SEEKING OUT CLEAR AND UNIQUE INFORMATION SYSTEMS CONCEPTS: A NATURAL LANGUAGE PROCESSING APPROACH

Research paper

Hassan, Nik Rushdi, University of Minnesota Duluth, USA, nhassan@d.umn.edu

Prester, Julian, UNSW Business School, Sydney, Australia, j.prester@unsw.edu.au

Wagner, Gerit, HEC Montreal, Montreal, Canada, gerit.wagner@hec.ca

Abstract

The goal of this study is to advance conceptual development and the growth of knowledge in the information systems (IS) field by placing the spotlight on a component of theory that is rarely discussed – the native IS concept. Beginning with the assertion that concepts are not the same as constructs, we build the argument that concepts, which are observable sets of ideas, should take priority over constructs which are unobservable fictions and hypothetical entities. Using natural language processing (NLP) based principles and techniques, we extract a sample of the most important concepts in the IS field from a corpus of 245 highly cited IS review articles and 1,293 citing articles from the Senior Scholars' Basket of Journals to illustrate the extent to which the field agrees on their usage, their clarity and distinctiveness and how the field can move forward in enhancing its conceptual formation.

Keywords: IS theory, native IS concepts, natural language processing, scientometrics, conceptual formation

1 Introduction

> What surprises me is how unusually mute our discipline seems to be on so important a subject. While management journals devote some space to discussions of constructs and their role in theory development … the coverage seems disproportionate to the importance of the topic. This anomaly is reflected, somewhat, in how we train graduate students, where considerable time is devoted to understanding how constructs are measured and operationalized but substantially less time is devoted to understanding how constructs are created and used in the research process.
>
> [Suddaby, (2010) "Editor's Comments: Construct Clarity in Theories of Management and Organization"]
>
> … management research has failed to cultivate any truly indigenous theories of management and organization … (i.e., roughly two-thirds) of the research conducted in management is rooted in theories borrowed from other disciplines. [Suddaby et al. 2011 "Where Are the New Theories of Organization?"]

These two quotes from the *Academy of Management Review* journal could have been about the information systems (IS) field by just replacing "management" with "IS" and they would still ring true. Because the IS field borrows constructs heavily from various other disciplines including management and adapts them into our research frameworks, we would be at a severe disadvantage if those borrowed constructs are, as Suddaby (2010) claims, not well constructed. It is therefore no surprise that historically, the IS field has struggled making our research relevant to practice, for "the solution to critical problems, with later transfer to practice, does not originate or occur within the academic com-

munity" (Senn, 1998, p. 23). In other words, because we rely on those borrowed constructs, the IS field effectively becomes second-level beneficiary of hand-me-down research that, if the management discipline is any indication (Suddaby, Hardy and Huy, 2011), two-thirds of which is rooted in other disciplines.

The management and IS fields are not alone in struggling with concepts. IT-related fields such as software engineering and healthcare informatics struggle to define their roles and their core concerns. Mary Shaw (1990, p. 15), one of the pioneers in the software engineering field argued that "the word 'engineering' to describe this activity [software engineering] takes considerable liberty with the common use of that term." Two decades later, she argues that although there has been progress, the software engineering field needs to measure this progress by recognizing "which class a given problem falls" (Shaw, 2009, p. 84) so the field can offer solutions to those problems. Such a goal can only be accomplished if the field can agree on its concepts. Similarly, the healthcare informatics debate the conceptual bases for different "informatics" fields such as medical informatics, nursing informatics and dental informatics (Staggers and Thompson, 2002; Wyatt and Liu, 2002).

The goal of this paper is to advance the discussion about this lack of focus on our own field's concepts – native IS concepts – which are critical components of theory (Bacharach, 1989), and how they impact the growth of knowledge in the field. The field's concepts are not to be confused with the field's "constructs," which will be discussed in the next section. Sociologists have always maintained that the difference between concepts and constructs is absolute (Willer and Webster Jr., 1970; Merton, 1973). Many IS authors have expressed the same concerns about IS concepts (Burton-Jones and Straub, 2006; Markus and Saunders, 2007; Furneaux and Wade, 2009; Hassan, 2011) but there have not been many efforts to broadly analyse and remedy the problem. In the history of disciplines, the study of concepts has always taken a front row seat (Schön, 1963) and those that do not prioritize its study of concepts will have its progress held back. Herbert Blumer (1954), one of the earliest exponents of interpretivism, blames the deficiencies in social theory on the vagueness of concepts. Concepts are the means by which theory is fruitfully linked to the empirical world and it is the concept that points to the empirical instances about which theories are constructed. If the concept is unclear, then references to the empirical instance are weakened and any theory that depends on the concept cannot be brought close to the empirical world it is attempting to explain. As Nobel laureate Sir George Thomson (1961, p. 4) noted:

> Science depends on its concepts. These are the ideas that receive names. They determine the questions one asks, and the answer one gets. They are more fundamental than the theories which are stated in terms of them.

Concepts, because they are subject to the rules of the discipline that produces them, demarcate that discipline's subject matter and declares to the world what the discipline represents. For example, there is broad agreement on the meaning of respiration and circulation, or relativity and quantization, and no one doubts that these concepts belong to the disciplines of biology and physics, respectively. The question is: What concepts belong to the IS field (Markus and Saunders, 2007; Hassan, 2011) and if we do have native IS concepts, are their meanings clear and distinctive? This effort is not to be confused with drawing up frameworks for IS research (e.g. Mason and Mitroff (1973) or Ives et al., (1980) or other classic IS research frameworks) or keyword classifications schemes (Barki and Rivard, 1988; Barki, Rivard and Talbot, 1993) which are also important. It has to do with identifying unique key concepts in the field so their role in IS theories can be made clear, because concepts are the building blocks of theories (Bacharach, 1989; Hassan, 2014). The next section argues why clear and unique IS concepts are important? The following three sections use a natural language processing (NLP) approach to analyse a large corpus of significant IS research to extract the most important concepts. Finally, the remaining sections discuss and compare the results with similar previous studies and conclude with how the IS field can move forward with these newly found insights.

2 Why Are Clear and Unique IS Concepts Important?

This study undertakes the journey to identify and clarify those IS concepts so that internally we are clear about our core concerns and, externally, others are clear about our subject matter. If we do not seriously address this issue, we will not only continue counterproductive practices of reproducing abstruse research that borrow heavily from other disciplines, our idiosyncratic categories will remain less relevant to our stakeholders.

