**Innovation in Co-Creation Practices: An Exploratory Study**

***By***

***Jainal A. Chisty***

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of
**Master of Applied Science in Technology Innovation Management**

Department of Systems and Computer Engineering,
Carleton University
Ottawa, Canada, K1S 5B6
January 2011

*Your file Votre référence*
*ISBN:* 978-0-494-79536-1
*Our file Notre référence*
*ISBN:* 978-0-494-79536-1

## Abstract

The study of value co-creation is an emerging marketing and innovation paradigm that recognizes the significance of firm-customer value. The objective of this thesis is to use research methodology utilized by Stephen Allen (2009) to provide value creation practices currently active in a sample dominated by open source firms. Data is collected concurrently using Google and Yahoo! search engines to extract two sets of value creation components, and a linear regression analysis is performed to explore the potential relationship between value co-creation components and the ability of firms to produce new offerings. Google value creation components are evaluated against components extracted from the Yahoo! data set, which show similar components related to mutual learning and co-production processes. Linear regression results show that activities surrounding modular products, access to internal firm resources, risk management, and flexible design processes enhance firm ability to develop new offerings.

# Acknowledgements

I am eternally grateful to my parents for giving me the perseverance to keep going when the road seemed never-ending, and to my fiancée for having the patience to keep on supporting me during difficult times. I would also like to thank my friends and family who believed in me. A special appreciation is extended to my thesis supervisor Professor Stoyan Tanev, and to Tony Bailetti, for all their help, advice, and the patient resolve with which they pushed me to complete the program successfully.

## Contents

# List of Figures

## List of Tables

# 1. Introduction

Value co-creation is an emerging business paradigm that describes how customers can be involved as active participants in the design and development of personalized products, services, and experiences (Prahalad & Ramaswamy, 2004; Etgar, 2008; Payne et al, 2008). The active involvement of customers in co-creation is enabled through multiple interaction channels, such as internet-based technological platforms designed specifically to facilitate co-creation practices (Sawhney et al, 2005; Nambisan & Nambisan, 2008; Nambisan & Baron, 2009). This new paradigm supports an understanding of the customer-centric view of traditional value networks, which is considered as a people driven web of potential value configurations that can be actualized on the basis of specific consumer demands (Normann & Ramirez, 1993; Flint & Mentzer, 2006; Prahalad & Krishnan, 2008).

The adoption of value creation practices leads to the need of "changing the very nature of engagement and relationship between the institution of management and co-creators of value – customers, stakeholders, partners or other employees" (Ramaswamy, 2009). While literature outlines several value co-creation streams, there is a lack of research studies that relate value creation efforts to a firm's ability to develop new products, processes and services. There have been some recent studies done to establish activities that are employed by firms to enable the creation of value (Allen, 2009; Tanev et al., 2010). Using a similar research methodology, this research uses web data mining and factor extraction methods to outline value creation components utilized by a sample of firms. Google and Yahoo! search engines are used to find two sets of value creation components that are linked to components outlined by Allen (2009). Furthermore, the relationship between specific value creation components and a firm's ability to develop new offerings is explored through linear regression methods.

## 1.1 Objectives

The objective of this research is to focus on the innovation outcomes of value co-creation practices by (i) using a research methodology pioneered by Stephen Allen (2009) in order to identify the key components of value co-creation practices, using data from two different search engines; and (ii) examining the relationship between a firm's involvement in specific co-creation components and firms' capacity to produce new offerings. For the purpose of this research, value co-creation is defined as a specific type of collaboration between firms, their value networks and their customers. The capacity of the firms to develop new offerings is measured by (i) the frequency of online comments about new products, services, and processes on the firms' websites, and (ii) the number of new products, services, and processes introduced by the firms in the past three years.

## 1.2 Deliverables

This research produces the following deliverables:

- An description of value creation components firms currently use to enable the creation of value, extracted through the use of Google and Yahoo! search engines
- An inventory of firms that are active in the creation of value through the use of these components
- Identification of value co-creation components that influence the capacity of the firms in the sample to develop new products, services, and processes

## 1.3 Contributions

This research makes the following contributions to the existing literature on value co-creation:

- Firm time and effort required to adopt value creation practices is reduced, leading to enhanced production of new products, services, and processes
- Academics can easily implement web data mining and factor extraction methods for their own research by using this research methodology

## 1.4 Relevance

This research is relevant to the following groups of people:

- Entrepreneurs and management teams could be interested in the innovation-related outcomes of co-creation practices
- Researchers and students can use the research methodology to study additional aspects of value co-creation and technology innovation management.

## 1.5 Document organization

This research thesis is organized into seven sections. Section 1 provides an introduction and an overview of the research. Section 2 presents a review of the existing literature on value co-creation and its emerging subcategories, web content data mining, and exploratory factor analysis processes. It also provides an analysis of the two search engines used for this research. Section 3 outlines the research strategy, while section 4 provides the research results. Section 5 provides an analysis of the results, along with the component interpretation. Section 6 presents the summary of the key findings of this research. And finally section 7 outlines the limitations taken into consideration, and offers proposals for future research opportunities.

# 2. Literature review

This section highlights the main research streams for this thesis. Section 2.1 delves into the existing academic literature on value co-creation. The evolution of value co-creation is reviewed, along with the concepts, frameworks and theories that define it. Section 2.2 outlines the literature on internet based data mining. The third section reviews the literature on the concepts used to analyze the data, the varied methods of exploratory factor analysis. Finally section 2.4 provides an overview of the lessons learned from reviewing the literature.

## 2.1 Value co-creation

A systematic review of the literature shows several streams of value co-creation research. The *General Marketing Perspective* (Prahalad & Ramaswamy, 2000, 2004; Jaworski & Kohli, 2006; Etgar, 2006, 2008: Nambisan & Nambisan, 2008; Payne et al., 2008; Ramaswamy, 2009; Ramaswamy & Gouillart, 2010) focuses on value creation practices and their implications; *Virtual Customer Environments* (Edvardssonet al., 2005; Nambisan & Baron, 2007; Nambisan & Nambisan, 2008; Nambisan & Baron, 2009; Nambisan, 2009; Kohler, Matzler & Füller, 2008; Bonsu & Darmody, 2008; Droge et al., 2010; Füller et al. 2006) deals with the partnership of customers and companies in a web-based forum, to collaborate in different phases of product innovation; *Service Dominant Logic of Marketing* (Vargo & Lusch, 2004; Edvardsson et al., 2005; Vargo, 2008; Ballantyne & Varey, 2008; Cova & Salle, 2008; Dong, Evans & Zou, 2008; Maglio & Spohrer, 2008; Kristensson et al., 2008; Bolton & Saxena-Iyer, 2009; Brohmanet al., 2009; Ng, Maull & Yip, 2009; Ferguson & Paulin, 2010; Ostrom et al., 2010) extends the traditional marketing model to include intangible resources as well as services as an inherent part of products; and *New Product Development and Innovation* (Prahalad & Ramaswamy, 2003; Sawhney et al., 2005; Prahalad & Ramaswamy 2008; Franke & Schreier, 2008; Kristensson et al., 2008; Midgley, 2009; Romero & Molina, 2009; Tanev et al., 2009; Nambisan, 2009; Bowonder et al., 2010), which focuses on the personalization of market offers, multiple stakeholder interactions, and access to global resources. The following subsections describe the General Management Perspective, the Virtual Customer Environments, and the New Product Development and Innovation streams in further detail, in order to sufficiently outline the literature relevant to this research on value co-creation.

### 2.1.1 General management perspective

The General Management Perspective of value co-creation is mainly concerned with improving the customer experience while co-creating value. It is mainly associated with value co-creation best-practices and its implications: the building blocks of value co-creation, the dimensions of choice in co-creation experiences and the motivation behind it. Prahalad and Ramaswamy's work is cited extensively as the pioneers behind value co-creation literature. The DART model they developed is geared towards providing guidance on how to build relationships with customers in order to strengthen brand loyalty. Otherwise there is an alarming lack of literature concerning frameworks that help companies manage co-creation (Payne et al, 2008, p. 85).

The traditional image of the market as an aggregation of customers, or as a locus of exchange where the company trades goods and services with a customer, is being challenged by the concept of value co-creation. In this approach to value co-creation, empowered consumers are looking to exercise their influence in all parts of the business system, with whole communities of professionals, service providers, and other consumers, in order to create value for themselves. The changing dynamic of the company-customer interaction as the locus of co-creation redefines the meaning of the value and the value creation process. Dialog, access, risk-benefits, and transparency (DART) are emerging as the basis of communication between the customer and the company. These building blocks of value co-creation challenge the traditional positions held by company managers on labelling laws, disclosure of risks, transparency of financial statements, and open access and dialogue with consumers and communities (Prahalad & Ramaswamy, 2004b, pp.6-9).

The four DART building blocks can be used individually or in combination by companies looking to co-create value with customers. Combining the four building blocks in different ways can create important new capabilities for companies as outlined in Table 1. In combination, access and transparency enables the customer to make more informed choices regarding a company's offerings. Pairing dialogue with risk assessment allows the customer to debate and co-develop policy choices with companies. Additionally, mixing access with dialogue leads to development of thematic communities, and combining transparency with risk assessment facilitates the co-development of trust between company and customer. The combination of the four building blocks to enable co-creation is further outlined in literature by Ballantyne (2004), and Jaworski and Kohli (2006). Ballantyne outlines the importance of dialogue as a factor that enables trust and strengthens the relationship between company and the consumer. This trust emerges through the use of open dialogue (Jaworski & Kohli, 2006).

Table 1: Moving from traditional to value co-creation exchange

| | Traditional Exchange | Co-Creation Exchange |
|---|---|---|
| Goal of interaction | Extraction of economic value | Co-creation of value through compelling co-creation experiences, as well as extracting economic value |
| Locus of interaction | Once at the end of the value chain | Repeatedly, anywhere, and at any time in the system |
| Company-customer relationship | Transaction Based | Set of interactions and transactions based on co-creation experiences |
| View of choice | A variety of products and services, features and functionalities, product performance, and operating procedures | Co-creation experiences based on interactions across multiple channels, options, transactions, and the price-experience relationship |
| Pattern of interaction between the company and the customer | Passive and firm-initiated, one-on-one | Active, initiated by either firm or customer, one-on-one or one-to-many |
| Focus of quality | Quality of internal processes and what firms have on offer | Quality of company-customer interactions and co-creation experiences |

(Source: Prahalad & Ramaswamy, 2004a)

The transformation of traditional business models emerges from the book by C.K. Prahalad and M.S. Krishnan "The New Age of Innovation: Driving Co-Created Value through Global Networks". The traditional business to business (B2B) transactions now become business to customer (B2C) interactions. The outcome of this shift may be that the existing management systems in place today can be a liability in the constantly reconfiguration of business models. The new business model has to be geared towards personalized experiences for individual consumers (Prahalad & Krishnan, 2006).

Michael Etgar (2007) presents a model that presents co-production as a dynamic process that extends over time and includes five different stages that consumers considering co-production should pass through. This model elaborates on the preconditions necessary, and presents the potential cost-benefit information that might motivate consumers to use value co-production. The following five stages summarize a customer strategy designed to achieve customization of marketing offers:

(i) Development of antecedent conditions: Customers willing to participate in co-production should meet certain preconditions. These are macro-environmental conditions, consumer linked, product linked, and situational linked conditions.

(ii) Development of motivations that prompt customers to engage in co-production: Customers engaging in co-production do so in order to achieve preset goals that reflect customer values and serve as motivational forces. Consumer behaviour based on economic and behavioural models list economic, psychological, and social benefits as reasons for engaging in co-production.

(iii) Calculation of cost-benefits of co-production: In this stage consumers should perform a cost-benefit analysis to evaluate the benefits they expect to accrue if they engage in co-production, and weigh them against the relevant costs and risks of co-production activities. They should base their decision to engage in co-production on this cost-benefit analysis.

(iv) Activation: Customers, when they become active in co-production, will need to choose the level of engagement and the activities they wish to be active in. They can select from the following stages: the consumption stage, distribution and logistics stage, the assembly phase, the manufacturing/construction stage, the design phase, and the initiating phase.

(v) Evaluation of outputs and results of the process: The activation stage ends with the creation of various outputs resulting in benefits consumers will need to evaluate. Customers then compare the values received with the goals they set up in the second stage of the process.

### 2.1.2 Virtual customer environments

A Virtual Customer Environment (VCE) is an IT-enabled customer co-innovation or value co-creation platform. This web-based forum has significantly changed the nature and level of customer involvement in product innovation and its support activities. The VCEs can offer facilities like online customer discussion forums to virtual product design and prototyping centers. VCEs assist in enhancing the customer experience extensively through interactions with the company and other customers. Satish Nambisan, an influential researcher of VCEs and co-creation, investigates the different customer roles within VCE and develops an understanding of VCE mechanisms that assist these roles.

The establishment of web-based forums has changed the extent of customer involvement in product innovation and support, and made the process more feasible and cost-effective (Nambisan, 2002). VCEs play a major role in facilitating customer discussion and customer-enabled product design. The user experience in interacting with other customers and the company has been greatly enhanced because of VCEs (Nambisan, 2002; Sawhney, Verona, & Prandelli, 2005). Management literature has routinely defined five customer roles in value creation: as resource, as co-producer, as buyer, as user, and as product (Finch, 1999; Gersuny and Rosengren, 1973; Kaulio, 1998; Lengnick-Hall, 1996). These customer roles can be extended into the VCE co-creation context as: product conceptualizer, product designer, product tester, product support specialist, and product marketer (Nambisan, 2009). All five roles may not be relevant to all companies, so many of them may pursue one form of VCE that is focused on a specific goal (Nambisan, 2008). Table 2 outlines the different customer roles in VCEs, along with their desired outcomes and the way the VCEs facilitate interaction between customers and the company.

Table 2: Customer value co-creation roles in the VCE

| | Customer as Ideator | Customer as designer | Customer as tester | Customer as marketer and support specialist |
|---|---|---|---|---|
| Primary focus | -Product conceptualization | -Product design and development | -Product testing and prototyping | -Product diffusion and after-sales support |
| Desired outcome | -Product improvement<br>-New product ideas | -Input on product features and design tradeoffs<br>-Production and delivery of products/services | -Identification of design flaws<br>-Input on prototype performance | -Delivery of product support services to peer customers<br>-Diffusion of new product information<br>-Product improvement ideas |
| Examples | Microsoft, Procter & Gamble, Ducati | Peugeot, Diesel, and Adidas-Salomon | Volvo, Ducati, and Microsoft | Microsoft, IBM, and Macromedia |
| Interaction facilities | Discussion forums, Message tools, Product knowledge-base | Discussion forums, User design toolkits, Virtual prototyping tools | Discussion forums, Virtual concept testing tools, Virtual product simulations | Discussion forums, Messaging tools, Product knowledge-base |

(Source: Nambisan, 2009)

Many theoretical frameworks may offer insight into the nature of customer motivation for engaging in value co-creation. The 'uses and gratifications' approach by Katz, Blumler, and Gurevitch (1974) is relevant in value co-creation and identifies four types of benefits from participation in VCEs: (i) Cognitive benefits that relate to information acquisition and strengthening the understanding of the environment; (ii) Social integrative benefits that refer to strengthening customer ties with other customers with similar interests; (iii) Personal integrative benefits that are connected with improving the credibility, status, and confidence of the individual; and (iv) Hedonic benefits that enhance aesthetic or pleasurable experiences (Nambisan & Baron, 2009).

Customer interaction experiences in VCE are based on three fundamental contextual factors: the product context where customer interactions are related to knowledge regarding the different features of the product life cycle; the community context where interactions occur in a social setting with peer customers and members of the host company; and lastly the technology mediated environment is where the interactions are mediated by the technological infrastructure of the environment. The company may also restrict the level of product-related knowledge that can be communicated in the VCE (Franke & Shah, 2003; Fuller et al., 2006; Hertel et al., 2003; Wasko & Faraj, 2000). The three contextual factors defined above imply four components of user experiences in VCEs: the pragmatic experience is based on acquiring information about a product, its technologies and uses; the sociability experiences enables the customer to identify themselves as members of a community, and the underlying social and relational aspects of the interactions shape the user experience; the usability experience deals with the simplicity with which the customer is able to interact and perform tasks in the VCE; and the hedonic experience is created through the mental stimulation and enjoyment of the VCE interactions (Nambisan & Nambisan, 2008). Figure 1 presents the typical customer experience profiles in different type of VCEs.

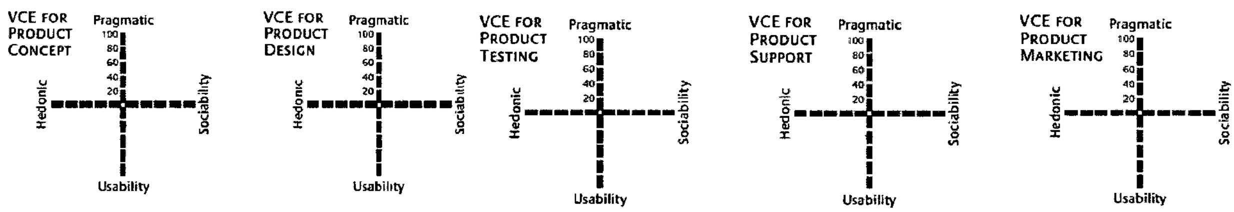



**Figure 1: Customer experience profiles in different VCEs (Source: Nambisan, 2008)**

The literature on the impact of customer involvement in innovation and value creation can be extended to the VCE context. The extensive value creation literature identifies two broad outcomes of customer participation in VCEs – customer relationship management (CRM) related outcomes and outcomes that are related to innovation. Studies have shown that customer involvement in co-production may lead to psychological outcomes, such as those related to brand loyalty and customer satisfaction (Bendapudi & Leone, 2003). Customers with a highly pragmatic experience in the VCE may feel a stronger connection to products and deem them valuable and useful. Similarly, customers with a high level of sociability experience may connect the product or service with social or community related features they can easily identify with.

Such experiences in the VCE may have future implications in customer attitudes (good and bad) towards the product and company, and for the product marketing. The other possible outcome of consumer involvement in co-creation relates to innovation, specifically innovation cost, time-to-market, and product/service quality. Customers with positive interactive experiences in the VCE may enhance their contributions to the innovation process, making an impact in cost, time, and quality of the innovation.

Nambisan's research into VCE and co-creation provides meaningful insight into strategies and practices companies can embrace to develop customer participation in VCEs. These implications are classified broadly into three main categories: design elements of VCEs, VCE integration with other parts of the innovation process, and customer relationship management.

The internet is viewed as a key element in the collaborative innovation process. Mohanbir Sawhney, Gianmarco Verona, and Emanuela Prandelli in 2005 outlined how the internet played a major role in value co-creation. The internet is viewed as a platform for customer engagement, and it enables capabilities such as enhanced reach, speed and flexibility with which companies can use to co-create value with users. Virtual environments can create important benefits in the collaborative innovation process:

- Interaction changes from one-way knowledge communication to an active dialogue that helps firms to learn from individual customers and customer communities
- Virtual environments allow companies to explore the social knowledge of the customers, which enriches the customer experience
- The size and scope of the customer audience increases since virtual environments facilitates direct as well as mediated interactions with current and potential customers (Sawhney et al, 2005)

Figure 2 displays the conceptual framework developed by Payne (2008). It highlights the customer and supplier processes and the encounter between customers and suppliers (Payne et al, 2008).

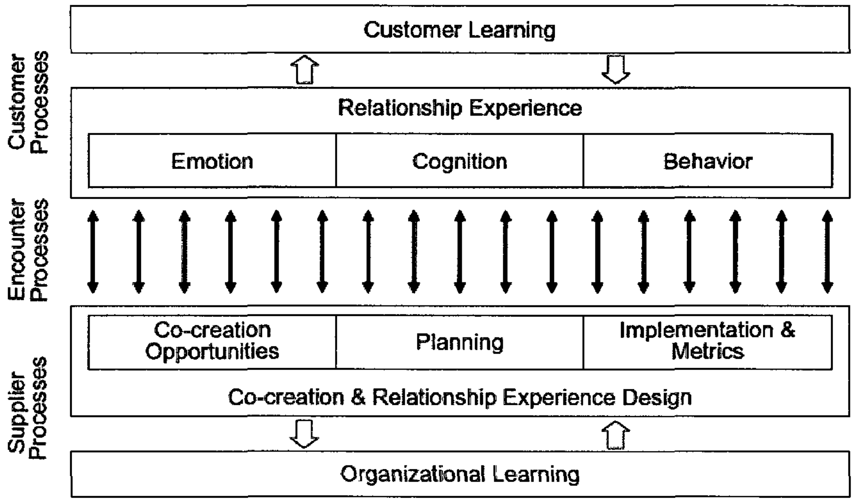



Figure 2: Conceptual value co-creation framework

### 2.1.3 New product development and innovation (NPDI)

This research stream emerges by means of terminology that oscillates between the semantics of two other paradigms – user-driven innovation (von Hippel, 2006; Bogers, Afuah & Bastian, 2010) and open innovation (Chesbrough, 2003). User-driven innovation is defined by its single firm-driven, product-centric, non-transactional and participatory approach to user involvement in the design of new products and services. However, its application of innovation toolkits (von Hippel, 2001) and innovation communities associates it to the value co-creation paradigm with its focus on customer participation platforms, personalization of market offers, multiple stakeholder interactions and access to global resources (Prahalad & Krishnan, 2008), customer-driven business models, and virtual customer experience environments. Alternately, the open innovation paradigm offers a more generic and broader vision of the innovation landscape. It articulates key mechanisms for inbound and outbound business and innovation processes, intellectual property, knowledge and resource flows used by firms to engage into a more proactive pursuit of new markets and innovations (Chesbrough, 2003).

