International Institute for Applied Knowledge Management

Refereed Paper Proceedings - KM Conference 2017 – Novo Mesto, Slovenia

A Publication of the International Institute for Applied Knowledge Management

Refereed Paper Proceedings

Sponsors

International Institute for Applied Knowledge Management

Refereed Paper Proceedings - KM Conference 2017 – Novo Mesto, Slovenia

A Publication of the International Institute for Applied Knowledge Management

Table of Contents

International Institute for Applied Knowledge Management

Refereed Paper Proceedings - KM Conference 2017 – Novo Mesto, Slovenia

A Publication of the International Institute for Applied Knowledge Management

Conference Chairs, Program Committee, and Editorial Team

Conference Co-Chairs

Bostjan Delak
Faculty of Information Studies, Slovenia
bostjan.delak@fis.unm.si

Meliha Handžic
International BURCH University, Bosnia & Herzegovina
meliha.handzic@ibu.edu.ba

Vili Podgorelec
University of Maribor, Slovenia
Vili.Podgorelec@um.si

KM2017 Conference Organizers and Coordinators

Shonda Brown
Middle Georgia State University, USA
Shonda.Brown@mga.edu

Michelle Ramim
Middle Georgia State University, USA
michelle.ramim@gmail.com

Nathan White
Central Washington University, USA
nathan.white@cwu.edu

Online Journal of Applied Knowledge Management (OJAKM)

Meir Russ – OJAKM Editor-in-Chief
University of Wisconsin - Green Bay, IL, USA
russm@uwgb.edu

Alex Koohang – Director of Publications
Middle Georgia State University, USA
alex.koohang@mga.edu

Yair Levy – Senior Editor
Nova Southeastern University, FL, USA
levyy@nova.edu

Program Committee Co-Chairs

Nitza Geri
The Open University of Israel, Israel
nitzage@openu.ac.il

Alex Koohang
Middle Georgia State University, USA
alex.koohang@mga.edu

Yair Levy
Nova Southeastern University, USA
levyy@nova.edu

Doctoral Consortium Co-Chairs

Zoran Levnajic
Faculty of Information Studies, Slovenia
zoran.levnajic@fis.unm.si

Vladislav Rajkovic
University of Maribor, Slovenia
vladislav.rajkovic@gmail.com

KM2017 Program Committee Members

Paul Alpar	Philipps University at Marburg, Germany
Alvaro Arenas	IE Business School, Spain
Graziano Aretusi	Marine Protected Area "Torre del Cerrano", Italy
Shauna Beaudin	Southern New Hampshire University, USA
Ofir Ben Assuli	Ono Academic College, Israel
Carlene Blackwood-Brown	Nova Southeastern University, USA
Ina Blau	The Open University of Israel, Israel
Steve Bronsburg	Nova Southeastern University, USA
Shonda Brown	Middle Georgia State University & CIGNA Healthcare, USA
Melissa Carlton	Nova Southeastern University, USA
Witold Chmielarz	Management Faculty of Warsaw University, Poland
Dimitar Christozov	American University of Bulgaria, Bulgaria
Karla Clarke	Nova Southeastern University, USA
Anat Cohen	Tel-Aviv University, Israel
Beata Czarnacka-Chrobot	Warsaw School of Economics, Poland
Carla Curado	ISEG - University of Lisbon, Portugal
Bostjan Delak	Faculty of information studies, Novo Mesto, Slovenia
Gary Delorenzo	California University of Pennsylvania, USA
Helena Dudycz	Wroclaw University of Economics, Poland

Monika Eisenbardt	University of Economics in Katowice, Poland
Yoram Eshet-Alkalai	The Open University of Israel, Israel
Ruti Gafni	Tel-Aviv Yaffo Academic College, Israel
John Girard	Middle Georgia State University, USA
Michal Golinski	Warsaw School of Economics, Poland
Jodi Goode	Nova Southeastern University, USA
Tanja Grublješic	University of Ljubljana, Slovenia
Jose Luis Guerrero-Cusumano	Georgetown University, USA
Robert Hambly	Nova Southeastern University, USA
Wilnelia Hernandez	University of Turabo, Puerto Rico
Jeretta Horn Nord	Oklahoma State University, USA
Anat hovav	Korea University, Korea
Pedro Isaias	University of Queensland, Australia
Jurij Jaklic	Ljubljana University, Slovenia
Dorota Jelonek	Czestochowa University of Technology, Poland
Oliver Jokisch	Leipzig University of Telecommunications, Germany
Deanna Klein	Minot State University, USA
Mirjana Kljajic Borštnar	University of Maribor, Slovenia
Benjamin Kralj	Institut 49, Slovenia
Gila Kurtz	The College for Academic Studies, Israel
Myungjae Kwak	Middle Georgia State University, USA
Anat Lerner	The Open University of Israel, Israel
Jay Liebowitz	Harrisburg University of Science and Technology, USA
Mike Lohle	University of Bridgeport, USA
Christiaan Maasdorp	Stellenbosch University, South Africa
Eliel Melon	University of Puerto Rico, Puerto Rico
Krzysztof Michalik	University of Economics in Katowice, Poland
Roisin Mullins	University of Wales Trinity Saint David, United Kingdom
Federico Niccolini	University of Pisa, Italy
Daryl Nord	Oklahoma State University, USA
Joanna Paliszkiewicz	Warsaw University of Life Sciences, Poland
Luka Pavlic	University of Maribor, Slovenia
Ilona Paweloszek	Czestochowa University of Technology, Poland
Nuno Pena	ISEG - University of Lisbon, Portugal
Paula Peres	Polytechnic Institute of Porto, Portugal
Donatella Persico	National Research Council, Italy
Winnie Ng Picoto	ISEG - University of Lisbon, Portugal
Michal Pietrzak	Warsaw University of Life Sciences, Poland
Vili Podgorelec	University of Maribor, Slovenia
Daphne Raban	University of Haifa, Israel

International Institute for Applied Knowledge Management

Refereed Paper Proceedings - KM Conference 2017 – Novo Mesto, Slovenia

A Publication of the International Institute for Applied Knowledge Management

Knowledge sharing or knowledge protection? The effects of cyber regulations and security policies on firms' market orientation and performance

[Research-in-Progress]

Anat Hovav, Korea University, South Korea, anatzh@korea.ac.kr

Itzhak Gnizy, Ono Academic College, Israel, itzikgn@gmail.com

Abstract

Firms today operate in data-rich environments. Information has become one of the major strategic business assets. Intra- and inter- organizational sharing of information and knowledge offer opportunities for companies to achieve competitive advantage and financial benefits. However, extensive data sharing between organizations also poses concerns about consumer privacy and data security. Scholars have also called on to identify directions for data analytics methods that focus on customers' privacy and data security. The collection of marketing big data, their examination via marketing analytics methods (e.g., examining their applications to structured & unstructured data generated internally or externally), and their potential to support marketing decisions relate to firms' marketing concept is termed market orientation (MO). MO is defined as the generation and dissemination of, and responsiveness to market knowledge. Although marketing research suggest that MO increases firm performance, recent studies argued that certain environmental conditions may moderate the relationship between MO and performance. In this research-in-progress, we propose a model that examines the moderating effects of privacy regulations and information security policies on the relationships between MO and firm's performance. Privacy regulations are external to the organization and are often beyond management control. Information security policies are internal to the organization and may be adjusted to meet organizational strategic goals. Therefore, we expect that privacy regulations will negatively affect all three components of MO's relationships with a firm's profits, while security policies will negatively affect the relationship between knowledge dissemination and responsiveness with a firm's profits.

Keywords: Market orientation, cyber regulations, security policies, and firm performance

Introduction

One of the most notable change caused by the advent of Web 2.0 technologies is the development of a 'river' of information (Day, 2011; Klingberg, 2009; Micu, Coulter, & Price, 2011), also known as big data (Leeflang, Verhoef, Dahlström, & Freundt, 2014), where businesses and individuals promote their products, services, opinions, reviews and blogs, creating a wealth of easily accessible information. Firms today operate in data-rich

environments. Information has become one of the major strategic business assets. Intra- and inter- organizational sharing of information and knowledge offer opportunities for companies to achieve competitive advantage and financial benefits. By utilizing various information assets, firms can create value for their customers and improve their innovation processes (Verhoef, Kooge, & Walk, 2016). However, extensive data sharing between organizations also poses concerns about consumer privacy and data security. Scholars have also called on to identify directions for data analytics methods that focus on customers' privacy and data security (Wedel & Kannan, 2016).

Data, and by implication big data, are the building blocks of firms' knowledge. The knowledge management process comprises of three phases, namely, knowledge creation, knowledge sharing, and knowledge application (Hayton & Cacciotti, 2014). In the marketing area, market orientation (MO) is a basic strategic concept employed by many firms. MO comprises of three phases: intelligence generation, intelligence dissemination, and responsiveness (Kohli, Jaworski, & Kumar, 1993). Firms often apply novel business models and profit structures, which rely on the availability, utilization and sharing of such knowledge. For example, companies such as Google.com and Facebook.com do not charge for the use of their direct services. However, these companies leverage the data/knowledge that they collect from customers and generate incomes from information sharing with their partners. While most governments in the developed world regulate firms in the healthcare and finance industries, firms in other industries are rarely regulated. That is, in most developed countries, the use and sharing of financial data and PII (personally identifiable information) collected by organizations is regulated. However, the use and sharing of marketing information such as buying habits, search history and recommendation analysis are not regulated in most countries. Yet, such information could infringe on the privacy of customers in much the same way that PII could. Therefore, we posit that privacy issues that are related to cyber regulations and security policies may affect firms' MO strategy.

Privacy issues may influence the reputation of companies as they are often blamed for neglecting to protect customer private information. In the Netherlands, the Dutch bank ING received strong negative publicity about their announced initiative to use their data for personalized marketing of partner firms, such as retailers. Consequently, ING had to retract this initiative (Verhoef et al., 2016). Privacy may also impacts data analytics, as firms may choose to collect less data from, or store less data of (i.e., shorter time horizons) their customers.

Market Orientation (MO)

In the context of marketing, the digitalization of business processes generates massive amounts of available data sources (Wortmann, Fischer, & Reinecke, 2016). From a marketing perspective, marketing intelligence (MI) is the organization's ability to acquire internal and external information regarding its customers, competitors, markets and industry. Furthermore, to achieve strategic value from MI, organizations should be able to fully analyze, assess and utilize the collected information to enhance the company's strategic decision-making process and gain competitive advantage (Huster, 2005). Much of the marketing intelligence comes from electronic

media such as social networks, daily customer transactions (Royle & Laing, 2014), as well as Internet of Things (IoT) and computer-generated information. From the knowledge application perspective, business-to-business (B2B) organizations use social media (e.g., Facebook, Twitter, & LinkedIn) to primarily attract new customers and cultivate relationships (Royle & Laing, 2014). MO comprises of organization-wide generation, dissemination, and responsiveness to market knowledge (Kohli & Jaworski, 1990). Intelligence generation refer to information collection activities used by the organization such as using market research firm, polling customers, conducting focus groups, detecting shift in consumer taste, regulations, technology and industry structure. Intelligence dissemination refers to the timely sharing of knowledge among business units, departments and partners. Responsiveness refers to the organization's ability to respond to the intelligence it has gathered and disseminated. Such response could be in the form of a change to pricing structure, product or service offerings, response to a competitor's strategy by launching a new and timely marketing campaign, and interaction with customers to resolve complaints (or churn). For MO to facilitate competitive advance and superior performance, these activities are expected to be timely and coordinated across business units, departments and strategic partners (Huster, 2005; Kohli & Jaworski, 1990).

MO that is based on big data can help the marketing organization to derive significant insights from consumer- and firm-generated digital as well as non-digital content about customers, competitors, and markets. These insights might change managerial decision-making. For example, supermarkets often collect and generate detailed consumer buying habits. This data is used as an operational and strategic tool. Operationally, these supermarkets print customized coupons and offer personalized sales. Furthermore, managers can decide on regional and local promotions. Strategically, the supermarket chain might offer new products or services based on changing consumer sentiments. The UK Retailer, Tesco, built a culture of customer data-driven decision-making (e.g., format management, category management, CRM systems, communication) into every level of the company to become one of the world's top retailers. Systematically turning loyalty card data into insights and insights into business decisions fueled Tesco's rise to the number one retailer in the UK. Tesco has created a powerful data collection engine through the combination of data obtained from loyalty cards, scanners, Web sites, and (additional) market research (Humby, Hunt, & Phillips, 2008). Tesco has been collecting 1600 million new data items monthly from 10 million cardholders and eight million transactions from 700 stores, and 50,000 stock keeping units (SKUs) weekly. Their customer insights are supported by an outsourced analytical and data storage partner. Thus, Tesco is an example of an organizational culture that seeks to use data to better understand customers.

MO and Firm Performance

Despite the fact that many meta-studies confirm the positive relationship between MO and firm performance (Ellis, 2007; Grinstein, 2008; Kirca, Jayachandran, & Bearden, 2005), the stream of research having firms' strategic orientations as the unit of analysis keeps challenging this long established positive relationship. This concerns predominantly the mixed results from the

empirical research on the MO-performance relationship when introducing institutional and environmental moderators (Cadogan, Diamantopoulos, & Siguaw, 2002; Cadogan, Sundqvist, Salminen, & Puumalainen, 2002; Rose & Shoham, 2002). In the marketing literature, it is argued that many firms need to adopt a market-oriented posture, to enable them to become more responsive to changes in consumer needs and wants (Murray, Gao, & Kotabe, 2011). The awareness of the benefits accruing from applying MO, when the firm takes into consideration customer wants and market conditions, is put forward as explanations for increased MO in firms. However, firms' ability to utilize information via MO and adapt to market changes depends on their understanding of the conditions that might impede the application of information and hinder its benefits. For example, security concerns and points of business law are barriers to adopting an online analysis approach (Royle & Laing, 2014). Whereas organizations aim to generate actionable insights from information in general and big data in particular, managers need to understand the mechanisms that may inhibit the use of MO. We posit that the challenge of privacy regulations (i.e., a firm's external factors), and security policies (i.e., a firm's internal factors) may play a role in the effects of MO on firms performance. The purpose of this study is to provide an understanding of the moderating effects of privacy regulations and security policies on the marketing concept of en-route firms' business performance, an issue that has been under-explored in the literature.

Hereunder, we postulate that besides the direct relationship between MO and firm performance, cyber regulations and security policies are argued to be institutional factors that inhibit firms from enhancing the positive effects of their market-oriented activities. After presenting our model, and hypotheses' analyses, we describe out research plan and conclude by discussing our expected contributions.

Cyber Regulations and Security Policies

Organizations often follow both security laws and internal cybersecurity policies (thereof policies). While laws are established and enforced by governments, policies are developed and enforced by the organization. Some policies are voluntary, others are driven by standardization organizations (e.g., NIST in the U.S., ISO in Europe & Asia) and consortia (e.g., PCI), and some are implemented to comply with governmental laws.

Cybersecurity and Privacy Regulations

We define cybersecurity and privacy regulations (thereof regulations) to be government laws and regulations that determine what organizations are allowed (or not allowed) to do with data and information. Such regulations often instruct organizations in the protection of data. That is, government regulations often describe what data has to be protected, in what manner, for how long, and the penalties for non-compliant behavior. Early cybersecurity laws mostly addressed the infrastructure used to manage information (i.e. computers, networks). The Health Insurance Portability and Accountability Act (HIPAA) of 1996 is the first known regulation in the U.S. to address information privacy. HIPAA was passed by the U.S. congress in 1996 and addresses

medical records and patients privacy across all healthcare related companies (providers, labs, insurance, intermediaries, etc.). As of 2017, most developed countries have adopted similar medical records and patient privacy laws. The Financial Services Modernization Act also known as the Gramm-Leach-Bliley (GLB) Act of 1999 followed with detailed regulations regarding financial information. The act calls financial institutions "to insure the security and confidentiality of customer records and information" (FFIEC, 1999). Similar regulations now exist in most developed countries and are often driven by the Basel Accords. Although most developed countries, established some security and privacy regulations, their scope and reach vary. For the sake of brevity, we will mention only some of the fundamental differences.

Opt-in versus Opt-out: in some countries, such as the U.S., individuals have to opt-out (notify any company they do business with if they do not want their data shared). Other countries such as all members of the EU have adopted the opt-in option where customers have to deliberately agree before their data can be shared. Global organizations need to comply with local rules. From a business perspective, the opt-out framework enables companies to provide personalized products and services, accurate recommendation systems and targeted marketing.

Localized versus universal laws: as noted above, most early cyber laws addressed the protection of relevant and sensitive infrastructure. However, as the U.S. government began to develop laws to protect private information such as HIPAA and GLB, they concentrated on particular industries. This model was followed by most governments but not by all. In some countries, such as India, privacy laws are content or infrastructure-based and therefore are more universal. Neither model is perfect as both fail to address new technology, new types of information and new types of breaches in a timely manner. From a business perspective, localized laws are beneficial to companies that are not regulated. Such companies have less incentive to invest in advanced security solutions than companies in regulated industries (Hovav & Gray, 2014).

Who owns the information: as mentioned above, Web 2.0 created novel business models, where information is a tradable commodity. Firms such as Google or Facebook use information they collect as a means to increase their profits. At present, these companies own the data they collect and with a few exceptions, can use it for their benefit or share it (as long as the Personally Identifiable Information (PII) data is scrubbed). As of 2017, the sharing of information depends on the governing laws of the country of residence. That is, EU citizens have to opt-in for companies to share their data. However, the company owns the data and can use it internally for product development or marketing purposes.

In May 2014, the European Court of Justice ruled against Google in a case brought by a Spanish man, who requested the removal of a link to an article (Bygrave, 2015). The "right to be forgotten" began and various countries ruled regarding digital data stored in the likes of Facebook and Google data warehouses. The EU is in the process of revising their privacy laws, asking Google to delink information not only from European versions of their website (for example, google.co.fr) but also from Google Inc. and other international domains. From a business prospective, companies whose business models rely on the value of the information

they collect will have to find a new business model. As the saying goes, "there is no free lunch." Thus, there is no guarantee that these new business models will not interfere with current technological advances such as big data analytics, cloud services and IoT.

Information Security Policies

An organizational security policy defines the rules and guidelines for the use of information assets (D'Arcy, Hovav, & Galletta, 2009). Such policies define what are the acceptable behaviors given a particular asset and a particular user role (Jeong & Hovav, 2015) and the repercussion for circumventing the above policies. While security policies address a gamut of organizational computing resources, in this study we are mostly interested in policies that address information and knowledge sharing. Knowledge sharing within the organization is often addressed in two ways. Most organizations have established access control policies. These policies determine who is allowed to access what information, when, and for what purpose (Whitman & Mattord, 2013). Access control policies are often implemented using a variety of hardware and software solutions (Hovav & Berger, 2009) to ensure the integrity and confidentiality of the information. Despite the existence of such systems, users tend to circumvent the above measures for a number of reasons (Bulgurcu, Cavusoglu, & Benbasat, 2010). The most notable reason for this circumventing behavior is that access control policies are often rigid and interfere with users' daily task. Policies are often revised periodically (Whitman & Mattord, 2013), usually once a year. All the while, increased competition and the rapid change of technology demand flexible work patterns. In summary, while at the strategic level, organizations advocate knowledge-sharing, security policies compartmentalize knowledge based on rigid "need to know" based policies.

In academic writing, information leakage has been discussed as a major issue for organizations in Siponen and Vance (2010). Industry acknowledged the risks of information leakage after the Snowden incident of 2013. Information leakage refers to the ability of users to share confidential information with external entities. Although most organizations have established policies regarding the sharing of information, enforcement of these policies is difficult for two reasons. First, once the information leaves the organization, it is practically impossible to control that information using traditional security mechanisms such as access control systems (Morin & Hovav, 2012). Second, some implicit organizational knowledge exists in peoples' heads and can be leaked verbally intentionally or accidentally. For example, an engineer might leak the details of a new design for monetary benefits (intentional) or a group of collaborators form various firms might discuss the market potential of a new technology and inadvertently reveal the launch of a new service based on that technology. While the latter is hard to control using a technical measure, organizations occasionally implement Enterprise Rights Management (ERM) systems. Such systems provide persistent controls for information assets. ERM systems may increase confidentially but reduce flexibility and knowledge-sharing capabilities (Jeon & Hovav, 2015).

Model Development

In this section, we detail the basis for our hypotheses. The first set of hypotheses (H1a, H1b, & H1c) relates to the relationships between the three dimensions of MO and firm performance. The second set of hypotheses (H2a, H2b, & H2c) relates to the moderating effect of cyber regulations on the relationship between MO and firm performance. Finally, the third set of hypotheses (H3a & H3b) depicts the moderating effect of cyber policies on the relationship between MO and firm performance. Our proposed model is depicted in Figure 1.

MO and Firm Performance

Strategic orientations employed by firms have been discussed extensively in prior research and have led to conclusions that cover a wide range of their effects on firms (Hakala, 2011). These orientations are related to firms' strategy and refer to companies' adoption of specific values, norms, and operation in certain ways. They also reflect patterns of behavior and firms' best practices. Specifically, strategic orientations are adaptive mechanisms of principles that direct firms' activities and generate the behaviors intended to ensure firms' viability and business performance (Hakala, 2011). Research identified orientations as tools that firms use to obtain resources, create capabilities, and thus achieve competitive advantages, which in turn enhance performance (Kohli & Jaworski, 1990). Strategic marketing practice and research have generally taken for granted that firms should embrace organization-wide mechanisms to achieve excessive business performance: an informal organizational mindset and culture of MO (Frösén, Luoma, Jaakkola, Tikkanen, & Aspara, 2016; Gebhardt, Carpenter, & Sherry, 2006; Narver & Slater, 1990). Significant changes with respect to data management and processing contributed to the data-oriented approach that marketing organizations began to adopt. This approach provides firms with sophisticated knowledge generation and dissemination mechanisms (Kumar, 2015). For example, the abundance and ease of data collection of detailed customer data (Kumar, 2015) and the exponential increase in storage, access and analytics capabilities enable firms to capture individual customer data and the identification of customers' lifetime behaviors for close to zero cost (Lauden & Traver, 2016). Indeed, the ability to generate and leverage customer insights is an essential challenge for ''digital'' marketers today (Leeflang et al., 2014).

MO is probably the most studied orientation among firms' strategic orientations, has attracted widespread attention and is usually recognized as a major driver that enhances firms' success and performance (Hakala, 2011). The vast majority of studies that have examined the effect of an orientation on performance demonstrated the superiority of MO in comparison to other orientations. MO is a critical concept in marketing and management and its positive effect on firm performance is well documented in various settings (e.g., Pelham & Wilson, 1996; Calantone & Knight, 2000). One of the most common conceptualizations of MO is that it comprises of organization-wide generation, dissemination, and responsiveness to market knowledge and intelligence (Kohli & Jaworski, 1990). Therefore, we hypothesize that:

H1: Market orientation is positively related to firm performance such that:

H1a: Intelligence generation is positively related to firm performance.

H1b: Intelligence dissemination is positively related to firm performance.

H1c: Intelligence responsiveness is positively related to firm performance.

Cyber Regulations and Policies

From the above discussion, it is clear that MO and cybersecurity address two, somewhat conflicting, organizational goals. The collection, sharing and use of information and knowledge enables organizations to create new business models, products and services, and customer experience. Conversely, whereas increase regulatory activity helps maintain users' privacy (a human right) such regulations could also hinder the use of information for the benefit of the organization and its customers. Similarly, organizational policies and controls limit users' ability to share information. Security policies are often rigid and are driven by standard operating procedures (SOPs). However, these policies may be ineffective in today's volatile competitive environment and for companies that rely on market orientation strategies and knowledge process management to gain competitive advantage. Therefore, we hypothesize that:

H_2: The relationship between market orientation and firm performance is negatively moderated by cyber regulations such that:

H_{2a}: Cyber regulations negatively moderate the intelligence generation-performance relationship.

H_{2b}: Cyber regulations negatively moderate the intelligence dissemination-performance relationship.

H_{2c}: Cyber regulations negatively moderate the intelligence responsiveness-performance relationship.

H_3: The relationship between market orientation and firm performance is negatively moderated by security policies such that:

H_{3a}: Security policies negatively moderate the intelligence dissemination-performance relationship.

H_{3b}: Security policies negatively moderate the intelligence responsiveness-performance relationship.

Research Plan

To measure the above model and hypotheses, we have developed a 7-point likert-based survey instrument. The target population is firms in Israel. We expect the survey to be answered by one or more top managers at each surveyed firm. The MO questions were adapted from Kohli et al. (1993), the questions for organizational performance were adapted from Gold, Malhotra, and Segars (2001), and the questions regarding security policies are loosely based on D'Arcy et al. (2009) and were adapted to the MO context. Similarly, the questions regarding the organizational regulatory environment were loosely adapted from D'Arcy et al. (2009). The survey questions

were translated to Hebrew. The translation was inspected and compared with the original questionnaire for consistency and accuracy in content and meaning. A pilot of 75 firms is planned for April 2017. If necessary, the questionnaire will be adjusted. Subsequent data collection targeting 250-300 firms is planned for May-June 2017. The data will be analyzed using WrapPLS 5.0 using linear and non-linear regressions for best model fit (Kock, 2015).

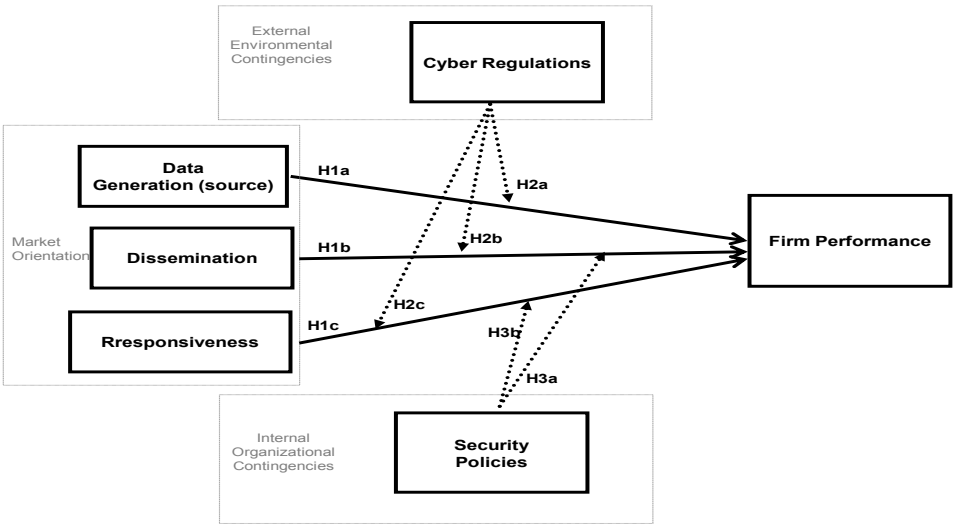


Figure 1. Proposed Research Model

Summary

Our study, when completed, will contribute to our body of knowledge in several ways. First, we extend MO's traditional view on intra-organizational data flow boundaries (e.g., information generation & sharing, dimensions of MO). While MO usually views data flow as intra-organizational, we view it broadly as intra- and inter-organizational. Second, we link between MO and digital data. There have been calls for research in marketing to increase relevance for current challenges and issues within industry (Brady, Fellenz, & Brookes, 2008; Reibstein, Day, & Wind, 2009). While marketers in the firm are blamed for not understanding technology or technology-oriented issues, which undermine their value in the organization (Keaveney, 2008), marketers are called on to acquire information technology (IT) skills and capabilities (Arons, van den Driest, & Weed, 2014). We contend that marketers should be aware of IT-oriented issues such as security regulations and policies that may impede the application of MO, which is considered a victory of marketing thought in the firm. Third, our study explores the relationship between the marketing and IT functions/processes in the firm, a topic that is scarcely studied. Companies today foster connections by putting marketing and other functions under a single

leader. For example, Motorola's Eduardo Conrado is the senior VP of both marketing and IT. "Marketing has become too important to be left just to the marketers" (Arons et al., 2014). Finally, we add new nuances to the under studied circumstances where MO might be less beneficial for firms. Specifically, our focus on cyber regulations and security policies provides additional nuance to Kohli and Jaworski's (1990) list of certain conditions where the strategic value of MO may be challenged. While the majority of research on MO points out its positive effects on performance, Murray et al. (2011) suggested that it is likely that important variables are missing from models of the relationship between MO and success/performance. Additionally, while firms are encouraged to employ MO and increase its use, there is little marketing literature on how current emerging issues such as customer privacy protection affects the diffusion of MO and its post adoption.

References

Arons, M. D. S., van den Driest, F., & Weed, K. (2014). The ultimate marketing machine. *Harvard Business Review, 7*, 1-12.

Bulgurcu, B., Cavusoglu, H., & Benbasat, I., (2010). Information security policy compliance: An empirical study of rationality-based beliefs and information security awareness. *MIS Quarterly, 34*(3), 523-548.

Brady, M., Fellenz, M. R., & Brookes, R. (2008). Researching the role of information and communications technology (ICT) in contemporary marketing practices. *Journal of Business and Industrial Marketing, 23*(2), 108–114.

Bygrave, L. A. (2015). A right to be forgotten? *Communications of the ACM*, *58*(1), 35-37.

Cadogan, J. W., Sundqvist, S., Salminen, R. T., & Puumalainen, K. (2002). Market-oriented behavior: Comparing service with product exporters. *European Journal of Marketing,* 36(9/10), 1076-1102.

Cadogan, J. W., Diamantopoulos, A., & Siguaw, J. A. (2002). Export market-oriented activities: Their antecedents and performance consequences. *Journal of International Business Studies*, 33(3), 615-626.

D'Arcy, J., Hovav, A., & Galletta, D. (2009). User awareness of security countermeasures and its impact on information systems misuse: A deterrence approach. *Information Systems Research*, 20(1), 79-98.

Day, G. S. (2011). Closing the marketing capabilities gap. *Journal of Marketing*, 75(4), 183-195.

Ellis, R. (2007). *Entropy, large deviations, and statistical mechanics*. New York, NY: Springer.

Gold, A. H., Malhotra, A., & Segars, A. H. (2001). Knowledge management: An organizational capabilities perspective. *Journal of management information systems, 18*(1), 185-214.

Grinstein, A. (2008). The effect of market orientation and its components on innovation

consequences: A meta-analysis. *Journal of the Academy of Marketing Science*, 36(2), 166-173.

Federal Financial Institutions Examination Council (FFIEC) (1999). Gramm-Leach-Bliley Bill Section 501(b). Retrived from: https://www.ffiec.gov/exam/infobase/documents/02-con-501b_gramm_leach_bliley_act-991112.pdf

Frösén, J., Luoma, J., Jaakkola, M., Tikkanen, H., & Aspara, J. (2016). What counts versus what can be counted: The complex interplay of market orientation and marketing performance measurement. *Journal of Marketing*, 80(3), 60-78.

Gebhardt, G. F., Carpenter, G. S., & Sherry, J. F. (2006). Creating a market orientation: A longitudinal, multifirm, grounded analysis of cultural transformation. *Journal of Marketing*, 70(4), 37-55.

Hakala, H. (2011). Strategic orientations in management literature: Three approaches to understanding the interaction between market, technology, entrepreneurial and learning orientations. *International Journal of Management Reviews*, 13(2), 199-217.

Hovav, A., & Gray, P. (2014). The ripple effect of an information security breach event: A stakeholder analysis. *Communications of the Association for Information Systems*, 34(50), 893-912.

Hovav, A., & Berger, R. (2009). Tutorial: Identity management systems and secured access control. *Communications of the Association for Information Systems*, *25*(1), 531-570.

Humby, C., Hunt, T., & Phillips, T. (2008). *Scoring points: How Tesco continues to win customer loyalty*. London, UK: Kogan Page Publishers.

Huster, M. (2005). Marketing intelligence: A first mover advantage, *Competitive Intelligence Magazine*, 8(2), 13-17.

Jeon, S., & Hovav, A. (2015). Empowerment or control: Reconsidering employee security policy compliance in terms of authorization. Hawaii International Conference on System Sciences (HICSS-48): January 5-8, 2015, IEEE, 3473-3482.

Keaveney, S. M. (2008). The blame game: An attribution theory approach to marketer–engineer conflict in high-technology companies. *Industrial Marketing Management, 37*(6), 653-663.

Kirca, A. H., Jayachandran, S., & Bearden, W. O. (2005). Market orientation: A meta-analytic review and assessment of its antecedents and impact on performance. *Journal of Marketing*, 69(2), 24-41.

Klingberg, T. (2009). *The overflowing brain: Information overload and the limits of working memory*. London, UK: Oxford University Press.

Knight, G. A., & Calantone, R. J. (2000). A flexible model of consumer country-of-origin

perceptions: A cross-cultural investigation. *International Marketing Review*, 17(2), 127-145.

Kock, N. (2015). A note on how to conduct a factor-based PLS-SEM analysis. *International Journal of e-Collaboration (IJeC)*, 11(3), 1-9.

Kohli, A. K., & Jaworski, B. J. (1990). Market orientation: The construct, research propositions, and managerial implications. *The Journal of Marketing*, 54(2), 1-18.

Kohli, A. K., Jaworski, B. J., & Kumar, A. (1993). MARKOR: A measure of market orientation. *Journal of Marketing research, 30*, 467-477.

Kumar, V. (2105). Evolution of marketing as a discipline: What has happened and what to look out for. *Journal of Marketing 79*(1), 1-9.

Laudon, K. C., & Traver, C. (2016). *E-Commerce 2016: Business, Technology, Society*. Boston, MA: Pearson Higher Ed.

Leeflang, P. S., Verhoef, P. C., Dahlström, P., & Freundt, T. (2014). Challenges and solutions for marketing in a digital era. *European Management Journal*, 32(1), 1-12.

Micu, C. C., Coulter, R. A., & Price, L. L. (2009). How product trial alters the effects of model attractiveness. *Journal of Advertising, 38*(2), 69-82.

Morin, J. H., & Hovav, A. (2012). Strategic value and drivers behind organizational adoption of enterprise DRM: The Korean case. *Journal of Service Science Research, 4*(1), 143-168.

Murray, J. Y., Gao, G. Y., & Kotabe, M. (2011). Market orientation and performance of export ventures: The process through marketing capabilities and competitive advantages. *Journal of the Academy of Marketing Science, 39*(2), 252-269.

Narver, J. C., & Slater, S. F. (1990). The effect of a market orientation on business profitability. *The Journal of Marketing, 54*(4), 20-35.

Pelham, A., & Wilson, D. T. (1996). A longitudinal study of the impact of market structure, strategy, and market orientation on small-firm business performance. *Journal of the Academy of Marketing Science, 24*(1), 27-44.

Reibstein, D. J., Day, G., & Wind, J. (2009). Guest editorial: Is marketing academia losing its way? *Journal of Marketing, 73*(4), 1-3.

Rose, G. M., & Shoham, A. (2002). Export performance and market orientation: Establishing an empirical link. *Journal of Business Research*, *55*(3), 217-225.

Royle, J., & Laing, A. (2014). The digital marketing skills gap: Developing a Digital Marketer Model for the communication industries. *International Journal of Information Management*, *34*(2), 65-73.

Siponen, M., & Vance, A. (2010). Neutralization: New insights into the problem of employee

information systems security policy violations. *MIS Quarterly, 34*(3), 487-502.

Verhoef, P. C., Kooge, E., & Walk, N. (2016). *Creating value with big data analytics: Making smarter marketing decisions.* Abingdon, UK: Routledge.

Wedel, M., & Kannan, P. K. (2016). Marketing analytics for data-rich environments. *Journal of Marketing, 80*(6), 97-121.

Whitman, M., & Mattord, H. (2013). *Management of information security*. Boston, MA: Cengage Learning.

Authors' Biographies

Anat Zeelim-Hovav is a full professor at Korea University Business School in Seoul, South Korea. Her research interests include behavioral information security, risk assessment, innovation management, and Futures research. Dr. Hovav has published in internationally refereed journals and conferences and is the winner of the 2013 citation of excellence award.

Dr. Itzhak Gnizy is a lecturer in the fields of Information Technology and Marketing and has a longstanding professional career in the High-Tech and Information Technology industry, having reached top management positions in leading companies.

