RESEARCH ARTICLE

WILEY

QCA and the harnessing of unstructured qualitative data

Rohit Nishant[1] ⓘ | M.N. Ravishankar[2] ⓘ

[1]Département de systèmes d'information organisationnels, FSA Ulaval, Quebec, Canada

[2]School of Business and Economics, Loughborough University, Loughborough, UK

Correspondence
Rohit Nishant, Département de systèmes d'information organisationnels, FSA Ulaval, Quebec, Canada.
Email: rohit.nishant@fsa.ulaval.ca

Abstract

This paper proposes qualitative comparative analysis (QCA) as a novel method to harness unstructured data sets such as publicly available reports and news articles. It shows how QCA and conventional qualitative IS research can complement each other. In particular, it demonstrates how qualitative IS research can combine typical qualitative coding techniques with a specific type of QCA, namely crisp-set QCA (csQCA). The paper illustrates how QCA offers qualitative IS research an innovative approach to explicate the combination of conditions associated with particular outcomes. Drawing on an empirical study of green IS, it showcases the potential of QCA to harness large unstructured qualitative material and generate deeper insights about emerging IS phenomena. The paper also highlights how QCA can contribute to the data collection and analysis stages of qualitative IS research.

KEYWORDS
crisp-set QCA (csQCA), green IS, qualitative IS research, research methods

1 | INTRODUCTION

The proliferation and abundance of unstructured textual material in electronic form (eg, reports, news articles, blogs, and social media posts) presents unique opportunities for qualitative researchers. In particular, such data can potentially provide a comprehensive view of a given phenomenon, thus opening up spaces for more robust and deeper analysis. However, typical qualitative approaches such as case studies and ethnographies do not always tap extensively into the full range of available textual material (Sarker, Xiao, Beaulieu, & Lee, 2018). Even when studies claim an engagement with diverse sources of data, it is often hard to see how the named sources were specifically utilized.

This paper suggests that IS research can better exploit the potential of unstructured textual data by embracing innovative qualitative research methods and techniques.

The paper introduces csQCA (crisp-set qualitative comparative analysis) a set-theoretic method for analysing qualitative data, which lends itself to both inductive and deductive approaches. For illustrative purposes, this paper mainly details an inductive QCA approach. QCA is a technique that originated in sociology (Ragin, 1987) and is typically used in conjunction with conventional case study approaches to code qualitative data and test specific relationships (Aversa, Furnari, & Haefliger, 2015). QCA can help test hypotheses similar to conventional empirical analysis. However, in contrast to conventional approaches that merely restrict themselves to identifying relationships between variables QCA facilitates deeper analyses and helps determine the combination of causal conditions relevant to a given outcome over time (Garcia-Castro & Francoeur, 2016). It takes into account patterns of interrelated elements, structures, attributes, and practices rather than just illuminating the impact of independent variables.[1] QCA's configurational analyses showcase complex causality and nonlinear relationships (Fiss, 2007) rather than assume unifinality (ie, the existence of a single optimal configuration). In other words, it highlights equifinality (ie, the existence of multiple configurations) and explores synergistic effects.

In order to illustrate how QCA can enrich qualitative IS research, this paper focuses on a specific variant of QCA known as csQCA and applies it to the green IS phenomenon—a setting particularly replete with unstructured data and one in which the first author has prior research experience.

The paper contributes to IS research in multiple ways. First, it shows how qualitative IS research can use the QCA method to explore configurations associated with specific outcomes of interest. It suggests that the method can be used inductively as well as to test hypotheses and engage in deductive reasoning, using unstructured qualitative data. Second, it demonstrates how QCA, in conjunction with unstructured data, can be used to conduct longitudinal qualitative studies. Finally, it shows how the application of QCA can help generate focused questions for further research, elicit better quality information from respondents, and reveal (in)consistencies between firms' documented IT initiatives and respondents' interpretation of the same initiatives.

The rest of the paper is structured as follows. The next section provides an overview of the QCA method and highlight opportunities for studying emerging IS phenomena. The subsequent sections of the paper present and explain the application of csQCA to unstructured green IS data. The concluding sections discuss the implications for research and practice.

2 | QUALITATIVE COMPARATIVE ANALYSIS

QCA is a set-theoretic method of data analysis which has its origins in sociology (Ragin, 2008). It is now emerging as a valuable approach in business and management research (Garcia-Castro & Francoeur, 2016). QCA examines possible configurations (ie, combinations of conditions) associated with a given outcome and thus helps find the likely configuration of variables that may lead to the observed outcome. QCA can help in the analysis and interpretation of large swathes of textual data. Two of the most commonly used QCA tools are the (a) fuzzy-set QCA (fsQCA) and (b) crisp-set QCA (csQCA). fsQCA transforms qualitative data into a fuzzy set (Ragin, 2008; Garcia-Castro & Francoeur, 2016) with values ranging between 0 to1 based on quantitative anchor points such as the 25th percentile, 50th percentile, and 75th percentile. These values are later used for determining full nonmembership and full membership of categories. Alternatively, the quantitative variables can be rank-ordered (Longest & Vaisey, 2008).

In contrast, csQCA draws on the coding of qualitative data "into" or "out of" specific categories. The word "crisp" is used here because of the clean and clear demarcation between categories, wherein each variable of interest is coded either as 1 or 0. For instance, in a recent study of Formula 1 (F1) racing firms' business models, Aversa et al. (2015) coded F1 firms' business model into two categories: firms that adopted the specific business model (1) and firms that did not adopt the specific business model (0). Similarly, Rey-Martí, Porcar, and Mas-Tur (2015) coded

women entrepreneurs' risk propensity into two categories: presence of risk propensity (1) and absence of risk propensity (0). In this article, we specifically illustrate the application of this csQCA version of QCA.

Building on the principles of "conjunctural causality" (Garcia-Castro & Francoeur, 2016), QCA utilizes the Boolean algebra and manipulation of logical statements to determine the equifinal combination of variables for a given outcome. QCA is known to be robust against small sample size (Greckhamer, Misangyi, & Fiss, 2013) and can, therefore, be used for large unstructured data sets common to individual-level studies as well as for smaller data sets seen in firm-level studies. It focuses on two aspects—whether a given outcome is observed when a particular configuration is present and whether the outcome is not observed when the configuration is absent. Thus, it has the potential to cover for both necessary and sufficient conditions for outcomes. A "necessary condition" indicates that a specific outcome can be observed only when a particular condition is present. A "sufficient condition" indicates that the presence of a specific condition implies a specific outcome (Ragin, 2008; Garcia-Castro & Francoeur, 2016). Initially, QCA generates a "truth table" that lists all potential configurations. Later, it applies the Quine-McCluskey algorithm[2] (Aversa et al., 2015) to arrive at a parsimonious set of logically relevant configurations. QCA also generates statistics on "consistency" and "coverage" to ascertain if tested configurations are statistically rigorous enough. Consistency ranges between 0 and 1 and indicates how closely a specific configuration and outcome relationship is approximated. This figure is seen as being similar to the significance score in conventional regression, and high consistency indicates that a specific configuration-outcome relationship is significant (Garcia-Castro & Francoeur, 2016). Studies often use the consistency score of 0.75 as a threshold value (Garcia-Castro & Francoeur, 2016). Coverage evaluates the degree to which a particular configuration accounts for instances of an outcome (Garcia-Castro & Francoeur, 2016).

QCA presents novel opportunities for qualitative research in IS. Studies following qualitative IS research traditions have typically drawn on sensitizing theoretical lenses to explore, explain, and illuminate a phenomenon. While this approach provides rich insights into the role played by different conditions in producing an outcome, QCA, by virtue of its potential to also inductively explicate in detail how a combination of several variables or conditions correspond to the said outcome (Liu, Mezei, Kostakos, & Li, 2017), can add a crucial new dimension to qualitative IS research. QCA's potential, of course, fundamentally hinges on access to large qualitative data sets about a given phenomenon. Several scenarios from our own published IS research illustrate the potential value of QCA. Drawing on the notion of organizational culture and subculture, Ravishankar, Pan, & Leidner (2011) focused on the role played by three subcultures, namely, enhancing, countercultural, and chameleon in influencing the outcome of an IS implementation. They found that within different subcultural environments, specific conditions such as routinization, cultural interpretations, and perceived ease of use influenced implementation outcomes. The authors were able to make a set of useful claims about how key conditions at different levels influenced implementation outcomes. In this particular case, the application of QCA could potentially also help explore the configurations of conditions at various levels (organization, unit, individual), which correspond to successful and unsuccessful implementation over time.

