Debates and Perspectives

On Being 'Systematic' in Literature Reviews in IS

Sebastian K Boell[1], Dubravka Cecez-Kecmanovic[2]

[1]Discipline of Business Information Systems, The University of Sydney Business School, Sydney, NSW, Australia;
[2]UNSW BusinessSchool, University of New South Wales, Sydney, NSW, Australia

Correspondence:
S K Boell, Discipline of Business Information Systems, The University Sydney Business School, Sydney, NSW 2006, Australia

Abstract
General guidelines for conducting literature reviews often do not address the question of literature searches and dealing with a potentially large number of identified sources. These issues are specifically addressed by so-called systematic literature reviews (SLRs) that propose a strict protocol for the search and appraisal of literature. Moreover, SLRs are claimed to be a 'standardized method' for literature reviews, that is, replicable, transparent, objective, unbiased, and rigorous, and thus superior to other approaches for conducting literature reviews. These are significant and consequential claims that – despite increasing adoption of SLRs – remained largely unnoticed in the information systems (IS) literature. The objective of this debate is to draw attention of the IS community to SLR's claims, to question their justification and reveal potential risks of their adoption. This is achieved by first examining the origins of SLR and the prescribed systematic literature review process and then by critically assessing their claims and implications. In this debate, we show that SLRs are applicable and useful for a very specific kind of literature review, a meta study that identifies and summarizes evidence from earlier research. We also demonstrate that the claims that SLRs provide superior quality are not justified. More importantly, we argue that SLR as a general approach to conducting literature reviews is highly questionable, concealing significant perils. The paper cautions that SLR could undermine critical engagement with literature and what it means to be scholarly in academic work.

Keywords: systematic literature review; SLR; systematic review; conducting literature reviews; literature review; narrative literature review; database searches; literature search

Introduction

Literature reviews are an important ingredient in all research. Literature reviews typically provide an overview, synthesis and a critical assessment of previous research, challenge or problematize existing knowledge and identify or construct novel research problems and promising research questions (Alvesson and Sandberg, 2011; LePine and Wilcox-King, 2010). The process of conducting a literature review can be daunting especially for novice researchers when faced with a potentially large, unbounded and continuously growing body of literature.There are several approaches and guidelines for conducting literature reviews in information systems (IS) (Bandara et al., 2011; Levy and Ellis, 2006; Schwarz et al., 2007; Webster and Watson, 2002; Wolfswinkel et al., 2013). Importantly, they provide a conceptual foundation for developing and constructing literature reviews. For instance, Webster and Watson (2002) suggest the use of a topic-centric approach for presenting, classifying and assessing relevant literature; Bandara et al. (2011) propose the use of thematic analysis and qualitative research software for analyzing a body of literature; Wolfswinkel et al. (2013) recommend the use of grounded theory for conducting and presenting literature reviews; and Schwarz et al. (2007) provide criteria for distinguishing review and framework articles. While providing insightful approaches and guidelines, these papers do not pay due attention to the role of literature search processes and their importance in conducting literature reviews.

The importance of literature searches, when conducting literature reviews, is discussed by an approach known as systematic literature review (SLR) or just systematic re-

Preprint of: Boell, Sebastian K; &Cecez-Kecmanovic, Dubravka (2015) On Being 'Systematic' in Literature Reviews in IS. *Journal of Information Technology*. 30(2). 161-173. The definitive publisher-authenticated version is available: http://dx.doi.org/10.1057/jit.2014.26

view. The SLR approach is part of the evidence-based movement – originating in medicine and spreading to social sciences – that aims to advance policy and practice by providing the best evidence available from research (Campbell Collaboration, 2007; Morrell, 2008). The distinct feature of SLR is a protocol that prescribes how to identify, select, assess and synthesize evidence from the literature. By adhering to such a protocol, SLRs, it is claimed, provide a 'standardized method' for literature reviews that is *replicable, transparent, objective, unbiased,* and *rigorous* (Atkins and Louw, 2000; Okoli and Schabram, 2009, 2010; Oates, 2011; Oates et al., 2012). SLRs are thus proposed as a superior alternative to so-called narrative or traditional literature reviews (Baumeister and Leary, 1997; Hjorland, 2011). The latter are by implication assumed to be unsystematic, biased, non-replicable, unscientific and non-rigorous (MacLure, 2005; Morrell, 2008).

As a particular approach and procedure for undertaking literature reviews, SLR was initially proposed in IS by Atkins and Louw (2000) and the first IS specific guidelines for conducting SLRs were published in 2010 (Okoli and Schabram, 2010). To assess the current state of SLRs in IS we conducted a search for currently published SLRs in IS.[1] We plotted the number of papers whichappeared in major IS conferences (ICIS, ECIS, AMCIS) and journals in Figure 1.

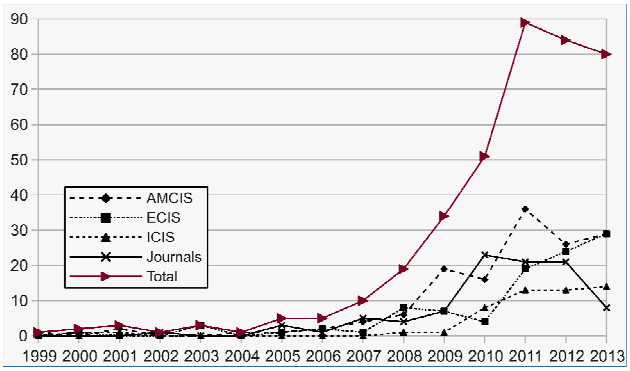

Figure 1 An illustration of the increasing number of papers referring to SLR in Information Systems[2]

Figure 1 illustrates a notable increase in the use of SLRs within the IS discipline. These SLRs use a rhetoric of scientific principles, elimination of researcher's interpretation and judgment as a source of bias, and the necessity for transparency (e.g. Dwivedi et al., 2008; Mohan and Ahlemann, 2011). By adopting the SLR protocol, they make claims to objectivity, replicability and rigor without providing further explanation of or justification for these claims (e.g. Grahlmann et al., 2012; Mohan and Ahlemann, 2011; Roztocki and Weistroffer, 2008; Wolfswinkel et al., 2010). Furthermore, some assert that SLR allow an analysis of 'the current state of play' through a 'comprehensive review'(Williams, et al., 2009), and that SLR lead to 'sound' reviews for IS (Cheung and Thadani, 2010).

While SLRs are making serious inroads into the IS literature, claiming superiority to traditional, narrative literature reviews, the IS research community remains silent, seemingly suggesting tacit approval. It is indeed puzzling that SLRs and their claims in the IS discipline have not been scrutinized especially given growing concerns and criticism of both SLRs and the evidence-based movement in management (Morrell, 2008), information science (Hjorland, 2011), education (MacLure, 2005), and nursing (Finfgeld-Connett and Johnson, 2013;Sandelowski, 2008). MacLure (2005), for instance, provides a thorough critique of SLR in education research (and more broadly in social sciences) arguing that it "constitutes a threat to quality and critique in scholarship and research" (p. 393).

These concerns and numerous criticisms raised in other disciplines, and the increasing adoption of SLR in the IS literature are yet to spark a serious debate about SLR in the IS discipline. It seems that currently SLRs are entering the field of IS without sufficient reflection on their assumptions, limitations and implications for IS scholarship. The aims of this debate paper are therefore to draw the attention of IS researchers, students and practitioners to the claims of SLRs to quality, to question justification of these claims, and to demonstrate implications for scholarship and potential risks of SLR adoption as a general approach to literature reviews. By critically engaging with SLR in IS and more broadly social sciences, the debate contributes to the understanding of SLR and the nature of claims made by their proponents and, most importantly, reveals significant perils and implications for scholarship.

The debate is structured as follows. The next section introduces the origins and intentions of SLRs in medicine and traces their adoption by the social sciences and IS. The following sections then discuss key differences between traditional narrative reviews and SLR and assumptions underlying the concept of SLRs and database searches. We continue by critically engaging with central claims made by SLR proponents, namely that they are replicable, unbiased, scientific and rigorous and thus superior to other approaches for conducting literature reviews. Before we finally discuss potential use of SLR in IS we argue that unreflective adoption of SLRs turns a blind eye to the risks and limitations of SLRs as a general approach for conducting literature reviews.

Origins and Procedures of Systematic Literature Reviews

SLRs are proposed as a novel way for conducting literature reviews. The development of SLRs can be traced back to the evidence-based medicine movement originating in early 1990s following the publication by the *Evidence-based Medicine Working Group* (1992) in the medical journal *JAMA* and a meeting between the *British Medical Journal* and the UK *Cochrane Centre*(Chalmers and Altman, 1995). The issue faced by medical researchers was that of inconclusive findings regarding medical treatments, such as prescriptions, and the role of literature reviews when combining the findings of multiple studies. Originally SLRswere, therefore, closely associated with *meta-analysis* (Clarke and Stewart, 1995; Eysenck, 1995; Thompson, 1995), a type of

Preprint of: Boell, Sebastian K; &Cecez-Kecmanovic, Dubravka (2015) On Being 'Systematic' in Literature Reviews in IS. *Journal of Information Technology*. 30(2). 161-173. The definitive publisher-authenticated version is available: http://dx.doi.org/10.1057/jit.2014.26

literature review that 'synthesizes and combines numerical results of earlier research' (Cruzes and Dybå, 2011). In particular, Knipschild (1995) critiqued literature reviews for providing only *spotty* coverage of the overall evidence that was published regarding a particular medical treatment. The spotty coverage can then lead to bias in assessing the treatment. This is a serious issue, as it may lead either to unnecessary studies on a medical treatment already shown to be unpromising, or even have more severe consequences when hindering treatment which could be life saving (Mulrow, 1995). SLRs are used in medicine to avoid such a bias by applying more rigorous methods when conducting literature reviews (Oxman, 1995).