The participatory platform nature of value co-creation practices enables a systematic positioning of customers across the entire innovation lifecycle, leading to a significant enhancement of the user-driven innovation potential. As a result, the development of value co-creation platforms is increasingly recognized a promising innovation strategy associated with an ongoing change of the nature of innovation itself (Prahalad et al., 2003; Nambisan, 2009; Romero & Molina, 2009; Midgley, 2009; Bowonder et al., 2010). The customer becomes an active stakeholder in defining the context of the interaction, including its specific personal meaning, through participation in the co-creative network (Prahalad & Ramaswamy, 2003). The personal nature of the interactive experiences enables new dimensions of value which are based on the quality of the interaction events as well as on the opportunity for customers to co-create their own unique end products, services and experiences (Franke & Schreier., 2008). These dimensions are critical for the emergence of experience innovation networks that place individuals at the heart of the co-creation experience through the development, access and dynamic reconfiguration of appropriately designed technological infrastructures (Prahalad et al., 2008). Ultimately, the value co-creation paradigm represents a specific market-driven approach to the adoption of an open innovation business philosophy.

Existing literature clearly emphasize that customer participation in value co-creation activities should impact their innovation outcomes, such as innovation cost, time-to-market, new product/service quality and development capacity (Kristensson, 2008; Prahalad & Ramaswamy, 2008; Nambisan, 2009; Midgley, 2009; Romero & Molina, 2009; Bowonder et al., 2010; Ramaswamy et al., 2010). It also reveals that firms tend to measure the performance of co-creation practices from an innovation perspective alone, neglecting side effects such as brand perception or customer-firm relationship quality, which may even exceed in value the actual innovation performance (Nambisan et al., 2007, 2008). Virtual customer environments enabling co-innovation with users can be considered as massive interactive marketing campaigns due to the sheer number of contact points with potential customers. With such advantages, proper management of collaborative innovation with consumers may become a cost-efficient or even costless way of innovating. Unfortunately, most of the existing studies focus on individual cases and there is little quantitative research focusing on the relationship between the degree of firms' involvement in value co-creation activities and their innovation related outcomes.

## 2.2 Web content data mining

Data mining has emerged as a result of the natural evolution of information technology. There is a fast-growing amount of data being collected in databanks and repositories in the technologically connected world of today; and without powerful tools to extract information from data, we are stuck in a "data-rich but information-poor" situation. Data mining techniques are used to discover patterns, trends, and anomalies in datasets; and turn them into information from which we can extract knowledge. Ultimately data mining is about solving problems by discovering patterns already present in datasets (Witten & Frank, 2005; Han & Kamber, 2006).

Data mining is not a new phenomenon, since humans have been looking for patterns for centuries in order to identify information and opportunities. The new approach for data mining in the technology-rich environment calls for a more automated approach. The internet-based data mining techniques build on the concept that it is possible to use keyword based search to build theories, and to test hypotheses in the field of management science. Some major contributions to literature behind internet based data mining were made by noted researchers such as Ferrier (2001), Opoku (2005), Hicks et al. (2006), McGinnis (2008), and Lombardi (2008).

Walter J. Ferrier published conducted research to study the characteristics and performance of different classes of competitive attacks within high tech firms. For his study, he manually conducted a search for keywords in thousands of press releases by sixteen competing firms over a period of seven years (Ferrier, 2001).

Robert A. Opoku used data mining techniques for his 2005 study on brand personality among business schools. For his research he used keywords that were representative of the dimensions of brand personality. The classification of the business schools were done according to the frequency of keywords that appeared on their websites (Opoku, 2005).

Diana M. Hicks, Dirk P. Libaers, Alan L. Porter, and David J. Schoenek use data mining in their effort to identify the technology commercialization strategies of high-tech small firms. Their data mining technique included searching for innovation terms on the firms' websites, and this method was capable of finding high tech strategies for small firms. They concluded that web data mining "holds promise" for gathering information on highly innovative firms (Hicks et al, 2006).

McGinnis' work in 2008 used web content analysis in order to classify the competitive actions of companies reliant on open source software for revenue. His work was similar to Ferrier's study of competitive actions, but he used automated data mining methods to collect the keyword frequencies on company sites. This sample of 55 keyword frequencies on 77 company website resulted in fourteen factors representative of the competitive actions of OSS firms being extracted from the data.

Stephen J. A. Lombardi, in his 2008 thesis *"Interactions between Eclipse Foundation Members and Eclipse Projects"*, used web data analysis of keywords from company sites to classify the interactions between Eclipse Foundation members and Eclipse projects. He collected 43 keywords frequencies on each of the 163 Eclipse member organizations. The keywords chosen were a suitable representation of the firms' interactions with Eclipse code; and this study resulted in four classes of interaction that correspond well with the theoretical basis of the research.

Stephen Allen's thesis, titled *"An Empirical Study of the Components of Value Co-creation"*, aims to provide an empirical identification of the components of value co-creation. His use of data mining and exploratory factor analysis resulted in a set of 29 keyword frequencies measured across 287 companies. Allen's research provided the underlying characteristics of value co-creation activities that enable value co-creation in his sample of companies. These activities are outlined by four distinguishing components that highlight the approaches used by organizations involved in value co-creation. The four components are outlined in Table 3.

Table 3: Value co-creation components defined

| Component | Factor Definition |
|---|---|
| 1. Community Driven Open Dialogue | Community forum designed to engage customers in an open dialog including networking, information sharing and learning activities with the organization, other customers or other members of the value network |
| 2. Partnerships for Resource Sharing | Partnerships enabling user access to company expertise & resources to create adaptable designs & processes aiming at reducing costs and based on trust, integrity & risk management |
| 3. Personalization through Options and Modularity | Personalization of offers through partnerships across the value network to provide choices/options enabled by product and process modularity, and integrated online services |
| 4. Co-production | Co-production of offers by user involvement in manufacturing, assembly and final beta trial activities; requiring disclosure and sharing of intellectual property |

(Source: Allen, 2009)

Tanev et al. (2010) developed on Allen's work and used web content search for keywords that indicate some form of value co-creation activity. Using the Principal Component Analysis of the data, this study was able to extract two versions of component sets that define value co-creation. The four component version seemed to have a high level of overlap between co-creation activities that was deemed unreliable. The three component version of co-creation activities seemed to distinctly outline the practices among the sample companies. The successful research methodology shared by the above research is outlined in table 4.

Table 4: Research methodology to study value co-creation

| Research Steps | Description |
|---|---|
| 1 | Develop a set of keywords that reasonably encompasses the dimensions of value co-creation |
| 2 | The value co-creation characteristics are then associated with specific keywords that are expected to be found on the company websites |
| 3 | Firms employing co-creation practices and firms that appear in value co-creation literature, with websites containing between 50 and 1.55 million webpages, are included in the sample |
| 4 | Application of factor analysis, for which around 300 firms are required to satisfy the criteria |
| 5 | The relevant data is collected using a keyword search tool based on the Google Atlas API |
| 6 | Factor analysis is done using SPSS with specific conditions |
| 7 | Factors are extracted based on the optimal factor extraction technique is SPSS |
| 8 | The meanings of the different factors are interpreted on the combination and loading of variables (keyword combination) within each of the factors |

According to the literature reviewed, data mining techniques have proven to be effective in research. Firms engaged in value co-creation tend to leverage the use of the internet in co-creation activities, and numerous tools are available to support the research.

The measured innovation performance of a given company can vary substantially depending on the innovation metrics used (Tidd et al, 2001). Multiple innovation performance indicators can offer a better understanding of innovation performance, overcoming the incompleteness of the individual measures (Saviotti & Metcalfe, 1984; Hagedoorn & Cloodt, 2003). The measurement of the innovation performance of any firm should equally include product, process and service types of innovation (Weerawardena et al, 2002). For example, Souitaris (2002) measured innovation performance by means of seven variables drawn from the OSLO Manual (Organization for Economic Co-operation and Development, 1997): the number of incrementally innovative products introduced in the last 3 years, the number of radically innovative products introduced in the last 3 years, the number of innovative manufacturing processes introduced in the last 3 years, the percentage of current sales due to incrementally innovative products introduced in the last 3 years, the percentage of current sales due to radically innovative products introduced in the last 3 years, the expenditure on innovation in the last 3 years over current sales (including R&D funding and capital expenditure on innovative manufacturing processes), and the number of patents acquired in the last 3 years.

The OSLO manual is still considered one of the most respected sources that provide practical guidelines for the measurement of firms' innovation performance. The adoption of a few of its suggested innovation metrics, or a combination of all of them, can be used in various research studies that involve determining the innovation performance of firms. However, the literature review reveals that an approach involving innovation metrics have not been applied to the study of innovation capacity of firms by means of keyword based online web search techniques. There are also no studies that consider the relationship between innovative outcomes of value co-creation practices in firms.

### 2.2.1 Search engines

Data mining requires the extraction of information from analyzing huge collections of data. There are various tools that enhance the web searches needed to collect the data for analysis. Search engines provide an essential means to collect data for information extraction. Since this research aims to evaluate the performance of two different search engines, it is imperative to know how they function.

The mode of operation for a search engine is widely known. While the exact details of the methods Google or Yahoo! search engines use is a closely guarded secret, search engines produce results based on three main operations (Thelwall, 2007): *Crawling*: Web crawling involves identifying, downloading, and saving as many potentially useful web pages as possible; *Results matching*: This is process used by a search engine to identify pages in its index to a user query; *Results ranking*: Search engines generally arrange the URLs that match a user query to maximize the probability that the relevant result appears in the first two pages.

Search engines are an important tool in the data collection process, but for reliable data extraction certain quality factors of search engines need to be considered (Lewandowski & Ho¨chsto¨tter, 2007):

- Index quality – The index quality refers to the importance of the search engine database. The important features to consider when selecting a search engine include web coverage (Gulli & Signorini, 2005), country bias (Vaughan & Thelwall, 2004), and updated indexing (Lewandowski et al, 2006).
- Result quality – The derivatives of classic retrieval tests are applied to determine the quality of the search engine results. However, it needs to be determined which measures are to be applied and whether new measures are needed to satisfy the uniqueness of search engines and their users (Lewandowski, 2007).
- Quality of search features – Search engines should offer a sufficient set of features and a sophisticated selection of query languages. These features should be reliable and produce quality results.

- Search engine usability – The user interface of the search engine should enable the users to conduct searches in an effective and efficient way.

Previous research concerning value co-creation by Lombardi (2008), Allen (2009), Allen et al (2009), and Tanev et al (2010) used Google's Ajax API. Google's popularity in the search engine market is astounding, accounting for more than 80% of the market share in Europe, and between 50%-60% of all queries in the United States is made on Google (Lewandowski, 2008). The website www.marketshare.hitslink.com provides a visual representation of the search engine market in figure 3.

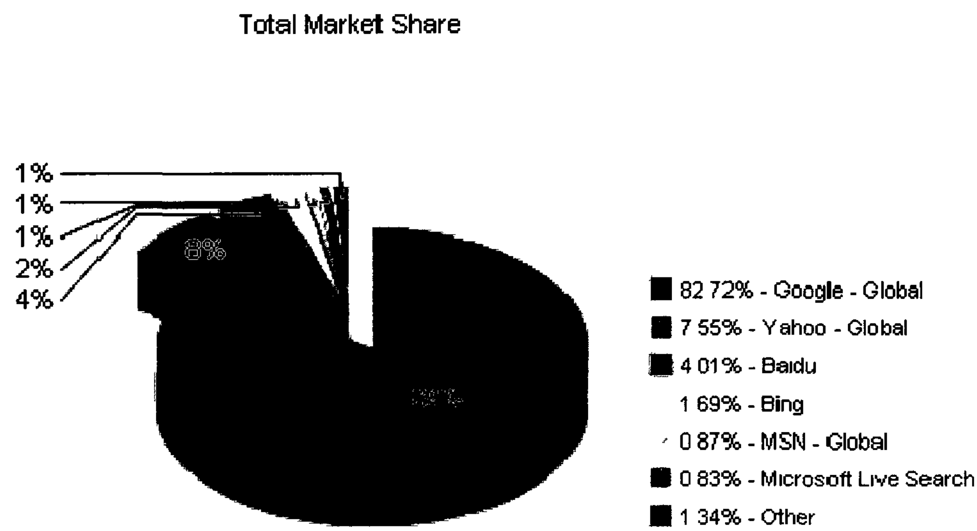


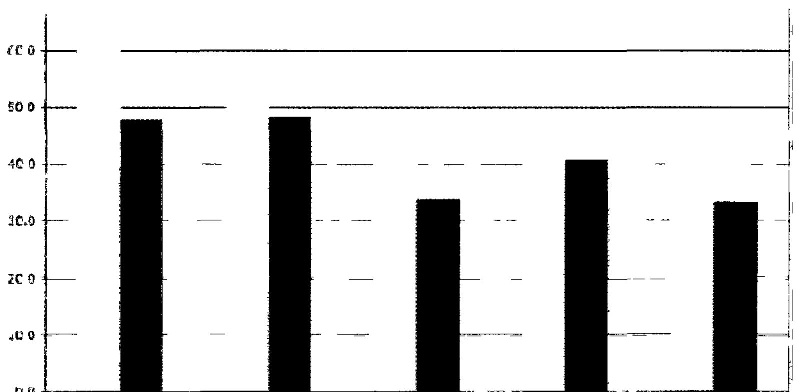

**Figure 3: Search engine market share - 2008 through September 2010**

While Google is the most popular search engine, studies show that the Yahoo! search engine produces comparable, and sometimes superior, results for user queries (Lawendowski, 2008; McCown, 2007; Thelwal, 2007). The following sections compare the attributes of the Google and Yahoo! search engines.

The indexable Web (Selberg, 1999) is the portion of the internet that is considered for indexing by the major search engines. Gulli and Signorini (2005) estimate the size of the indexable web to be around 11.5 billion pages (as of January 2005). The authors provide as estimate of the relative size of the major search engines, as well as their overlaps. The results of the research show that the search engines cover about 9.3 billion web pages of the 11.5 billion total pages. The findings also conclude that Google covers roughly 76.2% of the sampled data set, while Yahoo! covers about 69.3% and MSN and Ask/Teoma cover 61.9% and 57.6% respectively. One important note about the findings is that the overlap for Google and Yahoo! is more than 56%, which would account for search results unique to either engine for certain queries (Gulli & Signorini, 2005). While Google is the largest search engine, there are pages that are not covered in the Google index that other search engines do cover.

One major concern for this research is the relevancy of the results obtained through the search engines. For the keyword search queries to provide meaningful information, the results have to be relevant. Dirk Lewandowski, in his 2008 paper *"The retrieval effectiveness of web search engines: considering results descriptions"*, discusses the retrieval effectiveness of major search engines, including Yahoo! and Google.

For his research Lewandowski (2008) aims to assess the quality of results from five search engines, namely Google, Yahoo!, MSN, Ask.com, and Seekport. The research shows that Yahoo! provides the highest number of relevant results at 48.5%, followed by Google at 47.9%. In terms of result descriptions, Google provided about 60.2% relevant descriptions to Yahoo!'s 52.8 %. The relevant results and result descriptions are compared in figure 4. It can be concluded that while Google has a slight advantage when it comes to relevancy of results descriptions, the relevancy of results are very similar for Yahoo! and Google, and either search engine can sufficiently provide relevant results for user queries (Lawandowski, 2008).



| | Google | Yahoo | MSN | Ask | Seekport |
|---|---|---|---|---|---|
| Descriptions | 60.2 | 52.8 | 45.1 | 48.2 | 40.6 |
| Results | 47.9 | 48.5 | 33.9 | 40.7 | 33.5 |

**Figure 4: Comparison of search engines on the ratio of relevant results vs. relevant results descriptions**

Google, Yahoo!, and MSN all provide an application programming interfaces (API) for their search engines, along with the web user interface (WUI). Frank McCown and Michael L. Nelson (2007) attempt to find out whether the search engine APIs produce search results from an index that is smaller or older than the WUI results. Prominent scholars have discovered that search engine API results are consistently different from WUI results, and this distinction is reflected in the forums and listserves that cater to API communities. McCown and Nelson determine that the API results were significantly different from the WUI results. In the case of Google and Yahoo!, the top 100 API results vary as much as 14% from the WUI results. When considering the top 10 results, Google API produces results that are 20% different from the WUI; and Yahoo! API produces results that are 14% different that the WUI. The difference can be attributed to the API indexes being smaller for both search engines. In light of this information, researchers are cautioned when generalizing information obtained from using search engine API results to that using the engine's WUI. Since web 'scraping' for data is restricted by search engines, using their APIs is the optimal way to collect data for research and analysis (McCown & Nelson, 2007).

Aaron Wall (2006) offers a comparison of search engines from a marketing perspective, the highlights of which are outlined in table 5.

Table 5: Search engine algorithm comparison

| | Yahoo! | Google |
|---|---|---|
| Page content | Yahoo! is believed to place greater weight on page content than the Google search engine. Yahoo! search results include a lot of Yahoo! content, which makes it beneficial for them to bias their search towards commercial websites. | Google is the search engine leader in distinguishing real text patterns from ones that are spam. Google uses an algorithm with an aggressive filter that screens out documents that contain targeted phrases for search engine optimization. Google continues to scan millions of books that should help them differentiate manipulative text from real text patterns. |
| Query processing | Yahoo! search is heavily weighted towards common search terms in the query. The search engine produces results by matching text. The results are also biased towards commercial sites. | Google search produces results by matching the intent of a query, rather than matching the text directly. Common words like 'and' and 'to' may be de-weighted compared to other search terms to provide a better discrimination value. Google results are usually heavily biased towards informational sites. |
| Page vs. site | The search engine looks at both links to a page and links to a site to determine the relevancy of a page. Pages on newer sites rank well on Yahoo! as long as they have some descriptive inbound links. Domains do need necessarily need to have a lot of trust built up to rank well. | Google looks for natural link growth over time to determine the importance of web pages. As a result, Google tends to rank pages from older websites higher. Sites need to build up a certain amount of trust to rank competitively in Google. |

(Source: Thelwall, 2006)

The Google search engine has been previously utilized in order to provide a quantitative analysis of value co-creation activities (Allen, 2009; Allen et al, 2009; Tanev et al, 2010). However, the literature on search engines and their algorithms show that no single search engine indexes all pages of a website, and drawing hypotheses on value co-creation activities based on data from a single source would be unwise. Prior research shows that search engine indexes have a limited overlap between them and display international biases in coverage; it is therefore recommended that research should be done using a different search engine in addition to Google (Gulli & Signorini, 2005; Lewandowski, 2008). Search results provided by Google and Yahoo! are not significantly different, with Yahoo! performing as well as Google for most search queries. With a distinctive index for its search engine, Yahoo! is able to provide a new perspective for value co-creation activities that is equally valid as the one outlined by Google.

## 2.3 Exploratory factor analysis

Exploratory factor analysis (EFA) is primarily used to reduce data into a smaller set of summary variables, or to understand the underlying theoretical structure of some phenomenon. Recently EFA has been efficiently used as a tool to reduce massive amounts of data collected through web data mining techniques by extracting key dimensions of how small and medium firms carry out business functions (Hicks et al, 2006; Lombardi, 2008; McGinnis, 2008). Richard B. Darlington describes factor analysis as a means to discover simple patterns in the relationships among variables (Darlington, 2008). The main principle behind factor analysis is to closely examine these relationships to assess if they can be defined using smaller sets of variables called factors.

Recent texts and papers provide an excellent outlook on using the statistics software SPSS for factor analysis (Darlington, 2008; Field, 2009; Costello & Osborne, 2005; Reinard, 2006). Anna B. Costello and Jason W. Osborne make a few observations on how to perform EFA using SPSS. They particularly stress that there are differences between methods of factor analysis, and informed choices must be made to obtain the optimal conclusion for research (Costello & Osborne, 2005). Andy Field points out that the Principal Component Analysis (PCA) method is most favourable when the sample used for analysis is the population and findings cannot be extrapolated beyond the sample. PCA presents an excellent method when for exploring the data and applying the findings to the sample (Field, 2009).

## 2.4 Lessons learned

A thorough examination of the literature provided the following major insights.

Value co-creation has been increasingly attracting interest from companies and researchers. The rapidly expanding body of literature tends to focus on qualitative, case-based research and there is a need for new quantitative studies that further explore the specific outcomes of value co-creation.

The study of innovation-related outcomes of value co-creation is an emerging research stream. There are an increasing number of research studies that indicate firms' involvement in co-creation practices enhances innovation. However, the literature is lacking in research that explores the relationship between the firms' degree of involvement in co-creation practices and their innovation capacity.

The literature indicates that the use of multiple performance indicators can offer an enhanced understanding of the innovation performance of a firm. The innovation indicators should include product, service, and process types of innovation metrics.

Keyword-based data mining, coupled with exploratory factor analysis, can be effectively used to explore the commercialization strategies of firms (Hicks et al.2006; Lombardi, 2008). This approach has also been efficiently used to identify activities that outline value co-creation (Allen, 2009; Tanev et al, 2010).

Search engines provide an automated solution to collecting the online data that outline company strategies of value co-creation (Allen, 2009; Tanev et al, 2010). However, different search engine companies that have their own algorithms that are commercial secrets and prone to constant change. There is a lack of studies that compare analytical results based on information provided by different search engines.

# 3. Research strategy and method

This section covers the research strategy and the specific research steps undertaken for the research. The first section provides an overview of the methodology and the objectives this research aims to fulfill. The second section details the specific actions taken in order to obtain and analyze the data.

## 3.1 Research strategy

The objective of this research is to use a research methodology pioneered by Stephen Allen (2009) in order to identify the key components of value co-creation practices, using data from two different sources; and to examine the relationship between the degree of firms' involvement in specific co-creation components and firms' innovation capacity. This study is exploratory in nature and aims to examine (i) how time and difference between search engines would influence the composition of co-creation activities, and (ii) the potential relationship between co-creation and innovation.

The research methodology is based on the work of Allen (2009), who uses data mining techniques and website content analysis to develop an empirically driven method to identify distinct value co-creation activities used by firms. This approach follows the keyword-based data mining research of Ferrier (2001), Lombardi (2008), and McGinnis (2008), and uses the information extracted by EFA techniques to classify value co-creation activities currently employed by firms. To retain consistency with Allen's work and to accurately study the changes in value co-creation approaches within firms over time, this study utilizes the same keywords and same sample of firms. In addition, the data collection process is done using the same Google search API as in his research. There are two key differences from Allen's work. Firstly, the methodology in this research uses a second data set generated by the Yahoo! search engine.