Similarly, upcoming research areas such as the digital economy of emerging markets could usefully leverage QCA. In a recent study, Sandeep & Ravishankar (2016) found that frame bridging, frame amplification, frame extension, and frame transformation strategies help impact sourcing firms address tensions and misunderstandings with local communities during the process of providing IT outsourcing services from rural areas. Here, the application of QCA could help tease out particular configurations of frame elements that contribute to the alleviation of the local community's concerns at various phases of IT strategy implementation. IS studies could also apply QCA tools to explore the configurations of motivational factors or conditions that contribute to particular outcomes at different stages of IT implementation (Zimmerman and Ravishankar, 2016).

3 | APPLYING QCA TO AN EMERGING IS PHENOMENON

As indicated above, QCA can provide insights into how a complex configuration of conditions influences outcomes over time. In order to highlight and explain how QCA can be a useful addition to existing qualitative research

methods' toolkit for IS scholars, this paper applies csQCA to the green IS phenomenon. In particular, the paper draws on information readily available in the form of unstructured data sources and shows how csQCA can explain and elaborate on IT firms' green IS strategies that are associated with profitability.

Green IS is broadly concerned with the design and implementation of IS that contribute to environmental sustainability (Brocke, Watson, Dwyer, Elliot, & Melville, 2012). Applying csQCA to green IS strategies of firms illuminates the relationship between the combinatorial presence of different variables (in this case, different green IS business model configurations) and particular outcomes (in this case, increase and decrease in firm profitability). Conventional case study–based qualitative research may, of course, explore how firms deploy green IS business models. As we will argue in subsequent sections of the paper, QCA can complement such an exploration by identifying the specific configuration of green IS business model elements that closely correspond with improved and diminished firm performance.

4 | RESEARCH APPROACH AND ANALYSIS

To illustrate how QCA can help advance qualitative IS research methods, we conducted an in-depth exploration and analysis of various archived reports relating to 25 well-known IT companies. The use of archived reports facilitates a larger sample size compared with conventional qualitative studies in IS research. We conducted our research in three stages. In the first stage, we conceptualized the study from a business model perspective. Business models are concerned with the "design or architecture of the value creation, delivery, and capture mechanisms" (Teece, 2010). Drawing on Baden-Fuller and Mangematin's (2013) work on business models, we identified four central elements of a green IS business model—identifying customers (CI), customer engagement or value proposition (VP), monetization (MON), and value chain linkages (VCL).

In the second stage, we conducted text mining on our data sources (ie, we did a full and thorough search for "environment"-related content) (see Figure A2 in the appendix) to identify material corresponding to green IS. We specifically focused on descriptions of environment-friendly actions, and further examined the broader context of these depictions to make sure that we did not miss relevant words and paragraphs. We then used conventional open coding techniques to extract from the firms' self-reported green IS strategies empirical material corresponding to the above four green IS business model elements. Following this coding, we applied csQCA to empirically examine the different green IS-based configurations that corresponded with an increase and decrease in firm profitability. In other words, we produced cross-sectional snapshots of green IS business model configurations that were associated with positive and negative financial outcomes.

4.1 | Research sample

IT firms are increasingly introducing new environment-friendly products and technologies, and the idea of green IS/IT is gaining in significance (Gholami, Watson, Molla, Hasan, & Bjørn-Andersen, 2016). We arrived at our sample of green IT firms based on rankings such as the Newsweek Green Ranking, which has been ranking firms since 2009 for their greening initiatives. We specifically selected firms whose sustainability initiatives are publicly documented in the form of corporate sustainability and responsibility reports. Such reports follow the global reporting initiative (GRI) guidelines and are frequently audited, thus addressing credibility concerns. Some firms regularly respond to questionnaires from agencies such as the Carbon Disclosure Project (CDP). We also included such responses in our sample. Additionally, in our analysis, we included news articles and interviews featuring senior executives such as CEOs, Chief Sustainability Officers (CSOs), and third-party analyses of firms' sustainability initiatives (see Table 1 for the full list of data sources). The financial data pertaining to firm profitability was compiled from multiple data sources such as WRDS (Compustat) and Wolfram Alpha.

TABLE 1 Data sources

Unstructured Data Sources	Access
Annual sustainability reports, corporate responsibility reports, and environmental reports	Available on GRI report registry (http://database. globalreporting.org/), the Corporate Register (http://www.corporateregister.com/), and firms' websites
Firms' responses to CDP questionnaires (see https://www.cdp.net/en/guidance for details on the questionnaire)	Firms' websites, CDP database
News articles detailing firms' sustainability initiatives and interviews featuring senior executives such as CEOs and Chief Sustainability Officers (CSOs)	Available through FACTIVA search
Third-party analyses of firms' sustainability initiatives	(https://www.greenbiz.com/) (www.environmentalleader.com) (www.csrwire.com) (http://www.earthtimes.org/) (https://cleantechnica.com/) (https://www.treehugger.com/)

Our final sample comprised 25 IT firms, which have been engaged in sustainability initiatives for more than a decade (see Table 2 for the full list of sampled firms). Our sample exhibits an unbalanced panel structure since we have larger unstructured datasets for some firms. For instance, we have more information about firms such as IBM, which has focused on greening and environmental sustainability since the early 1990s.

4.2 | Coding

The two authors followed a conventional open coding scheme and coded the unstructured data obtained from the sources listed in Table 1. The authors ensured high interrater reliability and resolved any differences after detailed deliberation (for details of coding, please see Appendix). Drawing on Baden-Fuller and Mangematin (2013), we conceptualized green IS business models as being distinctive in terms of four key elements: identifying customers (CI), customer engagement or value proposition (VP), monetization (MON), and value chain linkages (VCL). In our coding, we drew on two frameworks extensively utilized in green IS literature—natural-resource-based view (NRBV) and Hart's sustainability framework[3]—to identify specific actions that could be mapped to each of the four business model elements. Firstly, we coded instances of targeting emerging clean technologies (eg, green data centres, green computing devices, smart grids, and carbon management systems) at new markets and specific customer segments as identifying customers (CI). Secondly, we coded instances of product stewardship, ie, the development of ecofriendly IS products that are easier to recover, reuse, or recycle as representative of value proposition (VP). Thirdly, we coded examples of crossfunctional coordination efforts (ie, collaboration with other firms, suppliers, and partners) to deliver green IS as VCL. Finally, we coded firms' initiatives that offered different pricing and payment mechanisms (eg, discounts and incentives) and financing instruments to help customers purchase green IS products as MON. After the coding and mapping of unstructured data to the four business model elements, we chose return on assets (ROA)—ratio of net income to total assets—as the outcome variable and the measure of firm profitability. We classified firms as high-performing firms if their profitability improved or was stable and as low-performing firms if their profitability declined (see Table 3 for an illustration of coding and Figure 1 for a step-by-step outline of the research approach adopted).

5 | RESULTS

As noted earlier, in our analysis, we specifically applied a variant of QCA termed as csQCA. In csQCA, every case is coded as either "in" or "out" of predecided categories. Thus, a firm is either "in" or 1 or "out" or 0 for different

TABLE 2 List of sampled firms

Firms	Availability of Sustainability Reports
International Business Machine (IBM)	2002-2015
Hewlett Packard (HP)	2001-2015[a]
Dell	1998-2015[a]
Cisco	2006-2015
Computer Associates	2008-2015[a]
Autodesk	2008-2015[a]
Apple	2008-2015
Adobe	2003-2015[a]
Analog Devices	2010-2015
Intel	2002-2015[a]
Microsoft	2003-2015[a]
EMC Corporation	2009-2015[a]
Xerox	2007-2015[a]
Qualcomm	2006-2015[a]
Applied Materials	2005-2015[a]
Motorola	2002-2015[a]
Symantec	2008-2015[a]
Texas Instruments	2005-2015[a]
Seagate	2005-2015[a]
SAP	2008-2015
Samsung	2000-2015[a]
Tata Consultancy Services	2007-2015
Ericsson	2007-2015[a]
TE Connectivity Limited	2010-2015
Xilinx	2005-2015[a]

[a]We could not find publicly available sustainability reports for all years. In their absence, we relied on other archived sources listed in Table 1.

categories. csQCA uses Boolean algebra and set theory to identify the combination of variables (business model elements, in this case) that correspond closely with a given outcome (increase and decrease in profitability, in this case). We applied csQCA to better explicate the relationship between green IS business model elements and profitability. First, using csQCA and following Aversa et al. (2015), we computed the aggregate configuration(s) for different years. Through this approach, we were able to ascertain the stability and evolution of the configurations corresponding to improved and reduced profitability. Our results show a temporal evolution of configurations and suggest that despite variations for different years, specific business model elements could play a central role in contributing to positive payoffs from green IS. Second, we computed the configurations that corresponded with improved and reduced profitability for each firm over the duration of the sample.