International Institute for Applied Knowledge Management

Refereed Paper Proceedings - KM Conference 2017 – Novo Mesto, Slovenia

A Publication of the International Institute for Applied Knowledge Management

Towards empirical exploration of employee's cybersecurity countermeasures awareness and skills: Differences in training delivery method and program type

[Research-in-Progress]

Jodi Goode, Nova Southeastern University, USA, jp1587@nova.edu

Yair Levy, Nova Southeastern University, USA, levyy@nova.edu

Abstract

The protection of an organization's information systems and assets from cybersecurity threats is increasingly important in today's world, especially as they become more reliant upon information technology for daily operations. Employees who lack knowledge and skillsets are recognized as the most significant threat vector for cyber-attacks. Therefore, they are being targeted with continually evolving threats, such as social engineering attacks. However, employees cannot be held responsible for cybersecurity practices if they are not provided the security education and training program (SETA) to acquire the knowledge as well as skills, which allow for identification of cybersecurity threats along with the proper course of action. In addition, awareness of the importance of cybersecurity, the responsibility of protecting organizational data, and of emerging cyber threats is quickly becoming essential. This work-in-progress research will be conducted in three phases and will utilize a mixed method approach combining an expert panel, developmental research, in addition to quantitative data collection. This study will empirically assess if there are any significant differences on employees' cybersecurity countermeasures awareness and cybersecurity skills based on the use of two SETA program types (traditional vs. socio-technical) and two SETA delivery methods (face-to-face & online). Recommendations for SETA program type and delivery method as a result of data analysis will be provided.

Keywords: Cybersecurity training, cybersecurity skills, cybersecurity countermeasures awareness, information security, security education, training, and awareness (SETA)

Introduction

Von Solms and Van Niekerk (2013) put forth the idea that the impact of cyber threats goes beyond that of traditional information security. Not only can an individual be personally harmed, but society as a whole can also be directly affected by cyber-attacks. As technology becomes increasingly critical for achieving business objectives, state of the art security systems can provide a false sense of protection to organizations. In addition, Hovav and Gray (2014) contended that cyber-attacks not only effect the attacked organization but ripple through the ecosystem impacting other connected organizations, stakeholders, as well as innocent

bystanders. Organizational perspective dictates that while technical solutions are imperative, focus must be placed on the actions of information security management and on advancement toward a secure business environment from the human-centric side of cybersecurity (Ransbotham & Mitra, 2009; Carlton, Levy, Ramim, & Terrell, 2016). Information security managers are tasked with aligning the practices of employees with the desired cybersecurity posture of the organization. Thus, research must encompass the human-centric lens, as employees are often the potential targets or unintentional facilitators of cyber-attacks.

Although systematic enhancements are essential to increase the security of information systems and to strengthen protection of data within organizations, it is also critical that emphasis be placed on ways in which employees' naive cybersecurity actions may be mitigated. The seminal work of D'Arcy, Hovav, and Galletta (2009) established that implementation of a security education, training, and awareness (SETA) program is critical to the mitigation of cyber threats within an organization. Likewise, the development of cybersecurity countermeasures awareness (CCA) as well as cybersecurity skills (CyS) through SETA initiatives are imperative, although additional research is needed to determine the most valuable program type and delivery method (D'Arcy et al., 2009). Therefore, this work-in-progress study will develop a roadmap for an empirical assessment of differences in CCA along with CyS based on SETA program types and delivery methods.

Theoretical Framework

A successful approach to cybersecurity must be comprised of defenses such as the establishment and promotion of policy, security awareness campaigns, as well as training opportunities for all employees (Furnell & Clarke, 2012). D'Arcy et al. (2009) found raising employee awareness of security policies and the implementation of SETA programs were beneficial in mitigating cybersecurity threats. SETA programs can be used to empower employees, who are often cited as the weakest link in information systems (IS) security due to limited knowledge and lacking skillsets (Albrechtsen, 2007). SETA programs not only focus on raising employee awareness of responsibilities in relation to their organizations' information assets, but also train on the consequences of abuse while providing the necessary skills to help fulfill these requirements (D'Arcy & Hovav, 2007). Therefore, development of CCA and CyS through SETA initiatives is critical to the mitigation of cyber threats (D'Arcy et al., 2009). Straub and Welke (1998) used the term security countermeasures to collectively describe a mix of procedural and technical controls to mitigate IS risk. Building upon previously used security countermeasures definitions, CCA can be said to include employee awareness of cybersecurity policies, SETA programs, computer monitoring, and computer sanctions (Choi, Levy, & Hovav, 2013; D'Arcy et al., 2009). Awareness of the importance of cybersecurity, the responsibility of protecting organizational data, as well as of emerging cyber threats is quickly becoming essential as the threat landscape is increasing in sophistication at an alarming rate.

Employees cannot be held responsible for cybersecurity practices if they are not provided the

education and training to acquire skills, which allow for identification of information security threats along with the proper course of action. Boyatzis and Kolb (1991) defined skill as a "combination of ability, knowledge and experience that enables a person to do something well" (p. 280). Skill is also described as the capability to utilize knowledge, intellectual capabilities, and past experiences to perform the best course of action well in a given situation (Levy, 2005). Accordingly, cybersecurity skill "corresponds to an individual's technical knowledge, ability, and experience surrounding the hardware and software required to execute IS in protecting their information technology against damage, unauthorized use, modification, and/or exploitation" (National Initiative for Cybersecurity Careers & Studies, 2014). While computing skills have been the focus of IS literature, studies have failed to evaluate the role of skills in the mitigation of cybersecurity threats, nor measure them in an objective non-self-reported way (Choi et al., 2013).

The ultimate purpose of organizational learning is to bring about a positive change in the work environment and employees' practices by providing them the relevant knowledge along with the experiences to develop skills. IS security training is designed to effect the decisions of the individual in relation to the secure use of IS, however, many SETA programs focus on the memorization of organizational IS security policies and procedures (Parrish & Nicolas-Rocca, 2012). These typical SETA campaigns often involve coercion, fear tactics, or perception of external pressures, which previous studies found to have no influence on employee compliance with organizational IS policies (Kranz & Haeussinger, 2014). Typical SETA programs fall short in that they do not employ socio-technical philosophies, providing a means for employees to see how the training materials provided correlate to their day-to-day practices. Socio-technical philosophies embrace social and technical elements for optimal design and use of organizational systems. Training and education efforts are more effective if they not only outline what is expected, but also provide an understanding of why this is important to the individual (Parsons, McCormac, Butavicius, Pattinson, & Jerram, 2014).

While early IS research on employee training focused on traditional training methods in a classroom environment, e-learning methods are increasingly being used as an approach for the enhancement of skills and knowledge (Levy, 2006). Both online and face-to-face training delivery methods have their advantages, and in previous research, both have been found to successfully produce a motivated employee who has the skills needed to apply their training to job-related tasks (Gupta, Bostrom, & Huber, 2010). Although some researchers have found no discernable difference in learning outcomes between training delivered face-to-face vs. online (McLaren, 2004), others have found variations by discipline (Smith, Heindel, & Torres-Ayala, 2008), and delivery type (Faux & Black-Hughes, 2000). However, there seems to be insufficient research in the field of IS to determine the most successful delivery method as well as the type of program for cybersecurity focused SETA.

The majority of employees are not aware of the importance of protecting personal and organizational information or IS. Therefore, their naive cybersecurity practices reflect this lack of understanding. To this point, a study by Vance, Siponen, and Pahnila (2012) noted that more

than half of IS security breaches were caused by naive actions on the part of the individual. Vance et al. (2012) addressed a gap in the body of knowledge by examining the influence of past behavior on individuals' compliance with information policies. Vance et al. (2012) utilized the full model of protection motivation theory (PMT) to investigate the impact of past information security compliance behavior on threat appraisal and coping responses. PMT suggests that past behavior will have a significant influence on the process of accessing threats and on an individuals' ability to cope with the threat (Boer & Seydel, 1996). Protection motivation processes attempt to influence individuals' established practices (i.e. habits) and typical response. However, the work of Vance et al. (2012) was limited by the use of intention as a dependent variable, and the measurement of compliance in only four scenarios, which might not work well for all employees or in all organizational situations. This study builds on previous research by D'Arcy et al. (2009), Levy (2005), Choi et al. (2013), Carlton et al. (2016), Vance et al. (2012), as well as Dinev, Goo, Hu, and Nam (2009). PMT will serve as the foundational theory for comparison of SETA delivery method and program type on the CCA and CyS of the employee. In addition, the Delphi methodology will be utilized to validate the assessment instruments developed to measure CCA as part of the SETA programs' delivery, while CyS measure will be adopted from Carlton and Levy (2015) as well as Carlton et al. (2016).

Problem Statement, Goals, and Hypotheses

The research problem that this proposed study will address is employees' naive cybersecurity practices, which can lead to organizational hazards including financial implications, impact to business reputation, loss of company information assets, and proprietary information leakage (D'Arcy et al., 2009; Vance et al., 2012). Employees' naive cybersecurity practice is defined as unintentional mistakes made by an employee that may expose an organization to potential loss of information assets (Gundu & Flowerday, 2012). These practices may include the use of weak passwords for critical systems, visiting malware infested Websites, responding to phishing attempts, storing login information in an insecure manner, or providing confidential information to unapproved requestors.

The main goal of this research study is to empirically assess if there are any significant differences on employees' cybersecurity countermeasures awareness (CCA) and cybersecurity skills (CyS) based on the use of two SETA program types (traditional vs. socio-technical) and two SETA delivery methods (face-to-face & online). Previous research has focused on the decisions made by the individual that cause damaging effects, not out of maliciousness, but because they lack the skill level required to respond to threats in a conscious way (Stanton, Stam, Mastrangelo, & Jolton, 2005). Employee practices are a key factor in the mitigation of cybersecurity threats within the organization. Consequently, there is a need to develop good cybersecurity practices on the part of the employee and to promote compliance with information security policies (Vance et al., 2012). CCA has been found to influence cybersecurity practices by producing employees that think through and anticipate 'what if' scenarios, preparing them to apply the acquired CyS when required (Ross, 2006). Therefore, this work-in-progress study will

seek to assess if there are any significant differences on employees' CCA and CyS based on SETA program type and delivery method.

The need for this work is demonstrated by the research of Dinev et al. (2009), which focused on the impact that computer self-efficacy and virtual working status had on the deterrent effectiveness of security countermeasures (security policies, SETA programs, & computer monitoring) on computer misuse intention. Choi et al. (2013) built upon their work by expanding the research to determine the role of computer self-efficacy, CCA, and CyS on computer misuse intention. Based upon empirical findings, Choi et al. (2013) recommended additional study on the role of SETA programs on cybersecurity skills development. However, Choi et al. (2013) has several limitations. First, the construct of computer self-efficacy provides measurement not of the skill of the individual, but is a self-assessment of his/her perceptions about their capability to execute certain courses of action (Bandura, 1997). Secondly, grounded empirical studies have found the basing of research upon intention to comply with information security policies and procedures to be a significant limitation, as intention does not necessarily translate to actual behavior (Vance et al., 2012). Finally, survey based self-assessment measures have been used in other studies and were found to be generally ineffective predictors of security practice (Vance, Anderson, Kirwan, & Eargle, 2014).

Additional challenges for the determination of SETA program outcomes competency are posed by the existing measures of CyS and CCA, which are dated and limited (Carlton & Levy, 2015). To address this, Carlton (2016) developed an iPad application to measure CyS index and a corresponding vignette-based assessment (MyCyberSkills™) of employee skills in relation to cybersecurity. Likewise, due to difficulties with prior construct measures, it is important that further research is conducted to develop and validate a measurement tool to properly assess the CCA level of employees. For the purposes of this research, vignette-based assessments of CCA and CyS will be utilized. The vignettes must appear plausible to participants and will be drafted using anonymized situations based on previous research and validated by cybersecurity SMEs (Barter & Renold, 1999).

The Delphi methodology will be employed to validate and improve upon the developed CCA vignette-based assessment, which in conjunction with the CyS assessment validated by Carlton (2016), will be applied as both pre- and post-assessments during SETA program delivery. The Delphi methodology is used in situations where prior information is unavailable and aims to achieve an informed judgment with consensus on a particular topic (Best, 1974). This methodology has been found to effectively utilize a group communication process to refine measures based on the input of an expert panel (Ramim & Lichvar, 2014).

While traditional training has been held in face-to-face format, online methods are increasing in popularity as they have proven to be cost effective, flexible options for organizations. However, more work is needed to determine the most successful delivery method for cybersecurity focused SETA programs. The SETA programs will be delivered via online and face-to-face methods. The

pre- and post-assessments will be used to determine if there are significant differences in the CCA and CyS of the employee based on delivery method.

The main research question (RQ) that this work-in-progress study will address is: Are there any significant differences in employees' CCA and CyS between two SETA program types and two SETA delivery methods? Development and validation of a measurement tool to properly assess the CCA level of employees is imperative to this research study due to the limitations of construct measurement in previous research. To address this need, the first four specific RQs focus on use of the Delphi methodology to determine Subject Matter Experts' (SMEs) approved measurement criteria for CCA, weights of the three CCA categories, as well as the development of two SETA programs with integrated vignette-based assessment.

RQ1: What are the SMEs' approved topics for the two SETA program types using the Delphi methodology?

RQ2: What are the SMEs' approved measurement criteria for CCA using the Delphi methodology?

RQ3: What are the SMEs' approved weights for the three CCA categories (awareness of policy, SETA, & monitoring)?

RQ4: What are the SMEs' approved two SETA programs with integrated vignette-based assessments for CCA and CyS using the Delphi methodology?

The next three research questions address the results of the pilot and main study in relation to CCA and CyS levels of employees. Pre- and post-assessment will allow for a better understanding of significant differences between two SETA program types and two SETA delivery methods. Examination of these research questions will expand the body of knowledge, providing insight into the most effective use of organizational resources as cybersecurity threats become an increasing concern to information assets, information systems, and day-to-day operations.

RQ5: Are there any significant differences between the two SETA program types and the two SETA delivery methods based on the vignette-based pre- and post-assessments of CCA and CyS using a pilot group of participants?

RQ6: Are there any significant differences between the two SETA program types and the two SETA delivery methods based on the vignette-based pre- and post-assessments of CCA and CyS using the main study group of participants?

RQ67a-e: Are there any significant differences between the two SETA program types, and the two SETA delivery methods based on the vignette-based pre- and post-assessments of CCA and CyS using the main study participants, when controlled for participants' (a) age, (b) gender, (c) role in the organization, (d) highest educational level, and (e) years since last attended formal education?

The specific hypotheses for RQ5 and RQ6 (in null form) are:

Ho1a: There will be no statistically significant mean differences in employee's pre- and post-assessment of cybersecurity countermeasures awareness (CCA) and cybersecurity skills (CyS) for the typical SETA program based on the two delivery methods (face-to-face & online).

Ho1b: There will be no statistically significant mean differences in employee's pre- and post-assessment of cybersecurity countermeasures awareness (CCA) and cybersecurity skills (CyS) for the socio-technical SETA program based on the two delivery methods (face-to-face & online).

Ho2: There will be no statistically significant mean differences on employee's cybersecurity countermeasures awareness (CCA) and cybersecurity skills (CyS) between the two SETA program types (typical vs. socio-technical).

Ho3: There will be no statistically significant interaction between the two SETA program types and the two delivery methods.

Experimental Research and Procedures

This work-in-progress research study will utilize a mixed method approach following the work of Carlton and Levy (2015), using both qualitative and quantitative research methods. Research will be conducted in three phases and multivariate analysis of variance (ANOVA), multivariate analysis of covariance (ANCOVA), as well as Spearman Correlation will be used to assess the seven research questions and three hypotheses. Quantitative methods will then be used to deploy two SETA program types via two delivery methods to randomized participants. In Phase 1, qualitative methods will require assistance of SMEs per the Delphi methodology to determine the topics to be covered in the SETA program, to validate and refine the measure of CCA, and to approve the content of the two SETA programs with integrated vignette-based assessments for CCA and CyS (See Figure 1).

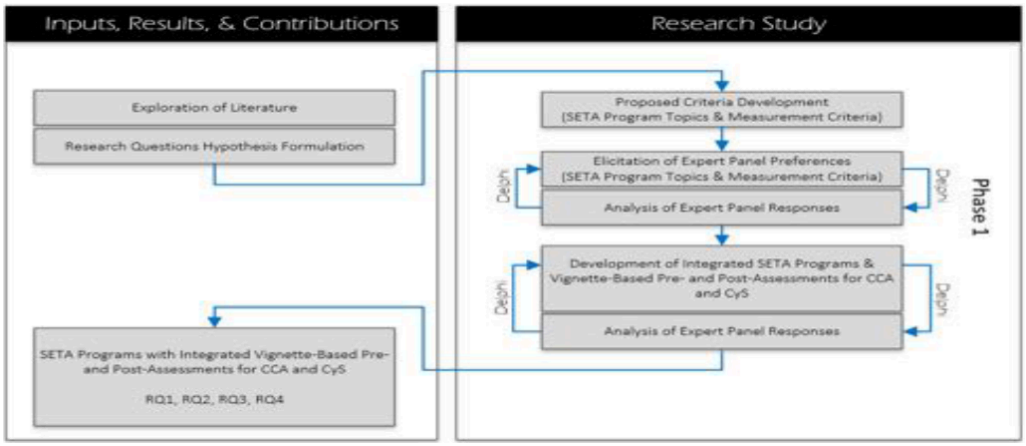


Figure 1. Overview Flowchart for the Proposed Research Phase 1.

Topics for SETA program inclusion will be developed and provided to the SMEs for input and revision per the Delphi methodology. After the topics have been confirmed, two SETA program types will be developed: 1) a traditional SETA program that informs the employee of organizational policies, along with actions that should and should not be taken, as well as 2) a socio-technical SETA program that also includes explanations of why certain actions may cause difficulties and the potential repercussions associated (See Figure 2). Pre- and post-assessments will be used to determine if there are significant differences in the CCA and CyS of the employee based on delivery method. An expert panel will be utilized to ensure validity of the two SETA programs' content per the Delphi methodology.

	Online	***Face-to-Face***
Typical SETA	Online Delivery of Typical SETA Content	Face-to-Face Delivery of Typical SETA Content
Socio-Technical SETA	Online Delivery of Socio-Technical SETA Content	Face-to-Face Delivery of Socio-Technical SETA Content

Figure 2. Proposed Experimental Factorial Design for SETA Program Types and Delivery Methods.

The measurement instrument for CCA will be developed based on the security countermeasures assessments of Hovav and D'Arcy (2012) along with the work of Vance et al. (2012). Although previous work presented these items in survey format, this study will utilize a vignette-based assessment of CCA. Proposed CCA vignettes cover awareness of policy, SETA, as well as monitoring and address key IS security policy topics (SANS Institute, 2014). The Delphi methodology will be used to obtain SMEs feedback on the adapted vignettes as well as the weights for the three CCA categories. The validated vignette-based assessment of CCA will then be integrated into the SETA program.

Phase 2 will consist of a pilot study with randomized participant group allocation into one of two developed SETA program types (typical vs. socio-technical) delivered via two delivery methods (face-to-face & online). Pilot data will be collected from both a pre- and post-assessment integrated with each SETA program and data analysis will be performed using ANOVA to ensure validity and reliability (See Figure 3). The SETA programs and the CCA instrument will be revised per the preliminary data analysis, providing validated measures for the main study.

Figure 3. Overview Flowchart for the Proposed Research Phase 2.

The main study will be Phase 3 of this research, with participants assigned randomly to two developed SETA program types (typical vs. socio-technical) delivered via two delivery methods (face-to-face & online). Main study data will be collected from both a pre- and post-assessment integrated with each SETA program and pre-analysis data screening will be completed (See Figure 4).

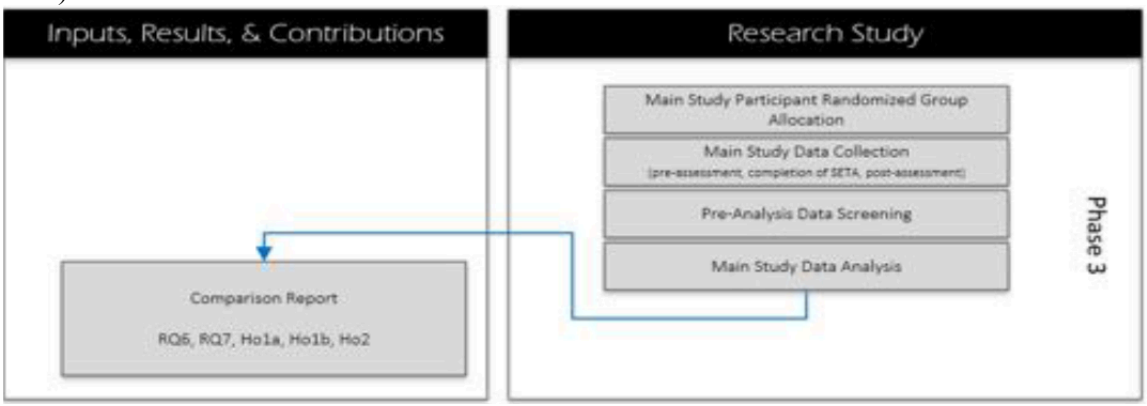


Figure 4. Overview Flowchart for the Proposed Research Phase 3.

Main study data analysis will empirically assess if there are any significant differences on employees' CCA and CyS based on the use of two SETA program types (traditional vs. socio-technical) and two SETA delivery methods (face-to-face & online). Pre- and post-analysis scores for each of the four program type and delivery method combinations will be completed using ANOVA. In addition, ANCOVA will be used to compare the groups, while also controlling for a variable that may exert an influence on the dependent variable, in this case a set of demographics variables such as: participants' (a) age, (b) gender, (c) role in the organization, (d) highest educational level, and (e) years since last attended formal education (Mertler & Vannatta, 2010).

Conclusions and Discussions

Despite considerable investment in organizational security, the majority of approaches and protection methods focus heavily on external attacks as well as technological defenses, and have not minimized the number of security incidents (Pahnila, Siponen, & Mahmood, 2007). However, Abawajy (2012) pointed out that the organization is only as secure as its weakest link. Given the importance of organizational focus on IS security with a human-centric lens, the significance of this study is substantial (Furnell & Clarke, 2012). This work-in-progress study will provide empirically validated data to expand the body of knowledge in relation to cybersecurity. Consequently, this knowledge will also assist practitioners as they determine how to best use training resources to increase organization efficiency and to decrease the chance for losses due to naïve employee cybersecurity behaviors.

References

Abawajy, J. (2012). User preference of cyber security awareness delivery methods. *Behaviour & Information Technology, 2*(4), 1-12.

Albrechtsen, E. (2007). A qualitative study of users' view on information security. *Computers & Security, 26*(4), 276-289.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Barter, C., & Renold, E. (1999). The use of vignettes in qualitative research. *Social Research Updates, 25*(9), 1-6.

Best, R. J. (1974). An experiment in Delphi estimation in marketing decision making. *Journal of Marketing Research*, 448-452.

Boer, H., & Seydel, E. R. (1996). *Protection motivation theory predicting health behavior: research & practice with social cognition models*. Buckingham, PA: Open University Press.

Boyatzis, R. E., & Kolb, D. A. (1991). Assessing individuality in learning: The learning skills profile. *Educational Psychology, 11*(3-4), 279-295.

Carlton, M. (2016). *Development of a cybersecurity skills index: A scenarios-based, hands-on measure of non-IT professionals' cybersecurity skills (Doctoral dissertation)*. Available from Proquest Dissertations Publishing. (UMI No. 10240271)

Carlton, M., & Levy, Y. (2015). Expert assessment of the top platform independent cybersecurity skills of non-IT professionals. *Proceeding of the IEEE SoutheastCon Conference*, Fort Lauderdale, Florida, 1-6.

Carlton, M., Levy, Y., Ramim, M. M., & Terrell, S. R. (2015). Development of the MyCyberSkills™ iPad app: A scenarios-based, hands-on measure of non-IT

professionals' cybersecurity skills. *Proceedings of the Pre-International Conference of Information Systems (ICIS) SIGSEC - Workshop on Information Security and Privacy (WISP) 2015*, Ft. Worth, Texas, 1-12.

Choi, M. S., Levy, Y., & Hovav, A. (2013). The role of user computer self-efficacy, cybersecurity countermeasures awareness, and cybersecurity skills influence on computer misuse. *Proceeding of the Pre-International Conference of Information Systems on Information Security & Privacy*, Milan, Italy, 1-16.

D'Arcy, J., & Hovav, A. (2007). Deterring internal information systems misuse. *Communications of the ACM, 50*(10), 113-117.

D'Arcy, J., Hovav, A., & Galletta, D. (2009). User awareness of security countermeasures and its impact on information systems misuse: A deterrence approach. *Information Systems Research, 20*(1), 79-98.

Dinev, T., Goo, J., Hu, Q., & Nam, K. (2009). User behaviour towards protective information technologies: The role of national cultural differences. *Information Systems Journal, 19*(4), 391-412.

Faux, T. L., & Black-Hughes, C. (2000). A comparison of using the Internet versus lectures to teach social work history. *Journal of Research on Social Work Practice, 10*(4), 454-466.

Furnell, S., & Clarke, N. (2012). Power to the people? The evolving recognition of human aspects of security. *Computers & Security, 31*(1), 983-988.

Gundu, T., & Flowerday, S. (2012). The enemy within: A behavioural intention model and an information security awareness process. *Proceeding of the Information Security for South Africa Conference*, Grahamstown, South Africa, 1-8.

Gupta, S., Bostrom, R. P., & Huber, M. (2010). End-user training methods: What we know, need to know. *Communications of the ACM, 41*(4), 9-39.

Hovav, A., & D'Arcy, J. (2012). Applying an extended model of deterrence across cultures: An investigation of information systems misuse in the US and South Korea. *Information & Management, 49*(2), 99-110.

Hovav, A., & Gray, P. (2014). The ripple effect of an information security breach event: A stakeholder analysis. *Journal of the Association for Information Systems, 34*(50), 893-912.

Kranz, J., & Haeussinger, F. (2014). Why deterrence is not enough: The role of endogenous motivations on employees' information security behavior. *Proceeding of the International Conference on Information Systems*, Auckland, Australia, 1-14.

Levy, Y. (2005). A case study of management skills comparison in online and on-campus MBA programs. *International Journal of Information & Communication Technology Education, 1*(3), 1-20.

Levy, Y. (2006). *Assessing the value of e-learning systems.* Hershey, PA: Information Science Publishers.

McLaren, C. H. (2004). A comparison of student persistence and performance in online and classroom business statistics experiences. *Decision Sciences Journal of Innovative Education, 2*(1), 1-10.

Mertler, C. A., & Vannatta, R. A. (2010). *Advanced and multivariate statistical methods: Practical application and interpretation.* Los Angeles, CA: Pyrczak.

National Initiative for Cybersecurity Careers & Studies. (2014). *Cyber glossary.* Retrieved from https://niccs.us-cert.gov/awareness/cybersecurity-101

Pahnila, S., Siponen, M., & Mahmood, A. (2007). Employees' behavior towards IS security policy compliance. *Proceeding of the 40th Annual Hawaii International Conference on System Sciences*, 156-174.

Parrish, J. L., & Nicolas-Rocca, S. (2012). Toward better decisions with respect to is security: Integrating mindfulness into IS security training. *Proceeding of the Pre-ICIS Workshop on Information Security & Privacy*, Orlando, FL, 1-17.

Parsons, K., McCormac, A., Butavicius, M., Pattinson, M., & Jerram, C. (2014). Determining employee awareness using the human aspects of information security questionnaire (HAIS-Q). *Computers & Security, 42*, 165-176.

Ramim, M. M., & Lichvar, B. T. (2014). Eliciting expert panel perspective on effective collaboration in system development projects. *Journal of Applied Knowledge Management, 2*(1), 122-136.

Ransbotham, S., & Mitra, S. (2009). Choice and chance: A conceptual model of paths to information security compromise. *Information Systems Research, 20*(1), 121-139.

Ross, C. (2006). Training nurses and technologists for trauma surgery. *Journal of Trauma Nursing, 13*(4), 193-195.

SANS Institute. (2014). *Information security policy templates.* Retrieved from https://www.sans.org/security-resources/policies/

Smith, G., Heindel, A., & Torres-Ayala, A. T. (2008). E-learning commodity or community: Disciplinary differences between online courses. *The Internet and Higher Education, 11*(3), 152-159.

Stanton, J. M., Stam, K. R., Mastrangelo, P., & Jolton, J. (2005). Analysis of end user security behaviors. *Computers & Security, 24*(2), 124-133.

Straub, D. W., & Welke, R. J. (1998). Coping with systems risk: Security planning models for management decision making. *Management Information Systems Quarterly, 22*(2), 441-470.

Vance, A., Anderson, B., Kirwan, C., & Eargle, D. (2014). Using measures of risk perception to predict information security behavior: Insights from electroencephalography (EEG). *Journal of the Association for Information Systems, 15*, 679-722.

Vance, A., Siponen, M., & Pahnila, S. (2012). Motivating IS security compliance: Insights from habit and protection motivation theory. *Information & Management, 49*(4), 190-198.

Von Solms, R., & Van Niekerk, J. (2013). From information security to cyber security. *Computers & Security, 38*, 97-102.

Authors' Biographies

Jodi Goode is Assistant Vice President for Information Technology Services and adjunct instructor of Computer Information Systems at Howard Payne University, Brownwood, Texas where she has worked since 2003. She earned a Bachelor's degree in Computer Information Systems and Master's degree in Information Systems from Tarleton State University, Stephenville, Texas. Jodi is currently ABD and pursuing her PhD in Information Systems with a concentration in Information Security at Nova Southeastern University.

Dr. Yair Levy is a Professor of Information Systems and Cybersecurity at the College of Engineering and Computing at Nova Southeastern University, the Director of the Center for e-Learning Security Research, and chair of the Information Security Faculty Group at the college along with serving as the director of the M.S. and Ph.D. programs in Cybersecurity and Information Assurance. He heads the Levy CyLab (http://CyLab.nova.edu/), which conducts innovative research from the human-centric lens of four key research areas Cybersecurity, User-authentication, Privacy, and Skills, as well as their interconnections. Levy authored one book, three book chapters, and numerous peer-reviewed journal as well as conference proceedings publications. His scholarly research have cited over 1,400 times. Dr. Levy has been an active member of the US Secret Service (USSS)'s - Miami Electronic Crimes Task Force (MECTF) and The South Florida Cybercrime Working Group (SFCWG). He was trained by the Federal Bureau of Investigation (FBI) on various topics, and actively serves as a member on of the FBI/InfraGard, and consults federal agencies, state and local government groups on cybersecurity topics. He is also a frequent invited keynote speaker at national and international meetings, as well as regular media interviews as a Subject Matter Expert (SME) on cybersecurity topics. Read more about Dr. Levy via: http://cec.nova.edu/~levyy/

International Institute for Applied Knowledge Management

Refereed Paper Proceedings - KM Conference 2017 – Novo Mesto, Slovenia

A Publication of the International Institute for Applied Knowledge Management

Cognitive data retrieval using a Wizard-of-Oz framework

[Research-in-Progress]

Markus Huber, InnoTec21 GmbH Leipzig and Brandenburg University of Technology Cottbus-Senftenberg, Germany, markus.huber@innotec21.de

Oliver Jokisch, Leipzig University of Telecommunications (HfTL), Germany, jokisch@hft-leipzig.de

Abstract

A suitable human-computer interaction plays an increasing role in various applications, such as smart home, learning or autonomous systems. The Universal Cognitive User Interface (UCUI) shall enable the user to handle different applications by intuitive actions via speech, gestures or a virtual keyboard and allows an adaptation to the users' communication style but also to their strategy in problem solving. In the system development, Wizard-of-Oz experiments are used to collect typical user inputs. In further steps, the user behavior is analyzed and integrated into the system model. To support a user-driven construction, a Wizard-of-Oz Framework (WoOF) was designed, which requires adequate methods for knowledge creation and management including the necessary data processing. The article is focused to an innovative WoOF concept, the underlying data structures and the processing methods. Subsequently, results of the first user experiments and test runs are summarized and discussed, followed by a short conclusion and the outlook to further research steps.

Keywords: Human-computer interaction (HCI), knowledge creation, Wizard-of-Oz framework (WoOF), cognitive user interfaces.

Introduction

Due to the rapid growth in complex technical systems, an appropriate design of human-computer interaction plays an important role, linked to practical aspects of knowledge management (KM). User interfaces and modalities have been already surveyed within the KM context, in particular from the perspective of mobile and multimodal applications (Karolić, 2013; Silber-Varod & Geri, 2014). To enable a user-friendly system handling, in current studies the control paradigms shift to intelligent interfaces with near-to-cognitive abilities, which can be observed in various research areas, for example in smart home applications or autonomous systems. Advanced knowledge creation, modeling and test runs include decided user-oriented paradigms such as Wizard-of-Oz – originated in concepts from the 1980ies; cf. Kelley (1984) – as e.g. demonstrated in Martelaro (2016) or in Hoffman (2016) for human-robot interaction.

The authors are contributing to a national research project, which is intended to create a cognitive user interface, and therefore they also employ the Wizard-of-Oz method. To support

the construction the *Wizard-of-Oz Framework* (WoOF) was build, which allows for the creation of different evolving simulators and also serves as an execution environment for these simulators (Huber, Mayer, Nowack, & Geßler, 2016). Besides brief descriptions of the project context and the Wizard-of-Oz method, the framework is described in more detail revealing its flexible design.

Project Context of the Cognitive User Interface

The joint research project *Universal Cognitive User Interface* (UCUI) from 2015 to 2018 aims at a single interface system, which allows a user to easily manage connected home appliances by corresponding intuitive actions. The user can control the system via speech, gestures or virtual keyboard, whereby a barrier-free application is supported. The system is required to operate widely autonomous, neither using an extensive database (big data) nor an internet connection. These principles also enable an improved data protection and privacy. The system collects user-specific data, which are processed by a *cognitive behavior control* to allow an adaptation to the users' communication style and to improve the strategy in problem solving. The underlying paradigm request the systems' adaptation to the user and not vice versa, assuming the fact, that such systems are mainly used by human beings who are less trained in the use of complex technical devices. In addition, user-specific data are not delivered to other users to avoid possible conclusion from these data (Weber, 2005; Manzeschke, Weber, Rother, & Fangerau, 2013).

To achieve an appropriate system behavior, a variety of possible human-machine interactions needs to be integrated into the UCUI system, since alternative input phrases may have an identical meaning in speech control. Therefore, all input and output modalities are fused on a *semantic processing level*. So-called Wizard-of-Oz experiments are suitable to collect typical user inputs. In further steps, the user behavior is analyzed and integrated into the system model. For the representation of semantic data so-called feature-values-relations are used which get processed by Petri net transducers – see Lorenz and Huber (2013) for an overview and Lorenz, Huber, and Wirsching (2014) as well as Huber, Römer, and Wolff (2017) for a more detailed introduction. In short, feature-values-relations are treelike non-sequential structures where a feature has a set of values which themselves can be features again. Petri net transducers are used to translate input signals into such structures and also for translating them into output signals.

Wizard-of-Oz Experiments

The described cognitive interface is developed user-driven, which poses a challenge, as the overall system is still under construction. The project partners need to evaluate and to optimize some system functions before their implementation. For this purpose, the *Wizard-of-Oz* (WoZ) method is used in the UCUI project. The main component in Wizard-of-Oz experiments is a human being (the wizard), who simulates the final system behavior. During the experiment, the test user interacts with the interface of an assumed technical system. All system reactions to the user are pretended by the wizard. The WoZ method is subject to several demands (Fraser & Gilbert, 1991; Green, Hüttenrauch, & Eklundh, 2004), e.g. to avoid any suspicion regarding the assumed computer-controlled system. Fraser and Gilbert (1991) postulate:

- The corresponding system has to be simulatable (e.g. considering human restrictions).
- The system behavior needs to be specified.
- The simulation has to be plausible.

Wizards have to react accurately and in short time on user inputs. This can be supported with predefined, frequent responses in rapid access, e.g. *please wait, your inquiry is processed* or similar statements, and by a suitable training of wizards to achieve constantly accurate responses.

User scenarios and tasks require a known goal of the actions which can be only achieved by the means of the system without restricting the user in his solution strategy, verbal utterances or gestures (Fraser & Gilbert, 1991; Green et al., 2004). The task construction has to consider the interaction variety of the user and should communicate the options to the user (Green et al., 2004). Within the UCUI project, the user is receiving written instructions beforehand, and the interface system is demonstrated on the basis of a simple vendor machine application by an investigator (not identical with the wizard). Furthermore, the task assignment is based on hypotheses with regard to the expected user and system behavior, cf. (Green et al., 2004).