Secondly, the methodology introduces a new research step which focuses on examining the relationship between co-creation and innovation.

A key objective for value co-creation is to produce innovative products, services, and processes by collaborating with customers and other partners in its value chain. The proposed measurement metric, to measure the innovation capability of firms online, is defined by the frequency of comments regarding new products, services, and processes on the firms' websites. The keyword set that outlines the innovation metric is founded on insights from the Oslo Manual ($3^{rd}$ edition, 2005), which provides guidelines for collecting and interpreting innovation data. An additional measure included in this research methodology assesses the innovation capabilities of firms' value co-creation activities by including the number of new products, services, and processes. The conventional innovation metric takes into account the actual number of innovations from a company instead of simply relying on the frequency of online comments which can be misleading. This extra step in the research continues the work done by Tanev et al (2010) by assuming to strengthen the innovation metric with a more accurate measurement of the innovation capacity of firms.

## 3.2 Research steps

The research method for this thesis is similar to the approach by Allen (2009), Lombardi (2008), and Tanev et al (2010). The steps undertaken to produce meaningful results are outlined in table 6 and described in the following subsections.

Table 6: Research steps

| Research Step | Activity |
|---|---|
| 1 | **Select keywords** – keywords sets are believed to be a comprehensive representative of value co-creation characteristics and innovation features |
| 2 | **Select sample** – the sample consists of firms known to engage in co-creation activities |
| 3 | **Gather data** – the Yahoo! and Google search engine APIs are used to collect keyword frequencies for the value co-creation and innovation keyword sets |
| 4 | **Validate data** – the co-creation data collected is validated using Kaiser-Meyer-Olkin measure of sampling adequacy, Bartlett's test of sphericity, and the determinant of the correlation matrix |
| 5 | **Factor extraction** – a variety of extraction methods on SPSS are used to analyze the search results and the optimal technique will be used to determine the numbers of factors to be extracted |
| 6 | **Value co-creation components** – The extracted factors comprise a set of variables that are used to explain the value co-creation components |
| 7 | **Firm rankings** – the firms' value co-creation activities are ranked according to the extracted components and cluster analysis performed to identify the firms with the same value co-creation approach |
| 8 | **Associate co-creation with innovation** – linear regression is used to explore the association between value co-creation components and innovation within firms |
| 9 | **Results and insights** – the results from the Google set analysis is linked to Allen's value co-creation factors to obtain a new model; evaluation of the Yahoo! and Google data is done to validate current results; insights are drawn from the results. |

### *Research step 1: Selecting keywords*

The search keywords to explore value co-creation activities have been adopted from Allen (2009). The set of keywords are to represent the breadth of activities of value co-creation, so that the web content data mining and exploratory factor analysis techniques would be able to discover the multiple co-creations aspects across the sample of companies. Allen's set of keywords were collected through an exhaustive review of co-creation literature and a contextual search of corporate web pages and developer's forums of four organizations, namely Facebook, Android/Open Handset Alliance, Taiwan Semiconductor Manufacturing Corporation (TSMC), and the online game Second-Life. While his initial set of more than 170 keywords were logically reduced to produce the comprehensive set of 29 keywords, the concern of other relevant terms being omitted remained an issue. In order to explore other possible keywords, this research uses the following keyword search tools to verify the completeness of the 29 keyword value co-creation set: 'Google sets' is a Google labs application that automatically creates a set of items from a few examples; the website www.synonymlab.com provides a synonym search tool that uses the Google tilde (~) search operator to provide a list of related synonymous terms; the www.gorank.com ontology lookup tool provides related keywords by checking the top 1000

Google results for related keywords. This additional step is assumed to discover other relevant value co-creation terms that should be included in the keyword set.

The expansive initial set of keywords was also condensed using Boolean logic. Operators such as "AND' and "OR" were used to combine keywords with similar definitions. Since the Yahoo! and the Google search APIs support the operators, a smaller keyword set provides a more comprehensive representation of value co-creation aspects. Table 7 provides an example of the Boolean operators as used in the keyword search terms.

Table 7: Boolean operator use within keywords

| Boolean operator | Keyword example | Function |
|---|---|---|
| AND | Integrated AND online AND services | The AND operator combines all the keywords and the results show pages with all the keywords |
| OR | Customer OR user | The OR operator will ensure that webpages with either of the keywords are displayed |
| "..." | "intellectual property" | The double quotation enables the search engine to look for the exact term |

(Source: Allen, 2009)

The online innovation metric, to measure the innovative capacity of the firms through the frequency of online comments regarding new products, services, and processes, consists of a composite keyword that aims to measures a firm's capability through the frequency of its online comments regarding new products, services, and processes. Boolean operators are used to create one complex keyword that synthesizes relevant innovation terms that would appear on a firm's website. An example of a keyword, as it would appear in the search, is described in table 8, showing its structure, source, and context.

Table 8: Keyword example

| Keyword: (customer OR user) AND (suggest OR suggestion OR input OR request OR demand) | | |
|---|---|---|
| Keywords | Source | Context |
| customer OR user | Research | Qualifier keywords used to remove pages that do not describe activities concerning customers or users |
| suggest<br>suggestion | Facebook Developers' Forum | "You can suggest your ideas through the suggestion form" |
| input | TSMC online newsletter | "TSMC will be even more diligent in seeking customer input..." |
| request | Secondlife Grid Developers' Forum | "To submit a request to participate in the RegAPI program..." |
| demand | Lattice Semiconductor News Release | "...and we are rapidly expanding our sales organizations to meet very strong customer demand" |

The primary set of keywords used for this research was organized by Allen (2009). A number of search engine tools are used to verify the completeness of the keyword list. The objective of this search is to attempt to find relevant terms synonymous with value co-creation keywords already defined. Table 9 shows the synonyms and related keywords produced by a sample keyword set from Allen's list.

Table 9: Synonyms and related keywords based on original keyword sample

| | **Google sets**: produces set of related words | **Google synonyms**: (www.synonymlab.com) provides synonyms and 'sub-synonyms' |
|---|---|---|
| customer OR user: | Customer, user, product, customerid, client, consumer, rma, view cart, invoices, open orders, close out order history | **User:** manual, users **Customer:** consumer, customers Related sub-synonyms **Technical:** software, tech, engineering, technology **Service:** support, care, to others, community, definition |
| learn OR learning: | Learning, learn, education, language, elearning, online, school, teaching, life, programming, tutorials, free | **Learn:** tutorial **Learning:** learner, courses, training, English, learning English, games, learning games Related sub-synonyms: **Know:** known, knows **Study:** education, survey, report, studies |

Table 9 provides words related to the keyword 'customer OR user AND learn OR learning'. In the context of value co-creation this keyword refers to the learning aspect of co-creation. The related words provided by the search engine tools do not provide any additional terms that are relevant in the co-creation context. Since the search for the relevant terms were conducted using tools based on the Google engine, it is unlikely that any additional terms are required to convey the learning aspect. An analysis of the other 28 keywords did not provide terms that were relevant for the purposes of this research. Ultimately Allen's original 29 keywords were used to measure the various aspects of co-creation. These keywords are listed in table 10.

Table 10: Value co-creation keywords

| | |
|---|---|
| V1 | (customer OR user) AND (suggest OR suggestion OR input OR request OR demand) |
| V2 | internal AND (expertise OR resource) |
| V3 | (customer OR user) AND (dialog OR dialogue OR communicate OR communication OR conversation OR contact OR feedback OR call OR interact OR "information exchange" OR "information sharing" OR connect OR access OR engage) |
| V4 | (customer OR user) AND (communities OR community OR network OR networking OR forum) |
| V5 | (customer OR user) AND (learn OR learning) |
| V6 | (customer OR user) AND experience |
| V7 | (customer OR user) AND (test OR trial OR beta) |
| V8 | ecosystem OR "value network" OR "value constellation" OR "multiple partners" OR "external contributor" OR "external source" |
| V9 | integrated AND (online AND services) |
| V10 | (product OR process) AND (modularity OR modular OR module) |
| V11 | (product OR process OR service) AND (evolution OR evolve) |
| V12 | (customer OR user) AND (produce OR assemble OR manufacture) |
| V13 | (customer OR user) AND (options OR choice OR choose) |
| V14 | (design OR process) AND (flexibility OR flexible OR adaptable) |
| V15 | lease OR rent OR license OR "self serve" OR "self service" |
| V16 | cost AND (reduce OR reduction OR saving) |
| V17 | (customer OR user) AND (negotiate OR negotiation) |
| V18 | (customer OR user) AND (risk OR manage OR management OR control OR assess OR reduce OR reduction OR potential OR exposure) |
| V19 | (customer OR user) AND (IP OR "intellectual property") |
| V20 | (customer OR user) AND language AND translation |
| V21 | (customer OR user) AND address AND concern |
| V22 | (customer OR user) AND (survey OR review OR voting OR vote OR rate OR rating) |
| V23 | trust OR honesty OR integrity |
| V24 | (customer OR user) AND (dashboard OR statistics) |
| V25 | customer AND (partnerships OR interaction OR relationship OR participate OR participation OR activity OR action) |
| V26 | customization OR customize OR customized OR personalize OR individualize OR "add feature" OR "added feature" |
| V27 | simulation OR simulate OR model OR modeling OR "virtual world" OR "reference design" OR "reference flow" OR "demo application" OR toolkit OR tutorial OR SDK OR "software development kit" |
| V28 | (customer OR user) AND (disclose OR inform OR disseminate OR reveal) |
| V29 | (customer OR user) AND (suggest OR suggestion OR input OR request OR demand) |

***Innovation measures***

The innovation measure for this research has to be representative of the innovative capabilities of the firms in the sample. It should be pointed out that this measure is not a traditional metric for innovation since it does not take into account the actual number of new products, services, and processes introduced by the firm. The metric to measure the perception of innovativeness is used in this research to explore how it relates to the value co-creation components. It is outlined in table 11. A traditional measure of innovation, which does account for the number of innovations from a firm, is also incorporated into this research. The linear regression for both measures of innovation against the value co-creation components should provide viable results as to which measure is able to define how co-creation influences the innovative process in firms.

Table 11: Innovation keyword

| |
|---|
| new AND product OR service OR process OR application OR solution OR feature OR release OR version OR launch OR introduction OR introduce OR "new product" OR "new service" OR "new process" OR "new solution" OR "product launch" |

(Source: Tanev et al, 2010)

***Research step 2: Selecting the sample***

The unit of analysis for this research is the website of a company that is known to be active in value co-creation. The sample of companies for this research is the same as that used by Allen (2009), in order to explore how value co-creation efforts have changed within a similar sample of firms. Allen's sample of companies were selected based on the company employing value co-creation as part of their key business strategy, and the organization website having between 50 and 1.55 million pages, in order to reduce the influence of outliers. The 287 companies include 61 Open Source Software (OSS) companies identified by McGinnis (2008), 140 Eclipse firms sourced from Lombardi's (2008) research, and 86 companies identified by Allen (2009) through his research of value co creation. The high level of co-creation activity in the sample is validated by the fact that OSS companies are presumed to enable customer co-creation, while companies from the OSS-driven Eclipse ecosystem are representative of many co-creation aspects. Allen's 86 companies are extracted from case studies within the literature and exemplify co-creation activities. This research re-examines the 287 companies to evaluate whether they qualify for the sample. The Google and Yahoo! APIs search the sample sites for the keywords, the sites which do not contain between 50 and 1.55 million pages are excluded, as are sites that do not produce any results for the keyword search. This might indicate that the disqualified companies went through some drastic change due to the recession, or their co-creation activities are not uncovered by the extent of value co-creation keywords identified, or that the firms are not active in co-creation as defined by the context of this research.

The sample contains a considerable number of firms that are involved in OSS projects. The significance of OSS in co-creation is entrenched in the observation that co-creation platforms operate similarly to the participation architectures that exist in firm-driven open source platforms. It is also important to note that the value co-creation building blocks outlined by the DART framework of Prahalad and Ramaswamy (2004) is represented sufficiently by OSS firms. OSS firms are observed to: (i) be active contributors to OSS development participation platforms and often engage in dialogue with contributors who are most often end users, (ii) provide access, often unrestricted, to their source code, internal resources, and development processes, (iii) share IP management and development risk with members of its value network, and (iv) demonstrate a high level of transparency through newsgroups and forums geared towards co-development of the product. The OSS firms were vetted to ensure that a main part of their revenue is generated by OSS projects, and sourced from the Canadian Companies Capabilities Directory of OS Companies database maintained by Industry Canada (http:// www. ic.gc.ca/app/ccc/sld/cmpny.do?lang=eng&profileId=1861&tag=025049) and other sites such as Network World (www.networkworld.com) and blogs (http://1vc.typepad.com/soaring_on_ ridgelift/2006/08/open_source_com.html) that list OSS companies that are defined by criteria similar to the requirements of this research.

An important thing to note about the sample in this research is the structure of the websites being searched by the APIs. While the search engine APIs are adept at searching through subdirectories within websites, they need to be instructed to search through the subdomains for each website. Some firms use multiple subdomains to separate resources for developers and online communities. While the main site http://www.android.com is included in the sample, the subdomain http://developer.android.com is not searched for keywords since it does not appear as a subdirectory. One proposed solution to this subdomain issue is to remove the 'http://www' from all the websites in the research sample. This theory was tested and a search of the websites without the 'http://www' results in more pages in some cases, and similar results for other websites. This leads to the conclusion that not all companies are using subdomains to organize their sites. While the method with the 'http://www' prefix seems to produce a more thorough search for keywords, this research aims to extend the methodology outlined by Allen (2009) and therefore the company sample, in its original form, is the only one considered for this research. This ensures that the new value co-creation model is based on a similar platform as that of Allen, and avoids the influence of outliers. This alternative approach, however, is significant to the research methodology to warrant further research.

***Research step 3: Gathering data using Google and Yahoo! search engines***

The data acquisition process was achieved with a keyword search tool. Following the research methods of Opoku (2005) and Hicks et al (2006), the keyword search tool looks for the presence of select keywords as indicators of value co-creation. The search tool was developed on the Eclipse Java platform and uses the Google search API, and later modified to include the Yahoo! search API. Similar to the IFAPI tool developed for Hicks' research at Georgia Tech, the API search tool takes text files containing user specified keywords and website URLs as input to automate the search of millions of web pages for the desired keywords (Hicks et al, 2006). The automated search outputs an excel file with two sets of values: (i) the number of web pages under each URL, and (ii) 29 sets of keyword counts normalized by the number of pages under the specific URL. The search tool was altered such that searches were restricted to pages in English only, since the sample firms operate internationally and pages in other languages would influence the resulting keyword counts significantly. As defined in the 'Selecting Keywords' subsection the Boolean operators in the keyword structures are supported by both Yahoo! and Google search APIs, and these must be fulfilled to increase the keyword 'hits'. If the AND operator is used all the keywords must be present on the page for a 'hit'. Similarly, if the OR operator is used, either of the words separated by the operator must be present for a 'hit'. It is important to note that the search tool may return normalized keyword counts greater than 100, but this only indicates that the keyword is counted numerous times on a single page. The user interface for the keyword search tool is shown in figure 5.

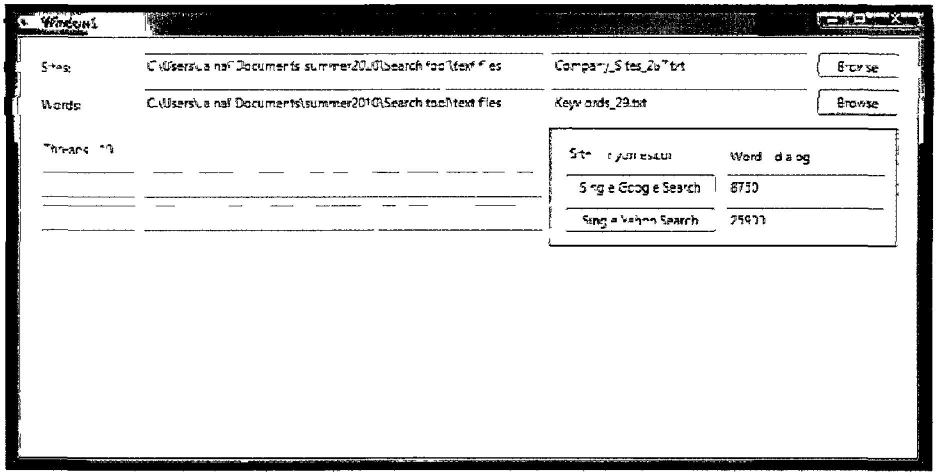



Figure 5: Keyword search tool user interface

***Research step 4: Preparing data for factor analysis***

Successful exploratory factor analysis requires the data to meet specific conditions for correlation, sampling adequacy, and statistical significance (Field, 2009). The data from the keyword search tool is analyzed using SPSS to verify whether it meets the criteria. Field (2009) suggests performing a few checks on the data set to ensure that factor analysis will produce relevant information.

The reliability of factor analysis depends on the size of the sample. While scholars suggest the use of at least 300 cases for factor analysis, between 5-10 participants per variable is regarded as acceptable. If the factor has four or more loadings greater than 0.6, it is argued to be reliable and the size of the sample is insignificant.

An alternative method of determining if the data is appropriate for factor analysis is the Kaiser-Meyer-Olkin measure of sampling adequacy (KMO). The KMO represents the ratio of the squared correlation between variables to the squared partial correlation between variables, and can be calculated for individual and multiple variables. A value close to 1 signifies that the correlation pattern is relatively compact, and factor analysis is likely to produce distinct and reliable factors. Field (2009) suggests that KMO value should ideally be higher than 0.5. A KMO value approaching 1 is considered superb and the data is highly expected to provide dependable factors.

An important part of factor analysis is to look at the intercorrelation between variables to ensure that they make sense. The solution to the factor analysis is unlikely to have any meaning unless the variables are sensible. SPSS is used to obtain the correlation matrix between the 29 keyword variables. The variables pose an issue if:

- The correlations are not high enough The Bartlett's test is used to examine whether the variables correlate badly with the other variables, and if the correlation matrix is significantly different from an identity matrix. This test is preferred to be significant (have a significance value of less than 0.5) for the factor analysis to be appropriate for the data set.
- The correlations are too high – It is advisable for factor analysis to avoid extreme multicollinearity (variables that are very highly correlated) and singularity (variables that are perfectly correlated). The presence of highly correlated variables can be checked by looking at the determinant of the matrix. A simple heuristic in this case is that the determinant should be greater than $1X10^{-5}$.

In both cases the data set is improved by removing the variables. Table 12 summarizes the validation criteria the variable correlation matrix must meet for factor analysis.

Table 12: Validation criteria for data set correlation matrix

| Metric | Measure |
|---|---|
| Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy | Must be greater than 0.5, preferably between 0.8 and 1 |
| Bartlett's test of sphericity | Should have a significance value less than 0.5 |
| Determinant of correlation matrix | Should have a value greater than $1X10^{-5}$ |

***Research step 5: Extracting factors***

Extracting factors from data to uncover the relationships between variables is the main purpose of exploratory factor analysis. In the context of this research, the correlation between 29 keyword variables produces subsets of large correlation coefficients measuring aspects of the same underlying dimension. These dimensions are known as factors. The factor extraction process essentially reduces the large data set of interconnected variables to a smaller set of factors in order to explain the maximum amount of common variance in a correlation matrix using the smallest number of explanatory constructs (Field, 2009, p. 629).

There are multiple methods for uncovering factors from data, but ultimately the choice of method depends on the purpose of the analysis. If the purpose of the analysis is to explore data and apply the findings to the sample, then one of the methods of factor analysis and extraction available in SPSS will suffice. SPSS supports the following kind of factor extraction methods:

- Principal Component Analysis (PCA)
- Common factor analysis methods based on:
    - Unweighted least squares
    - Generalized least squares
    - Maximum likelihood
    - Principal axis factoring
    - Alpha factoring
    - Image factoring

There is little information about the strengths and weaknesses of the mentioned extraction methods, often only in obscure references. Lombardi (2008) used a common factor analysis method for his research, selecting to use Alpha factoring over PCA. Allen (2009) used PCA after evaluating the different factor extraction methods, claiming PCA provided "the cleanest component loading table with relatively high loading values and minimal cross loadings as well as the best fit with prevalent theory" (Allen, 2009, pp. 43-44). PCA and principal axis factoring are the preferred methods for factor extraction, and they tend to produce similar solutions (Field, 2009). However, it is argued that PCA is more of a data reduction method than a factor analysis method, and its use should be restricted in favor of a true factor analysis technique. Other scholars have successfully used PCA and advocate its use (Costello et al, 2005).

This thesis utilizes SPSS for factor extraction. Multiple techniques were used on the data to make an educated decision regarding which technique best defines the co-creation components. Scholars seem to agree that the selection of the method should be an informed decision based on the data available and the desired outcome. There is consensus among the literature that a sample size of roughly 300 will provide similar results across different techniques. With a sample size of 287 sites for this research, all the methods are expected to produce similar results. The optimal method selected for factor extraction should be the one that provides the cleanest structure and best fits the theory (Reinard, 2006).

Factor analysis literature suggests that two main methods are to be explored to determine the number of factors to retain:

- Kaiser criterion: This states that all factors with eigenvalues over 1 should be retained. It is based on the idea that the eigenvalues represent the amount of variation explained by a factor and than an eigenvalue of 1 is representative of a substantial amount of variation.
- Scree plot: This graphs each eigenvalue against the factor with which it is associated. The graph shows the relative importance of each factor. An examination of the eigenvalues provides a threshold for factor selection.

Andy Field (2009) provides some guidelines on how to determine the number of factors to retain using the mentioned criteria. All factors with eigenvalues greater than 1 should be retained; or, with a sample size exceeding 200, the scree plot provides a reliable criterion for factor selection.

