Political Economy*, *124*(5), 1423-1465.

Hu, Y., Boyd-Graber, J., Satinoff, B., & Smith, A. (2014). Interactive topic modelling. *Machine Learning, 95*(3), 423-469.

Icha, O., & Edwin, A. (2016). Effectiveness of social media networks as a strategic tool for organizational marketing management. *Journal of Internet Banking and Commerce*, *21*(S2), 1-19.

Jack, S. L., & Anderson, A. R. (2002). The effects of embeddedness on the entrepreneurial process. *Journal of Business Venturing*, *17*(5), 467-487.

Kane, G. C., Alavi, M., Labianca, G., & Borgatti, S. (2014). What is different about social media networks? A framework and research agenda. *Management Information Systems Quarterly*, *38*(1), 275-304.

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of social media. *Business Horizons*, *53*(1), 59-68.

Katopol, P. (2016). Managing change with environmental scanning. *Library Leadership & Management*, *29*(1), 1-7.

Kenney, M., & Zysman, J. (2016). The rise of the platform economy. *Issues in Science and Technology*, *32*(3), 61.

Khavandkar, E., Theodorakopoulos, N., Hart, M., & Preston, J. (2016). Leading the diffusion of intellectual capital management practices in science parks. In Shipton H., Budhwar P., Sparrow P., Brown A. (Eds) *Human resource management, innovation and performance* (pp. 213-231). London: Palgrave Macmillan.

Kietzmann, J. H., Hermkens, K., McCarthy, I. P., & Silvestre, B. S. (2011). Social media? Get serious! Understanding the functional building blocks of social media. *Business Horizons*, *54*(3), 241-251.

Kuhn, K., Galloway, T., & Collins-Williams, M. (2016). Near, far, and online: small business owners' advice-seeking from peers. *Journal of Small Business and Enterprise Development, 23*(1), 189-206.

Lee, J. (2019a). Digital research methods (Week 3 topic modeling and sentiment analysis). Technical report. Rotterdam: Erasmus University Rotterdam.

Lee, J. (2019b). Digital research methods (Week 4 (social) network analysis). Technical report. Rotterdam: Erasmus University Rotterdam.

MacMillan, K., & Komar, J. (2018). Population ecology (organizational ecology): An experiential exercise demonstrating how organizations in an industry are born, change, and die. *Journal of Management Education*, *42*(3), 375-397.

McKenna, B., Myers, M. D., & Newman, M. (2017). Social media in qualitative research: Challenges and recommendations. *Information and Organization*, *27*(2), 87-99.

Mislove, A., Marcon, M., Gummadi, K. P., Druschel, P., & Bhattacharjee, B. (2007, October). Measurement and analysis of online social networks. In *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement* (pp. 29-42). ACM.

Mohr, J. W., & Bogdanov, P. (2013). Introduction—Topic models: What they are and why they matter. *Poetics*, *6*(41), 545-569.

Montanari, F., & Mizzau, L. (2007). The influence of embeddedness and social mechanisms on organizational performance in the music industry: The case of Mescal Music. *International Journal of Arts Management*, *10*(1), 32.

Moody, J., & White, D. R. (2003). Structural cohesion and embeddedness: A hierarchical concept of social groups. *American Sociological Review*, 68(1), 103-127.

Morse, E. A., Fowler, S. W., & Lawrence, T. B. (2007). The impact of virtual embeddedness on new venture survival: Overcoming the liabilities of newness. *Entrepreneurship theory and practice*, *31*(2), 139-159.

Mwai, G. M., Namada, J. M., & Katuse, P. (2018). Influence of organizational resources on organizational effectiveness. *American Journal of Industrial and Business Management*, *8*(06), 1634.

Neuman, W. (2014). *Social research methods: Qualitative and quantitative approaches.* Harlow: Pearson.

Ngai, E. W., Tao, S. S., & Moon, K. K. (2015). Social media research: Theories, constructs, and conceptual frameworks. *International Journal of Information Management*, *35*(1), 33-44.

Noble, H., & Smith, J. (2015). Issues of validity and reliability in qualitative research. *Evidence-based nursing*, *18*(2), 34-35.