Finally, a successful, user-driven system construction includes a series of WoZ experiments, whereby the tested system states should increasingly interact in autonomous mode with the user, i.e. less-controlled by the wizard, cf. Kelley (1984). Consequently, the UCUI project involves three consecutive test runs, followed by the overall evaluation of the optimized system.

WoOF: Wizard-of-Oz Framework

The *Wizard-of-Oz Framework* (WoOF) was build to support the user-driven construction. It allows for the creation of different evolving simulators and serves as an execution environment for these simulators (Huber, Mayer, Nowack, & Geßler, 2016).

The following subsections cover the requirements the framework has to comply with as well as some details of its implementation regarding the user interface and the core services.

Conceptual Specifications

The listed requirements include general ones on frameworks supporting WoZ experiments as well as special ones following from the project specifications.

Since the task is to simulate a real system which gives visual and audible feedback to the user, there has to be some mechanism to present visual objects on a monitor and to route audio data to the user. Eventually the system should be controllable via speech and touch-input (among other inputs) which imposes the necessity to interact with the visual objects and to route audio data from the user to the framework. Besides these basic functionalities a good support for the realization of the experiments has to be included. This covers creation of simulators and experiments as well as supporting the wizard during the experiments. For the UCUI project the first experiments consist of a series of scenarios, which are seen as tasks to the participants. The

wizard should be able to switch between scenarios. A single scenario is understood as a unit of a user task, the aim of the task, and possible visual and audible feedback.

The outcome of the experiments should be a collection of user behavior. Therefore all interactions with the system have to be recorded which includes the monitor content the participants see, all touch-events they trigger on it and all spoken input during the experiment. To support the integration of gestures control in following project phases the participants are additionally recorded by camera. To ease the evaluation of the recorded data the audio output of the system is recorded as well.

A form of session management to allow for data per participant has to be included. To respect the privacy of the participants is explicitly not a requirement on the framework. This has to be assured by the experimenters. The motivation behind is that there cannot be any algorithmic solution appropriate to all applications of the framework. So this burden is left to the user. A session as the execution of an experiment with a distinct participant should include some well-defined events. Also it should be possible to add new types of events to the framework.

The preparation of the collected data should be semi-automated. All data of a session should be cut into pieces corresponding to the scenarios. Audio data should be transliterated and phonetically transcribed. All those steps should be advised by a human being. The construction of the feature-values-relations and Petri net transducers is an ongoing research.

Implementation and Technical Details

WoOF is implemented in the Java programming language. Its primary site of operation is the *Cognitive Systems Lab* (CSL) of the Chair of Communications Engineering at the Brandenburg University Cottbus-Senftenberg. For the local audio hardware Java-drivers are provided. All graphical programs within the lab use *LCARSWT* from Wolff (2015) which is also written in Java. Therefore the use of the Java programming language was obvious.

The GUI

To integrate into the CSL software stack WoOF also uses LCARSWT for its graphical user interface. LCARSWT is a widget toolkit for the creation of interfaces which look like the LCARS interfaces known from the TV series *Star Trek: The Next Generation*. The most prominent property of such interfaces besides its touch-capability is the absence of windows and dialogs. An LCARS interface is more like a digital switch panel and is therefore seen as an ensemble consisting of several functional units. In LCARSWT an interface is simply called a *panel*. The units do not need to be connected areas nor to have a distinct shape. However, this peculiarity is only used very rare within WoOF.

Figure 1 shows the panel for the wizard. It is divided into three areas – above left, below left, and right – which can be filled with different so-called *boards*. A board combines several *controls* which are functional units providing for example different logs, a viewer for a webcam (below left in Figure 1), a clock (also below left in Figure 1), a level indicator for audio channels (above

left in Figure 1), a list of feedback items, a list of icons (above right in Figure 1), a list of scenarios (below right in Figure 1), or a preview of the participant panel (center right in Figure 1). Those controls are part of an extensible library which also imposes some sort of look & feel for a consistent appearance.

Figure 1. Wizard panel.

The different boards of the wizard panel can be configured individually. The small areas on the left can both hold boards of the same type but the large area on the right can only hold boards of a different type. Some buttons from the panel can be configured and used by large boards. An experiment can load more than one board for an area. Boards are organized by a stack, so only one board per area is shown at once. Boards can react to distinct events, making themselves visible or invisible. This way an experiment can provide arbitrary functionalities to the wizard. In the middle between the two small areas two so-called *informators* can be loaded which can display status information and make an indicator blink. The last components are so-called *controllers,* which can use some buttons of the wizard panel and also a status line and an indicator. There even exist stacks for informators and controllers within the panels. All components can react to the same set of events during an ongoing session.

The participant panel is depicted in Figure 2. It provides one large area for boards, which are also organized by a stack. The figure contains an example of a task in German language. The translation is given in the subsection *First Experiments*. Since LCARSWT has an integrated network support, it is possible to run the panels on one computer but displaying them on

different computers within the network. It is even possible to display additional panels on connected tablets.

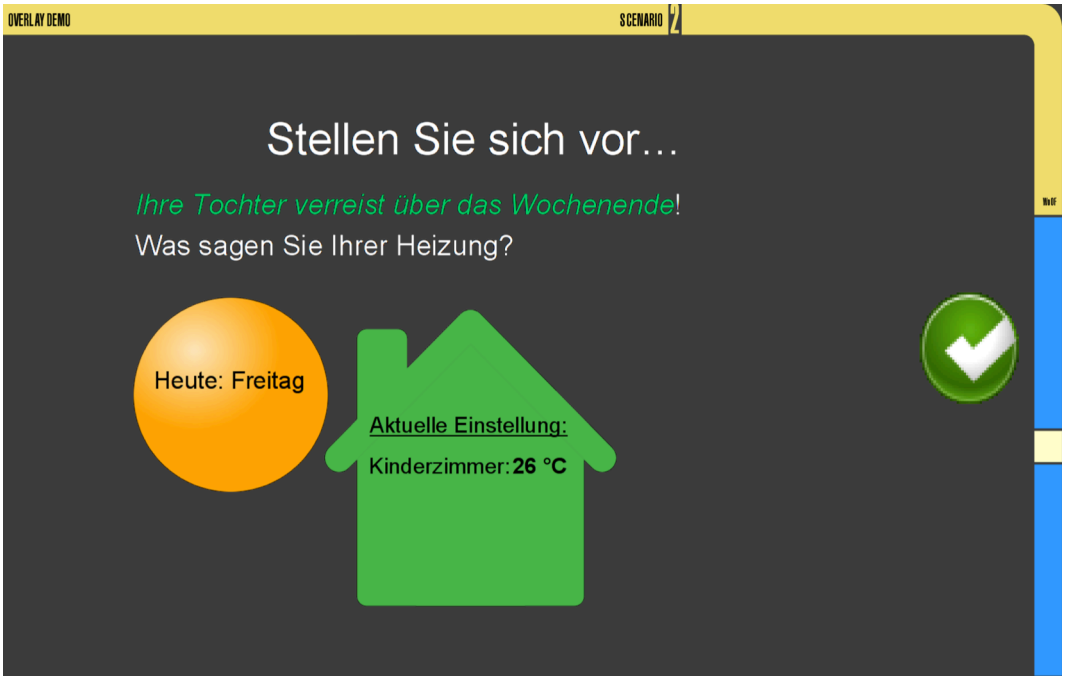


Figure 2. Participant panel.

These concepts are flexible enough to even implement a presentation system on top of WoOF, which includes a presenter view with additional information and a slide view for the audience.

The Framework

The framework itself mainly consists of the central class WoOF, which is implemented as a singleton. The only instance of this class maintains all resources and serves as service access point. During the startup phase it creates the startup log, loads a basic configuration from a text file, collects all available scripting languages and initializes possibly present extensions.

The core services provided by the framework involve handling of time, loading and managing of configurations from text files, handling of session information, abstracting away details of data storing, interfaces to scripting languages, access to protocol logs, maintaining of session state and notifications about state changes, an infrastructure for extensions and modules, and an infrastructure for sending, filtering, and receiving arbitrary objects. The last service makes it possible to use any object as an event and also changing data inside the object between sending and receiving. Any object can register itself as sender, filter, or receiver for a distinct data type. Furthermore some low-level services as threading, loading of objects and resources, and logging of exceptions are available through WoOF as well as all functions regarding the loading and execution of an experiment.

Experiments are organized into projects. This way common configuration options and resources can be provided. All options are loaded from text files with simple key-value syntax. The configuration service can convert the strings into primitive data types and is even capable of handling lists. These lists can be extended by later configuration options on both sides.

Every experiment maintains a list of modules. Modules contain code which runs autonomously reacting on events – e.g. clearing a displayed icon after a period of time, or changing some internal behavior, e.g. the folder layout for storing data or the logging of events. For some modules a setup exists allowing for runtime configuration via the GUI.

So-called *function-providers* are a mechanism to run code that was loaded at runtime while the calls to run the code reside inside the framework. This makes it for example possible to adapt WoOF to different hardware by having different modules for producing audio output.

WoOF includes a session management, which maintains a state and some data. Transitions between states are well-defined, e.g. a non-configured session cannot be started. Every transition is signaled as an event to all modules and panels. The panels in turn signal their informators, controllers, and boards, which again signal their controls. Since panels displayed on different computers all run on the same machine there is no need to dispatch events through the network.

First Experiments

At the time of writing two test runs were already finished. Sixty-two adults (28 men, 34 women; mean age: 31.4 years, SD = 13.38; age range: 18–76 years) took part in the first experiment, sixty participants (34 men, 26 women; mean age: 36.1 years, SD = 15.13; age range: 19–72 years) in the second. In both experiments, they received course credit for their participation. All were German native speakers, had normal or corrected-to-normal eye vision, and did not take medication.

In both experiments the tasks were presented to the participants via a graphic appearing on their panel as can be seen in Figure 2. All tasks refer to controlling a heating installation and the graphics follow the same schematic subdivision. An iconographic house shows the actual settings, a ball shows day and time, the headline says “Imagine that…” followed by an indirect task description, e.g. “your daughter goes away for the weekend!” followed by “What do you tell your heating installation?”

During the first test run, the setting of the WoZ experiments was somehow conventional. The wizard was hiding in a sound booth controlling the simulator. The participants were instructed by a third person. Throughout the experiment the participants were wearing headsets while the wizard was equipped with headphones and a stationary microphone. The participants should talk to the system addressing it as “Computer” and the wizard tried to react like the simulated system using pre-built answers. The aim was to gain some insights how human beings would talk to a system. It turned out that almost none of the participants used intuitive phrases like “Computer, I would love to have it warmer in here.” or “Computer, I’m going on a holiday for three days.”

Rather they used some sort of command language, e.g. “Computer, increase the temperature by 2 degrees.”. It was a really trying challenge for the wizard to direct the participants into uttering intuitive phrases using only the pre-built answers. Therefore the possibility to use the microphone for free interaction was integrated. However, the illusion of a machine was still dominant. The first run produced around 90 hours (45 hours each for wizard and probands) of audio- and video-data.

For the second run the setting was slightly adjusted. The wizard is no longer hidden but stands nearly in front of the participants acting as the operator of the system. The idea was that the participants should talk to a human being telling it what the system should do. The wizard pretends controlling the system and simulates visual feedback in form of some icons appearing on the screen. Already the first few sessions showed the expected effect. For instance, participants were more likely to use more indirect speech acts like “I’m away for a few days.” or “It’s hot in here.” The second run also includes touch-interaction with the system. Therefore a dynamic creation of buttons and sliders was integrated. Areas of distinct colors inside the graphics get overlaid by visual objects on the screen. These objects are touch-sensitive and all interactions get logged. The dynamically created objects can be displayed as rudimentary or full widgets.

Conclusion

This article has provided a proof of concept for a Wizard-of-Oz framework (WoOF), which supports the required knowledge creation and data processing for the advanced cognitive user interface UCUI. Both, user interaction and behavior of the interface system were demonstrated. The underlying processing methods, data structures and algorithms support the project goal of a near-to-cognitive interface and enable a semantic data processing. The first user experiments show that using natural language to interact with a technical system is not self-evident. Actually the reactions of the system have to provoke this behavior of the user. In the context of knowledge management, further research will be dedicated to the evaluation and to the optimization of the semantic processing level.

Acknowledgement

This work has been developed in the project Universal Cognitive User Interface (UCUI) which is partly funded by the German Federal Ministry of Education and Research (BMBF) within the research program IKT2020 (grant #16ES0297).

References

Fraser, N. M., & Gilbert, G. N. (1991). Simulating speech systems. *Computer Speech & Language, 5*, 81 - 99.

Green, A., Hüttenrauch, H., & Eklundh, K. S. (2004). Applying the Wizard-of-Oz framework to cooperative service discovery and configuration. In *Proceedings of the IEEE Intern. Workshop on Robot and Human Interactive Communication* (pp. 575 - 580), IEEE.

Hoffman, G. (2016). OpenWoZ: A runtime-configurable Wizard-of-Oz Framework for human-robot interaction. In *Proceedings of the AAAI Spring Symposium on Enabling Computing Research in Socially Intelligent Human-Robot Interaction* (pp. 121 - 126). Stanford, Californa, USA.

Huber, M., Mayer, W., Nowack, K., & Geßler, P. (2016). WoOF: Ein framework für Wizard of Oz experimente. In O. Jokisch (Ed.), *Proceedings of Elektronische Sprachsignalverarbeitung (ESSV)* (pp. 127 - 134). Dresden, Germany: TUDpress.

Huber, M., Römer, R., & Wolff, M. (2017). Little drop of Mulligatawny soup, Miss Sophie? Automatic Speech Understanding provided by Petri Nets. In J. Trouvain, I. Steiner, & B. Möbius (Eds.), *Proceedings of Elektronische Sprachsignalverarbeitung (ESSV)* (pp. 122 - 129), Dresden, Germany: TUDpress.

Karolić, B. (2013). Increasing the availability of information using modern technologies of the open Web to build user interfaces for mobile devices. *Online Journal of Applied Knowledge Management, 1*, 143 - 155.

Kelley, J. F. (1984). *An iterative design methodology for user-friendly natural language office information applications*. *ACM Transactions on Information Systems, 2*, 26 - 41.

Lorenz, R., & Huber, M. (2013). Realizing the translation of utterances into meanings by petri net transducers. In P. Wagner (Ed.), *Proceedings of Elektronische Sprachsignalverarbeitung (ESSV)* (pp. 103 - 110), Dresden, Germany: TUDpress.

Lorenz, R., Huber, M., & Wirsching, G. (2014). On weighted petri net transducers. In G. Ciardo, & E. Kindler (Eds.), *Proceedings of the Application and Theory of Petri Nets and Concurrency - 35th International Conference, PETRI NETS 2014, Tunis, Tunisia, June 23-27, 2014,* (pp. 233 - 252), Springer.

Manzeschke, A., Weber, K., Rother, E., & Fangerau, H. (2013). Ethische fragen im bereich altersgerechter assistenzsysteme. *Studie für VDI/VDE Innovation + Technik GmbH*.

Martelaro, N. (2016). Wizard-of-Oz interfaces as a step towards autonomous HRI. *In Proceedings of the AAAI Spring Symposium on Enabling Computing Research in Socially Intelligent Human-Robot Interaction* (pp. 147 - 150). Stanford, Californa, USA.

Silber-Varod, V., & Geri, N. (2014). Can automatic speech recognition be satisficing for audio/video search? Keyword-focused analysis of Hebrew automatic and manual transcription. *Online Journal of Applied Knowledge Management, 2*, 104 - 121.

Weber, K. (2005). Das recht auf informationszugang. Berlin: Frank & Timme.

Wolff, M. (2015). LCARS Widget Toolkit 2.1.0-rc.1. Retrieved February 21, 2016, from https://github.com/matthias-wolff/LCARSWT/releases

Authors' Biographies

Markus Huber is a researcher at Innotec21 GmbH Leipzig and Brandenburg University of Technology Cottbus-Senftenberg. He studied educational sciences and mathematics at the Catholic University of Eichstätt-Ingolstadt and worked as a computer scientist at the University of Augsburg. Markus holds a graduate degree each in education and mathematics. His current research interests include Petri nets based transductions, formal languages, abstract algebra, semantic processing and cognitive systems.

Oliver Jokisch is teaching as a professor for signal and system theory at the Leipzig University of Telecommunications (HfTL), Germany. He studied information technology at TU Dresden in Germany as well as at the Loughborough University in United Kingdom. Oliver graduated as a diploma engineer and holds a PhD degree in information technology from TU Dresden. His research is dedicated to different areas in audio and speech communication such as audio coding, speech prosody, speech synthesis and language learning systems. He served as local chair of the 6th Workshop on Speech and Language Technology in Education (SLaTE'15) and of the 27th Electronic Speech Signal Processing Conference (ESSV'16). Beyond he acted as publicity chair of the 16th Annual ISCA Conference (INTERSPEECH'15). Oliver co-founded and mentored the IT-oriented companies voiceINTERconnect GmbH, Linguwerk GmbH, COSEDA Technologies GmbH and ambisone GmbH as well as the education and knowledge management firm IBWM GmbH. Personal site: http://www1.hft-leipzig.de/ice/jokisch_en.php

International Institute for Applied Knowledge Management

Refereed Paper Proceedings - KM Conference 2017 – Novo Mesto, Slovenia

A Publication of the International Institute for Applied Knowledge Management

Autopoiesis of knowledge management systems supported by software agent societies

[Research-in-Progress]

Mariusz Zytniewski, University of Economic in Katowice, Poland,
mariusz.zytniewski@ue.katowice.pl

Abstract

The structure of autonomous information systems requires reference to the aspect of their possible self-organization and adaptation considered in terms of the system autopoiesis. Self-organization, which is a bottom-up process initiated in a particular system by autonomous individuals, can interact with mechanisms of adaptation initiated from above. An example of such a system is an organization knowledge management system supported by agent technologies. Such systems, equipped with autonomous agents, allow to model their self-organization and adaptability in response to changing environmental conditions. The aim of this paper is to analyze the concept of autopoiesis in knowledge management systems supported by agent systems. The paper will propose a concept of an agent system model which is supported by mechanisms regulating agent behaviour as part of a knowledge management system.

Keywords: Software agent, autopoiesis, agent societies, knowledge management

Introduction

The purpose of building distributed systems is to diversify their functionality, which is distributed between subsystems that constitute their components. This requires not only separating the functions of the system into subsystems, but also defining the relations between its components. These relations can rely on interaction relating to the communication between subsystems. Consequently, when defining distributed systems one must not only indicate their structure, but also their dynamism by defining the relations between the components of such systems.

Maturana and Varela (1998) defined an autopoietic system as:

> "a system organized (defined as a unity) as a network of processes of production, transformation and destruction of components that produces the components which: through their interactions and transformations regenerate and realize the network of processes (relations) that produced them and constitute it as a concrete unity in the space in which they exist by specifying the topological domain of its realization as such a network." (p. 136)

Fernandez, Maldonado, and Gershenson (2013) analysed autopoiesies in terms of system complexity. Gershenson (2015) argued that "autopoiesis considers systems as self-producing not in terms of their physical components, but in terms of their organization, which can be measured in terms of information and complexity" (p. 870). A self-organized system is composed of many locally functioning and interacting components (Kasinger, Bauer, & Denzinger, Holvoet, 2010). From the outside, the system may seem complicated, it is controlled by rules that define the interaction of subsystems in a quite simple way. This is due to defined relations between system components, which must be ordered and organized. The interaction of components allows to see the synergistic effect in such systems (Żytniewski, 2010). Distribution of system functions into its subsystems supports its analysis and changes. They relate to each of its components, which make them easier to implement and control.

Self-organization requires communication. Communication of system components is, however, the basis for the construction of social systems (Paetau, 1996). The basis of social systems is defined relations between its components that are established by communication. Communication of system components is what differentiates social systems from technical systems. A technical system, which may be implemented as part of an organization, constitutes a bounded totality. It is implemented and utilized. A social system, on the other hand, assumes that changes within its structure and function result from a change in the relation of its components that have a dynamic nature. Examples include normative systems, which can be applied in social systems. Definition of the standards and principles of the operation of a system may lead to new relations between the elements of the system as a result of the changing standards and principles. An example of such a mechanism is the model proposed further in the paper. It may also cause removal of a fragment of the system, e.g. a software agent from the society. This enforces change in the relations between its elements.

The present article focuses on the issues of self-organization and adaptability associated with the concept of autopoietic systems and automorphosis (Yolles, 2006), considered as an element of multi-agent systems and knowledge management. The aim of this paper is to analyze the concept of autopoiesis in knowledge management systems supported by agent systems. The next section will present the theory of autopoiesis analyzed from the angle of knowledge management systems supported by agent technologies. Then, a proposal for a model of an autopoietic agent system supported by mechanisms regulating agent behaviour will be featured. Following, an example of the operation of a software solution developed will be presented.

Autopoiesis in a Software Agent Society Against the Background of Knowledge Management Processes

Social systems considered in the context of technical solutions can be perceived as autopoietic solutions (Bourgine & Stewart, 2004; Letelier, Marin, & Mpodozis, 2002; Razeto-Barry, 2012). Furthermore, "a system is autopoietic, if it is able to reproduce itself as an autonomous and self-organising unit only by interaction of the internal elements of the system" (Paetau, 1996, p. 5).

The key element in this definition is reproduction (assuming the change in the structure of a system without changing its organization). As indicated by Thannhuber, Tseng, and Bullinger (2001), autopoiesis is a cycle in which a system-executed process defines the structure of the system being created. The system structure is determined by possible self-organization of an autopoietic system and impacts the execution of the process itself.

In principle, self-organization of the above-mentioned components, or more precisely subsystems, has a local nature. Subsystems functioning in a particular environment take actions to achieve a particular task. It is assumed that their structure is built as a bottom-up process where individuals organize themselves. It should be mentioned that self-organization is often mistakenly perceived as the concept of self-adaptation. The difference between the two approaches is related to the process of structuring a group of subsystems. In the case of autoadaptation we are dealing with a top-down approach (Cheng, de Lemos, Giese, Inverardi, & Magee, 2009).
A knowledge management system shares a range of characteristics with an autopoietic system. Within accepted assumptions, Jackson (2007) indicated that a learning organization could be perceived as an autopoietic system. At the same time, he pointed out a range of features that an entity being considered should have to be an autopoietic system (Jackson, 2007):

1. entity must have a boundary;
2. entity must have distinct components of a system;
3. components must be capable of satisfying relations that determine interactions and transformations – the system is made up of the interactions of its parts;
4. components creating boundaries must do so as a result of interactions with other components of the system;
5. components of the boundary must be produced from inside the system;
6. all other components must be produced from inside the system.

The above-defined characteristics of an autopoietic system indicate functional dispersion of such a system resulting from its division into components, necessity of modelling relationships between its elements and significance of defining boundaries between components understood as sets of rules and regulations. Maturana and Valera (1980) pointed out that an autopoietic system should have the following characteristics: autonomy, individuality, organizational closure and self-specification of boundaries. These characteristics can be examined at the level of the whole system or its elements. In the latter case these characteristics can also be noticed in the concept of software agent (Żytniewski & Klement, 2015). A software agent society has defined boundaries in the form of the society within which agents reside {1}, it is divided into separate components in the form of agents {2}, and agents interact with each other based on defined {relationships 3}. From the perspective of the fourth characteristic, it is necessary to define additional mechanisms defining the principles governing the system, e.g. reputation or trust. Then, it is possible to define the operating principles of such a system {4}.

The last two characteristics {5} and {6} indicate a strong link between the concept of software agent societies and knowledge management systems. Assuming that a software agent society is created as an element of a knowledge management system, the mechanisms for defining the principles governing the operation of agent societies should be built based on the knowledge defined in an organization's knowledge management system {5} and refer to the knowledge bases defined in an organization {6}. As a result, the interface of software agent societies and knowledge management systems in the autopoietic approach presented is the aspect of the sharing of organizational knowledge and rules and regulations governing an organization by an agent-based system. The rules and regulations transferred to agent relationships impact the behaviour of agent societies (Figure 1).

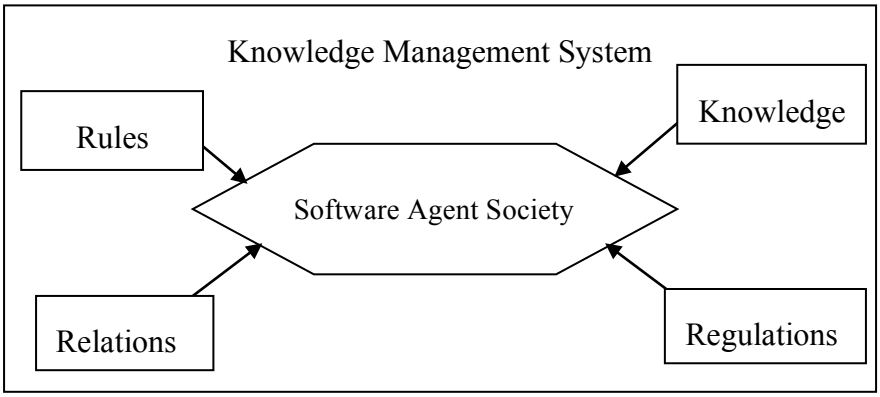


Figure 1. Software agent society as an element of a knowledge management system

Modelling of a software agent society as an element of a knowledge management system requires addressing a range of issues. The first of these is openness and autonomy. Multi-agent systems, and in particular a society of software agents, must adapt their structure or behaviour to the purposes of a given system. These purposes can be time-varying, which means that the state of agents, their knowledge and relations that they build will also change. As a result, achievement of the same objective may vary at two different points of time and bring different effects. The second problem is the lack of observation of the entire system by agents and the lack of ability to control the activities of other agents. As a result, it is not possible to optimally control such a society, but only to search for a local optimum associated with a particular action of the agent. One must be aware that software agents act in a certain time horizon. Even the mechanism of a multi-agent platform, which supervises the activities of agents, does not complete the knowledge about agents, because the time it takes to change the state of an agent, its knowledge, generate a message and receive it can result in subsequent events. The activities of agents should be monitored. The supervising component is the multi-agent platform or agents themselves. The problem that arises when using a monitoring and control mechanism is its failure rate. In the case of failure of the multi-agent platform, the operation of agents will not be possible, because the platform itself provides the environment in which agents are located. The

situation is different in the case of a single agent unit. If it ever fails, the society will continue to function but without the possibility to use the services that it offers.

In the case of an agent society, self-organization can be realized through the mechanisms of a multi-agent platform or by direct interaction of agents. The first type of self-organization is associated with the specificity of agents treated as software entities. Agents wishing to reside in a given environment are dependent on the prevailing rules imposed by the environment in which they are located. Such an environment may be the operating system or multi-agent platform. These restrictions affect the aspect of controlling agents and their behaviour. The multi-agent platform can impose rules prevailing in a given society or provide helpful information, used by agents in interactions with other individuals. In this case, information from the platform can be seen as the impulse, which influences the agent society. As a consequence, the agent society starts the adaptation process.

The second type of self-organization is related to the interactions between agents. Agents observing individual agent units have the ability to collect information about them. E.g. the lack of cooperation on the part of the agent may cause its exclusion from the society. This kind of self-organization is related to the Peer-to-Peer communication or interaction of robots. In order to analyse autopoiesis in the context of software societies that can be used in knowledge management systems, it is necessary to define the typology of software agents and the criteria indicating when a given society represents a specific type.

Referring to the indicated types of societies, it was pointed out (Sayama, 2014) that it is legitimate to refer to five main elements that define the society. These are: the agent's state, its observations, taken actions, the function of observing the agent and the function of changes in the agent. Based on that it is possible to specify four types of agent collectives (Sayama, 2014):

- Homogeneous collectives - this approach assumes that the behaviour of specific agents is determined by their observations with reference to the environment. The agent cannot determine its state. It only responds to stimuli from the environment and is considered in terms of a reactive system. The key aspect of the construction of this arrangement is the function that transforms the stimuli from the environment to the behaviour of the agent.
- Heterogeneous collectives - this type presupposes the existence of a mechanism defining the current state of the agent, changing under the influence of the observation concerning functions defining the behaviour of the agent. As a result, between successive iterations resulting from the change of time, the agent is able to process information about itself and about the changes of the environment in which it is located. This approach does not imply changes in the agent's state in subsequent iterations.
- Heterogeneous collectives with dynamic differentiation/re-differentiation - this is extension of the previous approach. The agent analyzes its environment and takes

action based on the collected knowledge. In addition, it makes continuous changes to its state (something that was not present in the previous approach).

- Heterogeneous collectives with dynamic differentiation/re-differentiation and local information sharing - this approach assumes the possibility of sharing information between agents about their states and observations. In addition, the assumptions from previous approaches are realized.

The typology indicated is hierarchical in character. Each subsequent type of society has the characteristics of the previous one, therefore, it is appropriate to make an attempt to build a model that has the features of the last type of society, as with the use of certain simplifications it will be able to be used in other types. The model that will be proposed in the next section provides specifications of a software agent society in line with the above-defined concept of "Heterogeneous collectives with dynamic differentiation/re-differentiation and local information sharing" (Sayama, 2014, p. 2), which has been extended with the reputation elements proposed in the article (Żytniewski & Klement, 2015). The practical elements shown in this article were defined by using the JADE platform extensions developed by the author (Żytniewski, 2017).

Model Proposition

Let m ∈ N, where m is the number of agents and l ∈ N, where l is the number of agents' actions (N is the set of natural numbers). *D* signifies a set of elements *d*. Upper case indicates the type of this element, e.g. *A* is an agent, *O* – an observation, *S* – a state, *RA* - reputation of an action, *RT* – reputation of a task. Let (1):

$$D_{k_j}^{(A)} = \{d_{1_{k_j}}^{(A)}, d_{2_{k_j}}^{(A)}, ..., d_{x}{}_{k_j}^{(A)}\} \qquad (1)$$

where k ∈ {1, 2, ..., m}, j ∈ {1, 2, ..., l}and $D_{k_j}^{(A)}$ is a set of behaviours of agents' k-th concerning j-th actions which constitute an element of the entire society of agents.

Let m ∈ N be the number of agents, t ∈ N be the number of agents observations, D signifies a set of elements, (O) represents an observation context and (2):

$$D_{k_i}^{(O)} = \{d_{1_{k_i}}^{(O)}, d_{2_{k_i}}^{(O)}, ..., d_{y}{}_{k_i}^{(O)}\} \qquad (2)$$

where k ∈ {1, 2, ..., m}, i ∈ {1, 2, ..., t}and $D_{k_i}^{(O)}$ is a set of expected behaviours for the k−th agent concerning the i−th observation of the agent. As a result, the approach proposed on the "homogeneous collectives" can be described as (3):

$$d_{y}{}_{k_i}^{(O)\prime} = F_t(d_{x}{}_{k_j}^{(A)}) \qquad (3)$$

where F_t is a function of observing agent (based on Sayama, 2014).

Let m ∈ N be the number of agents, h ∈ N be the number of agents states and (4):

$$D_{k_l}^{(S)} = \{d_{1_{k_l}}^{(S)}, d_{2_{k_l}}^{(S)}, ..., d_z{}_{k_l}^{(S)}\} \qquad (4)$$

where k ∈ {1, 2, ..., m}, l ∈ {1, 2, ..., h}and $D_{k_l}^{(S)}$ is a set of expected behaviours for the k−th agent concerning the l−th state of the agent. By defining a set of states of an agent in accordance with the "Heterogeneous collectives" approach, the function enabling the transformation of the state and action of the agent in the new action will be defined as (5):

$$d_y{}_{k_i}^{(O)\prime} = F_t(d_x{}_{k_j}^{(A)}, d_z{}_{k_l}^{(S)}) \qquad (5)$$

where, F_t is a new function of observing agent (based on Sayama, 2014).

In the third approach "heterogeneous collectives with dynamic differentiation/ redifferentiation", one can define another function that makes it possible to change the state of an agent defined as (6):

$$d_z{}_{k_h}^{(S)\prime} = G_t(d_x{}_{k_j}^{(A)}, d_z{}_{k_l}^{(S)}) \qquad (6)$$

where, G_t a function of changes in the agent state (based on Sayama, 2014). However, functions modelled in this way do not take into account the issue of self-organization resulting from the aspect of building a society of agents, because they do not indicate the mechanism that should define the change of agents' state. As a result, the proposed model of self-organization needs to be developed in the context of the mechanisms which can dominate in the society of agents. One of the mechanisms indicated in the literature is the use of trust and reputation of agents (Żytniewski & Klement, 2015). The reputation of agents will be built at the lowest level of the structure, i.e. an agent's reputation in the society as a performer of a specific action. Reputation of the action will be built based on the feedback from other agents that constitute their average. Based on (1) we can specify (7):

$$D_{k_j}^{(RA)} = \{d_{1_{k_j}}^{(RA)}, d_{2_{k_j}}^{(RA)}, ..., d_y{}_{k_j}^{(RA)}\} \qquad (7)$$

where $D_{kj}(RA)$ is a set of reputation for agent k–th concerning j−th action (behaviours) in the entire society of agents. Let (8):

$$| D_{k_j}^{(RA)} | =: y_{k_j} \qquad (8)$$

where ∀k ∈ {1, 2, ..., m}, ∀j ∈ {1, 2, ..., l}: $y_{kj} > 0$. A set of indicators y_{kj} of reputation located in the society of agents must be greater than zero, so that one could determine the reputation of the agent, in their absence the value of reputation is set to 0. To maintain such changes we need to specify function GRA defined as (9):

$$d_{y_{k_j}}^{(RA)\prime} = G_{RA}(d_{x_{k_j}}^{(A)}, d_{z_{k_l}}^{(S)}, d_{y_{k_j}}^{(RA)}) \qquad (9)$$

Function G_{RA} allows the system to specify a new agent reputation for a specific agent's action. As a result, the established indicator concerning reputation of agents' actions (taken in a given society) is expressed by the formula (10):

$$RA_k = \frac{\sum_{y=1}^{N} d_{y_{k_j}}^{(RA)}}{y_{k_j}} \qquad (10)$$

The use of such an indicator requires the use of a mechanism to assess actions taken by agents located in the multi-agent platform, where each action is assessed. Another indicator is the indicator concerning the reputation of a task's performance. Agents chosen to perform a given business process are assigned specific tasks. Each task requires an agent to perform a set of actions. One can therefore specify that a subset of a given set of indicators is a set concerning a specific task, which the agent has previously performed (11):

$$D_{k_j}^{(RT)} = \{d_{1_{k_j}}^{(RA)}, d_{2_{k_j}}^{(RA)}, ..., d_{\xi_{k_j}}^{(RA)}, d_{(\xi+1)_{k_j}}^{(RA)}, ..., d_{(\mu)_{k_j}}^{(RA)}\} \qquad (11)$$

where $d_{1_{k_j}}^{(RA)}, d_{2_{k_j}}^{(RA)}, ..., d_{\xi_{k_j}}^{(RA)}$ is a set of indicators concerning the reputation of actions related to a particular task. Then, similarly to the formula (10), the reputation of a given task will be defined by the formula (9), (10) and (11):

$$RT_k = \frac{\sum_{\xi=1}^{N} d_{\xi_{k_j}}^{(RT)}}{\xi_{k_j}} \qquad (12)$$

The indicators concerning reputation for the RP_k process of the k−th agent and the overall indicator concerning reputation are calculated analogously. In the case of the last indicator concerning general reputation, if it concerns only one society of agents, its value will be equal to the indicator concerning the reputation of tasks or actions depending on the adopted assumptions.

The indicated model allows for evaluation of purposeful behaviour in a given society based on the mechanism of reputation. This model makes it possible to analyze the actions of individuals and gives the system an opportunity to self-organize and adapt. According to the proposed model, the knowledge on the agents' reputation kept by the society enables the selection of agents to undertake actions, tasks and processes in which they may participate. The multi-agent JADE platform does not have this functionality, which is why the next chapter will concentrate on the developed software solution elements implementing the specified model.

System Example

For the realization of agents society, it was necessary to develop an adaptation mechanism and a mechanism to control the behaviour of agents. The latter was developed by extending the mechanism of agents behaviour through its ability to monitor and evaluate its activities, expanding the defined Behaviour class. The agents created on the platform implement specific actions induced by the mechanism of an agent's artificial intelligence. This means that every behaviour of the agent is analyzed and memorized by the multi-agent platform.

