Computational Design Science: A Critical Information Systems Research Area Contributing to Artificial Intelligence and Data Science

Xiao Fang, University of Delaware, xfang@udel.edu
Paul Jen-Hwa Hu, University of Utah, paul.hu@eccles.utah.edu
Michael Chau, University of Hong Kong, mchau@business.hku.hk
Hsinchun Chen, University of Arizona, hchen@eller.arizona.edu

Abstract

In the era of artificial intelligence (AI) and data science, the demand for advanced AI methods and systems capable of analyzing voluminous, diverse, and granular data to tackle grand challenges in businesses and society is paramount. Situated at the exciting intersection of information technology with business and social sciences, the information systems (IS) discipline is uniquely equipped to lead cutting-edge research that develops advanced AI solutions to address important business and societal challenges. But the potential of IS to contribute to this critical area has not been fully realized, largely due to its inadequate recognition of the relevance and importance of computational design science (CDS) research. As a specific genre of design science, CDS aims to develop novel computational algorithms and methods to solve business and societal problems with significant impact. In this article, we uncover the relationship of CDS with AI and data science, and stress that CDS is a vital IS area that can make substantial methodological contributions to not only the IS knowledge base but also AI and data science. We further identify issues that prevent the IS community from fully appreciating the relevance and significance of CDS research. By addressing these issues, we delineate what CDS research is and what its contribution is, clarifying that CDS research should be evaluated by its methodological contribution. Moreover, we explicitly elaborate the identity of CDS research and distinguish it from machine learning (ML) research in computer science and applied ML studies in other business fields. Finally, we propose a holistic template to provide IS scholars with clear and thorough guidance for preparing CDS papers.

Keywords: Computational design science, design science, information systems, artificial intelligence, data science, machine learning, methodological contribution

Introduction

Information systems (IS) research centers on the interaction between information technology (IT) with people, organizations, and society, with a focus on the development and use of IT artifacts to achieve their goals (Lee, 1999; Niederman & March, 2012). Distinct paradigms of IS research have prevailed, including behavioral science and design science (Chen et al., 2022). The former develops theories or tests theory-based models and hypotheses to explain phenomena pertinent to human-technology interactions in an organizational or social context. The latter designs and evaluates novel IT artifacts that enable people, organizations, and society to accomplish their objectives effectively and efficiently (Hevner et al., 2004).

Within the design science paradigm, different approaches have been taken to create IT artifacts, which Rai (2017) describes as different genres of design science research. These genres share important commonalities yet notably differ in the problem solved, the IT artifact designed, and the type of knowledge advanced (Rai, 2017). For example, the representation genre creates grammars and methods that faithfully represent complex phenomena (Clarke et al., 2016), whereas the optimization genre develops optimal solutions for operational or decisional problems (Menon & Sarkar, 2016). Hence, there is no single script for conducting and evaluating different genres of design science research (Rai, 2017). This article targets the computational genre of design science, or computational design science (CDS), *which aims at developing novel computational algorithms and methods to solve important business and societal problems with significant impact* (Chen, 2017). We emphasize that CDS is a unique and crucial area where IS research contributes to the IS knowledge base, while advancing the methodological core of artificial intelligence (AI) and data science simultaneously.

Despite the great value and impact of CDS, its relevance and significance have not been fully recognized in the IS field (Goes, 2014; Ram & Goes, 2021)[1], largely due to several confining issues. First, although several design science guidelines and templates exist, including Padmanabhan et al. (2022) and Abbasi et al. (2024), the IS community still lacks a shared, operationalized understanding of CDS as a distinct genre of design science, especially regarding what constitutes CDS research and how to assess its contribution. Second, few efforts have been devoted to elaborating the identity of CDS research. While many technical IS scholars have a good sense of CDS research, the community at large tends to confuse it with machine learning (ML) research in computer science or applied ML studies in other business disciplines (e.g., marketing), which precludes the IS community from fully appreciating its relevance and significance. Third, no holistic templates are available to guide IS researchers in reporting and disseminating important results and knowledge of CDS studies in the form of academic papers.

Due to these confining issues, authors, reviewers, and editors often have different views about what constitute a high-quality CDS paper, making reviews of CDS studies in IS journals challenging, vacillating, and uncertain. Consequently, many IS scholars find it difficult to contribute to exciting, fast-evolving AI and data science research, and the IS discipline collectively is missing the opportunity to advance the methodological core of AI and data science from the important IS perspective. Such difficulties and uncertainties might prompt CDS researchers to shift their publication choice from IS journals to outlets in other disciplines (e.g., computer science conferences and journals), thereby creating further drains of high-quality, impactful IS work that would otherwise contribute core design knowledge in AI and data science

[1] Goes (2014) and Ram and Goes (2021) report that design science articles account for 5 percent of the total publications in *MIS Quarterly* between 2006 and 2019, and the percentage of CDS publications should be even lower. Our analysis shows that the percentage of CDS articles in *MIS Quarterly* has increased to 9.3 percent between 2020 and 2023, which still marks a relatively lower presence in this prestigious journal.

from the IS perspective. The described pattern can create a vicious cycle that discourages IS scholars from conducting impactful CDS research and disseminating their work within the IS community, especially those in technical AI and data science area. Without conscious efforts, adequate anchors, and well-established guidelines that steer and foster CDS research, the IS discipline is likely to become marginalized in the methodological core of AI and data science over time, with its impact in this critical area substantially diminished.

This article aims to address these confining issues by:

(1) uncovering the relationship of CDS with AI and data science, and thereby providing a clear pathway for IS research to make substantial methodological contributions to AI and data science;

(2) delineating what CDS research is and what its contribution is;

(3) elaborating the unique identity of CDS research; and

(4) proposing a holistic template and substantive suggestions for crafting CDS papers.[2]

Artificial Intelligence, Data Science, and Computational Design Science

In essence, AI is a field that attempts to understand human intelligence and build intelligent entities (Russell & Norvig, 2020; National Science Foundation, 2022). Broadly, AI research can be categorized as foundational or use-inspired (National Science Foundation, 2022). The former seeks to develop theories and methods independent of any application domain, as exemplified by efforts that develop a theoretical framework for understanding neural networks, a building block of deep learning (Roberts et al., 2022). The latter targets grand challenges and pressing needs in

[2] Abbasi et al. (2024) identify major impediments for design science research for AI and propose ways to overcome them. However, their suggestions are not specific to CDS research. Instead of approaching design science in a general sense, this article focuses on its computational genre and delves deeper into the relationship of CDS, AI, and data science. Moreover, we also elaborate the relevance, importance, identity, and contributions of CDS research and provide a template to guide the presentation of CDS research.

different subject domains, such as business, healthcare, and various scientific or engineering areas. Use-inspired AI develops novel algorithms and methods to tackle these challenges and needs, thereby solving application problems in target domains while contributing to AI research simultaneously. Because of its development of novel AI algorithms and methods, rather than merely applying existing ones, use-inspired AI research is distinctive from applied AI (National Science Foundation, 2022).

Use-inspired AI research straddles AI and domain sciences, featuring several fundamental characteristics.

(a) Problem solving. Use-inspired AI research is application-driven and seeks to solve important problems in subject matter domains.

(b) Machine learning. Algorithms and methods created by use-inspired AI research are ML algorithms and methods. As Goodfellow et al. (2016, p. 8) note, "machine learning is the only viable approach to building AI systems that can operate in complicated real-world environments."

(c) Computational. Algorithms and methods developed by use-inspired AI research are implemented and executed in a computational environment.

(d) Interface science. Use-inspired AI research is situated at the junction of AI and subject matter domains; it develops AI algorithms and methods to solve problems in these domains. Simon (1996) refers to such a scientific field as an interface science. Because of its interfacing orientation (Benbasat & Zmud, 2003) and interdisciplinary, boundary-expanding nature (Argawal & Lucas, 2005), the IS discipline is well-suited for use-inspired AI research by solving challenging problems in business and society.