Palacios-Marqués, D., Merigó, J., & Soto-Acosta, P. (2015). Online social networks as an enabler of innovation in organizations. *Management Decision, 53*(9), 1906-1920.

Parveen, F., Jaafar, N. I., & Ainin, S. (2015). Social media usage and organizational performance: Reflections of Malaysian social media managers. *Telematics and Informatics*, *32*(1), 67-78.

Penni, J. (2017). The future of online social networks (OSN): A measurement analysis using social media tools and application. *Telematics and Informatics*, *34*(5), 498-517.

Podolny, J., Stuart, T., & Hannan, M. (1996). Networks, knowledge, and niches: Competition in the worldwide semiconductor industry, 1984-1991. *American Journal of Sociology, 102*(3), 659-689.

Porter, M.E. (2001). Strategy and the internet. *Harvard Business Review*, *79*(3), 63–78.

Ratajczak-Mrozek, M. (2017). *Network embeddedness: Examining the effect on business performance and internationalization*. Cham, Switzerland: Springer.

Ressler, S. (2006). Social network analysis as an approach to combat terrorism: Past, present, and future research. *Homeland Security Affairs*, *2*(2).

Robert Walters. (2018). Business Intelligence Market Guide. Retrieved from https://www.robertwalters.co.uk/content/dam/robert-walters/country/united-kingdom/files/whitepapers/Robert-Walters-BI-Market-Guide-2018.pdf

Schipper, D., & Spekkink, W. (2015). Balancing the quantitative and qualitative aspects of social network analysis to study complex social systems. *Complexity, governance & networks*, *2*(1), 5-22.

Schultz, R. J., Schwepker Jr, C. H., & Good, D. J. (2012). Social media usage: an investigation of B2B salespeople. *American Journal of Business*, *27*(2), 174-194.

Singh, J. V., & Lumsden, C. J. (1990). Theory and research in organizational ecology. *Annual review of sociology*, *16*(1), 161-195.

Smith, C., Smith, J. B., & Shaw, E. (2017). Embracing digital networks: Entrepreneurs' social capital online. *Journal of Business Venturing*, *32*(1), 18-34.

Statista. (2019). Number of Facebook users worldwide 2008-2019. Retrieved from https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/

Strategic Policy Forum on Digital Entrepreneurship. (2016). Big data and B2B digital platforms: The next frontier for Europe's industry and enterprises. Retrieve from http://edz.bib.uni-mannheim.de/daten/edz-h/gdb/16/Big%20data%20and%20B2B%20digital%20platforms.pdf

Stringfellow, A. (2016, November 17). 37 Marketing pros reveal how real-time marketing impacts and benefits brand (with real world examples). *NGData*. Retrieved from https://www.ngdata.com/real-time-marketing-examples/

Susarla, A., Oh, J. H., & Tan, Y. (2012). Social networks and the diffusion of user-generated content: Evidence from YouTube. *Information Systems Research*, *23*(1), 23-41.

Sveiby, K. E. (2001). A knowledge-based theory of the firm to guide in strategy formulation. *Journal of intellectual capital, 2*(4), 344-358.

Swaminathan, A. (2001). Resource partitioning and the evolution of specialist organizations: The role of location and identity in the US wine industry. *Academy of Management Journal*, *44*(6), 1169-1185.

Takac, C., Hinz, O., & Spann, M. (2011). The social embeddedness of decision making: Opportunities and challenges. *Electronic Markets*, *21*(3), 185.

Thackeray, R., Neiger, B. L., Hanson, C. L., & McKenzie, J. F. (2008). Enhancing promotional strategies within social marketing programs: use of Web 2.0 social media. *Health promotion practice*, *9*(4), 338-343.

Thelwall, M. (2017). The heart and soul of the web? Sentiment strength detection in the social web with SentiStrength. In Holyst J. (eds) *Cyberemotions* (pp. 119-134). Cham, Switzerland: Springer.

Uzzi, B. (1996). The sources and consequences of embeddedness for the economic performance of organizations: The network effect. *American Sociological Review*, *61*, 674-698.