In Tables 4 and 6, we report "overall," "raw," and "unique" coverage (Ragin, 2008; Aversa et al., 2015). Overall coverage is the aggregate of coverage corresponding to every identified configuration. It is plausible that more than one configuration could correspond to improved and reduced profitability. Raw coverage is defined as the number corresponding to a specific configuration and outcome divided by the total number of a particular outcome. Unique coverage indicates the proportion of outcome uniquely corresponding to a specific outcome.

5.1 | Aggregate green IS configurations for improved profitability

Table 4 details the aggregate business model configurations that corresponded with improvements in profitability across our sample of 25 IT firms. These yearly configurations can help analyse and potentially better understand the green IS actions of firms that may have produced profitability improvements in a given year as well as across several years. They also provide useful information about changes to configurations over a 12-year period (2004-2015). We consider four observations (from Table 4) for illustrative purposes and elaborate on the analytical and explanatory potential of the csQCA approach (See Table 5). From a methodological standpoint, it is important to recognize that unstructured qualitative data form the basic building blocks of these observations.

TABLE 3 Illustration of coding of unstructured data

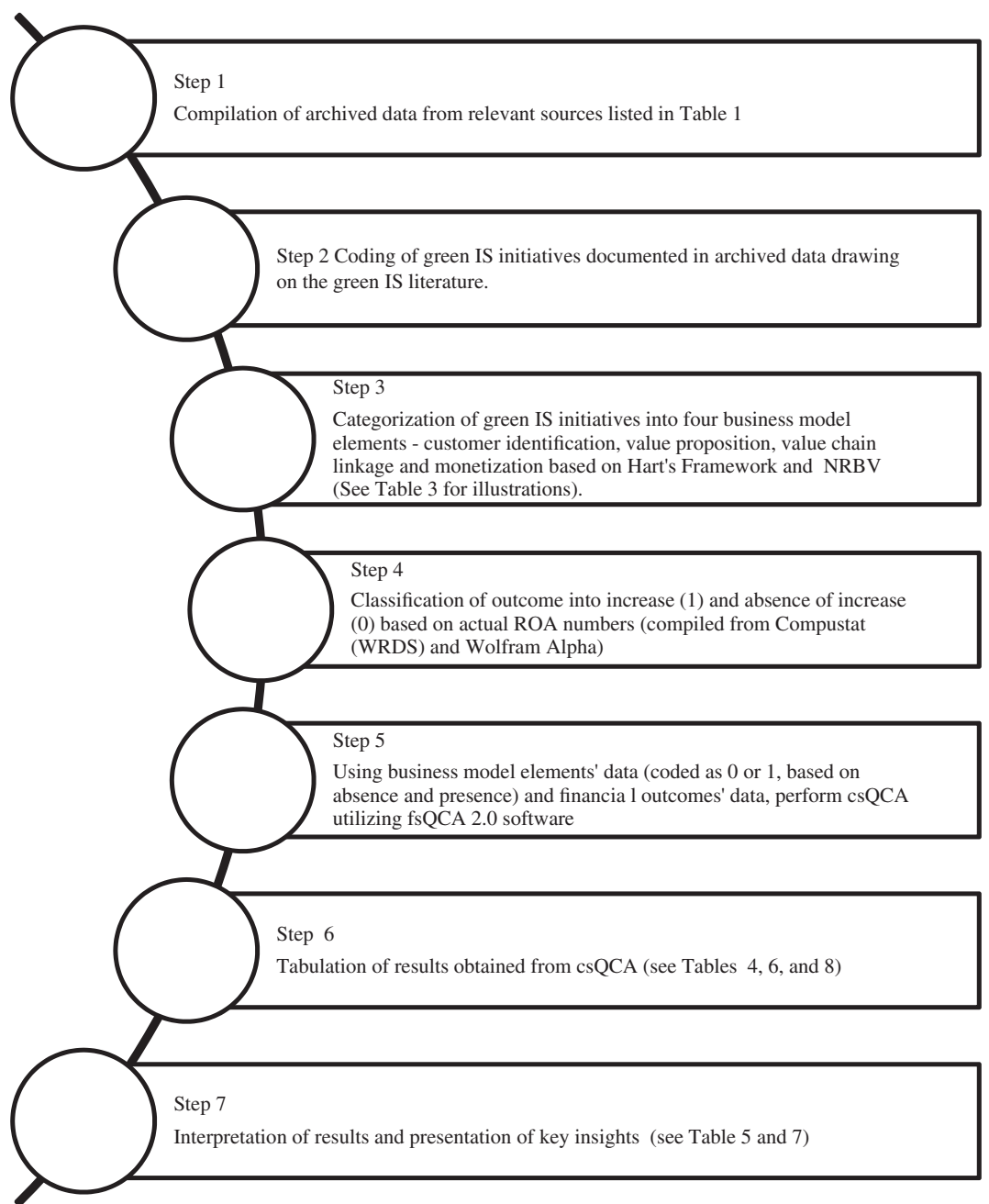| Firms	Excerpts from Archived Data Source	Coding
International Business Machine (IBM)	"IBM became the first computer manufacturer to establish a product recycling service in the United States for consumers and small businesses" (IBM Corporate Responsibility Report, 2002)	**VP:** The product recycling initiative launched by IBM offers a new value proposition to customers as it alleviates their concerns about e-waste. It is an illustration of product stewardship and is coded as value proposition.
Dell	"In June 2005, Dell introduced the OptiPlex GX520 and GX620 desktops which were designed with lead-free motherboards, chassis and power supplies. Dell will continue to launch new RoHS-compliant2 products in advance of the RoHS implementation deadline A marketing segment survey was conducted in 2003 to help determine which customer segments are now or soon will be ready to transition to lead-free solutions, and which segments prefer to wait until the new technology has a more proven track record. Dell's lead-free product roadmap is closely aligned with customer expectations. These surveys assist Dell's product development extended teams to design systems that will use lead-free components. (Dell Sustainability Report 2003, 2005)	**CI:** Dell followed a systematic scientific approach. It conducted a marketing survey to identify promising customer segments and subsequently drew on the findings of the survey to launch lead-free products.
Qualcomm	"Because building a sustainable, conflict free mineral supply chain requires coordinated effort and investment from all sectors of society, we continued collaborating as part of the Conflict Free Sourcing Initiative, the Public Private Alliance for Responsible Minerals Trade, the Responsible Sourcing Network's multi-stakeholder group and the ITRI Tin Supply Chain Initiative (iTSCi)" (Qualcomm Sustainability Report 2014)	**VCL:** Following Aversa, Furnari, & Haefliger's (2015) categorization of linkages between the focal firm and its suppliers and partners as value chain linkages, we categorized Qualcomm's collaborative efforts as value chain linkages
Intel	"Intel-backed PC purchase programs enable governments to provide PCs at a more affordable price, allowing thousands of teachers and students to gain access to them for the first time" (Intel Sustainability Report, 2010)	**MON:** Economic support and incentives, pricing programs constitute monetization (Aversa et al., 2015). Here, Intel runs a program that facilitates the discounted large-scale purchase of computing devices by governments, which would also help the company as PCs often use Intel processors.

Step 1
Compilation of archived data from relevant sources listed in Table 1

Step 2 Coding of green IS initiatives documented in archived data drawing on the green IS literature.

Step 3
Categorization of green IS initiatives into four business model elements - customer identification, value proposition, value chain linkage and monetization based on Hart's Framework and NRBV (See Table 3 for illustrations).

Step 4
Classification of outcome into increase (1) and absence of increase (0) based on actual ROA numbers (compiled from Compustat (WRDS) and Wolfram Alpha)

Step 5
Using business model elements' data (coded as 0 or 1, based on absence and presence) and financia l outcomes' data, perform csQCA utilizing fsQCA 2.0 software

Step 6
Tabulation of results obtained from csQCA (see Tables 4, 6, and 8)

Step 7
Interpretation of results and presentation of key insights (see Table 5 and 7)

FIGURE 1 Process of leveraging unstructured data

5.1.1 | Observation 1

The customer identification (CI) element dominates the configurations in Table 4. As noted earlier, CI cover actions of targeting emerging clean technologies at new markets and specific customer segments. Thus, the presence of CI in many of the parsimonious configurations indicates that firms focused heavily on selling emerging green technologies to new markets and customers over a 12-year period and that such a focus helped them improve their profits.