2.1 Constructs before concepts: Putting the cart before the horse

One reason why the IS field and its allied management fields do not place much emphasis on concepts is because we tend to use constructs in research rather than concepts. Constructs are not the same as concepts and the difference between the two is central to understanding why the IS field lacks focus on concepts. First, concepts are more general than constructs. In psychology, constructs are defined as "postulated attribute of people, assumed to be reflected in test performance" (Cronbach and Meehl, 1955, p. 283) or abstractions used to explain more observable phenomenon (Morgeson and Hofmann, 1999). In sociology, constructs could refer to abstract properties of particular persons, places or events (Willer and Webster Jr., 1970). Although both concepts and constructs are abstractions, concepts are not just about people, places or events. It could be anything that is observable that can be defined. “Fake news” is a very broad concept, and may include other concepts such as satire, parody, fabrication, photo manipulation, and propaganda (Tandoc, Lim and Ling, 2018). The more specific measures that need to be invented from all these concepts are called constructs. What really distinguishes constructs from concepts is that constructs are always explicitly or implicitly defined in the propositions of the theory it is associated with, whereas concepts are not.

In fact, the quotation at the beginning of the article specifically refers to “Construct Clarity in Theories of Management and Organization” rather than “Concept Clarity.” The confusion surrounding the differences between constructs and concepts exists even among trained scholars. As measures, constructs help researchers make sense of such observations by acting as heuristic devices, and together with other observables and constructs, form what is known as the "nomological network," defined as the "the interlocking system of laws which constitute a theory" (Cronbach and Meehl, 1955, p. 290). Concepts, on the other hand, are defined as observable sets of ideas associated with or elicited by given words that are treated according to logical rules (Sartori, 1975) and need not be associated with any theory necessarily.

It should be intuitively clear that before any unobservable, abstract construct can be constructed, the underlying concepts should be made clear. Constructs are in essence “fictions and hypothetical entities” (MacCorquodale and Meehl, 1948, p. 95) contrived from concepts to enable the use of some form of measurement or evaluation or to bridge several concepts in a study. It almost seems like management scholars have put the cart before the horse in addressing constructs before their concepts are made clear. Concepts flow between disciplines freely and are often borrowed to explain phenomenon in the borrowing discipline by using the familiarity afforded by the more popular concept in the lending discipline. Such a practice of “displacement of concepts” (Schön, 1963) or “traveling concepts” (Bal, 2002) has been the source of many classic studies and the growth in knowledge. For example, Herbert Spencer (1897), responsible for popularizing Darwin’s notion of evolution into other areas of knowledge, borrowed the concepts of “organism” from biology (as in “organic structure of society”) and “statics” from physics (as in “social statics”) to explain society. Many years later, management scholars (Burns and Stalker, 1961) would use the same basic understanding of those concepts to write about how more organic organizational structures lead towards innovative practices compared to mechanistic structures. Within each discipline, the concept takes on a different meaning and that discipline owns that concepts.

Another reason why the underlying concept should be the primary focus in our research is indicated by the process by which constructs are built. An important part of building constructs is ensuring they have validity. Construct validity—how closely a construct is interpreted as a measure of some qualita-

tive attribute which cannot be operationally defined – relies on a clear understanding of the concept. In IS, where many of its nomological networks are borrowed either from management or from social psychology (Benbasat and Barki, 2007), construct validity becomes even more critical because as the constructs are applied in a different context, the associated measures may not transfer. The notion of construct validity thus suggests a need to clearly define concepts.

Unfortunately in IS, constructs are often conflated with concepts and most researchers do not distinguish between them, and therefore the more important item, the one that is less abstract and is directly observed in the phenomenon of interest – the concept – is ignored. Furneaux and Wade (2009) identified 690 distinct "constructs" that were researched in *MIS Quarterly (MISQ)* and *Information Systems Research (ISR)* between 1999 and 2007. The results show a preponderance of research undertaken that are related to the research areas of technology acceptance, adoption, trust and organizational performance. Similar to Furneaux and Wade (2009), Shuraida et al., (2018) applied several classic frameworks and classifications as their bases to manually code 1,361 empirical studies from four IS journals: *Journal of the Association for Information Systems (JAIS), Journal of Management Information Systems (JMIS), ISR,* and *MISQ*, and extracted 64 different relationships amongst their constructs. Studies on these concepts of interest provide insights into not only the nature of the IS field, but also its progress in relation to other disciplines, its identity and its future direction. An earlier effort to define IS concepts was by IFIP in their series of FRISCO reports (Falkenberg, Lyytinen and Verrijn-Stuart, 2000; Stamper, 2000; Verrijn-Stuart, 2001), which, in part because of its formalisms, did not capture the imagination of the IS community.

2.2 The Lack of Clear Concepts

Why is there a lack of interest in something so fundamental to research? It is not surprising that the IS field rarely investigates the constructs and concepts it is researching because historically the vast majority of its research is positivistic in nature (Orlikowski and Baroudi, 1991; Chen and Hirschheim, 2004). Positivistic or quantitative-type research rely more on their data, measurement and statistical models to derive conclusions. Accordingly, they focus attention on the nature of the data and its quantitative measures. They spend less time on the concept and focus more on operationalizing and analysing datasets (Goertz and Mahoney, 2012). Arguing for conceptual clarity, Sartori et al., (1975) list the following impairments for fields of study that choose to ignore it:

2.2.1 Loss of Etymological Anchorage

The semantics of scientific terms are anchored in their Latin and Greek roots, which help to stabilize and at the same time clarify their conceptual foundations. Unfortunately, much of the modern terminology, especially those applied in IS are not anchored in any etymological roots. As a result, unlike "biology" which is a contraction of bios, "life," and logia, "discourse," to mean the "science of life," new terms created cannot be traced back to any tradition and are open to any interpretation. Instead what the IS field does is to merely attach the name of the field (“IS”) or the technology (“IT”) to a concept, like “IS success,” “IS infrastructure,” “IS Strategy,” “IT alignment” or “IT innovation.” As a result, the IS field is forever tied to other disciplines that have defined “success,” “infrastructure,” “strategy,” “alignment” or “innovation.” This practice is common in the IS field and can be seen in its keyword classification schemes (Barki and Rivard, 1988; Barki et al., 1993) where a large proportion of its terms consists of these prefixes.