Uzzi, B. (1997). Social structure and competition in interfirm networks: The paradox of embeddedness. *Administrative Science Quarterly*, *42*, 35-67.

Uzzi, B. (1999). Embeddedness in the making of financial capital: How social relations and networks benefit firms seeking financing. *American Sociological Review*, *64*, 481-505.

Vasileiadou, E. M., & Missler-Behr, M. (2011). Virtual embeddedness and social media as a basis for the relational capital Management of new ventures. *Electronic Journal of Knowledge Management*, *9*(3).

Vasudeva, G., Zaheer, A., & Hernandez, E. (2013). The embeddedness of networks: Institutions, structural holes, and innovativeness in the fuel cell industry. *Organization science*, *24*(3), 645-663.

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications* (Vol. 8). Cambridge, MA: Cambridge University Press.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small world' networks. *Nature*, *393*(6684), 440.

Wilson, C., Boe, B., Sala, A., Puttaswamy, K. P., & Zhao, B. Y. (2009, April). User interactions in social networks and their implications. In *Proceedings of the 4th ACM European conference on Computer systems* (pp. 205-218). ACM.

World Economic Forum. (2017). *Digital Transformation Initiative: Unlocking B2B Platform Value* [White paper]. Retrieved from World Economic Forum website: http://reports.weforum.org/digital-transformation/wp-content/blogs.dir/94/mp/files/pages/files/wef-platform-report-final-3-26-17.pdf

Wu, L. Y., Wang, C. J., Chen, C. P., & Pan, L. Y. (2008). Internal resources, external network, and competitiveness during the growth stage: A study of Taiwanese high–tech ventures. *Entrepreneurship theory and practice*, *32*(3), 529-549.

Yang, H., Lin, Z., & Peng, M. W. (2011). Behind acquisitions of alliance partners: exploratory learning and network embeddedness. *Academy of Management Journal, 54*(5), 1069-1080.

Zhang, M., Jansen, B. J., & Chowdhury, A. (2011). Business engagement on Twitter: a path analysis. *Electronic Markets*, *21*(3), 161.

Zhang, Y., Li, J., Jiang, W., Zhang, H., Hu, Y., & Liu, M. (2018). Organizational structure, slack resources and sustainable corporate socially responsible performance. *Corporate Social Responsibility and Environmental Management*, *25*(6), 1099-1107.