TABLE 4 Parsimonious solutions for improvements in profitability: csQCA

Configuration	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Parsimonious	CI*~MON	~MON	~VCL ~VP CI*~MON ~CI*MON	CI*~MON ~VP	~VP*~VCL CI*~VCL CI*~VP	CI*~VCL CI*~VP	CI*~VCL CI*~VP ~CI*MON	~VP*~VCL CI*~VCL ~CI*MON	CI*~VCL CI*~VP	CI*~VCL CI*~VP	CI*~VCL	~CI*VP*~VCL
Overall coverage	0.25	0.90	0.60	0.50	0.36	0.33	0.545	0.375	0.154	0.25	0.100	0.285
Raw coverage	0.25	0.90	0.20 0.20 0.20 0.10	0.30 0.20	0.18 0.09 0.09	0.165 0.165	0.181 0.090 0.272	0.125 0.062 0.187	0.077 0.077	0.125 0.125	0.100	0.285
Unique coverage	0.25	0.90	0.1 0.1 0.2 0.1	0.30 0.20	0.18 0.09 0.09	0.165 0.165	0.181 0.090 0.272	0.125 0.062 0.187	0.077 0.077	0.125 0.125	0.100	0.285
Consistency	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
N	7	12	13	14	21	22	23	23	23	23	23	23
Description	CI in combination with an absence of MON corresponds to improved profitability	Absence of MON corresponds to improved profitability	Four different configurations correspond to profitability improvements	Two different configurations correspond to profitability improvements	Three different configurations correspond to profitability improvements	Two different configurations correspond to profitability improvements	Three different configurations correspond to profitability improvements	Three different configurations correspond to profitability improvements	Two different configurations correspond to profitability improvements	Two different configurations correspond to profitability improvements	CI in combination with absence of VCL correspond to profitability improvements	Absence of CI, VCL, in combination with VP correspond to profitability improvements

Notes: The parsimonious solution is considered a more simplified and pragmatic solution (Baumgartner, 2015). Consistency statistics here refer to both the overall solution and the individual solutions. * indicates combination, for instance CI*MON indicates a configuration that is a combination of customer identification and monetization. ~ indicates the absence of specific elements. For example, CI*~MON indicates a configuration that is a combination of customer identification and the absence of monetization.

Abbreviations: CI, customer identification; MON, monetization; VCL, value chain linkages; VP, value proposition.

TABLE 5 Summary of csQCA for improved profitability

Illustrative observations from Table 4	CI is a dominant element	The ~CI*VP*~VCL configuration in 2015	~CI*MON configuration in 2010 and 2011	The ~VCL element
Interpretation	Targeting green technologies at new markets and customers have contributed to improvements in profitability	Focus on value proposition and staying away from (a) targeting green technologies at new markets and customers; (b) building collaborative links	Focus on financial incentives in difficult macroeconomic conditions	Staying away from partnerships
Opportunities for future research	Investigate the role of specific green IS products in delivering profits	Investigate the realization of competitive advantage in mature green IS markets	Investigate how IT firms deliver value proposition in times of economic turbulence	Investigate why IT firms decide to stay away from collaborative strategies

The implication, therefore, is that green technologies, as part of a firm's product portfolio in new markets, have contributed to improvements in firm profitability. Apart from providing initial evidence for the positive role of green IS solutions in IT firms' strategies, this insight can also be used to further advance green IS research. For instance, building on the importance of CI in this csQCA analysis, future research could investigate the degree to which deployments of specific green IS products such as green data centres, smart grids, personal computing devices with energy star labels, and green-packaged IT (Corbett, 2010) contribute to improved profits and competitive advantage.

5.1.2 | Observation 2

The csQCA analysis for 2015 (Table 4) shows that the group of firms whose profitability increased year-on-year adopted the configuration ~CI*VP*~VCL to a reasonably significant extent. In fundamental terms, this configuration points to clear evidence of an organizational emphasis on the development of ecofriendly IS products (VP), a simultaneous de-emphasis on crossfunctional coordination efforts (VCL), and new market-targeting and customer-targeting activities (CI). Thus, it suggests that firms' profitability could be closely linked to the delivery of a stronger value proposition such as green energy options that reduce upfront project development costs for customers and help them improve their profitability through cost-savings. For instance, qualitative data from Apple, which was one of the firms whose profitability improved in 2015, showed an explicit focus on "greening" clients' internal operations and reducing the operational costs of clients in the energy and utility sector:

> Apple worked with NV Energy and the Nevada Utility Commission to create a green energy option open to all commercial customers that does not require the customer to fund project development upfront (Apple Environmental Responsibility Report, 2015).

A possible interpretation of this ~CI*VP*~VCL configuration is that as green IS market matures and as clients' awareness levels increase, profitability could depend to a significant extent on the ability to deliver a stronger client value proposition. From a practical point of view, this insight also suggests that firms may increasingly need to (re) orient their efforts towards developing the VP dimension of their green IS business models.

Table 4 also indicates absence of the MON element in configurations corresponding to the years (2004-2007). One possible explanation for how firms were able to achieve improvements in profitability while ignoring monetization could be that customers of our sampled firms saw clear value in green IS and were therefore ready to invest in green IS products and services. There is also an explanation grounded in the notion of product life cycles.[4] 2004 to

2007 were early years for green IS. The early adopters of technology products often tend to be affluent customers, who may not care much about price discounts. Consequently, monetization may not have been predominantly featured in configurations. However, in the later stages of the product life cycle, competition tends to be more price-driven and, thus, monetization strategies can become increasingly important.

5.1.3 | Observation 3

The csQCA analysis for 2010 and 2011 highlights ~CI*MON as an additional configuration possibly leading to improved profitability. This configuration is particularly interesting because it suggests the use of financing instruments to help customers purchase green IS products (ie, MON) in the immediate aftermath of the 2008 financial crisis. This finding could help show how successful green IS-focused IT firms mitigated the possible impacts of big jolts in the macroeconomic environment on clients by offering them monetary incentives. For instance, initiatives detailed in Intel's 2010 sustainability report indicate that the company proactively promoted affordable paperless education (ie, the substitution of physical artefacts with computing products in the education sector) through various discounts:

> Intel-backed PC purchase programs enable governments to provide PCs at a more affordable price, allowing thousands of teachers and students to gain access to them for the first time. In 2010, Intel partnered with Microsoft, Internet provider Safaricom, Equity Bank, and the Kenya Institute of Education on an initiative designed to make laptop and Internet connectivity more affordable to some 260,000 teachers in Kenya. Intel also helped the department of education in South Africa to launch the Teacher Laptop Initiative, aimed at enabling 400,000 teachers across the country to purchase affordable laptops (Intel Sustainability Report, 2010).

In a broader sense, the presence of the MON element in the different configurations can help better understand how IT firms foster green IS purchase intentions among consumers even in the backdrop of challenging macroeconomic conditions.

5.1.4 | Observation 4

The green IS configurations in Table 4 also suggests that the green IS business model configuration element VCL was absent in several yearly configurations that corresponded with profitability over the 2004-2015 period. This observation is particularly intriguing since it indicates minimal collaboration with other firms, suppliers, and partners in delivering green IS (ie, minimal value chain linkages). However, except in 2006, the absence of VCL (indicated by ~VCL in Table 4), on its own does not correspond with profitability improvements. Rather, it is the combination of ~VCL with additional elements (eg, CI*~VCL in 2012) that appear to help firms. The implication here is that successful green IS business models of supplier firms seem to have depended little on collaborative efforts and linkages (VCL) in contrast with their efforts aimed at selling to new markets and customers (CI) and offering monetary incentives (MON).