Parallel to the development of AI, data science was first envisioned by John Tukey, one of the greatest statisticians of the 20th century, though he referred to it as the science of data analysis

(Tukey, 1962). This new science then was conceived as the science of learning from data, which underscores the interplay of statistics, computing, and the challenge of analyzing voluminous data in various application domains (Tukey, 1962; Donoho, 2015). Cleveland (2001) follows Tukey's conception of data analysis and identifies several technical areas in data science, including multidisciplinary investigations that emphasize data analysis for problem solving in subject matter areas, data analysis models and methods in conjunction with computational environments, and statistical theory for data analysis. While statisticians approach data science from the angle of statistics, computer scientists conceive it from the perspective of computing. For example, Jeffery Ullman, the 2020 Turing Award winner, views data science as the intersection between computer science and domain sciences (Ullman, 2020). As a subfield of computer science, ML plays a vital role in data science, because it empowers learning from data, while other computer science subfields (e.g., database) also contribute to data science. Domain sciences are crucial to data science too, because they help conceptualize problems in subject matter domains and provide contextual insights for analyzing and solving them. While statisticians and computer scientists have different views about the definition and scope of data science, they agree on a set of common features of data science.

(a) Problem solving. The objective of data science is to leverage data to solve important problems in subject domains.

(b) Machine learning. ML, which statisticians often refer to as predictive modeling (Donoho, 2015), provides essential models and methods for data analysis.

(c) Computational. ML models and methods render problem-solving utilities only if they are implemented as software or information systems deployed in a computational environment (Cleveland, 2001; Ullman, 2020).

(d) Interface science. Data science lies at the junction of domain sciences and IT, particularly

ML. As an interface science, it is intrinsically interdisciplinary. That is, no single discipline "owns" data science exclusively (Donoho, 2015; Ullman, 2020; Özsu, 2023); rather, different concerning disciplines can contribute to data science, including IS (Saar-Tsechansky, 2015).

Our analysis reveals that both use-inspired AI and data science center on the intersection between AI and domain sciences, with an intended focus of developing novel ML methods that can be implemented and deployed in a computational environment to solve problems in subject matter domains. Like use-inspired AI and data science, IS also is an interface discipline that straddles IT and domain sciences — business and social sciences (Lee, 1999; Niederman & March 2012). Therefore, IS research is uniquely endowed with great potential to advance use-inspired AI and data science. To achieve this noble goal, CDS is crucial and its significance must be comprehensively understood and fully recognized by the IS community.

Notably, CDS possesses all the featuring characteristics of use-inspired AI and data science. It emphasizes *problem solving* and tackles important business and societal problems. It is *computational* and develops novel IT artifacts in the form of computational algorithms and methods. These algorithms and methods leverage *machine learning* for effective problem-solving. Indeed, state-of-the-art CDS studies develop novel ML algorithms and methods to solve a wide array of important business and societal problems, ranging from cybersecurity and healthcare analytics to FinTech and social network analytics (Abbasi et al., 2010; Meyer et al., 2014; Bardhan et al., 2015; Fang & Hu 2018; Samtani et al., 2022; Zhao et al., 2023; Wang et al., 2023). Finally, CDS is *interdisciplinary* (Chen, 2017) and typically involves ML and business and social sciences. Because of these shared characteristics, CDS is central to use-inspired AI and data science in the domain of business and society. Advanced algorithms and methods produced by CDS research not only expand the knowledge base of IS but also add to the

knowledge base of use-inspired AI/data science. Hence, CDS represents a unique and critical area where IS research can make significant methodological contributions to AI and data science.

However, the IS community has not yet adequately recognized the relevance and importance of CDS research (Goes, 2014; Ram & Goes, 2021), due to the confining issues described previously. As a result, the potential of IS to contribute to the methodological core of AI and data science is impeded, resulting in missed opportunities that could have made the IS discipline more relevant and impactful. Many IS scholars are competent and innovative in designing novel algorithms and methods to tackle important business and societal problems, because they possess sound technical knowledge, rich research experience, and deep insights into domain sciences. We firmly believe that the IS discipline has great potential to produce more high-quality, impactful CDS research that expand the IS knowledge base while advancing AI and data science, than as it does now. To actuate such important contributions and to orient and facilitate researchers in conducting and publishing CDS research in IS outlets, we address the confining issues by providing guidelines for conducting and evaluating CDS research, elaborating its identity, and proposing a template for preparing CDS papers.

Contribution and Evaluation of Computational Design Science Research

Computational Design Science Research and Its Contribution

As we present in Figure 1, CDS research begins with an important *business or societal problem* that concerns people, organization, or society. To solve the problem, CDS research builds on the knowledge base of IS as well as that of use-inspired AI/data science. Both knowledge bases contain CDS knowledge, which consists of existing ML algorithms and methods[3] for business

[3] Cormen et al. (2022) define an algorithm as a sequence of computational steps that transforms the input into the output, while a computational method is a set of algorithms that jointly solve a problem.

and societal problems and principles for guiding the design and evaluation of CDS research. Because CDS is a subfield of IS, CDS knowledge is part of the broader IS knowledge base. In addition to CDS knowledge, the IS knowledge base contains other knowledge, such as IS theories, instruments, knowledge created by other genres of design science research, and general design science principles (Hevner et al., 2004). Meanwhile, as CDS is a core part of use-inspired AI/data science in the business domain, CDS knowledge is part of the use-inspired AI/data science knowledge base too. The use-inspired AI/data science knowledge base also includes knowledge beyond CDS knowledge, such as ML algorithms and methods for non-business domains (e.g., astronomy) and general-purpose ML algorithms and methods, which often are created by research outside IS.

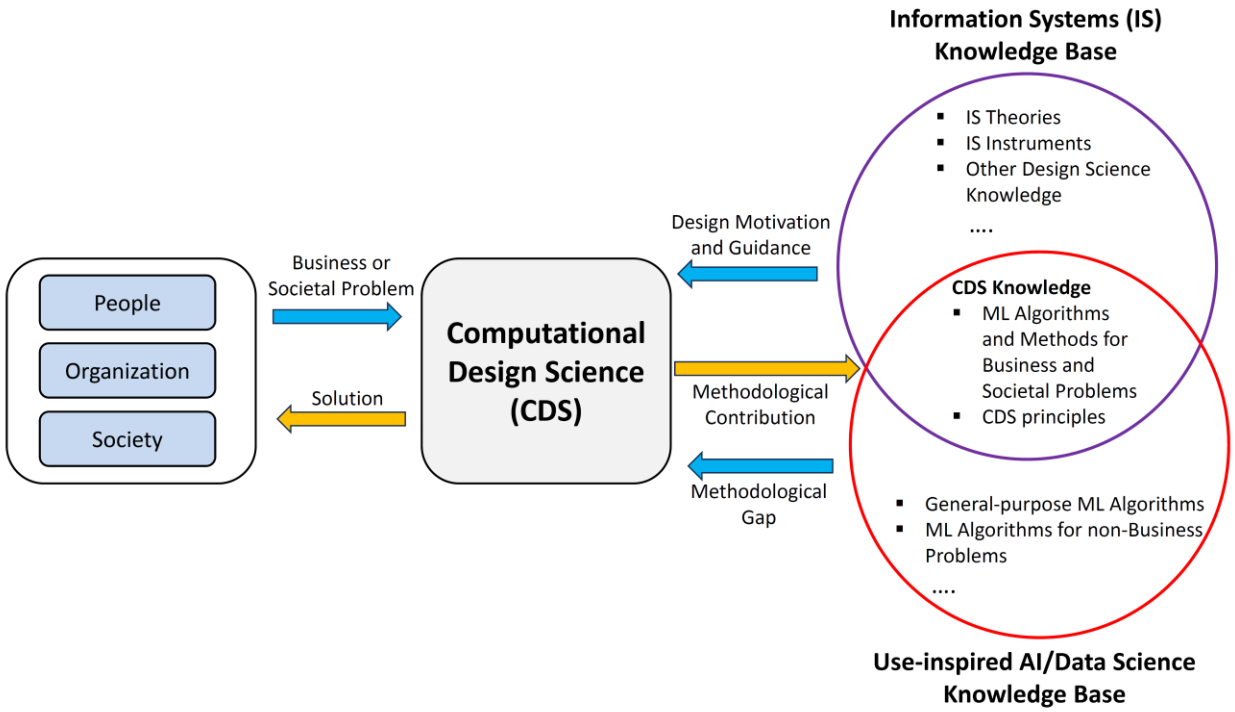

Figure 1: Computational Design Science

As Figure 1 illustrates, mindful observations of people, organization, or society lead to the

identification of important business or societal problems that need efficacious solutions by CDS research. The IS knowledge base *motivates and guides* CDS research in such efforts. More specifically, it provides domain and problem understanding, contextual insights, and design science principles that guide problem analysis and solution development. Meanwhile, the IS and use-inspired AI/data science knowledge bases jointly reveal *methodological gaps* of existing ML algorithms and methods for solving a focal problem. The identified gaps specify the deficiencies of existing ML algorithms and methods, which make them unsuited or ineffective for solving the problem. It is noted that existing ML algorithms and methods include those designed for the problem (i.e., those in CDS knowledge) and general-purpose ML algorithms that can be adapted to solve it (i.e., those in the use-inspired AI/data science knowledge base other than CDS knowledge). To address the identified methodological gaps, CDS research develops a novel ML algorithm or method. As a result, it solves the focal problem with a more effective *solution* and makes *methodological contribution* to both IS and use-inspired AI/data science knowledge bases. Specifically, methodological contribution refers to the proposed ML algorithm or method that advances CDS knowledge situated at the intersection of the two knowledge bases. Table 1 presents several representative CDS studies to illustrates the important aspects of CDS research.

