

Association for Information Systems AIS Electronic Library (AISeL)

PACIS 2014 Proceedings

Pacific Asia Conference on Information Systems
(PACIS)

2014

A CONTENT ANALYSIS OF CONTENT ANALYSES IN IS RESEARCH: PURPOSES, DATA SOURCES, AND METHODOLOGICAL CHARACTERISTICS

André Coners

South Westphalia University of Applied Sciences, coners.andre@fh-swf.de

Benjamin Matthies

South Westphalia University of Applied Sciences, matthies.benjamin@fh-swf.de

Follow this and additional works at: <http://aisel.aisnet.org/pacis2014>

Recommended Citation

Coners, André and Matthies, Benjamin, "A CONTENT ANALYSIS OF CONTENT ANALYSES IN IS RESEARCH: PURPOSES, DATA SOURCES, AND METHODOLOGICAL CHARACTERISTICS" (2014). *PACIS 2014 Proceedings*. 111.
<http://aisel.aisnet.org/pacis2014/111>

This material is brought to you by the Pacific Asia Conference on Information Systems (PACIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in PACIS 2014 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

A CONTENT ANALYSIS OF CONTENT ANALYSES IN IS RESEARCH: PURPOSES, DATA SOURCES, AND METHODOLOGICAL CHARACTERISTICS

André Coners, South Westphalia University of Applied Sciences, Hagen, Germany,
coners.andre@fh-swf.de

Benjamin Matthies, South Westphalia University of Applied Sciences, Hagen, Germany,
matthies.benjamin@fh-swf.de

Abstract

Qualitative research relies primarily on qualitative data in form of texts. The method of content analysis (CA) represents a scientifically well-founded and effective solution for making valid inferences from varied sources of textual information. This paper investigates the use of CA in the context of high-quality IS research with a focus on the underlying research questions, data sources, and methodological characteristics. Building on that, distinctive patterns, trends through time, and potentials will be discussed and compared with a reference discipline (organizational research). The results indicate that the general application of CA in high-quality IS research has steadily increased. In this context, CA is used in a very wide range of ways to understand and explain complex phenomena. Furthermore, CA is frequently used to categorize primary data collections, derived, for example, from transcribed interviews or open-ended surveys, and to transform qualitative data into quantitative outcomes. Besides other methodological issues, the increasing application of computational approaches seems to distinguish the IS discipline. Finally, this paper should also serve as an introduction to CA, bring transparency with respect to the methodological characteristics, and inspire researchers to carry out further CAs in the context of IS research.

Keywords: Content analysis, Text analysis, Qualitative data

1 INTRODUCTION

Qualitative research primarily involves the use of words rather than numbers (Kaplan and Maxwell, 2005). Through the analysis of natural language, relevant social phenomena, such as the interaction of people with information systems, can be understood and explained (Myers, 1997). Against this backdrop, the large quantity of textual information in the form of, for example, academic publications, corporate reporting, job advertisements, press releases, technical manuals, web contents, policies, or transcribed interviews, holds great potential for the information systems (IS) discipline (Freitas & Moscarola, 1998; Indulska et al., 2012; King, 2009). The research methodology of content analysis (CA), which has already been used in social science for more than 60 years, represents a scientifically well-founded and effective solution for making valid inferences from the varied qualitative data sources. In this context, CA can be understood as an “empirically grounded method” (Krippendorff, 2013, p.1) which is directed at the “objective, systematic, and quantitative description of the manifest content of communication” (Berelson, 1952, p.18).

As a method, CA has long been recognized in the context of IS research (see, e.g., Freitas & Moscarola, 1998; Lacity & Janson, 1994; Myers; 1997; or Remenyi, 1992), and its usage has expanded and taken root over time (see Mingers, 2003; or Weerakkody et al., 2009). In this context, two developments support the relevance of CA as a mode of analysis: First, the large and growing volume of textual information offers considerable potential for relevant analyses (Indulska et al., 2012). Second, computational text analysis methodologies make it possible to efficiently exploit this large volume of qualitative data (Crowston et al., 2012). This paper will therefore take this background to investigate the application of CA in IS research with a focus on the underlying research questions, data sources, and methodological characteristics. Investigating the use of such research methods in a discipline is an important task in order to uncover patterns and trends in their implementation and to highlight methodological lacks and potentials for further research (Scandura & Williams, 2000). A frequently discussed weakness in other research disciplines is, for example, the fact that the fundamental requirements of validity, objectivity, and reliability are not adhered to in the context of CA (see Beattie & Thomson, 2007; Duriau et al., 2007; Guthrie et al., 2004; Riffe & Freitag, 1997). This may result from, among other things, a shortcoming in well-founded analytical constructs, insufficient determination of samples, scarce consideration of accompanying methods, or the absence of necessary reliability checks. Another widespread criticism is that transparent documentation of such studies is frequently lacking (Beattie & Thomson, 2007; Riffe & Freitag, 1997). Some critics go so far as to cite a “pervasive lack of explanation in previous studies” (Beattie & Thomson, 2007, p.6). It is necessary, however, to safeguard these scientific standards in order to create credible results and to make them transparently usable for subsequent researchers.

The goal of this paper is to evaluate the use of CA within the IS research and, in particular, to determine underlying research questions and data sources along with the manner in which the method is used. In doing so, corresponding studies within the high-quality IS literature will be structurally evaluated based on predefined characteristics. A comparison with a reference discipline (organizational research) will improve the interpretation. Building on that, remarkable patterns, trends, and common rationales for using CA as a mode of analysis will be discussed. Furthermore, implications for future CAs will be outlined. To the best of our knowledge, this has not yet been done for the IS discipline. Within the framework of a systematic analysis, we will consider, in summary, the following research questions (RQ):

- RQ: How is the CA method used in the context of IS research?*
- a) Which research questions is the method used to pursue?*
 - b) What are the rationales for using CA as a mode of analysis?*
 - c) What are the underlying data sources?*
 - d) Are there patterns or trends in the methodological use?*

This paper is organized in the following manner: First, CA will be defined as a research methodology (Section 2). Then a systematic CA of CAs within the IS literature will be developed and deployed (Section 3). A discussion built onto that will portray CA as a method of IS research and address potential implications for further CAs (Section 4).