The second key element of the society is the mechanism of society adaptation. According to the concept of application of a society of software agents to support the activities of the organization, the structure of an agent-based society will be subject to changes by indicating its responsibilities. In the presented example, the society's task is to evaluate business processes that need to be performed and dynamically adapt its structure in order to achieve them. The JADE platform does not have this functionality, and thus it was necessary to extend its mechanisms by adding new agents to the platform.

The defined tasks of a business process undergo decomposition into actions of the society of agents. Then, they are analyzed by the process agent, which determines whether they can be realized with the current configuration of agents. On the basis of reputation indicators, the agents are invited to the selected society. The decision to join the society comes out of social confidence of the agent in relation to other agents in the society. After forming the society, the agents have to perform tasks.

The solution proposed is consistent with the presented concept of using the concept of software agent societies as a solution designed to support a knowledge management system. An agent system, during internalisation of its parameters, receives knowledge on the rules, relations and regulations connected with the business process being performed. On this basis, in accordance with the cycle of the operation of an autopoietic system, it performs the business process using self-organization mechanisms. As a result, new knowledge is provided to the knowledge management system and can be used in subsequent iterations of its operation.

The example process consists of three tasks. First, the selected agent saves the document (task 1), then the information on saving the document is recorded in the database (task 2), and the user is informed of this action (task 3). For each task, two agent actions need to be performed. Based on the prepared simulator, the tasks were assigned to a set of three agents, each of which can perform a selected task. Their implementation is monitored by the control system. After completing the tasks, the society is dissolved. This example attempts to implement the following business process (Figure 2).

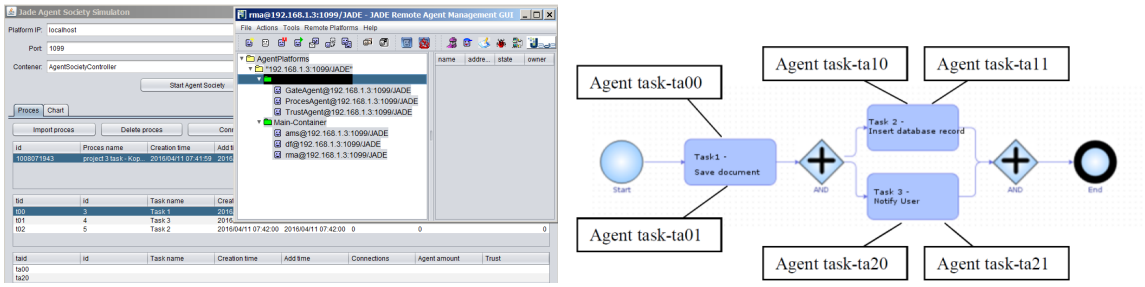


Figure 2. Agent Society Simulator Interface

Table 1 shows a set of initial parameters for a single process, assuming a cold start. In this approach, only with the result of the activities of agents it is possible to indicate the value at a given point of time. The following figures will present the changes in the value of reputation for the individual agents and their actions.

Table 1. Initial parameters of simulation

Level/Type	Reputation of agent n (initial)	Possible agent	Success probability
Action ta00	0,5	ServiceAgent01	0.3
Action ta01	0,5	ServiceAgent01	0.7
Action ta10	0,5	ServiceAgent02	0.7
Action ta11	0,5	ServiceAgent02	0.7
Action ta20	0,5	ServiceAgent03	0.7
Action ta21	0,5	ServiceAgent03	0.7

It can be stated that some agents have high level of success probability, whereas other entities demonstrate low level. It has been assumed that the simulation will cover 10 iterations of the indicated process. The effect of the developed simulator is demonstrated in Figure 3.

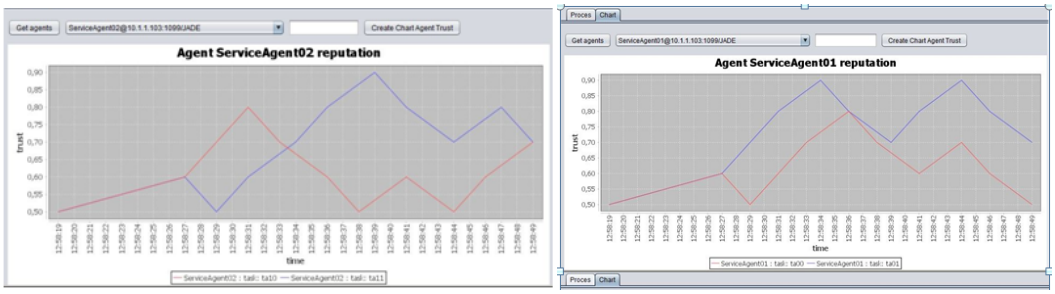


Figure 3. The effects of society activities

As shown in Figure 3 (based on equation (12), (13), & (14)), the correct execution of actions arising from the initiated process resulted in enhancement of reputation of the respective agent. An improper execution of the action ta00 by the ServiceAgent01 resulted in the decline of his reputation. More about the presented JADE extension can be found in Żytniewski (2017).

Conclusions

This paper presented the concept of using software agent societies as an autopoietic element of a knowledge management system. The characteristics of multi-agent systems require reference to the aspect of designing and creating mechanisms that support their own function in complex environments. The indicated issue of self-organization and adaptability of agent systems is an important element when building a society of software agents. From the perspective of autopoietic societies of software agents, both of the presented features are important, because the occurrence of one of the above-mentioned mechanisms can determine the effect of the other. In the case of the adaptive mechanism, the top-down indication of the desired system activities may lead to the necessity of reorganising its structure.

The concept of using a software agent society as an element of knowledge management system has a range of advantages. The first one is the possibility of using individual agents' knowledge about the operation of other units. Such knowledge may come from the information possessed by an agent as a result of its presence in another society. In the adopted model, reputation has a local character, which refers to a specific multi-agent platform. The second advantage is shortening of the time during which an agent will negatively affect a specific society and its operations. The third one is speeding up of the moment when an agent ceases to be a part of a given multi-agent platform. As a result, it releases its resources, contributing to a general improvement of the performance of the operation of the platform and agents. Further research of the author will address the issues of using the model in building autopoietic systems to support selected business processes. The research will involve the evaluation of the impact of the model on the operation of the system.

References

Bourgine, P., & Stewart J. (2004). *Autopoiesis and cognition. Artif Life, 10*(3), 327-345.

Cheng B. H. C., de Lemos, R., Giese, H., Inverardi, P., & Magee, J. (2009). *Software engineering for self-adaptive systems: A research road map.* In Software Engineering for Self-Adaptive Systems, Cheng, B. H. C., de Lemos, R., Inverardi, P., Magee, J. (Eds.), 5525, 1-26.

Jackson, T. W. (2007). Applying autopoiesis to knowledge management in organizations. *Journal of Knowledge Management, 11*(3), 78-91.

Kasinger, H., Bauer B., Denzinger J., & Holvoet T. (2010). Adapting environmentmediated self-organizing emergent systems by exception rules. *Proceedings of the Second International Workshop on Self-Organizing Architectures*, ACM, pp. 35-42.

Gershenson, C. (2015). Requisite variety, autopoiesis, and self-organization. *Kybernetes, 44*, 866–873.

Maturana, H. R., & Varela, F. J. (1998). *The tree of knowledge: The biological roots of human understanding.* Boston and London: Shambhala Publications.

Fernandez, N., Maldonado, C., & Gershenson, C. (2013). Chapter 2: Information measures of complexity, emergence, self-organization, homeostasis, and autopoiesis. In *M. Prokopenko (Ed.), Guided Self-organization: Inception* (pp. 19-51). Berlin, Germany: Springer.

Letelier, J. C., Marin G., & Mpodozis, J. (2002). Computing with autopoietic systems. In *R. Roy R., M. Köppen, S. Ovaska, T. Furuhashi, & F. Hoffmann (Eds.) Soft Computing and Industry* (pp. 67-80). London: Springer.

Paetau, M. (1996). Self-organization of social systems - A new challenge for organization sciences and systems design. *ACM SIGOIS Bulletin 17*(1), 4-6.

Razeto-Barry, P. (2012). Autopoiesis 40 years later. A review and a reformulation. *Orig Life Evol Biosph, 42*(6), 543-567

Sayama, H. (2014). Four classes of morphogenetic collective systems. *Proceedings of the 14th International Conference on the Synthesis and Simulation of Living Systems*, H. Sayama, J. Rieffel, S. Risi, R. Doursat, & H. Lipson (Eds.), MIT Press, pp. 320-327.

Thannhuber, M., Tseng, M., & Bullinger, H. (2001). An autopoietic approach for building knowledge management systems in manufacturing enterprises. *CIRP Annals - Manufacturing Technology, 50*(1), 313-318

Yolles, M. (2006). *Organizations as complex systems: An introduction to knowledge cybernetics.* Greenwich, UK: Information Age Publishing.

Żytniewski, M. (2010). Perfecting synergic effect in hybrid multi-agent systems. In *J. Korczak (Ed.), Business Intelligence and Data Mining*. Wroclaw, Poland: Publishing House of the University of Economics.

Żytniewski, M. (2017). Gossip and ostracism in modelling automorphosis of multi-agent systems. In *Goluchowski, J., Pankowska, M., Linger, H., Barry, C., Lang, M., and Schneider, C. (Eds.), Complexity in Information Systems Development,* pp. 135–150. Lecture Notes in Information Systems and Organisation 22, Springer.

Żytniewski, M., & Klement, M. (2015). Trust in software agent societies. *Online Journal of Applied Knowledge Management, 3*(1), 93-101.

Author's Biography

Mariusz Żytniewski, Ph.D. is employed at the University of Economics in Katowice as a lecturer in the Faculty of Informatics and Communication, Department of Informatics. He is taking part in the research into computer science, systems analysis and computer system design, information systems management, software agents and knowledge-based organizations.

International Institute for Applied Knowledge Management

Refereed Paper Proceedings - KM Conference 2017 – Novo Mesto, Slovenia

A Publication of the International Institute for Applied Knowledge Management

Assessing university quality ranking system in Kurdistan regional government higher education

[Complete Research]

Azad Ali, Indiana University of Pennsylvania, USA, azad.ali@iup.edu

Ava Omar Fatah, University of Sulaimani, Sulaimani, Iraq, ava.fatah@univsul.edu.iq

Abstract

The purpose of this paper was to assess the work that goes into preparing the university quality ranking report for the Kurdistan Regional Government (KRG) ministry of higher education (MHE). We used the word 'system' to describe this work because preparing the ranking report goes through multiple layers of data gathering, verification, calculation, and reporting. Thus, we deemed that the word 'system' is descriptive of the work involved in producing this report. The KRG-MHE started a system that ranks universities under their jurisdiction according to established criteria. The KRG-MHE issued their first report in 2015 and then a second in 2016. Being it new and after two years of reporting, we determined that it is helpful to assess the mechanism by which the ranking of universities implemented and reported. In this paper, we assess the work of the KRG-MHE that goes into producing the university quality ranking report. It compares the factors used by KRG-MHE in determining quality with similar factors used in established systems of ranking (Like systems in the United States & United Kingdom). The paper at the end submits recommendations to the KRG-MHE to make their system more consistent with the established systems we reviewed in this paper.

Keywords: University ranking, academic ranking, university ranking Kurdistan, academic quality ranking

Introduction

> "Nonetheless, just as democracy, according to Winston Churchill, is the worst form of government except for all the others, so quality rankings are the worst device for comparing the quality of … colleges and universities, except for all the others." (Webster, 1986, p. 6)

Democratic governing systems are not perfect, and all democratic governments are subject to wide range of criticism and skepticism. Yet, democracy has embedded in it a process that allow for critiques, reviews, and assessments that often turn into revisions, enhancements and improvements, which presumably intended to serve the will of the majority. For the reason of this process (and for other reasons), democracy persists and many nations call to adopt them in their governments.

In the same way, university quality ranking systems are not perfect. They are subject to wide range of criticism and skepticism reported in different studies (Avery, Glickman, Hoxby, & Metrick, 2004; Batesdo & Bowman, 2010; Coates, 2007; Hazelkorn, 2015). Yet, many countries and regions around the world call to implement them within their higher education. Similar to democracy, university ranking systems allow for reviews, critics, assessments, and then of revisions so that the system would more accurately reflect the quality of universities in their ranking reports.

This study assessed the work of the Kurdistan ranking report to determine if the report measures quality of their universities similar to the way that established ranking systems measures quality. This paper compares the quality factors used by the Kurdistan Regional Government (KRG) ministry of higher education (MHE) with similar factors used in established university quality ranking systems (like the systems used in the United States & United Kingdom). Then, this study concludes with recommendations to the KRG-MHE on how to make their ranking system more consistent to established system of ranking.

Literature Review - University Rankings

This section provides a literature review about university quality ranking systems and reports. It begins by defining the term "University Ranking" and the institutions that typically issue ranking reports. It then explains about the mechanism of developing the ranking reports and the factors that typically considered in measuring quality. It further lists some guidelines that university ranking systems typically adhere to. We determined that this introductory information contributes to giving background information about the topic so that gives to better understanding of the different ranking systems (Including KRG-MHE ranking) as we present them later on.

University Rankings – Definition

Different terms have been introduced to describe university quality ranking and various definitions have been introduced for them. Usher and Savino (2007) provided the following definition:

> "University rankings are lists of certain groupings of institutions (usually, but not always, within a single national jurisdiction), comparatively ranked according to a common set of indicators in descending order. University rankings are usually presented in the format of a ''league table'', much as sports teams in a single league are listed from best to worst according to the number of wins and losses they have achieved". (p. 6)

Usher and Savino's (2007) definition make assertions about a few interesting points. First, the ranking lists the universities according to their performance in descending order - from best to worst. The best universities listed at the top and can easily be viewed by potential students who look for such reports. While at the bottom and in places that cannot be easily be viewed,

universities with lower scores listed. Margison (2007) called such rankings "powerful" because it shows the top universities in easy-to-be-viewed places where they can be considered by potential students. Additionally, the definition noted here that common set of indicators used to measure the quality of all universities. These indicators eventually make up the score that each university given for their performance.

Hazelkorn (2008) used the term 'University League Tables Ranking System' and gave it the acronym LTRS. Hazelkorn (2008) explained that LTRS are "contemporary form, type are published by, inter alia, government and accreditation agencies, higher education, research and commercial organizations, and popular media, as a consumer information tool" (p. 193). The interesting description by Hazelkorn's (2008) is that it considered LTRS as a "consumer information tool". This in turn equated academic ranking with many other rankings that are published periodically in consumer magazines, automobile comparisons, the rating of movies, and other similar ratings publications.

Salmi and Saroyan (2007) called it "institutional ranking" and used the term "report cards" to describe academic ranking reports. Salmi and Saroyan described the "reports cards" as:

> "Constructed by using objective and/or subjective data obtained from institutions or from the public domain, resulting in a "quality measure" assigned to the unit of comparison relative to its competitors. For the most part, the unit consists of tertiary education institutions, primarily universities. However, rankings are also done of colleges or specific subject areas or programs across all institutions." (p. 33)

Ranking Institutions

Different institutions produce various university ranking reports for countries and regions around the world. Hazelkorn (2012) listed the following five groups of institutions that work on and produce ranking reports:

1- Government organizations

2- Accrediting agencies

3- Research institutions

4- Commercial organizations

5- Popular media (like newspapers and magazines)

Despite this variety of ranking organizations, in the most cases for developing countries that we reviewed, only government organizations do the task of ranking university quality. Salmi and Saryon (2007) explained that in developing countries; only government organizations oversee the work that goes into producing university ranking reports (as the case of KRG-MHE that we are studying here). This may be because higher education in most developing countries administered in large by the governments. Despite this extended use of government agencies as ranking organizations, we do not have data to support which group of organizations provide

more accurate ranking reports. Nevertheless, the large use of ranking reports of newspapers or magazines, such as the US News and World report in the US (Margison, 2007) lead us to believe that these reports are more reliable.

National, International, and Subject Ranking

Van Dyke (2005) noted that the first ranking report of universities in the US published in 1983 when the US News and World Reports issued their first annual report ranking universities in the USA. The report contained several pages but the magazine listed only the first 25 universities on their printed magazine. The report, which contains the rankings of more than 1000 universities, saved in digital format and distributed to participants.

Since 1983 ranking organizations have proliferated and they did not continue to only rank universities, instead there are organizations that rank programs as well (Merisotis & Sadlak, 2005). Furthermore, there are organizations that rank universities globally, that is ranking university quality around the world. Salmi and Saryon (2007) reported that there are at least 30 reports produced annually which rank universities in the US. They also added that there are "countless" numbers of program ranking reports – like the ranking reports for the MBA (Master of Business Administration) programs. An important point to note is that many reports are being generated produce different ranking (Hazelkorn, 2008). Having that many ranking reports and given that they often produce different rankings, this may question the validity and reliability of all the ranking reports (Bhattacharjee, 2011; Osterloh & Frey, 2015). Thus, assessment or the ranking report is helpful to compare these results.

Global ranking (that is ranking universities worldwide in a single report) is on the rise as well. Hazelkorn (2012), for example, listed some of the institutions that produce global ranking reports:

- Academic Ranking of World Universities (ARWU) (Shanghai Jiao Tong University)
- Webometrics (Spanish National Research Council, 2003)
- World University Ranking (Times Higher Education/QS)
- Performance Ranking of Scientific Papers for Research Universities (HEEACT)
- Leiden Ranking (Centre for Science & Technology Studies, U Leiden)
- SCImago Institutional Rankings
- Top University Rankings (QS)
- World University Ranking (Times Higher Education/Thomson Reuters [THE-TR])
- U-Multirank (European Commission)

Different factors explain the increase emphasis on global ranking. Stolx, Hendel, and Horn (2010) suggested that the notion of "World University" as one factor contributes to this emphasis. The world university implies that students can move to different universities around the world with relative ease as compared to previous years. Also, the competition for international students among universities in advanced countries led to increase importance on reports that lists universities from different place around the world (Hazelkorn, 2015, Shin & Toutkousian, 2011)

In this paper, we are going to focus on the national ranking – that is the university ranking conducted for a country or a region so that to make our assessment of the ranking report of the KRG-MHE more comparable. Although we mentioned subject and global rankings, but these were noted for illustration and to distinguish them from national and regional ranking.

Quality Ranking Report Mechanism

At the conclusion of most ranking reports, universities are listed in according to one score given to each university. This score is so significant that some called it "single, all-encompassing quality score" (Usher & Medow, 2009, P. 3). Clark (2002) called this score "one easy-to-digest number" (p. 446). Whether easy or all encompassing, this score is typically derived through a process that involves data collection, verification, calculation and then reporting.

The score for each university calculated based on selected quality criteria and then giving a weight to each quality criteria. The quality criteria called "Quality indicators" or "Determinants of Quality" (Avery, 2003) In other words; the quality indicators are individual factors that supposedly measure the quality of the work of the university. So the quality indicators need to be quantified so that the single quality score can be calculated. The steps that typically followed in calculating the final quality score are:

- Indicators of quality are determined that are applied to all universities
- Quality indicators are quantified and weights are given to all of them
- Scores are given to the universities for each indicator based on collected data
- The scores of all indicators are summed together to give the final score for each university

The scoring for each indicator typically works in the following: The highest performing university in the indicator is given 100% and then the lower performing universities are scored in proportion to the top level (Usher & Medow, 2009). All scores are converted to percentages and combined together. The combined score (typically out of 100 points) is the score that indicates the quality of the work of a given university (Van Dyke, 2005).

The Grouping of Ranking Indicators

The quality indicators are often too numerous to be able to list them in one easy-to-understand report. Usher and Medow (2009) for example surveyed the quality indicators in different systems and found there are more than 600 different indicators that measure quality in the surveyed systems. This large number of indicators make them difficult to compare and find similarities and difference among the systems producing ranking reports.

Grouping or categorizing the indicators is one solution suggested to solve this problem (Dill & Soo, 2005). By categorizing, some researchers combined quality indicators into groups according to similarities that they (the indicators) have. These individual indicators can then be combined together and a score is calculated for the group for easier comparison (Usher & Medow, 2009).

Various studies suggested different kinds of grouping of quality indicators to make the comparisons easier. Usher and Savino (2007), for example, combined the quality indicators into the following eight categories:

- Beginning characteristics
- Learning inputs – staff
- Learning inputs – resources
- Learning environment
- Learning outputs
- Final outcomes
- Research
- Reputation

Dill and Soo (2005) on the other hands, suggested four general categories that group the quality indicators under each one of them as listed below:

- Input
    - Faculty
    - Students
    - Financial Resources and facilities
- Process
- Output
    - Satisfaction
    - Graduation
    - Value added
    - Learning progress
    - Employment
- Reputation

The point that can derived from both categorizations listed above is that comparison between different ranking reports are possible but categorization helps in focusing the points to compare. Thus, we will follow this strategy in our study when we compare the work of the KRG-MHE.

Assessment of Ranking Systems

As the case in many institutions, there has to be some kind of principles to provide general guidelines to standardize the work of the institution. The same thing applies to the work involved in the systems of university quality ranking. There are different criteria established to standardize the ranking systems of reporting. The most notable published criteria is "The Berlin Principles on Ranking of Higher Education Institutions" (IREG, 2006).

According to Stolz, Hendel, and Horn (2010), these principles are set of standards from which measurements of academic quality are derived. These set of standards contain 16 principles that cover different doctrines ranging from formation, stating objectives to establishing criteria for

selecting quality indicators (Cheng & Liu, 2008). Two of these principles are of concern in this study; we list principle # 7 and #9 from this document below:

> #7: Choose indicators according to their relevance and validity. The choice of data should be grounded in recognition of the ability of each measure to represent quality and academic and institutional strengths, and not availability of data. Be clear about why measures were included and what they are meant to represent
>
> #9: Make the weights assigned to different indicators (if used) prominent and limit changes to them. Changes in weights make it difficult for consumers to discern whether an institution's or program's status changed in the rankings due to an inherent difference or due to a methodological change

Methodology

The goal of this paper was to assess the university ranking system of KRG-MHE. To achieve our goal, we compared the quality indicators used by other established systems with those used by KRG-MHE. We have chosen four ranking systems to select their quality indicators and compared them with those in the KRG-MHE system of ranking.

As noted earlier, the quality indicators are too many to be able to make the comparison clear. Therefore, to make this comparison more manageable, we go through the following steps:

- We established three general categories of quality indicators: Input, process and output.
- We look at the quality indicators for each ranking we selected for comparison and then allocate each indicator to one of our three general categories.
- We sum the scores assigned for the indicators within the categories and the sum of the scores represents the score for the given category.
- After this, we compare the scores for the three groups for all systems under consideration in one table to see the differences between the scores in all of them.

Having selected the three categories of input, process and output, we need to establish criteria on how to include indicators under each of the three categories.

For the *input* category, we considered any factor that contributes to the education prior to the arrival of the students to the university for study. This includes student scores prior to enrollment, admission rate, number of programs and environment at the university.

For the *process* category, we considered factors of the students when they wre enrolled at the university. Such factors include graduation rate, faculty/student ratio, expenditure, faculty/staff salary, and financing received whether for students or faculty/staff.

For the *output* category, we considered factors that influence the university/student after graduation and completion their degree. It also refers to sub categories that deal with what the university staff produced. Some of these factors include graduation rate, employment rate, salary of graduates, research record, and reputation.

Comparing the Reports

In this section, we are comparing the categories of the quality indicators of the KRG-MHE with similar categories in established systems. Ali, Fatah, and Kohun (2016) noted that KRG-MHE want to model their ranking system with those in established countries. Thus, we selected four established systems for the purpose of this comparison: two of them in the US and two from the United Kingdom. The following are the selected four established ranking systems:

- US News and World Reports of ranking for the US universities
- Wall Street Journal/Times ranking of US colleges and universities
- The Guardian report of ranking universities in the UK
- The Times university ranking report in the UK.

The remainder of this section explains about each of the four reports and then lists in a table the categories and the weight used in the production of their ranking report.

US News and World Reports

US News and World Reports is the most prominent organization that produce annual reports to rank universities around the US. They started issuing their first report in 1983 (Clarke, 2002). Usher and Medow (2009) compared the US News and World ranking system with others and divided their ranking indicators into five categories: beginning characteristics, learning, final outcomes, research, and reputation. We followed the same categories first and then allocated the scores into our three categories (input, process, & output) and combine them in Table 1 below.

The first column in Table 1 contains the original categories as listed Usher and Medow (2009). The second column lists the score for the category as listed by User and Medow (2009). The third column lists our assigned category - that is what we consider each item under the categories of input, process or output.

Table 1. US News and World Report's Ranking

Original category	Sub-category	Our assigned category
Beginning characteristics 15%	Beginning characteristics 15%	Input (15%)
Learning 60%	Learning input – staff 20%	Process 20%
	Learning inputs – resources 15%	Input 15%
	Learning input – environment 0%	Input 0%
	Learning outputs 25%	Output 25%
Final Outcomes 0%	Final Outcomes 0%	Output 0%
Research 0%	Research 0%	Output 0%
Reputation 25%	Reputation 25%	Output 25%

After tallying the totals for all three categories, we found that total for input is 30%, for process is 20%, and for output is 50%

The Wall Street Journal/Times Higher Education College Ranking

The Wall Street Journal (WSJ) along with the Times Magazine have issued their first ranking report in September 2016. They issued their report that lists the ranking of 500 universities in print format and on their web site (http://www.wsj.com/graphics/college-rankings-2016/). Their total ranking report includes more than 1000 universities. They elected to make the total report available only to members who subscribe to their journal or upon requesting the report. Although the WSJ is reporting for the first time in this category but the Times Magazine has been reporting on ranking colleges and universities in the UK this for a number of years. The reputation of the WSJ along with the experience of the Times Magazine in reporting academic ranking compelled us to select this report for comparison.

The report was elaborate about the methodology they selected for their reporting, the quality indicators and the categories they grouped them. They listed four top categories (Resources, engagement, Outcomes, & Environment) with each of them divided into different subcategories. Table 2 below lists the categories, subcategories along with the weight assigned to each sub category. The third column lists our assigned category for the items listed in the second column.

Table 2. WSJ/Higher Education Ranking

Original category	Sub-category	Our assigned category
Resources 30%	Finance per student 11%	Process 11%
	Faculty per student 11%	Process 11%
	Research per faculty 8%	Output 8%
Engagement 20%	Student engagement 7%	Process 7%
	Student recommendation 6%	Output 6%
	Interaction with teachers and students 4%	Process 4%
	Number of accredited programs 3%	Input 3%
Outcomes 40%	Graduation rate 11%	Output 11%
	Value added to graduate salary 12%	Output 12%
	Value added to the loan repayment rate 7%	Output 7%
	Academic reputation 10%	Output 10%
Environment 10%	Proportion of international students 2%	Input 2%
	Student diversity 3%	Input 3%
	Student inclusion 2%	Input 2%
	Staff diversity 3%	Input 3%

WSJ Total includes: 13% for input, 33% for process, and 54% for output.

The Guardian (UK)

The Guardian Magazine has been issuing ranking reports for universities in the UK since 1999 (Dill & Soo, 2005). Their selection of categories was simple: they had similar categories to what we used in our paper (input, process, & output) but added one more category, which is *reputation*. Table 3 below shows the original categories and subcategories in the first two columns as selected by the Guardian, and then we list our selected category in the third column.

Table 3. The Guardian (UK)

Original category	Sub-category	Our assigned category
Input 14%	Student/staff ratio 6%	Process 6%
	Per student spending 8%	Process 8%
Process 65%	Teaching assessment 65%	Output 65%
Output 15%	Student staff ratio 6%	Input 6%
	Per student spending 9%	Process 9%
Reputation 6%	Demand among high school students 6%	Input 6%

After adding the different percentages for our three assigned categories, the resulted totals are: input 12%, process 23%, and output 65%.

Times (UK)

The "Times Good University Guide" report of ranking universities in the UK has been producing at an annual rate since 1992 (Ibid, 2005). Although their categories are similar to those listed by the Guardian (Input, process, output and reputation) but they included a different set of sub categories. Thus, our assigned percentages to our selected categories is different from those listed in the previous table. Table 4 below lists the categories/subcategories as reported by the Times and then in the third column we list our assigned category based on our classification.

Table 4 - The Guardian (UK)

Original category	Sub-category	Our assigned category
Input 50%	Student/staff ratio 9%	Process 9%
	Research assessment 14%	Output 14%
	Average A and AS Levels 9%	Output 9%
	Library and computing spending/student 9%	Input 9%
	Facilities spending 9%	Input 9%
Process 23%	Teaching assessment 23%	Output 23%
Output 27%	Graduation rate 9%	Output 9%
	First and second upper degrees 9%	Output 9%
	Job prospects 9%	Process 9%

The totals revealed from each assigned categories included input 18%, process 18%, and output 64%. A note to be mentioned here is that the category of reputation was mentioned in the Times Magazine with zero percentage, thus we included it in the table above.

KRG-MHE

The KRG-MHE started their reporting from in 2015. They published their first ranking report in 2015 and then in second year in 2016. All data about the purpose of the ranking report, the reasons for producing the report, the methodology followed, and the final ranking reports listed on their web site (http://www.nur-krg.net/). We reviewed the web site and list below the categories, the percentage they assigned to each category and then the category that we list our assigned classification based on our list of categories we established for this paper.

1. Academic Staff (13%) (input)
2. Scientific Research (40%) (output)
3. International Activities (6%) (process)
4. Student Satisfaction (3%) (Output)
5. Quality (28%) (Process)
6. Cultural and Community Activities (4%) (process)
7. Library (6%) (Input)

After adding the scores for each category, we list below the total for of our assignment categories: input 19%, process 30%, and output 51%

Summary and Discussion

Table 5 represents a comparison among the five agencies, we studied in this paper in terms of their proportion of score in input, process and output.

Table 5. Comparison of Ranking Reports Scoring

Organization name	Input %	Process %	Output %
US News and World report	30%	20%	50%
WSJ & Times	13%	33%	54%
The Guardian	12%	23%	65%
Times (UK)	18%	18%	64%
KRG-MHE	19%	30%	51%

The result of the comparison in Table 5 above did not reveal major shift of the KRG-MHE scores from the established ranking systems we selected in this paper. The percentages allocated by KRG-MHE for each category did not differ significantly with the other four reports we compared in this study. Although KRG-MHE emphasized more on process and less on output than the others did, but this does not signal major shift. However, a closer look at the quality indicators listed on the KRG-MHE website, we noted the following three major differences between KRG-MHE reporting versus the others:

- More emphasis on research
- Higher number of sub categories (quality indicators)
- Often ambiguous naming of categories

The KRG-MHE allocated a score of 40% for research. No other report gave such a high score for research. The US News and World report gave zero percentage for research and the Wall Street Journal gave 8% for research for faculty. The Times (UK) puts 14% on research assessment and we did not find anything allocated by the Guardian (UK) on research productivity. We found a high research score in the international ranking reports such the Shanghai Jiao Ton University when they assigned 90% of the final score to research (Hazelkorn, 2012). Similarly, ranking Iberoamericano allocated 100% of the score for research so to rank universities only on research (Usher & Medow, 2009). However, we did not find ranking that emphasizes such a high percentage (40%) on research in the other ranking systems we reviewed.

The number of Quality indicators listed by the KRG-MHE is high. We counted from the website and found 47 sub-categories listed and most them have a weight of 1% or 2%. This in contrast to smaller number of indicators in the other systems and they all have higher percentages allocated for them. For example, the US News and World Report lists 15 sub-categories in their ranking while Times Magazine report has nine sub-categories.

We also noticed a number of ambiguous naming of categories and then contradicting sub categories in the KRG-MHE ranking system. For example, the category "Quality" does not explain what is meant by this term. Then when we examined the sub categories from the web site, we found the following list:

- University campus space
- Classroom space
- Accommodation space
- E-Management and e-QA

Recommendations

Based on the findings of our study, we offer recommendations to the KRG-MHE regarding their university ranking system. We feel that implementing these recommendations will make their ranking system more consistent with established systems of ranking that we examined in this study. As noted earlier, the goal of KRG-MHE is to establish a ranking system that model other established ranking system and our recommendations contribute to achieving this goal.

The following are our recommendations to the KRG-MHE regarding the design of the quality indicators for their university quality ranking system:

- Reduce the emphasis on research from the current percentage of 40% to somewhere below 20%
- Reduce the number of quality indicators from the 47 currently have to around 20. This can be accomplished by combining indicators to make the ranking process easily understandable
- Use more meaningful category/sub category naming and align the sub categories properly within each appropriate category

References

Ali, A., Fatah, A. O., & Kohun, F. (2016). Preparing for academic ranking reports in the Kurdistan regional government higher education. *Proceedings of Informing Science & IT Education Conference (InSITE) 2016,* 1-19. Retrieved from http://www.informingscience.org/Publications/3426

Avery, C., Glickman, M., Hoxby, C., & Metrick, A. (2004). *A revealed preference ranking of US colleges and universities (No. w10803).* National Bureau of Economic Research.

Bhattacharjee, Y. (2011). Saudi universities offer cash in exchange for academic prestige. *Science, 334*(6061), 1344-1345.

Clarke, M. (2002). Quantifying quality: What can the US News and World Report rankings tell us about the quality of higher education? *Education Policy Analysis Archives, 10*, 16.

Coates, H. (2007). Universities on the catwalk: Models for performance ranking in Australia. *Higher Education Management & Policy, 19*(2), 69-85.

Dill, D. D., & Soo, M. (2005). Academic quality, league tables, and public policy: A cross-national analysis of university ranking systems. *Higher education, 49*(4), 495-533.

Hazelkorn, E. (2008). Learning to live with league tables and ranking: The experience of institutional leaders. *Higher Education Policy, 21*(2), 193-215.

Hazelkorn, E. (2012). *Striving for world class excellence: Rankings and emerging societies.*

Hazelkorn, E. (2015). *Rankings and the reshaping of higher education: The battle for world-class excellence*. Springer.

IREG. (2006). *Berlin principles on ranking of higher education institutions.* Berlin, Germany: International Ranking Expert Group.

Liu, N. C., & Cheng, Y. (2005). The academic ranking of world universities. *Higher Education In Europe, 30*(2), 127-136. doi:10.1080/03797720500260116

Marginson, S., & Van Der Wende, M. (2007). *Globalisation and higher education.* OECD Education Working Papers, No. 8. OECD Publishing (NJ1).

Merisotis, J., & Sadlak, J. (2005). Higher education rankings: Evolution, acceptance, and dialogue. *Higher Education in Europe*, *30*(2), 97-101.

Osterloh, M., & Frey, B. (2015). Current dynamics of scholarly publishing. *Evaluation Review, 39*(1), 102-129.

Salmi, J., & Saroyan, A. (2007). League tables as policy instruments: Uses and misuses. *Higher Education Management & Policy, 19*(2), 31-68.

Shin, J. C., & Toutkoushian, R. K. (2011). The past, present, and future of university rankings. In *University Rankings* (pp. 1-16). Springer Netherlands.

Stolz, I., Hendel, D. D., & Horn, A. S. (2010). Ranking of rankings: Benchmarking twenty-five higher education ranking systems in Europe. *Higher education*, *60*(5), 507-528.

Usher, A., & Savino, M. (2007). A global survey of university ranking and league tables. *Higher Education in Europe*, *32*(1), 5-15. doi:10.1080/03797720701618831

Usher, A., & Medow, J. (2009). A global survey of university rankings. In *B. M. Kehm & B. Stensaker (Eds.), University Rankings, Diversity, and the New Landscape of Higher Education*, pp. 3-18.

Van Dyke, N. (2005). Twenty years of university report cards. *Higher Education In Europe, 30*(2), 103-125. doi:10.1080/03797720500260173

Webster, D. S. (1986). Academic quality rankings of American colleges and universities., Springfield, IL: Charles C. Thomas .

Authors' Biographies

Azad Ali, D.Sc., Professor of Information Technology at Eberly College of Business – Indiana University of Pennsylvania – has 30 years of combined experience in areas of financial and information systems. He holds a bachelor degree in Business Administration from the University of Baghdad, an MBA from Indiana University of Pennsylvania, and an MPA from the University of Pittsburgh, and a Doctorate of Science in Communications and In-formation Systems from Robert Morris University. Dr. Ali's research interests include ser-vice learning projects, web design tools, dealing with isolation in doctoral programs, and curriculum.