Table 1: Important Aspects of CDS Research, Illustrated by Representative CDS Studies

	Zhao et al. (2025)	Wei et al. (2025)	Zhang et al. (2024)	Samtani et al. (2022)
Problem	▪ Fintech problem of developing an industry classification system (ICS) that can identify economically related firms.	▪ Societal problem of mitigating biases in hate speech identification on social media.	▪ Resource management problem of proactively determining types and quantities of resources required for disaster relief.	▪ Cybersecurity problem of linking hackers' exploits to device vulnerabilities.
Design motivation and guidance	▪ Solution to focal problem is fundamental to firms' operational and strategic decision-making, as well as to business research. ▪ Key to ICS development is to represent the semantics of firms' 10-K documents. ▪ Study is guided by general and computational design science principles.	▪ Solution to the problem is critical for more accurately detecting hate speech on social media platforms with a small number of expert-labelled examples. ▪ Proposed method is guided by general and computational design science principles; it is also enlightened by research on differential language usage.	▪ Solution to focal problem is critical for reducing fatalities and damages by disasters, but this important problem has been neglected by prior research. ▪ Proposed solution follows general and computational design science principles.	▪ Focal problem is critical for protecting cybersecurity in organizations, but has been overlooked by IS research. ▪ Method design and evaluation are guided by general and computational design science principles, such as problem-solving orientation, emphasis on methodological contributions, and justification of benchmark methods.
Methodological gap	▪ Both existing CDS knowledge and general-purpose ML methods (e.g., document embedding methods) cannot account for two salient characteristics of 10-K documents: concept switching and document evolution.	▪ Conventional text classification methods suffer from biases in data, and existing debiasing methods require a large number of expert-labelled examples.	▪ Existing methods estimate resource requirements based on currently unfulfilled demands but ignore future demands arising between resource requesting time and arrival time.	▪ General-purpose deep learning algorithms for short text matching cannot solve focal problem effectively, due to their inability to capture overlapping text features between exploits and vulnerabilities.
Methodological contribution	▪ Proposed method makes two key methodological innovations: an adaptive gating mechanism for modeling concept switching and a dynamic embedding mechanism for capturing document evolution.	▪ Proposed method leverages large language models (LLMs), contrastive learning, and prompt-based learning. ▪ Design a novel pair generator for contrastive learning and customized knowledge injection in continuous prompts.	▪ Propose a deep learning-based temporal point process method for future demand prediction. ▪ Design a stochastic optimization model that produces an optimal resource requesting plan according to currently unfulfilled demands and predicted future demands.	▪ Develop the EVA-DSSM algorithm by incorporating an attention technique into a short text matching algorithm. ▪ Design a new metric to assess the severity of device vulnerability.
Solution	▪ Proposed method enhances firms' decision-making and benefits business research by allowing for more granular control in selecting comparable firms.	▪ Proposed method reduces inherent biases in LLMs and detects hate speech more effectively using a small number of labelled examples.	▪ Proposed solution enables disaster relief agencies to make effective resource requests by taking future demands into account.	▪ Developed algorithm and metric enable organizations to respond to hackers' exploits in a proactive and effective way.

Evaluation of Computational Design Science Research

Due to its intrinsic problem-solving nature (Hevner et al., 2004; Chen, 2017), CDS research should be evaluated by its methodological contribution. Novel IT artifacts produced by CDS research for solving focal problems are of foremost importance. These artifacts, in the form of computational algorithms and methods, constitute methodological contributions by CDS research (Chen, 2017; Padmanabhan et al., 2022). In contrast, behavioral science research develops or tests theories that explain phenomena emerged from the utilization of IT by people, organization, or society (Lee, 1999). Hence, it is conceivable to assess behavioral science research based on its theoretical contribution. Because CDS research differs from behavioral science research in objective (problem solving versus phenomena explanation) and merit (creating computational algorithms and methods versus developing or testing theories), we should not apply the standard for evaluating one type of research to the other. That is, behavioral science research is not expected to make methodological contribution[4], and CDS research should not be required to make theoretical contribution (Chen, 2017; Ram & Goes, 2021).

The significance of a CDS study's methodological contribution can be gauged using the design science knowledge contribution framework proposed by Gregor and Hevner (2013). According to this framework, a CDS study can make significant methodological contribution, if it develops a novel ML algorithm or method to solve a new problem (invention), or proposes a novel ML algorithm or method to address a known problem (improvement), or extends an existing ML algorithm or method to a new problem (exaptation). For both invention and improvement, methodological novelty of the ML algorithm or method developed by a CDS study

[4] Methodological contribution in behavioral and social sciences research entails the creation or significant refinement of generic research methods, such as survey and case study, which can be applied across different research contexts.

is essential for substantiating its methodological contribution. Because a CDS study aims to solve a business or societal problem, its methodological novelty should be evaluated relative to the problem context and the pertinent existing knowledge. Such knowledge includes existing ML algorithms and methods designed to address the focal problem (i.e., CDS knowledge) as well as general-purpose ML algorithms and methods that can be adapted to solve it. A detailed problem analysis can reveal key challenges of solving the problem and specify the methodological deficiencies of the existing knowledge in addressing these challenges. By overcoming these deficiencies, the proposed ML algorithm or method exhibits significant methodological novelty relative to the existing knowledge. For example, Zhao et al. (2025) tackle the problem of developing an Industry Classification System (ICS) that can identify economically related firms. The authors show that the key to ICS development is to represent the semantics of firms' 10-K documents, as these documents contain essential information central for measuring their economic relatedness. The existing knowledge relevant to this study includes document embedding methods specifically designed to represent the semantics of 10-K documents for ICS development (CDS knowledge) as well as general-purpose document embedding methods, including state-of-the-art methods built upon large language models, that are not specifically developed for 10-K documents. Their analysis of the problem uncovers that a key challenge for effective 10-K document representation is to accommodate two salient features of these documents: concept switching and document evolution. The former refers to the phenomenon that, while a 10-K document covers many concepts, the emphasis shifts from one subset of concepts to another across different parts of the document. The latter refers to the evolution of a firm's 10-K documents over time. None of the existing document embedding methods can model concept switching or document evolution. To address the deficiencies of the existing methods,

the authors propose a method with two key methodological innovations: an adaptive gating mechanism for modeling concept switching and a dynamic embedding mechanism for capturing document evolution.

For exaptation, an existing ML algorithm or method, or a combination of existing algorithms or methods, is extended to solve a new problem. To make nontrivial methodological contribution, such efforts must clearly articulate the specific challenges for solving the problem and compellingly demonstrate the extension is novel and substantial (Gregor & Hevner, 2013; Brown, 2024). For example, Samtani et al. (2022) extend a deep learning-based short text matching algorithm to tackle the problem of linking hackers' exploits to device vulnerabilities. The specific challenge to the attempted extension is that existing short text matching algorithms cannot model overlapping semantics between exploits and vulnerabilities. To address this challenge, the authors incorporate an attention technique into the short text matching algorithm. The underlying rationale is that the attention mechanism can link letter trigrams in exploits with those in vulnerabilities, thereby capturing overlapping semantics between them. As a result, the proposed method advances existing knowledge by providing a more effective solution for a critical cybersecurity problem.