2 CONTENT ANALYSIS

As was mentioned previously, the origins of CA lie in social science, and it has been developing for more than 60 years (Krippendorff, 2013). Over time, this research methodology has been given a whole array of definitions, theoretical concepts, and analytical procedures (Duriiau et al., 2007). One simple yet very clear definition describes CA as “a systematic, replicable technique for compressing many words of text into fewer content categories based on explicit rules of coding” (Stemler, 2001, p.1). This involves the use of varied types of qualitative data, such as documents, transcribed interviews or speeches, maps, or even images (Miles & Huberman, 1994). However, insofar as CA is an “empirically grounded method” (Krippendorff, 2013, p.1), safeguarding scientific standards such as validity, objectivity, and reliability should always be emphasized (Berelson, 1952; Krippendorff, 2013; Neuendorf, 2002; Weber, 1990). It is therefore necessary to create “replicable and valid inferences from data to their context,” according to Krippendorff’s (2013, p.24) interpretation. Berelson similarly defines CA as “a research technique for the objective, systematic, and quantitative description of the manifest content of communication” (Berelson, 1952, p.18). The method’s scientific claims become even clearer in Neuendorf’s definition: “Content analysis is a summarizing, quantitative analysis of messages, that relies on the scientific method (including attention to objectivity-intersubjectivity, a priori design, reliability, validity, generalizability, replicability, and hypothesis testing) and is not limited as to the types of variables that may be measured or the context in which the messages are created or presented” (Neuendorf, 2002, p.10). In this sense, the creation of quantitative statements (i.e., “measurement” in texts) refers to the identification, counting, and, where necessary, weighting of clearly predefined content, for example in the form of certain keywords, phrases, themes, or charts (Berg, 2007). Therefore, CA can be seen as a bridge between qualitative data and quantitative analyses. The fundamental assumption here is that the frequency with which found contents appear is a direct reflection of their importance (Weber, 1990). This is supposed to make structured insights into the authors’ thinking and perspectives possible. Generally, such studies can be limited to three central purposes (Holsti, 1968): (1) describing the characteristics of communication (e.g., communication styles, communication contents, or techniques of persuasion); (2) making inferences with respect to the antecedents of communication (e.g., strategies in marketing campaigns); and (3) making inferences with respect to the effects of communication (e.g., recipient responses to certain messages). In this context, a typical opening question in the framework of a CA can essentially be stated as follows: “Who says what, to whom, how, with what effect, and why?” (Holsti, 1968, p. 603).

3 A CONTENT ANALYSIS OF CONTENT ANALYSES IN IS RESEARCH

3.1 Research Approach

This paper is particularly focused on the methodological use of CA in IS research. This means that we are rather interested in generating quantitative statements about the distribution of methodological characteristics, than interpreting and discussing research outcomes. In keeping with the theme of this paper, the CA treated here will therefore serve as a research approach for this primarily exploratory study. In this regard, relevant CA literature of high-quality IS research (see Section 3.2) will be investigated based on predefined categories and coding rules (see Section 3.3). By this means, the existence or absence of specific methodological characteristics will be revealed and quantitative statements with respect to the methodological use of CA will be produced. The comparison of chronological horizons should give the analysis additional depth and make it possible to uncover trends over time. Moreover, the insights will be compared with similar results in a reference discipline (organizational research), thus improving the overall interpretation (Baskerville & Myers, 2002).

3.2 Scope of the Analysis

At the very beginning of our analysis, we wanted to explore the general distribution of CA in the context of IS research. For this reason, we initially conducted a keyword search in scientific literature

databases. On the one hand, this should reveal the academic interest in CA and, on the other hand, should confirm the relevance of our analysis. In order to ensure the connection to IS research, we focussed on three scientific databases, namely: ACM Digital Library, AIS Electronic Library, and CS Digital Library. We used the search term “content analysis” and restricted the search to the publication titles, abstracts, and keywords. The corresponding search revealed more than 850 articles probably dealing with CA published during the past twenty-three years (see Figure 1). Particularly interesting is the increasing trend during the more recent past (2008–2013). This indicates an increasingly growing interest in CA and supports the relevance of our study.

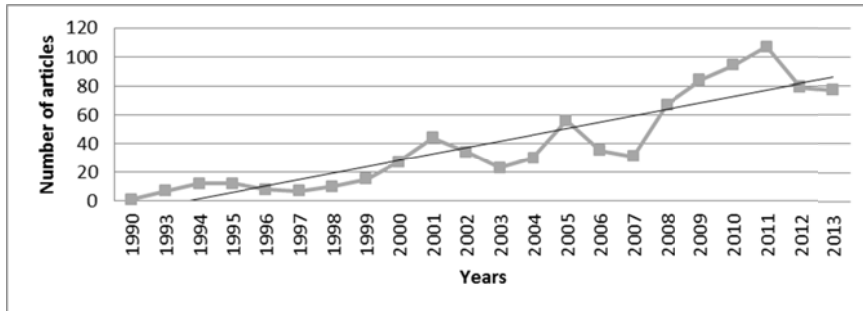


Figure 1. Results from database search

As we had to manually code the CA literature in order to uncover the methodological use of CA, it was necessary to limit the scope of our analysis to a manageable body of studies. This analysis will therefore consider articles from the eight IS journals of the Senior Scholars' Basket of Journals (December 2011), provided by the Association for Information Systems (AIS).¹ Such a focus on high-quality IS literature is a widespread research strategy (vom Brocke et al., 2009) and ensures the significance of the statements in the analysis. A corresponding database search was carried out (12/10/2013) using the search term “content analysis” (in “all fields”).

Journal	Database	Hits	Relevant
European Journal of Information Systems (EJIS)	Palgrave Macmillan	50	4
Information Systems Journal (ISJ)	Wiley Online Library	53	10
Information Systems Research (ISR)	INFORMS	31	3
Journal of the Association for Information Systems (JAIS)	AIS eLibrary (AISEL)	23	6
Journal of Information Technology (JIT)	ProQuest - ABI/INFORM	27	5
Journal of Management Information Systems (JMIS)	EBSCOhost (Business Source Premier)	14	4
Journal of Strategic Information Systems (JSIS)	ScienceDirect	28	6
Management Information Systems Quarterly (MISQ)	EBSCOhost (Business Source Premier)	11	4

Table 1. Results of the journal search

The corresponding search revealed a total of 237 articles (Table 1). Given that a relatively wide search was undertaken in “all fields,” the articles that were actually relevant to the analysis then had to be extracted in the following step. This was done based on the criteria that each article incorporated a well-founded CA, explicitly defined it as an applied mode of analysis, and, moreover, provided information that was necessary for our study (see Section 3.3). Ultimately we were left with a purposive sample of 42 relevant articles.

3.3 Coding of the Content Analysis Literature

In general, “there is no simple right way to do content analysis” (Weber, 1990, p.13). This paper, however, is particularly focused on the methodological use of CA. Fortunately, CA can be compared

¹ <http://start.aisnet.org/?SeniorScholarBasket>

very well, given that such studies ordinarily consist of a comparable “set of procedures” (Weber, 1990, p.9) that turn up in almost every study (Krippendorff, 2013). This paper follows on the concepts of Duriau et al. (2007) as well as Riffe and Freitag (1997), whereby the analytical categories and coding rules that they used were taken up, combined, and partly specified (described below).

Given that we had to implement a manual coding of the articles, we had to ensure the reliability of the results. First of all, detailed instructions were created with a description of the underlying categories and coding rules as well as concrete examples. Furthermore, the coding was done independently by one of the authors and by a research assistant who was trained in the concept of CA. In the first step, ten exemplary articles were examined independently and the agreement of the coding was tested. The inter-coder reliability was evaluated using Cohen’s kappa (Cohen, 1960) as the coefficient, which was on a reliable level of 0.81 (Landis & Koch, 1977). Then the differences were discussed and the instructions correspondingly revised. In the next step, all 42 articles were coded independently by the two coders. The final inter-coder agreement was 0.84. Once again, differences were discussed and final coding decisions were made.²

The transparent documentation of the underlying analytical construct is essential for a comprehensible CA (Beattie & Thomson, 2007). The theoretically founded categories and coding rules of this study (see Table 2) should therefore be described in greater detail in the following.