Appendix A: An example of the document-to-topic output of CB Insights

	Document	Fit to Topic1	Fit to Topic	Fit to Topic	Fit to Topic	Fit to Topic	Fit to Topic	Fit to Topic7
2	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_445-RS-ST.txt	0.968974844	0.006296309	0.00606703	0.005796866	0.005017596	0.004265597	0.003581758
3	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_465-RS-ST.txt	0.968974844	0.006296309	0.00606703	0.005796866	0.005017596	0.004265597	0.003581758
4	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_184-RS-ST.txt	0.96548628	0.007004286	0.006749226	0.006448683	0.00558179	0.004745234	0.003984502
5	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_547-RS-ST.txt	0.96548628	0.007004286	0.006749226	0.006448683	0.00558179	0.004745234	0.003984502
6	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_269-RS-ST.txt	0.963430274	0.007421536	0.007151282	0.006832836	0.005914301	0.005027911	0.004221861
7	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_272-RS-ST.txt	0.963430274	0.007421536	0.007151282	0.006832836	0.005914301	0.005027911	0.004221861
8	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_464-RS-ST.txt	0.963430274	0.007421536	0.007151282	0.006832836	0.005914301	0.005027911	0.004221861
9	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_131-RS-ST.txt	0.961113795	0.007891647	0.007604274	0.007265656	0.006288937	0.0053464	0.004489291
10	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_188-RS-ST.txt	0.961113795	0.007891647	0.007604274	0.007265656	0.006288937	0.0053464	0.004489291
11	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_280-RS-ST.txt	0.961113795	0.007891647	0.007604274	0.007265656	0.006288937	0.0053464	0.004489291
12	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_403-RS-ST.txt	0.961113795	0.007891647	0.007604274	0.007265656	0.006288937	0.0053464	0.004489291
13	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_36-RS-ST.txt	0.958483999	0.008425343	0.008118535	0.007757018	0.006714246	0.005707966	0.004792893
14	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_168-RS-ST.txt	0.951990415	0.009743164	0.009388368	0.008970305	0.007764431	0.006600758	0.005542557
15	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_172-RS-ST.txt	0.951990415	0.009743164	0.009388368	0.008970305	0.007764431	0.006600758	0.005542557
16	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_194-RS-ST.txt	0.947917242	0.010569782	0.010184885	0.009731355	0.008423172	0.007180772	0.006012792
17	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_46-RS-ST.txt	0.947917242	0.010569782	0.010184885	0.009731355	0.008423172	0.007180772	0.006012792
18	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_294-RS-ST.txt	0.943088852	0.011549666	0.011129086	0.010633509	0.009204052	0.007824619	0.006570215
19	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_305-RS-ST.txt	0.921162777	0.015999389	0.015416773	0.014730266	0.012750083	0.010839199	0.009101512
20	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_41-RS-ST.txt	0.921162777	0.015999389	0.015416773	0.014730266	0.012750083	0.010839199	0.009101512
21	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_438-RS-ST.txt	0.918436041	0.006296309	0.00606703	0.056135669	0.005017596	0.004265597	0.003581758
22	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_432-RS-ST.txt	0.909264737	0.063225828	0.006749226	0.006448683	0.00558179	0.004745234	0.003984502
23	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_462-RS-ST.txt	0.909264737	0.007004286	0.062370768	0.006448683	0.00558179	0.004745234	0.003984502
24	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_425-RS-ST.txt	0.903859574	0.007421536	0.007151282	0.006832836	0.005914301	0.06459861	0.004221861
25	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_120-RS-ST.txt	0.89776964	0.071235801	0.007604274	0.007265656	0.006288937	0.0053464	0.004489291
26	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_162-RS-ST.txt	0.893915405	0.021529026	0.02074505	0.019821275	0.017156711	0.014585395	0.012247137
27	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_220-RS-ST.txt	0.893915405	0.021529026	0.02074505	0.019821275	0.017156711	0.014585395	0.012247137
28	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_316-RS-ST.txt	0.893915405	0.021529026	0.02074505	0.019821275	0.017156711	0.014585395	0.012247137
29	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_440-RS-ST.txt	0.890856006	0.076053336	0.008118535	0.007757018	0.006714246	0.005707966	0.004792893
30	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_460-RS-ST.txt	0.874252372	0.005993409	0.053882668	0.005517993	0.052883718	0.00406039	0.003409448
31	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_354-RS-ST.txt	0.874252372	0.005993409	0.101990178	0.005517993	0.004776212	0.00406039	0.003409448
32	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_284-RS-ST.txt	0.873784621	0.009743164	0.009388368	0.0871761	0.007764431	0.006600758	0.005542557
33	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_48-RS-ST.txt	0.873784621	0.005743164	0.087594163	0.008970305	0.007764431	0.006600758	0.005542557
34	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_138-RS-ST.txt	0.871753393	0.026026631	0.025078875	0.023962115	0.0207409	0.017612414	0.014805672
35	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_158-RS-ST.txt	0.871753393	0.026026631	0.025078875	0.023962115	0.0207409	0.017612414	0.014805672
36	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_17-RS-ST.txt	0.871753393	0.026026631	0.025078875	0.023962115	0.0207409	0.017612414	0.014805672
37	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_292-RS-ST.txt	0.871753393	0.026026631	0.025078875	0.023962115	0.0207409	0.017612414	0.014805672
38	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_3-RS-ST.txt	0.871753393	0.026026631	0.025078875	0.023962115	0.0207409	0.017612414	0.014805672
39	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_383-RS-ST.txt	0.871753393	0.026026631	0.025078875	0.023962115	0.0207409	0.017612414	0.014805672
40	CB_169041553285088_2019_03_28_10_41_32_fullstats.tab_185-RS-ST.txt	0.837886432	0.032899662	0.031701626	0.030289955	0.026218091	0.022288726	0.018715507