5.2 | Aggregate green IS configurations for decline in profitability

Table 6 details the aggregate business model configurations that corresponded with a decline in profitability across our sample of 25 IT firms. These yearly configurations can help analyse and potentially better understand the green IS actions of firms that may have produced a decline in profitability in a given year as well as across several years. They also provide useful information about changes to configurations over a 12-year period (2004-2015). In detailing

TABLE 6 Parsimonious solutions for decline in profitability: csQCA

Configuration	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Parsimonious Solution	~CI*MON	~CI*MON	NO SOLUTION	VP*~VCL ~CI*MON	~CI*MON ~ CI*VP*~VCL ~CI*~VP*VCL	~CI*MON ~ CI*VP*~VCL ~CI*~VP*VCL	~VP*~VCL CI*MON	~CI*VP*~VCL	~CI*~VCL ~CI*MON	~CI*~VCL ~CI*~VP	~CI*VP ~CI*MON	CI*~VCL ~VP*VCL
Overall coverage	0.33	0.50		0.50	0.50	0.33	0.357	0.22	0.58	0.35	0.428	0.235
Raw coverage	0.33	0.50		0.17 0.33	0.167 0.167 0.167	0.11 0.11 0.11	0.142 0.214	0.22	0.33 0.25	0.235 0.235	0.285 0.142	0.058 0.176
Unique coverage	0.33	0.50		0.17 0.33	0.167 0.167 0.167	0.11 0.11 0.11	0.142 0.214	0.22	0.33 0.25	0.117 0.117	0.285 0.142	0.058 0.176
Consistency	1.00	1.00		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
N	7	12	13	14	21	22	23	23	23	23	23	23
Description	Absence of CI in combination with MON corresponds to a decline in profitability	Absence of CI in combination with MON corresponds to a decline in profitability		Two different configurations correspond to a decline in profitability	Three different configurations correspond to a decline in profitability	Three different configurations correspond to a decline in profitability	Two different configurations correspond to a decline in profitability	Only one configuration correspond to a decline in profitability	Two different configurations correspond to a decline in profitability	Two different configurations correspond to a decline in profitability	Two different configurations correspond to a decline in profitability	Two different configurations correspond to a decline in profitability

Notes: Consistency statistics here refer to both the overall solution and the individual solutions. * indicates combination, for instance CI*MON indicates a configuration which is a combination of customer identification and monetization. ~ indicates the absence of specific elements. For example, CI*~ MON indicates a configuration which is a combination of customer identification and the absence of monetization.

Abbreviations: CI, customer identification; MON, monetization; VCL, value chain linkages; VP, value proposition.

the configurations that correspond with a decline in profitability, Table 6 offers a useful comparison with Table 4, which details configurations corresponding to increase in profitability. In particular, the comparison can help understand whether the absence of key configurations (the presence of which contributes to improved profitability) corresponds to a decline in profitability. It also brings into sharp focus the differences between configurations that correspond with improved profitability and those that correspond with a decline, highlighting the need for a deeper analysis. Below, we consider four observations (from Table 6) for illustrative purposes and elaborate on the analytical and explanatory potential of the csQCA approach (See Table 7).

5.2.1 | Observation 1

The absence of the customer identification element (indicated by ~CI) is particularly noticeable in Table 6. This absence contrasts with Table 4, in which we saw that the presence of the CI element was a noteworthy contributor to increased profitability. The dominance of ~CI in Table 6 (seen in conjunction with the dominance of CI in Table 4) suggests that a failure to strategically identify and target markets positively disposed towards green technologies could be an important contributor to diminishing profits. One possible interpretation of the dominance of ~CI in Table 6 is that firms may have been excessively focused on other elements of the business model (eg, pricing and discounts—indicated by MON in Table 6) to the detriment of developing a clear sense of the market segments they could target (ie, CI). This plausible interpretation opens up further opportunities for research. For instance, future research could critically examine how firms embed market-centric approaches in the strategic planning and implementation of their green IS offerings.

5.2.2 | Observation 2

The csQCA configurations in Table 6 suggest that the MON element was heavily implicated in firms' profitability decline. The presence of the MON element in several profitability decline configurations indicates that the offering of discounts and financial incentives has not really worked for IT firms in our sample. In Table 6, we also see that the MON element often appears in combination with ~CI (ie, the absence of CI). This combination points to one possible reason for the firms' decline in profitability: they offered monetary incentives to many of their customers somewhat indiscriminately without a clear identification of target market segments. For instance, HP's 2009 sustainability highlights a big focus on discounts and incentives.

> Through our HP PartnerONE Diversity Network, we provided marketing and sales support to more than 100 diverse resellers in 2009. The support comes in the form of marketing subsidies, discounted products... (HP Global Citizenship Report, 2009).

From a practical point of view, the close correspondence of the ~CI*MON combination with profitability decline highlights the need for firms to first carefully identify their different target customer segments and then customize their financial incentive offerings to the different segments.

5.2.3 | Observation 3

The csQCA analysis for 2008 and 2009 (Table 6) show that the group of firms whose profitability declined year-on-year adopted the configuration ~CI*VP*~VCL. As mentioned in our observations linked to Table 4, this configuration indicates an organizational emphasis on the development of ecofriendly IS products (VP), a simultaneous de-

TABLE 7 Summary of csQCA for decline in profitability

Illustrative observations from Table 6	Absence of CI (~CI) is a dominant element	The MON element	The ~CI*VP*~VCL configuration in 2008 and 2009	The ~VCL element
Interpretation	Failure to strategically identify and target markets could be an important contributor to decline in profitability	The offering of discounts and financial incentives may not be enough to attract customers to green IS products and technologies	The configuration's contribution to a decrease in profitability could imply that customers did not respond positively towards value-added green IS products in the aftermath of the 2007 financial crisis	Limited collaborations, when combined with a poor focus on identifying target customer segments, can lead to a decline in profitability
Opportunities for future research	Investigate the processes through which firms embed market-centric approaches in the strategic planning and implementation of their green IS offerings	Investigate how firms can identify different target customer segments and customize their financial incentives to the differing needs of each segment	Investigate the impact of changes in the macroeconomic environment on customers' attitudes to green IS	Investigate the impact of combining specific green IS business model elements on financial performance

emphasis on crossfunctional coordination efforts (VCL), and new market-targeting and customer-targeting activities (CI). This configuration's presence in both Tables 4 and 6 is intriguing since it demonstrates the configuration's correspondence with both an increase as well as a decline in profitability. However, there is a time dimension to the puzzle that is potentially illuminating. The ~CI*VP*~VCL configuration appeared multiple times in Table 4. However, they were associated with a decline in profitability (Table 6) specifically in 2008 and 2009, which suggests that the green IS value proposition (VP) offerings did not work as well as the IT firms in our sample anticipated. A possible explanation is that in the aftermath of the global financial crisis of 2007 consumers (both B2B as well as B2C) were cautious in their spending and are less positively disposed towards unique characteristics of products.

Overall, correspondence of the ~CI*VP*~VCL configuration with a decrease in profitability in 2008 and 2009 provides some evidence for the lack of improvisation in firms' green IS strategies. The strategies appear to have failed to respond creatively to the impact of macroeconomic upheavals on customer spending, leading to a possible negative impact on the firms' financial performance. This configuration also highlights an opportunity for future research to closely examine how changes in the macroeconomic environments influence customers' attitudes towards green IS products.

5.2.4 | Observation 4

Similar to Table 4, it is noticeable in Table 6 that the green IS business model configuration element VCL was absent (indicated by ~VCL) in several yearly configurations. It suggests very limited collaboration with other firms in the process of delivering green IS products. This observation is interesting as it creates the impression that limited collaboration (ie, ~VCL) corresponded with both an increase as well as a decrease in profitability. However, results of the csQCA show that while an element like CI was often combined with the absence of VCL (indicated by ~VCL) in configurations associated with improvement in profitability (Table 4), the absence of CI (indicated by ~CI) combined with absence of VCL (indicated by ~VCL) in configurations associated with a decline in profitability (Table 6). In other words, the combination of ~CI and ~VCL in several Table 6 configurations helps demonstrate the fact that firms whose profitability declined neither had a cohesive strategy for identifying target customer segments nor did they

TABLE 8 Firm-level QCA analysis

Configuration	IBM
Parsimonious	~MON*VCL
	~MON*CI
Complex	~MON*VCL*VP*CI
Overall coverage	0.63
Raw coverage	0.63
Consistency	1.00

Notes: * indicates combination, for instance, CI*MON indicates a configuration which includes a combination of customer identification and monetization. ~ Indicates the absence of specific elements. Consistency statistics here refer to both the overall solution and the individual solutions.
Abbreviations: CI, customer identification; MON, monetization; VCL, value chain linkages; VP, value proposition.

build collaborative links. This insight reiterates that the point that more than individual business model elements, it is the configurations that are ultimately implicated in financial success (and failure) of green IS strategies.