Achieving a rigorous review – according to Chalmers and Altman (1995) – necessitates, firstly, developing a review protocol that clearly specifies the research question that should be addressed as well as inclusion and exclusion criteria for selecting and assessing publications. Secondly, every effort should be made to identify as many publications as possible relevant to this question. This also requires going beyond a set of specific journals or databases, to survey 'grey' literature to include foreign language publications as well as unpublished research (Knipschild, 1995). This also includes negative or inconclusive findings that are less often published than positive results. The process of identifying and obtaining such research can be tedious and time consuming. However, it is justified by the even greater time required for conducting additional medical studies, typically high costs, and potential risks associated with such studies. Thirdly, after an exhaustive list of publications is compiled each publication is screened for methodological rigor and soundness of its findings. This forms the basis for inclusion or exclusion of papers (evidence), based on criteria clearly specified in the protocol. For instance, there are different types of medical evidence, which are ranked according to a 'hierarchy of evidence' ranging from professional opinion (weak) to randomized controlled trials (strong) (Pawson, 2006). More often SLR would be limited to randomized controlled trials only. Fourthly, the results from all included studies are screened and often statistically processed in order to arrive at an overall assessment of the evidence regarding the effectiveness of a treatment. And fifth, the results of the study are prepared for dissemination through databases or publications.

Since the mid 1990s, the use of SLR spread from medicine through the field of health informatics (e.g. Ramsay et al., 2000; Shiffman et al., 1999) to psychology (e.g. Rose et al., 2003), nursing (Sandelowski, 2008), education (MacLure, 2005) library science (Hjorland, 2011), software engineering (Kitchenham, 2004; Kitchenham and Charters, 2007), and more recently into IS (Okoli and Schabram, 2010). Acknowledging that social sciences differ from medicine, Petticrew and Roberts (2006) developed a guideline for conducting SLR in the social sciences. At first glance, all of these SLR guidelines, including the ones proposed for IS (c.f. Okoli and Schabram, 2010) look similar to those initially proposed in medicine (see Figure 2). Authors first develop a protocol aiming to address a specific research question. In step 2 database searches are conducted according to the protocol in order to identify relevant publications. In step 3 retrieved publications are then screened (typically based on abstracts) for quality of method and findings (evidence) and in turn narrowed down to only those publications considered appropriate for the review. Based on the selected publications, a review is compiled by 'summarizing' the findings of the individual publications in step 4 and results are distributed in step 5.

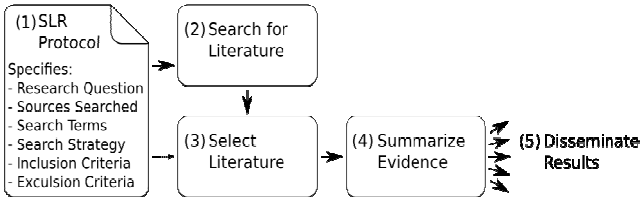

Figure 2 Overview of the SLR process

While these SLR guidelines look similar to those proposed and used in medicine, there is a considerable difference between them: SLRs were proposed in medicine as a particular means for the *summation of evidence regarding particular medical treatments*. SLRs therefore offer a particular means for conducting literature reviews for a specific purpose — that of summarizing clearly identifiable evidence from earlier research. However, SLR guidelines outside of medicine often fail to acknowledge this limitation and therefore are misunderstood as offering a general approach to conducting literature reviews.

Furthermore, SLRs in medicine focus on *the comprehensiveness of the coverage*, whereas SLR guidelines outside of medicine emphasize *the rigor of the search process*. Kitchenham, for instance, emphasizes that it is the rigor of the literature search process "that distinguishes systematic reviews from traditional reviews" (2004:7). Rigor here is seen as being directly related to the process of identification of literature. It is based on the assumption that "the quality of literature reviews is particularly determined by the literature search process" (Brocke et al., 2009:1). In medicine, however, rigor relates to comprehensiveness of the coverage (e.g. inclusion of 'grey' literature) and the methodological rigor of research in selected literature sources (e.g. randomized controlled trials) (Chalmers and Altman, 1995).

These differences are of key importance for understanding how the concept of SLR became translated and distorted as it travelled to other disciplines. The central focus of SLRs in IS and software engineering, for instance, is on *searching*, its rigor and replicability and therefore the reproducibility of document sets that are assessed as part of a review (e.g. Kitchenham, 2004; Oates, 2011; Okoli and Schabram, 2010). The claim to 'rigor' is preserved while its meaning has changed: rigor in medicine refers to comprehensiveness of the literature and the methodology of the selected research studies, *not* how they are identified or whether the process of identifying them is *reproducible* by others.

Preprint of: Boell, Sebastian K; &Cecez-Kecmanovic, Dubravka (2015) On Being 'Systematic' in Literature Reviews in IS. *Journal of Information Technology*. 30(2). 161-173. The definitive publisher-authenticated version is available: http://dx.doi.org/10.1057/jit.2014.26

Systematic Literature Reviews vs Traditional Reviews

SLRs by their very definition differ from traditional literature reviews: their scope is to address a highly specific research question (e.g. 'what works' or 'what works best') for which evidence from the literature is sought. SLRs thus do not aim to provide what traditional literature reviews do: an assessment of a state of knowledge in a problem domain and identification of weaknesses and needs for further research (Finfgeld-Connett and Johnson, 2013; Hart, 1989). However, in IS, SLRs are often portrayed as an alternative general approach to literature review (e.g. Mohan and Ahleman, 2011; Okoli and Schabram, 2009, 2010). To understand both the SLR approach and how it differs from traditional 'narrative' approaches, we examine the aims and underlying assumptions of both. Table 1 summarizes and compares key characteristics derived from the SLR literature (Knipschild, 1995; Atkins and Louw, 2000; Kitchenham, 2004; Kitchenham and Charters, 2007, Okoli and Schabram, 2010; Cruzes and Dybå, 2011) and those from a broad literature on literature reviews in IS and beyond (Bandara et al., 2011; Boell and Cecez-Kecmanovic, 2014; Feak and Swales, 2009; Finn, 2005; Hart, 1989; Levy and Ellis, 2006; Machi and McEvoy, 2009; Ridley, 2008; Schwarz et al., 2007; Webster and Watson, 2002; Wolfswinkel et al., 2013).

SLRs specify up-front inclusion/exclusion criteria which are then systematically and rigorously applied in the selection of relevant literature, with minimal judgement by individual researchers. Researchers' interpretation, imagination, creativity, and individuality in selecting and judging the studies and findings are to be minimized. This is why SLRs often engage several researchers and seek to achieve high inter-researcher reliability (Jalali and Wohlin, 2012). This is in contrast to traditional literature reviews that assume that researchers develop relevance criteria as part of their increasing understanding of the literature. The increasing breadth and depths of understanding the literature help researchers become more critical and reflective in assessing the studies, thus leading them to continually develop relevance criteria. Originality of views, reflexivity and imagination in comparing and contrasting findings and assessing the state of knowledge in a particular problem domain are seen as highly valuable and are encouraged.

SLRs introduce a discourse that regulates what counts as 'evidence' or 'scientific evidence' that is worthy of inclusion, thus rendering other potentially interesting evidence irrelevant. The privileged type of evidence allows and justifies a particular 'extraction' and 'summation' of results assuming that different studies address exactly the same question and that the aggregation of results is meaningful. This is in contrast to discourses in traditional literature reviews that promote insight into and engagement with literature, problematization of assumptions behind certain approaches or findings, and a critical assessment of existing knowledge based on which an argument is developed justifying further research.

Assumptions about Database Searches in SLR

A centralcharacteristic of SLRs is their focus on database searches as key processes for producing literature reviews. It is thus important to have a closer look at their assumptions regarding database searches. Often SLRs presume that good search terms can be known in advance, and therefore specified up front in a protocol in step 1 (Figure 2). Furthermore, it is assumed that well-specified search terms (and their synonyms and abbreviations) are key for identifying a 'complete' set of relevant documents and that the quality of database searches is predominantly determined by the number of relevant sources identified by a search. And finally, it is assumed that databases generally provide a good coverage of relevant literature.

The assumption that good search terms can be known and defined up front is problematic for several reasons. Generally, terms used to describe a topic are non-finite and cannot be known a priori (Bodoff, 2009). Furthermore, when engaging with research literature in an area and its body of knowledge, it is typically not clear what central, and therefore good, search terms are. Learning the concepts and the language used to discuss them is an important part of the actual process of reviewing the literature. Familiarity with concepts and relevant terminology used to describe a problem are the result of an engagement with the literature. Even when lists of search terms are open to adjustment at a later point (e.g. Kitchenham and Charters, 2007), it is not understood that when undertaking a literature review people are searching for certain ideas and *concepts* and not for *terms* (Fugmann, 2007). Concepts are not only expressed using specific terms, they are also described or paraphrased. This is of particular importance when looking at emerging phenomena or new research areas, as these are generally still establishing a specific terminology to talk about the research domain and its various aspects. Especially in these instances, phenomena are paraphrased and described in different ways rather than through a well-established terminology that can be used for searching. For example, those conducting a literature review on the 'Technology Acceptance Model' or simply TAM will not find Davis's (1989) landmark paper when searching for either the phrase or the abbreviation, as neither is used anywhere in the publication.