Appendix B: The list of 14 companies in the subsample and their labelled topics

Compay	Topic	Weight	TopicMembers
500 Startups	T_Accelerator program	0.117	batch - startup - day - company - accelerator - apply - program - join - kobe - investment
500 Startups	T_Startup event	0.095	miami - inclusion - startup - join - unity - diversity - founder - tech - team - early -
500 Startups	T_Growth	0.089	startup - grow - year - vietnam - global - southeast - check - asia - tip - cover -
500 Startups	T_Investor webinar	0.087	today - investor - program - learn - apply - early - unlocked - day - week - webinar -
500 Startups	T_CEO/founder talks	0.086	ceo - founder - company - strong - share - portfolio - investment - feature - christine - tsai -
500 Startups	T_Funding	0.075	million - raise - investor - round - sanfrancisco.premoney.co - ventures - funding - ticket - conference - fund -
500 Startups	T_Partnership	0.071	startup - blockchain - partner - venture - valley - silicon - program - business - announce - learn -
CB Insights	T_Healthcare market	0.172	trend - report - industry - healthcare - watch - wellness - biggest - top - trends - technology -
CB Insights	T_Tech trend	0.124	client - note - startups - fintech - future - blockchain - don - register - briefing - unicorns -
CB Insights	T_Funding	0.12	report - download - funding - deal - complete - global - publish - increase - moneytree - pdf -
CB Insights	T_Startup report	0.114	startup - company - tech - unicorn - year - list - investor - check - full - valuation -
CB Insights	T_Consumer advice	0.099	consumer - company - business - startup - digital - map - brand - impact - space - break -
CB Insights	T_Product introduction	0.084	market - today - product - insight - research - start - day - data - team - trial -
CB Insights	T_Tech companies	0.07	google - amazon - apple - market - world - acquire - smart - data - back - uber -
Constant Contact	T_Email marketing	0.464	email - marketing - customer - send - business - contact - learn - create - automation - time -
Constant Contact	T_Holiday business	0.256	business - small - customer - day - start - today - holiday - year - time - season -
Constant Contact	T_Online business	0.174	contact - social - constant - media - facebook - business - website - page - grow - list -
Constant Contact	T_Sharing sessions	0.136	event - free - marketing - business - today - expert - join - register - share - webinar -
Constant Contact	T_Email tip	0.123	email - open - rate - send - line - subject - bounce - friday - quiz - receive -
Constant Contact	T_Personalized email	0.099	make - people - feel - image - color - mistake - donor - content - nonprofit - audience -
Constant Contact	T_Internal activity	0.070	day - comment - giveaway - winner - win - december - purchase - visit - prize - unique -
DigitalOcean	T_Server	0.118	server - secure - set - datum - load - user - create - guide - infrastructure - intro -
DigitalOcean	T_Tutorial	0.090	learn - python - tutorial - install - web - open - tool - source - part - click -
DigitalOcean	T_Product introduction	0.080	droplet - plan - month - celebrate - dohq - black - happy - cpu - digitalocean - nyc -
DigitalOcean	T_Space storage	0.075	space - storage - object - digitalocean - datum - cover - store - back - backup - api -
DigitalOcean	T_Expert sharing	0.070	learn - cloud - webinar - deploy - janakiram - expert - build - docker - kubernetes - scale -
DigitalOcean	T_Advice	0.067	product - share - work - team - engineering - blog - week - director - build - comment -
DigitalOcean	T_Trend	0.056	developer - current - trend - survey - open - currents - currents-survey - sammy - engineer - source -
Docusign	T_Digital signature	0.186	digital - docusign - business - customer - experience - company - technology - process - work - cloud -
Docusign	T_Product efficiency	0.167	time - day - docusign - sign - document - signature - esignature - save - electronic - team -
Docusign	T_Business event	0.158	docusign - london - digital - business - momentum - join - esignature - event - learn - year -
Docusign	T_Tech influence	0.088	life - make - easier - device - people - smart - transfer - bring - app - tech -
Docusign	T_CEO/founder talks	0.063	docusign - ceo - dan - innovation - springer - release - read - announce - news - international -
Docusign	T_Planet wareness	0.054	impact - docusignforforest - calculator - space - water - forest - city - paper - tree - video -
Docusign	T_Security	0.052	docusign - datum - update - system - trust - customer - document - photo - security - access -
G2 Crowd	T_Community service	0.206	review - crowd - day - work - giveback - give - software - donate - proud - leave -
G2 Crowd	T_Marketing event	0.192	crowd - company - marketing - review - miss - don - booth - user - check - love -
G2 Crowd	T_Social media usage	0.181	facebook - crowd - tool - business - blog - marketing - customer - start - video - product -
G2 Crowd	T_Review blog	0.138	crowd - research - blog - review - explore - datum - technology - smallbusiness - series - life -
G2 Crowd	T_Networking	0.132	chicago - crowd - office - join - thecrowd - today - marketing - brand - work - cmo -
G2 Crowd	T_Best Software	0.090	software - list - company - crowd - top - make - congratulation - recognize - teams - slack -
G2 Crowd	T_CEO/founder talks	0.070	ceo - abel - godard - news - intent - buyer - learn - datum - woman - partner -