The way the extracted factors are interpreted can be improved through factor rotation. Rotation maximizes the loading of each variable on one of the extracted factors while minimizing the loading on all other factors. This makes the relationship between the variables and factors much clearer. SPSS supports three kinds of orthogonal rotations, namely Varimax, Quartimax, and Equamax; while Direct Oblimin and Promax are oblique rotation methods. In case of unrelated factors, an orthogonal rotation such as Varimax should be used. If the factors are expected to correlate either oblique methods of rotation can be selected (Field, 2009).

***Research step 6: Interpreting value co-creation components***

The criteria surrounding the factor extraction methods should provide a clear outlook on how many components are to be retained, and the variables associated with them. Using an approach similar to Allen (2009), value co-creation components are deduced based on the combination of variables in the component, coupled with research insights from literature. The value co-creation components are interpreted and analyzed according to the loadings of the variables, which determine the importance of the variable within the component. Reinard (2006) recommends the exclusion of variables that have a high degree of cross loadings, while other scholars believe that variables with loadings above 0.4 should be considered for component interpretation. Taking into account various suggestions from literature, and to ensure the components comprise of relevant co-creation aspects, the components should have a minimum of three variables with loadings greater than 0.5. The components were also to have clean variables, without 'cross loadings' with other components. This ensures that the components are defined by exclusive variables, and there is no misconception about the value co-creation aspects.

***Research step 7: Ranking firms***

The process of factor extraction provides interpreted value co-creation components consisting of three or more variables, each with their own loadings. These loadings enable the component score for each firm in the sample to be calculated so their level of co-creation activity can be assessed. The component data is normalized using SPSS, so the component score for the firms can be used to determine their level of activity for a specific component. The firms can then be classified according to the combinations of the value co-creation components they are considered to be active in. This approach is an extension of the method used by Allen (2009). Allen classified the firms using non-normalized component scores, with the top 75% of cases in each variable being active. This research normalizes the component scores and a single threshold value can be used to distinguish firms active in the different components of value co-creation.

The component scores for each firm is an important contribution to this research. The scores outline the value co-creation levels in the sample firms, and can be further analyzed to learn which activities are more pertinent in value co-creation. The sample firms are also selected from either Open Source Software based companies (Eclipse and OSS) or classified as General companies. The factor scores can provide a summary of how these distinct groups are engaged in co-creation. The aggregate scores of individual components for the firms are particularly relevant in this research, since they can be used to rank the firms by their degree of activity in any one component. The different combinations of co-creation components are seen as the firms' approach to co-creating value with other players in the value chain, based on the assumption that any specific combination of the components are deliberate and part of the company strategy.

***Research step 8: Examine the relationship between value co-creation and innovation***

One of the main contributions of this research is the analysis of innovation against value co-creation components as outlined by an updated data set. Allen (2009) proposed that the co-creation components be evaluated against some measure of innovation of the firms in the sample. Tanev et al (2010) provide a study of the relationship between value co-creation and innovation, with their innovation measure being the frequency of online comments regarding new products, processes, and services. The innovation metrics for this research consist of an updated data set, with recently collected frequencies of online comments about new products, processes, and services; as well as a more conventional measure of innovation, where the number of new products, processes, and services for each of the companies in the sample were obtained, in collaboration with some students from the University of Southern Denmark, through a thorough search of the news releases from the last three years. This conventional innovation metric is assumed to provide a more accurate perspective on how innovative the sample companies are.

SPSS is used to conduct a linear regression to determine the relationship between value co-creation components and the two distinct measures of innovation. The co-creation components are normalized for the purpose of this regression, as are the counts for new products, processes, and services. The regression is run for the value co-creation components as the dependent variable while the innovation measure is the independent variable. The conventional innovation measure provided independent counts for new products, processes, and services, and this is used to run linear regressions on value co-creation components against the innovation measures individually. This provides a unique look into how co-creation components are related to company innovation on a distinct level. The linear regression method is also used to examine the innovative capabilities of OSS firms and non-OSS firms, and explore how the different component combinations contribute to the innovation process in these two distinct groups.

***Research step 9: Discussion of results***

The value co-creation components are interpreted according to the variable loadings, and this allows the classification of firms active in co-creation with the use of two discrete search engine data. The Google data allows the value co-creation components to be linked to Allen's research (2009). The comprehensive set of value co-creation keywords are believed to represent an overall outlook of firms' value co-creation activities. With a similar methodology the value co-creation components that emerge from this research can be linked to the four components outlined by Allen.

The Yahoo! search engine data was collected at the same time as the updated Google data. The same research methodology was applied to extract value co-creation components from this data set. This enables the evaluation of the Google components against the Yahoo! co-creation components. The two search engines index websites in their own unique way, and the use of two distinct search engines provide a concrete method to examine whether two discrete data sets produce similar value creation components.

The linear regression between the innovation measures against value co-creation components provided by the Google and Yahoo! data will explore the relationship, if any, between co-creation activities and the innovation capacity of the firms. The regression results state the relationship between the co-creation activity and innovation capacity of the firms with innovation data (number of new products, processes, and services) in the sample.

# 4. Research Results

The research results are summarized in this section. Section 4.1 outlines the selection of firms for the research sample, and provides an explanation of their groupings. Section 4.2 describes the data collection process, and the methods used to refine the data for factor analysis. Section 4.3 explains how the data is prepared for factor analysis. Section 4.4 presents the factor extraction method selected to obtain the value co-creation components, and the techniques that define the number of factors to be retained. Section 4.5 shows the rates of component use for the two data sets. Finally, section 4.6 presents the linear regression results between value co-creation components and innovation measures.

## 4.1 Research Sample

The initial sample for this thesis consisted of the 287 firms used by Stephen Allen for his thesis research. This sample was re-examined to ensure that the company websites had not changed significantly. The data collection for the sample firms, using both search engines, was completed on the 9$^{th}$ of September, 2010.

The screening process resulted in a total of 262 firms for the research samples for both the Google and Yahoo! search engines. They were drawn from previous research (Lombardi, 2008; McGinnis, 2008; Allen, 2009) and are a reliable representation of firms engaged in value co-creation. The entire research sample can be found in Appendix A.

The remaining firms can be divided into the following types: (i) Open Source Software (OSS) firms are ones that draw a majority of their revenue from open source projects, (ii) Eclipse (ECL) members are firms associated with the business ecosystem driven by the Eclipse Foundation, (iii) OSS companies that are also associated with the Eclipse Foundation (OSS+ECL), and (iv) firms that are not associated with the Eclipse Foundation and are not involved in open source projects (GEN). The breakdown of the firms in the Google research sample is outlined in table13.

Table 13: Google research sample firm breakdown

| Type of firm | Firms of this type | Percentage of total |
|---|---|---|
| OSS | 69 | 26.34 |
| ECL | 105 | 40.08 |
| GEN | 63 | 24.05 |
| OSS+ECL | 25 | 9.54 |
| Total number of firms: 262 | | |

The breakdown of the types of firms in the Yahoo! research sample is displayed in table 14.

Table 14: Yahoo! research sample firm breakdown

| Type of firm | Firms of this type | Percentage of total |
|---|---|---|
| OSS | 66 | 25.19 |
| ECL | 107 | 40.84 |
| GEN | 63 | 24.05 |
| OSS+ECL | 26 | 9.92 |
| Total number of firms: 262 | | |

## 4.2 Data Collection

The keyword search tool allowed the value co-creation and innovation performance data to be gathered efficiently for both research samples. The search tool is designed to use the Google and the Yahoo! search APIs to look for the keywords specified within the websites of the research sample. The keywords and the websites for the sample are specified in text files, and the tool returns two values (i) the number of pages (the size) under the specified firm URL, and (ii) the number of times each keyword appears under the firm URL, normalized by the 'size' of the website. The descriptive statistics for the data collected from the final sample is presented in table 15.

Table 15: Descriptive statistics for the Google keyword search on the sample firms

| Keyword | N | Min | Max | Mean | Std. Deviation |
|---|---|---|---|---|---|
| customer+OR+user+dialog+OR+dialogue+OR+communicate+OR+communication+OR+conversation+OR+contact+OR+feedback+OR+call+OR+interact+OR+%22information+exchange%22+OR+%22information+sharing%22+OR+connect+OR+access+OR+engage | 262 | 0 | 96.08 | 19.02 | 17.07 |
| customer+OR+user+communities+OR+community+OR+network+OR+networking+OR+forum | 262 | 0 | 112.40 | 15.87 | 16.39 |
| lease+OR+rent+OR+license+OR+%22self+serve%22+OR+%22self+service%22 | 262 | 0 | 86.30 | 8.86 | 11.56 |
| customer+OR+user+cooperate+OR+cooperation+OR+collaboration+OR+partnership | 262 | 0 | 50.88 | 4.20 | 7.14 |
| customer+OR+user+suggest+OR+suggestion+OR+input+OR+request+OR+demand | 262 | 0 | 90.76 | 9.40 | 11.15 |
| internal+exertise+OR+resource | 262 | 0 | 48.34 | 1.79 | 4.00 |
| customer+OR+user+risk+manage+OR+management+OR+control+OR+assess+OR+reduce+OR+reduction+OR+potential+OR+exposure | 262 | 0 | 55.45 | 2.98 | 5.70 |
| customer+OR+user+IP+OR+%22intellectual+property%22 | 262 | 0 | 54.98 | 4.55 | 7.15 |
| customer+OR+user+learn+OR+learning | 262 | 0 | 57.72 | 6.52 | 8.74 |
| product+OR+process+OR+service+evolution+OR+evolve | 262 | 0 | 47.81 | 1.93 | 4.72 |
| customer+OR+user+experience | 262 | 0 | 37.53 | 5.36 | 6.43 |
| customer+OR+user+test+OR+trial+OR+beta | 262 | 0 | 113.18 | 9.70 | 12.24 |
| integrated+online+services | 262 | 0 | 47.39 | 1.88 | 4.46 |
| simulation+OR+simulate+OR+model+OR+modeling+OR+%22virtual+world%22+OR+%22reference+design%22+OR+%22reference+flow%22+OR+%22demo+application%22+OR+toolkit+OR+tutorial+OR+SDK+OR+%22software+development+kit%22 | 262 | 0 | 99.55 | 14.74 | 13.44 |
| product+OR+process+modularity+OR+modular+OR+module | 262 | 0 | 50.24 | 3.28 | 6.18 |
| customer+OR+user+produce+OR+assemble+OR+manufacture | 262 | 0 | 50.24 | 2.36 | 4.54 |
| customer+OR+user+options+OR+choice+OR+choose | 262 | 0 | 67.40 | 7.14 | 7.86 |
| design+OR+process+flexibility+OR+flexible+OR+adaptable | 262 | 0 | 56.87 | 4.67 | 7.09 |
| customer+partnerships+OR+interaction+OR+relationship+OR+participate+OR+participation+OR+activity+OR+action | 262 | 0 | 130.23 | 5.33 | 10.71 |
| cost+reduce+OR+reduction+OR+saving | 262 | 0 | 56.87 | 3.46 | 5.60 |
| customer+OR+user+survey+OR+review+OR+voting+OR+vote+OR+rate+OR+rating | 262 | 0 | 109.30 | 8.33 | 10.98 |
| trust+OR+honesty+OR+integrity | 262 | 0 | 50.71 | 3.89 | 6.84 |
| customer+OR+user+disclose+OR+inform+OR+disseminate+OR+reveal | 262 | 0 | 16.11 | 1.20 | 1.96 |
| customer+OR+user+dashboard+OR+statistics | 262 | 0 | 23.76 | 2.38 | 3.37 |
| customization+OR+customize+OR+customized+OR+personalize+OR+individualize+OR+%22add+feature%22+OR+%22added+feature%22 | 262 | 0 | 70.63 | 4.90 | 7.65 |
| customer+OR+user+negotiate+OR+negotiation | 262 | 0 | 15.17 | 0.44 | 1.43 |
| ecosystem+OR+%22value+network%22+OR+%22value+constellation%22+OR+%22multiple+partners%22+OR+%22external+contributor%22+OR+%22external+source%22 | 262 | 0 | 49.76 | 1.71 | 5.07 |
| customer+OR+user+language+translation | 262 | 0 | 22.75 | 0.80 | 2.19 |
| customer+OR+user+address+concern | 262 | 0 | 36.49 | 0.83 | 2.64 |
| Valid N (listwise) | 262 | | | | |

The search tool supports both the Google and the Yahoo! search APIs, and produces an excel page with the keyword frequency data for each firm in the sample. SPSS is then used on this data set to produce a correlation matrix to analyze how well the 29 variables correlate with each other. After the variables are screened the Google data set was left with 25 keyword variables, while the Yahoo! data set had 18 variables remaining for factor extraction.

## 4.3 Preparing data for factor analysis

The variable screening process produces a final data set that has to meet specific criteria to ensure that the consequent factor analysis will produce relevant information. SPSS is used to create correlation matrices for the final variable sets for Google and Yahoo! to examine the specific condition necessary for data validation. The data can be validated using the determinant of the correlation matrix, the Kaiser-Meyer-Olkin measure of sampling adequacy (KMO), and Bartlett's test of sphericity.

Table 16 presents the correlation matrix analysis of the Google data set. The matrix determinant value of $1.73X10^{-5}$ indicates that the quality of the data is adequate for extraction of factors. The KMO value of 0.893 is in the 'great' range and the data is expected to provide meaningful factors that define value co-creation. The Bartlett's test significance value of 0.00 shows that there is an adequate degree of correlation and factor analysis is appropriate for the data.

Table 16: Correlation matrix analysis for Google data

| | | Required | Actual |
|---|---|---|---|
| Determinant of correlation matrix | | >1x10^-5 | 1.74 x10^-5 |
| Kaiser-Meyer-Olkin measure of sampling adequacy | | >0.5 (preferably closer to 1) | 0.893 |
| Bartlett's test of sphericity | Approx. Chi – Square | | 2760.04 |
| | df. | | 300 |
| | Sig. | <0.5 | 0.00 |

Table 17 provides a similar correlation matrix analysis of the Yahoo! data set.

Table 17: Correlation matrix analysis for Yahoo! data

| | | Required | Actual |
|---|---|---|---|
| Determinant of correlation matrix | | >1x10^-5 | 1x10^-3 |
| Kaiser-Meyer-Olkin measure of sampling adequacy | | >0.5 (preferably closer to 1) | 0.854 |
| Bartlett's test of sphericity | Approx. Chi – Square | | 1739.27 |
| | df. | | 153 |
| | Sig. | <0.5 | 0.00 |

The matrix correlation analysis of the Yahoo! data shows values that validate the data set for factor extraction. The determinant value of 0.001 proves the data is free from multicollinearity, the KMO value of 0.854 indicates that the data is expected to produce meaningful value co-creation components, and the significance value of 0.00 for the Bartlett's test confirms that factor extraction is appropriate for the data set.

## 4.4 Factor Extraction

Research using exploratory factor analysis typically uses either Principal Component Analysis (PCA) or Common Factor Analysis (CFA) to extract factors from data sets. All seven extraction methods were used on SPSS to determine the optimal method to extract value co-creation components from the two data sets. With the Google and Yahoo! sample sizes relatively close to 300 firms, similar results were achieved from PCA and CFA methods.

The most favorable results were obtained using the PCA method. The CFA methods predictably provided results similar to PCA, but drawing from value co-creation literature insights, the outcome from the PCA method fit the value co-creation criteria best. Using PCA, the variable consistently had loadings close to 0.6, which is the necessary threshold towards obtaining reliable factors. Finally, the PCA method resulted in variable groupings with minimal cross loadings and made sense within the value co-creation context. The decision to use the PCA method was based on the component results befitting the theory of value co-creation insights collected from literature.

The factor analysis methods on SPSS provide a few ways to ascertain how many components to extract from the data. Statistics literature generally point to two methods to determine the number of components to extract. The first uses Kaiser's criterion, which states that all components with eigenvalues over 1 should be retained, and the second uses the Scree plot, which shows the relative importance of the components. Analyzing the communality table, the extracted factors are not greater than 0.7, so the Kaiser's criterion condition does not apply. Therefore the number of components to be extracted was determined by using the Scree plot method.

The Scree plot method confirms a three-component outline of value co-creation for both data sets. The presence of three components above the tangent line for the Google data set in figure 6 indicates that three components can explain the co-creation activities sufficiently. A similar Scree plot indicated that three components are to be extracted from the Yahoo! data. To conclusively verify the three-component set, the PCA method was rerun to extract more factors. Unfortunately, the resulting components consisted of variable loadings too low to convincingly define value co-creation, or were composed of variables that did not accurately match the value co-creation context.

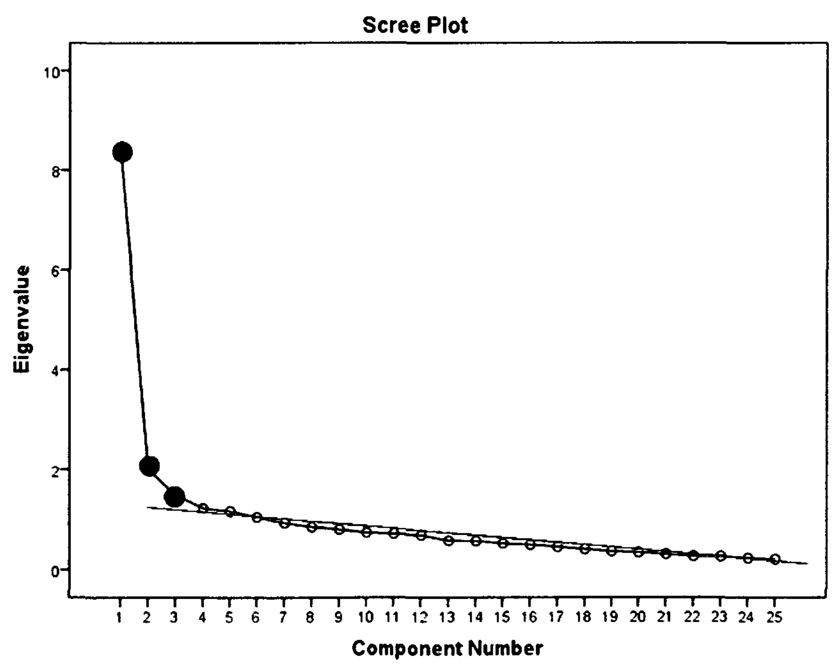



Figure 6: Scree plot for Google data set

A second method to verify the three-component model is necessary since the Kaiser and the Scree plot criteria are prone to providing different solutions. The eigenvalues associated with each component represent the variance explained by that particular linear component. SPSS also produces the eigenvalues in terms of the percentage of the variance explained. For the Google and Yahoo! data sets, we can see from tables 18 and 19 the percentage of total variance explained by the extracted components.

Table 18: Total variance explained by Google components

| Component | Initial Eigenvalues | | | Extraction Sums of Squared Loadings | | | Rotation Sums of Squared loadings | | |
|---|---|---|---|---|---|---|---|---|---|
| | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % |
| 1 | 8.344 | 33.376 | 33.376 | 8.344 | 33.376 | 33.376 | 4.946 | 19.784 | 19.784 |
| 2 | 1.999 | 7.998 | 41.374 | 1.999 | 7.998 | 41.374 | 3.562 | 14.25 | 34.034 |
| 3 | 1.491 | 5.962 | 47.336 | 1.491 | 5.962 | 47.336 | 3.326 | 13.303 | 47.336 |

Table 19: Total variance explained by Yahoo! components

| Component | Initial Eigenvalues | | | Extraction Sums of Squared Loadings | | | Rotation Sums of Squared loadings | | |
|---|---|---|---|---|---|---|---|---|---|
| | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % |
| 1 | 5.992 | 33.287 | 33.287 | 5.992 | 33.287 | 33.287 | 3.488 | 19.378 | 19.378 |
| 2 | 1.547 | 8.596 | 41.883 | 1.547 | 8.596 | 41.883 | 3.017 | 16.764 | 36.141 |
| 3 | 1.319 | 7.326 | 49.209 | 1.319 | 7.326 | 49.209 | 2.352 | 13.068 | 49.209 |

The table for the Google set demonstrates that 47.34% of the total variance is explained by three components. Similarly 49.21% of the total variance is explained by three components from the Yahoo! data set. This confirms the finding that the three components extracted from the updated data can sufficiently define the aspects of value co-creation.

The factor analysis using PCA resulted in three components being extracted from each data set. Reinard (2006) recommends that the keyword variables within the components should have loadings greater than 0.6, with cross loading values below 0.4; while Costello and Osborne (2005) suggest that a minimum loading of 0.32 is appropriate for variables.

For the purposes of this research, keyword variables had to be absent of cross-loadings, and a minimum threshold value of 0.48 was selected. This ensured that the keyword variables had high loading values, an attribute essential for the reliability of the components.

Table 20: Rotated component variables and loadings for Google data

| | Keyword Variables | Loading |
|---|---|---|
| **Component 1** | | |
| V11 | (customer OR user) AND experience | 0.775 |
| V1 | (customer OR user) AND (dialog OR dialogue OR communicate OR communication OR conversation OR contact OR feedback OR call OR interact OR "information exchange" OR "information sharing" OR connect OR access OR engage) | 0.773 |
| V9 | (customer OR user) AND (learn OR learning) | 0.723 |
| V5 | (customer OR user) AND (suggest OR suggestion OR input OR request OR demand) | 0.708 |
| V4 | (customer OR user) AND (cooperate OR cooperation OR collaboration OR partnership) | 0.698 |
| V19 | Customer AND (partnerships OR interaction OR relationship OR participate OR participation OR activity OR action) | 0.651 |
| V2 | (customer OR user) AND (communities OR community OR network OR networking OR forum) | 0.625 |
| V7 | (customer OR user) AND risk AND (manage OR management OR control OR assess OR reduce OR reduction OR potential OR exposure) | 0.538 |
| V17 | (customer OR user) AND (options OR choice OR choose) | 0.523 |
| **Component 2** | | |
| V15 | (product OR process) AND (modularity OR modular OR module) | 0.733 |
| V6 | Internal AND (expertise OR resource) | 0.628 |
| V8 | (customer OR user) AND (IP OR "intellectual property") | 0.617 |
| V3 | lease OR rent OR license OR "self serve" OR "self service" | 0.598 |
| V28 | (customer OR user) AND language AND translation | 0.493 |
| **Component 3** | | |
| V16 | (customer OR user) AND (produce OR assemble OR manufacture) | 0.645 |
| V22 | trust OR honesty OR integrity | 0.609 |
| V23 | (customer OR user) AND (disclose OR inform OR disseminate OR reveal) | 0.604 |
| V21 | (customer OR user) AND (survey OR review OR voting OR vote OR rate OR rating) | 0.603 |
| V29 | (customer OR user) AND address AND concern | 0.589 |
| V20 | cost AND (reduce OR reduction OR saving) | 0.586 |

(Method: PCA, Rotation: Varimax with Kaiser normalization)

Table 20 displays the rotated components extracted from the Google data set. The table shows that the three Google components each have at least four keyword variables with loadings of around 0.6. This indicates that the components are reliable, and the components dependably represent meaningful aspects of value co-creation.