The use of alternative terminology and the lack of precision of search terms refer to a well-known problem in information retrieval described by Blair (2006) as the 'indeterminacy of language'. For instance, Checkland and Holwell (1998:134), in their example of the soft systems around the 'Battle of Britain', provide a list of twelve different expressions used to describe the concept of a 'radar', without mentioning the term at all. Agrawal et al. (2010) provide another example of the shortcomings of search term-based literature reviews. For identifying relevant literature, Agrawal at al. (2010) used a set of 'often used keywords in HCI papers' for their searches. However, their search term driven approach found only seven percent of the literature that was identified by Zhang et al. (2009) in four top IS journals for the same topic in the same journals over the same time period.

Preprint of: Boell, Sebastian K; &Cecez-Kecmanovic, Dubravka (2015) On Being 'Systematic' in Literature Reviews in IS. *Journal of Information Technology*. 30(2). 161-173. The definitive publisher-authenticated version is available: http://dx.doi.org/10.1057/jit.2014.26

Table 1 Comparison of SLR and traditional narrative reviews

	Systematic Literature Reviews (SLRs)	*Traditional narrative literature reviews*
Description	SLR is a protocol-based approach to literature review which explicitly prescribes the steps and processes for searching, selecting and validating studies and summarizing their results (Atkins and Louw, 2000' Okoli and Schabram, 2010). 'Systematic' in SLRs assumes that the review process that follows the explicit procedure and rules is a guarantor that the literature review is 'objective', 'scientific', 'transparent', 'replicable' and 'rigorous'.	Traditional literature review is a creative process through which a researcher identifies and examines prior research and develops increasing understanding of a phenomenon under examination and in the process constructs the relevant body of knowledge. Various guidelines for literature reviews assist researchers in identifying, reading, analysing, interpreting, mapping, classifying and critically assessing the literature and writing a literature review (Boell and Cecez-Kecmanovic, 2014; Levy and Ellis, 2006; Webster and Watson, 2002; Wolfswinkel et al., 2013).
Purpose and aims	SLRs seek to provide answers to specific questions such as "what works" or "what works best" (e.g. related to the use of a software, tool, method); 'how one variable is related to another' (e.g. the relation between "ease of use" and "intention to use technology"); or questions regarding a particular hypothesis (e.g. that a product/tool/technique has positive effects on productivity). SLR aim to provide evidence that answers such type of questions. The presentation of the literature review and its outcomes follow a specific schema that is similar across different SLRs (e.g. Amrollahi et al., 2013; Gräning et al., 2011; Williams et al., 2009).	Traditional reviews typically address a topic (a phenomenon or research problem), aiming to develop comprehensive understanding and critical assessment of knowledge relevant for the topic. Reviews involve a dialogical interaction among the researcher and the literature (Wright Mills, 1978) and can cross disciplinary boundaries. Depending on their purpose literature reviews can be very different in their structure and presentation: it may build upon an existing classification framework (Schultze and Leidner, 2002), focus on an analysis of researched variables (Lacity et al., 2011), categorize and summarize streams of earlier research (Leonardi and Barley, 2010), etc.
Literature selection – relevance criteria	Relevance criteria are defined in advance: studies are included/excluded based on the validity of their findings (external and internal validity), which is ideally replicable by others. It is assumed that the validity of a study and its findings can be assessed abstractly based on research design (and method used) only, using "the credibility hierarchy of research designs" (e.g. randomized controlled trials being most credible) (Morrell, 2008).	Relevance criteria are gradually developed as the review progresses and researcher's understanding advances. It is assumed that inclusion/exclusion of studies in the final literature review is based on researcher's insight and contextually-sensitive judgement of studies' relevance and contribution. The validity of a study is assessed based on both explicit and tacit knowledge, standpoint and worldview, and also merits of research design and execution (Hart, 1998).
Role of researcher	SLRs presume and encourage minimal reliance on researcher's interpretation, judgement and discretion in order to eliminate subjectivity and bias. Given the procedural nature of the review process the role and influence of the researcher in the resulting literature review is minimized (see e.g. Staples and Niazi, 2007; Jalali and Wohlin, 2012).	It is assumed that a researcher develops better understanding of a phenomenon by extensive and in-depth reading, interpreting, comparing, contrasting, classifying, and questioning publications. A researcher uses judgement and discretion to classify, critically assess and determine importance of individual studies and their contributions to knowledge (see e.g. Lacity et al., 2011; Leonardi and Barley, 2010; Schultze and Leidner, 2002).
Literature review content	SLRs regulate what counts as evidence and prescribe a particular knowledge extraction and summation process: the results from selected studies are combined and aggregated assuming they address exactly the same phenomenon and the aggregation of results is meaningful. Aggregated evidence aims to provide conclusive assessment regarding the research question.	A literature review is expected to provide an overview, classification, comparison, and mapping of previous research results, as a basis for critical assessment of existing knowledge (potentially revealing weaknesses and poor understanding of a phenomenon or problematizing approaches, theories or claims). Critical assessment of literature is essential for developing an argument that certain areas or topics warrant further research.

Furthermore, high quality searches are not only effective in identifying relevant literature. Database searches also need to be efficient in avoiding (the identification of) *irrelevant* literature. Generally, search results provided by information retrieval systems can be assessed by two central measurements. These are recall and precision (Salton and McGill, 1983). *Recall* refers to the number of relevant documents retrieved by a specific query compared to the number of relevant documents that could potentially be found in a database. In contrast, *precision* reflects how

Preprint of: Boell, Sebastian K; &Cecez-Kecmanovic, Dubravka (2015) On Being 'Systematic' in Literature Reviews in IS. *Journal of Information Technology*. 30(2). 161-173. The definitive publisher-authenticated version is available: http://dx.doi.org/10.1057/jit.2014.26

many relevant documents are among the documents found. Ideally, a good query has a high recall as well as high precision. That is, it should find all relevant documents stored in a database while not picking up any irrelevant documents. However,"it is impossibly difficult for inquirers using a large full-text retrieval system to predict, by means of their search queries, the exact words, word combinations, and phrases that are used to express the content of the documents they desire but are *not used* to express the content of the documents they do not desire" (Blair, 2006:302).

This problem is well known in information retrieval as recall-precision trade-off (Buckland and Gey, 1994). Good search strategies therefore aim at an optimum trade-off between recall and precision. According to SLR guidelines, searches are optimized for retrieving as many (presumably) relevant documents as possible, thus leading to high recall. High recall generally happens at the expense of the accuracy of searches – resulting in very low precision. Large document sets that are typically retrieved pose a new problem: they need to be assessed for relevance. For instance, Mohan and Ahlemann (2011) report that their search in eight databases retrieved 22,291 documents, however,after spending 74 hours examining them (based on titles and abstracts) only 58 were selected and eventually only 17 were considered for inclusion in the literature review.

This example illustrates the typical outcome of high recall/low precision searches in step 2 of the SLR process. Search strategies advocated by SLRs result in large document sets that then need to be assessed for relevance. This is often done based on document titles and abstracts. However, titles and abstracts are often not successful in conveying the content of documents (Hartley and Betts, 2009). As a consequence, some relevant documents that are retrieved may be lost because their titles and abstracts do not sound promising during the initial 'screening'. Aiming for high recall searches with an immense number of results is therefore problematic as 'the bigger the results set, the greater the chance for error' in the selection of relevant literature (Blair, 2006). This further compromise the presumption of completeness of the selected document set.

Finally, there is the underlying assumption that databases generally provide a good coverage of relevant publications. Evidence from SLRs in medicine, however, tells a different story. Knipschild (1995) reports that only 36 percent of the literature was identified through database searches, and Greenhalgh and Peacock (2005) report only 25 percent. This is despite the availability of large scale literature databases (e.g. *Medline*, *Embase*, and *PubMed Central*)andthe fact that in medicine controlled vocabulary (e.g. *MeSH*) is much further developed than in the social sciences and IS. Generally, the coverage of relevant literature in databases is far from being complete (Hood and Wilson, 2001). As Levy and Ellis (2006) put succinctly: "A common mistake by novice researchers, specifically in IS, is to assume that the keyword search yields all that is available from the literature." (p.190).

Claims to Quality in Systematic Literature Reviews

SLR guidelines frequently claim*objectivity* and *replicability*as key distinguishing qualities of their literature reviews. This claim is justified by the transparency of the review process: search strategies are fully disclosed; the steps subsequently taken to narrow down the literature set are described (e.g. inclusion/exclusion criteria, analysis of titles and abstracts); and the final summation of results is detailed. This is most notable in the tables containing the numbers of documents obtained from different databases, and the visualization of the different steps of the review accompanied by the number of documents making it through each phase (e.g. Mohan and Ahlemann, 2011). However, as Oancea and Pring (2008) observe, in this context, transparency is treated as a purely technical matter. SLRs confuse 'procedural objectivity' of the process with objectivity of review outcome. SLRs thus do not guarantee "transparency or replicability of review outcomes" (Sandelowski, 2008:105). While the order of steps is "objective in that the steps taken are communicable and, therefore, repeatable as steps" (Sandelowski, 2008:106), the result coming out of this process depends on subjective decisions made along the way regarding quality criteria that serve as the basis for the selection and exclusion of studies, and the comparison and combination of studies. As a result SLRs are not and cannot be replicable as they are inevitably based on subjective judgment:"What is immediately transparent and replicable is only adherence to the tasks and to a style for reporting them, not the enactment of those tasks. Indeed, systematic reviews are reliably unreliable as any one review is a product of the comparability work that defines the unique interaction between reviewers and the body of research they created for review" (Sandelowski et al., 2007:244).