5.3 | Firm-level analysis

In our analysis, we also pooled data for individual IT firms. For illustrative purposes, we consider the csQCA results of one firm—IBM—in Table 8. It shows IBM's aggregate configuration that corresponded with an increase in profitability across the years. Although the configuration builds on qualitative data from all years and potentially leads to a loss of the data's longitudinal characteristics, it still provides useful insights into IBM's green IS strategies and their impact on profitability outcomes. Our sample size (N) for IBM is comparable with the sample size in recent QCA studies (eg, Aversa et al., 2015). Table 8 presents both the parsimonious and complex configurations (configurations ignoring counterfactuals) that correspond to improved profitability at IBM. It indicates that IBM combined CI, VP, and VCL, but did not use MON. A logical extrapolation from this configuration is that IBM appears to have leveraged its position as a leading global firm and focused on building strong collaborative links (VCL) and offering high VP to targeted customer segments (CI). The absence of the MON element (indicated by ~MON) suggests that the firm has offered its green IS products at a price premium and refrained from offering discounts and incentives. The qualitative material produced by IBM clearly illustrates this focus on premium offerings. For instance:

> "The PC era has ended, representing a fundamental shift in the technology requirements of the company's clients. IBM is well positioned to provide its enterprise clients the open technologies and high-value solutions they will need to compete. IBM is leveraging its leadership position in the convergence of software and services, in SOA, in virtualization, in high-performance chips, in open and modular IT—continuing its shift from commoditizing segments to higher value segments with better profit opportunity" (IBM 10 k filings, 2007).

6 | DISCUSSION

The analysis above shows how csQCA can provide detailed and potentially valuable insights into the green IS strategies of firms and their impact on firm profitability. They also suggest "equifinality" at play (ie, different configurations of green IS business model elements may underpin an increase [and decrease] in profitability). How can the csQCA approach

complement and advance existing methods in qualitative IS research? This question is probably best answered by considering the type of insights a typical qualitative multiple case study design may have generated in this green IS scenario. In the case study method, researchers would typically interview multiple respondents across firms and explore the various green IS initiatives. The questions would focus on the evolution of these initiatives. In addition, researchers would try to get a good sense of the type of initiatives that have succeeded and failed. Following the interviews, researchers would code the transcripts and use inductive reasoning to ascertain the conditions that could have contributed to improved profitability or diminished profitability. This approach often suffers from problems such as respondents' short memory and the associated issues around the quality of the gathered longitudinal data. Besides, the small sample size issues often encountered in firm-level case studies would make it hard to explicate the combination of business model elements associated with improvements and decline in profitability. Research teams may, of course, turn to archived reports as well, but in the absence of a systematic analysis of the green IS empirical material, they may find it difficult to make robust and credible claims about multiple business model configuration pathways to profitability.

As we have shown, csQCA draws on the inherent strengths of qualitative coding techniques and facilitates a deeper investigation of available qualitative material. In our case, conventional open coding techniques helped prepare the qualitative empirical material for the csQCA process. Subsequently, the csQCA process revealed how particular configurations of business model elements corresponded with improved profitability and diminished profitability. In our study, we did not formulate a priori hypotheses, as we adopted an exploratory approach. But it is certainly possible to formulate a priori hypotheses on relationships between configurations and outcomes of interests and test relationships. Thus, csQCA can also be utilized to test hypotheses that postulate specific relationships between different configurations and various outcome variables, although it is important to recognize that hypothesizing such relationships is more complex and will have to take into account the joint impact of several variables. With csQCA, hypothesized relationships are not restricted to the mere direction (positive or negative) of the relationship between an explanatory variable and outcome variable. Thus, csQCA facilitates a better understanding of the configurations that may contribute to both positive and negative payoffs from emerging digital technologies. The equifinality angle to csQCA approaches can strengthen the robustness claims of qualitative studies. Qualitative studies are often seen as taking a less definitive approach relative to quantitative studies. csQCA's revealing of several underlying paths to particular outcomes can, therefore, help improve qualitative approaches to enhance their analytical tightness. This approach also responds to calls for combining elements of qualitative and quantitative research and inductive and deductive reasoning (Sarker et al., 2018).

csQCA facilitates longitudinal qualitative studies of unstructured data obtained from archived data sources. Longitudinal data are notoriously difficult to obtain through primary interviews because of changes in the roles of executives and employee turnover. Archived reports fill this significant gap, and csQCA builds on such reports and help assess the contributions of different configurations to the realization of specific outcomes over time. When the rationales for certain configurations are unclear (eg, why was a particular combination of strategic IS deployments made?), it can inspire researchers to look for other credible sources or design more focused qualitative studies that may help better explain organizational decision-making. For instance, in our case, we turned to the 10-K filings of the IT firms in our sample to make better sense of some of the observed configurations.

6.1 | Implications for research

Our study offers two sets of implications. The first set of implications is from a methodological perspective. Qualitative research typically involves collecting data from respondents in firms through interviews and observations (Sarker, Xiao, & Beaulieu, 2013). Framing relevant and topical questions for data collection is often challenging as researchers have a limited knowledge of the particular research setting. Archived reports could obviously help, but the mere reading of such reports do not necessarily offer actionable insights for qualitative researchers. In such cases, the use of csQCA can help researchers develop more fruitful and deeper lines of inquiry in subsequent research projects. For instance, following the findings described in this paper, a research team conducting a case study of sustainability

|

strategies at IBM could focus specifically on the evolution of the company's green IS VCL, which appears to be a critical contributor to improved performance or focus on mechanisms underlying customer identification whose absence (~CI) appears to be a critical contributor to diminished performance. Similarly, researchers could also conduct more in-depth qualitative inquiries about elements that did not prove to be effective (eg, MON).

csQCA can help researchers reveal misalignments between perceptions of executives reported in primary interviews and findings arrived at via analyses of archival data. Such misalignments can sensitize researchers to the limitations of both archival data as well as primary interview data and thus motivate them to reconcile disparities. Knowledge of misalignments can help researchers undertaking action research projects to clearly show executives how organizational realities diverge from their expectations. This process may eventually encourage executives to open up and show greater involvement in the qualitative research process. csQCA also contributes to the data analysis phase of qualitative IS research. The dataset coded using conventional qualitative techniques could be utilized to test hypotheses grounded in existing theoretical frameworks or explore relationships to arrive at typologies that extend existing theories and help develop a deeper understanding of a given phenomenon. Although qualitative case studies can be used to test theories through (say) comparative case studies (Flyvbjerg, 2006), csQCA adds a further layer of nuance to theory testing in qualitative IS research by facilitating the discovery of the optimal combination of various independent variables contributing to a given outcome. Thus, the method also potentially contributes to theory building in qualitative IS research as it extends the focus from merely testing the significance of relationships to demonstrating how outcomes potentially stem from a configuration of contributing conditions. As we have noted earlier, csQCA can contribute to further enhancing the legitimacy of qualitative IS research and arguably, has a distinct advantage over quantitative studies that often restrict themselves to testing of the relationship between two variables, thus ignoring the role of configurations. Thus, we would argue that csQCA could be a powerful source for theory testing as well as extension.

csQCA can be utilized in different ways. In this study, we have computed effective configurations for improved and reduced profitability. Thus, csQCA offers the opportunity to conceptualize the presence and absence of a specific outcome as two distinct dimensions and to find out which configurations produce them (Aversa et al., 2015; Rey-Martí et al., 2015). On both sides of the quantitative-qualitative divide in IS research, there is an increasing emphasis on longitudinal studies. Consequently, many quantitative studies now use panel data. Collecting longitudinal data in qualitative research is clearly challenging. Interviewees often fail to recall events from the past with great precision. They may also remember events differently from how they unfolded in reality. Consequently, it is often difficult to develop persuasive and robust explanations of phenomena. By working on publicly archived reports that provide researchers access to information about processes, policies, product portfolios, and past performance, csQCA can help discover the optimal configuration of variables, the longitudinal evolution of configurations, the emergence of new configurations and the disappearance of older configurations in relation to particular outcomes. Qualitative methods are often recommended for answering "how" questions. csQCA, in conjunction with traditional qualitative methods, presents opportunities not only to answer "how" questions but also to explicate in detail the evolution of the underlying relationships over time. The focus on evolution of the underlying relationships over time is critical to developing a better understanding of the role IS plays in society and business. While it is known that IS and its roles are dynamic, many research projects are able to obtain only a snapshot of the relationship between variables—a problem that csQCA addresses by detailing the evolution of configurations.

