Design Science Research: Evaluation in the Lens of Big Data Analytics

Ahmed Elragal [1] and **Moutaz Haddara [2],***

[1] Department of Computer Science, Electrical and Space Engineering, Computer and Systems Science, Luleå University of Technology, 97187 Luleå, Sweden; ahmed.elragal@ltu.se
[2] Department of Technology, Kristiania University College, 0186 Oslo, Norway
* Correspondence: moutaz.haddara@kristiania.no

check for updates

Abstract: Given the different types of artifacts and their various evaluation methods, one of the main challenges faced by researchers in design science research (DSR) is choosing suitable and efficient methods during the artifact evaluation phase. With the emergence of big data analytics, data scientists conducting DSR are also challenged with identifying suitable evaluation mechanisms for their data products. Hence, this conceptual research paper is set out to address the following questions. Does big data analytics impact how evaluation in DSR is conducted? If so, does it lead to a new type of evaluation or a new genre of DSR? We conclude by arguing that big data analytics should influence how evaluation is conducted, but it does not lead to the creation of a new genre of design research.

Keywords: DSR; big data analytics; evaluation

1. Introduction

Design science research brings both practical as well as theoretical rigor to information systems research [1]. According to Rai [2], there exist identifiable genres of design science research (DSR) (see Figure 1). They vary with regard to the problem being addressed, the types of artifacts designed and evaluated, the search processes to create and refine the IT artifacts, and the types of knowledge contribution. These genres are neither exhaustive nor mutually exclusive [2]. These genres are shown below.

- *Computational Genre*:

  ○ This class of DSR research stresses an interdisciplinary approach in developing artifacts such as data representations, algorithms e.g., machine learning algorithms, analytics methods, and human–computer interaction (HCI) innovations.

- *Optimization Genre*:

  ○ This class of DSR research looks at the creation of IT artifacts that are intended to solve organizational problems such as: maximizing profits, utilities, welfare, etc. Additionally, it falls under the optimization of supply chain activities, internal operations, customer relationship management activities (e.g., the effective use of personalization technologies), and pricing decisions (e.g., price discrimination strategies enabled by analytics).

- *Representation Genre*:

  ○ This class of DSR research contributes by evaluating and refining existing modeling grammar or methods, the design of new modeling grammar or methods, the development

of software artifacts to support or instantiate such work, or the evaluation of these efforts using analytical methods. Such contributions face challenges from philosophy, linguistics, and psychology that unavoidably occur when experimenting representations.

-   *IS Economics Genre*:

    ○   This class of DSR research contributes by developing and refining models and algorithms that focus on the role of IS to solve problems related to the conduct of economic activities and attainment of objectives of an economic system. Additionally, the discovery and characterization of behaviors of economic participants. The instantiation of models and algorithms into software, technology platforms, and other artifacts. Lastly, the evaluation of the validity of causal mechanisms and the utility of solutions. This class mainly focuses on unfolding the relationship between IS and the design of economic systems.
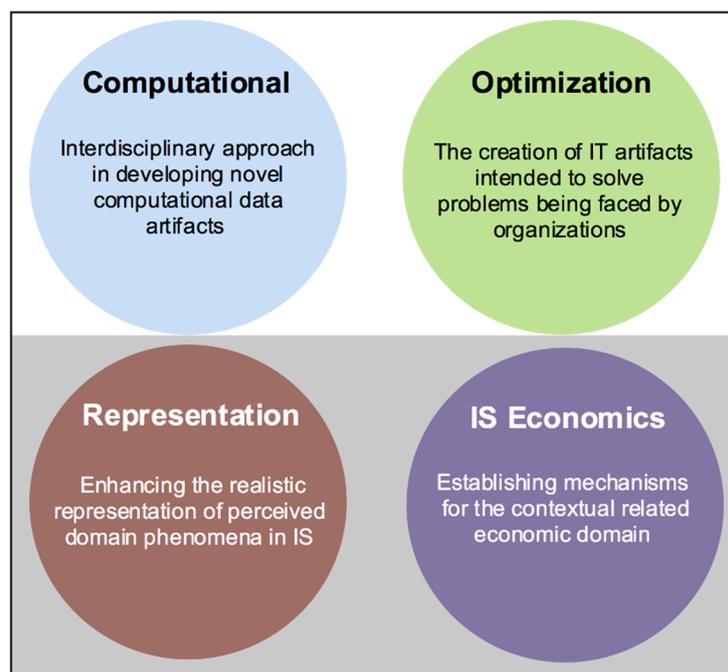
Figure 1. DSR genres.