Table 21: Rotated component loadings and variables for Yahoo! data

| | Keyword Variables | Loading |
|---|---|---|
| **Component 1** | | |
| V1 | (customer OR user) AND (dialog OR dialogue OR communicate OR communication OR conversation OR contact OR feedback OR call OR interact OR "information exchange" OR "information sharing" OR connect OR access OR engage) | 0.698 |
| V11 | (customer OR user) AND experience | 0.690 |
| | Keyword Variables | Loading |
| V9 | (customer OR user) AND (learn OR learning) | 0.689 |
| V19 | Customer AND (partnerships OR interaction OR relationship OR participate OR participation OR activity OR action) | 0.685 |
| V4 | (customer OR user) AND (cooperate OR cooperation OR collaboration OR partnership) | 0.601 |
| V2 | (customer OR user) AND (communities OR community OR network OR networking OR forum) | 0.573 |
| **Component 2** | | |
| V16 | (customer OR user) AND (produce OR assemble OR manufacture) | 0.684 |
| V17 | (Customer OR user) AND (options OR choice OR choose) | 0.659 |
| V20 | cost AND (reduce OR reduction OR saving) | 0.597 |
| V21 | (customer OR user) AND (survey OR review OR voting OR vote OR rate OR rating) | 0.572 |
| V5 | (customer OR user) AND (suggest OR suggestion OR input OR request OR demand) | 0.556 |
| **Component 3** | | |
| V7 | (customer OR user) AND risk AND (manage OR management OR control OR assess OR reduce OR reduction OR potential OR exposure) | 0.665 |
| V14 | simulation OR simulate OR model OR modeling OR "virtual world" OR "reference design" OR "reference flow" OR "demo application" OR toolkit OR tutorial OR SDK OR "software development kit" | 0.614 |
| V18 | (design OR process) AND (flexibility OR flexible OR adaptable) | 0.595 |

(Method: PCA, Rotation: Varimax with Kaiser normalization)

Table 21 displays the rotated components extracted from the Yahoo! data set. The keyword variable loadings for the Yahoo! components are similar to the three Google components. Since the Yahoo! set had fewer variables for analysis, these components each have at least three variables with loadings of around 0.6, indicating that the components extracted are meaningful. The average value of the variable loadings for all three Yahoo! components is also greater than 0.6, ensuring that the components extracted from the Yahoo! data set also accurately represent various aspects of value co-creation.

A preliminary evaluation of the variables included in the Google and Yahoo! components is presented is table 22. The highlighted variables demonstrate that 11 variables were common between the two data sets.

Table 22: Google and Yahoo! component similarity

| | Google component variables (sorted by loading values) | Yahoo! component variables (sorted by loading values) |
|---|---|---|
| Component 1: | V11, V1, V9, V5, V4, V19, V2, **V7**, **V17** | V1, V11, V9, V19, V4, V2 |
| Component 2: | V15, V6, V8, V3, V28 | **V16**, **V17**, **V20**, **V21**, V5 |
| Component 3: | **V16**, V22, V23, **V21**, V29, **V20** | **V7**, V14, V18 |

## 4.5 Rate of component use

The rate of use of the two distinct value co-creation component sets is analyzed to determine the types of firms adopting these approaches. The Google and the Yahoo! samples consist of the different types of firms in very similar numbers. The rate of use of the different components for the Google and Yahoo! search engines is expected to provide a comparable adoption rate.

### 4.5.1 Rate of Google component adoption

Figure 11 presents an adoption chart for the three Google components. The mean scores for the three components are displayed according to their adoption levels for OSS and GEN type of firms.

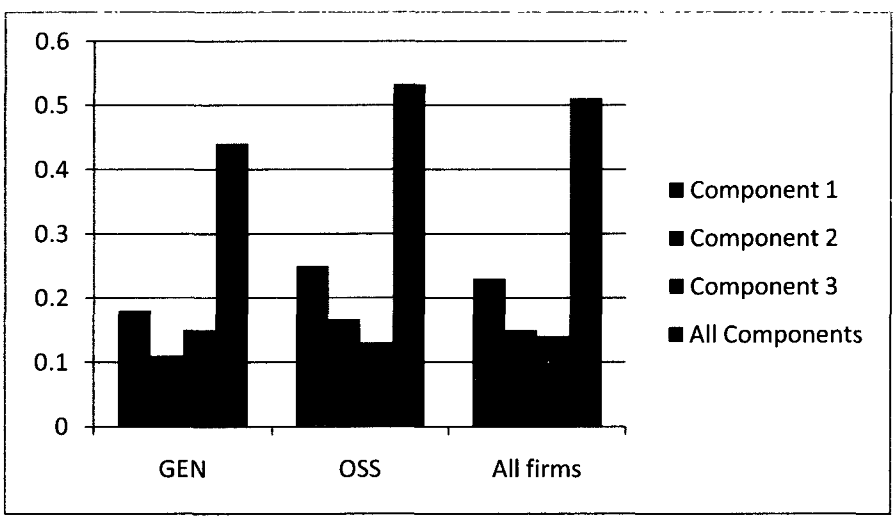



**Figure 7: Adoption rate of Google components**

### 4.5.2 Firms active in all three Google components

The variable loadings within the components enable the calculation of component 'scores' for all the firms in the sample to determine their level of co-creation activity. The firm scores for each component is the product of their keyword variable score weighted by the specific variable loading within each component. Figure 8 shows the firms active in all value co-creation components according to the Google data.

| | Firm | Type | Comp1 | Comp2 | Comp3 |
|---|---|---|---|---|---|
| 1 | http://www.queue.acm.org/ | ECL | 0.45 | 0.29 | 0.61 |
| 2 | http://www.atmel.com/ | GEN | 0.25 | 0.45 | 0.41 |
| 3 | http://www.brocade.com/ | OSS+ECL | 0.49 | 0.89 | 0.44 |
| 4 | http://www.convergys.com/ | GEN | 0.87 | 0.46 | 0.4 |
| 5 | http://www.freescale.com/ | ECL | 0.35 | 0.61 | 0.62 |
| 6 | http://www.intervoice.com/ | OSS+ECL | 0.5 | 1 | 0.3 |
| 7 | http://www.latticesemi.com/ | GEN | 0.33 | 0.54 | 0.43 |
| 8 | http://www.openmakesoftware.com/ | OSS+ECL | 0.46 | 0.33 | 0.44 |
| 9 | http://www.perforce.com/ | ECL | 0.27 | 0.29 | 0.25 |
| 10 | http://www.progress.com/ | OSS+ECL | 0.28 | 0.28 | 0.36 |
| 11 | http://www.radview.com | OSS | 0.53 | 0.26 | 0.28 |
| 12 | http://www.tibco.com/ | ECL | 1 | 0.25 | 0.43 |
| 13 | http://www.vyatta.com | OSS | 0.38 | 0.33 | 0.25 |

Figure 8: Firms active in all three components – Google

According to the normalized component scores, thirteen firms are active in all value co-creation components. Unsurprisingly, most of the firms are involved in OSS projects.

### 4.5.3 Rate of Yahoo! component adoption

Figure 9 presents the graphical representation of the rates of adoption of the Yahoo! components by different types of firms.

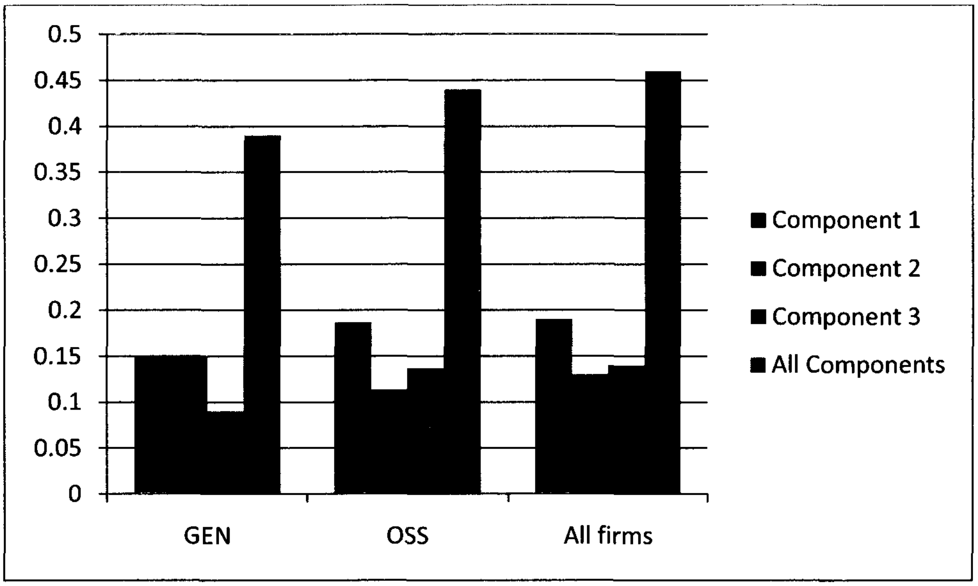



Figure 9: Adoption rate of Yahoo! components

### 4.5.4 Firms active in all three Yahoo! components

Figure 10 shows the firms that are active in all value co-creation components according to the Yahoo! data.

| | Firm | Type | Comp1 | Comp2 | Comp3 |
|---|---|---|---|---|---|
| 1 | http://www.altera.com/ | GEN | 0.34 | 0.37 | 0.27 |
| 2 | http://www.blackducksoftware.com/ | ECL | 0.44 | 0.35 | 0.56 |
| 3 | http://www.diyhappy.com/ | GEN | 0.4 | 1 | 0.41 |
| 4 | http://www.innovations-software.com/ | ECL | 0.45 | 0.32 | 1 |
| 5 | http://www.lynuxworks.com/ | ECL | 0.43 | 0.39 | 0.57 |
| 6 | http://www.mks.com/ | ECL | 0.45 | 0.37 | 0.54 |
| 7 | http://www.oracle.com/ | ECL | 0.38 | 0.34 | 0.42 |
| 8 | http://www.parasoft.com/ | ECL | 0.29 | 0.49 | 0.32 |
| 9 | http://www.pentaho.com | OSS | 0.35 | 0.31 | 0.3 |
| 10 | http://www.radview.com | OSS | 0.44 | 0.39 | 0.29 |
| 11 | http://www.sybase.com/ | ECL | 0.42 | 0.33 | 0.77 |
| 12 | http://www.tibco.com/ | ECL | 0.7 | 0.38 | 0.64 |

Figure 10: Firms active in all three components - Yahoo!

## 4.6 Linear Regression

Once the components for both search engine data sets have been extracted, a linear regression test can be conducted in order to explore the relationship between aspects of value co-creation and innovation capacity of firms. The regression attempts to discover the assumed connection between value co-creation and innovation in all firms, firms involved in OSS projects, and 'general' type of firms. The innovation capacity of the firms is defined by two discrete approaches. The first innovation metric, following Hicks et al (2006) and Ferrier's (2001) data mining approach, assumes to outline the innovation capability of a firm by measuring the frequency of comments regarding new products, services and processes on the firm's website. The second, more traditional, innovation metric is an actual count of the number of products, services, and processes. This information is collected from the news release pages of the sample firms' websites for the last three years. Due to time restrictions, the data set can be attributed as a joint effort with the students of the Product Development and Innovation program at the University of Southern Denmark[1]. The students in the program were familiar with the objective, relevance, and deliverables of this research and could extract relevant data concerning new products, processes, and services from the sample firms. The data from this research was combined with the one sourced from the Product Development and Innovation program to come up with three innovation related measures. The linear regression tests are conducted on both innovation measures to relate the outcome of the 'perception' metric to the more traditional metric.

Linear regression tests the relationship between value co-creation and innovation. The stepwise regression method on SPSS was used to analyze the value co-creation – innovation relationship, with the distinct innovation metric as the independent variable and the value co-creation components as the dependent variables. The variable data is normally distributed for the regression process. The Kurtosis and Skewness statistics of the variables indicate whether the variables are normally distributed. Once the Kurtosis and Skewness statistics were less than twice the value of their standard errors for all the variables, the stepwise linear regression was carried out to determine the value co-creation – innovation relationship for the Google and

Yahoo! data sets. Finally, the regression method explores the innovation capacity of different types of firms against the value co-creation components.

The regression results are outlined in the following sections. The R-square value from the SPSS analysis determines the explanatory power of the relationships between value co-creation and innovation. In case of multiple linear regressions, the adjusted R-square value is used to indicate the explanatory power.

### 4.6.1 Linear regression results for Google data

The linear regression results for the Google data set is presented below, in tables 23 and 24. The descriptive statistics for the normalized variables are presented in the initial tables, and the regression analysis to test the relationship between value co-creation components and innovation follow. While the online innovation metric is not an accurate measure of innovation within a firm, it is expected to estimate how innovative the firm is perceived to be.

Table 23: Descriptive statistics of normalized Google components

| | Mean | Std. Deviation | Skewness | | Kurtosis | |
|---|---|---|---|---|---|---|
| | | | Statistic | Std. Error | Statistic | Std. Error |
| $(\text{Component }1)^{1/3}$ | 3.669 | 1.10132 | -0.124 | 0.150 | -0.102 | 0.300 |
| $(\text{Component }2)^{1/3}$ | 2.127 | 0.79416 | 0.115 | 0.150 | 0.054 | 0.300 |
| $(\text{Component }3)^{1/3}$ | 2.191 | 0.71354 | 0.245 | 0.150 | -0.057 | 0.300 |

C1: Experience driven mutual learning; C2: Self service enabled by resources; C3: Customer involvement in co-production

Table 24: Descriptive statistics for normalized innovation metrics for Google

| | Mean | Std. Deviation | Skewness | | Kurtosis | |
|---|---|---|---|---|---|---|
| | | | Statistic | Std. Error | Statistic | Std. Error |
| $(\text{New products})^{1/5}$ | 1.0510 | 0.71911 | -0.180 | 0.156 | -0.277 | 0.310 |
| $(\text{New services})^{0.27}$ | 0.7478 | 0.74790 | 0.434 | 0.156 | -0.569 | 0.310 |
| $(\text{New processes} + 0.0005)^{1/3}$ | 0.4675 | 0.57438 | 0.985 | 0.156 | -0.580 | 0.310 |
| Total Innovation | 2.266 | 1.48700 | 0.209 | 0.156 | -0.769 | 0.310 |
| $(\text{Online metric} + 1)^{1/4}$ | 2.1314 | 0.43632 | 0.129 | 0.151 | 0.484 | 0.300 |

It is important to note that the value co-creation components and innovation metrics were normalized to allow linear regression. While the Skewness and Kurtosis statistics are not less than twice the standard error in every case, the skewness statistic was adjusted to be between the required range of -1 and 1.

***Regression results for online innovation metric vs. Google components***

The following tables provide the regression analysis results for the measurement of the online innovation metric against the value co-creation components. The model summary table consists of the statistics required to assess the explanatory power of the regression, while the coefficient table outlines the regression models.

Table 25: Regression results for Google data (online innovation metric)

| Model Summary | | | | | |
|---|---|---|---|---|---|
| | Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
| All firms | 1 | .456 | .208 | .205 | .389 |
| | 2 | .483 | .233 | **.277** | .384 |
| GEN firms | 1 | .268 | **.072** | .057 | .515 |
| OSS firms | 1 | .540 | .291 | .287 | .331 |
| | 2 | .554 | .307 | **.299** | .328 |

| Coefficients | | | | | | | |
|---|---|---|---|---|---|---|---|
| Model | | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
| | | | B | Std. Error | Beta | | |
| All firms | 1 | (Constant) | 1.468 | .084 | | 17.465 | .000 |
| | | Comp 1 | .181 | .022 | .456 | 8.246 | .000 |
| | 2 | (Constant) | **1.399** | .086 | | 16.233 | .000 |
| | | Comp 1 | **.139** | .026 | .351 | 5.367 | .000 |
| | | Comp 2 | **.104** | .036 | .190 | 2.909 | .004 |
| GEN firms | 1 | (Constant) | 1.631 | .175 | | 9.334 | .000 |
| | | Comp 3 | .163 | .075 | .268 | 2.174 | .034 |
| OSS firms | 1 | (Constant) | 1.378 | .092 | | 14.951 | .000 |
| | | Comp 1 | .209 | .023 | .540 | 8.971 | .000 |
| | 2 | (Constant) | **1.324** | .095 | | 13.937 | .000 |
| | | Comp 1 | **.175** | .028 | .451 | 6.147 | .000 |
| | | Comp 2 | **.083** | .040 | .153 | 2.089 | .038 |

C1: Experience driven mutual learning; C2: Self service enabled by resources; C3: Customer involvement in co-production

The regression analysis in table 25 provides the relationship models between value co-creation and the online innovation metric. The significant values for each model are highlighted in bold. For this research the model with the higher explanatory power is considered to define the relationship better. In the case of 'GEN' firms, the R-square value is below the minimum threshold value of 0.1, which indicates that there is no statistical significance between the co-creation component in this model and the perception metric. The models for the perception metric are summarized in the table below.

Table 26: Online innovation metric vs. value co-creation for Google data

| Type of firms | Model | Linear regression model | R-square | Relationship |
|---|---|---|---|---|
| All firms | 2 | Online metric = 1.4 + (0.14*C1) + (0.1*C2) | 0.277 | Online metric ~ C1 & C2 |
| GEN firms | N/A | | | |
| OSS firms | 2 | Online metric = 1.32 + (0.18*C1) + (0.08*C2) | 0.299 | Online metric ~ C1 & C2 |

C1: Experience driven mutual learning; C2: Self service enabled by resources; C3: Customer involvement in co-production

***Regression results for new products, services, and processes vs. Google components***

The regression results for the conventional innovation metrics against the Google value co-creation components are presented. This allows an examination of the relationship between value co-creation and the distinct innovative capabilities of firms. Similar to table 25, the R-square values define the explanatory power of the regression models. Some of the R-square values fall below the 0.1 threshold but were pretty close to be considered.

Table 27: Regression results for Google data (Conventional innovation metrics)

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **New Products vs. Value co-creation** | | | | | | | |
| **Model Summary** | | | | | | | |
| | Model | | R | R Square | Adjusted R Square | Std. Error of the Estimate | |
| All firms | No statistical significance | | | | | | |
| GEN firms | 1 | | .306 | **.094** | .078 | .858 | |
| OSS firms | No statistical significance | | | | | | |
| **Coefficients** | | | | | | | |
| Model | | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig |
| | | | B | Std. Error | Beta | | |
| All firms | No statistical significance | | | | | | |
| GEN firms | 1 | (Constant) | **.369** | .241 | | 1.532 | .131 |
| | | Comp 2 | **.294** | .119 | .306 | 2.471 | .016 |
| OSS firms | No statistical significance | | | | | | |
| **New Services vs. Value co-creation** | | | | | | | |
| **Model Summary** | | | | | | | |
| | Model | | R | R Square | Adjusted R Square | Std. Error of the Estimate | |
| All firms | No statistical significance | | | | | | |
| GEN firms | 1 | | .402 | **.161** | .147 | .764 | |
| OSS firms | No statistical significance | | | | | | |
| **Coefficients** | | | | | | | |
| Model | | | Unstandardized Coefficients | | Standardized | t | Sig |

| | | | | | Coefficients | | |
|---|---|---|---|---|---|---|---|
| | | | B | Std. Error | Beta | | |
| All firms | No statistical significance | | | | | | |
| GEN firms | 1 | (Constant) | **.165** | .214 | | .770 | .444 |
| | | Comp 2 | **.357** | .106 | .402 | 3.368 | .001 |
| OSS firms | No statistical significance | | | | | | |
| **New Processes vs. Value co-creation** | | | | | | | |
| **Model Summary** | | | | | | | |
| | Model | | R | R Square | Adjusted R Square | Std. Error of the Estimate | |
| All firms | No statistical significance | | | | | | |
| GEN firms | 1 | | .297 | **.088** | .073 | .577 | |
| OSS firms | No statistical significance | | | | | | |
| **Coefficients** | | | | | | | |
| Model | | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
| | | | B | Std. Error | Beta | | |
| All firms | No statistical significance | | | | | | |
| GEN firms | 1 | (Constant) | **.130** | .162 | | .805 | .424 |
| | | Comp 2 | **.191** | .080 | .297 | 2.390 | .020 |
| OSS firms | No statistical significance | | | | | | |
| **Total Innovation (new products + new services + new processes) vs. Value co-creation** | | | | | | | |
| **Model Summary** | | | | | | | |
| | Model | | R | R Square | Adjusted R Square | Std. Error of the Estimate | |
| All firms | No statistical significance | | | | | | |
| GEN firms | 1 | | .498 | **.248** | .235 | 1.375 | |
| OSS firms | No statistical significance | | | | | | |
| **Coefficients** | | | | | | | |
| Model | | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
| | | | B | Std. Error | Beta | | |
| All firms | No statistical significance | | | | | | |
| GEN firms | 1 | (Constant) | **.665** | .386 | | 1.721 | .090 |
| | | Comp 2 | **.842** | .191 | .498 | 4.414 | .000 |
| OSS firms | No statistical significance | | | | | | |

C1: Experience driven mutual learning; C2: Self service enabled by resources; C3: Customer involvement in co-production

The regression results table above indicates that most of the relationships between co-creation components and the traditional innovation metrics are statistically insignificant, especially in the cases involving the entire sample and the firms involved in OSS projects. The other regression models are summarized below.