Category	Coding rules
Research topic	Short description of the investigated studies’ research purposes
Data source	Annual reports (AR), company documents (CD), contracts (CO), interview transcripts (IN), job advertisements (JA), open-ended surveys (OS), newspaper articles (NA), other field data (FD), policies (PO), press releases (PR), scholarly articles (SA), trade magazines (TM), web content (WC)
Sampling	Convenience sampling (CS), purposive sampling (PS), or representative sampling (RS)
Analytical construct	Inductive (0), deductive (1), or both (2)
Longitudinal design	Yes (1), no (0)
Analysis techniques	Qualitative approaches (QA), frequency counts (FC), statistical analyses (SA), advanced features (AF)
Multimethods	Single (0), conjunction (1), cross-validation (2)
Reliability check	Yes (1), no (0)
Computer-aided text analysis	Yes (1), no (0)

Table 2. Categories and coding rules

Research topic. The purpose of the analysis should be clearly described in any CA. Generally, corresponding questions should address a concrete problem that permits the examination of several possible answers, which in turn should be supported by the texts that are being examined and, moreover, should be verifiable (Krippendorff, 2013).

Data source. The various sources of qualitative data (i.e., texts) should be distinguished. The precise context and informative content of the texts does not always need to coincide with the researcher’s ideas. Therefore, the actual collection of texts always involves a certain degree of uncertainty and has a direct impact on the results of the analyses (Neuendorf, 2002). The suitability of the texts to the research topic should therefore ideally be (empirically) proven already (Krippendorff, 2013). In the context of our analysis, the following data sources will be distinguished: *annual reports* (AR), *company documents* (CD), *contracts* (CO), *interview transcripts* (IN), *job advertisements* (JA), *newspaper articles* (NA), *open-ended surveys* (OS), *other field data* (FD), *policies* (PO), *press releases* (PR), *scholarly articles* (SA), *trade magazines* (TM), and *web contents* (WC).

² The instructions and the results of the reliability check cannot be described in detail here due to the limited length of this paper. We would, however, gladly make them available upon request.

Sampling. In every study the question of the envisioned population arises, i.e., the field (e.g., an entire research discipline) for which the researcher is trying to generalize the results (Neuendorf, 2002). The universe of available texts in a population is usually too large to investigate in its totality. Each individual CA therefore requires a “manageable body of texts” (Krippendorff, 2013, p.112) to be defined. In this regard, one fundamental requirement is “to sample their texts to give their research question a fair chance of being answered correctly” (Krippendorff, 2013, p.114). Based on the study of Riffe and Freitag (1997), we distinguish between a *convenience sampling* (CS), a *purposive sampling* (PS), and a *representative sampling* (RS) in this paper. In doing so, CS (also known as opportunity sampling) stands for the use of a small selection of data that is available to the researchers and does not systematically derive from the underlying population. Typically no external selection process is applied in these cases, which is why the results seldom permit broad generalizations (Neuendorf, 2002). Examples are interviews of one’s own students or case studies conducted in a small and exemplary selection of companies. PS stands for the stepwise selection of a very detailed excerpt of a population, which is narrowed down based on specific criteria and serves as a basis for a very specific analytical goal. This is often a matter of investigating specialized problems (e.g., from a very detailed research area) or expert knowledge (Holsti, 1968). A typical example is a stepwise database search of scholarly articles based on specific selection criteria (Krippendorff, 2013). RS stands for a probability sampling of texts (e.g., through random sampling) that can be regarded as representative of a population (Krippendorff, 2013). In that process, each element of the population has the same chance of being selected, and the sampling is thus a valid reflection of all characteristics of the population for which the results are to be generalized (e.g., an IS profession). Furthermore, we assume RS in cases where all data relevant for a research topic are comprehensively analysed (e.g., Huang, 2007, who conducted a comprehensive study of all available U.S. county e-government portals).

Analytical construct. As Krippendorff (2013, p.91) emphasizes: “Content analyses succeed or fail, however, based on the validity (or invalidity) of the analytical constructs.” In the context of CA, an analytical construct can be described as a framework of conceptual categories and coding rules for systematically generating inferences from text contents. Such a construct helps to transform the collected data into useful information and to support inferences. Typical examples are dictionary approaches, which categorize significant terms or keywords into classes with similar meanings (Beattie & Thomson, 2007). Fundamentally, a distinction can be made between *inductively* (0) and *deductively* (1) developed analytical constructs (Berg, 2007, Mayring, 2000), as well as the use of *both* (2) approaches. An inductively developed construct originates primarily during the research process, and categories and coding rules are reductively and often iteratively deduced based on the underlying data (Mayring, 2000). For example, this could be the case in research endeavors with the purpose of generating new insights toward the formulation and grounding of a theory through the CA results. By contrast, a deductively developed analytical construct is primarily based on theories that already exist, and the corresponding (a priori) categories represent facets of the theory to be explored in a more confirmatory way. In this context, Krippendorff (2013) principally recommends the consideration of profound “sources of certainty” such as established theories, previous successes or failures of comparable analyses, expert knowledge and experiences, or embodied practices. By this means, the validity as well as objectivity of analytical constructs should be safeguarded.

Longitudinal design. One particular characteristic of CA is the use of data from distinct time periods and the analysis of developments over time (Holsti, 1968). Such a comparison of temporal horizons gives a CA an additional depth and makes trend analyses possible. Even though some studies select data for several years, a longitudinal analysis goes beyond merely collecting data from multiple time periods and explicitly requires a comparative juxtaposition of time frames (Duriiau et al., 2007; Scandura & Williams, 2000). Therefore, our analysis determines whether the particular CA *includes* (1) such a longitudinal analysis or *not* (0).

Analysis techniques. CA relies on several analytical techniques for generating inferences from texts. Neuendorf’s (2002, p.167) emphasis in this matter is clear: “The most common form of presentation – simple descriptive frequencies – is not the only option.” Researchers have a whole array of statistical and advanced exploratory techniques at their disposal. Our analysis will therefore distinguish between *qualitative approaches* (QA), *frequency counts* (FC), *statistical analyses* (SA), and *advanced features*

(AF). QA stands for a qualitative-interpretative mode of analysis, which primarily aims at the qualitative categorization of particular themes and main ideas, as well as interpretive discussions on their shared commonalities (Mayring, 2000). FC stands for general quantitative statements in the form of descriptive frequency counts or cross-tabulations of specific textual contents (e.g., keywords or themes). SA stands for the use of further statistical techniques, such as, for example, regression or variance analyses. AF stands for advanced analyses or graphic visualization, such as, for example, the establishment of co-occurrence matrices.

Multimethods. The use of multiple methods is a valuable attribute of a research project (Mingers, 2003; Scandura & Williams, 2000). Using multiple quantitative or qualitative modes of analysis that support, expand, or even cross-validate the results of a CA is particularly recommended (Bowen, 2007; Krippendorff, 2013; Neuendorf, 2002). The potential of CA in this context becomes evident in the statement of Bhattacharjee and Premkumar (2004, p.246): “The purpose of [CA] was to qualitatively triangulate and validate our earlier quantitative findings, and possibly gain additional insights into the nature and causes of the hypothesized associations.” Referring to Duriau et al. (2007), our analysis will therefore distinguish between studies with the CA as the *sole method* (0); studies which use CA in *conjunction* (1) with other methods, where multiple methods establish or complement each other in order to investigate a more complex research endeavor (e.g., CA in conjunction with other interpretive techniques such as ethnomethodology); and studies which use CA explicitly for a *cross-validation* (2) of results in the sense of triangulation (e.g., quantitative CA results in comparison with quantitative survey results).