2.2.2 Loss of the Extensions in Discourse

When concepts and words are borrowed from other disciplines, often uncritically or speculatively, at some distance from the discourse from which they came, they add little to the creation of new concepts. So, what the term “information security” adds to the well-known concept of security, defined as “the state of being free from danger or threat,” is unclear (Anderson, 2003). Schön (1963) suggests that the production of new concepts is closely related to understanding how to work with metaphors and analogies. He noted that “the new concept grows out of the making, elaboration, and correction of

the metaphor" (p. 53). He calls this process the displacement of concepts, where words and terms from previous research undergo transposition (i.e., applying an old concept to a new situation), interpretation (i.e., assigning that concept to a specific aspect of the new situation), correction (i.e., an adjustment resulting from adaptation and modification), and spelling out (i.e., resolving commonalities and differences) as a way of addressing problems or improving understanding. Thus, security as it relates to IT needs to be transposed, reinterpreted, corrected or spelled out within the IS field. Presumably, information security would not be same as computer security. This same approach should be applied to other common terms like "organization," "information," or "systems." They all have different meanings in IS from the meanings in the disciplines they are originally defined.

2.2.3 Loss of Historical Anchorage

The technological revolution, in many ways, has also severed the relationship between the terms used to describe a technological phenomenon from historical experience. Many concepts in the natural and social sciences such as DNA, "protein synthesis," political "authority" and "democracy" acquire their meaning from a long history of learning, observation or experimentation. The rapid change in information technology has somehow caused these terms to lose their connection with the past and even if the past is inferred, words such as "Web 2.0," are misused, causing the field to endlessly arbitrarily add subsequent versions to the term such as Web 3.0, Web 4.0 and so on. As Schön (1963) explains, concepts can be extended and translated into different meanings as long as their relationship with its history is maintained. These extensions contribute to the growth of knowledge within the field. What do not contribute to knowledge are arbitrary terms that cannot be linked back to any history of research and serve only to confuse.

2.2.4 Novitism and the Chaos of Jargon

Scholars of the past were very careful of creating terms and of being perceived as being novel at whatever costs because they understood their role in transmitting knowledge. Perhaps these days, the easiest way to innovate is to "play with musical chairs with words" (Sartori et al., 1975, p. 9) resulting in one appearing more creative than others, and the process continues as another string of words are written to contradict the former, only to the detriment of true understanding. There is, as Sartori et al., (1975) notes, a major difference between the dynamics of language and the dynamics of science. Just like a card game, the game can only be played because the cards and the rules of their combination are stable. Similarly, only a disciplined use of concepts and terms permit the scientist to play the game. By contrast, if we invest more and more time in altering the cards, it is not the field that is being furthered, it is confusion. This state of confusion will lend itself when the field is subject to the chaotic changing of duplicative jargon produced weekly by industry.

2.3 The Need for Unique IS Concepts

Another side of the coin of conceptual clarity has to do with the relevance of the concepts. If the field is to embrace relevance to practice as its underlying philosophy, the IS field needs to approach research by first examining its concepts of interest. If the IS field considers the development of native theories as an important goal, unique constructs therefore will be critical since theories are essentially a system of concepts (Bacharach, 1989). If the theories are to be considered native, at least some of those concepts will also need to be native. To be relevant, these concepts should be those in which society is interested in and their understanding contributes to the individual, business or society as a whole. Ideally these concepts should not be the same concepts that are already the purview of other fields, or at the very least, not viewed in IS in the same way as it is viewed in other fields. For instance, borrowing concepts from social psychology without any extension or expansion in meaning, which was the case with the acceptance and adoption area of IS research, only serves to restrict IS research to within the discipline of social psychology. Both the aspects of clarity and distinctiveness of concepts in concept formation are what we seek to explore in our study.

3 Research Method: Finding Clear and Unique IS Concepts

Now that the importance of clear and unique IS concepts is understood, how can the IS field identify which concepts are the most important to the field? These important concepts can then be evaluated to see if they represent native IS concepts, that is, concepts that are developed within IS research and not borrowed uncritically from other disciplines (Markus and Saunders, 2007). There are several methods how concepts are identified. We propose instead to use scientometrics to extract the most important concepts from the IS field because not only does it provide a means for evaluating the importance of the concepts, it also provides an objective way of identifying the important terms from a corpus of work or from an entire field of study especially how those terms emerged from previous studies.

Scientometric scholars distinguish different roles citations play in the growth of knowledge. One of those roles is the use of citations to associate ideas from the cited works with the current study. Small (1978) suggests that by analysing the citations and the context surrounding the citations, it is possible to extract the ideas behind the citation, which represent the ideas and concepts the author is discussing. The cited works represent symbols for those ideas and by embedding those symbols in the citing study, the citing author gives further meanings to the current research and at the same time, imparts meanings to the cited works. Some of those meanings are "standard" and have achieved consensus within the research community, and these are the meanings we are interested in discovering for the IS field. An application of Small's proposal is co-citation analysis (Small, 1973) which has been applied in the IS field (e.g. see (Culnan and Swanson, 1986; Culnan, 1987)). By identifying shared references – serving as a heuristic for the influence of the same key concepts – between two or more citing works, co-citation analysis is able to map the intellectual structure of a field or area of research. What co-citation analysis lacks is the context that surrounds in-text citations, which contain important key terms, representing the meanings which the authors imparted on the cited documents and its concepts. An extension of Small's (1978) original method called citation context analysis (Small and Greenlee, 1980) solves this problem. By including the words around the citation in the analysis, it is possible to distinguish "standard symbols" that are shared by most of the researchers in the community as represented by the most-cited documents, from less standard concepts. Researchers carry a repertoire of these standard symbols with them as they conduct research because they are the researchers' tools-of-the-trade and provide the underlying conceptual and methodological framework for their research. By collecting the most-cited documents in the field and extracting the terms around the citations they use, it is possible to gather the most important concepts and the meanings that have been imparted on them by the field. To extract these terms, we will use natural language processing (NLP) techniques.