Table 28: Traditional innovation vs. value co-creation models for Google data

| New Products | | | | |
|---|---|---|---|---|
| Type of firms | Model | Linear regression model | R-square | Relationship |
| All firms | N/A | | | |
| GEN firms | 1 | New products = 0.369 + (0.294 * C2) | 0.094 | New products ~ C2 |
| OSS firms | N/A | | | |
| **New Services** | | | | |
| All firms | N/A | | | |
| GEN firms | 1 | New services = 0.165 + (0.357 * C2) | 0.161 | New services ~ C2 |
| OSS firms | N/A | | | |
| **New Processes** | | | | |
| All firms | N/A | | | |
| GEN firms | 1 | New processes = 0.130 + (0.191 * C2) | 0.088 | New processes ~ C2 |
| OSS firms | N/A | | | |
| **Total Innovation (new products + new services + new processes)** | | | | |
| All firms | N/A | | | |
| GEN firms | 1 | Total innovation = 0.665 + (0.842 * C2) | 0.248 | Total innovation ~ C2 |
| OSS firms | N/A | | | |

C1: Experience driven mutual learning; C2: Self service enabled by resources; C3: Customer involvement in co-production

### 4.6.2 Linear Regression results for Yahoo! data

The regression analysis for the Yahoo! data is similar to the method used for the Google data set. The Yahoo! regression analyses can be found in Appendix B. Since the Yahoo! search engine produces different search results for the keywords the descriptive statistics for Yahoo! are presented.

Table 29: Descriptive statistics of the normalized Yahoo! components

| | Mean | Std. Deviation | Skewness | | Kurtosis | |
|---|---|---|---|---|---|---|
| | | | Statistic | Std. Error | Statistic | Std. Error |
| (Component 1)$^{1/3}$ | 3.5892 | 1.20679 | -.080 | .150 | -.413 | .300 |
| (Component 2)$^{1/3}$ | 2.6277 | .95866 | .281 | .150 | -.086 | .300 |
| (Component 3)$^{1/4}$ | 1.8369 | .50970 | .090 | .150 | -.030 | .300 |

C1: Experience driven mutual learning; C2: Customer involvement in co-production; C3: Managing risk through process flexibility

Table 30: Descriptive statistics for normalized innovation metrics for Yahoo!

| | Mean | Std. Deviation | Skewness | | Kurtosis | |
|---|---|---|---|---|---|---|
| | | | Statistic | Std. Error | Statistic | Std. Error |
| (New products)$^{1/5}$ | 1.0341 | .71979 | -.172 | .156 | -.329 | .310 |
| (New services)$^{0.27}$ | .7383 | .73405 | .360 | .156 | -.809 | .310 |
| (New processes + 0.005)$^{1/3}$ | .4653 | .57454 | .997 | .156 | -.560 | .310 |
| Total Innovation | 2.2377 | 1.48152 | .233 | .156 | -.722 | .310 |
| (Online metric + 1)$^{1/3}$ | 3.1387 | .74009 | -.291 | .151 | .047 | .300 |

Similar to the Google normalization process, the Skewness and Kurtosis statistics might not be less than twice their standard errors. The Skewness value was corrected to an acceptable value between -1 and 1, in the event that the statistics did not meet the initial normalization criteria.

### *Regression results for online innovation metric vs. Yahoo! data*

The regression results for the Yahoo! data are identical to the Google results. The significant information from the model summary and the coefficient tables are presented in the regression models. The following table outlines the regression tests carried out to explore how the online innovation metric relates to the value co-creation components extracted from the Yahoo! data.

Table 31: Online innovation metric vs. value co-creation for Yahoo! data

| Type of firms | Model | Linear regression model | R-square | Relationship |
|---|---|---|---|---|
| All firms | 2 | Online metric = 1.93 + (0.21*C1) + (0.17*C2) | 0.285 | Online metric ~ C1 & C2 |
| GEN firms | 1 | Online metric = 2.42 + (0.27*C2) | 0.145 | Online metric ~ C2 |
| OSS firms | 2 | Online metric = 1.50 + (0.33*C1) + (0.21*C3) | 0.390 | Online metric ~ C1 & C3 |

C1: Experience driven mutual learning; C2: Customer involvement in co-production; C3: Managing risk through process flexibility

### *Regression results for new products, services, and processes vs. Yahoo! components*

The regression results for the Yahoo! data set explore the relationship of the conventional innovation metrics against value co-creation components. Similar to the Google results, the regression models outline the connection of value co-creation components and the firms' innovative capacities. The regression results show how the co-creation components associate with new products, processes, and services for different types of firms.

Table 32: Conventional innovation vs. value co-creation models for Yahoo! data

| Type of firms | Model | Linear regression model | R-square | Relationship |
|---|---|---|---|---|
| **New Products** | | | | |
| All firms | N/A | | | |
| GEN firms | 1 | New products = 0.016 + (0.539 * C3) | 0.099 | New products ~ C3 |
| OSS firms | N/A | | | |
| **New Services** | | | | |
| All firms | N/A | | | |
| GEN firms | 1 | New services = -0.417 + (0.748 * C3) | 0.222 | New services ~ C3 |
| OSS firms | N/A | | | |
| **New Processes** | | | | |
| All firms | N/A | | | |
| GEN firms | N/A | | | |
| OSS firms | N/A | | | |
| **Total Innovation (new products + new services + new processes)** | | | | |
| All firms | N/A | | | |
| GEN firms | 1 | Total innovation = -0.201 + (1.456 * C3) | 0.232 | Total innovation ~ C3 |
| OSS firms | N/A | | | |

C1: Experience driven mutual learning; C2: Customer involvement in co-production; C3: Managing risk through process flexibility

# 5. Analysis of results

The analyses of the results are presented in this chapter. In Section 5.1, the interpretation of the value co-creation components extracted by using the Google and the Yahoo! search engines summarize the co-creation activities of the firms. Section 5.2 presents the similarity between the Google components and the components extracted by Allen. Section 5.3 will examine the Google and Yahoo! co-creation components. Section 5.4 presents the results of the linear regression tests.

## 5.1 Interpretation of value co-creation components

The factor analysis method identified three factors that represent value co-creation components for the Google and Yahoo! data sets. The components include a set of keyword variables that outline distinct co-creation activities within the firms. The insights from value co-creation literature, coupled with the context of the keywords themselves, are used to interpret the explanation for each component.

### 5.1.1 Interpretation of Google components

This section provides the interpretation of the components extracted from the Google data.

***Component 1: Experience driven mutual learning***

Component 1 extracted from the Google data contains the keywords shown in table 33.

Table 33: Google component 1 keyword variables

| | Keyword Variables | Loading |
|---|---|---|
| V11 | (customer OR user) AND experience | 0.775 |
| V1 | (customer OR user) AND (dialog OR dialogue OR communicate OR communication OR conversation OR contact OR feedback OR call OR interact OR "information exchange" OR "information sharing" OR connect OR access OR engage) | 0.773 |
| V9 | (customer OR user) AND (learn OR learning) | 0.723 |
| V5 | (customer OR user) AND (suggest OR suggestion OR input OR request OR demand) | 0.708 |
| V4 | (customer OR user) AND (cooperate OR cooperation OR collaboration OR partnership) | 0.698 |
| V19 | Customer AND (partnerships OR interaction OR relationship OR participate OR participation OR activity OR action) | 0.651 |
| V2 | (customer OR user) AND (communities OR community OR network OR networking OR forum) | 0.625 |
| V7 | (customer OR user) AND risk AND (manage OR management OR control OR assess OR reduce OR reduction OR potential OR exposure) | 0.538 |
| V17 | (customer OR user) AND (options OR choice OR choose) | 0.523 |

This component is interpreted as "Experience driven learning based on dialogue, cooperation, partnerships, and networking"

***Component 2: Resources and processes enabling lease, rent, and self-service***

The loading of the variable "(product OR process) AND (modularity OR modular OR module)" shows the relative significance of the modularization of products of processes within component 2.

Table 34: Google component 2 keyword variables

| | Keyword Variables | Loading |
|---|---|---|
| V15 | (product OR process) AND (modularity OR modular OR module) | 0.733 |
| V6 | Internal AND (expertise OR resource) | 0.628 |
| V8 | (customer OR user) AND (IP OR "intellectual property") | 0.617 |
| V3 | lease OR rent OR license OR "self serve" OR "self service" | 0.598 |
| V28 | (customer OR user) AND language AND translation | 0.493 |

Component 2 is interpreted as "Resources and processes enabling lease, rent, licensing, and self service based on modularity and IP sharing".

***Component 3: Customer involvement in Co-production***

The variables in component 3 extracted from the Google data is listed in table 35.

Table 35: Google component 3 keyword variables

| | Keyword Variables | Loading |
|---|---|---|
| V16 | (customer OR user) AND (produce OR assemble OR manufacture) | 0.645 |
| V22 | trust OR honesty OR integrity | 0.609 |
| V23 | (customer OR user) AND (disclose OR inform OR disseminate OR reveal) | 0.604 |
| V21 | (customer OR user) AND (survey OR review OR voting OR vote OR rate OR rating) | 0.603 |
| V29 | (customer OR user) AND address AND concern | 0.589 |
| V20 | cost AND (reduce OR reduction OR saving) | 0.586 |

Component 3 is interpreted as "Customer involvement in production and manufacturing based on trust and focusing on cost reduction".

### 5.1.2 Interpretation of Yahoo! components

The value co-creation components extracted from the Yahoo! data is summarized in this section.

***Component 1: Experience driven mutual learning***

The first component extracted from the Yahoo! data reflects Google component 1 very closely. The extraction of the similarly structured component 1 validates the importance of the mutual learning experience in value co-creation.

Table 36: Yahoo! component 1 keyword variables

| | Keyword Variables | Loading |
|---|---|---|
| V1 | (customer OR user) AND (dialog OR dialogue OR communicate OR communication OR conversation OR contact OR feedback OR call OR interact OR "information exchange" OR "information sharing" OR connect OR access OR engage) | 0.698 |
| V11 | (customer OR user) AND experience | 0.690 |
| V9 | (customer OR user) AND (learn OR learning) | 0.689 |
| V19 | Customer AND (partnerships OR interaction OR relationship OR participate OR participation OR activity OR action) | 0.685 |
| V4 | (customer OR user) AND (cooperate OR cooperation OR collaboration OR partnership) | 0.601 |
| V2 | (customer OR user) AND (communities OR community OR network OR networking OR forum) | 0.573 |

Component 1 extracted from the Yahoo! data is translated as "Experience driven learning based on dialogue, cooperation, partnerships, and networking focusing on risk management and providing multiple customer choices".

.

***Component 2: Customer involvement in Co-production***

The following table outlines the keyword variables that define Yahoo! component 2.

Table 37: Yahoo component 2 keyword variables

| | Keyword Variables | Loading |
|---|---|---|
| V16 | (customer OR user) AND (produce OR assemble OR manufacture) | 0.684 |
| V17 | (Customer OR user) AND (options OR choice OR choose) | 0.659 |
| V20 | cost AND (reduce OR reduction OR saving) | 0.597 |
| V21 | (customer OR user) AND (survey OR review OR voting OR vote OR rate OR rating) | 0.572 |
| V5 | (customer OR user) AND (suggest OR suggestion OR input OR request OR demand) | 0.556 |

This component is interpreted as "Customer involvement in production assembly manufacturing focusing on cost reduction enabled by multiple choices".

***Component 3: Risk management through design flexibility***

The final component extracted from the Yahoo! data is outline in table 38.

Table 38 Yahoo! component 3 keyword variables

| | Keyword Variables | Loading |
|---|---|---|
| V7 | (customer OR user) AND risk AND (manage OR management OR control OR assess OR reduce OR reduction OR potential OR exposure) | 0.665 |
| V14 | simulation OR simulate OR model OR modeling OR "virtual world" OR "reference design" OR "reference flow" OR "demo application" OR toolkit OR tutorial OR SDK OR "software development kit" | 0.614 |
| V18 | (design OR process) AND (flexibility OR flexible OR adaptable) | 0.595 |

Yahoo! component 3 is interpreted as "Tools and processes enabling customer risk management through design and process flexibility".

## 5.2 Google and Allen's co-creation components

Exploratory factor analysis of the updated Google data produces a new set of value co-creation components. This is evaluated against the findings of Allen's (2009) research in order to study the change in value co-creation activities within the research sample.

Table 39 differentiates Allen's co-creation components against the ones extracted from the updated Google data.

Table 39: Google and Allen's co-creation components

| Component | Jainal Chisty (2010) | Stephen Allen (2009) |
|---|---|---|
| 1 | Experience driven learning based on dialogue, cooperation, partnerships, and networking | Community forum designed to engage customers in an open dialog including networking, information sharing and learning activities with the organization, other customers or other members of the value network |
| 2 | Resources and processes enabling lease, rent, licensing, and self service based on modularity and IP sharing | Partnerships enabling user access to company expertise & resources to create adaptable designs & processes aiming at reducing costs and based on trust, integrity & risk management |
| 3 | Customer involvement in production and manufacturing based on trust and focusing on cost reduction | Personalization of offers through partnerships across the value network to provide choices/options enabled by product and process modularity, and integrated online services |
| 4 | | Co-production of offers by user involvement in manufacturing, assembly and final beta trial activities; requiring disclosure and sharing of intellectual property |

The set of co-creation components extracted from the Google data are similar in multiple aspects to the components from Allen's work. Component 1 in both cases essentially describes the information sharing process with other members of the value network. Component 1 from the updated Google component set emphasizes the importance of customer experiences in the mutual learning process, in order to provide more options and choices to users.

Component 3 as outlined by the updated data focuses on co-production related to reducing costs. This is very similar to component 2 in Allen's set, which refers to co-production. The Google data suggests that the sample firms, while assessing new business strategies for value co-creation have retained several practices used to enable co-creation with the value network.

Table 40: Google and Allen's value co-creation components

| | Chisty (2010) | | | Allen (2009) | | | |
|---|---|---|---|---|---|---|---|
| Keyword combinations | 1 | 2 | 3 | 1 | 2 | 3 | 4 |
| (customer OR user) AND experience | 0.78 | | | | | | |
| (customer OR user) AND (dialog OR dialogue OR communicate OR communication OR conversation OR contact OR feedback OR call OR interact OR "information exchange" OR "information sharing" OR connect OR access OR engage) | 0.77 | | | 0.56 | | | |
| (customer OR user) AND (learn OR learning) | 0.72 | | | 0.74 | | | |
| (customer OR user) AND (suggest OR suggestion OR input OR request OR demand) | 0.71 | | | 0.67 | | | |
| (customer OR user) AND (cooperate OR cooperation OR collaboration OR partnership) | 0 70 | | | | 0.55 | | |
| Customer AND (partnerships OR interaction OR relationship OR participate OR participation OR activity OR action) | 0 65 | | | | 0 65 | | |
| (customer OR user) AND (communities OR community OR network OR networking OR forum) | 0.63 | | | 0 71 | | | |
| (customer OR user) AND risk AND (manage OR management OR control OR assess OR reduce OR reduction OR potential OR exposure) | 0.54 | | | | 0.53 | | |
| (customer OR user) AND (options OR choice OR choose) | 0.52 | | | | | 0.68 | |
| (product OR process) AND (modularity OR modular OR module) | | 0 73 | | | | 0.46 | |
| Internal AND (expertise OR resource) | | 0.63 | | | 0.72 | | |
| (customer OR user) AND (IP OR "intellectual property") | | 0 62 | | | | | 0.50 |
| lease OR rent OR license OR "self serve" OR "self service" | | 0 60 | | | | | |
| (customer OR user) AND language AND translation | | 0 49 | | | | | |
| (customer OR user) AND (produce OR assemble OR manufacture) | | | 0 65 | | | | 0.57 |
| trust OR honesty OR integrity | | | 0 61 | | 0.53 | | |
| (customer OR user) AND (disclose OR inform OR disseminate OR reveal) | | | 0.60 | | | | 0.58 |
| (customer OR user) AND (survey OR review OR voting OR vote OR rate OR rating) | | | 0 60 | | | | |
| (customer OR user) AND address AND concern | | | 0 59 | | | | |
| cost AND (reduce OR reduction OR saving) | | | 0.59 | | 0.70 | | |
| (design OR process) AND (flexibility OR flexible OR adaptable) | | | | | 0.65 | | |
| integrated AND online AND services | | | | | | 0.66 | |
| customization OR customize OR customized OR personalize OR individualize OR "add feature " OR "added feature " | | | | | | 0.59 | |
| ecosystem OR "value network" OR "value constellation" OR "multiple partners" OR "external contributor" OR "external source" | | | | | | 0.46 | |
| (customer OR user) AND (test OR trial OR beta) | | | | | | | 0 44 |

The three co-creation components, based on recent firm data from Google, also indicate that Allen's "personalization" component seems to have been phased out by firms. The data implies that firms have shifted focus to providing modular products. It also shows that firms provide users with more options as a means to enhance their experience with the firm, rather than provide multiple options as part of a personalization process. The four co-creation components identified by Allen have changed to three main components of value co-creation that is active in the sample firms.

## 5.3 Google and Yahoo! co-creation components

The collection of the second set of value co-creation data using the Yahoo! search engine enables the creation of a second set of co-creation components. The Google and Yahoo! search engine APIs were used to look for the same keyword variables within very similar firm samples. The breakdown of the sample firms is presented in table 41. The data collection was done on the same day, and the components extracted from the two distinct search engines are expected to outline similar themes of value co-creation. The unique search ranking algorithms and indexes of the two search engine APIs is accountable for the differences in components extracted.

Table 41: Breakdown of sample firms

| Types of firms | Google | | Yahoo! | |
|---|---|---|---|---|
| | Count | Percentage | Count | Percentage |
| OSS | 69 | 26.34 | 66 | 25.19 |
| ECL | 105 | 40.08 | 107 | 40.84 |
| GEN | 63 | 24.05 | 63 | 24.05 |
| OSS+ECL | 25 | 9.54 | 26 | 9.92 |
| **Total** | 262 | 100 | 262 | 100 |

The component interpretations for the Google and Yahoo! results are presented in table 42.

Table 42: Component interpretations – Google vs. Yahoo!

| Component | Google | Yahoo! |
|---|---|---|
| 1 | **Experience driven learning** based on dialogue, cooperation, partnerships, and networking | **Experience driven learning** based on dialogue, cooperation, partnerships, and networking focusing on risk management and providing multiple customer choices |
| 2 | Resources and processes enabling lease, rent, licensing, and self service based on modularity and IP sharing | **Customer involvement in production and manufacturing focusing on cost reduction** enabled by multiple choices |
| 3 | **Customer involvement in production and manufacturing based on trust and focusing on cost reduction** | Tools and processes enabling customer risk management through design and process flexibility |

The analysis of the Google and Yahoo! components lead to the following findings.

The two component sets are comprised of similar value co-creation activities, with two of the three components describing similar approaches to enable co-creation in firms. With the results based on very similar samples of firms, the unique components that emerge from the Google and Yahoo! data sets can be attributed to the difference in search engine algorithms and indexes.

The firms that were identified as most active in all value co-creation components for the Google and Yahoo! data show that the same types of firms are active for both component sets. With the exception of two 'General' firms among the active firms, firms active in OSS projects are found to be most active in co-creation activities.

However, only two common firms are identified as active in all co-creation components for both Google and Yahoo! data sets. This also points to the difference between the Google and Yahoo! search engine results, and the assumption that their indexes and algorithms are inherently different. Ultimately, further analysis is required on each case to determine the level of co-creation within the firms, and whether these results are representative of deliberate business strategies employed to enable co-creation.

The use of the Yahoo! search engine to extract a second set of value creation components enables the verification of the components from the Google data. The second value creation component set was intended to verify the components identified by Google used to create customer value. Since both search engines use the same keyword set and a similar sample of firms, literature suggests the use of a secondary source of data to confirm the initial findings from the Google search.

The value creation components extracted from the Google and Yahoo! data were not integrated into a single set of components because of their unique compositions. Each keyword variable within a component was identified to be closely related to the other keyword variables within the component by the Principal Component Analysis process. This factor extraction method determined that the set of activities within each value creation component are highly correlated and representative of a set of value creation activities according to the firm data from a specific search engine. So while certain Google and Yahoo! components are interpreted to be similar, they actually represent a distinct set of value creation activities found to be employed by the firms in the research samples.

The variables that make up the Google and Yahoo! components, along with their loading values, are presented in table 43.