Reliability checks. Safeguarding reliability is a central scientific requirement in the context of CA. Reliability, in this case, means the degree to which a measurement produces the same results upon repeated implementation (including measurements made by other researchers) (Neuendorf, 2002). That means in essence, that “different people should code the same text in the same way” (Weber, 1990, p.12). However, in most cases, coding, i.e., identifying, counting, and assessing predefined content, is dependent upon the subjective interpretation, experience, and care of the coder, particularly when the detailed coding of texts is manually executed. To ensure reliability, use of the most precise instructions possible (also known as coding rules) is highly recommended (Berg, 2007; Holsti, 1968; Neuendorf, 2002). Besides that, various procedures are available for checking reliability, namely ones which examine the agreement of the results of two or more independent coders (inter-coder reliability) (see, e.g., Hayes & Krippendorff, 2007). The deployment and explanation of such reliability checks is integral to a well-founded CA (Beattie & Thomson, 2007), which is why the *presence* (1) or *absence* (0) of such reliability checks and their resulting coefficients (e.g., Cohen’s kappa) is included as an additional characteristic of our analysis.

Computer-aided text analysis. “Several compelling analogies can be made between the way computers work and what content analysts do” (Krippendorff, 2013, p.209). The various possibilities of a “computer-aided text analysis” (CATA) represent a far-reaching evolution in CA execution (Crowston et al., 2012; Indulska et al., 2012; Krippendorff, 2013). As such, CATA stands for a wide diversity of computational approaches to text analysis (e.g., text mining techniques). CATA’s benefits are particularly centered on efficient processing of very large collections of texts as well as diverse and, to some extent, highly exploratory analysis techniques (Indulska et al., 2012). Another benefit of CATA is its reliability, given that “[CATA] cannot *not* process text reliably” (Krippendorff, 2013, p.210). This means that, under stable conditions, computers always produce the same results and, on the other hand, are not influenced by the human coder’s experience, interpretation, or care. A distinction should therefore be made as to whether any support of CATA was *used* (1) or *not* (0).

3.4 Results

First of all, the 42 identified studies were arranged chronologically. The results (Figure 2) clearly show that the general use of CA in high-quality IS research has steadily increased. This corresponds to our findings in Section 3.2 and is furthermore consistent with other researchers’ insights, according to which CA has increased in relevance, driven by the clear increase in available text documents (Bowen, 2009; Neuendorf, 2002).

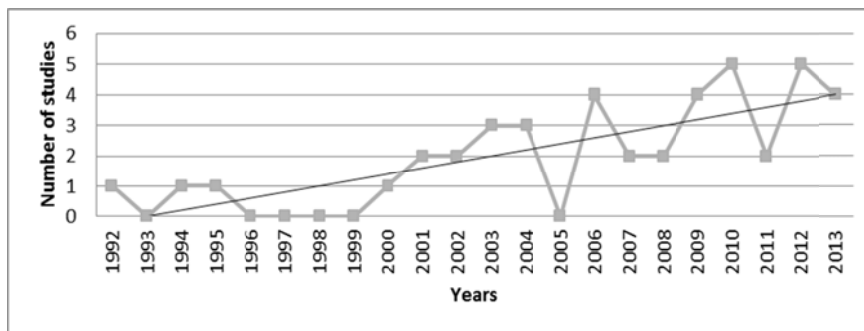


Figure 2. Number of CAs per year

In the next step, our collection of 42 articles was investigated based on the categories and coding rules presented above (the results of the coding are shown in the appendix):

- **Research topics:** The investigated studies' research topics varied drastically, which indicates a very diverse usage of the method. Although there was no discernible focus on specific phenomena, several studies share some commonalities. For example, the investigation of IS research trends based on scholarly articles was taken up repeatedly (Aksulu & Wade, 2010; Arnott & Pervan, 2012; Indulska et al., 2012; Heinrich & Riedl, 2013; Raub & Rüling, 2001). Furthermore, five studies examined social phenomena in online networks (Ghose, 2009; Krasnova et al., 2010; Kudaravalli & McGill, 2008; Moser et al., 2013; Pavlou & Dimoka, 2006). Three studies sought to explore required skills for IS professions (Klendauer et al., 2012; Todd et al., 1995; Wade & Parent, 2002). The underlying rationales for using CA as the mode of analysis will be discussed in the following section.
- **Data sources:** By far the most frequently used data source was interview transcripts (15 of 42 studies), followed by scholarly articles (9), web contents (6), and open-ended surveys (5).
- **Sampling:** 50.0% of the studies (21 of 42) relied on convenience sampling. Purposive sampling (38.1%; 16 of 42) was mostly used in the context of secondary sources of data, such as, for example, scholarly articles or job advertisements. Only five data collections could be classified as representative or at least comprehensive (Aksulu & Wade, 2010; Arnott & Pervan, 2012; Ghose, 2009; Huang, 2007; Raub & Rüling, 2001).
- **Analytical construct:** Many of the studies relied on purely theoretically founded categories and measures (38.1%; 16 of 42). Fourteen studies (33.3%) inductively develop their own constructs in order to pursue primarily exploratory goals as well as the establishment and evaluation of theories based on the CA results (see, e.g., Aksulu & Wade, 2010). Another twelve studies used a mix of deductively and inductively developed categories and measures. In most of these cases, a preliminary construct was deduced from theory and further specified on the basis of the investigated data.
- **Longitudinal design:** In total, only ten studies (23.8%) performed longitudinal analyses, i.e., a chronological comparison and the examination of trend trough time.
- **Analysis techniques:** Several studies (57.1%; 24 of 42) used qualitative-interpretive categorization of particular themes in order to reveal and discuss common concepts and opinions. However, most of the studies (90.5%, or 38 of 42) also showed quantitative measurement results in the form of descriptive frequency counts. 33.3% of the CAs (14) draw on this with statistical analyses (e.g., regression analyses). Only one analysis (2.4%) used advanced features.
- **Multimethods:** Many of the studies (54.8%) pursued a multimethod design (23 of 42), i.e., CA was one component of a more complex research endeavor using multiple methods. In this context, five studies (11.9%) also used CA explicitly for the purpose of cross-validation. Thus, 45.2% (19 of 42) of the studies relied upon CA as their sole method.
- **Reliability checks:** 64.3% of the studies carried out reliability checks (27 of 42). With respect exclusively to the CAs with primarily manual coding, reliability checks were performed in 73.0% of the analyses (27 of 37).
- **Computer-aided text analysis:** Altogether, 33.3% of the studies (14) utilized the support of computational approaches to carry out the analyses.

4 DISCUSSION, IMPLICATIONS, AND OUTLOOK

4.1 The Use of Content Analysis in IS Research

As has been shown, CA is used to investigate IS themes in a very wide range of ways, and the studies investigated here generally state that CA could provide valuable contributions to explore complex social phenomena based on natural language. In this context, five rationales for using CA as a mode of analysis were commonly stated: First, CA is a reliable technique for the identification of scalable structures and regularities, thereby allowing for objective inferences based on these structures and regularities (Berelson, 1952; Myers, 1997). For example, Watts and Henderson (2006) reasoned the use of CA with its ability to reliably synthesize the deep structure of text collections and to generate findings, which are valid, transferable, and relatively free from researcher bias. Second, the ability of CA to generate quantitative statements from narrative descriptions, which enables researchers to proceed with deeper statistical analyses and, for example, to verify causal relationships, is another common rationale. In this regard, Alonso-Mendo et al. (2009, p.268) stated that such statistical analyses enable the “identification of overall trends, new leads, and helps identify unexpected differences in a mass of answers.” Third, CA can be implemented to support both inductive and deductive research (Duriiau et al., 2007). As such, CA offers an analytical flexibility and is applicable to a wide range of relevant phenomena. Fourth, many researchers used CA in conjunction with other methods. For example, Koh et al. (2004) used CA techniques for the initial development of valid survey categories. Krasnova et al. (2010) used CA to complement theoretical findings obtained from a literature review. Fifth, CA allows a faster production of inferences compared with often time-consuming interpretive readings (Freitas & Moscarola, 1998). Nevertheless, central limitations of CA have been discussed as well: On the one hand, CA often tends to analyze texts in isolation with a limited scope, without considering further inherent contexts (see, e.g., Kaganer et al., 2010). On the other hand, the underlying assumption that the simple frequency with which the content appears directly reflects its relevance (Weber, 1990) is another discussed limitation (see, e.g., Sherif & Menon, 2004).