On the other hand, big data analytics (BDA) research focuses more on predictive research than exploratory research [3]. In such types of research, machine-learning techniques in addition to other techniques are used to unfold hidden patterns and elucidate knowledge from large and complex datasets [4]. Big data analytics research has been used in several domains such as predicting medical risks in healthcare [5], and detecting potential bottlenecks in supply chains [6]. Accordingly, BDA is leaning itself more toward the first two genres of DSR: computational and optimization. The only thing observed to be different is the "data." That is, big data analytics enabled vast amounts of data to be readily available for DSR and other research paradigms. Hence, the way DSR has been conducted should incorporate techniques that have the power to verify the data in order to avoid the golden rule stipulating "garbage-in, garbage-out." The DSR community would benefit from adopting the emerging big data and analytics platforms to create more sophisticated statistical models as well as machine learning models.

In summary, big data analytics has the potential to impact how evaluation in DSR is conducted. However, it does not lead to a new genre of DSR. Instead, BDA is closer to both the computational as well as the optimization genres.

2. Design Science Research

Design science research has been part of the engineering and information systems research, under different names, the last 30 years [7,8]. The design science epitome is an outcome-based research paradigm that pursues creating novel information technology (IT) artifacts [9]. For design science, the information system (IS) is one problem domain of many possible domains. In principal, design science aims at understanding and refining the search among probable constructs and components in order to develop artifacts that are projected to solve an existing organizational problem or challenge in their natural setting [10]. DSR attempts to improve the functional performance of the designed artifacts through the continuous focus on the development process and performance evaluation of those artifacts [9]. Such artifacts may vary from software applications, formal logic, and rigorous mathematical models and equations, to informal narratives and descriptions in a natural language [9]. According to March and Smith [11], design science outputs can be categorized in four main types, which include constructs, models, methods, and implementations. However, several research studies have extended these categories to include architecture, design principles, frameworks, instantiations, and theories [12,13]. From an abstract level, artifacts can be product or process artifacts [14]. Several IT artifacts have some level of abstraction, but can be potentially transformed to a more material form, like an algorithm converted to an operational model or software [14]. Product artifacts are usually technical or socio-technical. However, process artifacts are always socio-technical. Socio-technical artifacts are ones whose target users/humans must interact with, in order to provide their intended functions [15].

At the core of the DSR is the concept of learning through the systematic building and creation of knowledge, in which the research outcomes should deliver new, innovative, true, and interesting knowledge, designs, or artifacts within the respective community of interest [13]. During this creation process, the researcher must be aware of progressing both the design process and the design artifact as part of the research [9].

Since the artifacts are purposeful, the continuous evaluation of these artifacts during their creation process is deemed critical in conducting rigorous DSR [10,11,15,16]. The artifact evaluation mechanisms will provide a better interpretation of the problem and feedback in order to advance the quality of the designed artifact, as well as, the design effort and process [9]. In addition, the rigorous evaluation demonstrates evidence that the innovative artifact fulfills the purpose that it was designed for [15]. The evaluation mechanisms are usually iterative during the artifact creation process [17], and are conducted after the creation of the artifact. Artifact evaluation, however, may be a complex task, since the performance of the artifact is tightly coupled to its desired use and purpose, and the purpose and desired use of an artifact might cover a wide range of tasks [11]. Given the wide variety of evaluation methods and strategies, choosing the proper and suitable evaluation methods and efficient strategies may be a troublesome task for the design science researcher. For example, ex-ante and ex-post artifact evaluations can be quantitative or qualitative in nature. A mathematical basis for design allows diverse types of quantitative evaluations of an IT artifact, including optimization proofs, simulations, and comparison techniques with alternative designs. The evaluation of a new artifact in a given organizational context may allow for the possibility to implement empirical and qualitative evaluation methods [9].