Ava Omar Fatah, MBA, Lecturer of Business Management at School of Administration and Economics – University of Sulaimani – has 7 years of experience as Lecturer and 2 years as Quality Assurance Administrator. She holds a bachelor degree in Business Administration from the University of Sulaimani, an MBA from University of Wales. Meanwhile is a Ph.D. student in Human Resource Management field in University of Sulimani. She is a member of quality assurance committee at ministry of higher education in Kurdistan region-Iraq.

International Institute for Applied Knowledge Management

Refereed Paper Proceedings - KM Conference 2017 – Novo Mesto, Slovenia

A Publication of the International Institute for Applied Knowledge Management

The use of fuzzy logic to assess the knowledge gap in innovation processes

[Research-in-Progress]

Magdalena Jurczyk-Bunkowska, The Opole University of Technology, Poland, m.jurczyk-bunkowska@po.opole.pl

Przemysław Polak, Warsaw School of Economics, Poland, ppolak@sgh.waw.pl

Abstract

This article discusses the problem of the estimation of the knowledge gap size in innovation processes. This gap is defined as the knowledge that the company must acquire and deploy to implement an innovation process. This parameter allows to characterize the innovation process at the stage of planning. It depends on the novelty and the scope of innovation implementation, but these are usually defined in a descriptive and vague manner. The authors propose a method for assessing the knowledge gap in innovation processes based on fuzzy logic. The article also presents an example of the use of that method to estimate the knowledge gaps in innovation processes.

Keywords: Innovation process, fuzzy logic, knowledge gap, and knowledge management

Introduction

The importance of knowledge in management sciences has been discussed for a long time (Hayek, 1945; Polanyi, 1966; Orlikowski 1996). Nowadays, the issue is rapidly gaining in importance due to the implementation of the knowledge-based economy. Knowledge plays a key role in creating innovation (Hall & Andriani, 2003). Organizations that effectively manage and transfer their knowledge are more innovative and perform better (Riege, 2007). That is why the ability of organizations to survive in the market is determined by the ability to identify, capture, create, share or accumulate knowledge (Jang, Hong, Bock, Kim, 2002). A strong relationship between knowledge management and innovation is well described and justified through many studies conducted in recent years (Cavusgil, Calantone, & Zhoa, 2003; Plessis, 2007). It is also emphasized by a definition indicating that innovation should be seen as a creation of new knowledge and ideas to facilitate new business outcomes aimed at improving internal business processes and structures and to create market driven products and services (Plessis, 2007).

The implementation of innovation is dictated by the need to improve competitiveness (Dosi, 1988), to rise in profit (Cooper, Edgett, & Kleinschmidt, 2001) and to expand business (Johne, 1999). Although the purpose of innovation is well recognized, there remains the problem of defining the very concept. It is clear that it is something different form invention (Schoen, Mason, Kline, & Bunch, 2005), which need not be implemented in practice. Whereas innovation

must entail measurable economic effects. Thus, innovation can be defined as the effective application of processes and products new to the organization and designed to benefit it and its stakeholders (West & Anderson, 1996). This definition emphasizes that innovation can be any implementation of knowledge which is new from the point of view of a company. However, innovation is not always perceived in this way in everyday, official and scientific language. The definition by Mortensen and Bloch (2005) suggests that the significant novelty of implemented knowledge is required to call it innovation: the implementation of a new or a significantly improved product (a good or a service), or a process, a new marketing method in business practices, a workplace organization or an external relation. This definition also highlights the concept of innovation, which includes new products, processes, organizational and marketing methods. The high degree of novelty of the implemented knowledge is indicated also by defining innovation as any novel product, service, or production process that departs significantly from a prior product, service, or production process architectures (McKinley, Latham, & Braun, 2014). However, the word 'significantly' is vague and can be interpreted in different ways. Therefore, Rogers's (2003) proposal should be accepted that the innovation is every idea, practice, or object that is perceived as new by an individual or other unit of adoption. Furthermore, the entire space (spectrum) of innovative solutions ranging from incremental innovation, to radical innovation should be taken into consideration. The latter describes spectacular solutions, allowing a company to achieve a long-term success (Schepers, Schnell, & Vroom, 1999). However, only continuous innovation based primarily on long-standing innovative efforts, including the implementation of improvements characterized by a small degree of novelty, is the way to get a winning competitive position (Hoffman, 1999).

Given the dynamic changes in the market, managers often face the question how risky is the implementation of a new technology, a product, an organizational or marketing method. The answer to this question allows assessing whether to take that risk alongside the benefits expected from the implementation of an innovation. Currently, managers are primarily guided by their own intuition and inclination to take on challenges. Often this is not enough, especially in cases where the decision must be justify to superiors. Innovation processes, particularly those with a high degree of novelty, are unstructured. They often involve actions of an experimental nature. However, it is not possible to precisely define them and to identify the beginning of a process. Thus, it is not possible to identify risk factors in the individual phases of an innovation process. Moreover, the higher level of novelty implies the lower supply of knowledge, the higher cost and the bigger uncertainty of its acquisition. However, even the success in acquiring knowledge does not guarantee the effective implementation of innovation. Equally important is the assimilation and implementation of knowledge in daily procedures. It depends largely on the human factor, and thus it also entails the risk directly proportional to the number of involved employees. The scope of innovation implementation indicates the number of departments within the organization that will be involved in creating and assimilating new knowledge. Unfortunately, both the novelty and the scope of an innovation implementation are determined in a descriptive and thus imprecise way. Therefore, fuzzy logic was chosen for this assessment. Fuzzy logic allows the construction and operation of the model, even if the knowledge, on which the model is based, is

not very accurate and precise. The research purpose is the development of a method for assessing the size of a knowledge gap in innovation processes. As a result, the size of the gap can be used as a parameter that allows estimating the risk of an innovation process. The method was originally developed for a company in Poland in cooperation with its managers.

The paper is organized into four sections. The next explains the concept of knowledge gaps in innovation processes and highlights its importance for management. The third section shows the theoretical basis of the use of fuzzy logic in assessing knowledge gaps. And the fourth section presents an example of the assessment of two innovation processes by showing the difference in the input parameters that describe the novelty and scope of innovation. The article ends with a summary, which indicates the purpose and application of the proposed method of assessing knowledge gaps.

The Concept of Knowledge Gaps in Innovation Processes

An innovation process can be defined as a sequence of actions leading to the transformation of an idea into reality (Brown, Lamming, Bessant, & Jones, 2013). Utterback and Abernathy (1975) described innovation as an iterative process, where "a basic idea underlying the innovation is developed over time in a predictable manner with initial emphasis on product performance, then emphasis on product variety and later emphasis on product standardization and costs" (p. 642). Van der Ven, Polley, Garud, and Venkataraman (1999) used a term 'innovation journey' to emphasize the uncertain character of innovation processes. An innovation process starts off with little information and many uncertainties. Knowledge, which is necessary to implement an innovation, is gradually acquired and assimilated as a result of a number of linked activities. Therefore, when planning an innovation process, one needs to be aware that for its completion it will be necessary to fill specific knowledge gaps.

The concept of organizational knowledge gaps, which emerged from the analysis of both empirical and the secondary data is discussed in detail by Haider (2003), who defined the knowledge gap as "organizational knowledge which a company currently lacks but is identified to be critically important for its survival and growth and, hence, needs to be filled". In this paper, it is proposed to define a knowledge gap in innovation processes as knowledge that must be acquired and implemented by a company in order to implement a specific innovation. The size of a knowledge gap determines the time and resources that are required to carry out the process of innovation. It also determines the degree of process complexity, which results from the number of experimental activities and links between individuals inside and outside an organization. Thus, knowledge gaps have a decisive influence on the risk of innovation processes.

Innovation depends upon knowledge. Therefore, when implementing an innovation process, a company must assess possessed knowledge and the ability to create, acquire and implement necessary knowledge (Plessis, 2007). The experience and the ability to transform and use knowledge in a company determine its level of innovation, i.e. the ability to implement innovations. That is why a specific knowledge gap and a risk resulting from it are acceptable to

some companies, and not to others. Thus, the determination of the size of a knowledge gap is not the final stage of the decision about the acceptance of risks associated with an innovation process, as shown in Figure 1.

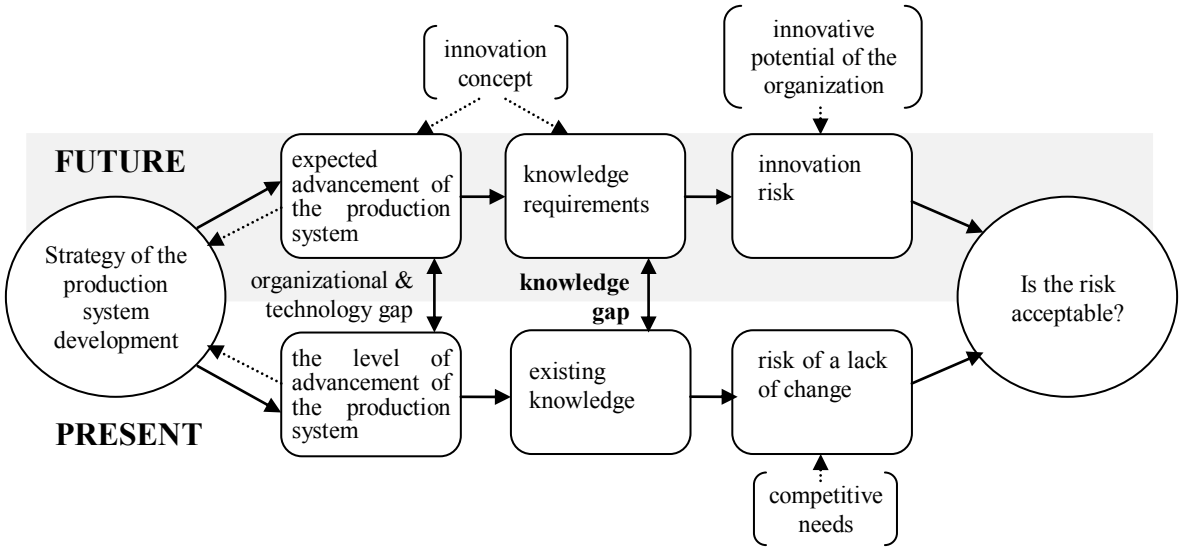


Figure 1. A knowledge gap as a primary, however not the final factor of risk assessment

A knowledge gap is also related indirectly to several other characteristics of innovation processes:

- the number of external sources of knowledge and the need for cooperation with external units,
- the necessary degree of formalization, associated with the need for knowledge transfer,
- the required competencies of a leader and a team,
- the scope of training,
- the scope of control and the method of process control.

A Fuzzy Logic Model for the Assessment of Knowledge Gaps

A human being with his ability to create knowledge is a key resource in innovation processes. Therefore, these processes are characterized by multidimensionality, a complex system of feedback and interactions, and thus a high degree of unpredictability. A decision concerning the adoption or rejection of an innovation process requires the comparison of potential benefits with the risk of failure. Due to the subjectivity and imprecision of the assessment of novelty and the scope of knowledge gaps, the use of fuzzy logic is proposed for the estimation of these elements (Dubois, Prade, & Yager, 2014).

The basis for assessing the size of the knowledge gap in innovation processes is the novelty and the scope of implemented changes. They are generally described in words, but point values can

be assigned to verbal expressions, as proposed in Table 1 and 2. The numerical values associated to verbal descriptions provide the range of variables related to innovation and the scope of an innovation process.

For the assessment of knowledge gaps, the following linguistic variables were used:

- n – the assessment of the degree of knowledge novelty,
- s – the assessment of the scope of knowledge implementation,
- g – the assessment of a knowledge gap.

Individual variables were assigned to the following sets of linguistic values (L) used to assess linguistic variables:

- $L(n) = \{N_1, N_2, N_3\} = \{low, average, high\}$,
- $L(s) = \{S_1, S_2, S_3\} = \{narrow, medium, wide\}$,
- $L(g) = \{G_1, G_2, G_3, G_4, G_5,\} = \{$ *very small, small, medium, large, very large* $\}$.

This assessment in the form of points (Table 1 and 2) provides a crisp input value which is converted into the membership degree of the value of linguistic variables n and s to fuzzy sets.

Table 1. The assessment of the novelty of innovation processes (variable n)

Point value (n)	Characteristics of the process in terms of the novelty
1	Improvement
2	Innovation on a plant scale
3	Innovation on an enterprise scale
4	Local innovation
5	State (country) innovation
6	Regional innovation (e.g. UE, USA)
7	Innovation for an industry segment (e.g. wood furniture)
8	Innovation on a branch scale
9	Innovation on an industry scale (e.g. wood industry)
10	Radical innovation

Table 2. The assessment of the scope of innovation processes (variable s)

Point value (s)	Characteristics of the process in terms of the scope
1	Single job
2	Set of connected jobs
3	Functional cell
4	Set of functional cells
5	Department of a company
6	Single process
7	Connected processes
8	Outside connected processes
9	Production plant
10	The whole system

A key element of the method is a membership function which is used to quantify linguistic terms. It should be tailored to the advancement of knowledge management in a company implementing the method. The trapezoidal and triangular membership functions $\mu(n)$ and $\mu(s)$ (Figure 2) were developed in consultation with the management of two companies in Poland. Their form suggests that there is no significant difference from the point of view of a company between the improvement and the acquisition of new knowledge. This is primarily due to the access to knowledge via the Internet. On the other hand, the implementation of innovation on an industry scale was considered equally demanding as the implementation of radical innovations because of the need for experimental studies.

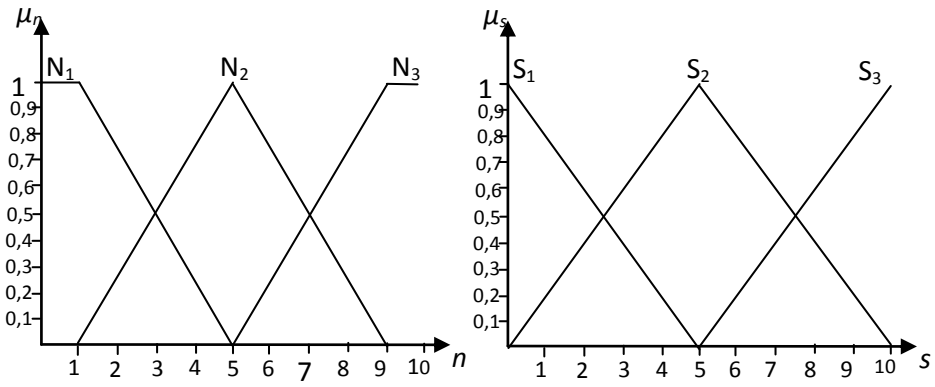


Figure 2. The membership functions of linguistic variables describing the novelty and the scope of innovation

A typical fuzzy model is composed of three building blocks: fuzzyfication (1), interference (2) and defuzzyfication (3). In the discussed case, the type 2IN/1OU model was used. The elements of the inference module include rules, the membership function of a linguistic variable and the inference engine. The inference module for assessing knowledge gaps used the following set of fuzzy rules:

- *R1: if (n= N_1) and (s = S_1) then (g = G_1),*
- *R2: if (n = N_1) and (s = S_2) then (g = G_2),*
- *R3: if (n = N_2) and (s = S_1) then (g = G_2),*
- *R4: if (n = N_2) and (s = S_2) then (g = G_3),*
- *R5: if (n = N_3) and (s = S_1) then (g = G_3),*
- *R6: if (n = N_1) and (s = S_3) then (g = G_3),*
- *R7: if (n = N_2) and (s = S_3) then (g = G_4),*
- *R8: if (n = N_3) and (s = S_2) then (g = G_4),*
- *R9: if (n = N_3) and (s = S_3) then (g = G_5).*

Membership functions of the fuzzy sets of output variable *g* were created by dividing the numeric space of variable *g* into four equal segments. This allowed the formulation of the five triangular fuzzy sets (See Figure 3).

The inference engine developed for the fuzzy model was reduced to the implementation of the following three steps:

1. Calculating the power of rules (*R1,.., R9*).
2. Determining the degree of fulfillment of conditions (*h*).
3. The aggregation of active rules and the creation of the membership function.

The degree of membership to the respective fuzzy set is determined for each variable in the rule assumptions. It takes a value in a range [0,1]. If the rule power is zero, it is considered that there was no activation of the rule. To determine the degree of the rule power was applied PROD operator, because it is responding very well to changes in the inputs of the model:

- $h_1 = PROD\ [\mu_{N1}(n^*), \mu_{S1}(s^*)] = \mu_{N1}(n^*) \cdot \mu_{S1}(s^*)$
- $h_2 = PROD\ [\mu_{N1}(n^*), \mu_{S2}(s^*)] = \mu_{N1}(n^*) \cdot \mu_{S2}(s^*)$
- $h_3 = PROD\ [\mu_{N2}(n^*), \mu_{S1}(s^*)] = \mu_{N2}(n^*) \cdot \mu_{S1}(s^*)$
- $h_4 = PROD\ [\mu_{N2}(n^*), \mu_{S2}(s^*)] = \mu_{N2}(n^*) \cdot \mu_{S2}(s^*)$
- $h_5 = PROD\ [\mu_{N3}(n^*), \mu_{S1}(s^*)] = \mu_{N3}(n^*) \cdot \mu_{S1}(s^*)$
- $h_6 = PROD\ [\mu_{N1}(n^*), \mu_{S3}(s^*)] = \mu_{N1}(n^*) \cdot \mu_{S3}(s^*)$
- $h_7 = PROD\ [\mu_{N2}(n^*), \mu_{S3}(s^*)] = \mu_{N2}(n^*) \cdot \mu_{S3}(s^*)$
- $h_8 = PROD\ [\mu_{N3}(n^*), \mu_{S2}(s^*)] = \mu_{N3}(n^*) \cdot \mu_{S2}(s^*)$
- $h_9 = PROD\ [\mu_{N3}(n^*), \mu_{S3}(s^*)] = \mu_{N3}(n^*) \cdot \mu_{S3}(s^*)$

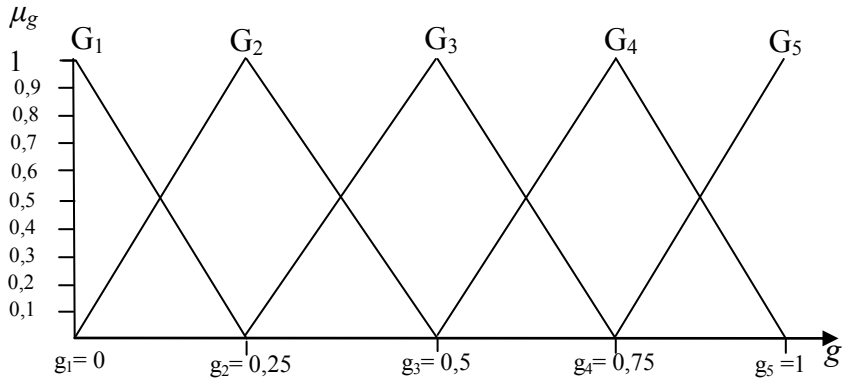


Figure 3. Fuzzy sets of linguistic variable *g* describing a knowledge gap in an innovation process

The values *n**, *s** are the input values of the evaluation model of knowledge gaps in innovation processes (See Tables 1 and 2). They are converted to fuzzy values according to the membership functions (See Figure 2).

The second step of the procedure is determining the degree of fulfillment of conditions (*h*). It is the basis for determining a fuzzy set $\mu_{Gx}(g)$, which is the result of rule activation. This operation is carried out for those rules that have been activated. The creation of modified membership functions $\mu_{Gx}^*(g)$ is performed using the operator MIN:

$\mu_{G1}^*(g) = MIN(h_1, \mu_{G1}(g)), \qquad h_1 > 0,$

$\vdots$

$\mu_{G5}^*(g) = MIN(h_9, \mu_{G5}(g)), \qquad h_9 > 0.$

The aggregation of active rules and the creation of the output membership function $\mu_{res}(g)$ involve the aggregation of output fuzzy sets from all the rules:

$$\mu_{res}(g) = \sum_{r=1}^{9} \mu_{G_x*}(g),\ x \in \{1,..,5\}$$

The output function becomes a basis for calculating the model output value (g), i.e. the evaluation of a knowledge gap in a particular innovation process. In this way, for crisp input values, the output is also a crisp value. A height method was used for defuzzification. It allows to take into account all rules when calculating a crisp output value (describing the size of a knowledge gap). Its characteristic feature is the replacement of the output fuzzy sets with crisp values placed at the points where they have the maximum values, as shown in Figure 4.

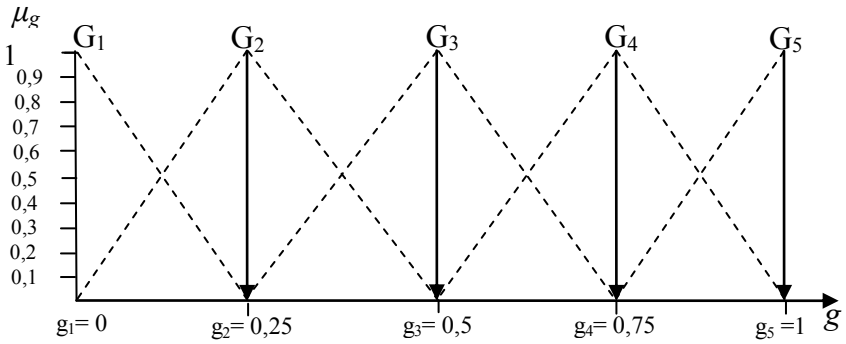


Figure 4. Replacing the fuzzy sets of variable g with single element sets

The end result of the procedure is determined using center of gravity method:

$$g = \frac{\sum_{r=1}^{n} g_x \mu_{Gx*}(g)}{\sum_{r=1}^{n} \mu_{Gx} *(g)}, x \in \{1,..,5\}$$

where: g – the crisp value of the fuzzy model (defuzzified value),
n – the number of conclusions ($h_1..h_9$<>0) which are different from 0,
g_x – the value of r-th fuzzy membership function,
$\mu_{Gx*}(g)$ – the fuzzy grade level.

An Example of Estimating a Knowledge Gap in Innovation Processes

Suppose that an entrepreneur wants to compare the risk of two innovations identified as P_1 and P_2. In the case of P_1 process the novelty is in the scale of a plant and has a value $n_1=2$, and the scope is related to outside connected processes and has a value $s=3$ (See Tables 1 and 2). The implementation of ERP system in a daughter company is an example of such innovation process. In the case of P_2 process the novelty is in the scale of industry segment and has a value $n_1=7$, and the scope is of a single job and has a value $s=1$ (See Tables 1 and 2). The development of a solution that enables automatic identification of an object location in a warehouse is an example of such innovation process. To assess the necessary degree of involvement in the management of the innovation process, the knowledge gap can be assessed using fuzzy inference system as shown in Figure 5. The fuzzyfication parameters of values: $n_1=2$, $s_1=8$, $n_2=7$ *and* $s_2=1$ are shown in Figure 6.

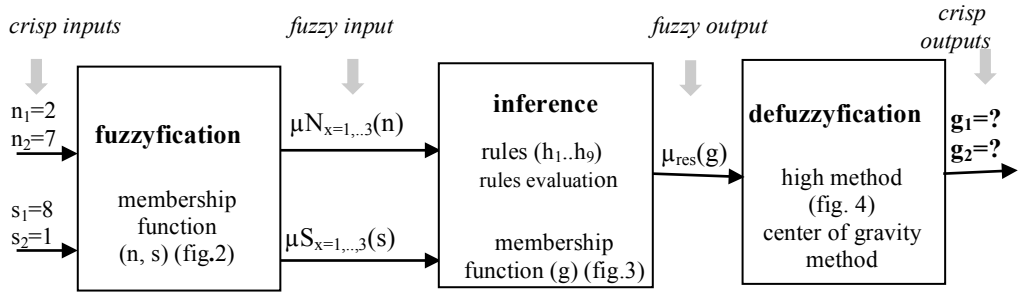
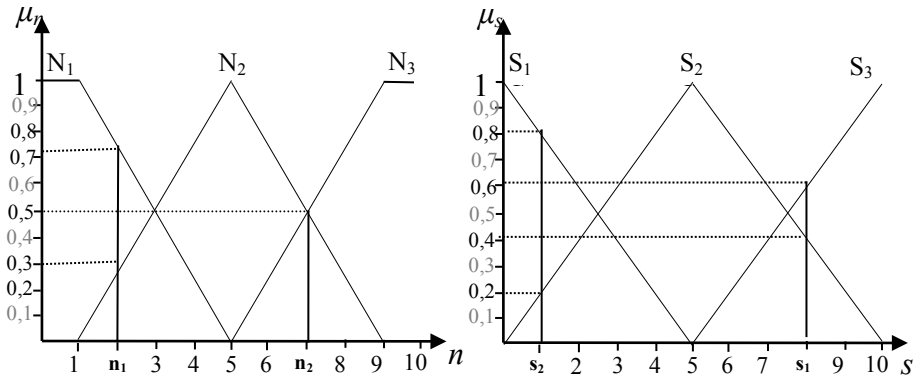


Figure 5. An inference system for the assessment of knowledge gaps of the innovation processes



$\mu_{low\ novelty}(n_1) = \mu_{N1}(n_1) = 0{,}75$
$\mu_{average\ novelty}(n_1) = \mu_{N2}(n_1) = 0{,}25$
$\mu_{high\ novelty}(n_1) = \mu_{N3}(n_1) = 0$
$\mu_{low\ novelty}(n_2) = \mu_{N1}(n_2) = 0$
$\mu_{average\ novelty}(n_2) = \mu_{N2}(n_2) = 0{,}5$
$\mu_{high\ novelty}(n_2) = \mu_{N3}(n_2) = 0{,}5$

$\mu_{narrow\ scope}(s_1) = \mu_{S1}(s_1) = 0$
$\mu_{medium\ scope}(s_1) = \mu_{S2}(s_1) = 0{,}4$
$\mu_{wide\ scope}(s_1) = \mu_{S3}(s_1) = 0{,}6$
$\mu_{narrow\ scope}(s_2) = \mu_{S1}(s_2) = 0{,}8$
$\mu_{medium\ scope}(s_2) = \mu_{S2}(s_2) = 0{,}2$
$\mu_{wide\ scope}(s_2) = \mu_{S3}(s_2) = 0$

Figure 6. The membership of $n_1=2$, $s_1=8$, $n_2=7$ and $s_2=1$ to fuzzyfication functions N_1, N_2, N_3, S_1, S_2, S_3

The activation of the rules and the determination of the fulfillment degree of the conditions (power rules) for P_1 process:

- $h_2=PROD\ [\mu_{N1}(2), \mu_{S2}(8)] = 0{,}75 \cdot 0{,}4 = 0{,}3\ (G2_{P1})$,
- $h_6=PROD\ [\mu_{N1}(2), \mu_{S3}(8)] = 0{,}75 \cdot 0{,}6 = 0{,}45\ (G3_{P1})$,
- $h_4=PROD\ [\mu_{N2}(2), \mu_{S2}(8)] = 0{,}25 \cdot 0{,}4 = 0{,}1\ (G3_{P1})$,
- $h_7=PROD\ [\mu_{N2}(2), \mu_{S3}(8)] = 0{,}25 \cdot 0{,}6 = 0{,}15\ (G4_{P1})$.

The activation of the rules and the determination of the fulfillment degree of the conditions (power rules) for P_2 process:

- $h_4=PROD\ [\mu_{N2}(7), \mu_{S2}(1)] = 0{,}5 \cdot 0{,}8 = 0{,}4\ (G3_{P2})$

- h_5=*PROD* [$\mu_{N2}(7)$, $\mu_{S3}(1)$]= 0,5 · 0,2 = 0,1 ($G3_{P2}$)
- h_7=*PROD* [$\mu_{N3}(7)$, $\mu_{S2}(1)$]= 0,5 · 0,8 = 0,4 ($G4_{P2}$)
- h_8=*PROD* [$\mu_{N3}(7)$, $\mu_{S3}(1)$]= 0,5 · 0,2 = 0,1 ($G4_{P2}$).

As a result of the conclusion of all the rules for P_1 process was obtained output function $\mu_{res}(g)$, marked in Figure 7 in bold blue line.

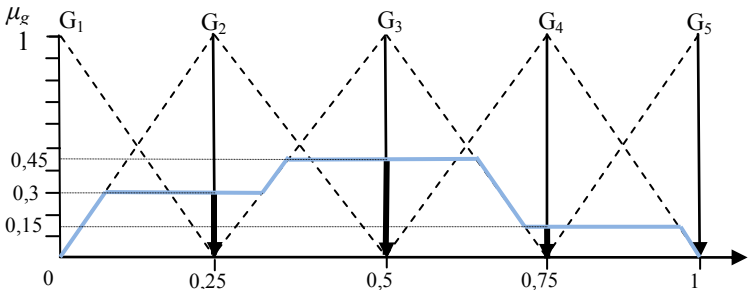


Figure 7. The conclusion active rules as a function $\mu_{res}(g)$ for P_1

The center of gravity method was used in order to obtain a crisp value describing the knowledge gap in the innovation process. It calculated the center of gravity for the area under the curve (See Figure 7). The final result of defuzzification procedure for P_1 was received in the following transformation:

$$g(P1) = \frac{0{,}25 \cdot 0{,}3 + 0{,}5 \cdot (0{,}45 + 0{,}1) + 0{,}75 \cdot 0{,}15}{0{,}3 + 0{,}45 + 0{,}1 + 0{,}15} = \frac{0{,}075 + 0{,}275 + 0{,}1125}{1} = 0{,}4625$$

The same procedure led to the following transformation for P_2 process:

$$g(P2) = \frac{0{,}5 \cdot (0{,}4 + 0{,}1) + 0{,}75 \cdot (0{,}4 + 0{,}1)}{0{,}4 + 0{,}1 + 0{,}4 + 0{,}1} = \frac{0{,}25 + 0{,}375}{1} = 0{,}625$$

The determination of the knowledge gap indicator provides information on the innovation process already at the planning stage. It shows how much effort the company has to make in order to provide the missing knowledge. The parameters allow to compare the various innovation processes and on that basis to decide on the implementation process, taking into account the scale of estimated difficulties.

Conclusions

We proposed the inference method based on fuzzy logic aimed at enabling the estimation of a knowledge gap in the planning stage of an innovation process. Knowing its estimated size, the complexity of the innovation process can be determined. Thus, the risk of innovation implementation can be estimated and collated with the expected benefits. However, it can also be applied to other parameters of the innovation process, such as process duration, or cooperation and commitment in obtaining knowledge.

The use of the knowledge gap estimation method for decisions making is important at the early stage of the innovation process called the frond-end phase. The method can serve both companies as well as institutions that support innovations financially. However, it requires further investigation. The biggest difficulties are expected when comparing parameters estimated from knowledge gaps, e.g. the time and the cost of an innovation process, with real values. Such studies are currently being conducted in two companies. However, at the present their scope is limited to innovation in manufacturing. The next study is planned for Information Technology (IT) projects, which involve the creation and transfer of new knowledge. The biggest concern raises the proper selection of the membership function. The final goal of the research is the development of a component of a computer-aided innovation planning system.

References

Brown, S., Lamming, R., Bessant, J., & Jones, P. (2013). *Strategic operations management*. London: Routhledge.

Cavusgil, S. T, Calantone, R. J., & Zhoa, Y. (2003). Tacit knowledge transfer and firm innovation capability. *Journal of Business and Industrial Marketing, 18*(1), 6-21.

Cooper, R. G., Edgett, S. J., & Kleinschmidt, E. J. (2001). Portfolio management for new product development: Results of an industry practices study. *R&D Management, 31*(4), 361-380.

Dosi, G. (1988). Sources, procedures, and microeconomic effects of innovation. *Journal of Economic Literature, 26*(3), 1120-1171.

Dubois, D., Prade, H., & Yager, R. R. (Eds.) (2014). *Readings in fuzzy sets for intelligent systems*. San Mateo: Morgan Kaufmann.

Haider, S. (2003). The concept of organizational knowledge gaps: Concept and implications. *DRUID Summer Conference 2003*, pp. 100-125.

Hall, R., & Andriani, P. (2003). Managing knowledge associated with innovation, *Journal of Business Research, 56*, 145–152.

Hayek, F. A. (1945). The use of knowledge in society. *The American Economic Review, 35*(4), 519-530.

Hoffman, R. C. (1999). Organizational innovation: Management influence across cultures. *Multinational Business Review, 7*(1), 31-49.

Jang, S., Hong, K., Bock, G. W., & Kim, I. (2002). Knowledge management and process innovation: The knowledge transformation path in Samsung SDI. *Journal of Knowledge Management, 6*, 479–485.

Johne, A. (1999). Successful market innovation. *European Journal of Innovation Management, 2*(1), 6-11.

McKinley, W., Latham, S., & Braun, M. (2014). Organizational decline and innovation: Turnarounds and downward spirals. *Academy of Management Review, 39*(1)*,* 88-110.

Mortensen, P. S., & Bloch, C. W. (2005). *Oslo Manual – Guidelines for collecting and interpreting innovation data*. Organisation for Economic Cooperation and Development, OECD.

Orlikowski, W. J. (1996). Improvising organizational transformation over time: A situated change perspective. *Information Systems Research, 7*(1), 63-92.

Plessis, M. D. (2007). The role of nowledge management in innovation. *Journal of Knowledge Management, 11*(4), 20-29.

Polanyi, M. (1966). *The tacit dimension*. London: Routledge and Kegan Paul.

Riege, A. (2007). Actions to overcome knowledge transfer barriers in MNCs. *Journal of Knowledge Management, 11*(1), 48-67.

Rogers, E. (2003). *Diffusion of Innovations* (5th Ed.). New York, NY: Free Press.

Schepers, J., Schnell, R., & Vroom, P. (1999). From ideas to business – how siemens bridges the innovation gap. *Research Technology Management, 42*(3), 26-31.

Schoen, J., Mason, T. W., Kline, W. A., & Bunch, R. M. (2005). The innovation cycle: A new model and case study for the invention to innovation process. *Engineering Management Journal, 17*(3), 3-10.

Utterback, J., & Abernathy, W. (1975). A dynamic model of process and product innovation. *OMEGA, 3*(6), 639-656.

Van de Ven, A. H., Polley, D. E., Garud, R., & Venkataraman, S. (1999). *The innovation journey*. New York, NY: Oxford University Press.

West, M. A., & Anderson, N. (1996). Innovation in top management teams. *Journal of Applied Psychology, 81*(6), 680–693.

Authors' Biographies

Magdalena Jurczyk-Bunkowska is an assistant professor in the Faculty of Production Engineering and Logistics at the Opole University of Technology. Her research interests include operational management as well as the innovation process management.

Przemysław Polak is a senior lecturer and a director of the Postgraduate Studies in Business Analysis in the Institute of Information Systems and Digital Economy at the Warsaw School of Economics. He is also an independent consultant in the field of information systems.

International Institute for Applied Knowledge Management

Refereed Paper Proceedings - KM Conference 2017 – Novo Mesto, Slovenia

A Publication of the International Institute for Applied Knowledge Management

Cybersecurity vital signs: The role of anomaly detection on insider threat triage

[Research-in-Progress]

Karla Clarke, Nova Southeastern University, USA, kc1127@nova.edu

Yair Levy, Nova Southeastern University, USA, levyy@nova.edu

Abstract

Detecting cybersecurity insider threat has become progressively challenging, as an over saturation of data has made it increasingly difficult to parse and consume information. In the past intrusion detection systems (IDS) were used to identify anomalies and potential misuse. However, IDSs do not specialize in the identification of anomalous activities. Thus, the development of anomaly detection systems (ADS) tailored toward the identification of deviations in behaviors is more sufficient. Though the use of anomaly detection systems have grown within cybersecurity, cybersecurity analysts face the problematic task of focusing on the right information in order to identify potentially malicious insider threats. In this paper, we will provide empirical evidence toward the identification and validation of cybersecurity vital signs that will aid cybersecurity analysts with triage for potentially malicious insider threats. A comparison of IDS and anomaly detection systems will be presented to depict the importance of separating anomaly detection from intrusion detection systems. We will also present the development of a prototype focused on effectively visualizing cybersecurity vital signs.

Keywords: Anomaly detection, cybersecurity, vital signs, intrusion detection, insider threat, visualization.