It is thus no surprise that there are various reports of 'problems' with inter-rater reliability (Staples and Niazi, 2007), and replicability of SLR (Jalali and Wohlin, 2012). The reason for this is that "[d]ifferent readers will extract different meanings from texts, according to their background assumptions and their current priorities and beliefs" (MacLure, 2005:409). These problems are of particular importance to literature reviews in the social sciences as they underline the limitation of using SLRs as an approach for theory development (Finfgeld-Connett and Johnson, 2013).

A further aspect of objectivity is implied by the notion of a 'summation of evidence' from earlier research that SLRs propose. The idea of the summation of evidence, as Hjorland observes, is questionable in the social sciences as it is "too narrow, too formalist, and too mechanical" (2011:1308) as a general means for conducting literature reviews. Unlike medicine where, for instance, it is possible to summarize evidence on the effects of a particular treatment, in social sciences research problems are rarely so well defined and 'evidence' is rarely presented in a form that can be counted and summarized. In addition, in contrast to medicine where well-established quality criteria

Preprint of: Boell, Sebastian K; &Cecez-Kecmanovic, Dubravka (2015) On Being 'Systematic' in Literature Reviews in IS. *Journal of Information Technology*. 30(2). 161-173. The definitive publisher-authenticated version is available: http://dx.doi.org/10.1057/jit.2014.26

exist for medical trials, in the social sciences methodological diversity does not clearly place some research approaches above others (Denyer and Tranfield, 2009). Furthermore, the aim of literature reviews, as numerous works and guidelines on literature reviews explain, is not merely the 'extraction of facts' from a set of publications, but the development of understanding that will allow the formulation of insightful research questions that are thoroughly argued on a basis of a critical assessment of relevant knowledge (Boell and Cecez-Kecmanovic, 2014; Davies and Beaumont, 2007; Feak and Swales, 2009; Finn, 2005; Hart, 1998; Levy and Ellis, 2006; Machi and McEvoy, 2009; Ridley, 2008; Schwarz et al., 2007; Webster and Watson, 2002). It is therefore fair to say that it is impossible to 'objectively' select and integrate results from different and diverse studies as part of a review, in a way that is independent of researchers' subjective judgment and understanding (Hammersley, 2001).

Regarding *bias*, it is important to recall the reasons for the development of SLRs in medicine in the first place. It was based on the observation that, for instance, relying on the prestige of top journals introduces bias when assessing the literature (Chalmers and Altman, 1995). However, SLRs in IS (e.g. Agarwal et al., 2011; Dwivedi, et al., 2008) or software engineering (e.g. Stol et al., 2009) often choose a particular set of journals or conferences as a foundation for their searches. Clearly, the selection of any set of journals or even databases enforces the very bias that SLRs in medicine were initially developed to minimize. In addition, databases cannot be regarded as providing an objective or unbiased selection of literature. For instance, databases only cover a subset of all academic journals and generally do not have a good coverage of book chapters and conference proceedings. Furthermore, databases exclude some journals with relevant publications and may only selectively index other journals, omitting many publications not fitting the scope of a particular database. As this discussion suggests, the simple act of choosing databases or a set of journals and conferences introduces systematic errors and biases.

Furthermore, the translation of SLRs from medicine to other disciplines, including IS, can lead to an empiricist/positivist bias towards scientific knowledge, its accumulation, collection and review (Finfgeld-Connett and Johnson, 2013; Hammersley, 2001; Hjorland, 2011). Proponents of SLRs consider research-based knowledge (empirical findings from research studies) as 'evidence' and treat it as "science at a meta level" (Hjorland, 2011:1302). However, this limits a review to a type of research that can provide 'valid evidence' which is clearly extractable and comparable (e.g. studies including a particular variable) and can thus be summarized as part of the review (Denyer and Tranfield, 2009). SLRs ideally provide "a concise summary of the best available evidence that uses explicit and rigorous methods to identify, critically appraise, and synthesize relevant studies on a particular topic" (Cruzes and Dybå, 2011:3). Importantly, in medicine, critical appraisals are based on a hierarchy of scientific methods that establishes a priori the veracity of research findings irrespective of the content and arguments provided in a paper. At the top are quantitative, positivist methods (in the form of randomized controlled trials) while case studies and qualitative methods are at the bottom (Denyer and Tranfield, 2009; Pawson, 2006; Tranfield et al., 2003). Privileging particular studies (as high quality evidence) based on their approach and method, however, introduces a particular systematic bias, thus contradicting SLR's major claim to non-bias (Murray et al., 2007). Some SLR authors have thus argued that SLRs need to change to also include the extensive body of non-positivist research in the social sciences (Denyer and Tranfield, 2009) and IS (Oates, 2011; Oates et al., 2012).

The claim that the SLR approach is *scientific* and *rigorous* is based on its methodical commitment to a prescribed protocol that is supposed to ensure scientific objectivity, replicability and control of the review process and thus guarantee high quality in the resulting literature review. Some SLR proponents, as discussed above, emphasize the rigor of the literature search process as a distinguishing quality of SLRs (Brocke et al., 2009; Kitchenham, 2004). As the above analysis shows, SLR protocols do not exclude subjectivity of literature selection, do not ensure comprehensiveness of coverage of relevant literature, nor do they produce a replicable document set. SLR thus cannot claim to be rigorous according to their conception of rigor.

In summary, we propose that SLR's notion of rigor of a literature review is narrowly conceived as the adherence to a prescribed protocol and a particular search process. It is worth reminding the reader that rigor in traditional literature reviews is a much deeper concept that apart from comprehensiveness of the literature encompasses an appropriate breadth and depth of analysis and classification, consistency and logic of argumentation, criticality of assessment of a body of knowledge and creativity and effectiveness of synthesis and presentation (Boell and Cecez-Kecmanovic, 2014; Feak and Swales, 2009; Finn, 2005; Hart, 1989; Levy and Ellis, 2006; Machi and McEvoy, 2009; Ridley, 2008; Schwarz et al., 2007).

Perils of Systematic Literature Reviews

The key focus of this debate is the uncritical adoption of SLR as a general, 'standard' approach to literature review that disregards the conditions for and limitations of their adoption. We draw attention here to the perils of SLRs as a general approach to literature reviews. As discussed above, SLRs claim particular quality criteria for literature review – scientific objectivity, transparency, replicability and rigor. These claims to quality are problematic for several reasons. SLR's ideal of scientific objectivity and replicability of literature reviews is ill-conceived and misleading. Technically, if the claims by SLRs held up, someone without any understanding of the literature should still be able to execute the literature review according to a protocol. In this sense, SLRs are problematic as they promote an ideal of scholarship as a replicable and mechanical pursuit in which an academic has no personality, let alone passion or scepticism, but is degraded to an executing technician (Constantinides et al., 2012).

Preprint of: Boell, Sebastian K; &Cecez-Kecmanovic, Dubravka (2015) On Being 'Systematic' in Literature Reviews in IS. *Journal of Information Technology*. 30(2). 161-173. The definitive publisher-authenticated version is available: http://dx.doi.org/10.1057/jit.2014.26

SLRs' claims to scientific objectivity, transparency, replicability and rigor as essential qualities of literature reviews significantly differ from those established in the literature on traditional literature reviews (see Table 1 and Table 2). More broadly, SLRs promote a discourse of transparency, accountability, certainty, and quality assurance, that is antithetical to traditional discourses on literature reviews. Furthermore, claims to "scientific objectivity and ... accountability (are) insinuating that other reviewing practices are both unscientific and shady" (MacLure, 2005:395). Such views have started to infiltrate the reviewing processes in IS as we observed recently in top IS conferences. Some reviewers critique papers for not following SLR protocol and for their lack of objectivity, transparency and rigor.

Furthermore, SLRs require research questions to be established before the literature review commences. However, this may only be possible in some cases concerned with the identification and summation of evidence. In contrast, numerous research methods texts argue that literature reviews are of essential importance for motivating an inquiry and formulating meaningful research questions (e.g. Davies and Beaumont, 2007; Hart, 1998; Neuman, 2011; Schwarz et al. 2007; Webster and Watson, 2002). Only after a fair understanding of the research area is developed are researchers able to identify shortcomings in current research. This then allows the formulation of meaningful and relevant research questions. SLRs, however, put this process on its head. As a result, SLRs are to be preceded by a more general literature review 'scoping' the literature (Denyer and Tranfield, 2009), thus showing that they cannot offer a general means for conducting literature reviews. This does not deny the fact that researchers initially may have some preliminary research questions before they engage with the literature. Nevertheless, such questions are likely to be amended and changed in the light of the literature and the contributions made by earlier research. However, the protocol driven approach of SLR does not allow a question to be amended at a later point in the light of what is learned during the literature review process. As MacLure (2005) puts it, "diversions into unanticipated areas are not encouraged ... learning from adjacent areas is not recommended either" (p. 399).