Table 43: Value creation components - Google vs. Yahoo!

| Keyword combinations | Google | | | Yahoo! | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 1 | 2 | 3 |
| (customer OR user) AND experience | 0.78 | | | 0.69 | | |
| (customer OR user) AND (dialog OR dialogue OR communicate OR communication OR conversation OR contact OR feedback OR call OR interact OR "information exchange" OR "information sharing" OR connect OR access OR engage) | 0.77 | | | 0.70 | | |
| (customer OR user) AND (learn OR learning) | 0.72 | | | 0.69 | | |
| (customer OR user) AND (suggest OR suggestion OR input OR request OR demand) | 0.71 | | | | 0.56 | |
| (customer OR user) AND (cooperate OR cooperation OR collaboration OR partnership) | 0.70 | | | 0.60 | | |
| Customer AND (partnerships OR interaction OR relationship OR participate OR participation OR activity OR action) | 0.65 | | | 0.69 | | |
| (customer OR user) AND (communities OR community OR network OR networking OR forum) | 0.63 | | | 0.57 | | |
| (customer OR user) AND risk AND (manage OR management OR control OR assess OR reduce OR reduction OR potential OR exposure) | 0.54 | | | | | 0.67 |
| (customer OR user) AND (options OR choice OR choose) | 0.52 | | | | 0.66 | |
| (product OR process) AND (modularity OR modular OR module) | | 0.73 | | | | |
| Internal AND (expertise OR resource) | | 0.63 | | | | |
| (customer OR user) AND (IP OR "intellectual property") | | 0.62 | | | | |
| lease OR rent OR license OR "self serve" OR "self service" | | 0.60 | | | | |
| (customer OR user) AND language AND translation | | 0.49 | | | | |
| (customer OR user) AND (produce OR assemble OR manufacture) | | | 0.65 | | 0.68 | |
| trust OR honesty OR integrity | | | 0.61 | | | |
| (customer OR user) AND (disclose OR inform OR disseminate OR reveal) | | | 0.60 | | | |
| (customer OR user) AND (survey OR review OR voting OR vote OR rate OR rating) | | | 0.60 | | 0.57 | |
| (customer OR user) AND address AND concern | | | 0.59 | | | |
| cost AND (reduce OR reduction OR saving) | | | 0.59 | | 060 | |
| simulation OR simulate OR model OR modeling OR "virtual world" OR "reference design" OR "reference flow" OR "demo application" OR toolkit OR tutorial OR SDK OR "software development kit" | | | | | | 0.61 |
| (design OR process) AND (flexibility OR flexible OR adaptable) | | | | | | 0.60 |

## 5.4 Linear regression test results

The linear regression method is used in this research to explore the relationship between the value co-creation components, extracted from the Google and Yahoo! data, and the innovation measures. The online innovation metric is assumed to provide an approximate measure of the innovative capacity of the firm, while the actual number of new products, services, and processed collected for each firm over the last three years represent a more conventional measure of the firm's innovative capacity. The regression results are presented below, for the data sets from both search engines.

Table 44 outlines the relationships between value co-creation components and the different innovation metrics for the Google data.

The analysis shows that the in the case of the entire research sample, the only statistically significant relationship is between Google components 1 and 2 and the online innovation metric. This implies that experience driven learning, combined with resource-enabled self service, enhance the innovation capacity within firms. OSS firms have a similar relationship with innovation, but with a higher explanatory power. This finding suggests that firms involved in OSS projects are able to articulate their innovation outcomes online better than more traditional firms.

Table 44: Google regression results

| | All firms | GEN firms | OSS firms |
|---|---|---|---|
| Online Innovation metric | = const + (0.14***C1**) + (0.1***C2**)<br>R-Square = 0.28 | N/A | = const+(0.18***C1**) +(0.08***C2**)<br>R-square = 0.30 |
| Total Innovation (new products+new services+new processes) | N/A | = const + (0.842 * **C2**)<br>R-Square=0.25 | N/A |
| New Products | N/A | = const + (0.294 * **C2**)<br>R-Square=0.09 | N/A |
| New Services | N/A | = const + (0.357 * **C2**)<br>R-Square=0.16 | N/A |
| New Processes | N/A | = const+ (0.191 * **C2**)<br>R-Square=0.09 | N/A |

C1: Experience driven mutual learning; C2: Self service enabled by resources; C3: Customer involvement in co-production

In the case of the conventional innovation metrics, the statistically significant relationships exist between co-creation components and new products, new services, and new processes in the case of GEN firms. The relationship between GEN firms and total innovation indicate that there is significant correlation between innovation and component 2, defined as self service enabled through firm resources. The lower explanatory power of these models suggests that the component does not assertively contribute to defining the innovative capacity of general type of firms.

The Yahoo! regression analysis is presented in table 45. The results show that for the online innovation metric, the relationships between innovation and value co-creation components are statistically significant for GEN and OSS firms.

The model that describes the relationship between value co-creation components and the online innovation metric has the highest explanatory power in the case of OSS firms. Similar to the regression analysis for the Google data, this suggests that OSS firms are in a better position to communicate their innovative capacities online. It is also indicative of the roles that 'experience driven mutual learning' and 'risk management through process flexibility' activities play in the articulation of the innovative capacity of OSS firms.

Table 45: Yahoo! regression results

| | All firms | GEN firms | OSS firms |
|---|---|---|---|
| Online Innovation metric | = const +(0.21***C1**)+(0.17***C2**)<br>R-Square = 0.29 | = const + (0.27***C2**)<br>R-Square = 0.15 | = const +(0.33***C1**)+(0.21***C3**)<br>R-square = 0.39 |
| Total Innovation (new products+new services+new processes) | N/A | = const + (1.456 * **C3**)<br>R-Square=0.23 | N/A |
| New Products | N/A | = const + (0.539 * **C3**)<br>R-Square=0.10 | N/A |
| New Services | N/A | = const + (0.748 * **C3**)<br>R-Square=0.22 | N/A |
| New Processes | N/A | N/A | N/A |

C1: Experience driven mutual learning; C2: Customer involvement in co-production; C3: Managing risk through process flexibility

For the conventional innovation metrics, significant relationships all involve component 3 and general types of firms. Yahoo! component 3 relates to the management of risk through process and design flexibility. However, the lower explanatory power of the model describing the relationship between total innovation outcome of a firm and this co-creation component suggests that the activities within this 'risk management' component do not compellingly define the innovative capacity of the firm. A thorough examination of the GEN firms in question is needed to verify this finding.

## 6. Summary of Key Findings

Analysis of the research results leads to the following key findings:

The research methodology involving data mining techniques along with exploratory factor analysis previously used by Hicks (2006), McGinnis (2008), Allen (2009), and Tanev (2010) provide a valuable tool in studying the co-creative activities of firms over time.

The difference in search engines appears to influence, to a certain extent, the results provided by the research methodology. The influence of the unique search engine algorithms and indexes on the research results can be summarized as follows:

- The composition of the value co-creation activities are unique for the distinct data sets, and two co-creation component sets, that are similar only to an extent, are produced by the Google and Yahoo! data
- The difference in search engines do not affect the types of firms identified as being active in co-creation, it affects the pecific selection of firms that were identified as most active according to the Google and Yahoo! data
- The specific components that are found to be relevant to a firm's innovation capacity is affected, since Google and Yahoo! identify different components as being influential in the innovative outcomes

The regression analysis implies a positive association between specific value co-creation components and the innovative capacities of firms. This relationship, however, is dependent on the type of innovation metric and on the specific search engine.

In the case of OSS firms the online innovation metric was found to be positively associated with the 'experience based mutual learning', 'self service enabled by resources', and 'managing risk through process flexibility' components. This finding suggests that OSS firms are better positioned to articulate the innovative aspects of their products, services, and processes.

In the case of 'general' type of firms, it was found that there is a positive association between the total number of innovations and the co-creation components 'self service enabled by resources' and 'managing risk through process and design flexibility'. The specific component that relates to the innovative capacity of firms varied according to the search engine, suggesting that the 'general' type of firms can profit from development of web resources that enhance their ability to convey their innovative capacity online.

# 7. Limitations and Future Work

The limitations of this research thesis, as well as recommendations for future research, are outlined in the following section.

The finding that conventional innovation metrics are not related to value co-creation activities raises some concerns, and need to be explored further. This research finding might be related to the typical nature of OSS firms, focused on the new releases and updates of existing products and services rather than new ones. However, it could also be related to the complexity of the data collection process, which was done by multiple researchers across the globe with differing views on what constitutes new innovations.

Future studies related to this research should address the following issues:

The interpretation of the co-creation components is subjective, and biased towards supporting the research hypothesis. Future studies should focus on a more qualitative approach that addresses the contextual manifestations of the co-creation components.

The relevance of the co-creation keywords appear to be dated, and not inclusive of new approaches to value co-creation practices. Future research should consider a more comprehensive set of keywords that outline present co-creation efforts.

It is impossible to explain why the 'general' type of firms demonstrates a positive relationship between certain co-creation activities and the innovative capacity of firms without a case-based analysis of their co-creation efforts. Future studies might be able to describe specific aspects of co-creation components that enhance the innovation capacity of the firms.

# References

**Allen, S. 2009.** An Empirical Study of the Components of Value Co-creation. Master Thesis, Technology Innovation Management, Carleton University, Ottawa, Ontario.

**Allen, S., Tanev, S. and Bailetti, T. 2009.** Components of co-creation. *Open Source Business Review Online Journal*, 2009 (November). Available at: www.osbr.ca/ojs/index.php/osbr/article/view/991/954 (Accessed: Jan 12th, 2010).

**Ballantyne, D., 2004.** Dialog and its role in the development of relationship specific knowledge. *Journal of Business and Industrial Marketing, 19(2),* 114-123.

**Ballantyne, D. and Varey, R., 2008.** The service-dominant logic and the future of marketing. *Journal*
*of the Academy of Marketing Science, 36(1),* 11–14.

**Bendapudi, N., and Leone, R. 2003.** Psychological implications of customer participation in coproduction.*Journal of Marketing, 67(1),* 14–28.

**Bharat, K. and Broder, A. 1998**. A technique for measuring the relative size and overlap of public web search engines. In 7th WWW, 1998.

**Bilgram, V., Brem, A. and Voigt, K.-I. 2008.** User-centric innovations in new product development - systematic identification of lead users harnessing interactive and collaborative online-tools. *International Journal of Innovation Management, 12(3),* 419–458.

**Bogers, M., Afuah, A. and Bastian, B. 2010**. Users as innovators: a review, critique, and future research directions. Journal of management, 36(4), 857-875.

**Bolton, R. & Saxena-Iyer, S. 2009.** Interactive services: a framework, synthesis and research directions. *Journal of interactive marketing, 23,* 91-104.

**Bonsu, S. & Darmody, A. 2008.** Co-creating Second Life – Market-consumer cooperation in contemporary economy. *Journal of macromarketing, 28(4),* 355-368.

**Bowonder, B., Dambal, A., Kumar, S. & Shirodkar, A. 2010.** Innovation strategies for creating competitive advantage. *Research-technology management, 53(3),* 19-32.

**Broder, A., Kumar, R., Maghoul, F., Raghavan, P., Rajagopalan, S., Stata, R., et al. 2000.** Graph structure in the Web. *Journal of Computer Networks, 33(1–6)*, 309–320.

**Brohman, M., Piccoli, G., Martin, P., Zulkernine, F., Parasuraman, A., Watson, R. 2009.** A design theory approach to building strategic network-based customer service systems. *Decision sciences, 40(3),* 403-430.

**Caillet, P. 2006.** Yahoo Introduction and Its Ranking Algorithm, Irka Web Productions. Available at: http://www.irkawebpromotions.com/search-engines/yahoo/ (2010. 05. 26)

**Caillet, P. 2006a.** Understanding the Google Ranking Algorithm, Irka Web Productions. Available at: http://www.irkawebpromotions.com/search-engines/google/4-google-algorithm.html (2010.05.26)

**Cattell, R. B. 1966.** The Meaning and Strategic Use of Factor Analysis. In *Handbook of Multivariate Experimental Psychology*. Chicago: Rand McNally.

**Chesbrough, H. W. 2003.** The Era of Open Innovation. *MIT Sloan Management Review 44(3),* 35-41.

**Cova, B. & Salle, R. 2008.** Marketing solutions in accordance with the S-D logic: Co-creating value with customer network actors. *Industrial marketing management, 37,* 270–277.

**Costello, A. and Osborne J. 2005.** Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most from Your Analysis. *Practical Assessment Research & Evaluation, 10(7),* 1-9.

**Crystal, A. and Greenberg, J. 2006.** Relevance criteria identified by health information users during web searches. *Journal of the American Society for Information Science and Technology, 57(10),* 1368-1382.

**Darlington, R. 2008.** Factor Analysis. Available at: http://www.psych.cornell.edu/Darlington/factor.htm (Accessed: Jan 11th, 2010).

**Dong, B., Evans, K. & Zou, S. 2008.**The effects of customer participation in co-created service recovery. *Journal of the academy of marketing science, 36,* 123–137.

**Droge, C., Stanko, M. and Pollitte, W. 2010.** Lead users and early adopters on the web: The role of new technology product blogs. *Journal of product innovation management, 27*, 66-82.

**Edvardsson, B., Enquist, B. and Johnston, R. 2005**. Cocreating customer value through hyperreality in the prepurchase service experience. *Journal of service research, 8,* 149-161.

**Etgar, M. 2006.** Co-production of services: A managerial extension. In: Lusch, R., & Vargo, S. (Eds.), The service-dominant logic of marketing: Dialog, debate and directions, 128–138. N.Y: M. E. Sharpe.

**Etgar, M. 2008.** A descriptive model of the consumer co-production process. *Journal of the academy of marketing science, 36(1),* 97-108.

**Eysenck H. J. 1953.** The Logical Basis of Factor Analysis. *American Psychologist, 8(3), 105-114.*

**Ferguson, R. & Paulin, M. 2010.** Customer sociability and the total service experience: antecedents of positive word-of-mouth intentions. *Journal of Service Management, 21(1),* 25-44.

**Ferrier, W. 2001.** Navigating the Competitive Landscape: The Drivers and Consequences of Competitive Aggressiveness, *Academy of Management Journal, 44(4),* 858-877.

**Field, A. 2009**. *Discovering Statistics using SPSS*, Sage Publishing.

**Finch, B. J. 1999**. Internet discussions as a source for consumer product customer involvement and quality information. *Journal of Operations Management, 17(5),* 535–556.

**Flint, D. and Mentzer, J. 2006.** Striving for integrated value chain management given a service-dominant logic of marketing. Ch. 11 in Lusch, R. & Vargo, S., Eds., *The service-dominant logic of marketing*, Armonk, NY: M. E. Sharpe Inc, 139-149.

**Franke, N., and Shah, S. 2003.** How communities support innovative activities: An exploration of assistance among users of sporting equipment. *Research Policy, 32(1)*, 157–178.

**Franke, N. & Schreier, M. 2008.** Product uniqueness as a driver of customer utility in mass customization. *Marketing letters, 19(2),* 93-107.

**Fuller, J., Bartl, M., and Muhlbacher, H. 2006.** Community based innovation: How to integrate members of virtual communities into new product development. *Electronic Commerce Research, 6(1)*, 57–73.

**Gersuny, C., and Rosengren, W. R. 1973.** *The service society.* MA: Schenkman Press.

**Google sets.** From Google Labs, available at: http://www.googlelabs.com/show_details?app_key=agtnbGFiczlwLXd3d3lTCxlMTGFic0FwcE1vZGVsGIYxDA (Accessed: Jan 18th, 2010).

**Google synonym Search Tool.** Avaiable at: www.synonymlab.com (Accessed Jan 29th, 2010).

**Gulli, A. and Signorini, A. 2005.** The indexable web is more than 11.5 billion pages. *Proceedings of the 14th International Conference on World Wide Web*, Chiba, Japan, 902-903.

**Hagedoorn, J. and Cloodt, M. 2003**. Measuring innovative performance: is there an advantage in using multiple indicators, *Research Policy, Vol. 32,* 1365-1379.

**Han, J. and Kamber, M. 2006.** *Data Mining: Concepts and Techniques.* CA: Morgan Kaufmann.

**Henzinger, M. 2006.** Finding near-duplicate Web pages: A large-scale evaluation of algorithms. Presented at: *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Revelopment in Information Retrieval,* 284–291.

**Hertel, G., Niedner, S., and Hermann, S. 2003.** Motivation of software developers in open source projects. *Research Policy, 32(7),* 1159–1177.

**Hearn, G and Pace, C. 2006.** Value-creating ecologies: understanding next generation business systems. *Foresight, 8(1),* 55-65.

**Hicks, D. M.; Libaers, D. P.; Porter, A. L. and D. J. Schoeneck. 2006.** Identification of the Technology Commercialization Strategies of High-tech Small Firms. *Small Business Research Summary.*December 2006.

**Jaworski, B. and Kohli, A. K. 2006.** Co-creating the voice of the customer, Ch. 8, *The Service Dominant Logic of Marketing,* R. Lusch & S. Vargo, Eds., Armonk, NY: M. E. Sharpe.

**Katz, E., Blumler, J. G., and Gurevitch, M. 1974.** Utilization of mass communication by the individual. In J. G. Blumler & E. Katz (Eds.), *The uses of mass communications: current perspectives on gratifications research (pp. 19–32).* Beverly Hills: Sage.

**Kaulio, M. A. 1998.** Customer, consumer, and user involvement in product development: A framework and a review of selected methods. *Total Quality Management, 9(1),* 141–149.

**Kohler, T., Matzler. K, and Fuller, J. 2008.** Avatar-based innovation: Using virtual worlds for real-world innovation. *Technovation, 29,* 395–407.

**Kristensson, P., Matthing, J and Johansson, N. 2008.** Key strategies for the successful involvement of customers in the co-creation of new technology-based services. International journal of service industry management, 19(4), 474-491.

**Langville, A. N. and Meyer, C. D. 2008.** Google's PageRank and Beyond: The Science of Search Engine Rankings. *The Mathematical Intelligencer, 30(1)*, 68-69.

**Lengnick-Hall, C. A. 1996.** Customer contributions to quality: A different view of the customer oriented firm. *Academy of Management Review, 21(3),* 791–810.

**Lewandowski, D., Wahlig, H. and Meyer-Bautor, G. 2006.** The freshness of web search engine Databases. *Journal of Information Science, 32(2),* 133-50.

**Lewandowski, D. 2007.** Mit welchen Kennzahlen la¨sst sich die Qualita¨t von Suchmaschinen messen?, in Machill, M. and Beiler, M. (Eds), Die Macht der Suchmaschinen (The Power of Search Engines), von Halem, Ko¨ln, pp. 243-58.

**Lewandowski, D. and Ho¨chsto¨tter, N. 2007.** Web Searching: a quality measurement perspective. In Spink, A. and Zimmer, M. (Eds), *Web Searching: Interdisciplinary Perspectives,* Springer, Dordrecht.

**Lewandowski, D. 2008.** The retrieval effectiveness of web search engines: considering results descriptions. *Journal of Documentation, 64(6),* 915-937.

**Lombardi, S. 2008.** Interactions between Eclipse Foundation Members and Eclipse Projects. Master's Thesis, Technology Innovation Management, Carleton University, Ottawa, Ontario.

**Maglio, P. & Spohrer, J. 2008.** Fundamentals of service science. *Journal of the academy of marketing science, 36,* 18–20.

**McCown, F. and Nelson, M. F. 2007.** Agreeing to Disagree: Search Engines and Their Public Interfaces. Presented at: *Proceedings of the 7th ACM/IEEE-CS joint conference on Digital libraries, Vancouver, Canada.* 309-318.

**McGinnis, G. 2008**. Competitive Actions of Companies whose Revenue Relies on Open Source Software. Master's Thesis, Technology Innovation Management, Carleton University, Ottawa, Ontario.

**Midgley, D. 2009.** Co-creating the innovation with customers. Ch. 5 in *The innovation manual – Integrating strategies and practical tools for bringing value innovation to the market.* Chichester, UK: John Wiley & Sons, 143-190.

**Nambisan, S. 2002.** Designing virtual customer environments for new product development: Toward a theory. *Academy of Management Review, 27 (3),* 392–413.

**Nambisan, S. and Baron, R. 2007.** Interactions in virtual customer environments: implications for products support and virtual customer relationship management. *Journal of interactive marketing, 21(2),* 42-62.

**Nambisan, S., and Nambisan, P. 2008.** How to profit from a better virtual customer environment. *MIT Sloan Management Review, 49(3),* 53–61

**Nambisan, S., 2008.** Transforming government through collaborative innovation. *The Public Manager, 37(3),* 36-41.

**Nambisan, S. and Baron, R., 2009.** Virtual Customer Environments: Testing a Model of Voluntary Participation in Value Co-creation Activities. *The Journal of Product Innovation Management, 26(4):* 388-406.

**Nambisan, S. 2009.** Virtual Customer Environments: IT-Enabled Customer Co-innovation and Value Co-creation. Nambisan, S., (Ed.), *Information Technology and Product Development, Annals of Information Systems, 5(2),* 109-127.

**Ng, I., Maull, R. and Yip, N. 2009.** Outcome-based contracts as a driver for systems thinking and service-dominant logic in service science: Evidence from the defence industry. *European management journal, 27,* 377- 387.

**Normann, R. and Ramirez, R. 1993.** From Value Chain to Value Constellation: Designing Interactive Strategy. *Harvard Business Review, 71(4*), 65-77.

**Normann, R. and Ramirez, R. 1993a.** Richard Normann and Rafael Ramirez Respond: *Harvard Business Review, 71(5),* 50-51.

**Ontology Finder.** From goRank SEO tools. Available at: http://www.gorank.com/seotools/ontology/ (Accessed: Jan 24th, 2010).

**Ostrom, A., Bitner, M., Brown, S., Burkhard, K., Goul, M., Smith-Daniels, V., Demirkan, H. & Rabinovich, E. 2010.** Moving forward and making a difference: research priorities for the science of service. *Journal of service research, 13(4),* 4-36.

**Opoku, R. 2005.** Communication of Brand Personality by Some Top Business Schools Online, Licentiate Thesis, Lulea University of Technology, ISSN: 1402-1757.

**Oslo Manual: Guidelines for Collecting and Interpreting Innovation Data, 3rd Edition, 2005.** Available from: The Organisation for Economic Co-Operation and Development. http://www.oecdbookshop.org/oecd/display.asp?CID=&LANG=EN&SF1=DI&ST1=5LGPBVQFQ4G5 (Accessed: Jan 10th, 2010).

**Payne, A., Storbacka, K. and Frow, P. 2008.** Managing the co-creation of value. *Journal of the Academy of Marketing Sciences. 36(1),* 83-96.

**Piller, F. and Walcher, D., 2006.** Toolkits for idea competitions: a novel method to integrate users in new product development. *R&D Management, 36(3),* 307-318.

**Prahalad, C.K. and Ramaswamy V. 2000.** Co-Opting Customer Competence. *Harvard Business Review, 78(1),* 79-87.

**Prahalad, C. K. and Ramaswamy, V. 2003.** The new frontier of experience innovation. *MIT Sloan Management Review, 44(4),* 12-18.

**Prahalad, C. K. and Ramaswamy, V. 2004.** ***The** future of competition - co-creating unique value with customers.* Boston: Harvard Business School Press.

**Prahalad, C. K. and Ramaswamy, V. 2004a.** Co-creating unique value with customers. *Strategy & Leadership, 32(3),* 4-9.