Introduction

Anomaly detection refers to models of intended user and application behaviors that detect deviations from normal behaviors, these deviations are referred to as anomalies. An anomaly detection system aids with the identification of abnormal behaviors based on complex data correlations (Patcha & Park, 2007). The overall goal of anomaly detection is to use complex data to identify patterns that do not conform to expected behaviors (Chandola, Banerjee, & Kumar, 2009). Misuse detection and anomaly detection are distinct in that misuse detection focuses on encoding and matching intrusion patterns in the data, while anomaly detection focuses on finding a normal pattern and identifying deviations from that pattern within the data (Chouhan & Richhariya, 2015). Within cybersecurity anomaly detection is used for "fraud detection for credit cards, insurance, or health care, intrusion detection, fault detection in safety critical systems, and military surveillance for enemy activities" (Chandola et al., 2009, p. 2). Vulnerabilities within networks and mitigation of cyber attacks are critical subjects for anomaly detection (Hong, Liu,

& Govindarasu, 2014). Anomaly detection has also been related to noise accommodation and novelty detection, these are all distinct notions. Noise detection refers to removing unwanted data that can be a hindrance to analysts. Novelty detection refers to the detection of previously unobserved patterns. These patters are often incorporated into normal models of expected behaviors (Chandola et al., 2009). Anomaly detection approaches include: statistical approach, Proximity-Based, Density-Based, and Clustering-Based (Chouhan & Richhariya, 2015). Anomaly detection is important because anomalies in data may often translate to critically, actionable information in a wide variety of applications (Chandola et al., 2009).

Visual analytics is often used within medicine for monitoring patients' vital signs to detect anomalies (Dutta, Maeder, & Basilakis, 2013). In cybersecurity insider threat triage will require investigation to identify anomaly metrics and attack patterns (Agrafiotis, Nurse, Buckley, Legg, Creese, & Goldsmith, 2015). Intrusion detection systems (IDS) are a long-standing technology within cybersecurity used to monitor network activities continuously and compare them with stored data. Therefore, IDSs can only detect known attacks (Imani, Rajabi, Taheri, & Naderi, 2015). While IDSs are good at detecting intrusions, they do not focus on identifying patters to detect deviations or anomalies (Gandomi & Haider, 2015).

Anomaly Detection vs. Intrusion Detection

Anomaly detection systems are a subset of IDSs, however, they are more adept at discovering unknown attacks and they make it more difficult for an attacker to go unnoticed (Patcha & Park, 2007). Anomaly detection is based on event correlation techniques to systematically establish the relationship between statistical data sets from various sources (Ten, 2010). There are two common types of IDSs, one is signature based, while the other is anomaly based (Ye, Emran, Chen, & Vilbert, 2002). IDSs are often signature based and require constant updates of rules and known attacks to stay effective (Patcha & Park, 2007). Signature based anomaly detection uses, intrusion signatures that are profiles of intrusion characteristics, if an intrusion signature is present then and intrusion has occurred (Ye et al., 2002). However, signature based detection is reactive and outdated (Jackson, 2012). Anomaly detection techniques use identified data to develop a baseline of normal activities (Ye et al., 2002). Some anomaly detection processes includes the identification of a score, this indicates "the degree of irregularity of a specific event", when the score exceeds the established baseline of normal activity, then, the occurrence will be flagged as an anomaly (Garcia-Teodoro, Verdejo, Fernandez, & Vazquez, 2009).

Riad, Elhenawy, Hassan, and Awadallah (2011) identified two problems with traditional IDSs, detection techniques, and user interfaces (UIs) that enable administrators to quickly recognize and respond to attacks. Anomaly detection stemmed from IDSs, however, the goal of an anomaly detection system is to detect new or unknown attacks (Yu, 2012). Unknown attacks are detected by creating a baseline of normal system, network, or program activity and if any activities deviate from the baseline then it is identified as a potentially malicious activity (Patcha & Park, 2007). Thus, activities that exceed the baseline as well as activities that are significantly below

the baseline can trigger further investigations. The baseline may be a running average for each pertinent category, if the data sets activities are above average, it will be deemed above the baseline and vice versa. The baseline may be organizational-dependent and could rely upon factors like the number of employees, type of data being collected, number of data sources, etc. (Legg, Buckley, Goldsmith, & Creese, 2015). Like organizations, individuals may be drastically different. However, there are established 'normal' baselines for each vital sign in relation to whether an individual is an infant or an adult. For instance, the vitals of a healthy adult are: "blood pressure: 90/60 mm/Hg to 120/80 mm/Hg, respiratory rate: 12–20 breaths per minute, pulse rate: 60–100 beats per minute, and temperature: 36°C–37.4°C" (Mok, Wang, & Liaw, 2015). These vital signs are typically displayed on an electrocardiogram monitor or EKG as depicted in Figure 1. As such relevant vital signs as well as their baselines may be established within cybersecurity for organizations, and monitored using a specific cybersecurity visualization interface.

Figure 1. Electrocardiogram Monitor (EKG)

Anomaly Detection Techniques

The most important step for anomaly detection is profiling the system and user activities, based on the delineated data sources. These data sources can include shell commands, system events, audit events, user keystrokes, and packages that traverse the network (Jyothsna, Prasad, & Prasad, 2011). Based on the anomaly detection technique data sources can be identified. Prior

research denotes anomaly detection techniques as: statistical, cognition, and machine learning based, as depicted in Figure 2 (Jyothsna et al., 2011; Garcia-Teodoro et al., 2009).

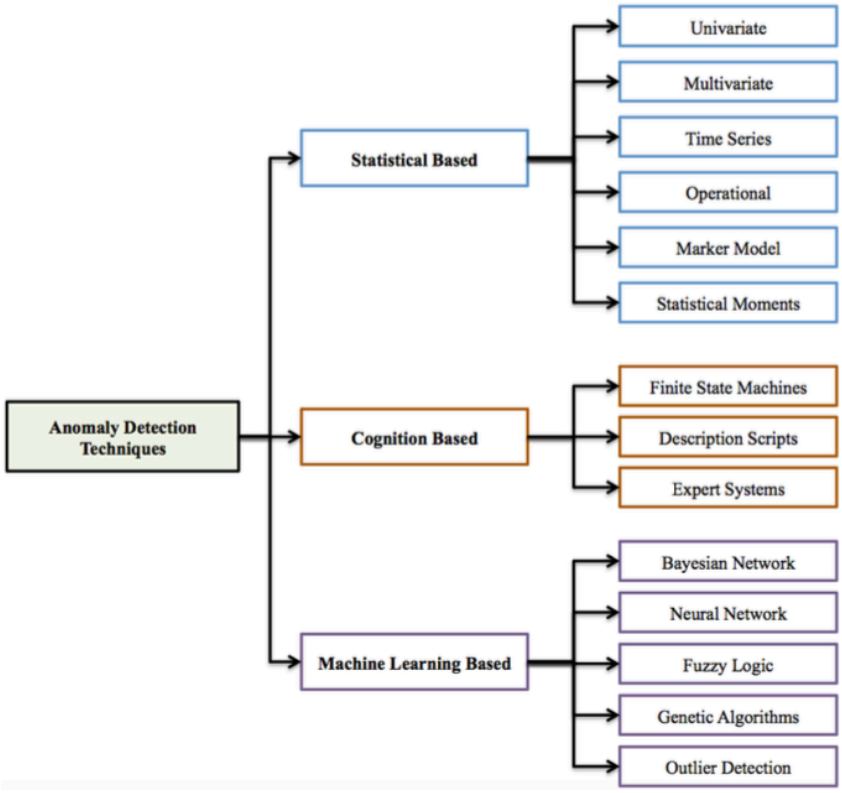


Figure 2: Anomaly Detection Techniques

(Garcia-Teodoro et al., 2009; Jyothsna et al., 2011)

Statistical based anomaly detection consists of techniques that capture network traffic and develop a behavioral profile. The developed profile based on metrics like traffic rate, number of packets, number of connections etc. for each protocol (Garcia-Teodoro et al., 2009). There are two steps involved with statistical anomaly detection. First 'normal behavior' is characterized, and then a time frame where behavior does not seem to be 'normal' is determined (Wang & Paschalidis, 2015). Within statistical based techniques, prior research started with univariate models that utilized a single metric, later multivariate models were developed using correlations of two or more metrics to determine deviations (Jyothsna et al., 2011). Time series models use time intervals with event counters, then consider the order and time frames of each activity and their value, so if at a given time the traffic observed is too low it may be identified as an anomaly (Garcia-Teodoro et al., 2009). The operational model also referred to as the threshold metric, counts events as they occur over a period of time, then an alert is triggered if the number of events is higher or lower than the specified thresholds (Jyothsna et al., 2011). The marker model is also known as the 'Markov model' which was broken into two approaches: Markov chains and

hidden Markov models (Garcia-Teodoro et al., 2009). Hidden Markov models (HMM) used one time series data modeling, this was identified as insufficient and Cao, Li, Coleman, Belatreche, and McGinnity (2015) developed and adaptive HMM (AHMAS) with anomaly states to detect anomalies based on a sequence of data and not a single value at a point in time. A Markov chain is a set of states connected by transition probabilities. Anomalies are detected by comparing the associated probability with the observed sequence (Garcia-Teodoro et al., 2009). With statistical moments in relation to the statistically based models, all identified correlations are termed 'moments'. If an event occurs above or below a moment it is identified as anomalous, the system determines the confidence interval based on observed user data that varies by user, this is an identified limitation of threshold models (Jyothsna et al., 2011).

Cognition based anomaly detection techniques utilize expert input to manually construct the desired model, this approach uses human input to determine legitimate behaviors (Garcia-Teodoro et al., 2009). A finite state machine (FSM) captures actions in states, each state contains information about the past, and an action is a description of an activity to be performed at a given moment (Jyothsna et al., 2011). Description scripts are scripting languages developed by the Intrusion Detection community, which describe signatures of attacks and can be used to identify attacks based on sequences of specific events (Jyothsna et al., 2011). Expert systems classify audit data according to rules, first different attributes are identified, then classification parameters and procedures are determined, and then finally the data is classified accordingly (Garcia-Teodoro et al., 2009).

Machine learning based anomaly detection techniques are based on models that allow patterns to be analyzed and categorized (Garcia-Teodoro et al., 2009). A series of mathematical inputs and outputs are utilized as a predictive method when detecting anomalies (Patan, 2015). Bayesian networks allow the ability to integrate prior knowledge and data to identify problematic relationships (Garcia-Teodoro et al., 2009). The Baysian network model is a mathematical framework that combines known information to postulate unknown information (Zaknich, 1998). Neural network models mimic the brain by attaining knowledge through learning (Garcia-Teodoro et al., 2009). Instead of utilizing crisp or precise rules, using fuzzy logic can improve detection accuracy by using approximate rules (Xu, You, & Liu, 2005). Fuzzy logic can be used to match any input or output of data this helps with understanding vague or ambiguous information (Dutta et al., 2013). Genetic algorithms use techniques stemming from biology, including inheritance, mutation and selection to identify deviations with no prior knowledge of the systems behaviors (Garcia-Teodoro et al., 2009). Unusual activities that defer from normal activities are considered outliers (Zhang & Zulkernine, 2006). Outlier detection consists of grouping data observations according to a given similarity, each new data point is grouped based on its identified similarity, points that do not belong to a cluster are determined to be outliers (Garcia-Teodoro et al., 2009).

Over the past decade after performing a review of anomaly detection systems and hybrid intrusion detection systems, Patcha and Park (2007) found that "today's intrusion detection approaches will not be able to adequately protect tomorrow's networks against intrusions and

attacks" (p. 3465). Therefore, anomaly detection methods will need to be advanced to address this problem. Jyothsna et al. (2011) denoted that identifying features to characterize user and system patterns would be the best way to clearly distinguish anomalous activities.

Proposed Experimental Research and Procedures

Prior research that examined detection of malicious insider cyber threats has mainly been focused on anomaly detection methods, malicious behaviors, or detection techniques used by cyber analysts (Agrafiotis et al., 2015; Azaria, Richardson, Kraus, & Subrahmanian, 2014; Legg et al., 2015; Santos et al., 2012). Thus, this work proposes an experimental research that will develop a visualization prototype, (QUICK.v™ displayed in Figure 3) which aims to develop a novel and effective detection method for the identification of anomalous activities when mitigating malicious insider cyber threats. QUICK.v™ was initially introduced within Hueca, Clarke, and Levy (2016). This proposed experimental research is a continuation of Hueca, Clarke, and Levy (2016) that will consist of three phases. Upon completion of these phases conclusions and recommendations for the development and visualization of cybersecurity vital signs will be presented.

Phase one of this proposed research will consist of identifying using SMEs critical cyber visualization categories that will be refereed to as 'cybersecurity vital signs'. Carlton and Levy (2015), utilized subject matter experts (SMEs) to successfully develop a list of top platform independent skills to form the basis for the set of scenarios that was later used to capture potential cybersecurity threats through hands-on tasks. Within this proposed research SMEs will be utilized for three different experiments. Within the first experiment SMEs will be utilized to identify cybersecurity vital signs. By identifying cybersecurity vital signs we intend to specify the SMEs' identified critical cyber visualization categories that should be displayed when using applications to detect potentially malicious insiders cyber threats. Phase one will also consist of identifying using SMEs' the rank order of the critical cyber visualization categories or vital signs that will be used to develop the prototype. In ranking the cybersecurity vial signs we be able to identify critical cyber visualization categories that should be displayed when developing a cybersecurity specific dashboard visualization prototype, QUICK.v™.

Phase two of this proposed research will consist of another experiment utilizing SMEs, to identify the most valid presentation of complex data correlations using the identified critical visualization categories over multiple visualization techniques. When visualizations are designed often times there are no explicit connections stated as to why the designer chose to use a particular method or technique, Mckenna et al. (2015) identified the lack of developmental research utilizing cyber analyst or SME input throughout the design process. By identifying critical visualization categories using SMEs we will ensure that the visualization methods used for displaying the cybersecurity vital signs are optimal for ensuring effective triage of potentially malicious insiders cyber threats. Within phase two we will also apply SMEs' identified critical visualization categories and techniques to develop QUICK.v™.

Within phase three of this proposed research we will conduct an experimental study using SMEs' to assess the effectiveness of the QUICK.v™ prototype performing cybersecurity triage. Inibhunu et al. (2016) sought to increase the effectiveness of cyber visualization tools by developing a system to provide adaptive level of detail in the interface. While the system was introduced the effectiveness of the system developed was not determined (Inibhunu et al., 2016). The SMEs' effectiveness or rating of satisfaction and value of the QUICK.v™ prototype will be obtained to determine the applicability of QUICK.v™ for SMEs when mitigating malicious insiders cyber threats. The effectiveness of QUICK.v™ will be based on the results of the data analysis performed on the quantitative data gathered from the cybersecurity experts.

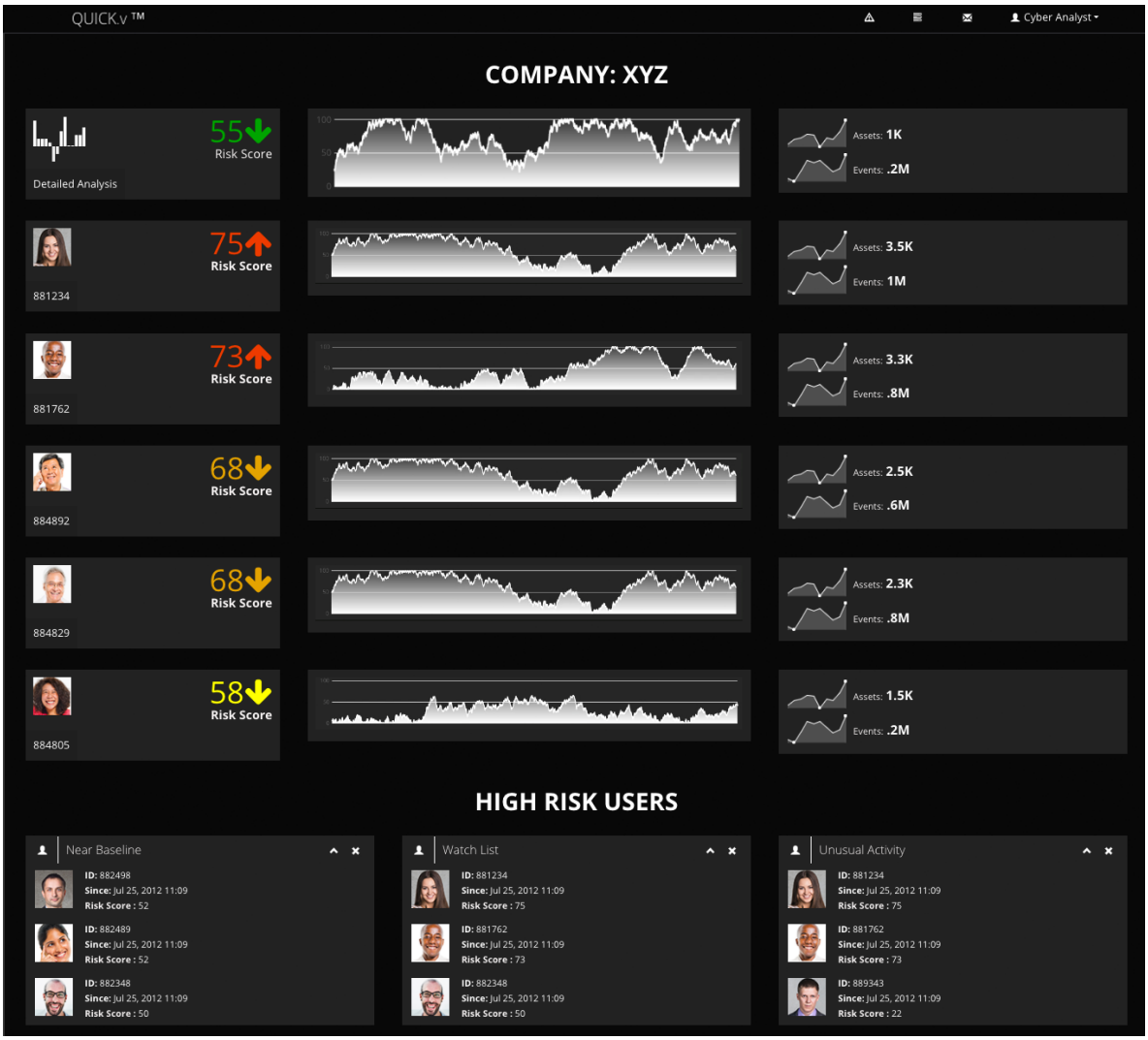


Figure 3: Proposed Quality User Insider CheckIng visualization (QUICK.v™) Prototype

Conclusion

This research agenda depicts the impotence of utilizing anomaly detection over intrusion detection within cybersecurity. Then a proposed experimental research was presented, that intends to identify cybersecurity vital signs and an effective visualization display to be used by cybersecurity analysts. Overall, identifying and validating cybersecurity vital signs that will aid cybersecurity analysts with triage for potentially malicious insider threats. This will be beneficial for focusing cybersecurity analysts during triage. Outlined within the research agenda is an update on an experimental research study in progress that will develop and validate using SMEs a cyber insider threat dashboard visualization QUICK.v™, as well as use it to conduct an experimental study, which aims to assess the effectiveness of enhancing the presentation of complex data correlations when mitigating malicious insiders cyber threats. This proposed work would aid cybersecurity practitioners with mitigating malicious cybersecurity insider threat.

References

Agrafiotis, I., Nurse, J., Buckley, O., Legg, P., Creese, S., & Goldsmith, M. (2015). Identifying attack patterns for insider threat detection. *Computer Fraud & Security, 7*, 9-17.

Azaria, A., Richardson, A., Kraus, S., & Subrahmanian, V. (2014). Behavioral analysis of insider threat: A survey and bootstrapped prediction in imbalanced data. *IEEE Transactions on Computational Social Systems, 1*(2), 135-155.

Cao, Y., Li, Y., Coleman, S., Belatreche, A., & McGinnity, T. M. (2015). Adaptive hidden Markov model with anomaly states for price manipulation detection. *IEEE transactions on neural networks and learning systems, 26*(2), 318-330.

Carlton, M., & Levy, Y. (2015). Expert assessment of the top platform independent cybersecurity skills for non-IT professionals. *Proceedings of the 2015 IEEE SoutheastCon,* Ft. Lauderdale Florida (pp. 1-6).

Carlton, M., Levy, Y., Ramim, M., & Terrell, S. (2015). Development of the MyCyberSkills™ iPad app: A scenarios-based, hands-on measure of non-IT professionals' cybersecurity skills. *Proceedings of the International Conference of Information Systems (ICIS) SIGSEC - Workshop on Information Security and Privacy (WISP) 2015*, Ft. Worth, Texas.

Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. *ACM Computing Surveys, 41*(3), 1-58. doi:10.1145/1541880.1541882

Chouhan, P., & Richhariya, V. (2015). A survey: Analysis of current approaches in anomaly detection. *International Journal of Computer Applications*, *111*(17), 32-36.

Dutta, S., Maeder, A. J., & Basilakis, J. (2013). Using fuzzy logic for decision support in vital signs monitoring. *Joint Proceedings of AIH/CARE* (pp. 29-33).

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management, 35*(2), 137-144. doi:10.1016/j.ijinfomgt.2014.10.007

Garcia-Teodoro, P., Diaz-Verdejo, J., Maciá-Fernández, G., & Vázquez, E. (2009). Anomaly-based network intrusion detection: Techniques, systems and challenges. *Computers & Security*, *28*(1), 18-28.

Hong, J., Liu, C. C., & Govindarasu, M. (2014). Integrated anomaly detection for cyber security of the substations. *IEEE Transactions on Smart Grid*, *5*(4), 1643-1653.

Hueca, A. L., Clarke, K., Levy, Y. (2016). Exploring the motivation behind cybersecurity insider threat and proposed research agenda. *Proceeding of the Knowledge Management (KM) 2016 Conference*, University of Lisbon, Portugal, (pp. 2-15)

Imani, M., Rajabi, M. E., Taheri, M., & Naderi, M. (2015). A novel approach to combine misuse detection and anomaly detection using POMDP in mobile ad-hoc networks. *International Journal of Information and Electronics Engineering, 5*(4), 245-249. doi: 10.7763/IJIEE.2015.V5.538

Inibhunu, C., Langevin, S., Ralph, S., Kronefeld, N., Soh, H., Jamieson, G. A., ... & White, M. (2016). Adapting level of detail in user interfaces for cybersecurity operations. *Proceedings of the Resilience Week (RWS)*, pp. 13-16.

Jackson, G. M. (2012). *Predicting malicious behavior: Tools and techniques for ensuring global security*. John Wiley & Sons.

Jyothsna, V., Prasad, V. R., & Prasad, K. M. (2011). A review of anomaly based intrusion detection systems. *International Journal of Computer Applications, 28*(7), 26-35.

Legg, P., Buckley, O., Goldsmith, M., & Creese, S. (2015). Automated insider threat detection system using user and role-based profile assessment. *IEEE Systems Journal, PP*(99), 1-10.

McKenna, S., Staheli, D., & Meyer, M. (2015). Unlocking user-centered design methods for building cyber security visualizations. *Proceedings of the IEEE Symposium on Visualization for Cybersecurity (VizSec)*, pp. 1-8.

Mok, W. Q., Wang, W., & Liaw, S. Y. (2015). Vital signs monitoring to detect patient deterioration: An integrative literature review. *International Journal of Nursing Practice, 21*, 91-98. doi:10.1111/ijn.12329

Patan, K. (2015). Neural network-based model predictive control: Fault tolerance and stability. *IEEE Transactions on Control Systems Technology, 23*(3), 1147-1155.

Patcha, A., & Park, J. M. (2007). An overview of anomaly detection techniques: Existing solutions and latest technological trends. *Computer Networks, 51*(12), 3448-3470.

Riad, A. E. D., Elhenawy, I., Hassan, A., & Awadallah, N. (2011). Data visualization technique framework for intrusion detection. *International Journal of Computer Science Issues*, *8*(5), 440-443.

Santos, E., Nguyen, H., Yu, F., Kim, K., Li, D., Wilkinson, J., . . . Clark, B. (2012). Intelligence analyses and the insider threat. *IEEE Transactions on Systems Management and Cybernetics, 42*(2), 331-347.

Ten, C. W., Manimaran, G., & Liu, C. C. (2010). Cybersecurity for critical infrastructures: attack and defense modeling. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 40*(4), 853-865.

Wang, J., & Paschalidis, I. C. (2015). Statistical traffic anomaly detection in time-varying communication networks. *IEEE Transactions on Control of Network Systems, 2*(2), 100-111.

Xu, J., You, J., & Liu, F. (2005). A fuzzy rules based approach for performance anomaly detection. *Proceedings of IEEE Networking, Sensing and Control*, pp. 44-48.

Ye, N., Emran, S. M., Chen, Q., & Vilbert, S. (2002). Multivariate statistical analysis of audit trails for host-based intrusion detection. *IEEE Transactions on Computers, 51*(7), 810-820.

Zaknich, A. (1998). Introduction to the modified probabilistic neural network for general signal processing applications. *IEEE Transactions on Signal Processing, 46*(7), 1980-1990.

Zhang, J., & Zulkernine, M. (2006). Anomaly based network intrusion detection with unsupervised outlier detection. *IEEE International Conference on Communications, 5*, 2388-2393.

Authors' Biographies

Karla Clarke is a Senior Associate in KPMG LLP's Cyber practice and a Ph.D. student in Information Systems at Nova Southeastern University. She holds a Bachelor of Arts in Anthropology from the University of Florida, and a Master of Science in Information Systems from Boston University. She is a member of the Information Protection practice at KPMG focused on the areas of identity and access management, privileged user management, logging monitoring and analytics. Prior to joining KPMG Karla work for another international consulting firm focused on infrastructure security and specializing in project management and security strategy implementation. Karla is a member of ACM, IEEE, and ISACA.

Dr. Yair Levy is a Professor of Information Systems and Cybersecurity at the College of Engineering and Computing, at Nova Southeastern University, the Director of the Center for e-Learning Security Research (CeLSR), and chair of the Information Security Faculty Group at the college along with serving as the director of the Ph.D. program in Information Assurance. He joined the university in 2003, was promoted to an Associate Professor in 2007, and to full

Professor in 2012. During the mid to late 1990s, Dr. Levy assisted NASA to develop e-learning platforms as well as manage Internet and Web infrastructures. He earned his undergraduate degree in Aerospace Engineering from the Technion (Israel Institute of Technology). He received his Masters of Business Administration (MBA) with Management Information Systems (MIS) concentration and Ph.D. in MIS from Florida International University. He heads the Levy CyLab, which conducts innovative research from the human-centric lens of four key research areas Cybersecurity, User-authentication, Privacy, and Skills (CUPS), as well as their interconnections. He authored over 60 articles, three book chapters, one book, and his publications have been cited for over 1,400 times by other scholarly research. Dr. Levy has been an active member of the US Secret Service (USSS)'s - Miami Electronic Crimes Task Force (MECTF) and The South Florida Cybercrime Working Group (SFCWG). He was trained by the Federal Bureau of Investigation (FBI) on various topics, and actively serves as a member on of the FBI/InfraGard, as well as consults federal, state, and local agencies. Dr. Levy serves on the national Joint Task Force of Cybersecurity Education (http://csec2017.org/), as well as other national initiatives related to cybersecurity workforce, education, and research. He is also a frequent invited keynote speaker at national and international meetings, as well as regular media interviews as a Subject Matter Expert (SME) on cybersecurity topics. Find out more about Dr. Levy and his research lab via: http://cec.nova.edu/~levyy/

International Institute for Applied Knowledge Management

Refereed Paper Proceedings - KM Conference 2017 – Novo Mesto, Slovenia

A Publication of the International Institute for Applied Knowledge Management

Balancing industry professional and researcher: The industry professional perspective

[Research-in-Progress]

Shonda D. Brown, Middle Georgia State University, USA, shonda.brown@mga.edu

Abstract

The advancement of Ph.D. programs have enabled industry professionals to maintain their jobs while pursuing a doctoral degree. Industry professionals pursue this degree for various reasons that range from a personal achievement to an academic position. However, industry professionals not already in academic positions can be challenged with continuing research upon degree completion. This paper provides an argument for the fact that in order to intensify this motivation, industry professionals must allocate time and become associated with an academic/research association. Moreover, it's argued that connection to an academic/research association will allow the industry professional to meet other scholars and encourage collaboration. Collaboration between the industry professional and academic scholar is essential to the integration of both scientific and practice knowledge towards the creation of new theories and knowledge production. The perspective taken in this paper is from the point of view of and industry professional, which completed a Ph.D. program while working full time, and had been engaged as an active member of an academic/research association. These endeavors were accomplished while balancing the demand from industry, time commitment needed for program completion, and active participation in an academic/research association.

Keywords: Researcher, work life balance, philosophy doctorate, industry professionals, knowledge sharing, knowledge production, hybrid Ph.D. programs

Introduction

With the evolution of Ph.D. programs to include hybrid methods combining both on-campus and online modalities, industry professionals have the ability to pursue a doctoral degree while maintaining their current occupation. However, as with any endeavor, this does not come without a price, trying to balance both the occupation and education. One of the primary purposes of the doctoral degree is the qualification to conduct research and the expectation to continue to add to the body of knowledge. However, unlike academic scholars who have the tenacity to continuously conduct research, this could be more of a challenge for industry professionals. Industry professionals aspire to obtain a doctoral degree for reasons such as: a) personal achievement; b) advance understanding of subject area; c) job promotion and d) academic position. Except for pursuing the doctoral degree to obtain an academic position, the remaining reasons may not influence the motivation to conduct research. As a result, time allocation and connection to an academic/research association will help to mitigate this challenge.

Time Allocation: While pursuing the Ph.D., a significant amount of time allocation is required to achieve the focus necessary for completing the doctoral degree requirements, especially the research. Once the doctoral degree is achieved, most industry professionals take a break, but it then becomes a struggle to resume the researcher mindset. Therefore, industry professionals must make a conscious effort to devote time to read literature. As a result, this time allocation will become a natural part of the balance for work, life, and research. While more time is most likely required when conducting the actual research, having this time already allocated, provides a foundation to build upon.

Academic/Research Association: Connection to an academic/research association that promotes research is critical to influencing the researcher mindset of an industry professional. Being actively engaged with other scholars that are continually conducting research will assist in stimulating the researcher mindset. Moreover, there are opportunities to collaborate on research with other scholars in the association. This will encourage knowledge production and sharing between the industry professional and academic scholar.

Zinskie and Rea (2016) noted that research conducted by industry professionals would typically focus on improving particular results experienced in practice in contrast to traditional research that tends to be more focused on generalization of results. "The research landscape is changing from one that frames the gap between theory and practice as a problem of knowledge transfer (seeing theory and practice as two distinct forms of knowledge) to a knowledge production problem with researchers and practitioners involved in the co-production of knowledge" (Williams & Schubert, 2017, p. 5400). In addition, Williams and Schubert (2017) stated when researchers are engaging in research from practice perspective, they are challenged with drawing on "both scholarly and professional knowledge to create robust theorizations" (p. 5400) when engaging in practice based research. Therefore, continued collaboration between the industry professional and academic scholar is important to combine both scientific and practice knowledge towards the creation of new theories.

In summary, industry professionals who were not already scholars may have a challenge to continue research upon completion of their Ph.D. degree. With a commitment of time allocation and connection with an academic/research association, industry professionals have the opportunity to engage with academic scholars. This partnership benefits both the industry professional and academic scholar from both a knowledge production and sharing perspective to advance the body of knowledge by helping to bridge the gap between scientific and practice based research.

Acknowledgments

I would like to thank the anonymous referees for their careful review and valuable suggestions. I also would like to thank the accepting editors, Dr. Nitza Geri, Dr. Bostjan Delak, Dr. Yair Levy, and Dr. Alex Koohang, for their recommendations and constructive comments.

References

Williams, S. P., & Schubert, P. (2017). Connecting industry: Building and sustaining a practice-based research community. *Proceedings of the 50th Hawaii International Conference on System Sciences*, 5400-5409.

Zinskie, C. D., & Rea, D. W. (2016). By practitioners, for practitioners: Informing and empowering practice through practitioner research. *National Youth-At-Risk Journal*, *1*(2), 1-7.

Author's Biography

Dr. Shonda D. Brown is a manager at a Fortune 500 company and part-time professor at the School of Information Technology for Middle Georgia State University. She has served as a referee research reviewer and editor for several international scientific journals and conference proceedings. A number of her papers won the 'best paper' award in national and international peer-review conference proceedings. Dr. Brown's research interests include information security and privacy. Moreover, she received a Bachelor's degree in Information Systems from Howard University, Master's degree in Information Systems from Drexel University, and Ph.D. in Information Systems with a concentration in Information Security from Nova Southeastern University. She is also a member of AIS, IEEE, ACM, and Upsilon Pi Epsilon Honor Society (UPE).

International Institute for Applied Knowledge Management

Refereed Paper Proceedings - KM Conference 2017 – Novo Mesto, Slovenia

A Publication of the International Institute for Applied Knowledge Management

Do digital natives have knowledge of mobile technology's acceptability to surveillance?

[Complete Research]

Scott C. Spangler, Middle Georgia State University*, GA, USA* scott.spangler@mga.edu

Abstract

This paper first seeks to meaningfully understand how the digital native culture utilizes innovations and mobile technologies to gather, transmit, sustain, or pool knowledge and information. Past scholars, and more recent researchers have continued to suggest digital natives are more sophisticated and understand technologies and innovations such as Smartphones and Internet sharing capabilities. However, a few dissenting voices have concurred that the culture is lacking in meaningful knowledge and abilities. Particularly, the literature lacks research on the culture's knowledge of mobile innovations and security concerns. Therefore, this paper seeks to clarify the culture's knowledge about mobile technologies, knowledge protection, and surveillance. Particularly, the paper seeks to answer the question: Do digital natives utilize mobile (Smartphones) to create knowledge and understand the devices are not secure? The research will shed a new understanding about the culture: digital natives lack awareness about device security.

Keywords: Digital native, higher education, acceptance of IT, mobile computing

Introduction

This paper seeks to understand how technology is changing, sustaining, and altering our everyday lives–particularly in the digital native generation. Digital natives are described as individuals growing up in a culture free of telephone cords and hardwired computers (Prensky, 2001a). Past scholars have quipped about cultural knowledge and sophistication for a decade. Today, scholarly work still reflects the digital native culture being sophisticated with technology and innovations. It appears that only three scholars, Bauerlein (2009); Rodi, (2014); and Spangler (2015), put forward a contrasting opinion.

Spangler (2015) and Bauerlein (2009) appear to be the most direct in questioning and point out the culture is and has been "significantly lower than adults in the previous decades" with abilities to read and have a "brazen disregard for books and reading" (pp. 39-40). The regard collectively is in response to past scholar's reflections and notations of the culture generating a sophisticated ire. Directly, Bauerlein (2009) cited past scholars' reflections and tots his research demonstrates a culture "embroiled in the swirl of social gatherings and contests, and it threats their intellectual development" while engaged with innovations and technologies (p. xii). This paper seeks to understand how mobile technologies are being used in the digital native culture. Particularly, the researcher seeks to understand–through a pilot study– if today's digital natives can navigate

through the world with his or her mobile device to generate, store, and or share knowledge. This question is derived through Tapscott's (2009) "new content creators" study question. In his study, Tapscott (2009) stated, "80 percent of the Net Geners under the age of 28 regularly visit blogs, the most popular way to create and share content… and some 64 percent of Net Geners engage in some form of content creation" or contribution (p. 45).

In greater hope, the scholar seeks to shed light onto misconceptions and misinformation about the ever-changing digital culture. The survey and data recovered in this project will be a guide for additional research and scholarly direction in the field of information technology and knowledge management based on the ever-changing digital native culture. Particularly, the research looks to understand the digital natives' generation deeper.