With respect to the investigated data sources, CA is frequently used to categorize primary data collections (45.2%), derived from transcribed interviews or open-ended surveys, and to transform qualitative data into quantitative outcomes. In this context, exploratory research relied most frequently on such primary data in order to discover unknown concepts in a body of text. The secondary data collections that are investigated are noticeably specific, which is particularly indicated by the frequent use of purposive samples based on often very specific selection criteria.

Despite the diversity of analytical purposes, several patterns and trends are discernible in the method’s application (see Table 3). The individual methodological characteristics will now be addressed in greater detail. First, there seems to be a tendency towards theoretically grounded analytical constructs (67%). This can be discussed from two contrary perspectives: On the one hand, several researchers have stated that deductively developed categories and coding rules provide a more objective and valid framework for CA (e.g., Huang, 2007; Rose & Schlichter, 2013). On the other hand, such approaches have their limitations in exploring new phenomena and limit the data to a predefined scope (Indulska et al., 2012). Second, longitudinal analyses can only be found in 24% of the studies. It should be noted that several other studies used data from distinct time periods without comparing them in their analyses; consequently, these are not classed as longitudinal. Nevertheless, with regard to the advantages of longitudinal analyses in order to explore and verify phenomena over time, 24% seems to be low. Third, the use of multiple methods is a central quality characteristic (Bowen, 2007; Mingers, 2003; Scandura & Williams, 2000). As indicated, 55% of the studies incorporated a multimethod design, and in most cases CAs were integrated in more complex research endeavors. For example, several researchers used CAs to develop and validate concrete categories and measures before designing surveys or interviews (e.g., Koh et al., 2004). The results of CA, however, are less frequently cross-validated in the sense of triangulation with the (quantitative) results of other methods (11.9%; see, e.g., Ghose, 2009). Fourth, a significant consistency in the use of quantitative statements (90%) is remarkable. This corresponds to the quantitative focus of the method (Neuendorf, 2002) and, thus, CA frequently serves as a bridge between qualitative data and quantitative outcomes. 33% of the

studies draw on this with statistical analyses in order to examine in greater depth the phenomena. Fifth, the studies ensure the reliability of their analyses in a large and growing portion, particularly where manual coding is concerned (73%). This is an important quality aspect of CA.

Years	<i>N</i>	Deductive ^a construct development	Longitudinal design	Multi- methods	Qualitative analyses	Frequency counts	Statistical analyses	Reliability ^b checks	CATA
2013 - 2011	11	73%	45%	64%	73%	100%	18%	89%	55%
2010 - 2006	17	71%	6%	53%	53%	88%	53%	93%	35%
2005 - 2001	10	70%	30%	50%	60%	80%	20%	44%	20%
2000 - 1996	1	0%	0%	0%	33%	100%	0%	0%	0%
1995 - 1990	3	33%	33%	67%	0%	100%	33%	33%	0%
TOTAL	42	67%	24%	55%	57%	90%	33%	73%	33%

Notes:
N: number of studies; ^a including deductive/inductive approaches; ^b in the context of human coding

Table 3. *Patterns and trends in the use of CA in IS research*

Sixth, the IS discipline appears to be increasingly reliant upon CATA. Correspondingly, statements on the potential of computational approaches to carrying out CA are more and more evident in the IS literature (e.g., Evangelopoulos et al., 2012; Indulska et al., 2012; King, 2009). One could in fact say that there has been a methodological evolution in executing CA (Crowston et al., 2012; Krippendorff, 2013). The analytical possibilities of corresponding tools have been continually developed in recent years and contain tremendous potential particularly for advancing exploratory CA. As an example, we need only name the latent semantic analysis (LSA) technique, which automatically reveals latent semantic relationships in texts based on statistical calculations (see, e.g., Evangelopoulos et al., 2012). Such advanced and exploratory analyses remained largely unconsidered in the studies reviewed here. Nonetheless, computational approaches in the context of a CA also have clear limits: Even complete reliability still does not ensure the validity of the results (Krippendorff, 2013). An experienced analyst should always evaluate the semi-automatic results in terms of their correctness and their suitability for answering the questions that were asked at the outset. The following statement aptly summarizes this: “To justify the use of CATA software, content analysts must assure themselves, as well as the community of peers, that the way a software package processes the data is compatible with what is known about the context of the texts, how texts are read, what they mean, and what role they play” (Krippendorff, 2013, p.212).

Additionally, our CA results allow several comparisons with patterns in the reference discipline of organizational research (OR) (see Duriau et al., 2007). In comparison, several methodological aspects have been considered to similar degrees, such as, for example, the employment of reliability checks (IS: 64.3%; OR: 62.5%) or the use of descriptive frequency counts (IS: 90.5%; OR: 84.7%). At this point, we will focus on some distinctive differences in order to complement our results discussed above. First of all, a less frequent consideration of longitudinal designs is noticeable (IS: 23.8%; OR: 34.7%). This insight supports the demand for more longitudinal analyses in the context of CA. Second, the IS research appears to have distinguished itself through a stronger exploratory focus and use of purely inductive designs (IS: 33.3%; OR: 17.3%). This seems to be related to the much more frequent use of primary data, such as interview transcripts and open-ended surveys (IS: 45.2%; OR: 20.4%). On the other hand, the use of annual reports as a source of data is at a remarkable low level compared with OR (IS: 2.4%; OR: 30.6%). Lastly, the consideration of multimethod designs seems to be less frequent (IS: 54.8%; OR: 66.3%).

As a last point, we should mention some distinctive gaps in consistent documentation of the studies being investigated. For instance, it is recommended in principle that researchers show examples of the investigated passages and explain their categorization and coding (Beattie and Thomson, 2007; Neuendorf, 2002). This gives the reader valuable information about the reliability and validity of the measurements. Such examples, however, were only very infrequent in the studies that we looked at (see, e.g., Kudaravalli & Faraj, 2008; Tan, 1994; Zmud et al., 2010). Moreover, although reliability checks were often described in the studies, the coefficients of those checks were frequently not shown (see Mohdzain & Ward, 2007; Rivard & Lapointe, 2012; or Rose & Schlichter, 2013). Furthermore,

especially against the backdrop of increasing use of computational approaches, we should emphasize at this point the particular requirements of comprehensive documentation in the context of CATA. In light of the complex transformation procedures from unstructured text information to structured data (see, e.g., Crowston et al., 2012) and various analytical possibilities, transparent documentation of all extensive adjustments has become a central theme of CA (Krippendorff, 2013). Therefore, just as a literature review (vom Brocke et al., 2009), CA requires rigor in documenting the research process.