**Prahalad, C. K. and Ramaswamy, V., 2004b.** Co-creation experiences: the next practice in value creation. *Journal of Interactive Marketing, 18(3*), 5-14.

**Prahalad, C. K. and Krishnan, M. S. 2008.** The new age of innovation. New York: McGraw Hill.

**Ramaswamy, V. 2009.** Leading the transformation to co-creation of value. *Strategy & Leadership, 37(2),* 32-37.

**Ramaswamy, V. and Gouillart, F. 2010.** Building the co-creative enterprise. *Harvard business review, October*, 100-109.

**Reinard, J. 2006.** *Communication Research Statistics,* Thousand Oaks, CA, Sage Publications.

**Romero, D. and Molina, A. 2009.** Value co-creation and co-innovation: Linking networked organisations and customer communities. Leveraging Knowledge for Innovation in Collaborative Networks, *IFIP Advances in Information and Communication Technology, Vol. 307,* 401-412.

**Saviotti, P. and Metcalfe, J. 1984.** A theoretical approach to the construction of technological output indicators, *Research Policy, Vol. 13,* 141–151.

**Sawhney, M., Verona, G., and Prandelli, E. 2005.** Collaborating to create: The Internet as a platform for customer engagement in product innovation. *Journal of Interactive Marketing, 19(4),* 4–17.

**Selberg, E. 1999.** Towards Comprehensive Web Search. PhD thesis, University of Washington.

**Souitaris, V. 2002.** Technological trajectories as moderators of firm-level determinants of innovation", *Research Policy, Vol. 31,* 877-898.

**Tanev, S., Knudsen, M. and Gerstlberger, W. 2009.** Value co-creation as part of an integrative vision for innovation management. Special Issue on Value Co-creation, *Open source business review online journal, December*. Available at: http://www.osbr.ca/ojs/index.php/osbr/article/view/1014/975

**Tanev, S., Durchev, P., Milyakov, H., Ruskov, P., Allen, S., and Bailetti, T. 2010.** Value Co-creation in Open Source Firms ... and Beyond. *Open Source Innovation (Beyond Software) Workshop*, Strasbourg, France, Feb. 24-25, 2010.

**Tidd, J., Bessant, J. and Pavitt, K. 2001.** *Managing innovation – Integrating Technological, Market and Organizational Change*, Second Edition, Chichester: John Wiley & Sons.

**Vargo, S. L. and Lusch, R. F. 2004.** Evolving to a new dominant logic for marketing. Journal of Marketing, *68 (January),* 1-17.

**Vargo, S. L. (2007).** On a theory of markets and marketing: From positively normative to normatively positive. *Australasian Journal of Marketing, 15(1),* 53-60.

**Vaughan, L. and Thelwall, M. 2004.** Search engine coverage bias: evidence and possible causes. *Information Processing & Management, 40(4),* 693-707.

**Vaughan, L. and Zhang, Y. 2007.** Equal Representation by Search Engines? A Comparison of Websites across Countries and Domains. *Journal of Computer-Mediated Communication, 12(3),* 888-909.

**Venkatesh, A., Penaloza, L., and Firat, F. 2006.** The market as a sign system and the logic of the market. In R. F. Lusch, & S. L. Vargo (Eds.), *The service-dominant logic of marketing: Dialog, debate, and directions (pp. 251–265).* Armonk, NY: ME Sharpe.

**von Hippel, E. 2001.** User toolkits for innovation. *Journal of product innovation management, 18 (4),* 247-257.

**von Hippel, E. 2006.** Why many users want custom products. Ch. 3 in Democratization of innovation, Cambridge: MIT Press, 33-43.

**von Hippel, E. 2007.** Horizontal innovation networks-by and for users. *Industrial and Corporate Change, 16(2),* 293-315.

**Wasko, M. M., and Faraj, S. 2005.** Why should I share? Examining social capital and knowledge contribution in electronic networks of practice. *MIS Quarterly, 29(1)*, 1–23.

**Weerawardena, J. and McColl-Kennedy, J. 2002.** New service development and competitive advantage: a conceptual model, *Australian Marketing Journal, Vol. 10,* 13-23.

**Witten, I. H. and Frank, E. 2005.** *Data Mining: Practical Machine Learning Tools and Techniques.* CA: Morgan Kaufmann.

# Appendix A: Research Sample

| Organization URL | Organization Type |
|---|---|
| http://2bits.com | OSS |
| http://apatar.com | OSS |
| http://aws.amazon.com/ | |
| http://aws.typepad.com/ | |
| http://bpocanada.com | OSS |
| http://code.google.com/android/ | OSS |
| http://craigslistfoundation.org/ | |
| http://developer.amd.com/ | Eclipse |
| http://developer.apple.com/ | |
| http://eracks.com | OSS |
| http://extjs.com | OSS |
| http://ez.no | OSS |
| http://factory.lego.com/ | |
| http://fedoraproject.org/ | OSS |
| http://forum.developers.facebook.com/ | OSS |
| http://groups.google.com/group/android-developers?pli=1 | OSS |
| http://ifs.hsr.ch/ | Eclipse |
| http://kiva.org/ | |
| http://liblime.com | OSS |
| http://namethis.com/ | |
| http://nikeplus.nike.com/nikeplus/?locale=en_us | |
| http://openmethods.com/ | OSS Eclipse |
| http://osuosl.org/ | OSS Eclipse |
| http://queue.acm.org/ | Eclipse |
| http://rcp-company.com/ | OSS Eclipse |
| http://redmondmediagroup.com/ | Eclipse |
| http://secondlife.com/ | |
| http://secondlifegrid.net/ | OSS Eclipse |
| http://software-support.biz/ | Eclipse |
| http://trolltech.com | OSS |
| http://us.etrade.com/ | |
| http://wso2.com | OSS |
| http://www.access-company.com/ | OSS Eclipse |
| http://www.accurev.com/ | Eclipse |

| | |
|---|---|
| http //www adacore com/ | Eclıpse |
| http //www adıdas com/ | |
| http //www adobe com/ | Eclıpse |
| http //www akazaresearch com | OSS |
| http //www aldon com/ | Eclıpse |
| http //www alfresco com | OSS |
| http //www alkacon com | OSS |
| http //www altera com/ | |
| http //www ancıtconsultıng com/ | Eclıpse |
| http //www anddev org/ | OSS |
| http //www andrena de/ | OSS Eclıpse |
| http //www androıd com/ | OSS |
| http //www anyware-tech com/ | OSS Eclıpse |
| http //www apache org/ | OSS |
| http //www aptana com/ | Eclıpse |
| http //www aras com | OSS |
| http //www archıpelago com/ | |
| http //www arm com/ | Eclıpse |
| http //www astaro com | OSS |
| http //www atmel com/ | |
| http //www avantsoft com/ | Eclıpse |
| http //www bandxı com/ | OSS Eclıpse |
| http //www basecamphq com/ | |
| http //www bırt-exchange com/ | Eclıpse |
| http //www blackducksoftware com/ | Eclıpse |
| http //www bluage com/ | Eclıpse |
| http //www bluenog com/ | Eclıpse |
| http //www borland com/ | Eclıpse |
| http //www bredex de/en/ | Eclıpse |
| http //www brocade com/ | OSS Eclıpse |
| http //www buıldabear com/ | |
| http //www bzmedıa com/ | Eclıpse |
| http //www ca com/ | Eclıpse |
| http //www cardınal com/ | |
| http //www cemexusa com/ | |
| http //www cenıt de/ | Eclıpse |
| http //www cıgnex com | OSS |
| http //www cısco com/ | |

| | |
|---|---|
| http://www.cleversafe.com | OSS |
| http://www.cloudsmith.com/ | Eclipse |
| http://www.coach.com/ | |
| http://www.collab.net/ | Eclipse |
| http://www.compeople.de/ | Eclipse |
| http://www.compiere.com | OSS |
| http://www.compuware.com/ | Eclipse |
| http://www.convergys.com/ | |
| http://www.crowdspring.com/ | |
| http://www.curl.com/ | Eclipse |
| http://www.cypress.com/ | |
| http://www.db4o.com | OSS |
| http://www.ddci.com/ | Eclipse |
| http://www.deere.com/ | |
| http://www.dell.com/ | |
| http://www.dfki.de/web/ | Eclipse |
| http://www.digium.com | OSS |
| http://www.diyhappy.com/ | |
| http://www.doityourself.com/did-it-myself | |
| http://www.dsdm.org/ | Eclipse |
| http://www.eads.com/ | Eclipse |
| http://www.earthweb.com/ | |
| http://www.eclipse.org/ | Eclipse |
| http://www.embarcadero.com/ | Eclipse |
| http://www.empolis.com/ | Eclipse |
| http://www.emusoftware.com | OSS |
| http://www.enea.com/ | Eclipse |
| http://www.enfoldsystems.com | OSS |
| http://www.enterprisedb.com | OSS |
| http://www.ericsson.com/ | Eclipse |
| http://www.eteration.com/ | Eclipse |
| http://www.etsy.com/storque/ | OSS |
| http://www.excelsior-usa.com/ | Eclipse |
| http://www.exist.com/ | OSS Eclipse |
| http://www.fokus.fraunhofer.de/ | Eclipse |
| http://www.freescale.com/ | Eclipse |
| http://www.froglogic.com/ | Eclipse |
| http://www.fujitsu.com/ | Eclipse |

| | |
|---|---|
| http://www.funambol.com | OSS |
| http://www.gehealthcare.com/usen/ | |
| http://www.gepower.com/ | |
| http://www.gigaspaces.com | OSS |
| http://www.greenplum.com | OSS |
| http://www.groundworkopensource.com | OSS |
| http://www.guardiandigital.com | OSS |
| http://www.harley-davidson.com/ | |
| http://www.healthnet.com/ | |
| http://www.hitachi.com/ | Eclipse |
| http://www.hp.com/ | Eclipse |
| http://www.ibm.com/ | Eclipse |
| http://www.ibs.net/ | Eclipse |
| http://www.icesoft.com | OSS |
| http://www.ideo.com/ | |
| http://www.ikea.com/ | |
| http://www.ikea.com/ms/en_US/ | |
| http://www.ilog.com/ | Eclipse |
| http://www.ingres.com/ | OSS Eclipse |
| http://www.innocentive.com/ | |
| http://www.innoopract.com/ | Eclipse |
| http://www.innovations-software.com/ | Eclipse |
| http://www.instantiations.com/ | Eclipse |
| http://www.instinctools.com/ | Eclipse |
| http://www.instructables.com/ | |
| http://www.intalio.com/ | Eclipse |
| http://www.intel.com/ | Eclipse |
| http://www.intervoice.com/ | OSS Eclipse |
| http://www.iona.com/ | Eclipse |
| http://www.itemis.com/ | Eclipse |
| http://www.iwaysoftware.com/ | Eclipse |
| http://www.jitterbit.com | OSS |
| http://www.jumpbox.com | OSS |
| http://www.kestral.com.au/ | Eclipse |
| http://www.klocwork.com/ | Eclipse |
| http://www.knithappens.com/ | |
| http://www.knowledgetree.com | OSS |
| http://www.kpitcummins.com/ | Eclipse |

| | |
|---|---|
| http://www.krugle.com/ | Eclipse |
| http://www.landsend.com/ | |
| http://www.laszlosystems.com/ | OSS Eclipse |
| http://www.latticesemi.com/ | |
| http://www.leapfrog.com/ | |
| http://www.lego.com/ | |
| http://www.lifung.com/ | |
| http://www.linux.org/ | OSS |
| http://www.linux.org/groups/usa/ | OSS |
| http://www.linuxit.com | OSS |
| http://www.lombardisoftware.com/ | Eclipse |
| http://www.lsi.com/ | Eclipse |
| http://www.lulu.com/ | |
| http://www.lynuxworks.com/ | Eclipse |
| http://www.macraigor.com/ | Eclipse |
| http://www.makezine.com/ | |
| http://www.making-greeting-cards.com/ | |
| http://www.mbproject.org/ | Eclipse |
| http://www.mdstec.com/ | Eclipse |
| http://www.medsphere.com | OSS |
| http://www.medtronic.com/ | |
| http://www.mentor.com/ | Eclipse |
| http://www.microdoc.com/ | Eclipse |
| http://www.microfocus.com/ | Eclipse |
| http://www.mirant.com/ | |
| http://www.misysbanking.com/ | OSS Eclipse |
| http://www.mks.com/ | Eclipse |
| http://www.motorola.com/ | Eclipse |
| http://www.mozilla.com/en-US/ | OSS |
| http://www.mulesource.com | OSS |
| http://www.mvista.com/ | OSS Eclipse |
| http://www.mysql.com/ | OSS Eclipse |
| http://www.nec.com/ | Eclipse |
| http://www.netapp.com/ | Eclipse |
| http://www.netapp.com/us/ | |
| http://www.netflix.com/ | |
| http://www.nexaweb.com/ | OSS Eclipse |
| http://www.ninesigma.com/ | |

| | |
|---|---|
| http://www.nokiausa.com/ | |
| http://www.novell.com/ | Eclipse |
| http://www.nuxeo.com | OSS |
| http://www.obeo.fr/ | Eclipse |
| http://www.objectweb.org/ | OSS Eclipse |
| http://www.ocsystems.com/ | Eclipse |
| http://www.ocwconsortium.org/ | |
| http://www.omg.org/ | Eclipse |
| http://www.onstar.com/ | |
| http://www.openbravo.com | OSS |
| http://www.openhandsetalliance.com/ | OSS |
| http://www.openlogic.com | OSS |
| http://www.openmakesoftware.com/ | OSS Eclipse |
| http://www.opensocial.org/ | OSS |
| http://www.opensystems-publishing.com/ | Eclipse |
| http://www.open-xchange.com | OSS |
| http://www.oracle.com/ | Eclipse |
| http://www.orangehrm.com | OSS |
| http://www.osgi.org/ | Eclipse |
| http://www.palamida.com/ | OSS Eclipse |
| http://www.paragent.com | OSS |
| http://www.parasoft.com/ | Eclipse |
| http://www.pentaho.com | OSS |
| http://www.perforce.com/ | Eclipse |
| http://www.polarion.com/ | Eclipse |
| http://www.ponoko.com/ | |
| http://www.progress.com/ | OSS Eclipse |
| http://www.project-open.com | OSS |
| http://www.prosyst.com/ | OSS Eclipse |
| http://www.protecode.com/ | Eclipse |
| http://www.qnx.com/ | OSS Eclipse |
| http://www.quest.com/ | Eclipse |
| http://www.radview.com | OSS |
| http://www.reardencommerce.com/ | |
| http://www.redhat.com/ | OSS |
| http://www.rei.com/ | |
| http://www.remain.nl/ | Eclipse |
| http://www.replaysolutions.com/ | Eclipse |

| | |
|---|---|
| http://www.rim.com/ | Eclipse |
| http://www.rpath.com | OSS |
| http://www.rsmart.com | OSS |
| http://www.rtcgroup.com/ | Eclipse |
| http://www.safe.com | OSS |
| http://www.salesforce.com/ | Eclipse |
| http://www.sap.com/ | Eclipse |
| http://www.sas.com/ | Eclipse |
| http://www.scalix.com | OSS |
| http://www.serena.com/ | Eclipse |
| http://www.sigs-datacom.de/ | Eclipse |
| http://www.sixapart.com | OSS |
| http://www.skywaysoftware.com/ | OSS Eclipse |
| http://www.slickedit.com/ | Eclipse |
| http://www.snaplogic.com | OSS |
| http://www.softwareag.com/ | Eclipse |
| http://www.sonyericsson.com/ | Eclipse |
| http://www.soyatec.com/ | OSS Eclipse |
| http://www.spikesource.com/ | OSS Eclipse |
| http://www.spotrunner.com/ | |
| http://www.springsource.com/ | OSS Eclipse |
| http://www.st.com/ | Eclipse |
| http://www.starstandard.org/ | Eclipse |
| http://www.sugarcrm.com | OSS |
| http://www.sumerset.com/ | |
| http://www.sybase.com/ | Eclipse |
| http://www.symbian.com/ | OSS |
| http://www.talend.com | OSS |
| http://www.teamprise.com/ | Eclipse |
| http://www.tensilica.com/ | Eclipse |
| http://www.tesco.com/ | |
| http://www.thalesgroup.com/ | Eclipse |
| http://www.thinkquest.org/library/ | |
| http://www.threadless.com/ | |
| http://www.threadless.com/news | |
| http://www.ti.com/ | Eclipse |
| http://www.tibco.com/ | Eclipse |
| http://www.tietoenator.com/ | Eclipse |

| | |
|---|---|
| http://www.timesys.com/ | OSS Eclipse |
| http://www.toshiba.com/ | Eclipse |
| http://www.travelocity.com/ | |
| http://www.trixbox.com | OSS |
| http://www.tsmc.com | |
| http://www.tutorvista.com/ | |
| http://www.ubuntu.com/ | OSS |
| http://www.umc.com/ | |
| http://www.varien.com | OSS |
| http://www.virtualiron.com | OSS |
| http://www.virtuallogix.com/ | Eclipse |
| http://www.virtutech.com/ | Eclipse |
| http://www.vmware.com/ | |
| http://www.vyatta.com | OSS |
| http://www.wavemaker.com | OSS |
| http://www.weiglewilczek.com/ | Eclipse |
| http://www.windriver.com/ | Eclipse |
| http://www.xilinx.com/ | |
| http://www.xwiki.com | OSS |
| http://www.zend.com/en/ | OSS Eclipse |
| http://www.zenoss.com | OSS |
| http://www.zensar.com/ | Eclipse |
| http://www.zeroc.com | OSS |
| http://www.zimbra.com | OSS |
| http://www.zmanda.com | OSS |

## Appendix B: Yahoo! regression results

### *Yahoo! online innovation metric vs. co-creation components*

| Model Summary | | | | | |
|---|---|---|---|---|---|
| | Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
| All firms | 1 | .524 | .274 | .271 | .632 |
| | 2 | .539 | .291 | **.285** | .626 |
| GEN firms | 1 | .376 | **.141** | .127 | .817 |
| OSS firms | 1 | .615 | .379 | .375 | .550 |
| | 2 | .630 | .397 | **.390** | .543 |

| Coefficients | | | | | | | |
|---|---|---|---|---|---|---|---|
| Model | | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
| | | | B | Std. Error | Beta | | |
| All firms | 1 | (Constant) | 1.988 | .123 | | 16.196 | .000 |
| | | Comp 1 | .321 | .032 | .524 | 9.893 | .000 |
| | 2 | (Constant) | **1.932** | .124 | | 15.626 | .000 |
| | | Comp 1 | **.214** | .054 | .350 | 3.964 | .000 |
| | | Comp 2 | **.167** | .068 | .216 | 2.449 | .015 |
| GEN firms | 1 | (Constant) | **2.419** | .246 | | 9.812 | .000 |
| | | Comp 2 | **.271** | .086 | .376 | 3.141 | .003 |
| OSS firms | 1 | (Constant) | 1.769 | .131 | | 13.464 | .000 |
| | | Comp 1 | .371 | .034 | .615 | 10.955 | .000 |
| | 2 | (Constant) | **1.506** | .169 | | 8.900 | .000 |
| | | Comp 1 | **.334** | .037 | .555 | 9.114 | .000 |
| | | Comp 3 | **.209** | .086 | .147 | 2.421 | .016 |

### *Yahoo! conventional metrics vs. co-creation components*

| **New Products vs. Value co-creation** | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Model Summary** | | | | | | | |
| | Model | | R | R Square | Adjusted R Square | Std. Error of the Estimate | |
| All firms | No statistical significance | | | | | | |
| GEN firms | 1 | | .314 | **.099** | .084 | .855 | |
| OSS firms | No statistical significance | | | | | | |
| **Coefficients** | | | | | | | |
| Model | | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig |
| | | | B | Std. Error | Beta | | |
| All firms | No statistical significance | | | | | | |
| GEN firms | 1 | (Constant) | **.016** | .364 | | .044 | .965 |
| | | Comp 3 | **.539** | .212 | .314 | 2.543 | .014 |
| OSS firms | No statistical significance | | | | | | |
| **New Services vs. Value co-creation** | | | | | | | |

| Model Summary | | | | | |
|---|---|---|---|---|---|
| | Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
| All firms | No statistical significance | | | | |
| GEN firms | 1 | .471 | **.222** | .209 | .735 |
| OSS firms | No statistical significance | | | | |

| Coefficients | | | | | | | |
|---|---|---|---|---|---|---|---|
| Model | | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig |
| | | | B | Std. Error | Beta | | |
| All firms | No statistical significance | | | | | | |
| GEN firms | 1 | (Constant) | **-.417** | .313 | | -1.33 | .188 |
| | | Comp 3 | **.748** | .182 | .471 | 4.106 | .000 |
| OSS firms | No statistical significance | | | | | | |

**New Processes vs. Value co-creation**

| Model Summary | | | | | |
|---|---|---|---|---|---|
| | Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
| All firms | No statistical significance | | | | |
| GEN firms | No statistical significance | | | | |
| OSS firms | No statistical significance | | | | |

| Coefficients | | | | | |
|---|---|---|---|---|---|
| Model | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
| | B | Std. Error | Beta | | |
| All firms | No statistical significance | | | | |
| GEN firms | No statistical significance | | | | |
| OSS firms | No statistical significance | | | | |

**Total Innovation (new products + new services + new processes) vs. Value co-creation**

| Model Summary | | | | | |
|---|---|---|---|---|---|
| | Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
| All firms | No statistical significance | | | | |
| GEN firms | 1 | .482 | **.232** | .219 | 1.390 |
| OSS firms | No statistical significance | | | | |

| Coefficients | | | | | | | |
|---|---|---|---|---|---|---|---|
| Model | | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
| | | | B | Std. Error | Beta | | |
| All firms | No statistical significance | | | | | | |
| GEN firms | 1 | (Constant) | **-.201** | .591 | | -.340 | .735 |
| | | Comp 3 | **1.456** | .344 | .482 | 4.226 | .000 |
| OSS firms | No statistical significance | | | | | | |