We start with a corpus of IS studies consisting of 245 IS research review articles published between 2000-2014 collected from 40 IS journals identified by Lowry et al (2013). We chose review articles because they (1) aggregate the key theories, concepts and ideas of a discipline, (2) address the main questions and problems and summarize the major issues and debates, (3) synthesize the fragmented body of knowledge of a discipline into a coherent whole and (4) often make new contributions to an area of knowledge (Hart, 1998; Schryen, Wagner and Benlian, 2015). These benefits of review articles are acknowledged by IS scholars (Webster and Watson, 2002) and have produced several methodological guides for writing reviews (Rowe, 2014; Paré, Trudel, Jaana and Kitsiou, 2015). Literature reviews make new contribution to knowledge (Prester, Wagner and Schryen, 2018) since every research paper is required to offer a substantial review section as evidence that the study has indeed contributed knowledge. Therefore, using the corpus of IS review articles that cover a broad range of research topics should help us identify the most significant concepts and ideas of the IS field and how they are used in top-tier IS research.

We extracted 70,047 citations to the 245 IS review articles published between the years 2000 and 2017. For the content and NLP-based analyses, we restricted the set of citing papers to the AIS Senior Scholars' Basket of Journals (Association for Information Systems, 2011) resulting in 1,293 citing articles. This restriction to the AIS Senior Scholars' Basket of Journal was made to ensure that the extracted citations represented the IS field and that the homogeneity of this corpus allowed for sufficient accuracy of document structure annotation and citation sentence extraction. Metadata and PDFs of

citing articles were acquired from Google Scholar, Web of Science and other institutional repositories. Using the list of cited review articles, we examine where each was cited in the citing articles and collected the unigram, bigrams and trigrams around the in-text citations to identify the concept. For these analyses, we considered the citation context, which is commonly defined as the citing sentence including its predecessor and successor sentences (Small, 1978). Common NLP pre-processing procedures were applied, including the removal of stop-words, numbers, symbols and reference markers. The goal of the technique is to tell us what the citing authors thought of the papers when they cited them, or what the cited documents symbolize for the citing authors. The analysis shows a pattern of common usage emerging and we can estimate the uniformity of usage by calculating the number of equivalent usages divided by the total citations sampled. Small and Greenlee (1980) call this measure the "percent uniformity." The n-grams around the citations are compared with one another and the concept or idea expressed by the largest number of citation contexts is determined and each context is coded for its presence or absence. The ratio of the number of contexts expressing the most prevalent concept to the total citations examined is calculated.

4 Results

We provide an overview of the results we collected using several descriptive statistics involving the most cited review articles, the percentage of these most cited review articles compared to the total number of all citations in the corpus and other descriptive statistics that are reviewed below.

4.1 Descriptive Statistics

Figure 1 shows the 20 most cited IS review articles from the corpus. Except for the first five review articles, the rest of the top 50 review articles received between 341 to 994 citations which amounts to between 0.5% to 1.4% of total citations to the corpus. These percentages are consistent with the proportions that Small (1978) found in this study of the discipline of chemistry, indicating that the sample we have from the IS field is significant (i.e. only the top cited articles are selected). The top five review articles are especially significant because they give us an indication of the research theme that a large majority of IS community has preferred to invest their time in by researching, citing and following up on them (Hassan and Loebbecke, 2017). Alavi and Leidner (2001), the review of knowledge management received 13.8% of all the citations to the corpus or 9,679 citations followed by the update on IS success by DeLone and McLean (2003) with 11.7% or 8,222 citations, Legris et al (2003) with the review of TAM research with 4.9% or 3,423 citations, Melville et al (2004) with 4% or 2,853 citations on a review of business value of IT research, and Wade and Hulland (2004) on a review of resource-based view research with 3.2% or 2,260 citations.

To further emphasise how significant these cited review articles are, of the 245 review articles cited, 73.9% of them are considered citation classics (Price, 1965; Garfield, 1984), that is, articles that are cited at least 4 times a year within 10 years. These citation classics are research that finds their way into the canons of the discipline and represent evidence of cumulative tradition demonstrated by how members of the discipline jointly contribute to the subject matter they consider significant. These cited review articles have also captured the attention of the IS community over a long period of time and are not just "flash in the pan." The average age of the cited review articles that are cited 40 or more times in the past five years is 10.18 years compared to 11.02 years for those cited 40 or more times in the past 10 years. That means, all of the review articles have consistently been cited over time as major references for the IS community.

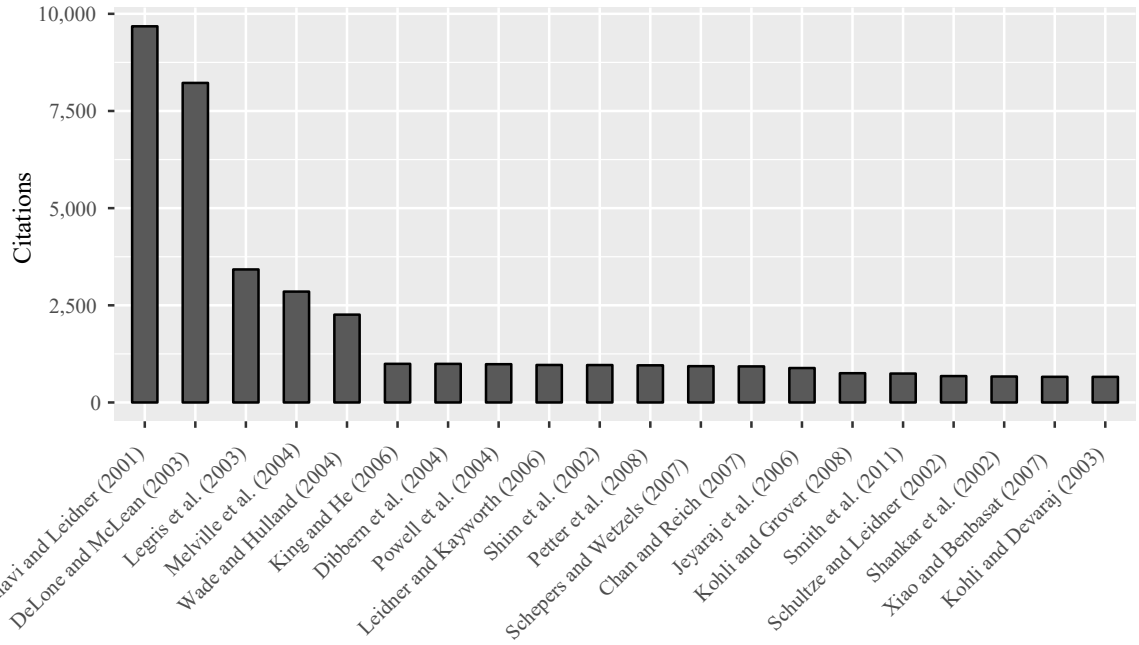

Figure 1: Most Cited IS Review Articles

Table 1 lists the most frequent unigrams and bigrams extracted from the context surrounding the citations to the IS review articles. Because the citations to the top five most-cited review articles made up nearly 38% of the total citations, it is not surprising that the concepts extracted from the corpus consist of concepts related to the five research themes of the those articles, namely: knowledge management, IS success model, TAM research, business value of IT and the resource-based view.