4.2 Implications

First of all, this paper intended to reveal the status quo of CA in the context of IS research with a focus on the underlying research topics, data sources, and methodological employment. Nevertheless, this paper should not only meet the goals of a pure review of previous studies. The methodological characteristics of a CA addressed here should also serve as an introductory guide, bring transparency with respect to the requirements of methodological quality, and inspire readers to carry out further CAs. Therefore, building on our discussions above, we want to emphasize central findings of our analysis and give researchers interested in using CA some concrete recommendations:

- Corporate documents, such as, for example, internal project documentations or annual reports, represent rich and largely unconsidered secondary data sources.
- A stronger consideration of longitudinal analyses can be recommended. A longitudinal design gives CAs an additional depth, makes trend analyses possible, and thus improves the verification of phenomena over time.
- The implementation of additional statistical analyses and advanced features are meaningful extensions of the quantitative CA results and improve the profound explanation of causal relationships.
- The investigated studies have shown the potential for using CA in conjunction with other methods. Furthermore, CA has proven itself as a useful means for cross-validation purposes (e.g., CA results in comparison with survey results).
- Reliability checks should be part of every manually executed CA.
- Recent advances in CATA represent an exciting evolution in CA. On the one hand, computational approaches to CA hold a great and largely untapped potential for exploratory analyses. On the other hand, CATA enables researchers to efficiently exploit large and thus more comprehensive text collections.
- The transparent documentation, especially in the context of CATA, has become an essential requirement of CA.

With respect to the limitations of this analysis, we must point out that the examined studies based on eight journals only represent an extract of all of the CAs undertaken in the IS discipline. However, because these derived from the Senior Scholars' Basket of Journals, we think that a representative view of the usage of CA in the context of high-quality IS research was produced. Nevertheless, we see potential for future research to extend the analysis to other than the selected journals. Furthermore, a manual coding of texts always involves reliability concerns. We tried to overcome these concerns by establishing precise coding rules and verification by means of reliability checks.

4.3 Outlook

All in all, the studies investigated here confirm a general applicability of CA to explore complex phenomena in the context of qualitative IS research. Against the backdrop of the large and growing amount of available texts, which certainly contain diverse and relevant IS contents, the increasing use of CA can only be welcomed. When sufficiently implemented, CA offers significant advantages and can overcome the limits of purely interpretive readings by creating quantitative statements, and thus can serve as a bridge between qualitative data and quantitative analyses. The method's corresponding potential was summarized in 1998 by Freitas and Moscarola (p.943): "When sufficiently defined and detailed, content analysis allows us to pass of the simple description and to reach the objective of every scientific research: the discovery of explanations and causal relationships."

References

- Alonso-Mendo, F., Fitzgerald, G., and Frias-Martinez, E. (2009). Understanding web site redesigns in small- and medium-sized enterprises (SMEs): a U.K.-based study on the applicability of e-commerce Stage Models. *European Journal of Information Systems*, 18 (3), 264-279.
- Aksulu, A. and Wade, M.R. (2010). A comprehensive review and synthesis of open source research. *Journal of the Association for Information Systems*, 11 (11), 576-656.
- Arnott, D. and Pervan, G. (2012). Design science in decision support systems research: An assessment using the Hevner, March, Park, and Ram Guidelines. *Journal of the Association for Information Systems*, 13 (11), 923-949.
- Baskerville R. and Myers, M.D. (2002). Information systems as a reference discipline. *MIS Quarterly*, 26 (1), 1-14.
- Beattie, V. and Thomson, S.J. (2007). Lifting the lid on the use of content analysis to investigate intellectual capital disclosures. *Accounting Forum*, 31 (2), 129-163.
- Berelson, B. (1952). *Content analysis in communication research*. Free Press, Glencoe, Ill.
- Berg, B.L. (2007). *Qualitative research methods for the social sciences*. 6th Edition. Allyn and Bacon, Boston, MA.
- Bhattacharjee, A. and Premkumar, G. (2004). Understanding changes in belief and attitude toward information technology usage: a theoretical model and longitudinal test. *MIS quarterly*, 28 (2), 229-254.
- Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative research journal*, 9 (2), 27-40.
- Chen, Y. and Bharadwaj, A. (2009). An empirical analysis of contract structures in IT outsourcing. *Information Systems Research*, 20 (4), 484-506.
- Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20, 37-46.
- Corvera Charaf, M., Rosenkranz, C., and Holten, R. (2013). The emergence of shared understanding: applying functional pragmatics to study the requirements development process. *Information Systems Journal*, 23 (2), 115-135.
- Crowston, K., Allen, E.E., and Heckman, R. (2012). Using natural language processing technology for qualitative data analysis. *International Journal of Social Research Methodology*, 15(6), 523-543.
- Cukier, W., Ngwenyama, O., Bauer, R., and Middleton, C. (2009). A critical analysis of media discourse on information technology: preliminary results of a proposed method for critical discourse analysis. *Information Systems Journal*, 19 (2), 175-196.
- Duriau V.J., Reger, R.K., and Pfarrer, M.D. (2007). A content analysis of the content analysis literature in organization studies: Research themes, data sources, and methodological refinements. *Organizational Research Methods*, 10 (1), 5-34.
- Evangelopoulos, N., Zhang, X., and Prybutok, V.R. (2012). Latent Semantic Analysis: five methodological recommendations. *European Journal of Information Systems*, 21 (1), 70-86.
- Freitas, H. and Moscarola, J. (1998). Content analyzing qualitative data on Information Systems. In *Proceedings of the 4th Americas Conference on Information Systems*, pp. 941-943, Baltimore, MD.
- Gallivan, M.J. (2001). Striking a balance between trust and control in a virtual organization: a content analysis of open source software case studies. *Information Systems Journal*, 11 (4), 277-304.
- Gallivan, M.G. and Depledge, G. (2003). Trust, control and the role of interorganizational systems in electronic partnerships. *Information Systems Journal*, 13 (2), 159-190.
- Gallivan, M.J. and Keil, M. (2003). The user-developer communication process: a critical case study. *Information Systems Journal*, 13 (1), 37-68.
- Ghose, A. (2009). Internet exchanges for used goods: An empirical analysis of trade patterns and adverse selection. *MIS Quarterly*, 33 (2), 263-291.
- Goel, S. and Chengalur-Smith, I.N. (2010). Metrics for characterizing the form of security policies. *Journal of Strategic Information Systems*, 19 (4), 281-295.
- Guthrie, J., Petty, R., Yongvanich, K. and Ricceri, F. (2004). Using content analysis as a research method to inquire into intellectual capital reporting. *Journal of Intellectual Capital*, 5 (2), 282-293.
- Hayes, A. F. and Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, 1(1), 77-89.