In addition to methodological contribution, CDS research must be impactful by providing research insights and pragmatic value. Research insights enlighten how a CDS study's problem modeling and computational method design can inform future efforts that target similar or related problems. For example, a study can benefit future CDS research by offering insights into problem formulation and modeling, architecture design, or method conceptualization and development. Equally important are practical impacts at individual, organizational, or societal level, with actionable strategy and policy implications (Abbasi et al., 2024). That is, the

pragmatic value of a CDS study can be demonstrated by meaningful benefits of its proposed method to key stakeholders, as measured by qualitative and quantitative metrics.

Role of Theory in Computational Design Science Research

CDS research does not have to be theory-based (Ram & Goes, 2021). The design of an algorithm or method needs to be grounded in sound rationales, justifications, and mechanisms. Rationales establish the legitimacy and appropriateness of the overall design, justifications defend key design choices, and mechanisms explain why the proposed algorithm or method is effective for a focal problem. Together, these elements form the foundation of the design and ensure the rigor and novelty of the designed algorithm or method. Rationales, justifications, and mechanisms can be derived through a creative and logical analysis of the focal problem and a thorough review of pertinent knowledge. Continuing with the example of Zhao et al. (2025), the authors analyze the focal problem, review pertinent knowledge, and then specify that existing document embedding methods cannot model concept switching or document evolution, which motivates the design of a novel method for 10-K documents. The underlying design rationales include representing a document as a weighted combination of concepts, adapting these weights across different parts of a document to model concept switching, and dynamically updating them over time to cope with document evolution. Guided by these rationales, the authors model each concept as a learnable vector and represent a document as a weighted sum of concept vectors. Next, they design a gating module to adjust concept weights based on the semantics of local context and propose a dynamic embedding module that models concept weights as a function of time. These design choices align with the elaborated rationales and are justified by findings that local contextual semantics provide critical signals for inferring local concept weights and that the semantics of a firm's 10-K documents evolve over time as its business activities change.

Mechanism-wise, the dynamic embedding module learns time-varying concept weights, which in turn are adjusted by the gating module at each time step, thereby capturing both concept switching and document evolution and ultimately leading to the superior performance of the designed method. This mechanism is further demonstrated by showing the contribution of each module to the method's performance empirically.

Design rationales, justifications, and mechanisms can also be informed by theory. Theory here should not be limited to business and social science theory; rather, they also include computational and mathematical theory. For example, the ML methods developed by Yin et al. (2023) and Fang and Hu (2018) are grounded in social science theory, while those proposed by Abbasi et al. (2010), Li et al. (2017), and Mao et al. (2026) are premised in computational and mathematical theory. An illustrating case in point is the top persuader prediction method by Fang and Hu (2018), which uses social network theories to conceptualize distinct social forces for persuasion and then integrates these forces to predict top persuaders in a social network. The misinformation detection method by Mao et al. (2026), in contrast, is built on a machine learning theory the authors develop. The objective is to learn a model from labelled source domain data (e.g., true or false news in the entertainment domain) to predict unlabeled target domain data as true or false (e.g., unlabeled news in the COVID domain). The authors show theoretically that the domain adaptation error incurred when applying a model trained on source domain data to a target domain is bounded by the sum of the source domain error and the covariate and concept shifts between the two domains. The theoretical analysis guides the design of a method that reduces domain adaptation error by mitigating both source domain error and covariate and concept shifts.

Having discussed the role of theory in algorithm or method design in CDS research, we now

turn to its role in algorithm or method evaluation. Typically, a designed algorithm or method is evaluated by comparing its performance with that of well-justified benchmark methods. Performance is usually measured by commonly accepted ML metrics such as precision, recall, and F1-score (Samtani et al., 2022; Mao et al. 2026) or problem-specific measures such as deprivation cost in disaster relief (Zhang et al., 2024). Theory is not required for such performance evaluations. Some CDS studies also conduct user or case studies to demonstrate the proposed method's utilities and values in business or sciential contexts. The use of theory is beneficial for these efforts, even though it is not necessarily required. For example, drawing on the technology acceptance model (Davis 1989), Samtani et al. (2022) design a user study to examine cybersecurity experts' perceived usefulness of their proposed algorithm. Similarly, Zhao et al. (2025) perform a case study to demonstrate the business value of the proposed ICS through the lens of industry momentum theory (Moskowitz and Grinblatt 1999). In contrast, the case study by Zhang et al. (2025) does not involve theory.

Identity of Computational Design Science Research

As a subfield of the IS discipline, CDS situates at the intersection of ML and business and social sciences. This intersection defines the identity of CDS research. Studies that focus on ML and design general-purpose ML algorithms without business or social contexts are not CDS research. Likewise, studies that develop or test business or social science theories, while viewing ML artifacts as a black box or variable, are not CDS research either. This identity makes CDS research unique and distinct from research involving ML in other disciplines, especially ML research in computer science (CS) and applied ML studies in other business fields such as marketing and finance.

CDS research designs novel ML algorithms and methods in business or social contexts.

Therefore, the confluence of ML with business and social sciences is central to CDS research, while ML is the dominant focus of ML research in CS (Padmanabhan et al., 2022). Because of this distinction, CDS research and ML research in CS proceed in different ways. Aiming at developing computational solutions for business and societal problems, CDS research requires a deep understanding of a focal problem, which can be achieved through a thorough review of problem-specific literature (e.g., Samtani et al., 2022; Zhao et al., 2023). Such reviews, however, are not necessary for ML research in CS, because it typically seeks to improve the performance of ML algorithms (e.g., Ganin et al., 2016). In a related sense, CDS research designs problem-specific ML algorithms and methods by incorporating salient characteristics of a focal problem (Padmanabhan et al., 2022). In contrast, algorithms and methods created by ML research in CS seldom consider problem-specific characteristics. Consequently, they are either unsuitable or ineffective for solving business and societal problems targeted by CDS research. Moreover, to demonstrate the utilities for solving a focal business or societal problem, CDS research evaluates proposed ML algorithms or methods using problem-specific measures, in addition to conventional ML metrics such as F1-score and AUC (Padmanabhan et al., 2022). For example, Zhao et al. (2023) perform a case study of tax filing by firms to demonstrate the economic value of their designed firm-industry assignment method, measured by reductions in tax-related cost. We acknowledge that some recent ML studies in CS are inspired by business problems. Nevertheless, CDS research still differs from most ML work in CS because it is at the intersection between ML and business and social sciences. Such identity of CDS research is "inherited and specialized from the social-technical identity of the IS discipline" (Padmanabhan et al., 2022, p. viii). Table 2 summarizes important distinctions between CDS research and ML research in CS.

Table 2: Important Distinctions between CDS Research and ML Research in CS

	CDS Research	ML Research in CS
Scope	Confluence of ML with business and social sciences	Dominant Focus on ML
Objective	Problem solving	Algorithm improvement
Literature survey	Literature on focal business or societal problem and relevant ML methods	Primarily ML literature
Artifact design	Designing ML algorithms and methods by incorporating problem-specific characteristics	Designing general-purpose ML algorithms and methods
Artifact evaluation	Using both problem-specific measures and conventional ML metrics	Using conventional ML metrics

Moreover, CDS research differs from ML studies in other business fields such as finance and marketing. One type of these studies investigates how ML impacts individuals, firms, markets, or society, such as credit markets (Fuster et al., 2022) and human decision-making (Boyacı et al., 2024). These studies view ML algorithms as a black box and treat them as a variable in their theoretical models, without considering methodological details. In contrast, CDS research solves business and societal problems by developing novel ML algorithms and methods. Instead of treating ML algorithms as a black-box or a variable, a CDS study examines details of existing ML algorithms relevant to a focal problem, identifies their limitations for solving the problem, and designs a novel algorithm or method that overcomes these limitations. Another type of ML studies in other business fields apply existing ML algorithms to solve business problems. For example, Anand and Lee (2023) use a conditional generative adversarial network (GAN) to preserve customer privacy when transferring customer data from firms to researchers, whereas Gu et al. (2020) compare the performance of different ML algorithms (e.g., random forest and neural network) for asset pricing. CDS research differs from these studies in that it develops

novel ML algorithms and methods rather than merely applying existing ones. Methodological novelty of the designed ML algorithm or method distinguishes CDS research from this type of ML studies in other business fields.