Unigram	Frequency	Bigram	Frequency
IS	1,657	IS Success	618
IT	1,461	Information Systems	488
Model	1,229	Systems success	314
Success	1,205	Success model	261
Research	1,205	Business value	182
Information	1,057	IS research	172
Systems	794	User satisfaction	164
Studies	751	Firm performance	143
Use	743	Competitive advantage	129
Literature	719	Information quality	120
Organizational	545	Information technology	115
Performance	541	System quality	108

Table 1: Most frequent unigrams and bigrams

4.2 Most Important and Unique IS Concepts

To overcome the bias caused by the top five review articles, we extracted the unigrams, bigrams and trigrams from the top 18 review articles in the corpus and listed some selected terms in Table 2. To illustrate the process, we expand on the concepts extracted from these review articles. Since the focus is on exploring the meaning imparted upon important concepts from a broad range of topics, we select other highly cited reviews regardless of topics. For the sake of brevity, Table 2 is shortened.

Exemplary review article	Unigrams (percent uniformity)	Bigrams (percent uniformity)	Trigrams (percent uniformity)	Key concept and variations	Definition
Alavi and Leidner (2001) Also: Schultze and Leidner (2002)	Knowledge (85%) Management (34%) Organizational (30%) Information (28%)	Knowledge management (28%) Knowledge creation (12%) Knowledge sharing (8%) Information systems (7%) Knowledge transfer (6%)	Knowledge management systems (5%) Tacit explicit knowledge (4%) Knowledge management (KM) (3%) Information systems (IS) (2%)	Knowledge management (KM) - KM process - Knowledge creation - Knowledge transfer - KM system	"Knowledge management refers to identifying and leveraging the collective knowledge in an organization to help the organization compete […]." (p. 113)
DeLone and McLean (2003) Also: Petter et al. (2008)	Success (72%) Model (64%) IS (46%) Information (44%) Systems (36%)	IS success (34%) Information systems (25%) Systems success (25%) Model information (19%) Success model (18%)	Information systems success (21%) Model information systems (18%) IS success model (13%)	Success - Success model - System success - IS Success	Technical (system quality), semantic (information quality) and effectiveness (individual and organizational impact) (p. 10)
Legris et al. (2003) Also: King and He (2006), Schepers and Wetzels (2007)	Technology (58%) Use (55%) IS (42%) Acceptance (35%)	Technology acceptance (29%) Intrinsic motivation (19%) Perceived usefulness (16%) IS research (16%) IS use (16%)	Perceived ease use (10%)	Adoption - IS use - Technology adoption - Technology usage - TAM research	No unique definition stated. "Davis [8] and Davis et al. [10] proposed TAM to address why users accept or reject information technology." (p. 192)
Melville et al. (2004). Also: Kohli and Devaraj (2003), Kohli and Grover (2008)	IT (74%) Business (50%) Value (42%) Performance (38%) Firm (35%)	Business value (27%) IT business (15%) Firm performance (14%) Organizational performance (13%) Competitive advantage (10%)	IT business value (12%) Business value IT (7%) Complementary organizational resources (4%)	Business value - IT impact - IT performance impact - Business value of IT - IS business value	"IT business value […][is] the organizational performance impact of information technology at both the intermediate process level and the organization wide level" (p. 287)
Wade and Hulland (2004)	IT (55%) Resources (47%) Firm (45%) Capabilities (41%) IS (36%)	Competitive advantage (20%) Firm performance (13%) Resource-based view (12%) IT resources (12%) IS resources (8%)	IS technical skills (4%) Resource-based view firm (3%) IS infrastructure (IS) (3%)	Technology - IT resources, assets - Information systems resource - Resource capability	"[IT resources are information technology] assets and capabilities that are available and useful in detecting and responding to market opportunities or threats." (p. 109)
Dibbern et al. (2004)	Outsourcing (61%) IT (33%) Literature (24%) Studies (23%)	IT outsourcing (17%) Transaction cost(s) (15%) Outsourcing literature (7%) IS outsourcing (7%)	Transaction cost economics (4%) Transaction cost theory (3%) IS outsourcing literature (3%) Cost theory agency (2%)	Outsourcing - IT outsourcing - Offshore outsourcing - Sourcing	"[IS sourcing refers to] the organizational arrangement instituted for obtaining IS services and the management of resources and activities required for producing these services." (p. 11)

Table 2: Table of concepts for the most impactful IS review articles

For each review, we report the top meaningful unigrams, bigrams and trigrams extracted from three sentences surrounding the in-text citation, along with the percent uniformity for each unigram. Some of these concepts are defined or described in other disciplines such as the well-known concepts "firm performance," "strategic alignment" and "competitive advantage." At the same time, many concepts are unique, that is, they were possibly transposed, reinterpreted, corrected or spelled out within the IS field. For example, "IT", "knowledge management," "decision support systems," and "technology acceptance" are well known within the IS research community and are not defined elsewhere to any theoretical depth. Their presence in this list in Table 2 and the relatively higher percent uniformity figures indicated within parentheses next to them suggest that they are important to IS researchers.

In the case of the natural or biological sciences, which Small (1978) and Small and Greenlee (1980) targeted, it is not uncommon to find percent uniformity figures as high as 90% or 100% for many concepts ("relativity" or "respiration" cannot be defined any other way) to represent agreement among its scholars. However, for the social or human sciences like IS, we can expect lower percent uniformity figures because unlike the natural sciences, the same concept is often represented by different terms. For example, the bigrams "IS success" and "systems success" refer to the same concept "success" when related to DeLone and McLean's (1992, 2003) works, so, if "IS success" has 20% uniformity and "systems success" has 30% uniformity, we can safely conclude that the combined 50% uniformity represents an agreement by IS scholars.