Gust	T_CEO/founder talks	0.209	investor - startup - david - rose - pitch - ceo - share - founder - angel - entrepreneur -
Gust	T_Teamwork	0.126	equity - startup - founder - team - work - time - company - dilution - ownership - right -
Gust	T_Success	0.120	startup - business - company - legal - founder - question - launch - make - success - answer -
Gust	T_Funding	0.112	startup - entrepreneur - zwilling - focus - note - martin - angels - funding - market - convertible -
Gust	T_Invest valuation	0.105	valuation - investor - startup - venture - make - capital - financial - pre-money - tip - business -
Gust	T_Startup report	0.101	startup - tech - accelerator - nyc - report - ecosystem - check - change - invest - york -
Gust	T_Accelerator program	0.075	apply - startup - accelerator - mentorship - deadline - investor - funding - offer - application - access -
HubSpot	T_Marketing strategy	1.187	business - marketing - company - make - work - change - year - facebook - people - day -
HubSpot	T_Advice	0.955	marketer - brand - social - world - teach - learn - media - good - favorite - thing -
HubSpot	T_CEO/founder talks	0.324	hubspot - customer - inbound - shah - dharmesh - representation - set - hire - join - online -
HubSpot	T_Mood	0.276	mood - relate - create - man - question - hour - marketing - industry - week - make -
HubSpot	T_Online emotion	0.245	laugh - emotion - catch - emojus - messenger - online - interaction - today - conservations - personalization -
HubSpot	T_Trust	0.242	favorite - influencer - trust - sit - technological - real - biometric - datum - video - fake -
HubSpot	T_Change	0.234	brand - video - change - huge - test - movie - media - market - meme - time - grow
Marketo	T_Customer experience	0.277	customer - learn - marketing - brand - experience - marketer - datum - content - create - company -
Marketo	T_Advice	0.120	marketing - strategy - learn - impact - email - team - roi - marketer - business - increase -
Marketo	T_Marketing webinar	0.103	learn - marketing - join - start - webinar - website - process - abm - digital - management -
Marketo	T_Marketing strategy	0.097	marketing - sale - learn - content - lead - team - marketo - strategy - guide - secret -
Marketo	T_Adobe Summit	0.096	marketo - learn - adobesummit - join - marketing - ticket - team - partner - find - event -
Marketo	T_CEO/founder talks	0.089	marketer - marketing - year - cmo - learn - market - list - steve - leader - marketober -
Marketo	T_Sharing sessions	0.067	marketo - share - champion - fearless - edition - latest - fearlessforum - top - practice - member -
NGDATA	T_Customer experience webi	0.203	customer - marketing - free - webinar - experience - learn - make - datum - guide - brand -
NGDATA	T_Product introduction	0.195	customer - experience - video - business - journey - deliver - analytic - lily - company - brand -
NGDATA	T_CDP Market	0.130	datum - company - customer - data - cdp - platform - technology - ngdata - machine - intelligence -
NGDATA	T_Marketing strategy	0.119	marketing - video - strategy - marketer - effective - round - single - acquisition - important - campaign -
NGDATA	T_Bank industry	0.096	ngdata - report - service - market - banking - world - bank - industry - analytics - financial -
NGDATA	T_CEO/founder talks	0.083	ngdata - ceo - burgelman - luc - talk - digital - gdpr - noels - steven - intelligence -
NGDATA	T_Data management	0.079	datum - company - machine - lake - business - tool - process - leverage - analytic - intelligence -