While artifact evaluation is considered a crucial issue in DSR, it is difficult to find solid guidelines on how to select, design, and implement an appropriate evaluation strategy [15]. This led to several efforts in developing and shaping various DSR evaluation frameworks and guidelines [15,18]. While these evaluation guidelines aid researchers in choosing and designing appropriate strategies, the diversity of evaluation techniques still make this process challenging.

3. Evaluation

According to [8], the evaluation ranges between the options below.

- If the DSR aims to develop design theories, then hypotheses proposed could be tested via experiments, conceptually, or instantiated-oriented;
- In case DSR aims to design an artefact, system, or method, then demonstration is an acceptable evaluation mechanism;
- Design oriented research could utilize lab experiments, pilot testing, simulations, expert reviews, and field experiments;
- Explanatory design, on the other hand, could be evaluated via hypothesis testing and experimental setup;
- Actions design research views design and evaluation as sequential. That is, they look at one process where organizational intervention and evaluation are required.

4. Big Data Analytics

The notion of big data goes far beyond the rapidly increasing quantity and variety of data, and focuses on the analysis of these data to enable fact-based decision-making [19]. Independent from the specific peta, exa, or zettabytes scale, the key feature of the paradigmatic change is that the analytic treatment of data is systematically placed at the forefront of fact-based decision-making [19]. Hence, several scientists and practitioners argue that the full name of "big data" is "big data analytics." Big data analytics is the use of advanced data mining and statistical techniques in order to find hidden or previously unknown patterns in big data. Thus, BDA is where advanced techniques operate on big datasets. A significant amount of these techniques rely on software tools such as data mining tools like RapidMiner, SAS Enterprise Miner, R, and others.

We produce data using our mobile phones, social networks interactions, and GPS. Most of such data, however, are not structured in a way so as to be stored or processed in traditional database management systems (DBMSs) [20]. This rather calls for big data analytics techniques in order to make sense out of such unstructured and semi-structured data. According to IBM, exponential data growth is leaving most organizations with serious blind spots. A survey conducted by IBM found that one in three business leaders admit to frequently making decisions with no data to back them up [21]. Conversely, survey results from a study by Brynjolfsson et al. [22] argue that top-performing organizations in the market use analytics five times more than lower performers.

With the emergence of big data, the data science domain has evolved. The data science approach is fundamentally different from the traditionally used approach, wherein the traditional business/data analysts choose an appropriate method based on their knowledge of techniques and the data [23]. In other words, BDA enable an entirely new epistemological approach for making sense of the world instead of by testing a theory by analyzing relevant data. New data scientists seek to gain insights that are a 'byproduct' from the data [24]. Thus, inductive and ensemble approaches can be employed to build multiple solutions [25]. This means that we can apply hundreds of different algorithms to a dataset in order to determine the best model or a composite model or explanation [26]. The 'best' model is usually chosen based on the best performance evaluation, which could be attained by several evaluation methods (e.g., cross validation). Specifically, data science regards computers and techniques as credible generators and testers of hypotheses by amending some of the known errors associated with statistical induction. Machine learning, which is characterized by statistical induction aimed at generating robust predictive models, becomes central to data science. Hence, unlike the traditional approach which asks "what data satisfy this pattern?" data science is about asking "what patterns satisfy these data?" [27]. Precisely, the data scientist aim is to uncover interesting and robust patterns that satisfy the data analyzed, where interesting is usually something unexpected and actionable. Robust means a pattern that is expected to occur in the future, and actionable refers to the predictive power in that the response related to an action can be reliably predicted from historical data, and acted upon with a high degree of confidence [27].