Based on the quantitative analyses of unigrams, bigrams and trigrams with the highest percent uniformity, we qualitatively extracted the key concept from the review article and its variations (5th column) cited by subsequent research. For example, from Alavi and Leidner (2001) and Schultze and Leidner (2002), we select the key concept symbolised by those articles (e.g. knowledge management) highlighted by its percent uniformity (Knowledge – 85%, Management – 34%, Knowledge Management - 28%). These concepts are "standard" concepts that IS researchers use in performing their research, and in extending and creating knowledge in the field. Concepts with lower percent uniformity are either variations of the same concepts or transpositions, reinterpretations, corrections or spelling out of the key concepts. Each cited review article may become the source of many different concepts depending on how the citing authors apply them in their writings.

To better see how clearly the key concepts are transposed, reinterpreted, corrected or spelled out and how clearly they are defined we omit such terms as "IS" or "IT" from the bigrams and trigrams like IS success, IT impact, IT resources, IT capability, IT culture and we find a definition for each key concept from the corpus to evaluate its distinctiveness. The results in the sixth column in Table 2 show that many key concepts do not have clear (e.g. Adoption, Culture, Alignment) or distinctive definitions (e.g., Alignment, Trust, Strategy). Other studies outside the corpus of review articles have found similar issues. For example, Burton-Jones and Straub (2006) found inconsistencies and weaknesses in the how the IS field defines the key concepts of adoption and use.

5 Discussion

The small sample of the most important IS concepts extracted from Table 2 are consolidated and summarized in Table 3 and compared with the closest constructs or research areas identified by Furneaux and Wade (2009), Sidorova et al., (2008) and Shuraida et al., (2018).

Most Important IS concepts	Furneaux and Wade's (2009) construct components	Sidorova et al's (2008) research themes	Shuraida et al's (2018) construct categories
Knowledge/ Information	Knowledge/Information	Knowledge management (IT & Org)	Knowledge (IT Management)
Success/Value/ Performance	Performance	Success/Value (IT & Org)	Impact
Adoption	Use	Adoption (IT & In-	Use (Use)

		dividual, IT & Org, IT & Markets)	
Technology	Technology	Development	IT Artefact
Outsourcing	Outsourcing	Planning & Out-sourcing	Outsourcing (IT Management)
Virtual teams	-not ranked-	Virtual team (IT & Groups)	-not ranked-
Culture	-not ranked-	Culture (IT & Org)	-not ranked-
Decision support	-not ranked-	Decision support systems (IT & Groups)	-not ranked-
Alignment/Strategy	Integration	Strategic alignment (IT & Org)	Governance (IT Management)
Privacy	-not ranked-	Personalisation and privacy (IT & Individual, IT & Markets)	-not ranked-
Trust	Trust	Trust (IT & Individual, IT & Markets, IT & Group)	-not ranked-

Table 3: Summary of Most Significant IS Concepts

The construct components from Furneaux and Wade (2009) represent 27% of all the relationships antecedents and consequents found in *MISQ* and *ISR* between 1999 and 2007 inclusively. High-loading terms used by Sidorova et al., (2008) to factor analyse research themes were matched to the most significant concepts extracted from the corpus of IS review articles. Constructs from Shuraida et al., (2018) represent underlying concepts from four major IS journals. Several observations can be made when comparing the most important concepts in the first column with the other three studies in Table 3.

5.1 Important IS Concepts Ignored

Because Furneaux and Wade (2009) and Shuraida et al., (2018) collected only constructs antecedent and consequent in the hypotheses of empirical studies, several important IS concepts such as virtual teams, decision support, culture and trust found in this study were not ranked in their studies. This is because those constructs were not "measured" when putting together the nomological network for their samples. Focusing on constructs places blinders on researchers and ultimately on the field because their view of the research is limited by the framework or theory which they have borrowed from their "reference discipline." Concepts that are relevant to their research interests may be left out or not investigated. Counterintuitively, focusing on just constructs (hypotheticals) tends to reproduce research that is abstruse because constructs are more abstract and further removed from empirical data than concepts.

5.2 Loss of Conceptual Depth

This first problem is partly addressed by viewing research in terms of its research areas or themes, as what Sidorova et al., (2008) did. However, viewing research in terms of research areas without focusing on the detailed concepts in those areas risks losing the necessary depth to help make the research relevant to stakeholders. For example, IT adoption is not only a major research theme, according to Sidorova et al., (2008), it is the dominant research theme within most of the research areas in IS including IT and individuals, IT and organizations and IT and markets. Yet, the concepts of adoption, use, appropriation or other cognate concepts in IS were never clearly defined or agreed on by IS researchers (Burton-Jones and Straub, 2006).

5.3 Need to Link to Historical Tradition

Focusing on hypothetical constructs or topical research areas or just heavily borrowing terms from industry divorces the concept of interest from its historical traditions. Masking the constructs or terms by prefixing them with the term "IS" or "IT" does not make them distinctive or unique to the IS field. For example, strategy in IS or culture in IS do not carry any significant theoretical tradition beyond what management scholars defined strategy to be or what has been defined as organizational culture. How different is IS culture from organizational culture? Is strategy in IS limited to strategic alignment? Or is there something more associated with strategy in the IS field? What is needed is the link of these concepts to its historical traditions that built those concepts. If IS culture is distinct from organizational culture, then, it is its historical traditions that will distinguish it.

6 Conclusion

This article began with a rare observation – constructs are not the same as concepts – and the common practice of focusing on constructs obscures the more important bigger picture surrounding the important concepts that distinguish the IS field from other disciplines. Without the detailed analyses required to develop native IS concepts, our field resorts to prefixing the terms it borrows from other disciplines with the name of the IS field itself or the name of the technology it is researching in the effort to domesticate those terms to the IS field. Therefore, the IS field misses the opportunity to enhance its practice of conceptual formation. Using NLP-based scientometric principles and techniques, we extract a sample of the most important concepts that represent the core concerns of the IS field. Although many of these concepts receive the agreement of the IS field, we also find that many of them are either absent from the list of constructs that IS researchers commonly use in their hypotheses or not clearly defined.