We notice recent efforts to design innovative ML algorithms for solving important problems in marketing and operations. Yet, CDS research remains distinctive from most work in these business fields because it straddles between ML and business and social sciences, with an eminent focus on developing novel ML algorithms and methods rather than applying existing ones. Table 3 summarizes the main differences between CDS research and ML studies in other business fields. As shown, studies examining the impact of ML on business typically treat ML algorithms as a variable, disregarding their differences and methodological specifics, resulting in a high level of abstraction of ML. Applied ML studies in business focus on how ML algorithms process inputs and generate outputs. In contrast, CDS research emphasizes methodological details of ML algorithms, identifies their methodological deficiencies, and develops new algorithms or methods to address these deficiencies. Consequently, the level of ML abstraction is medium for Applied ML studies but low for CDS research.

Table 3: Main Differences between CDS Research and ML Studies in Other Business Fields

	CDS Research	**ML Impact on Business**	**Applied ML in Business**
Level of ML abstraction	Low	High	Medium
Designing novel ML algorithm or method	Yes	No	No
Primary contribution	Methodological	Other contributions (e.g., theoretical)	Other contributions (e.g., theoretical)

At a broader level, CDS research helps strengthen the identity of the IS discipline. Benbasat

and Zmud (2003, p.186) advocate that "IS scholars and IS practitioners strive to increase their collective understanding of (1) how IT artifacts are conceived, constructed, and implemented, (2) how IT artifacts are used, supported, and evolved, and (3) how IT artifacts impact (and are impacted by) the context in which they are embedded." They emphasize the criticality of IT artifacts and caution against treating IT artifacts as a black box or as being synonymous with a more generic entity. Not only does CDS research align with this kernel view of IT artifacts in an organizational or societal context, but it also echoes the transformational impact of IT, especially in the fast-emerging era of AI, and responds to calls for research with a stronger macro focus, to provide increased visibility and impact (Agarwal & Lucas, 2005). Like IS, CDS research is interfacing; it can strengthen the IS identity in academia and enhance its relevance and impact in a broader way by designing compelling IT artifacts and demonstrating practical utilities for solving important problems for key stakeholders.

A Template for Computational Design Science Publication

Due to notable differences among different genres of design science research, there is no universal script suitable for all genres (Rai, 2017). Thus, templates for general design science research do not offer specific guidance for authors to prepare CDS papers. As described previously, CDS research, built on both the IS and the use-inspired AI/data science knowledge bases, strives to make methodological contributions to both knowledge bases. This unique feature demands special considerations for writing CDS papers, including what literature to review in the Related Work section, how to clearly elaborate methodological novelty of a designed algorithm or method in the Method section, and how to demonstrate the validity and value of a designed method in the Evaluation section. Next, we suggest a template that helps authors develop each section of a CDS paper, with illustrating example papers in Table 4.

Introduction. The objective of this section is to motivate a CDS study being conducted. First, this section describes the business or societal problem to be solved and explains its significant impacts on real-world practices. Ram and Goes (2021) advocate the IS discipline to focus on high-impact studies that "solve grand challenges and generate solutions that impact the world beyond academics" (Ram & Goes, 2021, p. 479). In response to this call, a CDS study must justify the practical significance of the business or societal problem it addresses (e.g., Chau et al., 2020), in addition to its research contributions. Second, this section identifies research gaps and highlights the methodological innovations and contributions of a reported CDS study by situating it in the relevant IS literature and other related research (Gregor & Hevner 2013; Brown 2024). Research gaps may be revealed by pinpointing important problems overlooked in the existing literature (e.g., Samtani et al., 2022; Zhang et al., 2024). Addressing such problems necessitates the development of innovative methods, which in turn leads to methodological innovations and contributions. Alternatively, a CDS study may focus on an existing problem and uncover deficiencies and challenges of existing methods for solving it — that is, research gaps (Kim et al., 2023; Zhao et al., 2023). By addressing these gaps, the CDS study designs an innovative method and makes methodological contributions to the literature.

Related work. This section reviews literature closely related to a CDS study, identifies research gaps to be addressed, and discusses methodological novelty of the study. One stream of related research is specific to a focal business or societal problem, which includes literature on the problem and literature on ML methods for solving the problem (i.e., CDS knowledge in Figure 1). To illustrate, Samtani et al. (2022) review cybersecurity literature on hacker community and vulnerability assessment, which closely relates to their targeted problem of linking hackers' exploits to devices' vulnerabilities. Another stream of related research is extant

literature on general-purpose ML algorithms and methods that can be adapted to solve a focal problem. In the same example, Samtani et al. (2022) survey deep structured semantic models and the attention mechanism, which can be adapted to address their investigated problem. While the Introduction section outlines research gaps of existing literature and methodological novelties of a CDS study, this section provides detailed elaborations of these two aspects. It delineates essential considerations and properties of an effective method for a focal problem and conveys that existing methods fall short of these considerations and properties. By proposing a method that overcomes the identified limitations, a CDS study develops a novel method significantly different from existing ones methodologically and solves the focal problem effectively.

Problem formulation. The business or societal problem tackled by a CDS study is described in the Introduction section and needs to be formally defined in the Problem Formulation section. A typical problem formulation begins by defining the important constructs of a focal problem. For example, Zhao et al. (2023) introduce important constructs of the industry assignment problem, including industry and firm-industry assignment. Use of illustrating examples is desirable and can help readers understand key problem constructs. Problem formulation defines the scope of a problem and the goal of the attempted problem-solving (Gregor & Hevner, 2013). Thus, it is necessary to specify the objective and constraints of a problem as well as what is given for solving the problem, based on problem constructs. The problem is then formulated as optimizing the objective by taking what is given as input, subject to the specified constraints.

Method. This section presents the method designed to solve the problem defined in the Problem Formulation section. The main contribution of a CDS paper is methodological, so the Method section is critical for authors, editors, reviewers, and readers. The designed method should be novel and rigorous. The goal of method design is to optimize the objective of a focal problem. This is achieved by designing novel modules, each instantiating one or more methodological novelties specified in the Related Work section. Therefore, to clearly convey the novelty of the designed method, it is important to link its modules with their respectively realized methodological novelties. Rigor is critical too (Sun et al., 2025). It is manifested in sound rationales that guide authors to conceptualize and design a method, appropriate justifications for key design choices, clear mechanisms that explain the effectiveness of the designed method. Rigor can be accomplished through a thorough analysis of problem characteristics and an effective use of the knowledge bases shown in Figure 1. Although not mandated, formally establishing important properties of a designed method, such as optimality (Li et al., 2017) and convergence (Fang & Hu, 2018), helps strengthen rigor. Authors should prepare the Method section in a logical, cohesive, and easy-to-understand manner. It is suggested to detail essential method components, such as key equations and algorithms, with examples. Due to the fast progress of ML literature, reviewers of a CDS paper need to keep abreast of the up-to-date literature relevant to the paper they review.

Evaluation. The objective of this section is to demonstrate the validity of a designed method (Chen, 2017). In design science, validity entails that a designed IT artifact "does what it is meant to do" (Gregor & Hevner, 2013, p. 351). Accordingly, for CDS research, validity entails that a designed method can effectively achieve the objective of solving the problem defined in the Problem Formulation section. Therefore, to establish the validity of a designed method, it is

essential to demonstrate its superior performance over well-justified benchmark methods, using appropriate data and evaluation metrics (Padmanabhan et al., 2022).[5] Evaluation metrics reveal a method's efficacy of accomplishing the problem-solving objective, which include both problem-specific measures and common ML metrics. In addition, it is necessary to justify the choice and appropriateness of benchmark methods (Brown, 2024), which include both state-of-the-art ML methods developed for a focal problem and advanced general-purpose ML methods that can be adapted to solve the problem. With demonstrated performance improvements, a designed method "could produce significant business value by reducing cost, increasing profit, or enabling new business models" (Padmanabhan et al., 2022, p. vii). Hence, it is beneficial to showcase the practical impact of the method at the individual, organizational, or societal level through a case study that simulates a real-world scenario. For example, Samtani et al. (2022) conduct a case study to evaluate the usefulness of the proposed algorithm perceived by cybersecurity experts; Zhao et al. (2023) perform a case study to demonstrate the reduction in firm-level tax-related costs achieved by the designed firm-industry assignment method; and Zhang et al. (2024) use a case study to show the effectiveness and fairness of the developed resource management method in fulfilling demand during disaster response. In addition to demonstrating the superior performance of a designed method, explaining why it outperforms benchmarks is important too. Towards that end, an ablation study can demonstrate the contribution of each novel module of the method to its observed performance (e.g., Hu et al., 2024). However, establishing the causal relationship between the novelties of the method and its performance advantages is not required.