- Heinrich, L.J. and Riedl, R. (2013). Understanding the dominance and advocacy of the design-oriented research approach in the business informatics community: a history-based examination. *Journal of Information Technology*, 28 (1), 34-49.
- Holsti, O.R. (1968). Content analysis. In *Handbook of social psychology* (Gardner L. and Aronson, E. Ed.), pp. 596-692, Addison-Wesley, Reading, MA.
- Huang, Z. (2007). A comprehensive analysis of U.S. counties' e-Government portals: development status and functionalities. *European Journal of Information Systems*, 16 (2), 149-164.
- Indulska, M., Hovorka, and Jan Recker, J. (2012). Quantitative approaches to content analysis: identifying conceptual drift across publication outlets. *European Journal of Information Systems*, 21 (1), 49-69.
- Kaganer, E.A., Pawlowski, S.D., and Wiley-Patton, S. (2010). Building legitimacy for IT innovations: The case of computerized physician order entry systems. *Journal of the Association for Information Systems*, 11 (1), 1-33.
- Kaplan, B. and Maxwell, J. A. (2005). Qualitative research methods for evaluating computer information systems. In *Evaluating the Organizational Impact of Healthcare Information Systems* (Anderson, J.G. and Aydin, C. Ed.), pp. 30-55, Springer, New York.
- King, W.R. (2009). Text analytics: boon to knowledge management? *Information Systems Management*, 26 (1), 87.
- Klendauer, R., Berkovich, M., Gelvin, R., Leimeister, J. M., and Krcmar, H. (2012). Towards a competency model for requirements analysts. *Information Systems Journal*, 22 (6), 475-503.
- Koh, C., Ang, S., and Straub, D.W. (2004). IT outsourcing success: A psychological contract perspective. *Information Systems Research*, 15 (4), 356-373.
- Krasnova, H., Spiekermann, S., Koroleva, K., and Hildebrand, T. (2010). Online social networks: why we disclose. *Journal of Information Technology*, 25 (2), 109-125.
- Krippendorff, K. (2013). *Content Analysis. An Introduction to its Methodology*. 3rd. Edition. SAGE Publications, Thousand Oaks, CA.
- Kudaravalli, S. and Faraj, S. (2008). The structure of collaboration in electronic networks. *Journal of the Association for Information Systems*, 9 (10), 706-726.
- Lacity, M.C. and Janson, M.A. (1994). Understanding qualitative data: a framework of text analysis methods. *Journal of Management Information Systems*, 11 (2), 137-155.
- Landis, J.R. and Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159-174.
- Li, F. (2010). Textual analysis of corporate disclosures: A survey of the literature. *Journal of Accounting Literature*, 29, 143-165.
- Mayring, P. (2000). Qualitative content analysis. *Forum: Qualitative Social Research*, 1 (2). Retrieved December 7, 2013 from <http://nbn-resolving.de/urn:nbn:de:0114-fqs0002204>.
- McLure Wasko M. and Faraj, S. (2000). It is what one does: why people participate and help others in electronic communities of practice. *Journal of Strategic Information Systems*, 9 (2), 155-173.
- Miles, M. B. and Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. 3rd. Edition. SAGE Publications, Thousand Oaks, CA.
- Mingers, J. (2003). The paucity of multimethod research: a review of the information systems literature. *Information Systems Journal*, 13 (3), 233-249.
- Mohdzain, M.B. and Ward, J.M. (2007). A study of subsidiaries' views of information systems strategic planning in multinational organisations. *Journal of Strategic Information Systems*, 16 (4), 324-352.
- Moser, C., Ganley, D., and Groenewegen, P. (2013). Communicative genres as organising structures in online communities-of team players and storytellers. *Information Systems Journal*, 23 (6), 551-567.
- Mustonen-Ollila, E. and Lyytinen, K. (2003). Why organizations adopt information system process innovations: a longitudinal study using Diffusion of Innovation theory. *Information Systems Journal*, 13 (3), 275-297.
- Myers, M. D. (1997). Qualitative Research in Information Systems. *MIS Quarterly*, 21 (2), pp. 241-242. MISQ Discovery, archival version, http://www.misq.org/discovery/MISQD_isworld/. MISQ Discovery, updated version, last modified: January 4, 2008 <http://www.qual.auckland.ac.nz/>
- Neuendorf, K.A. (2002). *The Content Analysis Guidebook*. Sage Publications, Thousand Oaks, CA.
- Oliver, D. and Romm, C. (2002). Justifying enterprise resource planning adoption. *Journal of Information Technology*, 17 (4), 199-213.

- Pavlou, P.A. and Dimoka, A. (2006). The nature and role of feedback text comments in online marketplaces: Implications for trust building, price premiums, and seller differentiation. *Information Systems Research*, 17 (4), 392-414.
- Raub, S. and Ruling, C.-C. (2001). The knowledge management tussle – speech communities and rhetorical strategies in the development of knowledge management. *Journal of Information Technology*, 16 (2), 113-130.
- Remenyi, D. (1992). Researching information systems: data analysis methodology using content and correspondence analysis. *Journal of Information Technology*, 7 (2), 76-86.
- Ridley, G. and Young, J. (2012). Theoretical approaches to gender and IT: examining some Australian evidence. *Information Systems Journal*, 22 (5), 355-373.
- Riffe, D. and Freitag, A. (1997). A content analysis of content analyses: Twenty-five years of *Journalism Quarterly*. *Journalism & Mass Communication Quarterly*, 74 (3), 515-524.
- Rivard, S. and Lapointe, L. (2012). Information technology implementers' responses to user resistance: nature and effects. *MIS Quarterly*, 36 (3), 897-920.
- Rose, J. and Schlichter, B. R. (2013). Decoupling, re-engaging: managing trust relationships in implementation projects. *Information Systems Journal*, 23(1), 5-33.
- Scandura, T.A. and Williams, E. (2000). Research methodology in management: Current practices, trends, and implications for future research. *Academy of Management Journal*, 43 (6), 1248–1264.
- Scheepers, R., Scheepers, H., and Ngwenyama, O.K. (2006). Contextual influences on user satisfaction with mobile computing: findings from two healthcare organizations. *European Journal of Information Systems*, 15 (3), 261-268.
- Schubert, P. and Legner, C. (2011). B2B integration in global supply chains: An identification of technical integration scenarios. *Journal of Strategic Information Systems*, 20 (3), 250-267.
- Sherif, K. and Menon, N.M. (2004). Managing technology and administration innovations: Four case studies on software reuse. *Journal of the Association for Information Systems*, 5 (7), 247-281.
- Stemler, S. (2001). An overview of content analysis. *Practical Assessment, Research & Evaluation*, 7 (17). Retrieved December 7, 2013 from <http://PAREonline.net/getvn.asp?v=7&n=17>.
- Tan, M. (1994). Establishing mutual understanding in systems design: An empirical study. *Journal of Management Information Systems*, 10 (4), 159-182.
- Tay, K. B. and Chelliah, J. (2011). Disintermediation of traditional chemical intermediary roles in the Electronic Business-to-Business (e-B2B) exchange world. *Journal of Strategic Information Systems*, 20 (3), 217-231.
- Todd, P.A., McKeen, J.D. and Gallupe, R.B. (1995). The evolution of IS job skills: A content analysis of IS job advertisements from 1970 to 1990. *MIS Quarterly*, 19 (1), 1-27.
- vom Brocke J., Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., Cleven, A. (2009). Reconstructing the giant: On the importance of rigour in documenting the literature search process. In *Proceedings of the 17th European Conference On Information Systems, ECIS 2009*, pp. 2206-2217, Verona, Italy.
- Wade, M.R. and Parent, M. (2002). Relationships between job skills and performance: A study of webmasters. *Journal of Management Information Systems*, 18 (3), 71-96.
- Wang, W. and Benbasat, I. (2008). Attributions of trust in decision support technologies: A study of recommendation agents for E-commerce. *Journal of Management Information Systems*, 24(4), 249-273.
- Watts, S. and Henderson, J.H. (2006). Innovative IT climates: CIO perspectives. *Journal of Strategic Information Systems*, 15 (2), 125-151.
- Weber, R.P. (1990). *Basic Content Analysis*. 2nd Edition. Sage Publications, Newbury Park, CA.
- Weerakkodyl, V., Dwivedi, Y.K. and Irani, Z. (2009). The diffusion and use of institutional theory: a cross-disciplinary longitudinal literature survey. *Journal of Information Technology*, 24 (4), 354-368.
- Zhou, L., Shi, Y., Zhang, D., and Sears, A. (2006). Discovering cues to error detection in speech recognition output. *Journal of Management Information Systems*, 22 (4), 237-270.
- Zmud, R.W., Shaft, T., Zheng, W., and Croes, Henry. (2010). Systematic differences in firm's information technology signaling: Implications for research design. *Journal of the Association for Information Systems*, 11 (3), 149-181.