Based on our analyses, we recommend that IS researchers apply a more judicious and disciplined use of concepts in their research by (1) clearly declaring the concepts of interest in their research instead of heavily borrowing constructs from other disciplines, (2) linking the concepts to empirical data especially to phenomena that are of concern to the field's stakeholders and to society, (3) working with metaphors and analogies to develop the concepts of interest (Schön, 1963; Hassan, Mathiassen and Lowry, 2019), and (4) linking the concepts to existing concepts both internally within the IS field and externally to other disciplines. This approach will not only help alleviate the internal communication deficit that exists within the IS field, it will help make the IS field intellectually and socially relevant and influential because concepts are more fundamental than the theories which are stated in terms of them.

References

Alavi, M. and D. E. Leidner. (2001). "Knowledge management and knowledge management systems: Conceptual foundations and research issues." *MIS Quarterly*, *25*(1), 107–136.

Anderson, J. M. (2003). "Why we need a new definition of information security." *Computers & Security*, *22*(4), 308–313.

Association for Information Systems. (2011). "Senior Scholars' Basket of Journals." Retrieved from https://aisnet.org/general/custom.asp?page=SeniorScholarBasket (visited on November 26, 2019)

Bacharach, S. B. (1989). "Organizational Theories: Some Criteria for Evaluation." *Academy of Management Review*, *14*(4), 496–515.

Bal, M. (2002). *Travelling Concepts in the Humanities: A Rough Guide*. Toronto, CA: University of Toronto Press.

Barki, H. and S. Rivard. (1988). "An Information Systems Keyword Classification Scheme." *MIS Quarterly*, *12*(2), 298–322.

Barki, H., S. Rivard and J. Talbot. (1993). "A Keyword Classification Scheme for IS Research Literature: An Update." *MIS Quarterly*, *17*(2), 209–226.

Benbasat, I. and H. Barki. (2007). "Quo vadis TAM?" *Journal of the Association for Information Systems*, *8*(4), 211–218.

Burns, T. and G. M. Stalker. (1961). *The Management of Innovation*. London: Tavistock Institute.

Burton-Jones, A. and D. W. Straub. (2006). "Reconceptualizing system usage: An approach and empirical Test." *Information Systems Research*, *17*(3), 228–246.

Chen, W. and R. Hirschheim. (2004). "A Paradigmatic and Methodological Examination of Information Systems Research from 1991 to 2001." *Information Systems Journal*, *14*(3), 197–235.

Cronbach, L. J. and P. E. Meehl. (1955). "Construct validity in psychological tests." *Psychological Bulletin*, *52*(4), 281–302.

Culnan, M. J. (1987). "Mapping the intellectual structure of MIS, 1980-85: A co-citation analysis." *MIS Quarterly*, *11*(3), 340–353.

Culnan, M. J. and E. B. Swanson. (1986). "Research in management information systems, 1980-1984: Points of work and reference." *MIS Quarterly*, *10*(3), 289–302.

Delone, W. H. and E. R. McLean. (1992). "Information System Success: The Quest for the Dependent Variable." *Information Systems Research*, *3*(1), 60–95.

Delone, W. H. and E. R. McLean. (2003). "The DeLone and McLean Model of Information Systems Success: A Ten-Year Update." *Journal of Management Information Systems*, *19*(4), 9–30.

Dibbern, J., T. Goles, R. Hirschheim and B. Jayatilaka. (2004). "Information systems outsourcing." *ACM SIGMIS Database*, *35*(4), 6–102.

Falkenberg, E. D., K. Lyytinen and A. A. Verrijn-Stuart. (2000). "Information System Concepts: An Integrated Discipline Emerging, IFIP TC8/WG8.1 International Conference on Information System Concepts: An Integrated Discipline Emerging (ISCO-4), September 20-22, University of Leiden, The Netherlands." Boston: Kluwer Academic.

Furneaux, B. and M. Wade. (2009). "Theoretical Constructs and Relationships in Information Systems Research." In: Y. K. Dwivedi, B. Lal, M. D. Williams, S. L. Schneberger, & M. Wade (Eds.), *Handbook of Research on Contemporary Theoretical Models in Information Systems* (pp. 1–17). Hershey, PA: IGI Global.

Garfield, E. (1984). "The 100 Most-Cited Papers Ever and How We Select Citation Classics." *Current Contents*, *23*, 3–9.

Goertz, G. and J. Mahoney. (2012). "Concepts and measurement: Ontology and epistemology." *Social Science Information*, *51*(2), 205–216.

Hart, C. (1998). *Doing a Literature Review*. Thousand Oaks: Sage Publications.

Hassan, N. R. (2011). "Is information systems a discipline? Foucauldian and Toulminian insights." *European Journal of Information Systems*, *20*(4), 456–476.

Hassan, N. R. (2014). "Useful products in theorizing for information systems." In: *International Conference for Information Systems (ICIS), Dec 14-17*. Auckland, NZ.

Hassan, N. R. and C. Loebbecke. (2017). "Engaging scientometrics in information systems." *Journal of Information Technology*, *32*(1), 85–109.

Hassan, N. R., L. Mathiassen and P. B. Lowry. (2019). "The process of information systems theorizing as a discursive practice." *Journal of Information Technology*, *34*(3), 198–220.

Ives, B., S. Hamilton and G. B. Davis. (1980). "A framework for research in computer-based management information systems." *Management Science*, *26*(9), 910–934.

King, W. R. and J. He. (2006). "A meta-analysis of the technology acceptance model." *Information and Management*, *43*(6), 740–755.

Kohli, R. and S. Devaraj. (2003). "Measuring information technology payoff: A meta-analysis of structural variables in firm-level empirical research." *Information Systems Research*, *14*(2), 127–145.

Kohli, R. and V. Grover. (2008). "Business value of IT: An essay on expanding research directions to keep up with the times." *Journal of the Association for Information Systems*, *9*(1), 23–39.

Legris, P., J. Ingham and P. Collerette. (2003). "Why do people use information technology? A critical review of the technology acceptance model." *Information & Management*, *40*(3), 191–204.

Lowry, P. B., G. D. Moody, J. Gaskin, D. F. Galletta, S. Humphreys, J. B. Barlow and D. Wilson.

(2013). "Evaluating journal quality and the Association for Information Systems (AIS) Senior Scholars' Journal Basket via bibliometric measures: Do expert journal assessments add value?" *MIS Quarterly*, *37*(4), 993–1012.

MacCorquodale, K. and P. E. Meehl. (1948). "On a distinction between hypothetical constructs and intervening variables." *Psychological Review*, *55*(2), 95–107.