As in DSR, explaining and evaluating the BDA artifacts (during and after model creation) are non-trivial tasks. In addition, ensuring the high (big) data quality during the model building is

problematic. Thus, some researchers have called for the utilization of artificial intelligence and crowdsourcing concepts to aid organizations in structuring the unstructured (or semi-structured) data via generating tags on data that can facilitate the data use and definition, and can serve as master data management guidelines [28]. These approaches have the potential to enhance the data quality and eventually enhance the artifact creation process. Moreover, one of the important aspects of DSR-developed artifacts, is how to communicate knowledge contributions [7] or, in other words, the DSR impact [1]. In the BDA realm, this is accomplished via the constant interaction with the business domain experts throughout the artifact design and evaluation process, and through demonstrating artifacts' validations, evaluations, and interpretations to the research community.

5. Evaluation in the Lens of BDA

When designing novel IT artifact or introducing an artifact into an application domain, the artifact must demonstrate measurable improvements in order to illustrate technology advancement and evolution [1]. Thus, evaluation is a core activity in conducting design science research and artifact construction.

In their seminal article, Venable et al. [15] suggested an evaluation framework. The framework includes the choice of ex-ante (prior to artifact construction) versus ex-post evaluation (after artifact construction) and naturalistic (e.g., field setting) versus artificial evaluation (e.g., laboratory setting). Their framework distinguishes product artifacts from process artifacts. Product artifacts are technologies such as tools, diagrams, software, and more. On the other hand, process artifacts are methods, procedures, etc.

In the era of big data, and mainly in the two genres of DSR (computational and optimization), data play a vital role in artifact construction. Hence, evaluation in the lens of big data should also include "veracity" i.e., correctness of the data. That is, data now comes from external sources that neither the organization nor the researcher have control over. Overlooking that, data veracity could lead to problems in the way DSR is conducted. Concerned with the theoretical knowledge needed to appropriately apply BDA within the frame of DSR, researchers should seek to address the following questions [3]:

- What kind of data [or datasets] about the world are available to a data scientist or researcher?
- How can these data [sets] be represented?
- What rules govern conclusions to be drawn from these datasets?
- How to interpret such a conclusion?

Big data analytics starts with acquiring the data through copying, streaming, and more. Such acquisition requires good understanding of the business domain as well as the data. Datasets, from which we source data, should be described in terms of: required data to be defined, background about the data, list of data sources, the method of acquisition or extraction for each data source, and reporting the problems encountered in data acquisition or extraction. One of the challenges associated with big data acquisition is: on one hand, there exist too much data while, on the other hand, all acquisitions require time, effort, and resources. The selection by the researcher might be attributable to: personal preference, technical abilities, or access to data and the 'streetlight' effect. In practice, researchers seek technological solutions, i.e., tools to acquire and compress the data, and focus on available data. However, such solutions do not really address the epistemological problem: we know that sampling in data collection is crucial and requires a great deal of reflection pertaining to the impact of data acquisition decisions on the result of the research.

On the other hand, preprocessing activities include: checking keys, referential integrity, and domain consistency, identifying missing attributes and blank fields, replacing missing values, data harmonization e.g., checking different values that have similar meanings such as customer, client, checking spelling of values, and checking for outliers. In result, preprocessing provides a description of the dataset including: background (broad goals and plan for pre-processing), rationale

for inclusion/exclusion of datasets, description of the pre-processing, including the actions that were necessary to address any data quality issues, detailed description of the resultant dataset, table by table and field by field, rationale for inclusion/exclusion of attributes, and the discoveries made during pre-processing and their potential implications for analytics.