APPENDIX: RESULTS OF THE CONTENT ANALYSIS

Author	Year	Research topic	Data source	Sampling	Analytical construct	Longitudinal design	Analysis techniques	Multi-methods	Reliability check	CATA
Akshu & Wade	2010	Review and synthesis of open source research	SA	RS	0	0	QA, FC	0	1	0
Alonso-Mendo et al.	2009	Understanding web site redesigns in SMEs	IN	PS	0	0	QA, FC	1	1	1
Arnott & Pervan	2012	Assessment of design science in decision support systems research	SA	RS	1	1	FC	0	1	0
Bhattacharjee & Premkumar	2004	Elaboration of changes in belief and attitude toward IT usage	OS	CS	1	0	FC	1, 2	1	0
Chen & Bharadwaj	2009	Assessment of contract structures in IT outsourcing	CO	PS	2	0	FC, SA	0	1	0
Convera Charaf et al.	2013	Analysis of shared understanding in requirements development	IN, FD	CS	2	0	QA, FC	0	1	0
Cukier et al.	2009	Analysis of media discourse on information technology	NA, SA, TM	PS	1	0	QA, FC	1	0	1
Gallivan	2001	Examination of open source software case studies	SA, TM	PS	1	0	QA	0	0	0
Gallivan & Depledge	2003	Examination of electronic partnerships	SA, TM	PS	1	0	FC	1	1	0
Gallivan & Keil	2003	Analysis of the user-developer communication process	IN, OS	CS	0	1	QA, FC	1	0	0
Ghose	2009	Analysis of trade patterns and adverse selection in online markets	WC	RS	1	0	FC, SA	1, 2	1	0
Goel et al.	2010	Examination of characteristics of security policies	PO	PS	1	0	FC, SA	1, 2	0	1
Heinrich & Riedl	2013	Analysis of the design-oriented research approach in Business Informatics	OS	CS	2	1	QA, FC	0	0	1
Huang	2007	Analysis of the services and functions of U.S. county e-government portals	WC	RS	1	0	FC, SA	0	1	0
Induska et al.	2012	Investigation of conceptual changes in published literature	SA	PS	0	1	FC, SA, AF	1, 2	0	1
Kagamer et al.	2010	Examination of the legitimacy for IT innovations	PR	PS	2	1	QA, FC	0	1	1
Klendauer et al.	2012	Exploring a competency model for requirements analysts	IN	CS	2	0	QA, FC	1	1	1
Koh et al.	2004	Identifying psychological contract obligations in IT outsourcing	IN	CS	0	0	QA	1	1	0
Krasnova et al.	2010	Understanding information disclosure on online social networks	IN	CS	0	0	QA, FC	1	1	0
Kudravalil & McGill	2008	Examination of the collaboration in electronic networks	WC	CS	2	0	FC, SA	0	1	0
Mc Lure Wasko & Farej	2000	Investigation of the participation in electronic communities	OS	CS	0	0	QA, FC	0	0	0
Mohdzaïn & Ward	2007	Study of subsidiaries' views of IS strategic planning	IN	CS	1	0	QA	0	1	1
Moser et al.	2013	Examination of organising structures in online communities	WC	CS	0	0	QA, FC	1	1	0
Mustonen-Ollila & Lyytinen	2003	Analysis of IS process innovations adoption decisions	CD, IN	CS	1	1	FC	0	0	0
Oliver & Romm	2002	Justifying enterprise resource planning adoption	CD	PS	0	0	QA, FC	0	0	1
Pavlou & Dimoka	2006	Analysis of feedback text comments in online marketplaces	WC	PS	1	0	FC, SA	1	1	0

Author	Year	Research topic	Data source	Sampling	Analytical construct	Longitudinal design	Analysis techniques	Multi-methods	Reliability check	CATA
Raub & Rittling	2001	Analysis of the knowledge management discourse	SA	RS	2	1	QA, FC, SA	0	0	0
Remenyi	1992	Analysis of influences in the formulation and implementation of SIS	IN	CS	0	0	FC	1	0	0
Ridley & Young	2012	Examination of the gender gap in the IT industry	NA	PS	1	1	QA, FC	0	1	1
Rivard & Lapointe	2012	Examination of IT implementers' responses to user resistance	SA	PS	2	0	QA, FC, SA	1	1	1
Rose & Schlichter	2013	Analysis of trust relationships in IS implementation projects	IN	CS	1	1	FC	1	1	1
Scheepers et al.	2006	Investigation of user satisfaction with mobile computing	IN	CS	0	0	QA, FC	1	1	1
Schubert and Legner	2011	Development of a framework for explaining B2B integration	CD, SA	PS	2	0	QA, FC	1	1	0
Sherif & Meron	2004	Management of technology and administration innovations	IN	CS	1	0	QA, FC, SA	0	1	1
Tan	1994	Establishing mutual understanding in systems design	IN	CS	0	0	FC, SA	1	1	0
Tay & Chelliah	2011	Disintermediation of traditional chemical intermediary roles in e-B2B	WC	CS	0	0	QA, FC	1	0	0
Todd et al.	1995	Analysis of IS job skills over time	JA	PS	1	1	FC	0	0	0
Wade & Parent	2002	Determination of skills demanded of Webmasters	JA	PS	2	0	FC	1	0	0
Wang & Benbasat	2008	Study of recommendation agents for e-commerce	OS	CS	1	0	FC, SA	1, 2	1	0
Watts & Henderson	2006	Exploration of the concept of innovative IT climates	IN	CS	2	0	QA	0	0	0
Zhou et al.	2006	Discovering cues to error detection in speech recognition output	IN, FD	CS	0	0	QA, FC, SA	0	1	0
Zmud et al.	2010	Examination of firms' IT signaling propensities	AR, PR	PS	2	0	FC, SA	1	1	0

Notes:

Data source: Annual reports (AR), company documents (CD), contracts (CO), interview transcripts (IN), job advertisements (JA), open-ended surveys (OS), newspaper articles (NA), other field data (FD), policies (PO), press releases (PR), scholarly articles (SA), trade magazines (TM), web content (WC)

Sampling: Convenience sampling (CS), purposive sampling (PS), or representative sampling (RS)

Analytical construct: Inductive (0), deductive (1), or both (2)

Longitudinal design: Yes (1), no (0)

Analysis techniques: Qualitative approach (QA), frequency count (FC), statistical analysis (SA), advanced features (AF)

Multimethods: Single (0), conjunction (1), cross-validation (2)

Reliability check: Yes (1), no (0)

CATA: Yes (1), no (0)