Moreover, the type of research question that can be answered by SLRs is limited to closed research questions that can be answered from literature unambiguously by summarizing evidence. However, questions of 'how' and 'why' which are fundamental to academic endeavours, especially in the social sciences (DiMaggio, 1995; Sutton and Staw, 1995; Whetten, 1989), are not considered. SLRs are ill equipped to produce rich answers to these types of questions as the methodology requires the provision of clear boundaries of what is covered through inclusion criteria and

Table 2 Criteria for good literature review practice

Criteria	*References*
Comprehensiveness of literature, breadths and depths of understanding - Literature review process is an understanding process that involves searching for literature, selecting, reading, comparing, classifying, and critically assessing earlier research, leading to the creation of a body of literature relevant for a phenomenon studied; literature reviews are assessed based on the comprehensiveness of and insight into the body of literature analyzed and breadth and depths of its understanding.	Boote and Beil, 2005; Hart, 1998; Perry, 1998; Schwarz et al., 2007; Webster and Watson, 2002
Argument development - The aim of literature reviews is to develop a solid argument by assessing prior research and contributions to knowledge of specific papers/streams/approaches, and based on that, identify weaknesses, under-researched phenomena or research gaps that warrant further research.	Feak and Swales 2009; Kwan et al., 2012; Levy and Ellis, 2006; Machi and McEvoy, 2012; Ridley, 2008
Ongoing engagement - Literature reviews are complex iterative processes through which a researcher acquires ever increasing understanding of the relevant literature and a potential to contribute to the literature; with broader and deeper understanding of literature, a researcher's ability to assesses relevance and value of individual studies, streams of research, and different domains of literature are increasing; literature reviews are therefore integrative to the complete research process, as engagement with the literature informs all stages of research, up until the final write up.	Combs et al., 2010; Dellinger, 2005; Dong, 1996; Goodfellow, 1998; Kwan, 2008; Onwuegbuzie et al., 2007; Ridley, 2008; Wolfswinkel, 2013
Criticality - Central to the quality of literature reviews is the critical engagement with knowledge claims made by earlier research and the assumptions underlying these knowledge claims; high quality literature reviews are characterized by criticality of assessment of existing knowledge about a target phenomenon, including contributions of specific approaches and streams of research.	Alvesson and Sandberg, 2011; Finn, 2005; MISQ, 2006; Ridley, 2008
Originality and innovative views - Literature reviews are judged by the novelty in approaching and investigating earlier research that brings out a researcher's voice. The aim is the generation of new insights, perspectives and understanding by problematizing knowledge claims, approaches and assumptions in earlier research.	Alvesson and Sandberg, 2011; Green et al. 2006; Hart, 1998; Khoo et al., 2011; MacLure, 2005

Preprint of: Boell, Sebastian K; &Cecez-Kecmanovic, Dubravka (2015) On Being 'Systematic' in Literature Reviews in IS. *Journal of Information Technology*. 30(2). 161-173. The definitive publisher-authenticated version is available: http://dx.doi.org/10.1057/jit.2014.26

what is ruledout by exclusion criteria. While this willinclude literature that explicitly addressed this question before, it does not cover literature that only engages with 'why' and 'how' questions implicitly. This will exclude much of the literature which will require adaptation of solutions from other contexts in creative, innovative, and new ways. Finfgeld-Connett and Johnson (2013) thus argue that strict protocol driven SLRs are not well suited for theory generating and knowledge building reviews.

What makes SLRs particularly risky is their adoption as primary means for training novice researchers in conducting literature reviews. More experienced researchers would not be easily persuaded by the claims to quality of SLRs. However, novice researchers are likely to be attracted by the simplicity and certainty of the SLR protocol and claims to quality. Importantly, SLRs are not directing or requiring novice researchers to read critically and engage independently and creatively within a wider body of knowledge, abstracting ideas from at first seemingly unrelated areas. From this perspective we agree with several authors who consider SLRs a potential threat to critical thinking and scholarship (Holmes et al., 2008; MacLure, 2005; Sandelowski et al., 2007). This raises the concern that promoting SLRs as a general means for conducting literature reviews may in the long run undermine the pursuit of scholarship in academia and in particular negatively affect the education of young researchers, their development as critical thinkers and their learning to become 'scholarly in their work' (Hart, 1998). In this sense, SLRs may create more serious problems for academia than they propose to solve.

Usefulness and Limitations of SLR

While we raise numerous questions regarding SLR adoption as a general approach to literature review, we would also like to draw attention to the conditions under which SLR as a very specific kind of literature reviews can be useful. Similar to their use in medicine, SLRs may be useful and justified for practitioners aiming to answer a narrow and unequivocal question, for instance, to find out whether the applications of a particular tool, method, process or software product is associated with a specific outcome across different studies. In addition, the application of SLRs in a research context can be justified when a research question is narrowly defined so that it is answerable by extracting and aggregating evidence from studies in a summative form. It is important to recognize that in both cases – practitioners' searches for evidence and scholarly work – the application of SLR is conditional upon:

1. A clearly delimited topic, described by highly discriminating terms and a research question that is very specific and closely formulated so that a) it is potentially possible to identify *all* literature addressing *this particular topic or question* and b) research questions can be answered by extracting evidence regarding particular relations and their strength (e.g. studies involving one or more variables hypothesizing a causal relation "If X is applied then Y will occur").
2. Systematic mapping of literature for a high level bibliometric type analysis interested in identifying key characteristics of the literature, such as number of publications over journals and over time, productivity of authors and countries, used research methods, etc. (e.g. Weerakkody et al., 2009; Williams et al, 2009).

These conditions make SLRs suitable *only* for a very specific kind of literature reviews. While SLRs offer a means for summarizing evidence from literature regarding a particular and a well-specified question, many SLR proponents and published literature reviews that adopt SLRs do not discuss these conditions and limitations (e.g. Mohan and Ahleman, 2011; Okoli and Schabram, 2009, 2010) and imply SLR's general applicability.

To illustrate the adoption of SLRs in IS and how their practice can be improved we looked into the full text of all identified 101 journal articles (included in Figure 1) and selected 14 that explicitly claim they adopted SLR. We examined whether articles discuss conditions for the application and limitations of SLRs, whether they report search details, selection criteria, use of guidelines and protocols and also whether they describe SLRs and make claims to quality. The summary of the analysis of these 14 journal articles is presented in Table 3.

Looking at the adoption of SLR in the selected articles it seems that the label SLR is currently used without a clear understanding of what SLRs entail and what the conditions for their application are. A common feature among all 14 articles is that they describe the search procedure and the literature selection process, which is frequently linked to claims about quality of the resulting review as being 'less biased', 'scientific', 'well guided', 'reproducible', 'comprehensive', 'rigorous', and 'reliable' (Table 3). While these articles label themselves as SLR, more than half of them do not define what SLRs entail. Instead, Grahlmann et al. (2012), Gräning et al. (2011), Guillemette and Paré (2012), Hummel et al. (2013) and Mettler et al. (2014)all refer to Webster and Watson (2002) as presumably offering SLR guidelines (which they are not). Furthermore, studies that do describe SLRs are often vague in their description, with only one SLR including the use of a protocol, arguably the most central part for a SLR. Instead, by implication articles assert that if database searches and literature selection procedures are reported, resulting reviews will be SLRs that are of higher quality than traditional narrative literature reviews.

Articles 1-12 report the conduct of some form of SLR that would imply the fulfilment of condition 1, that is, a clearly delimited topic, described by highly discriminating terms and a research question that is very specific and closely formulated so as to be answerable by extracting evidence from selected publications. The articles however do not meet this condition,they do not reflect or discuss the nature or limitations of SLR and assume SLR's general applicability. Consider for instance the research questions

Preprint of: Boell, Sebastian K; &Cecez-Kecmanovic, Dubravka (2015) On Being 'Systematic' in Literature Reviews in IS. *Journal of Information Technology*. 30(2). 161-173. The definitive publisher-authenticated version is available: http://dx.doi.org/10.1057/jit.2014.26

Table 3 Overview of SLR articles in IS journals

	Reports search details	*Uses selection criteria*	*Cites SLR guide-lines*	*Des-cribes SLR*	*Uses pro-tocol*	*Claims to quality*
1 - Amrollahi et al., 2013	✓	✓	✓	✓		
2 –Basten&Sunyaev, 2014	✓	✓	✓		✓	reliable; reproducible
3 –Goeken&Patas, 2010	✓	✓	✓	✓		scientific; evidence based
4 - Grahlmann et al., 2012	✓	✓	✓			
5 - Gräning et al., 2011	✓	✓				completeness; scientific
6 –Guillemette&Paré, 2012	✓	✓				comprehensive
7 - Hummel et al., 2013	✓	✓				structured
8 - Leite&Cappelli, 2010	✓	✓	✓	✓		guided by well established criteria
9 –Merschbrock&Munkvold, 2012	✓	✓				
10 - Mettler et al., 2014	✓	✓		✓		explicit; rigorous; transparent
11 - Tamm et al., 2011	✓	✓				objective; less biased; comprehensive
12 - Wang et al., 2013	✓	✓		✓		
13 - Weerakkody et al., 2009	✓	✓				comprehensive; current state of play
14 - Williams et al., 2009	✓	✓				comprehensive; current state of play

in articles no 2, 7, 8, 11 (in Table 3) respectively: "What evidence exists regarding the factors that affect the accuracy of software development effort estimation?" (BastenandSunyaev, 2014:52); "What is the role of communication in SD [systems development] projects employing agile practices?" (Hummel et al., 2013:344); "[H]ow should we build software systems supporting the demand for transparency?" (Leite and Cappelli, 2010:128); "How does EA [Enterprise Architecture] lead to organisational benefits?" and related sub-questions: "What are the potential benefits of EA? What are the mechanisms through which EA leads to these benefits? What are the theoretical underpinnings, and empirical evidence, that support these expectations?" (Tamm et al., 2011:142). These questions are neither narrow enough nor precise enough to be answerable by collecting and summarizing particular evidence as SLR implies. They are to a larger or lesser extent too broad to meet the condition 1 to justify SLRs.