[5] For a computational design science study, it is sufficient to demonstrate the performance advantages of its designed method over benchmarks using archival data or simulations. Conducting field experiments to show the performance advantages of the designed method is not necessary.

Establishing the generalizability of a designed method is also essential. A CDS study develops a method by abstracting and generalizing from real-world problem instances. A well-designed method captures generalized problem characteristics rather than instance-specific details. As a result, the method possesses inherent generalizability by design. To demonstrate its generalizability, a method can be evaluated with multiple datasets, in different and representative real-world scenarios, and under varying conditions. For example, Mao et al. (2026) evaluate the performance of their proposed method using multiple datasets (e.g., both English and Chinese infodemic datasets), various application scenarios (e.g., information veracity prediction and diffusion virality prediction), and varying conditions (e.g., changing closeness between source and target domains). Additionally, the generalizability of a method can be demonstrated analytically by proving its general properties, such as convergence (Zhang et al. 2025) and performance guarantees (Zhang et al. 2024).

Conclusion. The Conclusion section summarizes a CDS study and highlights its contribution to both the IS and the use-inspired AI/data science knowledge bases. Specifically, the novel ML method designed by a CDS study adds to CDS knowledge located at the intersection of the two knowledge bases (see Figure 1) and represents the study's methodological contributions to both knowledge bases. When deployed in real-world settings, the designed method empowers key stakeholders to solve the focal problem effectively, thereby enhancing business practices. Hence, it is also important to discuss the impact and implications of the designed method on business and society. The section concludes by identifying promising future research directions that extend the designed method.

Table 4: Template for CDS Paper: Illustrated with Exemplary CDS Papers

	Zhao et al. (2025)	Wei et al. (2025)	Zhang et al. (2024)	Samtani et al. (2022)
Introduction	▪ Describe the industry classification problem, which is fundamental to both firms' decision-making and business research. ▪ Identify deficiencies of existing industry classification systems and document embedding methods, and highlight methodological contributions.	▪ Elaborate the need for hate speech detection on social media and prevalent biases in training data sets due to language use differences. ▪ Identify the needs for expert-labeled training data and for methods that do not require a large number of labeled instances.	▪ Describe a novel resource management problem for disaster response, which aims to decide types and quantities of requested disaster relief resources. ▪ Indicate how existing literature overlooks the problem.	▪ Describe the problem of linking hackers' exploits to system vulnerabilities, the solution to which enables firms to address cybersecurity proactively. ▪ Indicate how existing IS studies overlook the problem.
Related Work	▪ Review FinTech research on industry classification and general-purpose document embedding methods. ▪ Elaborate methodological novelties: a novel method with two key methodological innovations -- the adaptive gating mechanism for modeling concept switching and the dynamic embedding mechanism for capturing document evolution.	▪ Review prior research on debiasing methods for hate speech detection, including adversarial training and contrastive learning. ▪ Review representative work in prompt-based learning, knowledge injection, and fine-tuning in large language models. ▪ Discuss methodological novelties: a novel debiasing method that combines a novel pair generator for contrastive learning and an enhanced continuous prompt for prompt-based learning.	▪ Review problem-specific research on resource management for disaster response and inventory control models, as well as general-purpose deep learning-based temporal point process (TPP) methods. ▪ Elaborate methodological novelties: a novel TPP method that predicts types and quantities of requested disaster relief resources and an innovative stochastic optimization method that decides optimal quantities of resources requested.	▪ Review problem-specific research on hacker community and vulnerability assessment, as well as general-purpose ML algorithms such as deep structured semantic model (DSSM) and attention mechanism. ▪ Elaborate novelties: a method that links exploits to vulnerabilities by capturing and prioritizing overlapping text features between them and a new metric that measures the severity of system vulnerability.
Problem Formulation	▪ *Given* a collection of document sequences, each containing a sequence of 10-K documents filed by a firm over years, the *objective* is to learn a representation of each document that captures both concept switching and document evolution.	▪ *Given* a small number of expert-labeled unbiased examples, the *objective* is to classify whether each social media post contains hate speech in an unbiased way.	▪ *Given* the quantity of remaining resources or unfulfilled demands for each resource type, the *objective* is to decide the quantity of requested resources for each resource type, such that the cost incurred by delays in satisfying unfulfilled demands is minimized, while satisfying the *constraint* of transportation capacity.	▪ *Given* a collection of exploits and a set of vulnerabilities, the *objective* is to link exploits to vulnerabilities.

Method	▪ To realize the specified methodological novelties, they develop two novel modules of their method: the adaptive gating mechanism and the dynamic embedding mechanism. ▪ The underlying design rationales include representing a document as a weighted combination of concepts, adapting these weights across different parts of a document to model concept switching, and dynamically updating them over time to capture document evolution.	▪ Two specific methodological novelties: a novel pair generator for contrastive learning and an enhanced continuous prompt with relevant external knowledge for prompt-based learning. ▪ Method design is guided by previous research on language use differences and is enlightened by the social identity theory.	▪ To realize the specified methodological novelties, two novel modules are designed: a TPP module for future demand prediction and a stochastic optimization module that decides an optimal resource requesting plan. ▪ Method design is guided by the rationale of proactively deciding optimal quantities of requested resources based on both currently unfulfilled demands and future demands. It is grounded in literature on deep learning-based TPP models and stochastic optimization.	▪ The proposed method consists of two modules: a deep learning-based method that integrates attention mechanisms into DSSMs to capture and prioritize overlapping text features between exploits and vulnerabilities and a new metric for measuring the severity of system vulnerability. ▪ Method design is guided by the rationale that, to effectively identify connections between exploits and vulnerabilities, it is critical to capture overlapping text features between them. It is grounded in ML literature on neural information retrieval and attention.
Evaluation	▪ Compare the proposed method with well-justified benchmarks using problem-specific metrics. ▪ Conduct an ablation study to show the respective contribution of the adaptive gating mechanism and the dynamic embedding mechanism. ▪ Perform two case studies: stock portfolio construction and mergers and acquisitions prediction.	▪ Compare the proposed method with benchmark methods using real-world datasets and ML metrics such as false positive rate, F1 score, and average statistical parity. ▪ Conduct ablation analysis to examine the contribution of each key component of the proposed method. ▪ Demonstrate the generalizability of the approach by applying the model in a different application domain.	▪ Demonstrate the superior performance of the designed method over well-justified benchmarks using problem-specific metrics such as deprivation cost. ▪ Conduct an ablation study to show the respective contribution of each methodological novelty. ▪ Perform a case study of a real-world humanitarian resource management system with multiple stakeholders and objectives.	▪ Demonstrate the superior performance of the designed method over well-justified benchmarks using ML metrics such as NDCG. ▪ Evaluate the contribution of each method component to the performance of the method. ▪ Demonstrate the practical value of the designed method with two case studies: one on publicly accessible devices in top U.S. hospitals and the other on SCADA devices.
Conclusion	▪ Discuss the contributions of the study, namely, a novel document embedding method and a new ICS based on the method, and their impact on business research and practices.	▪ Highlight the methodological contributions to contrastive learning and prompt-based learning, and its implications for AI fairness and algorithmic bias.	▪ Discuss the methodological contributions of the study and its implications for disaster response management practices.	▪ Articulate the contributions of the study, namely, the designed method and metric, and discuss their practical implications.