Orcale	T_Product introduction	0.439	oracle - cloud - learn - data - business - autonomous - discover - join - customer - webcast -
Orcale	T_AI	0.212	oracle - company - read - cloud - change - business - ai - find - learn - make -
Orcale	T_Security	0.093	cloud - find - oracle - management - bot - people - money - security - week - darkbot -
Orcale	T_Empowering women	0.084	mark - hurd - oracle - ceo - share - announce - winner - womenshistorymonth - future - tag -
Orcale	T_Marketing skills	0.062	oracle - deliver - experience - marketing - american - information - read - skill - turn - check -
Orcale	T_Education	0.057	oracle - make - chocolate - school - tech - developer - high - work - design - beer -
Orcale	T_Internal activity	0.054	sailgp - team - francisco - san - app - australia - race - oracle - day - sydney -
seoplus+	T_SEO strategy	0.200	seo - website - shopify - store - google - strategy - lindsay - expert - sale - blog -
seoplus+	T_Social media usage	0.190	blog - instagram - read - day - post - business - facebook - make - learn - change -
seoplus+	T_Marketing strategy	0.168	marketing - digital - content - ottawa - media - social - specialist - top - year - business -
seoplus+	T_Expert sharing	0.145	brock - murray - article - search - feature - read - seo - share - coo - check -
seoplus+	T_Internal activity	0.137	happy - day - seoplus - team - director - today - office - cuteness - month - friend -
seoplus+	T_Tech event	0.111	google - event - seoplus - business - mobile - time - free - great - website - register -
seoplus+	T_Recruitment	0.095	team - web - seoplus - digital - designer - specialist - ottawa - amazing - office - join -
Sprout Social	T_Customer connection	0.193	social - brand - customer - people - time - connection - strategy - build - connect - media -
Sprout Social	T_Social media usage	0.158	instagram - content - facebook - social - increase - campaign - brand - engagement - make - marketer -
Sprout Social	T_Marketing strategy	0.142	marketing - social - strategy - brand - customer - pinterest - experience - top - business - put -
Sprout Social	T_Video content	0.127	social - media - video - marketer - guide - don - stay - content - facebook - research -
Sprout Social	T_Product introduction	0.082	sprout - social - team - data - marketer - work - index - meet - learn - attention -
Sprout Social	T_Sharing sessions	0.078	day - agency - partner - year - join - focus - sproutsession - march - share - talk -
Sprout Social	T_Holiday business	0.076	holiday - list - hashtag - people - brand - matter - content - drive - compile - promote -
The Startup Garage	T_Success	0.384	startup - business - entrepreneur - check - success - top - start - year - make - startups -
The Startup Garage	T_CEO/founder talks	0.159	startup - san - diego - garage - capital - entrepreneur - community - jensen - event - founder -
The Startup Garage	T_Funding	0.151	investor - capital - people - raise - founder - relationship - money - deal - year - cover -
The Startup Garage	T_Innovation	0.150	tech - great - time - insights - company - article - silicon - innovation - build - list -
The Startup Garage	T_Internal activity	0.118	startupsavvy - startuplife - tsg - sandiego - getfunded - thestartupgarage - events - day - entrepreneurs - great -
The Startup Garage	T_Startup event	0.083	event - hera - pitch - venture - ceo - tsg - summit - startupsavvy - startuplife - fast -
The Startup Garage	T_Pitch	0.067	pitch - competition - company - quick - business - investor - coast - technology - finalist - give -