Markus, M. L. and C. S. Saunders. (2007). "Editorial comments: Looking for a few good concepts...and theories...for the information systems field." *MIS Quarterly*, *31*(1), iii–vi.

Mason, R. O. and I. I. Mitroff. (1973). "A Program for Research on Management Information Systems." *Management Science*, *19*(5), 475–487.

Melville, N., K. Kraemer and V. Gurbaxani. (2004). "Review: Information Technology and Organizational Performance: An Integrative Model of IT Business Value." *MIS Quarterly*, *28*(2), 283–322.

Merton, R. K. (1973). *The Sociology of Science: Theoretical and Empirical Investigations*. Chicago, IL: The University of Chicago Press.

Morgeson, F. P. and D. A. Hofmann. (1999). "The structure and function of collective constructs: Implications for multilevel research and theory development." *Academy of Management Review*, *24*(2), 249–265.

Orlikowski, W. J. and J. J. Baroudi. (1991). "Studying information technology in organizations: research approaches and assumptions." *Information Systems Research*, *2*(1), 1–28.

Paré, G., M. C. Trudel, M. Jaana and S. Kitsiou. (2015). "Synthesizing information systems knowledge: A typology of literature reviews." *Information and Management*, *52*(2), 183–199.

Petter, S., W. DeLone and E. McLean. (2008). "Measuring information systems success: Models, dimensions, measures, and interrelationships." *European Journal of Information Systems*, *17*(3), 236–263.

Prester, J., G. Wagner and G. Schryen. (2018). "Classifying the ideational impact of IS review articles: A natural language processing based approach." In: *International Conference on Information Systems (ICIS 2018)* (pp. 1–17).

Price, D. J. de S. (1965). "Networks of scientific papers." *Science*, *149*(3683), 510–515.

Rowe, F. (2014). "What literature review is not: Diversity, boundaries and recommendations." *European Journal of Information Systems*, *23*(3), 241–255.

Sartori, G. (1975). "The Tower of Babel." In: G. Sartori, F. W. Riggs, & H. Teune (Eds.), *Tower of Babel: On the Definition and Analysis of Concepts in the Social Sciences* (Vol. No. 6, pp. 7–37). Pittsburgh, PA: International Studies Association.

Sartori, G., F. W. Riggs and H. Teune. (1975). *Tower of Babel : On the Definition and Analysis of Concepts in the Social Sciences*. Pittsburgh, PA: International Studies Association, University of Pittsburgh.

Schepers, J. and M. Wetzels. (2007). "A meta-analysis of the technology acceptance model: Investigating subjective norm and moderation effects." *Information and Management*, *44*(1), 90–103.

Schön, D. A. (1963). *The Displacement of Concepts*. London: Tavistock Publications.

Schryen, G., G. Wagner and A. Benlian. (2015). "Theory of knowledge for literature reviews: An epistemological model, taxonomy and empirical analysis of IS literature." In: *International Conference on Information Systems (ICIS 2015)* (pp. 1–22).

Schultze, U. and D. E. Leidner. (2002). "Studying Knowledge Management in Information Systems Research: Discourses and Theoretical Assumptions." *MIS Quarterly*, *26*(3), 213–242.

Senn, J. A. (1998). "The Challenge of Relating IS Research to Practice." *Information Resources Management Journal*, *11*(1), 23–28.

Shaw, M. (1990). "Prospects for an Engineering Discipline of Software." *IEEE Software*, *7*(6), 15–24.

Shaw, M. (2009). "Continuing Prospects for an Engineering Discipline of Software." *IEEE Software*, *26*(6), 64–67.

Shuraida, S., H. Barki and A. Luong. (2018). "Empirical Research in Information Systems: 2001–2015." *Foundations and Trends in Information Systems*, *2*(3), 237–295.

Sidorova, A., N. Evangelopoulos, J. S. Valacich and T. Ramakrishnan. (2008). "Uncovering the

intellectual core of the information systems discipline." *MIS Quarterly*, *32*(3), 467–482.
Small, H. (1973). "Co-citation in the scientific literature: A new measure of the relationship between two documents." *Journal of the American Society for Information Science*, *24*(4), 265–269.
Small, H. (1978). "Cited documents as concept symbols." *Social Studies of Science*, *8*(3), 327–340.
Small, H. and E. Greenlee. (1980). "Citation context analysis of a co-citation cluster: Recombinant-DNA." *Scientometrics*, *2*(4), 277–301.
Spencer, H. (1897). *The Principles of Sociology*. New York: D. Appleton.
Staggers, N. and C. B. Thompson. (2002). "The Evolution of Definitions for Nursing Informatics." *Journal of the American Medical Informatics Association*, *9*(3), 255–261.
Stamper, R. K. (2000). "Information Systems as a Social Science: An Alternative to the FRISCO Formalism." In: E. D. Falkenberg, K. Lyytinen, & A. A. Verrijn-Stuart (Eds.), *Information Systems Concepts: An Integrated Discipline Emerging* (pp. 1–51). Boston: Kluwer Academic Publishers.
Suddaby, R. (2010). "Editor's Comments: Construct Clarity in Theories of Management and Organization." *Academy of Management Review*, *35*(3), 346–357.
Suddaby, R., C. Hardy and Q. N. Huy. (2011). "Where are the new theories of organization?" *Academy of Management Review*, *36*(2), 236–246.
Tandoc, E. C., Z. W. Lim and R. Ling. (2018). "Defining "Fake News": A typology of scholarly definitions." *Digital Journalism*, *6*(2), 137–153.
Thomson, G. P. (1961). *The Inspiration of Science*. London: Oxford University Press.
Verrijn-Stuart, A. A. (2001). "A Framework of Information Systems Concepts: The Revised FRISCO Report (Draft January 2001)." IFIP.
Wade, M. and J. Hulland. (2004). "The Resource-Based View and Information Systems Research: Review, Extension, and Suggestions for Future Research." *MIS Quarterly*, *28*(1), 107–142.
Webster, J. and R. T. Watson. (2002). "Analyzing the past to prepare for the future: Writing a literature review." *MIS Quarterly*, *26*(2), xiii–xxiii.
Willer, D. and M. Webster Jr. (1970). "Theoretical Concepts and Observables." *American Sociological Review*, *35*(4), 748–757.
Wyatt, J. C. and J. L. Y. Liu. (2002). "Basic Concepts in Medical Informatics." *Journal of Epidemiology and Community Health*, *5*(6), 808–812.