In addition, reviews also set themselves research objectives that are suitable for general literature reviews rather than SLRs. For instance: "[t]he objective of this paper is to develop an approach for the structuring of empirical RE [Requirements engineering] research and to undertake an evaluation of existing work," (Goeken and Patas, 2010:182). Another example from article 5: "The paper provides a comprehensive and structured overview of current XBRL research …" (Gräning et al., 2011:231) - shows a scope much beyond SLR's limitations. The research question in article 12 seems potentially suitable to a SLR: "what are the factors that have a direct and indirect impact on digital fluency?" (Wang et al., 2013:1). However 'digital fluency' is not a well established concept in the literature as the authors themselves acknowledge. They defined it as "the ability to reformulate knowledge to express oneself creatively and appropriately, and to produce and generate information rather than simply to comprehend it" (p. 2). The issue is that this definition is too broad to make the topic precisely defined and researchable by means of SLR.

These examples indicate that selected SLRs from the IS journals listed in Table 3 fail to acknowledge the conditions for using SLR and instead use them unreflectively, as a general approach to conducting literature reviews. Furthermore, some articles misattributed SLRs, conducting a traditional literature review but calling it a 'systematic review' without any further discussion of what they meant by this (e.g. Guillemetteand Paré, 2012; Merschbrockand Munkvold, 2012).

Articles 13 and 14 are different from the rest in Table 3, for they fulfil condition 2 necessary for SLR. Both articles focus on a particular topic conducting a bibliometric analysis of the articles they identified. They focus in their analysis on general bibliometric characteristics of the literature, using*Web of Science* as data source, a database frequently used for bibliometric analysis.

Based on our analysis we suggest that the use of SLRs in IS can be improved by clearly communicating what they are good for, and what their limitations are. From the 14 articles we reviewed only five that clearly stated their research question. We therefore also suggest that SLR articles need to justify upfront why SLR is adopted and how the conditions (such as those defined above) are met. The justification should include a clearly stated research question that the review is addressing, allowing readers to judge if the question is sufficiently narrow and precise to be answerable by a SLR. Furthermore, authors need to clearly describe their literature review approach and what it entails.

Preprint of: Boell, Sebastian K; &Cecez-Kecmanovic, Dubravka (2015) On Being 'Systematic' in Literature Reviews in IS. *Journal of Information Technology*. 30(2). 161-173. The definitive publisher-authenticated version is available: http://dx.doi.org/10.1057/jit.2014.26

In particular they need to include a protocol (as an appendix) clearly stating how literature was selected for the review to enable readers to judge the exact selection process and inclusion and exclusion criteria used by a SLR. We also noted that all but four SLRsin Table 3 fail to disclose the list of references included in the review thus limiting the potential of others to build on the review. As the literature selection process is particularly laborious and in large part dependent on interpretation and judgement of the reviewers, SLRs should provide a clear listing of the publications covered by the review. Finally, when conditions for their adoption are met SLRs as a specific type of literature review may be useful, but without grand and, as we have seen, unjustified claims to quality. Like any other literature reviews, SLRs have to justify any claim to quality by clearly demonstrating how it was achieved.

Conclusion

This paper engages in a debate about the use and limitation of SLRs for literature reviews in IS and the social sciences. We argue that SLR can have a place in research, but that claims that SLRs provide a general and superior approach to literature review are highly questionable. The conditions for applying SLRs make them suitable only for a very specific kind of review. However, the papers in IS that adopted SLR as a general approach to literature review do not recognize or reflectively consider these conditions and limitations. When SLR guidelines are adopted for general literature review they endorse an impoverished notion of literature review by degrading them to repeatable literature searches and selection processes. In addition, we provide evidence showing that claims that SLRs ensure 'objective', 'replicable', 'unbiased', 'scientific' and rigorous' reviews do not hold up. When claimed as inherent to SLRs, these quality criteria portray 'other' literature reviews as unscientific, biased, non-transparent and non-rigorous. If SLR quality criteria are taken for granted it may be hard to defend 'traditional' or 'narrative' literature reviews. This is potentially damaging for scholarly assessment of knowledge and literature reviews in the social sciences where in-depth analysis, criticality, imagination, and creativity are desirable qualities, particularly in IS where diversity in research topics and research methods is common.

To prevent misunderstanding, we would like to clarify that we believe literature reviews should be systematic inquiries, but this should not be confused with the prescriptions and protocols of SLRs. We also believe that SLRs do have a role in specific kinds of literature reviews as discussed above. Nevertheless, when introduced as part of research training, making novice researchers aware of these limitations is important. To propose that the quality of a literature review is determined or guaranteed by any method or technique (SLR included) used to search literature and conduct a review cannot be justified. This disregards the intellectual work and a very personal dialogue and critical engagement with the literature by a researcher.

There is no one objectively correct understanding of earlier research and knowledge contributions (Kuhn, 1962). There is therefore no single best way of looking at the literature and for creating literature reviews (Schwarz et al., 2007). Instead, diversity of views and outcomes of literature reviews is desirable and important for scholarly argument and constructive debates (Hart, 1998). Literature reviews do need to be systematic – as discussed by many sources on literature reviews (e.g. Hart, 1998; Levy and Ellis, 2006; Machi and McEvoy, 2009; Schwarz et al., 2007; Webster and Watson, 2002) – in order to provide a reasonably comprehensive account. Being systematic enables critical assessment of knowledge on a particular topic in order to convincingly demonstrate weaknesses and inadequacies in the current body of knowledge and potentially problematize its foundations. Being systematic in a literature review will not only help in covering the breadth of relevant works, but also, and importantly, in providing the depth of insight into a body of literature that will allow the production of competent, original and critical reviews. Developing a distinct and original perspective on a literature and arguing for its relevance and value is a quality of literature reviews that should not be seen as something that needs to be eradicated, but as something desirable that needs to be embraced (Hart, 1998).

Notes

1 On 22nd February 2014 we searched for the phrases "Systematic Literature Review" and "Systematic Review" (SR) in the *AIS eLibrary* and in *Scopus*. Our search specifically included all search fields and no limitation regarding publication date. However, in Scopus we limited our search to articles published in the AIS basket of 8 journals: *European Journal of Information Systems, Information Systems Journal, Information Systems Research, Journal of Information Technology, Journal of Management Systems, Journal of Strategic Management Systems, Journal of the AIS,* and *MIS Quarterly*. Our search returned 498 results from the AIS eLibrary and 42 from Scopus, with only 3 publications overlapping.

2 The search was conducted in February 2014, potentially returning incomplete results for journal publications from 2013.

References

Agrawal, A., Boese, M. and Sarker, S. (2010). A Review of the HCI Literature in IS: The Missing Links of Computer-mediated Communication, Culture, and Interaction, In *AMCIS 2010 Proceedings*, Paper 523.

Alvesson, M. and Sandberg, J. (2011). Generating Research Questions Through Problematization, *Academy of Management Review* 36(2): 247–271.

Amrollahi, A., Ghapanchi, A. H. and Talaei-Khoei, A. (2013). A Systematic Literature Review on Strategic Information Systems Planning: Insights from the Past Decade, *Pacific Asia Journal of the Association for Information Systems* 5(2): 39–66.

Atkins, C. and Louw, G. (2000). Reclaiming Knowledge: A Case for Evidence Based Information Systems Reclaiming Knowledge: A Case for Evidence-Based Information Systems, In *ECIS 2000*, Paper 28.

Bandara, W., Miskon, S. and Fielt, E. (2011). A Systematic, Tool-Supported Method for Conducting Literature Reviews in Information Systems, In *ECIS 2011 Proceedings*, Paper 221.

Basten, D. and Sunyaev, A. (2011). The Nature of Adherence to Planning as Criterion for Information System Project Success,*Sprouts: Working Papers on Information Systems* 11(142).

Basten, D. and Sunyaev, A. (2014). A Systematic Mapping of Factors Affecting Accuracy of Software Development Effort Estimation, *Communications of the Association for Information Systems* 34: 51–86.

Preprint of: Boell, Sebastian K; &Cecez-Kecmanovic, Dubravka (2015) On Being 'Systematic' in Literature Reviews in IS. *Journal of Information Technology*. 30(2). 161-173. The definitive publisher-authenticated version is available: http://dx.doi.org/10.1057/jit.2014.26

Baumeister, R. E. and Leary, M. R. (1997). Writing Narrative Literature Reviews, *Review of General Psychology* 1(3): 311–320.

Blair, D. (2006). *Wittgenstein, Language and Information. Back to the Rough Ground!*, Dordrecht: Springer.

Bodoff, D. (2009). Emergence of Terminological Conventions as a Searcher-indexer Coordination Game, *Journal of the American Society for Information Science and Technology* 60(12): 2509–2529.

Boell, S. K. and Cecez-Kecmanovic, D. (2014). A Hermeneutic Approach for Conducting Literature Reviews and Literature Searches, *Communications of the Association for Information Systems* 34: 257–286.

Boote, D. N. and Beile, P. (2005). Scholars Before Researchers: On the Centrality of the Dissertation Literature Review in Research Preparation, *Educational Researcher* 34(6): 3–15.

Brocke, J., Simons, A., Niehaves, B., Reimer, K., Plattfaut, R. and Cleven, A. (2009). Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process, In *ECIS 2009 Proceedings*, Paper 161.

Buckland, M. and Gey, F. (1994). The Relationship Between Recall and Precision, *Journal of the American Society for Information Science* 45(1): 12–19.

Campbell Collaboration (2007). Retrieved on 10 May 2014 from http://www.campbellcollaboration.org

Chalmers, I. and Altman, D. G. (1995). *Systematic reviews*, London: BMJ.