Conclusion

Ram and Goes (2021) call on IS researchers to take the lead in producing high-impact research to address the grand challenges faced by business and society. To achieve this noble goal, "the design science paradigm has to be considered as the face value" of IS publications (Ram & Goes, 2021, p. 479), and "design science … will need to integrate with data science" (Ram & Goes, 2021, p. 481). The integration of design science and data science is the CDS that we seek to foster with this article. The IS community should respond to this clarion call by placing significantly greater emphasis on CDS research. We are excited that the IS field is starting to recognize the criticality of CDS research, as reflected by recent publications in top IS journals, which propose novel, effective AI methods to address significant business and societal challenges (e.g., Wei et al., 2022; Zhao et al., 2023) or tackle fundamental issues in the development and use of AI methods, such as bias, privacy, interpretability, and human-in-the-loop (e.g., Kim et al., 2023; Lee & Ram, 2024).

Advents in generative and agentic AI and their increasing prevalence further opportunities for CDS research to maximizing their benefits while minimizing potential harms (Patterson, 2024). These exciting developments also raise growing concerns, partially because foundational AI capabilities are created outside the IS community and frequently extended by large technology firms with abundant resources. On balance, we believe that these advances will not hinder CDS research; rather, they can create more opportunities. In the past, while most traditional ML algorithms (e.g., decision trees, neural networks, and SVM) originated outside the IS field, many CDS studies still proposed new computational methods based on these algorithms for various application domains (e.g., Abbasi et al., 2010; Chau et al., 2020; Zhao et al., 2023). With powerful AI building blocks, such as LLMs and AI agents, we anticipate ample opportunities for

CDS researchers to design computational algorithms and methods that harness such advanced AI models to solve grand business and societal problems, while contributing to both IS and AI methodologically. We already witness multiple CDS studies that leverage generative AI to make both domain and methodological contributions (e.g., Wei et al., 2025; Wang & Ge, 2026; Wang et al., 2026).

Meanwhile, CDS research should be mindful of potential harms associated with the use or misuse of AI technologies, such as bias, privacy violation, and misinformation (Dathathri et al., 2024; Fang et al., 2024). Rapid advances of generative AI make these issues more severe, which also creates new opportunities for CDS research to devise new methods to combat these issues and alleviate their negative impacts (Wei et al., 2025). In addition, like other research areas, CDS research can benefit from the use of generative and agentic AI tools in the research workflow. Such tools can assist in various tasks such as formulating research problems and proposing novel design insights. This could potentially accelerate the knowledge creation process in CDS and IS research in general.

Because of its problem-solving nature and intent, CDS research has broader impacts that extend beyond academia to benefit both business organizations and the society as large. Long-term, high-impact CDS research programs not only generate academic publications but also advance the design, deployment, and use of pragmatic IT artifacts that make positive differences to the world (Ram & Goes, 2021). We, therefore, call on IS scholars to pursue high-impact CDS research programs exemplified by group decision support systems (Nunamaker et al., 2017), law enforcement analytics systems (Chen et al, 2003), and cyber security methods (Samtani et al., 2022), instead of publishing CDS research papers only. We anticipate and celebrate the flourishing of CDS research in the IS discipline. With this article, we seek to advocate and

facilitate this exciting move by clarifying the relevance, importance, identity, and contribution of CDS research, and by suggesting a template to guide its presentation in publications.

We conclude this article with takeaways for authors, reviewers, and editors.

For authors:

- CDS research designs novel ML algorithms and methods to solve important business and societal problems with impact, thereby making methodological contribution to IS as well as to AI and data science.
- Design of ML algorithms and methods should be grounded in sound rationales, justifications, and mechanisms, which may or may not be informed by theory.
- The ML algorithm or method developed in a CDS study should exhibit significant methodological novelty relative to existing ML algorithms and methods that are designed to address the same problem, as well as general-purpose ML algorithms and methods that can be adapted to solve the problem.

For Reviewers:

- CDS research should be evaluated by its methodological contribution. Therefore, methodological novelty of designed ML algorithms and methods is critical.
- CDS research need not be theory-based. The rigor of CDS research lies in the soundness of design rationales, justifications, mechanisms. Such rationales, justifications, and mechanisms may stem from a logical and creative analysis of the focal problem together with a thorough review of relevant knowledge, while they may also be informed by theory.
- Reviewers should stay abreast of the rapidly evolving ML literature relevant to a CSD study they review.

For editors:

- The significance of CDS in IS should be recognized. CDS represents a unique and crucial area that contributes to the IS knowledge base, while simultaneously advancing the methodological core of AI and data science.
- The identity of CDS should be acknowledged. Within the design science paradigm, CDS differs from other genres of design science research in its focus on the development of novel ML algorithms and methods. More broadly, CDS is different from ML research in other disciplines, such as computer science, and strengthens the identity of the IS discipline.

References

Abbasi, A., Zhang, Z., Zimbra, D., Chen, H. and Nunamaker Jr, J.F. 2010. "Detecting Fake Websites: The Contribution of Statistical Learning Theory," *MIS Quarterly*, 34(3), pp. 435-461.

Abbasi, A., Parsons, J., Pant, G., Sheng, O. R. L., and Sarker, S. 2024. "Pathways for Design Research on Artificial Intelligence," *Information Systems Research*, 35(2), pp. 441-459.

Agarwal, R. and Lucas H. C. 2005. "The Information Systems Identity Crisis: Focusing on High-Visibility and High-Impact Research," *MIS Quarterly*, 29(3), pp. 381-398.

Anand, P. and Lee, C. 2023. "Using Deep Learning to Overcome Privacy and Scalability Issues in Customer Data Transfer," *Marketing Science*, 42(1), pp. 189-207.

Bardhan, I., Oh, J.H., Zheng, Z., and Kirksey, K. 2015. Predictive Analytics for Readmission of Patients with Congestive Heart Failure. *Information Systems Research*, 26(1), pp. 19-39.

Benbasat, I. and Zmud, R. W. 2003. "The Identity Crisis within the IS Discipline: Defining and Communicating the Discipline's Core Properties," *MIS Quarterly*, 27(2), pp. 183-194.

Boyacı, T., Canyakmaz, C. and de Véricourt, F. 2024. "Human and Machine: The Impact of Machine Input on Decision Making under Cognitive Limitations," *Management Science*, 70(2), pp.1258-1275.

Brown, S. 2024. "Maximizing the Chances of Securing a Revision by Focusing on the Essentials," *MIS Quarterly*, 48(3), pp. iii-viii.

Cleveland, W.S. 2001. "Data Science: An Action Plan For Expanding the Technical Areas of the Field of Statistics," *International Statistical Review*, 69(1), pp. 21-26.

Chau, M., Li, T. M. H., Wong, P. W. C., Xu, J. J., Yip, P. S. F., and Chen, H. 2020. "Finding People with Emotional Distress in Online Social Media: A Design Combining Machine Learning and Rule-based Classification," *MIS Quarterly*, 44(2), pp. 933-955.

Chen, G., Ren, M., Wei, Q., Guo, X., Yi. C. 2022. "Data-Intelligence Empowerment: A New Leap of Information Systems Research", *Management World*, 38(1), pp.180-196.

Chen, H., Zeng, D, Atabakhsh, H., Wyzga, W., Schroeder, J. 2003. "COPLINK: Managing Law Enforcement Data and Knowledge," *Communications of the ACM*, 46(1), pp. 28–34.

Chen, H. 2017. "Computational Genre of Design Science Research," *MIS Quarterly*, 41(1), pp. vii-ix.

Clarke, R., Burton-Jones, A., and Weber, R. 2016. "On the Ontological Quality and Logical Quality of Conceptual-Modeling Grammars: The Need for a Dual Perspective," *Information Systems Research*, 27(2), pp. 365-382.

Cormen, T.H., Leiserson, C.E., Rivest, R.L., and Stein, C. 2022. *Introduction to Algorithms*. MIT press.

Dathathri, S., See, A., Ghaisas, S., Huang, P.S., McAdam, R., Welbl, J., Bachani, V., Kaskasoli, A., Stanforth, R., Matejovicova, T. and Hayes, J. 2024. "Scalable Watermarking for Identifying Large Language Model Outputs," *Nature*, 634(8035), pp.818-823.

Davis, F.D. 1989. "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology," *MIS Quarterly*, *13*(3), pp.319-340.

Donoho, D. 2015. "50 Years of Data Science," Working Paper Based on Presentation at the Tukey Centennial Workshop.