The second set of implications relates to making contribution claims at multiple levels in qualitative IS research. Qualitative studies in IS are typically pegged at the firm-level (Sarker et al., 2013). However, unstructured data sources such as archived reports and online posts offer flexibility to examine different units of analysis. For instance, online posts or tweets could be used to operationalize studies using the individual as a unit of analysis. csQCA could be utilized in such a scenario to ascertain, for instance, how a configuration of specific personality traits as manifested in these posts produce certain outcome variables (eg, the number of followers). Likewise, unstructured data such as technology patent applications could be coded and analysed using csQCA to understand what configuration of patent application characteristics correspond to actual patents. Unstructured data and csQCA offer novel possibilities to examine phenomena at the individual, firm, and industry levels. This process can facilitate a deeper

understanding of the extent to which configurations at different levels converge or diverge. Researchers could then collect more field data through interviews and observations to complement csQCA and explore the underlying reasons for convergence and divergence. In some contexts, text mining techniques can be used to generate continuous variables (eg, sentiment analysis). In such situations, other variants of QCA such as fsQCA can be used.

6.2 | Limitations and future research

While csQCA potentially introduces a new sophistication to qualitative IS research, it has a set of key limitations. The quality of insights generated from csQCA is largely dependent on the initial coding schemes and their reliability. In this sense, csQCA is as much about human interpretation as it is about rigorous data analysis. csQCA is also limited by the initial choice of a theoretical frame. Thus, although it can reveal how a complex configuration of variables correspond to a particular outcome, these variables are not exhaustive. For instance, the application of csQCA to the business model frame provides valuable insights into green IS strategies and firm performance, but this interpretation may not be in a position to take into account the influence of macroeconomic variables like the state of the global economy. In this paper, we identified four central elements of a green IS business model. Although they helped illustrate the value of the csQCA method, we acknowledge these elements are at a somewhat broad level. A more systematic classification of the specific green IS actions underpinning each of the four business model elements can offer deeper insights into combinations associated with firm profitability.

The csQCA method involves coding into only two categories. Arguably, other QCA techniques (eg, fuzzy set qualitative comparative analysis [fsQCA]) that handle continuous data can produce more precise combination of variables corresponding with profitability outcomes. However, we chose the simpler option of utilizing text mining and classifying our huge volume of unstructured data into discrete categories based on the presence of specific words or terms, rather than rating them on a scale. While past studies such as Montabon, Sroufe, and Narasimhan (2007) have utilized secondary data to rate initiatives on a scale of 1 to 10, we felt that such an approach entailed a high subjectivity bias compared with coding into two categories. In other words, relying on the frequency or counts of specific terms to rate might not be accurate, as they may not be the best criteria to evaluate the strength of any specific element in practice.

Future research could develop an index of divergence that examines disparities between QCA analyses of firms' archived initiatives and organizational respondents' perception of such initiatives. An index of divergence could be useful in comparing firms and could help researchers examine if such divergence has any impact on variations in outcomes. More broadly, there is a need for future research to facilitate complementarities between QCA and data analyses adopted in typical qualitative approaches. Further, QCA's linkages with innovative text mining techniques (Larsen, Monarchi, Hovorka, & Bailey, 2008; Korfiatis, Stamolampros, Kourouthanassis, & Sagiadinos, 2019), natural language processing (Müller, Junglas, Brocke, & Debortoli, 2016), sentiment analysis (Abbasi, Sarker, & Chiang, 2016), semantic keyword similarity method (Chiang, Grover, Liang, & Zhang, 2018), and the implications for IS qualitative research is a fruitful area of inquiry.

In conclusion, this paper introduced csQCA (a specific variant of QCA) as a useful addition to the qualitative IS research toolkit. Drawing on green IS data from IT firms, it illustrates the effectiveness of csQCA and its potential contribution to qualitative IS research. The paper demonstrates how csQCA can help researchers harness unstructured qualitative data and offer a more in-depth analysis and interpretation of emerging IS phenomena.

ORCID

Rohit Nishant https://orcid.org/0000-0001-7201-4901
M.N. Ravishankar https://orcid.org/0000-0003-3826-9403

ENDNOTES

[1] We thank anonymous reviewer 2 for this insight.

[2] The Quine-McCluskey algorithm is a method used for the minimization of the Boolean function.

[3] NRBV focuses on pollution prevention, product stewardship, and sustainable development. Hart's sustainability framework focuses on pollution prevention, product stewardship, and clean technology as strategies to enhance environmental sustainability.

[4] We thank anonymous reviewer 2 for this suggestion.

REFERENCES

Abbasi, A., Sarker, S., & Chiang, R. H. (2016). Big data research in information systems: Toward an inclusive research agenda. *Journal of the Association for Information Systems*, *17*(2), I–XXXII. https://doi.org/10.17705/1jais.00423

Aversa, P., Furnari, S., & Haefliger, S. (2015). Business model configurations and performance: A qualitative comparative analysis in Formula One racing, 2005–2013. *Industrial and Corporate Change*, *24*(3), 655–676.

Baden-Fuller, C., & Mangematin, V. (2013). Business models: A challenging agenda. *Strategic Organization*, *11*(4), 418–427.

Baumgartner, M. (2015). Parsimony and causality. *Quality & Quantity*, *49*(2), 839–856.

Brocke, J. V., Watson, R. T., Dwyer, C., Elliot, S., & Melville, N. (2012). Green information systems: Directives for the IS discipline. *Communications of the AIS*, *33*(30), 509–520.

Chiang, R. H. L., Grover, V., Liang, T. P., & Zhang, D. (2018). Special issue: Strategic value of big data and business analytics. *Journal of Management Information Systems*, *35*(2), 383–387. https://doi.org/10.1080/07421222.2018.1451950

Corbett, J. (2010). Unearthing the value of green IT. *ICIS*, *198*, 1–22.

Fiss, P. (2007). A set-theoretic approach to organizational configurations. *Academy of Management Review*, *32*(4), 1180–1198.

Flyvbjerg, B. (2006). Five misunderstandings about case-study research. *Qualitative Inquiry*, *12*(2), 219–245.

Garcia-Castro, R., & Francoeur, C. (2016). When more is not better: Complementarities, costs and contingencies in stakeholder management. *Strategic Management Journal*, *37*, 406–424.

Gholami, R., Watson, R. T., Molla, A., Hasan, H., & Bjørn-Andersen, N. (2016). Information systems solutions for environmental sustainability: How can we do more? *Journal of the Association for Information Systems*, *17*(8), 521–536.

Greckhamer, T., Misangyi, V. F., & Fiss, P. C. (2013). Chapter 3: The two QCAs: From a small-N to a large-N set theoretic approach. In *Configurational Theory and Methods in Organizational Research* (pp. 49–75). Bingley UK: Emerald Group Publishing Limited.

Korfiatis, N., Stamolampros, P., Kourouthanassis, P., & Sagiadinos, V. (2019). Measuring service quality from unstructured data: A topic modeling application on airline passengers' online reviews. *Expert Systems with Applications*, *116*, 472–486.

Larsen, K. R., Monarchi, D. E., Hovorka, D. S., & Bailey, C. N. (2008). Analyzing unstructured text data: Using latent categorization to identify intellectual communities in information systems. *Decision Support Systems*, *45*(4), 884–896.

Liu, Y., Mezei, J., Kostakos, V., & Li, H. (2017). Applying configurational analysis to IS behavioural research: a methodological alternative for modelling combinatorial complexities. *Information Systems Journal*, *27*, 59–79.

Longest, K. C., & Vaisey, S. (2008). fuzzy: A program for performing qualitative comparative analyses (QCA) in Stata. *Stata Journal*, *8*(1), 79–104.

Montabon, F., Sroufe, R., & Narasimhan, R. (2007). An examination of corporate reporting, environmental management practices and firm performance. *Journal of Operations Management*, *25*(5), 998–1014.

Müller, O., Junglas, I., Brocke, J. V., & Debortoli, S. (2016). Utilizing big data analytics for information systems research: Challenges, promises and guidelines. *European Journal of Information Systems*, *25*(4), 289–302.

Ragin, C. (1987). *The comparative method: Moving beyond qualitative and quantitative methods*. Berkeley: University of California.

Ragin, C. (2008). *Redesigning social inquiry: Fuzzy sets and beyond* (p. 240). Chicago: University of Chicago Press.

Ravishankar, M. N., Pan, S. L., & Leidner, D. E. (2011). Examining the strategic alignment and implementation success of a KMS: A subculture-based multilevel analysis. *Information Systems Research*, *22*(1), 39–59.