Preprocessing primarily aims for big data cleansing and harmonization, while quite often overlooking the importance of 'traditional' data collection by the researcher. Big data self-confidence tends to drive preprocessing toward the assumption that big data are a substitute for, rather than a supplement to, traditional data collection and analysis. The core challenge is that most big data in focus are not the output of instruments that were designed to produce valid and reliable data amenable for rigorous knowledge discovery [3].

We revisit our research question by asserting that, in the lens of big data analytics, big data analytics impact how evaluation in DSR is conducted by requiring more evaluation mechanisms pertaining to data acquisition and data preprocessing.

6. Conclusions

We intended to highlight the need for further research pertaining to evaluation in DSR, in the lens of big data analytics. We adopt the criteria postulated by Rai [29], in evaluating research based on: aesthetic, scholarly, and practical utility. Our idea to reconsider evaluation is expected to create scholarly value for both DSR and data science communities. In addition, enabling practical utility results in a broader impact on business and society conducting problem-solving approaches. Information systems are socio-technical systems that include people, processes, and information technology, and BDA open up many new and exciting research domains, questions, and possibilities [30]. However, there are several cases where BDA has been invasive, which violates and sometimes even harms people (see [31]). Hence, due to the nature of data used in BDA projects, and due to the results that might be drawn from such analytics, we second Myres &Venable's [32] call for considering ethical guidelines and principals (e.g., privacy) during the creation of desired artifacts.

Lastly, BDA has the potential to impact how evaluation in DSR is conducted. However, BDA does not lead to a new genre of DSR. Instead, BDA is closer to both the computational as well as the optimization genres.

References

1.　Baskerville, R.; Baiyere, A.; Gregor, S.; Hevner, A.; Rossi, M. Design science research contributions: Finding a balance between artifact and theory. *J. Assoc. Inf. Syst.* **2018**, *19*, 358–376. [CrossRef]
2.　Rai, A. Editor's Comments: Diversity of Design Science Research. *Manag. Inf. Syst. Q.* **2017**, *41*, iii–xviii.
3.　Elragal, A.; Klischewski, R. Theory-driven or process-driven prediction? Epistemological challenges of big data analytics. *J. Big Data* **2017**, *4*, 19. [CrossRef]
4.　Haddara, M.; Su, K.L.; Alkayid, K.; Ali, M. Applications of Big Data Analytics in Financial Auditing-A Study on The Big Four. In Proceedings of the Americas Conference on Information Systems (AMCIS), New Orleans, Louisiana, USA, 16–18 August 2018.
5.　Wang, Y.; Kung, L.; Byrd, T.A. Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. *Technol. Forecast. Soc. Chang.* **2018**, *126*, 3–13. [CrossRef]
6.　Wang, G.; Gunasekaran, A.; Ngai, E.W.; Papadopoulos, T. Big data analytics in logistics and supply chain management: Certain investigations for research and applications. *Int. J. Prod. Econ.* **2016**, *176*, 98–110. [CrossRef]

7.　Gregor, S.; Hevner, A.R. Positioning and presenting design science research for maximum impact. *MIS Q.* **2013**, *37*, 337–355. [CrossRef]

8.　Peffers, K.; Tuunanen, T.; Niehaves, B. Design science research genres: introduction to the special issue on exemplars and criteria for applicable design science research. *Eur. J. Inf. Syst.* **2018**, *27*, 129–139. [CrossRef]

9.　Von Alan, R.H.; March, S.T.; Park, J.; Ram, S. Design science in information systems research. *MIS Q.* **2004**, *28*, 75–105.

10.　Baskerville, R. *What Design Science Is Not*; Springer: Berlin, Germany, 2008.

11.　March, S.T.; Smith, G.F. Design and natural science research on information technology. *Decis. Support Syst.* **1995**, *15*, 251–266. [CrossRef]

12.　Chatterjee, S. Writing My next Design Science Research Master-piece: But How Do I Make a Theoretical Contribution to DSR ? In Proceedings of the 23rd European Conference on Information Systems (ECIS 2015), Munster, Germany, 26–29 May 2015.