Fang, X., and Hu, P. J. 2018. "Top Persuader Prediction for Social Networks," *MIS Quarterly*, 42 (1), pp. 63-82.

Fang, X., Che, S., Mao, M., Zhang, H., Zhao, M., and Zhao, X. 2024. "Bias of AI-generated Content: An Examination of News Produced by Large Language Models," *Scientific Reports*, *14*(1), 5224.

Fuster, A., Goldsmith-Pinkham, P., Ramadorai, T. and Walther, A. 2022. "Predictably Unequal? The Effects of Machine Learning on Credit Markets," *Journal of Finance*, 77(1), pp.5-47.

Ganin, Y., Ustinova, E., Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., March, M. and Lempitsky, V. 2016. "Domain-adversarial Training of Neural Networks," *Journal of Machine Learning Research*, 17(59), pp.1-35.

Goes, P. 2014. "Design Science Research in Top Information Systems Journals," *MIS Quarterly*, 38 (1), pp. iii-viii.

Goodfellow, I., Bengio, Y. and Courville, A., 2016. *Deep Learning*. MIT press.

Gregor, S. and Hevner, A.R. 2013. "Positioning and Presenting Design Science Research for Maximum Impact," *MIS Quarterly*, 37(2), pp. 337-355.

Gu, S., Kelly, B. and Xiu, D. 2020. "Empirical Asset Pricing via Machine Learning," *Review of Financial Studies*, 33(5), pp.2223-2273.

Hevner, A.R., March, S.T., Park, J. and Ram, S. 2004. "Design Science in Information Systems Research," *MIS Quarterly*, 28 (1), pp. 75-105.

Hu, Y. J., Rombouts, J., and Wilms, I. 2024. Fast Forecasting of Unstable Data Streams for On-Demand Service Platforms. *Information Systems Research*. Forthcoming.

Kim, B. R., Srinivasan, K., Kong, S. H., Kim, J. H., Shin, C. S., and Ram, S. 2023. "ROLEX: A Novel Method for Interpretable Machine Learning Using Robust Local Explanations," *MIS Quarterly*, 47(3). pp. 1303-1329.

Lee, A. S. 1999. "Inaugural Editor's Comments," *MIS Quarterly*, 23(1), pp. v-xi.

Lee, K. and Ram, S. 2024. "Explainable Deep Learning for False Information Identification: An Argumentation Theory Approach," *Information Systems Research*, 35(2), pp. 890-907.

Li, Z., Fang, X., Bai, X., and Sheng, O. 2017. "Utility-based Link Recommendation for Online Social Networks," *Management Science*, 63(6), pp. 1938-1952.

Mao, M., Zhao, X. and Fang, X. 2026. "Early Detection of Misinformation for Infodemic Management: A Domain Adaptation Approach," *Management Science, forthcoming*.

Menon, S. and Sarkar, S. 2016. "Privacy and Big Data: Scalable Approaches to Sanitize Large Transactional Databases for Sharing," *MIS Quarterly*, 40(4), pp. 963-982.

Meyer, G., Adomavicius, G., Johnson, P., Elidrisi, M., Rush, W., Sperl-Hillen, J., and P. O'Connor, P. 2014. "A Machine Learning Approach to Improving Dynamic Decision Making," *Information Systems Research*, 25(2), pp. 239-263.

Moskowitz, T. J., and Grinblatt, M. 1999. "Do Industries Explain Momentum?" *The Journal of Finance* 54(4), pp.1249-1290.

Natioanl Science Foundation. 2022. National Artificial Intelligence (AI) Research Institutes Accelerating, https://new.nsf.gov/funding/opportunities/national-artificial-intelligence-research/nsf22-502/solicitation

Niederman, F. and March, S. 2012. "Design Science and the Accumulation of Knowledge in the Information Systems Discipline," *ACM Transactions on MIS*, 3(1), pp. 1-15.

Nunamaker, Jr., J. F., Twyman, N. W., Giboney, J. S., and Briggs, R. O. 2017. "Creating High-Value Real-World Impact through Systematic Programs of Research," *MIS Quarterly*, 41(2), pp. 335-351.

Özsu, M.T., 2023. "Data Science—A Systematic Treatment," *Communications of the ACM*, 66(7), pp.106-116.

Padmanabhan, B., Fang, X., Sahoo, N., and Burton-Jones, A. 2022. "Machine Learning in Information Systems Research," *MIS Quarterly*, (46: 1) pp.iii-xix.

Patterson, D. 2024. "An Agenda to Maximise AI's Benefits and Minimise Harms," *The Economist*, December 3, 2024.

Rai, A. 2017. "Diversity of Design Science Research," *MIS Quarterly*, 41(1), pp. iii-xviii.

Ram, S., and Goes, P. 2021. "Focusing on Programmatic High Impact Information Systems Research, Not Theory, to Address Grand Challenges," *MIS Quarterly*, 45(1), pp. 479-483.

Roberts, D.A., Yaida, S. and Hanin, B., 2022. *The Principles of Deep Learning Theory* (Vol. 46). Cambridge, MA, USA: Cambridge University Press.

Russell, S.J. and Norvig, P., 2020. *Artificial intelligence: a Modern Approach*. London.

Samtani, S., Chai, Y., and Chen, H. 2022. "Linking Exploits from the Dark Web to Known Vulnerabilities for Proactive Cyber Threat Intelligence: An Attention-Based Deep Structured Semantic Model Approach," *MIS Quarterly*, 46(2), pp. 911-946.

Saar-Tsechansky, M. 2015. "Editor's Comments: The Business of Business Data Science in IS Journals," *MIS Quarterly*, 39(4), pp. iii-vi.

Simon, H., 1996. *The Science of the Artificial*. 3rd Edition, MIT Press.

Sun, H., Wen, W., Thatcher, J. B., Chau, M., and Brown, S. A. (2025). "Rebalancing Novelty with Rigor and Relevance in Information Systems Research," *MIS Quarterly*, 49(3), pp. iii-xi.

Tukey, J. W. 1962. "The Future of Data Analysis," *Annals of Mathematical Statistics*, 33(1), pp.1-67.

Ullman, J.D. 2020. "The Battle for Data Science," *IEEE Data Engineering Bulletin*, 43(2), pp.8-14.

Wang, Y., and Ge, Y. 2026. Align Generative Artificial Intelligence with Human Preferences: A Novel Large Language Model Fine-Tuning Method for Online Review Management. *Information Systems Research*, forthcoming.

Wang, Z. Jiang, C., and Zhao, H. 2023. "Depicting Risk Profile over Time: A Novel Multi-Period Default Prediction Approach," *MIS Quarterly***,** 47(4), 1455-1486.

Wang, Z., Wu, C., Jiang, C., and Zhao, H. 2026. "Generative AI as an Information Intermediary: A Novel Deep Learning Method for Financial Distress Prediction," *MIS Quarterly*, forthcoming.

Wei, X., Zhang, Z., Zhang, M., Chen, W., and Zeng, D. D. 2022. "Combining Crowd and Machine Intelligence to Detect False News on Social Media," *MIS Quarterly*, 46(2) pp.977-1008.

Wei, D., Chau, M., and Li, Z. 2025. "Mitigating Bias in Hate Speech Detection with a Small Number of Expert Annotations: A Prompt-Based Learning Approach," *MIS Quarterly* 49(4), pp. 1483-1512.

Yin, K., Fang, X., Chen, B., Sheng, O. 2023. "Diversity Preference-aware Link Recommendation for Online Social Networks," *Information Systems Research*, 34(4), pp. 1398-1414.

Zhang, H., Zhao, X., Fang, X. and Chen, B. 2024. "Proactive Resource Request for Disaster Response: A Deep Learning-Based Optimization Model," *Information Systems Research*, 35(2), pp.528-550.

Zhang, H., Qian, W. and Fang, X. 2025. "Latent Similarity-Enhanced Credit Risk Prediction," *MIS Quarterly*, forthcoming.

Zhao, X., Fang, X., He, J., and Huang, L. 2023. "Exploiting Expert Knowledge for Assigning Firms to Industries: A Novel Deep Learning Method," *MIS Quarterly*, 47(3), pp.1147-1176.

Zhao, X., Fang, X., He, J. and Sheng, O. 2025. "A Novel Document Embedding Approach to Industry Classification," *Available at SSRN 5158103*.