Rey-Martí, A., Porcar, A. T., & Mas-Tur, A. (2015). Linking female entrepreneurs' motivation to business survival. *Journal of Business Research*, *68*(4), 810–814.

Sandeep, M. S., & Ravishankar, M. N. (2016). Impact sourcing ventures and local communities: A frame alignment perspective. *Information Systems Journal*, *26*(2), 127–155.

Sarker, S., Xiao, X., & Beaulieu, T. (2013). Guest editorial: Qualitative studies in information systems: A critical review and some guiding principles. *MIS Quarterly*, *37*(4), iii–xviii.

Sarker, S., Xiao, X., Beaulieu, T., & Lee, A. S. (2018). Learning from first-generation qualitative approaches in the IS Discipline: An evolutionary view and some implications for authors and evaluators (PART 1/2). *Journal of the Association for Information Systems*, *19*(8), 752–774.

Teece, D. J. (2010). Business models, business strategy and innovation. *Long Range Planning*, *43*(2), 172–194.

Zimmermann, A., & Ravishankar, M. N. (2016). A systems perspective on offshoring strategy and motivational drivers amongst onshore and offshore employees. *Journal of World Business*, *51*(4), 548–567.

APPENDIX: Coding and identifying business model elements

We drew on archived reports listed in Table 1. Firms' sustainability reports on an average are about 50 pages long. However, for some firms, their length exceeded 100 pages. These reports often exhibited a common general

TABLE A1 Keywords associated with specific business model elements

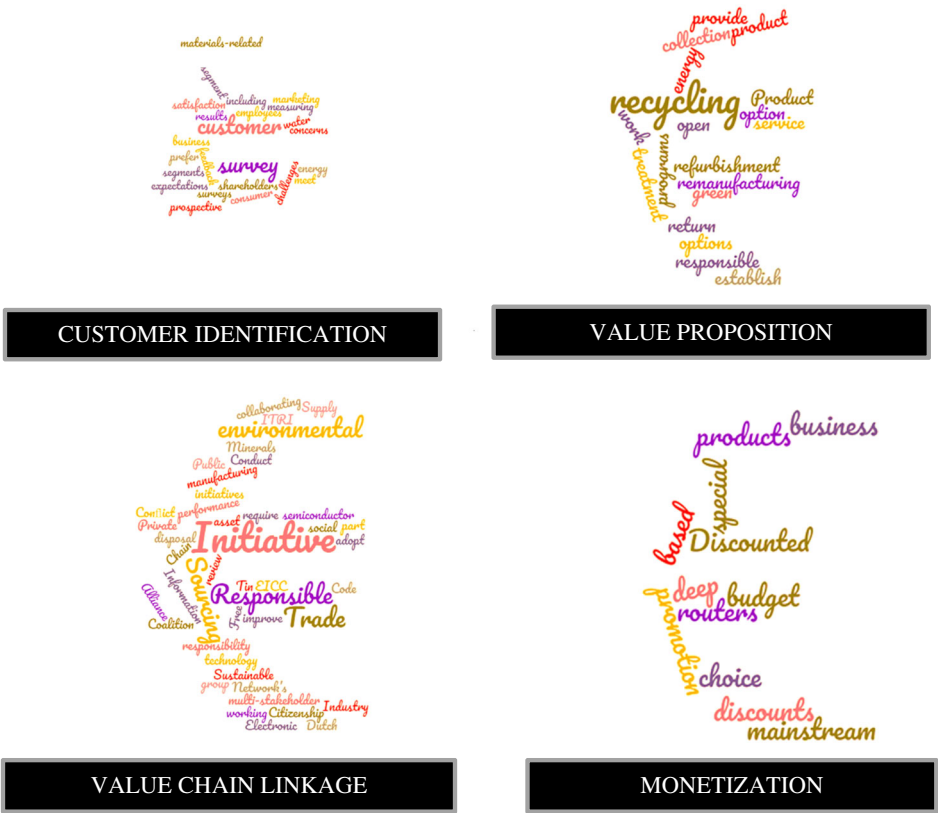| Keywords	Business Model Elements
Customer, survey, concerns, segments, consumers, perspective (consumer)	CI
Recycling, collection (product), refurbishment, product stewardship,	VP
Initiative, consortium, alliances, partnership, stakeholders, supply chain	VCL
Discounts, promotions, incentives	MON

CUSTOMER IDENTIFICATION

VALUE PROPOSITION

VALUE CHAIN LINKAGE

MONETIZATION

FIGURE A1 Word clouds showing words linked to business model elements

Distribution of keywords (Frequency)

RENEWABLE 12.8%
ENERGY 22.5%
FACILITIES 10.6%
DATA 8.1%
REDUCE 1.9%
SOLAR 2.1%
ENVIRONMENT 2.6%
EFFICIENCY 2.6%
ELECTRICITY 2.8%
CONSUMPTION 2.8%
CENTER 6.2%
WASTE 3.6%
EMISSIONS 4.9%
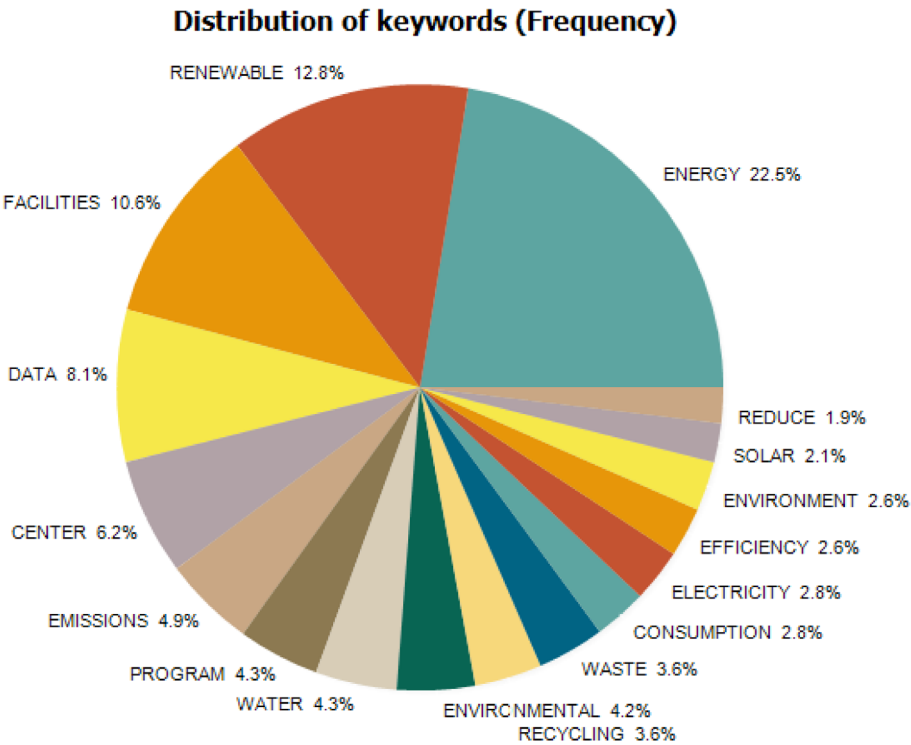PROGRAM 4.3%
ENVIRONMENTAL 4.2%
WATER 4.3%
RECYCLING 3.6%

FIGURE A2 Distribution of words in the archived report related to green IS actions

structure and followed Global Reporting Initiative (GRI) guidelines. The initial portion of such reports contained broad visions and performance snapshots. Firms' sustainability initiatives were often organized under broad headings such as pollution prevention and product stewardship. The structure of reports, therefore, helped us in understanding and interpreting, as the headings aligned closely with terms conceptualized in NRBV and Hart's sustainability framework. We (the two authors) individually mapped the unstructured green IS data to business model elements. We reviewed our respective coding schemes and resolved inconsistencies. We looked for keywords that indicated presence of specific business model elements (see Table A1 and Table 3). We examined these specific words and their use in sentences and paragraphs to ascertain the green IS actions taken by firms.

Motivated by the large size of our unstructured data, we also used conventional text mining on the different archived reports. We used the text-mining package in the open-source statistical analysis software R as well as WordStat for STATA to tabulate terms mentioned in different sections of the archived reports and matched this information with our coding. The rationale behind this approach was to ensure that our coding scheme was appropriate and robust, given the large size of unstructured data. Figure A1 provides a visual snapshot of the words (from the unstructured data set) extracted from the unstructured texts and linked with different business model elements. Figure A2 shows the distribution of keywords in a dataset linked to green IS actions.