Checkland, P. and Holwell, S. (1998). *Information, Systems and Information Systems: Making Sense of the Field*, Chichester: Wiley.

Cheung, C. M. K. and Thadani, D. R. (2010). The Effectiveness of Electronic Word-of-Mouth Communication: A Literature Analysis, In *BLED 2010 Proceedings*, Paper 18.

Clarke, M. J. and Stewart, L. A. (1995). Obtaining Data from Randomized Controlled Trials: How Much Do We Need for Reliable and Informative Meta-analyses?, In *Systematic Reviews*, London: BMJ, pp. 37–47.

Combs, J. P., Bustamante, R. M. and Onwuegbuzie, A. J. (2010).An Interactive Model for Facilitating Development of Literature Reviews, *International Journal of Multiple Research Approaches* 4(2): 159–182.

Constantinides, P., Chiasson, M. W. and Introna, L. D. (2012). The Ends of Information Systems Research: A Pragmatic Framework, *MIS Quarterly* 36(1): 1–19.

Cruzes, D. S. and Dybå, T. (2011). Research Synthesis in Software Engineering: A Tertiary Study, *Information and Software Technology* 53(5): 440–455.

Davies, W. M. and Beaumont, T. J. (2007). *Literature Reviews*, *Business*, Melbourne.

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology, *MIS Quarterly* 13(3): 319–339.

Dellinger, A. B. (2005). Validity and the Review of the Literature, *Research in the Schools* 12(2): 41–54.

Denyer, D. and Tranfield, D. (2009). Producing a Systematic Review, In D. A. Buchanan & A. Bryman (Eds.), *The Sage Handbook of Organizational Research Methods*, Thousand Oaks, CA, USA: Sage, pp. 671–689.

DiMaggio, P. J. (1995). Comments on "What Theory is Not", *Administrative Science Quarterly* 40(3): 391–397.

Dong, Y. R. (1996). Learning How to Use Citations for Knowledge Transformation: Non-Native Doctoral Students' Dissertation Writing in Science, *Research in Teaching of English* 30(4): 428–457.

Dwivedi, Y., Williams, M. D., Lal, B. and Schwarz, A. (2008). Profiling Adoption, Acceptance and Diffusion Research in the Information Systems Discipline, In *ECIS 2008 Proceedings*, Paper 112.

Evidence-based Medicine Working Group (1992). Evidence-based Medicine: A New Approach to Teaching the Practice of Medicine, *Journal of the American Medical Association* 268(17): 2420–2425.

Eysenck, H. J. (1995). Problems with Meta-Analysis, In I. Chalmers & D. G. Altman (Eds.), *Systematic Reviews*, London: BMJ, pp. 64–74.

Feak, C. B. and Swales, J. M. (2009). *Telling a Research Story: Writing a Literature Review*, Ann Arbor, Mich: University of Michigan Press.

Finfgeld-Connett, D. and Johnson, E. D. (2013).Literature Search Strategies for Conducting Knowledge-Building and Theory-Generating Qualitative Systematic Reviews, *Journal of Advanced Nursing* 69(1): 194–204.

Finn, J. A. (2005). *Getting a PhD: An Action Plan to Help Manage Your Research, Your Supervisor and Your Project*, London: Routledge.

Fugmann, R. (2007). Informationstheorie: Der Jahrhundertbluff. Eine-ZeitkritischeBetrachtung (Teil 1), *Information Wissenschaft und Praxis* 58(8): 449–458.

Gacenga, F., Cater-Steel, A., Toleman, M. and Tan, W.-G. (2011). Measuring the Performance of Service Orientated IT Management, *Sprouts: Working Papers on Information Systems* 11(162).

Goodfellow, J. (1998). *Constructing a Narrative*, In J. Higgs (Ed.), Writing Qualitative Research, Sydney: Hampden Press, pp. 175–187.

Grahlmann, K. R., Helms, R. W., Hilhorst, C., Brinkkemper, S. and Van Amerongen, S. (2012). Reviewing Enterprise Content Management: A Functional Framework, *European Journal of Information Systems* 21(3): 268–286.

Gräning, A., Felden, C. and Piechocki, M. (2011). Status Quo and Potential of XBRL for Business and Information Systems Engineering, *Business and Information Systems Engineering* 3(4): 231–239.

Green, B. N., Johnson, C. D. and Adams, A. (2006). Writing Narrative Literature Reviews for Peer-reviewed Journals: Secrets of the Trade, *Journal of Chiropractic Medicine* 5(3): 101–117.

Greenhalgh, T. and Peacock, R. (2005). Effectiveness and Efficiency of Search Methods in Systematic Reviews of Complex Evidence: Audit of Primary Sources, *British Medical Journal* 331(7524): 1064–1065.

Guillemette, M. G. and Paré, G. (2012). Toward a New Theory of the Contribution of the IT Function in Organizations, *MIS Quarterly* 36(2): 529–551.

Hammersley, M. (2001). On "Systematic" Reviews of Research Literatures: A "Narrative" Response to Evans &Benefield, *British Educational Research Journal* 27(5): 543–554.

Hart, C. (1998). *Doing a Literature Review. Releasing the Social Science Research Imagination*, Thousand Oaks: SAGE Publications.

Hartley, J. and Betts, L. (2009). Common Weaknesses in Traditional Abstracts in the Social Sciences, *Journal of the American Society for Information Science and Technology* 60(10): 2010–2018.

Hjorland, B. (2011). Evidence-Based Practice: An Analysis Based on the Philosophy of Science, *Journal of the American Society for Information Science* 62(7): 1301–1310.

Holmes, D., Murray, S. J., Perron, A. and McCabe, J. (2008). Nursing Best Practice Guidelines: Reflecting on the Obscene Rise of the Void, *Journal of Nursing Management* 16(4): 394–403.

Hood, W. W. and Wilson, C. S. (2001). The Scatter of Documents Over Databases in Different Subject Domains: How Many Databases Are Needed?,*Journal of the American Society for Information Science and Technology* 52(14): 1242–1254.

Hummel, M., Rosenkranz, C. and Holten, R. (2012). The Role of Communication in Agile Systems Development: An Analysis of the State of the Art, *Business and Information Systems Engineering* 5(5): 343–355.

Jalali, S. and Wohlin, C. (2012). Systematic Literature Studies: Database Searches vs. Backward Snowballing, In *International Conference on Empirical Software Engineering and Measurement*, Lund, Sweden: ACM, 29–38.

Khoo, C. S. G., Na, J.-C. and Jaidka, K. (2011). Analysis of the Macro-level Discourse Structure of Literature Reviews, *Online Information Review* 35(2): 255–271.

Kitchenham, B. (2004). *Procedures for Performing Systematic Reviews*, Keele, Eversleigh: Keele University and NICTA, Technical Report.

Kitchenham, B. and Charters, S. (2007). Guidelines for Performing Systematic Literature Reviews in Software Engineering. Retrieved on 10 May 2014 from http://www.dur.ac.uk/ebse/resources/guidelines/Systematic-reviews-5-8.pdf

Knipschild, P. (1995). Some Examples of Systematic Reviews, In I. Chalmers & D. G. Altman (Eds.), *Systematic Reviews*, London: BMJ, pp. 9–16.

Kuhn, T. S. 1962.*The Structure of Scientific Revolutions*, Chicago: University of Chicago Press.

Kwan, B. S. C. (2008). The Nexus of Reading, Writing and Researching in the Doctoral Undertaking of Humanities and Social Sciences: Implications for Literature Reviewing, *English for Specific Purposes* 27(1): 42–56.

Kwan, B. S. C., Chan, H. and Lam, C. (2012). Evaluating Prior Scholarship in Literature Reviews of Research Articles: A Comparative Study

Preprint of: Boell, Sebastian K; &Cecez-Kecmanovic, Dubravka (2015) On Being 'Systematic' in Literature Reviews in IS. *Journal of Information Technology*. 30(2). 161-173. The definitive publisher-authenticated version is available: http://dx.doi.org/10.1057/jit.2014.26

of Practices in Two Research Paradigms, *English for Specific Purposes* 31(3): 188–201.

Lacity, M. C., Solomon, S., Yan, A. and Willcocks, L. P. (2011). Business Process Outsourcing Studies: A Critical Review and Research Directions, *Journal of Information Technology* 26(4): 221–258.

Leonardi, P. M. and Barley, S. R. (2010). What's Under Construction Here? Social Action, Materiality, and Power in Constructivist Studies of Technology and Organizing, *The Academy of Management Annals* 4(1): 1–51.

Leite, J. C. S. do P. and Cappelli, C. (2010). Software Transparency, *Business and Information Systems Engineering* 2(3): 127–139.

LePine, J. A. and Wilcox-King, A. (2010). Editors' Comments: Developing Novel Theoretical Insight from Reviews of Existing Theory and Research, *Academy of Management Review* 35(4): 506–509.

Levy, Y. and Ellis, T. J. (2006). A Systems Approach to Conduct an Effective Literature Review in Support of Information Systems Research, *Informing Science Journal* 9: 181–212.

Machi, L. A. and McEvoy, B. T. (2009). *The Literature Review: Six Steps To Success*, Thousand Oaks: Corwin Press.

MacLure, M. (2005). "Clarity Bordering on Stupidity": Where's the Quality in Systematic Review?,*Journal of Education Policy* 20(4): 393–416.

Matthias, G. and Patas, J. (2010). Evidence-Based Structuring and Evaluation of Empirical Research in Requirements Engineering: Fundamentals, Framework, Research Map, *Business and Information Systems Engineering* 2(3): 175–185.