13.　Kuechler, B.; Vaishnavi, V. Design Science Research in Information Systems. 2004. Available online: http://www.desrist.org/design-research-in-information-systems (accessed on 22 May 2019).

14.　Gregor, S.; Jones, D. The anatomy of a design theory. *J. Assoc. Inf. Syst.* **2007**, *8*, 312.

15.　Venable, J.; Pries-Heje, J.; Baskerville, R. A comprehensive framework for evaluation in design science research. In Proceedings of the International Conference on Design Science Research in Information Systems, Las Vegas, NV, USA, 14–15 May 2012; pp. 423–438.

16.　Hevner, A.; Chatterjee, S. Design science research in information systems. In *Design Research in Information Systems*; Springer: Boston, MA, USA, 2010; pp. 9–22.

17.　Markus, M.L.; Majchrzak, A.; Gasser, L. A design theory for systems that support emergent knowledge processes. *MIS Q.* **2002**, 179–212.

18.　Pries-Heje, J.; Baskerville, R.; Venable, J.R. Strategies for Design Science Research Evaluation. In Proceedings of the ECIS, Galway, Ireland, 9–11 June 2008; pp. 255–266.

19.　Hilbert, M. Big data for development: A review of promises and challenges. *Dev. Policy Rev.* **2016**, *34*, 135–174. [CrossRef]

20.　Elragal, A.; Haddara, M. Big Data Analytics: A text mining-based literature analysis. In Proceedings of the Norsk Konferanse for Organisasjoners Bruk AV IT, Fredrikstad, Norway, 17–19 November 2014.

21.　Berman, S.; Korsten, P. Embracing connectedness: Insights from the IBM 2012 CEO study. *Strategy Leadersh.* **2013**, *41*, 46–57. [CrossRef]

22.　Brynjolfsson, E.; Hitt, L.M.; Kim, H.H. Strength in Numbers: How Does Data-Driven Decisionmaking Affect Firm Performance? Available online: https://ssrn.com/abstract=1819486 (accessed on 22 May 2019).

23.　Haddara, M.; Larsson, A.O. Big Data. How to find relevant decision-making information in large amounts of data? In *Metodebok for Kreative Fag*; Næss, H.E.P., Lene, Eds.; Universitetsforlaget: Oslo, Norway, 2017.

24.　Kitchin, R. Big Data, new epistemologies and paradigm shifts. *Big Data Soc.* **2014**, *1*. [CrossRef]

25.　Seni, G.; Elder, J.F. Ensemble methods in data mining: improving accuracy through combining predictions. *Synth. Lect. Data Min. Knowl. Discov.* **2010**, *2*, 1–126. [CrossRef]

26.　Siegel, E. *Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die*; John Wiley & Sons: Hoboken, NJ, USA, 2013; ISBN 1-118-35685-3.

27.　Dhar, V. Data science and prediction. *Commun. ACM* **2013**, *56*, 64–73. [CrossRef]

28.　O'Leary, D.E. Embedding AI and crowdsourcing in the big data lake. *IEEE Intell. Syst.* **2014**, *29*, 70–73. [CrossRef]

29.　Rai, A. Editor's Comments: Avoiding Type III Errors: Formulating IS Research Problems that Matter. *Manag. Inf. Syst. Q.* **2017**, *41*, iii–vii.

30.　Bichler, M.; Heinzl, A.; van der Aalst, W.M. Business analytics and data science: Once again? *Bus. Inf. Syst. Eng.* **2017**, *57*, 77–79. [CrossRef]

31.　Shim, J.P.; French, A.M.; Guo, C.; Jablonski, J. Big Data and Analytics: Issues, Solutions, and ROI. *CAIS* **2015**, *37*, 39. [CrossRef]

32. Myers, M.D.; Venable, J.R. A set of ethical principles for design science research in information systems. *Inf. Manag.* **2014**, *51*, 801–809. [CrossRef]