Literature Review

Scholars have agreed to the most part that youth today have utilized innovations in information technology over the last decade to enhance their lives. Speed is everything according to Tapscott (2009). The culture thrives on adaptability, and utilizing the next innovation, regardless of the conceptual understanding (Boyd, 2014; Howe & Strauss, 1993, 2000; Gautschi & Manafy, 2011; Jukes, McCain, & Crockett, 2010; Palfrey & Gasser, 2010; Prensky, 2012; Rodi, Spangler, Delorenzo, & Kohun, 2014; Spangler, 2015; Spangler, Delorenzo, Kohun, & Rodi, 2015; Spangler, Kovacs, & Kovalchick, 2014; Tapscott, 1999, 2009;). Prensky's (2001b) digital natives and digital immigrants (those learning the technology) coining has led the popular scholarly discussion for over a decade. Prensky's discussion has voiced an opinion of domain sophistication. Pointedly, Prensky (2001a, 2001b) and his followers have coined digital native technology sophistication. Following a preliminary discussion of Howe and Strauss (1993) ethnographic voyage in teen technology acceptance discovery, Presnky (2001a, 2001b) and Tapscott (1999) viewed the culture as being born into technology. Hence the culture has no barriers in knowledge and flexibility to learn new innovations. This is reflected later in Boyd's (2014) examination of high school aged digital natives. But, Boyd's (2014) research uncovered rich data suggesting a surface knowledge rather than a deep seeded mindfulness.

The mindfulness continuium was "ignorantly" discussed first in Bauerlein's (2009) exposure of digital native culture. Bauerlein (2009) proclaimed a culture of stupidity and arrogance. Bauerlein (2009) cited "Generation Next" "20-Somethings" calling themselves in New York Times article people who "don't suffer from literacy; they just suffer from e-literacy. We can't spell and we don't know synonyms because there's less need to know. What smart person would devote hours to learning words?" (p. 66). Spangler (2015) ethographic cultural examiniation dissertation, proclaimed a deep-seated inteligence indifference. He claimed the culture was driven by self gratification acknolwedged in social media "likes" and posts. Spangler (2015) quoted participants stating they only had a surface knowledge of technologies and inovations and were "addicted to communication technologies," which caused them "suffer from anxieties casued by the threat that they will be disconnected from technology" (p. 171). Interestingly, both

scholars Boyd and Spangler observed innovations being used for Internet social gatherings and more importantly to discourage abuse, race complications, and human rights abuse, which was first accorded in Howe and Strauss (1993) as well as Palfrey and Gasser (2010). Spangler (2015) stated, "the culture don't understand collaborative software and devices and had a diminshed skill set in utilizing technologies" (p. 172).

Because of the limitation in recent research in the digital native culture, a gap has been presented. First, the literature navigates around how digital natives use of mobile technologies. New research fails to requestion basic concerns from the past literature: if they understand and mobile technology surveillance. Past literature focuses on the prolific use of technology and negates the actual questioning of natives navigation abilities and surveillance understanding (Gautschi & Manafy, 2011; Howe & Strauss, 1993, 2000; Jukes, McCain, & Crockett, 2010; Palfrey & Gasser, 2010; Prensky, 2012; Rodi, Spangler, Delorenzo, & Kohun, 2014; Boyd, 2014). Spangler (2015) stated that the culture's sophistication is a myth and individuals have an "aversion to saftey" on the Internet with technologies. One example explained how a female failed to recognize a stalker "spying in on her private channel for more information about where she lived and what her plans were for the evening" (Spangler, 2015, p. 180). The latest research touches on mobile technology usage and a lack of understanding with repsect to software programming and hardware construction. This was relfected by Rodi et al. (2014) on higher education librarian's perspectives on digital natives. They captured similar data from librarians, who witnessed higher education students breaking down into tears and constantly seeking aid for rudimentary search engige issues. The librarians discussed how students would leave the library or forgo doing his or her research rather than continue to receive aid in understanding the elementary search engine software. Librarians remarked students would leave the library and quit his or her project before being anxiety stricken from the shame of needing librarian's help. The anxiety came from the digital natives never being trained how to use library software, which librarians blamed from cutbacks in budgets for library research education in higher education.

Mobile Surveillance Sophistication

More importantly, recent literature fails to question the culture's knowledge about technology's darker side. As an example, the literature has a gap on digital natives understanding about mobile technology surveillance capabilities. Boyd (2014) discovered the culture "sharing inappropriate conent…sexting or of inappropratie sexual images" with considerable disregard for personal saftey and considering the actions as "ephemeral gestures" (p. 64). However, some past literature does construct knowledge about the culture's understanding of the web's sinister abilities. Again, Boyd (2014) noted that children understanding means to pass "hidden messages" in "social steganography or hidden messages in plain site by leveraging shared knowledge and cues embedded in a particular social context" (p. 65).

Only a few scholars have sought answers about digital natives knowledge of sinister actions on the web. The past research conducted surface meaning into the culture's ability to understand the darker side of the Internet (Gautschi & Manafy, 2011; Howe & Strauss, 2000; Palfrey & Gasser,

2010; Smith, Skrbis, & Western, 2013; Tapscott, 1999). Palfrey and Gasser (2010) exercised a caution stating, "A young child going online dosen't have to do much to find himself exposed to images with graphic violence or sex that could cause him psycological harm" (p. 86). Takahashi (2011) stated that the culture is changing to one of "freedom from constraints of time-space and control of teachers and parents" leaving the youth open to unmonitored negative constructions and constraints (p. 73).

Howe and Strauss (2000) claimed the negatives were diminishing because of greater diseminsation of knowledge to the youth. Particuarly to point out the reduction of "gambling" in the culture showing a 47% "oppose gambling" and 70 percent of the youth now find the actions "damaging family and community life" (p. 211). Mobile security issues are still new in the literature. The literature navigates around questions concerning how deep digital natives understand actual mobile device surveillance and negative forces (Boyd, 2014; Spangler, 2015; Tapscott, 2010). Boyd (2014) had reflection into understanding how the culture can use the Internet for sinister action. His ethnography uncovered digital "bullying". His work shed light on a technology darkside in the culture. But, the research leaves out the deeper meaning about how much digital natives understand surveillance. Spangler (2015) touched on the surface meanings in the culture, but barley breaks ground on the culture's understanding of mobile surveillance knowledge. Particularly, Spangler's (2015) work remarks on the foolish and naïve nature in the culture rather than the direct understanding of surveillance.

Methodology

This study is constructed through an approved Institutional Review Board (IRB) survey from Robert Morris University. The 20-question survey seeks to understand how the digital native culture views and understands mobile technologies and computer technologies. The survey was administered through QuestionPro to 118 higher education students at Robert Morris University. The survey captured 79 full returns. The quantitative and qualitative answers will be used to construct the paper's meanings and conclusions.

The Tool Design

The survey first builds off Howe and Strauss's (2000) framing tool. The survey is designed in the same manner. The tool first stages questions according to Howe and Strauss's (2000) seven distinguishing traits to questions how the digital native culture views technology and understands it today. The general building questions tally conceptual meanings about technology and innovations. The survey seeks to reflect on past research about how youth understand innovations and its capabilities. Then the survey builds upon innovations to again seek differences or meaning comparisons in the generation gaps about the culture's knowledge of the Internet. Specifically, this expands upon the past survey tool to understand how today's digital natives are utilizing mobile technologies. The questions are framed using the past tools subscriptions. The reframed tool adds the questions with the replaced modifier being "mobile technologies" and not simply computer innovations or Internet access. The tool can be found in Appendix A.

The survey questioned how the digital natives view the Internet and understand its complexities and nature directly in consideration of mobile technologies. The questions build on each other to create a surface view about the culture's ability to understand mobile and Internet safety. Hence, the new adaptive tool based functionally on Howe and Strauss's (2000) model, examines the similar escalating scale questions but in relation to mobile innovations. As an example, the 2016 survey questions the demographics' openness to mobile technologies and thoughts towards leaving traditional computer media. This question has yet to be uncovered in scholarly literature. By reframing the original question by Howe and Strauss's (2000) on students perceptions of leaving books to utilize Internet capabilities, the tool effectively revisits perceptions in the culture (Howe & Strauss, 2000). The survey also constructs a series of questions about the participants' social web experiences and orientations that have yet to be expanded upon since Howe and Strauss' (2000) original survey. I copy of the enhanced survey can be observed here:

Additionally, the survey tool expounded upon the answer by mimicking (Howe & Strauss, 2000) survey tool's "Sidebar Voices" model. The model allowed traditional survey tools to be expounded upon by participants through qualitative response to each quantitative question. The qualitative responses or "voices" give direct inflection and expansion upon the participants' thoughts and unexpressed views in the quantitative answers.

Analysis of the Data

The survey material was first analyzed through QuestionPro, then frequency graphics to understand common threads were created. A series of visualizations of the data were generated to easily understand and communicate the frequency aspects to each question. From these threads of commonality, additional questions were driven for future research. This first round of surveyed participant responses helped determine the gap in meanings and need for a further deeper study.

The survey questions the culture's meanings and understandings through two parts. First, the survey utilized a traditional Likert style scale of questions. Then, to aid in understanding the responses, participants had the opportunity to elaborate on the responses qualitatively. Each quantitative question offered an opened qualitative response area. Additionally, any survey not completed fully was released from the data set.

Data Visualization and Construction Methodology

This case study followed a "Digital Ethnography" methodology. The fluid ethnography method utilizes captured qualitative the qualitative digital data in a quantitative study similar to how Howe and Strauss's (2000) survey vexed information into groups or pools of conjoining thoughts on a subject. These topics or frequencies of like-minded qualitative captures were utilized to impart deeper reflections on the quantitative discoveries Howe and Strauss generated in 2000. The method seeks to deconstruct the qualitative responses gathered in a spiral methodology into themes. The themes can be representations of "like" minded phrases or conceptions about a

topic. The spiral will aid the researcher's ability to find "clues" from the artifacts. The artifact clues may come from word frequencies or theme frequencies. Spangler's (2015) tool first utilized this thematic frequency to create the catalog of cultural clues. His cultural thematic cultural characteristic catalog tool is again the basis for discovering the themes represented in the captured artifacts.

Spangler's (2015) cultural characteristic catalog aids in discovering past themes in the literature (pp. 256-259). Additionally, the cultural characteristic catalog aids in uncovering new themes in this research. The frequencies guide in understanding meanings inside of the qualitative captures. The fluid-digital methodology focuses on "pools" of digital information that form in cyber domains, surveys, and artifacts similar to a metaphoric bend in a river. The qualitative knowledge pools in a traditional survey are webs of meanings. These meanings are reflected inside Howe and Strauss' (2000) "Sidebar Voices".

Researchers should consider these qualitative responses or digital conversations inside the traditional quantitative survey an ethnographic artifact. Understanding this concept, this paper utilized the qualitative construction as a traditional set of artifacts and details towards cultural meanings. To conduct this culture meaning in the data, this study followed the stream of qualitative captures in a fluid method to join participants' thoughts together in a digital stream (conversation) of thought. Hence, the captured qualitative data was viewed as a stream of interconnected thoughts and discussions as if the researcher was questioning the participants in a live social media chat. From this qualitative tracking discovery, the flow process can be uncovered. The process charted in Figure 1 demonstrates how digital natives are utilizing mobile devices to transfer information and respond to instruments such as this survey. Recent literature on social media posting elaborates on the pooling effect and collecting effect in the culture (Howe & Strauss, 2000; Nelson, 2015; Spangler, 2015).

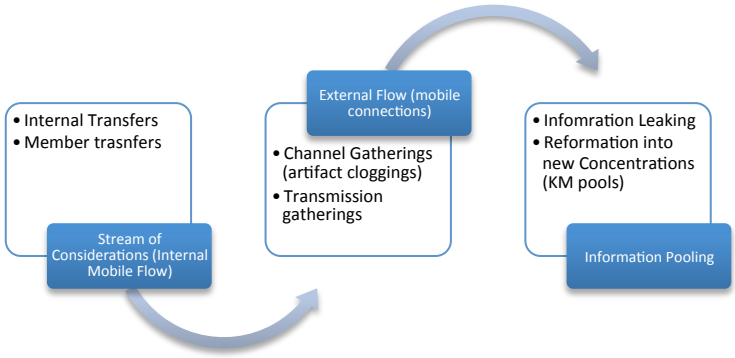


Figure 1. Qualitative Digital Theme Gathering

Goffman's (1986) construction of frame analysis trapped meanings in stages (Streams of Consideration). The fluid artifact process captures "frames of artifacts" in motion or action similar to Goffman's (1986) views of frames in stages. It creates an action capture to observe and create themes from the ongoing or fluid transmission in the digital world online. The stage in this

capture is not a proverbial theatre, but rather a virtual construct observed in the qualitative responses in the survey. These constructions will later form a revision or need to revise the survey's questions.

Research Questions

RQ1: Do digital natives in 2017 construct knowledge through their mobile devices?

RQ2: Do digital natives understand their mobile devices can be under surveillance?

Results and Discussion

Out of the 79 students surveyed, or 37% considered themselves experts in computer information science. However, 40% considered their abilities sophisticated enough to navigate and understand the basic components of a computer. Although, qualitative captures concluded only two participants regards themselves as being able to comprehend visual modeling or actual computer hardware.

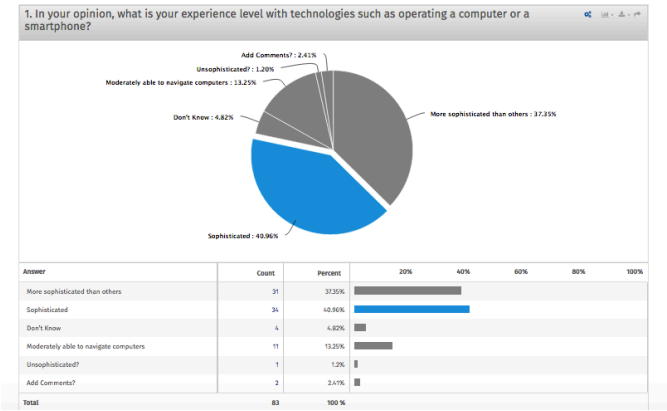


Figure 2. Technology Experience

The survey captured a "telling clue" that the 65% participants in the higher education demographic utilized the technology as a medium for communication and not business or educational purposes (Figure 2). This was originally reflected in Howe and Strauss' (2000) survey and later in Boyd (2014) and Spangler (2015). Spangler (2015) reflected a deeper meaning about the communications abilities stating the participants used the medium as "selfies" to highlight an individuals' emotional "wellness" more than actual traditional communication (p. 181).

Cultural Communication Change in Innovations

The participant responses constructed a theme demonstrating social media sites and "text messages" on Smartphones are the main media for communication in the culture. This reflection is observed throughout the literature and deeper in scholar's thoughts such as Turkle (2011). Turkle (2011) declared the culture living "alone together" because of the frequency of texting use and silent nature in communication across the observations. In this survey, qualitative responses aided the researcher in understanding and confirming this continuing fact.

Pointedly, the participants stated they prefer not generating "detailed emails". The culture prefers fast media such as text and "yes or no" answers to questions. Furthermore, many of the participants stated they had no use or understanding of "complex computer programs" outside of social media sites. Although the participants stated that they didn't know how to use sophisticated software for communication or data transfer, 53% of the participants contended they could communicate globally through technology. But interestingly, the qualitative responses contradicted the participants' quantitative analysis. Participants stated, "I don't even understand basic [Skype] communication let alone [how to] use technology to communicate globally." The overall consensus in the participants (79%) stated they use a "smartphone" to connect and converse (Figure 3). The data constructions showed no prescribed methods, or person to whom the culture connects with globally. Past research by Spangler (2015) demonstrated the culture connecting on computers and Smartpads through interactive global games such as digital word Scrabble. This survey could not confirm or deny the fact. Interestingly, the qualitative data constructs a meaning in the culture or new theme possibly. The fluid digital ethnography spiral determined a theme suggesting the demographic primarily utilizes the "smartphone"

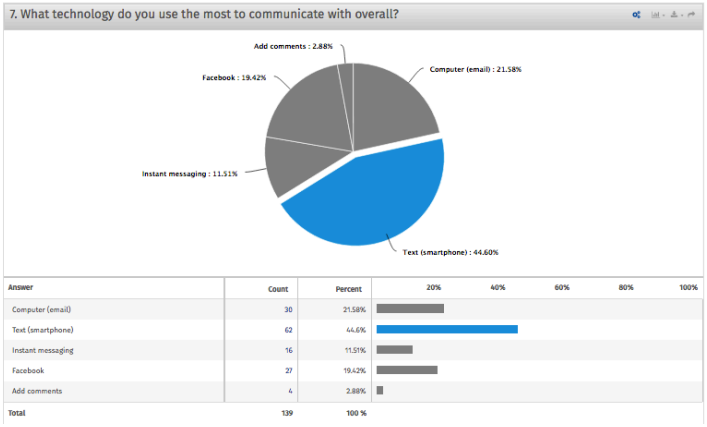


Figure 3. Technology Usage

technologies to "communicate to my dad or mother" in text forms. It also demonstrated a move in the digital immigrants' culture to communicating in text with his or her child rather than actual computer or email.

The participants responded that they generally use technology to communicate. The participants remarked as a theme qualitatively they have no understanding about how to actually build a computer or a "complex program". Additionally, inside of the qualitative thread, participants stated they had no use for innovations to aid themselves in business or other industry related activities. This disregard for computers again shadowed a movement away from email to simplistic text chattering. This foreshadows a changing view in the culture in technology use overall.

Understanding the Innovations Usage

Participants elaborated deeper, stating "texting" was a channel of communication for their culture and they find this channel most effective. Quantitative data reflected this sentiment: emails represented 21% of culture communication use. Texting represented the greatest methodology of 40%. But qualitative conversations furthered the idea that the culture was unfamiliar with how the texting software or mobile devices actually worked. Additionally, the qualitative remarks introduced a change in the culture's abilities. Participants reflected a lack in their abilities to use basic web tools such as website building software. Only 50% of the participants considered themselves able to manipulate or create a web page. But the participants remarked unanimously, they could only generate a webpage if they "they were using an online program that self generated the code" and visual basics. Additionally, only 34% of the participants considered themselves able to write HTML code or understand its meaning and representation. Nevertheless, the qualitative spiral created a common theme of misunderstanding about the culture's knowledge. This suggests a deeper change in the culture from Howe and Strauss (2000) and Boyd (2014). However, participants remarked they would need a "YouTube" video link to guide and tutor them through HTML data construction. This theme of using "YouTube" maybe a reflection of deeper meanings not uncovered in this survey or the literature.

Understanding Surveillance

Overall, 60% of the participants found they understood what computer cookies are and how they function. This action was one never documented in the literature and should be considered new cultural representation. This data represents an interesting suggestion that the culture can perceive that there is a system in the computer's hardware that generates a "history". Participants stated, "Cookies track your information and track your browsing history." But interestingly the researcher found again in the qualitative discussion, participants lack the ability to understand the fact "cookies" are also the media for tracking them physically and when they enter silent WiFi ranges with sleepers in the terms of service agreements. Additionally, a theme was uncovered that suggested the participants had "no idea" cookies could be on a mobile device. Participants remarked that cookies are only on computer websites and act as a surveillance mechanic on "home computers". Mobile devices are understood to be "free" of computer viruses and "watching" mechanics.

Limited Knowledge of New Surveillance Innovations

Participants considered mobile devices as "separate technologies" from home computers. Hence, the mobile devices "free" of commonplace computer alignments and constructs. The qualitative data shadows a theme of willful ignorance. This was discovered again in Spangler's (2015) and Boyd's (2014) reflections. Boyd's (2014) referred the laissez-faire or free spirited use of technology as "Digital Flaneurs" (p. 203). In essence, the participants become interested individuals wanting to "be part of public life, and visibility of their online activities creates tremendous consternation" (Boyd, 2014, p. 203). Additionally, the qualitative data demonstrated the culture's ability to understand, comprehend, and know about the existence and presence of super cookies on mobile devices. Although, one participant remarked that she thought marketers used super cookies to send her advertisements. But the participant did not discuss or point out the fact marketers or organizations utilize the super cookies to "track" consumers' virtual movements and track usage of the participants (Figure 4).

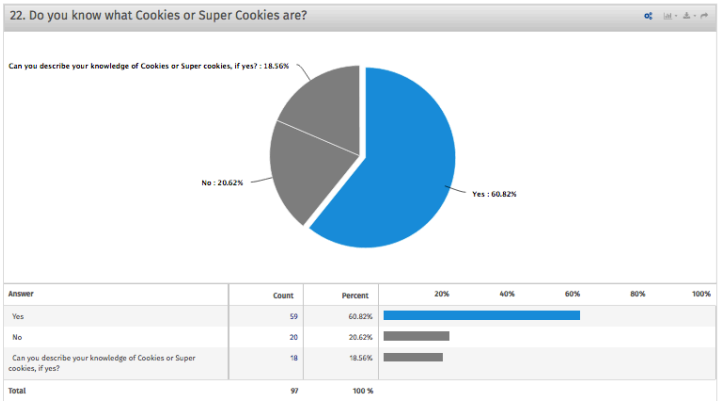


Figure 4. Understanding-Tracking Actions in Cookies

Interestingly, 73% of the participants remarked that they understood their smartphone could track their actions and motions on the web and physically. This thought was furthered through the qualitative theme uncovered in the data. Qualitative data represented a second theme in the forum. First, the participant's knowledge is viewed in the culture as superficial. Reflecting this thought, oddly, participants remarked the concept "Google" was tracking them on their computers and not mobile devices. This theme was uncovered throughout the data. More importantly, participants considered mobile devices only being tracked by his or her parents. Particularly, participants remarked they thought their parents used the tracking function like "Find my Phone" on an iPhone to track their motions, actions, and sometimes locations, through global positioning applications. Participants also stated they use such a function such as "Friend Finder" to locate friends in a non-surveillance or stalking methodology (Figure 5).

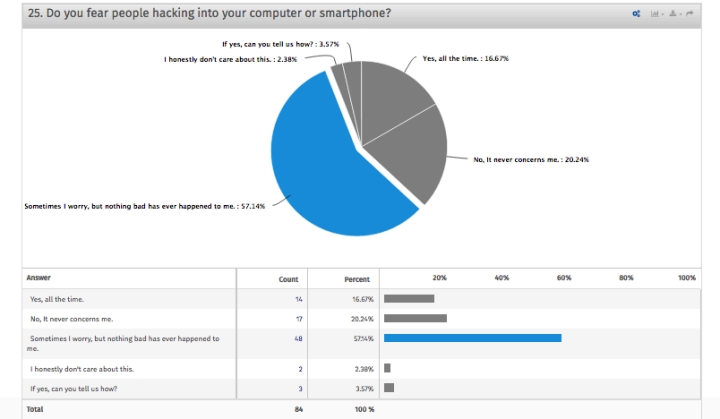


Figure 5. Understanding Surveillance

The final question outright asked participants if they understood innovation surveillance. Only 57% of the participants remarked, felt, or believed "sometimes" they had a person or organization tracking them or breaking into their mobile devices. Two percent of the participants stated, "They don't care" if they are being tracked. And interestingly, the survey's quantitative data found 20% of the participants have no concern in the matter. The qualitative responses created a theme of misunderstanding and lack of concern. The participants remarked they discovered, and had no knowledge of, deeper sinister actions being conducted through his or her mobile devices. This discovery reflects past literature's naïve youth attitude towards technology's darker side. This theme can be observed in the literature through the past decade, which introduces a notion–we are not educating our youth about the dangers of technology use, which was originally introduced in Tapscott's (1998) growing up digital; the rise of the Net Generation and later reflected in Howe and Strauss' (2000), Boyd's (2014), as well as Spangler's (2015) research. Tapscott (1999) remarked the youth are indiscriminately utilizing the technology to support "self-esteem" so they can "adopt another self in the real world" and no longer be characterized as a "nerd, nose-picker, fatty, or creep" and "kids get an opportunity to test the waters…before entering the elements of their personal life" (p. 92).

Conclusion and Furthering the Research

Because of the limitation and geographic constructs of this survey, cautions should be considered when recognizing the results. To further the research, this study prescribes another survey to first understand demographic differences in the culture's misconceptions on mobile technology and Internet based surveillance. The current survey failed, like past scholar's research to separate the genders for deeper meanings. Thus, this study's limited scope of participants and gender forgetfulness is also a weakness. The study prescribes a need for a larger generational research across the United States and its comparative countries around the globe to fully comprehend the culture's knowledge and misconceptions about technology surveillance. Therefore, it was

concluded that RQ1 (Do digital natives in 2017 construct knowledge through their mobile devices?) is not clear from the presented data and in need of a larger participant pool.

Although, it can suggested from the participants' responses to RQ2 (Do digital natives understand their mobile devices can be under surveillance?) that it has uncovered some meanings. This paper does suggest from the quantitative and qualitative responses that the culture is limited in knowledge about mobile device security. This prescribes a need for deeper research and points out a gap in the current literature.

Unfortunately, the survey failed to capture gender data, which again limits the comparative meanings. The gender segments were only captured in the qualitative data constructs. Therefore, the need to construct in the next cycle of surveys quantitative gender analysis for richer meanings is warranted. More importantly, this survey's limitations fail to complete the needed assessment on whether or not the culture contends its shared communications are protected knowledge management pools. A small glimmer was indicated in the survey, which demonstrated the culture does not understand its limitations in protecting data, and protection from its data being captured, stolen, or under constant strain of surveillance.

Interestingly, this research points out larger gaps in the literature review and a need for further questions. This paper did shed light into the need for a demographic question, which the literature too, neglects. Therefore, a further study on gender differences in the culture on mobile device knowledge security is suggested. Additionally, a survey to truly understand how the culture shares knowledge, stores knowledge, and understands the need for protecting its personal knowledge is warranted.

References

Bauerlein, M. (2009). *The dumbest generation: How the digital age stupefies young Americans and jeopardizes our future (or, don't trust anyone under 30).* New York, NY: Tarcher.

Bauerlein, M. (2011). *The digital divide.* New York, NY: Tarcher.

Boyd, D. (2014). *It's complicated: The social lives of networked teens (1st ed.).* New Haven, CT: Yale University Press.

Cortesi, S., Haduong, P., Gasser, U., & Beaton, M. (2013). *Youth news perceptions and behaviors online: How youth access and share information in a Chicago community affected by gang violence (SSRN Scholarly Paper No. ID 2342308).* Rochester, NY: Social Science Research Network. Retrieved from http://papers.ssrn.com/abstract=234230

D'Aurora, S., Spangler, S., & Wydra, C. (2014). Knowledge sharing systems and cultural identity in organizations. *Issues in Information Systems - A Journal of IACIS*, 15(1), 248–256.

Gautschi, H., & Manafy, M. (2011). *Dancing with digital natives: Staying in step with the generation that's transforming the way business is done.* Chicago, IL: Information Today, Inc.

Goffman, E. (1986). *Frame analysis: An essay on the organization of experience.* Northeastern.

Howe, N., & Strauss, W. (1993). *13th gen: Abort, retry, ignore, fail?* New York, NY: Vintage Books.

Howe, N., & Strauss, W. (2000). *Millennials rising: The next great generation.* New York, NY: Vintage Books.

Jukes, I., McCain, T. D. E., & Crockett, L. (2010). *Understanding the digital generation: Teaching and learning in the new digital landscape.* Kelowna, BC: Thousand Oaks, CA.: 21st Century Fluency Project; Co-published with Corwin, a SAGE Co.

Palfrey, J., & Gasser, U. (2010). Born digital: Understanding the first generation of digital natives (1st ed.). Basic Books.

Prensky, M. (2001a). Digital natives, digital immigrants. *On the Horizon, 9*(5), 1–6.

Prensky, M. (2001b). Digital natives, digital immigrants (part ii): Do they really think differently? *On the Horizon, 9*(6), 6-9.

Prensky, M. (2012). *From digital natives to digital wisdom: Hopeful essays for 21st century learning.* Thousand Oaks, CA: Corwin.

Rodi, A., Spangler, S., Delorenzo, G., & Kohun, F. (2014). A case study: Are digital natives dead? What are the key factors and perceptions librarian's view of the digital native culture in higher education? *Issues in Information Systems - A Journal of IACIS, 15*(2), 207–213.

Smith, J., Skrbis, Z., & Western, M. (2013). Beneath the "digital native" myth understanding young Australians' online time use. *Journal of Sociology, 49*(1), 97–118. https://doi.org/10.1177/1440783311434856

Spangler, S. C. (2015). What is the cultural experience of the digital native student today (2015)? (3702831 D.Sc.), Robert Morris University, Ann Arbor. *ProQuest Dissertations & Theses Global database*.

Spangler, S., Delorenzo, G., Kohun, F., & Rodi, A. (2015). Case study: Can digital natives adapt to technology's changes and speed? *Issues in Information Systems - A Journal of IACIS, 16*(3), 27–32.

Spangler, S., Kovacs, P., & Kovalchick, L. (2014). A case study: What are the practices librarians use in implementing and determining storing knowledge into digital data warehouses and archives. *Issues in Information Systems - A Journal of IACIS, 15*(2), 181–189.

Takahashi, T. (2011). Japanese youth and mobile media. In Michael Thomas (ed.), *Deconstructing digital natives: Young people technology and the new literacy* (pp. 67-82). New York, NY: Taylor & Francis.

Tapscott, D. (1998). *Growing up digital: The rise of the net generation.* New York, NY: McGraw-Hill.

Tapscott, D. (1999). Educating the net generation. *Educational Leadership, 56*, 5, 6–11.

Tapscott, D. (2009). *Grown up digital: How the net generation is changing your world.* New York, NY: McGraw-Hill.

Tapscott, D. (2010). Teachers must let students collaborate. In I. Jukes, T. D. E., McCain, & L. Crockett (Eds.), *Understanding the digital generation: Teaching and learning in the new digital landscape* (pp. 105–112). Kelowna, BC; Canada: SAGE Publications, Inc.

Turkle, S. (2011). Alone together: Why we expect more from technology and less from each other (1st ed.). Basic Books.

Verčič, A., & Verčič, D. (2013). Digital natives and social media. *Public Relations Review, 39*(5), 600–602.

Author's Biography

Scott C. Spangler, D.Sc. is an assistant professor of information technology at Middle Georgia State University. Spangler has taught before joining the MGAU faculty information science, information technology, and communication at Robert Morris University, Education Management Corporation and Oakbridge Academy. Spangler worked as a journalist and photojournalist for the past 20 years capturing and recording world histories prior to becoming a professor.

Appendix A

This appendix offers a view of the new survey tool created from the past. It will focus questions on literature holes. In addition to the mobile security holes, this survey will seek to find missing meanings about the digital native's cultures understand of the Dark Web and greater sinister actions in Internet and mobile technology use. This survey will add in gender and Spangler's (2016) digital native charting model.

What is your age demographic?
A. 10-15 years old Digital Juniors (individuals who have navigated youth in elementary and are in junior high school range with moderate IT knowledge);
B. 16 to 24-years old (Digital Constructionists, individuals who have circumnavigated surface IT tech scope and dominated in application knowledge or building);

C. 26 to 65-years-olds (Digital Inhabitants, individuals who have learned naturally or through advanced knowledge seeking interfaces and innovations
D. 65 to ? (Digital Emeritus, who have learned from others or their own IT and innovations for personal prosperity).

1. In your opinion, what is your experience level with technologies such as operating a computer or a smartphone? Add Comments with your answer?
2. Do you use a computer to do research (find answers) on the Internet? Add Comments with your answer?
3. Do you use technology to communicate with peers (family and friends) in your culture? Add Comments with your answer?
4. Do you purchase and shop on the Internet? Add Comments with your answer?
5. Do you use technology (the Internet or Smartphone innovations) to communicate globally? Add Comments with your answer?
6. Do you communicate believe your communications are safe? Add Comments with your answer?
8. Has technology allowed you to communicate with people or cultures that you would normally feel not comfortable to communicate with normally? Add Comments with your answer?
9. Are you able to write a computer program: HTML, JAVA, C, ect... Add Comments with your answer?
10. Can you create a web page, or design an online forum such as a blog site? Add Comments with your answer?
11. Do you know how to secure your internet site from hackers? Add Comments with your answer?
12. Do you understand there are different levels of the Internet? Add Comments with your answer?
13. Do you know what the Dark Web (Dark Internet) is currently? Add Comments with your answer?
14. Do you use your mobile device (smartphone) to find information? Add Comments with your answer?
15. Do you use your mobile device (smartphone) to connect on Social media? Add Comments with your answer?
16. Do you use your mobile device (smartphone) to conduct banking or business affairs? Add Comments with your answer?
17. Do you use your mobile device (smartphone) help you answer tests questions? Add Comments with your answer?
18. Do you feel the digital native culture is communicating across gender lines through technology? Can you share how you are doing this?
19. Do you feel the digital native culture is utilizing technology to break race lines? Add Comments with your answer? Can you tell us how?

20. Do you feel the digital native culture is utilizing technology to communicate across nationality or global cultural lines? Add Comments with your answer? Tell us how if you do?
22. Do you know what Cookies or Super Cookies are? If, Yes, can you describe them?
23. Do you know how much digital surveillance you're under daily from using your smartphone?
24. Are you aware your smartphone can track you? If yes, can you describe your knowledge of a tracking incident
25. Do you fear people hacking into your computer or smartphone? If yes, can you tell us how?
26. Have you ever had your identity stolen because of your smartphone or computer use?

International Institute for Applied Knowledge Management

Refereed Paper Proceedings - KM Conference 2017 – Novo Mesto, Slovenia

A Publication of the International Institute for Applied Knowledge Management

How the influential determinants of BI&A use intentions shift to socio-organizational determinants?

[Complete Research]

Tanja Grublješič, Faculty of Economics, University of Ljubljana, Slovenia,
tanja.grubljesic@ef.uni-lj.si

Abstract

Research and practice highlight that the use of Business Intelligence and Analytics (BI&A) can create competitive advantages for organizations. However, in order to create value for organizations, users need to accept BI&A and use it effectively. Identifying significant influential determinants of individual's BI&A use intentions is thus of great importance for organizations, since these can be proactively influenced by management action. Studies in the BI&A context have recognized the importance of socio-organizational determinants in explaining BI&A use intentions but a deeper understanding of how the basic acceptance determinants shift to socio-organizational motivations in influencing use intentions is however still missing. In response, we conduct a quantitative survey-based study to examine the relationships between result demonstrability, social influence, compatibility and performance perceptions as what we demonstrate to be significant elements of BI&A use intentions. The model is empirically tested through partial least squares (PLS) approach to structural equation modeling (SEM). We reinforce the importance and significance of socio-organizational considerations by showing that in addition to having strong direct impact on use intentions; these also have interaction effects by positively strengthening the perceived relevance of compatibility in impacting use intentions.

Keywords: Socio-organizational drivers, business intelligence & analytics, use intentions, compatibility, social influence, result demonstrability

Introduction

Fact-based (data-driven) decision-making using Business Intelligence and Analytics (BI&A) is regularly emphasized as a foundation for innovation and agility (Davenport, Barth, & Bean, 2012; Chen & Siau, 2011). Hence, understanding what are the relevant drivers that impact employees' decisions to use BI&A is of great value for organizations in order to better manage the organizational work environment to foster positive perceptions (Agarwal & Prasad, 1997). Researchers have also pointed out that individual level acceptance in the BI&A context is still under researched (Yoon, Ghosh, & Jeong, 2014).

The specific context of a BI&A use environment should be considered to fully understand successful BI&A acceptance (Hong, Chan, Thong, Chasalow, & Dhillon, 2014). People using

BI&A have high competences, experience and skills associated with IT/IS use and are usually higher educated (Luo, 2016). Motivations for accepting and using BI&A that enable and support organizational agility can no longer be based on assessments of the individualistic ease of use or utilitarian benefits of IT/IS use but should arise from socio-organizational recognition and approval of this behavior, visibility of the results of BI&A use and compatibility of its use with the work environment.

Previous case studies in the BI context already recognize the prevailing importance of socio-organizational drivers for acceptance of BI&A (Grublješič & Jaklič, 2015). Research in the BI&A use environment has also provided empirical evidence that individualistic considerations of effort and performance perceptions have no significant direct effect on individual's use intentions, but that socio-organizational considerations including social influence and result demonstrability influence their use intentions (Grublješič, Coelho, & Jaklič, 2014). The deeper understanding of this shift to socio-organizational drivers of BI&A use intentions and how these socio-organizational motivations interrelate and interact in influencing use intentions is however still missing. Thus, the main goal of this paper is to address this gap by conducting a quantitative survey-based study to examine the relationships between result demonstrability, social influence, compatibility and performance perceptions as what we demonstrate to be significant elements of BI&A use intentions, with providing additional evidence that BI&A use intentions are predominantly driven by socio-organizational considerations. Theoretical foundations provide a basis for our model development and identifying the interrelationships and interactions between the explored construct. The model is empirically tested through partial least squares (PLS) approach to structural equation modeling (SEM) on data collected from 195 medium and large sized organizations.