Merschbrock, C. and Munkvold, B. E. (2012). A Research Review on Building Information Modeling in Construction — An Area Ripe for IS Research, *Communications of the Association for Information Systems* 31: 307–228.

Mettler, T., Eurich, M. and Winter, R. (2014). On the Use of Experiments in Design Science Research: A Proposition of an Evaluation Framework, *Communications of the Association for Information Systems* 34: 223–240.

MISQ (2006).*Objectives of the MISQ Theory and Review*. Retrieved on 10 May 2014 from http://www.misq.org/skin/frontend/default/misq/pdf/TheoryReview/TRObjectives.pdf

Mohan, K. and Ahlemann, F. (2011). Understanding Acceptance of Information System Development and Management Methodologies by actual Users: A Review and Assessment of Existing Literature, In *Wirtschaftsinformatik Proceedings 2011*, Paper 41.

Morrell, K. (2008). The Narrative of "Evidence Based" Management: A Polemic, *Journal of Management Studies* 45(3): 613–635.

Mulrow, C. D. (1995). Rationale for Systematic Reviews, In I. Chalmers & D. G. Altman (Eds.), *Systematic Reviews*, London: BMJ, pp. 1–8.

Murray, S. J., Holmes, D., Perron, A. and Rail, G. (2007). No Exit? Intellectual Integrity Under the Regime of "Evidence" and "Best-practices", *Journal of Clinical Practice* 13: 512–516.

Neuman, W. L. (2011).*Social Research Methods: Quantitative and Qualitative Approaches*, 7th ed., Boston, Mass, USA: Pearson.

Oancea, A. and Pring, R. (2008). The Importance of Being Thorough: On Systematic Accumulations of "What Works" in Education Research, *Journal of Philosophy of Education* 42: 15–39.

Oates, B. (2011). Evidence Based Information Systems: A Decade Later, In *ECIS 2011 Proceedings*, Paper 222.

Oates, B. J., Edwards, H. M. and Wainwright, D. W. (2012). A Model-Driven Method for the Systematic Literature Review of Qualitative Empirical Research, In *ICIS 2012 Proceedings*, pp. 1–18.

Okoli, C. and Schabram, K. (2009). Protocol for a Systematic Literature Review of Research on the Wikipedia, *Sprouts: Working Papers on Information Systems* 9(65).

Okoli, C. and Schabram, K. (2010). A Guide to Conducting a Systematic Literature Review of Information Systems Research,*Sprouts: Working Papers on Information Systems* 10(26).

Onwuegbuzie, Anthony J Collins, K. M. T., Leech, N. L., Dellinger, A. B. and Jiao, Q. G. (2007). Mixed Methods + Literature Reviews = Mixed Research Syntheses: A Framework for Conducting and Writing Rigorous, Comprehensive, and Insightful Literature Reviews, In *The World of Educational Quality. Annual Meeting of the American Educational Research Association*, Chicago.

Oxman, A. D. (1995). Checklists for Review Articles, In *Systematic Reviews*, London: BMJ, pp. 75–85.

Pawson, R. (2006). Digging for Nuggets: How "Bad" Research Can Yield "Good" Evidence, *International Journal of Social Research Methodology* 9(2): 127–142.

Perry, C. (1998). A Structured Approach for Presenting Theses, *Australasian Marketing Journal* 6(1): 63–85.

Petticrew, M. and Roberts, H. (2006). *Systematic Reviews in the Social Sciences. A Practical Guide*, Malden: Blackwell.

Ramsay, C. R., Grant, A. M., Wallace, S. A., Garthwaite, P. H., Monk, A. F. and Russell, I. T. (2000).Assessment of the Learning Curve in Health Technologies - A Systematic Review, *International Journal of Technology Assessment in Health Care* 16(4): 1095–1108.

Ridley, D. (2008). *The Literature Review. A Step-by-Step Guide for Students*, Thousand Oaks: SAGE.

Rose, S., Bisson, J., Wessely, S. (2003). A Systematic Review of Single-Session Psychological Interventions ('Debriefing') Following Trauma, *Psychotherapy and Psychosomatics* 72(4): 176–184.

Roztocki, N. and Weistroffer, H. R. (2008). Event Studies in Information Systems Research: A Review, In *AMCIS 2008 Proceedings*, Paper 248.

Salton, G. and McGill, M. J. (1983). *Introduction to Modern Information Retrieval*, New York: McGraw-Hill.

Sandelowski, M. (2008). Reading, Writing and Systematic Review, *Journal of advanced nursing* 64(1): 104–10.

Sandelowski, M., Voils, C. I. and Barroso, J. (2007). Comparability Work and the Management of Difference in Research Synthesis Studies, *Social Science& Medicine (1982)* 64(1): 236–47.

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

Schwarz, A., Mehta, M., Johnson, N. and Chin, W. W. (2007). Understanding Frameworks and Reviews: A Commentary to Assist us in Moving Our Field Forward by Analyzing Our Past, *DATA BASE for Advances in Information Systems* 38(3): 29–50.

Shiffman, R. N., Liaw, Y., Brandt, C. A. and Corb, G. J. (1999). Computer-based Guideline Implementation Systems: A Systematic Review of Functionality and Effectiveness, *Journal of the American Medical Informatics Association* 6(2): 104–114.

Staples, M. and Niazi, M. (2007).Experiences Using Systematic Review Guidelines, *Journal of Systems and Software* 80(9): 1425–1437.

Stol, K.-J., Babar, M. A., Russo, B. and Fitzgerald, B. (2009). The Use of Empirical Methods in Open Source Software Research: Facts, Trends and Future Directions, In *FLOSS'09, May 18, 2009, Vancouver, Canada*, pp. 19–24.

Sutton, R. and Staw, B. M. (1995). Forum What Theory is Not, *Administrative Science Quarterly* 40(3): 371–384.

Tamm, T., Seddon, P. B., Shanks, G. and Reynolds, P. (2011). How Does Enterprise Architecture Add Value to Organisations?,*Communications of the Association for Information Systems* 28: 141–168.

Thompson, S. G. (1995). Why Sources of Heterogenity in Meta-Analysis Should be Investigated, In I. Chalmers & D. G. Altman (Eds.), *Systematic Reviews*, London: BMJ, pp. 48–63.

Tranfield, D., Denyer, D. and Smart, P. (2003). Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review, *British Journal of Management* 14(3): 207–222.

Wang, Q. E., Myers, M. D. and Sundaram, D. (2013). Digital Natives and Digital Immigrants: Towards a Model of Digital Fluency, *Business and Information Systems Engineering* 5(6): 409–419.

Webster, J. and Watson, R. T. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review, *MIS Quarterly* 26(2): xiii–xxiii.

Whetten, D. A. (1989). What Constitutes a Theoretical Contribution?,*Academy of Management Review* 14(4): 490–495.

Williams, M. D., Dwivedi, Y. K., Lal, B. and Schwarz, A. (2009). Contemporary Trends and Issues in IT Adoption and Diffusion research, *Journal of Information Technology* 24(1): 1–10.

Wolfswinkel, J. F., Furtmueller, E. andWilderom, C. P. M. (2013). Using Grounded Theory as a Method for Rigorously Reviewing Literature, *European Journal of Information Systems* 22(1): 45–55.

Wolfswinkel, J., Furtmueller, E. andWilderom, C. (2010). Reflecting on E-Recruiting Research Using Grounded Theory, In *ECIS 2010 Proceedings*, Paper 52.

Wright Mills, C. (1978[1959]) *The Sociological Imagination*, Oxford: Oxford University Press.

Preprint of: Boell, Sebastian K; &Cecez-Kecmanovic, Dubravka (2015) On Being 'Systematic' in Literature Reviews in IS. *Journal of Information Technology*. 30(2). 161-173. The definitive publisher-authenticated version is available: http://dx.doi.org/10.1057/jit.2014.26

Zhang, P., Li, N. (Lina), Scialdone, M. J. and Carey, J. (2008). The Intellectual Advancement of Human-Computer Interaction Research: A Critical Assessment of the MIS Literature (1990-2008), *AIS Transactions on Human-Computer Interaction* 1(3): 55–107.

About the Authors

Sebastian K Boellis Lecturer in Business Information Systems at the University of Sydney Business School. Sebastian has a PhD in Information Systems (IS) from the University of New South Wales (UNSW) in Australia and a MA in Information Science form Saarland University in Germany. His PhD dissertation was awarded the *ASIST ProQuest Doctoral Dissertation Award*. Sebastian's research encompasses conceptualizations of 'information' and 'information systems'. In particular, he is interested in advancing theoretical engagement with these concepts in IS research, practice and education. Sebastian has also worked on research in the areas of information retrieval, bibliometrics, scientometrics, and on telework and telecommuting. http://skb.unifind.de

Dubravka Cecez-Kecmanovic is professor of Information Systems at the UNSW BusinessSchool. Her research interests include theoretical and methodological developments in understanding information systems (IS) and organizing. She has published in *MIS Quarterly,Information Systems Journal, European Journal of Information Systems, Information Technology and People, Journal of Information Technology, Communication of AIS, Information and Management, Decision Support Systems, International Journal of Actor Network Theory and Technological Innovation*, and others. She is Senior Editor of Journal of AIS and Associate Editor of Information Systems Journal.

Preprint of: Boell, Sebastian K; &Cecez-Kecmanovic, Dubravka (2015) On Being 'Systematic' in Literature Reviews in IS. *Journal of Information Technology*. 30(2). 161-173. The definitive publisher-authenticated version is available: http://dx.doi.org/10.1057/jit.2014.26