The structure of the paper is as follows. In the next section, the theoretical foundations with specifics of the BI&A use environment are elaborated. The research model is then conceptualized and hypotheses are developed. Further on, the research design, methodology, and results of the estimation are outlined. This is followed by a discussion of the results, including implications for research and practice with suggestions for future research.

Theoretical Foundations

Information Technology Acceptance

Significant theories representing theoretical foundations in technology acceptance research include (Mao & Palvia, 2006) the technology acceptance model (TAM) (Davis, 1989), innovation diffusion theory (IDT) (Moore & Benbasat, 1991; Rogers, 1983) and unified theory of acceptance and use of technology (UTAUT) (Venkatesh, Morris, Davis, & Davis, 2003).

The accumulated evidence has consistently proven that performance perceptions are the main and the strongest driver of IT/IS acceptance (Venkatesh et al., 2003; Venkatesh, Thong, & Xu, 2012). Yet this user acceptance research has mostly considered the technological aspects of IS

with focusing on the individual, utilitarian view of IS usage (i.e. Davis, 1989; Venkatesh & Davis, 2000; Venkatesh & Bala, 2008). Normative and other socio-organizational aspects of IS acceptance have thereby been considered marginally, i.e. only as potential additional predictors of IT/IS acceptance and in many cases also found not to be statistically significant (Venkatesh & Bala, 2008; Mao & Palvia, 2006, Venkatesh et al., 2003). Petter, Delone, and McLean, (2013) thus argued that the cultural and people aspects are underrepresented in IS success models. Shin (2015) further elaborated that, although traditional technology acceptance models have so far proved to be robust, they require modifications in the case of new and emerging trends and technologies.

Specifics of the BI&A Use Environment

Business intelligence and analytics (BI&A) is usually referred to as the techniques, data processing and analytical technologies, systems, business-centric practices and methodologies, and applications that analyze critical business data to help an enterprise better understand its business and market and make timely business decisions (Popovič, Coelho, & Jaklič, 2009). In the late 2000s, business analytics was introduced to represent the key analytical component in business intelligence. According to Gartner Research (2016), business intelligence and analytics remains first among the top business and technology priorities of Chief Information Officers (CIO) in 2016, pointing to their growing strategic importance and the need for greater attention in research studies. The use of BI&A has created competitive advantages for many organizations (Audzeyeva & Hudson, 2015), pointing to the importance of understanding their acceptance.

When studying the drivers of BI&A use intentions and their interactions, it is important to understand the specific usage context (Hong et al., 2014). BI&A is mainly used for effective decision-making and strategic goals including analytics, explanation, and prediction of business problems and trends (Li, Po-An Hsieh, & Rai, 2013). The use of BI&A is mostly voluntary compared to operational systems use, where the use is necessary for carrying out business processes (Grublješič & Jaklič, 2015). Further, the structure of users is different where BI&A users are generally more educated workers and mostly managers and/or expert business analysts (Luo, 2016). With the data in BI&A systems being more aggregated and integrated at the level of the entire organization, there is greater sharing of information connected to their use (Bose, 2009), which relates to the need for an improved organizational and information culture (Davenport et al., 2012). With operational IS use, information quality problems entail traditional problems of data quality such as accuracy and completeness, whereas in the context of BI&A use the focus is more on the relevance of the information provided by BI&A systems (Popovič, Coelho, & Jaklič, 2009). In general, compared to operational IS, the benefits of BI&A use are much more indirect and long-term (Grublješič & Jaklič, 2015). When BI&A is introduced, users need to adapt to different ways of carrying out business processes (Deng & Chi, 2013). The processes in which BI&A is used are less structured and there is a lower number of enforced procedures, i.e. well-defined business rules within business processes where BI&A is most commonly used. Therefore, the use of BI&A is more innovative and research oriented, where compatibility with users' needs is important.

Previous empirical research provides evidence that the end user's individualistic considerations of performance and effort perceptions, proven to be the main and strongest motivational drivers of technology acceptance throughout the past decades, have no direct effect in predicting BI&A use intentions. Instead, use intentions of BI&A are driven by socio-organizational considerations of social influence and result demonstrability. These include end user's perceptions of their output effectiveness, visible and recognized in an organizational environment (Grublješič et al., 2014).

Research Model and Hypotheses

IDT and UTAUT are the key theories that contributed to the development of the research model. The model combines UTAUT's performance expectancy, social influence and behavioral intention constructs (Venkatesh et al., 2003) and IDT's result demonstrability and compatibility (Moore & Benbasat, 1991; Rogers, 1983). All the indicators used in our research measure individual's perceptions based on their experience, since respondents in our study are familiar with or/and experienced users of BI&A (see Table 1 for demographic profiles of respondents in our survey). Our model thus includes all relevant use intentions determinants from IDT theory, missing in UTAUT. Performance expectancy from UTAUT includes IDT's relative advantage. Image and visibility from ITD are included in UTAUT's social influence (Venkatesh et al., 2003; Moore & Benbasat, 1999). Ease of use or effort expectancy was dropped as proven to be not statistically significant in our previous study (Grublješič et al., 2014). Previous case studies in the BI&A context already provided suggestions that effort perceptions are not a major consideration for BI&A users (Grublješič & Jaklič, 2015). Trialability and voluntariness of use (Moore & Benbasat, 1999) were not included as not being relevant in our study based on the explained context of use and profiles of respondents (users are experience with BI&A use, BI&A use is predominantly voluntary).

The proposed model builds on the authors' previous research findings (Grublješič et al., 2014) and adds compatibility from IDT theory (Moore & Benbasat, 1999; Rogers, 1983) to the previously researched model of BI&A use intentions. The additional construct is reasoned to have an interrelated relationship with the existing predictors. The reasons for this extension and an upgrade are to provide better explanation and prediction of BI&A use intentions, as well as to demonstrate the mediation and interaction effects between the already tested relationships. The proposed model and relationships are based on theoretical reasoning described below.

Use intention or behavioral intention, as an established predictor of both self-reported and actual usage and a meaningful surrogate for behavior (Agarwal & Prasad, 1997; Mao & Palvia, 2006), is the dependent variable. The determinant was operationalized as use intentions, measuring continuous use intentions based on individual's perceptions, since the respondents were experienced users of BI&A. Ajzen and Fischbein (2005) stressed out that "when volitional control is high, intentions are good predictors of behavior" (p. 192).

Performance perceptions, as an individualistic utilitarian criterion, represent the extent to which someone believes that using BI&A enhances his or her work performance (Venkatesh et al., 2003; Venkatesh & Bala, 2008). They have consistently been recognized as the key driver of IT/IS acceptance (Venkatesh et al., 2003). In the BI&A context, this relates to the individual's perception of less time and effort he/she spends on accessing and analyzing information and the consequences of improvements in data soundness and data access quality for individuals' work on one side, and his/her perceptions of the quality of their work output on the other. Since traditional models posit that performance perceptions have a direct impact on use intentions, we put the same hypothesis, although we might find the hypothesis not to be significant, based on previous findings in the BI&A context (Grublješič et al., 2014).

H1: Performance perceptions positively impact BI&A Use intentions.

Venkatesh and Bala (2008) defined result demonstrability based on Moore and Benbasat (1991) conceptualization as "the degree to which an individual believes that the results of using a system are tangible, observable and communicable" (p. 277). Since in the BI&A context the benefits of its use are more indirect and long-term and connected to organizational performance (Popovič et al., 2009; Grublješič & Jaklič, 2015), we would expect that when the results of BI&A system use are actually visible and recognized inside the organization, this should be reflected in individual intentions to use BI&A systems, which was already demonstrated in Grublješič et al. (2014).

H2: Result demonstrability positively influences BI&A Use intentions.

Venkatesh et al. (2003) defined social influence as "the degree to which an individual perceives that important others believe he or she should use the new system" (p. 451). The reasons why social influence plays an important role in the BI&A use context is that BI&A use is mostly voluntary and therefore use is importantly motivated by recognition and appreciation of use of BI&A by respected others. The other reason is that the benefits of use are typically not instantly visible, but more indirect and long-term compared to operational IS use. Accordingly, if users perceive that the organization and colleagues promote its use (Moore & Benbasat, 1991) they will be more internally motivated to use it and embed it into their routines.

H3: Social influence positively influences BI&A Use intentions.

Compatibility as one of the direct predictors of innovation acceptance behavior in IDT theory (Moore & Benbasat, 1991) has been proven to directly impact behavioral intentions in previous studies (Agarwal & Prasad, 1997; Mao & Palvia, 2006). Compatibility is defined as (Moore & Benbasat, 1991) "the degree to which an innovation is perceived as being consistent with existing values, needs and past experiences" of individuals (p. 195). Agarwal and Prasad (1997) interpreted compatibility as "perceptions of innovation being compatible with innovator's work behavior" (p. 568) and in their study find that compatibility significantly impacts acceptance behavior. Compatibility was also found to be an important driver of acceptance in BI&A context

case studies accounting task-technology fit to individual's work style as well as compatibility with the organizational work environment (Grublješič & Jaklič, 2015).

H4: Compatibility positively influences BI&A Use intentions.

Karahanna et al. (1999) in their study find out that compatibility and performance perceptions load to the same factor. They explain these results with arguing that compatibility as defined by Rogers (1983) as well as Moore and Benbasat (1991) is a multidimensional construct as it implies two types of compatibility (Karahanna, Straub, & Chervany, 1999): "normative compatibility, referring to compatibility with what people feel or think about an innovation and practical or operational compatibility, referring to compatibility with what people do" (p. 193). Karahanna et al. (1999) explained that for a personal technology in an organizational context "task-centric beliefs that focus on the ability of the technology to facilitate one's job (i.e. perceived usefulness and operational compatibility beliefs) may be inextricably linked in the user's minds" (p. 193) and concluded that "consequently it is unlikely that individuals would view an innovation as useful if it is not compatible with their work style" (p. 193). Venkatesh et al. (2003) also explained that compatibility is operationalized in such a way that includes aspects of technological and organizational environment that are designed to remove barriers to use. They explain that the construct "incorporates items that tap the fit between the individual's work style and the use of the system in the organization" (Venkatesh et al., 2003, p. 453). Based on the explained reasoning, users' perceptions of compatibility include individual's considerations of their personal task-technology fit with their existing work style, as well as socio-organizational considerations, such as novel situations, accustoming to new tasks, other's use, changes in system environment and managers' or organizational request to engage in adaptation cycles to achieve a better fit between the system and the new context (Sun, 2012).

When organizations implement and adopt BI&A systems, their execution of business processes changes fundamentally and users need to adapt to these changes (Deng & Chi, 2013), and some time lag is evident before users routinize, learn and adapt to a new system. Performance perceptions, i.e. utility evaluations of a technology, and usability evaluations of applying the technology to a specific task (Stern, Royne, Stafford, & Bienstock, 2008) are less emphasizes due to the lower structuredness of processes in which BI&A are used (Grublješič & Jaklič, 2015). We cannot exclude individualistic considerations of performance perceptions by default, as these were throughout the past three decades of technology acceptance research the strongest and most powerful driver of acceptance (i.e. Venkatesh et al., 2003; Venkatesh & Davis, 2008). But we do expect that these only have an indirect impact in the BI&A use intentions context. Individuals should perceive performance perceptions as significant influential determinant of their use intentions only, if they see this additional usefulness on and through compatibility with the new organizational work environment along with the fit to the according new work style adapted.

H5: Performance perceptions positively influence Compatibility.

H6: The impact of Performance perceptions on BI&A Use intentions is mediated by Compatibility.

Socio-organizational considerations of result demonstrability and social influence were demonstrated to be the statistically significant predictors of BI&A use intentions (Grublješič et al., 2014). Following the explanation above that compatibility perceptions also include socio-organizational considerations, it is reasonable to believe that result demonstrability and social influence would strengthen the relationship between compatibility and use intentions. The more the results of BI&A use are apparent to the individual and communicable in the organizational environment, the higher the perceptions of compatibility will be. Accordingly, the higher the organizational support, management incentives, visibility of BI&A use as well as peer support, the higher the perceptions of compatibility are to an individual.

H7: The higher the level of Result demonstrability, the stronger the relationship between compatibility and BI&A use intentions.

H8: The higher the level of Social influence, the stronger the relationship between compatibility and BI&A use intentions.

Research Design and Methodology

The questionnaire was used as a research instrument as perception determinants regarding use intentions are most commonly researched by pre-developed survey item scales in order to provide generalizability of the results (Venkatesh et al., 2003). The questionnaire was developed by building on previous theoretical basis to assure content validity. To ensure face validity (Cooper & Schindler, 2003), pre-testing was conducted using a focus group involving selected university staff and IS academics from the field who were not included in the subsequent research. We used five items to measure performance perceptions, which were adapted from Venkatesh et al. (2003), Davis (1989), as well as Venkatesh and Bala (2008) in order to fully capture and reflect the context specific performance perceptions. Four items for measuring social influence were adapted from Venkatesh et al. (2003). Result demonstrability was measured by three validated items adopted from Moore and Benbasat (1991). One item was dropped in our previous analyses due to the inadequate loading (Henseler, Ringle, & Sinkovics, 2009). Compatibility was measured by Moore and Benbasat (1991) three item scale. The indicator items for measuring BI&A use intentions were operationalized based on Wixom and Todd's (2005) behavioral intention construct measurements, as these provided the most suitable basis for developing the measurement of the BI&A use intentions based on volitional state. Wixom and Todd (2005) adapted the measurement scale from the technology acceptance theories (Davis, 1989; Venkatesh et al., 2003).

Detailed questionnaire with the indicators of the measurement model can be obtained from the authors upon request. Our proposed measurement model involved 18 manifest or observable variables loading on to 5 latent constructs: (1) *Performance perceptions*; (2) *Social influence*; (3)

Result demonstrability; (4) *Compatibility*; (5) *Use intentions*. The interactions between compatibility and both socio-organizational constructs result demonstrability and social influence were modeled to create new constructs, having as indicators the product of the standardized indicators relative to the constructs involved in the interaction (Henseler & Fassott, 2010). We used a structured questionnaire with seven-point Likert scales.

The data were collected in 2013 through a survey of 2173 medium- and large-sized business organizations in Slovenia, EU, representing the entire population listed in an official database. The questionnaires were administered by regular post mail and electronically. The procedure and the survey aims were explained in the introductory letter. Questionnaires were addressed to users of BI&A: top management, heads of departments and divisions, IS managers, etc.. The two rounds of call-up were conducted yielding altogether a sample of 195 completed surveys.

To conduct the data analysis, partial least squares (PLS), a component-based structural equation modeling (SEM) technique, was used. This is a widely-used methodology in the IT and IS field (Chin, 1998). The estimation and data manipulation were performed using SmartPLS (Ringle, Wende, & Will, 2005) and SPSS. To verify the mediating effects, we used bootstrapping and Sobel's test (1982) following the reasoning and procedures presented in Rucker, Preacher, Tormala, & Petty (2011). The interaction effects were tested following the Henseler and Fassott (2010) techniques.

Results

The demographic profiles of respondents are given in Table 1. The respondents were employed in companies from all business areas according to the national classification.

Table 1: Demographic profiles of respondents

Age	Max: 65	Min: 22	Average: 44.66
Gender	Male: 61.14%	Female: 29.53%	
Education	Elementary school: 0% High school: 4.64% Professionally oriented higher education: 15.46% Higher vocational education: 6.19% B.Sc.: 51.03% M.Sc.: 20.10% PhD: 2.58% Average: B.Sc.		
Experience	Min: 1 month	Max: 288 months (24 years)	Average: 33.66 months (2.75 years)

We have examined the reliability and validity measures for our reflective measurement model (see Table 2). All Cronbach's alphas exceeded the 0.7 threshold. The latent variables composite reliabilities were higher than 0.9 showing the high internal consistency of indicators measuring each construct. The Average Variance Extracted (AVE) was generally around 0.7 or higher, demonstrating the convergent validity of the constructs. All standardized loadings of the indicators in the model exceeded the 0.7 threshold at the 0.001 significance level, thus confirming the high indicator reliability and convergent validity.

Table 2: Means and standard deviations and reliability and validity measures of the measurement model

Construct	Indicator	Mean	Standard deviation	Loadings	T-statistics	Cronbach's Alpha	Composite reliability	AVE
Performance perceptions	PP1	6.051	0.990	0.887	39.259	0.9282	0.9458	0.7775
	PP2	5.747	1.239	0.861	15.578			
	PP3	5.853	1.139	0.938	64.310			
	PP4	5.763	1.257	0.892	24.030			
	PP5	5.723	1.246	0.828	14.117			
Compatibility	C1	5.337	1.207	0.865	20.796	0.8695	0.9197	0.7926
	C2	5.467	1.150	0.925	73.224			
	C3	5.483	1.150	0.890	32.563			
Result demonstrability	RD1	5.646	1.093	0.898	49.082	0.8558	0.9123	0.7763
	RD2	5.607	1.076	0.902	43.783			
	RD3	5.720	1.123	0.842	13.387			
Social influence	SI1	5.058	1.584	0.785	13.719	0.8042	0.8714	0.6290
	SI2	5.123	1.522	0.800	13.752			
	SI3	5.576	1.252	0.822	18.436			
	SI4	5.695	1.242	0.765	11.009			
Use intentions	UI1	5.769	1.387	0.965	133.843	0.9630	0.9760	0.9312
	UI2	5.665	1.415	0.974	136.402			
	UI3	5.665	1.473	0.965	81.135			

The results of the assessment of the indicator loadings on their corresponding constructs indicated that manifest variable correlations with their theoretically assigned latent variables are an order of magnitude larger than other loadings to other constructs, meeting the first criteria of discriminant validity. Further, the square roots of AVE were significantly higher than the correlations between the constructs, thus confirming that they are sufficiently discriminable (see Table 3) (Henseler et al., 2009).

Table 3: Correlations between the latent variables and square roots of the average variance extracted

	Use intentions	Performance perceptions	Result demonstrability	Social influence	Compatibility
Use intentions	**0.9650**				
Performance perceptions	0.4141	**0.8818**			
Result demonstrability	0.4549	0.5717	**0.8811**		
Social influence	0.4021	0.4375	0.3632	**0.8118**	
Compatibility	0.3988	0.3468	0.4333	0.2152	**0.8903**

We further tested the significance of the hypothesized relationships between the constructs by bootstrapping with 1,000 replicates. The structural model was then assessed by examining the coefficient of determination (R^2) of the endogenous latent variable, the estimates for the path coefficients of relationships in the structural model and their significance levels (via

bootstrapping) (Chin, 1998). The influence of performance perceptions, compatibility, social influence, and result demonstrability explain 32.3 % of the variance in use intentions. The influence of performance perceptions explains 12.0% of the variance in compatibility. Since the exogenous variables explain a moderate to high proportion of the variance of the endogenous variable, we may conclude that the model holds sufficient explanatory power and is capable of explaining the constructed endogenous latent variable (Henseler & Fassott, 2010).

The direct impact of performance perceptions on use intentions is not statistically significant. Therefore we do not support our H1. The path coefficients associated with H2 and H3 are statistically significant at the 1% significance level, thus supporting these two hypotheses. As indicated by the path loadings, socio-organizational considerations, including result demonstrability (H3: $\hat{\beta}$=0.210; p<0.01) and social influence (H3: $\hat{\beta}$=0.227; p<0.01) have a significant direct and positive influence on use intentions. The direct positive impact of compatibility on use intentions is statistically significant ($\hat{\beta}$=0.218; p<0.01) supporting H4. H5 is confirmed as the path is statistically significant at 0.1% significance level ($\hat{\beta}$=0.347), where performance perceptions have direct positive influence on compatibility. Besides testing statistical significance of the impact of performance perceptions on compatibility via bootstrapping for the proof of mediation effect, we also tested the mediation by Sobel's test, where the results were statistically significant at 5% statistical significance proving that compatibility does fully mediate the effect between performance perceptions and usage intentions (Rucker et al., 2011) confirming H6. We have compared the variance explained without performance perceptions construct (R^2=31.5%) and the one with it and the additional explanatory power of indirect effect of performance perceptions on use intentions is low (R^2 changes only by 0.7%).

The moderating effect of result demonstrability on the relationship between compatibility and use intentions is significant and positive ($\hat{\beta}$=0.262; p<0.01). The size of the moderating effect is f=0.06. Also the interaction effect of social influence on the relationship between compatibility and use intentions is significant and positive ($\hat{\beta}$=0.193; p<0.05). The size of the moderating effect is f=0.04. The variance explained in use intentions increases to 36.1% with the interaction effects, showing a meaningful interaction effects (Chin et al., 2003).

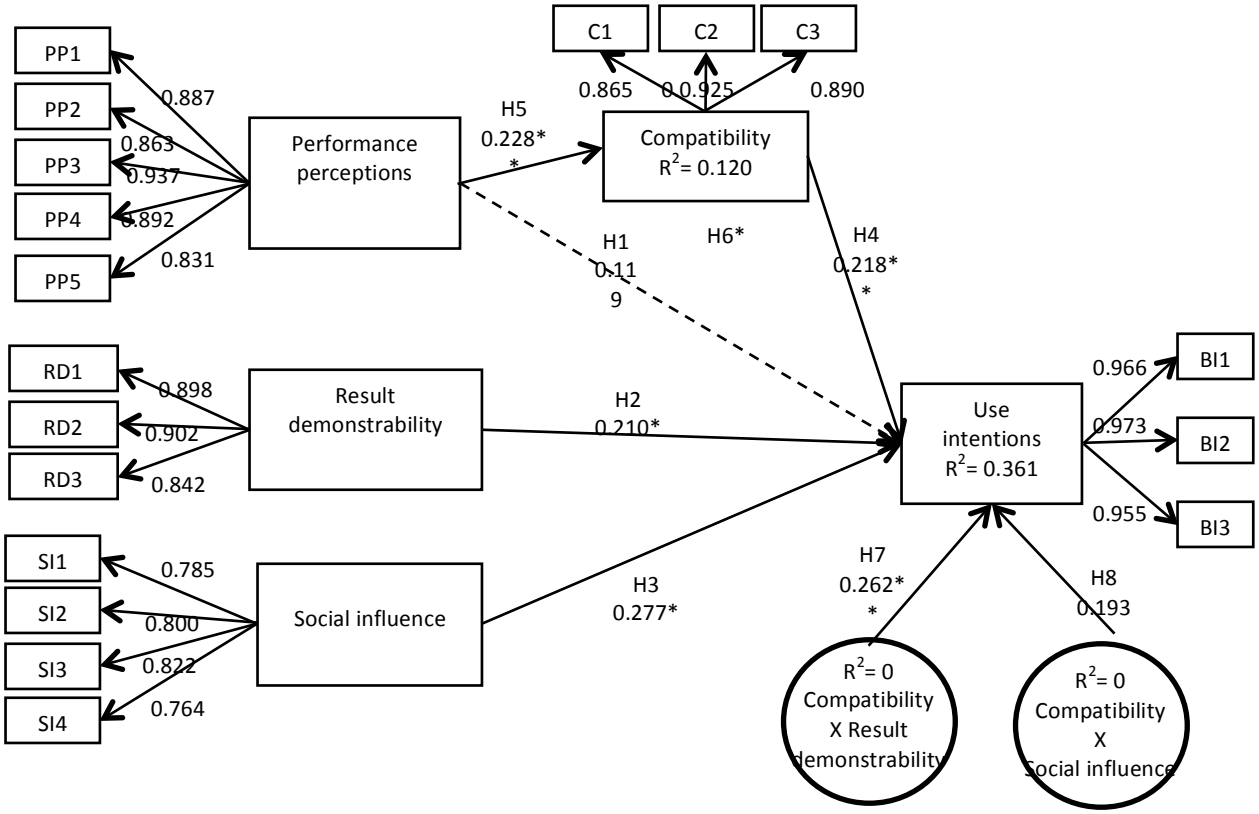


Figure 2: The research model and its results

Discussion with Implications for Research and Practice

The findings of our study provide some interesting new insights of the interrelationships between determinants of BI&A use intentions as explanatory and predictive variables. We confirm the established relationships in previous empirical study in the context of BI&A utilization (Grublješič et al., 2014) and enrich the explanatory power of BI&A use intentions with the inclusion of compatibility determinant and interrelationships between compatibility and the previously researched predictors in this context.

The analysis of the demographic profiles of our respondents reveals, that the majority of the respondents are male (61.14 %) and the average age of the respondents is 44.66 years. We also see that most of BI&A users are higher educated, as namely 89.17 % have professionally oriented higher education and higher, 73.53 % having B.Sc. or higher education, confirming that BI&A user are usually higher educated and more skilled workers (Luo, 2016). Respondents also have on average 2.75 years of experience with BI&A use. All these specific demographic

characteristics contribute to the postulation that the focus of significant drivers of BI&A use intentions might be different than traditionally established ones.

We provide evidence that performance perceptions, the strongest direct predictor of intentions throughout the past decades (Venkatesh et al., 2003), only have an indirect impact on use intentions mediated by compatibility. We therefore cannot neglect the consistently proven strongest predictor in explaining intentions, but it has a weak added explanatory power in predicting use intentions (only 0.7 %) in the contemporary BI&A use environment. Performance perceptions do have a meaningful effect on compatibility perceptions, thus, if users find BI&A useful their personal operational task-technology fit perceptions increase. Research on gender differences shows that men tend to be highly task oriented (in Venkatesh et al., 2003). Thus, performance perceptions which focus on task accomplishment are reasoned to be especially salient for men (Venkatesh et al., 2003). Given the fact that the majority of the respondents in our study are men, but also irrespective of it in the context of task-technology fit perceptions of compatibility, this could reason the fact that performance perceptions powerfully impact compatibility perceptions that are oriented on individual's personal efficiency increase in task accomplishment. On the other point, since even the mediation effect of performance perceptions through compatibility on use intentions is very small, this finding again underscores the idea that the focus of motivations to use BI&A has shifted from individualistic utility or operational gains to socio-organizational considerations.

We further corroborate this importance and significance of socio-organizational considerations in predicting BI&A use intentions. The emphasis goes in line with the encouragement of Junglas et al. (2013) that "future IS research should consider the inclusion of a social component into its utilization and acceptance models". Including the compatibility determinant advanced the variance explained in BI&A use intentions by direct significant positive impact of compatibility on use intentions. The compatibility perceptions of BI&A use include the fit to individual's work style and the use of the system in the organization. Since BI&A implementation significantly changes the organizational work environment, i.e. carrying out business processes (Deng & Chi, 2013) users need to adapt their work styles to these external stimuli. This includes changes in work environment by engaging in adaptation cycles to achieve a better fit between the system and the new context accustoming new tasks and carrying out these tasks (Sun, 2012). With the data in BI&A systems being more aggregated and integrated at the level of the entire organization, there is greater sharing of information connected to their use (Bose, 2009). The compatibility perceptions thus also include other's use, and managers' or organizational promotional incentives and support (Sun, 2012). Compatibility perceptions driving BI&A use intentions are therefore in big part promoted by socio-organizational motivation.

The third part of the discussion underscoring that the shift is underway to the socio-organizational considerations in the context of BI&A use intentions relates to the significant direct and moderating impacts of result demonstrability and social influence perceptions. In addition to having strong direct positive impact on use intentions and high explanatory power, result demonstrability and social influence also moderate the relationship between compatibility

and use intentions and by that strengthen the perceived relevance of compatibility in impacting use intentions. These interaction effects are statistically significant and positive. Thus, if the results of BI&A use are demonstrable, communicable, visible, supported by peers, management and organization in general, the higher the individual's perceptions of compatibility of BI&A use with their work style are. This "fit" is predominantly explained by the triggers from the socio-organizational work environment. These findings go in line with May & Finch (2009) reasoning that undoubtedly individuals do have preferences they act upon, but there are always social factors that promote or constrain a particular behavior. Cooper and Zmud (1990) further expound that adoption is better explained by rational task-technology fit, and later implementation stages are better explained by more socio-political and learning approaches (in Karahanna et al., 1999).

The findings are of value from theoretical and practical point of view for several reasons. From a theoretical perspective, it is beneficial to verify and find a specific interrelated set of influential perceived characteristics in a contemporary BI&A use environment as to provide a more parsimonious model. Further, if the available BI&A systems are not used appropriately this is of little value to the organizations as anticipated productivity gains cannot be realized. Hence, examining the importance of these perceptions is pragmatically important, since "they can be proactively influenced by management action" (Agarwal & Prasad, 1997, p. 559).

Conclusion

Our findings provide an enriched understanding of the interrelationships between the dimensions of BI&A use intentions as well as better explanatory power of these use intentions. Specifically, we find that performance perceptions, consistently proven to be the most rigorous and powerful direct predictor of use intentions, in the BI&A context only have an indirect impact on use intentions mediated through compatibility. Mediation effect is statistically significant but has a weak added explanatory power in predicting use intentions. Further, we advance the variance explained in BI&A use intentions by providing evidence about the direct significant positive impact of compatibility. Moreover, result demonstrability and social influence, in addition to having significant direct impact on use intentions with high explanatory power of use intentions; these also positively strengthen the relationship between compatibility and use intentions. By all these findings we, further corroborate the predominant significance and importance of socio-organizational drivers in predicting BI&A use intentions. If we use only traditional well-researched technology acceptance models (Davis, 1989; Venkatesh et al., 2003) in the context of BI&A use intentions, we might overlook important contributors and consequently focus on the inaccurate set of factors that do not lead to their effective utilization. Our results thus provide important basis and input for future studies of successful acceptance and further long-term use of BI&A.

Our research has some limitations that should provide basis for future research. In our context theorizing, we combined the relevant core constructs from general technology adoption theories (Davis, 1989; Moore & Benbasat, 1991; Venkatesh et al., 2003) based on the profiles and

demographic analysis of our respondents. Here we omitted some core constructs that might be important, such as voluntariness of use, trialability, and experience. Future research on different survey respondents could capture these beliefs.

References

Agarwal, R., & Prasad, J. (1997). The role of innovation characteristics and perceived voluntariness in the acceptance of information technologies. *Decision Sciences, 28*(3), 557-582.

Ajzen, I. & Fischbein, M. (2005). The influence of attitudes on behavior. In D. Albarracin, B. T. Johnson, M. P. Zanna, *Handbook of Attitudes and Attitude Change: Basic Principles*, Eds., Erlbaum, Mahwah, NJ.

Audzeyeva, A. & Hudson, R. (2015). How to get the most from a business intelligence application during the post implementation phase & quest: Deep structure transformation at a UK retail bank. *European Journal of Information Systems, 1*(18), 25-29.

Bose, R. (2009). Advanced analytics: Opportunities and challenges. *Industrial Management & Data Systems, 109*(2), 155 – 172.

Chin, W. W. (1998). Issues and opinions on structure equation modeling. *MIS Quarterly*, *22*(1), vii-xvi.

Chen, X., & Siau, K. (2011). Impact of business intelligence and IT infrastructure flexibility on competitive performance: An Organizational Agility Perspective. *Proceedings of the International Conference on Information Systems (ICIS)*, Shanghai, China.

Cooper, D. R., & Schindler, P. S. (2003). *Business research methods* (8th ed.). McGraw-Hill/Irwin.

Cooper, R. B., & Zmud. R. W. (1990). Information technology implementation research: A technological diffusion approach. *Management Science, 36*(2), 123-139.

Davenport, T. H., Barth, P., & Bean, R. (2012). How big data is different. *MIT Sloan Management Review, 54*(1), 43-46.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319–339.

Deng, X. N. & Chi, L. (2013). Understanding post-adoptive behaviors in information systems use: A longitudinal analysis of system use problems in the business intelligence context. *Journal of Management Information Systems, 29*(3), 291-325.

Gartner Summits. (2016). Gartner Business Intelligence & Analytics Summit 2016, October 10-11 2016, Munich, Germany, Available at http://www.gartner.com/events/emea/business-intelligence-de.

Grublješič, T., Coelho, P. S., & Jaklič, J. (2014). The importance and impact of determinants influencing business intelligence systems embeddedness. *Issues in Information Systems, 15*(1), 106-117.

Grublješič, T., & Jaklič, J. (2015). Business intelligence acceptance: The prominence of organizational factors. *Information Systems Management, 32*(4), 299-315.

Karahanna, E., Straub, D. W., & Chervany, N. L. (1999). Information technology adoption across time: A cross-sectional comparison of pre-adoption and post-adoption beliefs. *MIS Quarterly, 23*(2), 189–213.

Henseler, J., Ringle, C. M., & Sinkovics, R. R. (2009). The use of partial least squares path modeling in international marketing. *Advances in International Marketing, 20*, 227-319.

Henseler, J., & Fassott, G. (2010). Testing moderating effects in PLS path models: An illustration of available procedures. In *V. Esposito Vinzi et al. (eds.), Handbook of Partial Least Squares, Springer Handbooks of Computational Statistics*, Berlin Heidelberg: Springer-Verlag.

Hong, W., Chan, F. K. Y., Thong, J. Y. L., Chasalow, L. C., & Dhillon, G. (2013). A framework and guidelines for context-specific theorizing in information systems research. *Information Systems Research*, 25(1), 111–136.

Li, X., Po-An Hsieh, J. J., & Rai, A. (2013). Motivational differences across post-acceptance information system usage behaviors: An investigation in the business intelligence systems context. *Information Systems Research, 24*(3), 659-682.

Luo, W. (2016). Responsibility and skills requirements for entry level analytics professionals. *Journal of Organizational and End User Computing, 28*(4), 1-14.

Mao, E., & Palvia, P. (2006). Testing an extended model of IT acceptance in the Chinese cultural context. *Data Base, 37*(2&3), 20-32.

May, C., & Finch, T. (2009). Implementing, embedding, and integrating practices: An outline of normalization process theory. *Sociology, 43*(3), 535-554.

Moore, G. C., & Benbasat, I. (1991). Development of an instrument to measure the perceptions of adoption and information technology innovation. *Information Systems Research, 2*(3), 192-222.

Petter, S., DeLone, W. H., & McLean, E. R. (2013). Information systems success: The quest for the independent variables. *Journal of Management Information Systems, 29*(4), 7–62.

Popovič, A., Coelho, P. S., & Jaklič, J. (2009). The impact of business intelligence system maturity on information quality. *Information Research, 14*(4).

Ringle, C. M., Wende, S., & Will, S. (2005). *SmartPLS 2.0 (M3) Beta*. Hamburg, URL: http://www.smartpls.de.

Rogers, E. M. (1983). *Diffusion of innovations* (3rd ed.). New York, NY: Free Press.

Rucker, D. D., Preacher, K. J., Tormala, Z. L., & Petty, R. E. (2011). Mediation analysis in social psychology: Current practices and recommendations. *Social and Personality Psychology Compass, 5*(6), 359-371.

Shin, D.-H. (2015). Demystifying big data: Anatomy of big data developmental process. *Telecommunications Policy, 40*(9), 837–854.

Sobel, M. E. (1982). Confidence intervals for indirect effects in Structural equations models. *Sociological Methodology, 13*(1982), 290-312.

Stern, B. B., Royne, M. B., Stafford, T. F., & Bienstock, C. C. (2008). Consumer acceptance of online auctions: An extension and revision of the TAM. *Psychology & Marketing, 25*(7), 619-633.

Sun, H. (2012). Understanding user revisions when using information system features: Adaptive system use and triggers, *MIS Quarterly, 36*(2), 453-478.

Venkatesh, V., & Davis, F. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science*, *46*(2), 186-204.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Towards a unified view. *MIS Quarterly*, *27*(3), 425-478.

Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. *Decision Sciences, 39*(2), 273-315.

Venkatesh, V., Thong, J. Y. L., & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Quarterly, 36*(1), 157-178.

Wixom, H. B. & Todd, P. A. (2005). A theoretical integration of user satisfaction and technology acceptance. *Information Systems Research, 16*(1), 85-102.

Yoon, T. E., Ghosh, B., & Jeong, B.-K. (2014). User acceptance of business intelligence (BI) application: Technology, individual difference, social influence, and situational constraints. *Proceedings of the 47th Hawaii International Conference on Systems Sciences*, pp. 3758-3766.

Author's Biography

Tanja Grublješič, Ph.D. is a Research and Teaching Fellow at the Faculty of Economics of the University of Ljubljana. She holds a B.Sc. degree in the field of Management and Organization, an M.Sc. in International Economics and a PhD in Information Management from the Faculty of Economics, University of Ljubljana. Her research and teaching interests include the topics of the adoption, acceptance, embeddedness, use, and success of Business Intelligence Systems in organizations.