

Re-Visiting IS Design Science Artifacts: Making a Case for Critical Realism

Elin Uppström
Stockholm University
elinu@dsv.su.se

Abstract

In this paper critical realism is suggested as a suitable philosophical assumption to guide a separate, stand-alone retrospective evaluation of design science projects and artifacts. A main contribution of the paper is to argue that knowledge can be gained in retrospective evaluations of design science projects regardless of the success or non-success of the project itself. Thereby, retrospective evaluation complements current evaluations that are mostly means-end focused. The argumentation is supported through re-visiting two e-government design science projects, which can be considered both as failures and successes depending on the framing. Critical realism puts focus on knowing through making and widens the use of design science in areas where utility is not the main goal. Future research should focus on providing more details on how a critical realism retrospective in design science should be carried out.

1. Introduction

The aim of design science research (DSR) is to extend the boundaries of human and organizational capabilities by creating new and innovative artifacts; thus, the goal of design science is utility [1], [2]. This paper argues that in order to reach the aim of extending boundaries and capabilities in modern organizations, the goal of DSR should be more extensive; thus, DSR evaluation should be based on a wider range of goals than utility. Furthermore, [3] makes a call for future research that “illustrates the benefits of appropriate reflection in the context of past or ongoing DSR projects in greater detail” [3, p. 165]. Retrospective evaluation in and of DSR is therefore suggested and discussed in this paper.

Retrospective evaluation can be carried out both during and after finishing a DSR project, as a stand-alone effort. Furthermore, design science researchers should be self-critical in a retrospective where the researcher critically examines the design process and

the developed artifact(s). Critical realism (CR) is proposed as a suitable philosophical assumption to guide retrospective evaluation in DSR. Thereby, this research follows [4]–[6], who argue for moving beyond the limited positivistic notion of the DSR originator [7] by adopting a CR philosophy [8]. Furthermore, aligning DSR with CR highlights both rigor and relevance, including the promotion of emancipatory change, and provides a suitable basis for attempting to develop a more integrated and coherent body of theory and research [5]. Emancipatory elements in critical realism can thereby facilitate ethical and normative awareness in design science in line with what [9] suggests as suitable for information systems (IS) research.

The aim with CR retrospective is to improve the understanding of consequences of choices made and actions taken, and thereby, enrich the knowing through making process of DSR. By adopting CR as a philosophical basis the author envisions DSR to become the methodological choice in research when the knowing through making is fundamental but the goal is not primarily on achieving utility. Adopting CR thereby makes design science more suitable in some areas of research such as e-government. Design science is considered valuable for e-government; however, existing frameworks for design science do not provide specific guidance on how to apply DSR in an e-government context [10]. Following this line of thought, the argumentation made in this paper is illustrated and supported by re-visiting two e-government DSR projects where success and failure differ depending on the framing.

This research thereby complements and extends existing work on DSR since (i) it explicates that DSR can have different and/or complementary goals in parallel with the traditional utility goal advocated by e.g. [1], [2]. (ii) Retrospective evaluation can extend and be used in parallel to any of the existing DSR methodologies such as [1], [2], [11]–[13]. (iii) It encourages putting more focus on evaluation of DSR research projects as a stand-alone research endeavor and argues that valuable knowledge can be gained regardless of the success or non-success of a DSR project. By advocating critical realism as an explicit

underlying philosophical assumption, this research also addresses the need expressed by [14] to explicate underlying beliefs and assumptions when doing DSR.

By focusing on providing a philosophical basis for stand-alone evaluation of DSR projects and artifacts, this paper differs from other researchers who relate CR to DSR by suggesting a separate methodology based on CR [15], and from researchers focusing on DSR only as a philosophy beneficial to combine with CR philosophy without mentioning any DSR method or suggesting how this should be done [4]–[6]. All these authors also leave emancipatory and critical elements of CR out of their scope, although they mention their importance.

This article is structured accordingly: after the introduction, related work on design science is presented and CR is further explained. Then design science evaluation is discussed followed by the re-visitation of the two illustrative e-government cases. The article ends with a discussion and conclusions and suggestions for future research.

2. Theoretical Background

2.1 Design Science Knowledge

Design science is the science of the artificial and has its roots in engineering and science of the artificial [7]. Two of the most important works that has defined design science research are [1] and [2]. The aim of design science as described by these authors is to change reality and make it better through the design of Information Technology (IT) artifacts such as constructs, models, methods, and instantiations [1]. Another important outcome from design science is knowledge, such as design theories [16] and abstractions that can be used in designing and implementing IS initiatives [15].

More recently, a research stream that has a broader focus and addresses IS artifacts as socio-technical systems instead of only considering IT artifacts has emerged [15]. Socio-technical DSR differs from traditional DSR since it is more firmly embedded in the context where it is developed [15]. Although it can be argued that socio-technical DSR differs from traditional IS, they share the focus of knowing through making and the focus on change. The research processes suggested for both are also similar, although the focus on artifact output (e.g. [1], [2]) or theory output (e.g. [11], [12]) differs. DSR is thereby regarded as relying on an iterative process that contains three main parts: defining or finding a problem to solve, designing an IT/IS artifact that solve the problem, and evaluation to prove that the

artifact did in fact improve status quo (e.g. [7], [1], [2], [11], [13], [17]). The evaluation should be judged by the utility of the artifact [1], [2].

The need to produce knowledge (i.e. design theories) during the design and evaluation activities is also an important part of design science (e.g. [11], [12], [15]). The knowledge production of DSR has been articulated as knowledge of how to achieve desired ends or prescriptive knowledge [1], [16], [18], [19]. In [19] Gregor compares design theories with recipes and states that following a provisioned recipe will cause an artifact of a certain type to come into existence.

In [20] the prescriptive knowledge sought in design science is considered problematic from the perspective of management studies. The authors argue that design science is beneficial and relevant for management research “if it stands for investigating the complex nature of a design process in which ‘designers’ creatively use resources and engage with evolutionary patterns in order to reshape environmental constraints and create novel artifacts” [20, p. 183]. When design science is applied as prescriptive research, different from explanatory social research, the application of design science for management is considered very limited and not as fruitful for generating understanding [20].

In human-computer interaction, [21] also see weaknesses in prescriptive knowledge especially in the form of methods, and propose meals and ingredients rather than recipes as an analogy to argue for their standpoint. Applying this analogy on DSR would imply that apart from prescriptive design knowledge (the recipe), what has made up the recipe—the ingredients (e.g. the people working in the project, the context, people touched by the artifact, external factors affecting it, etc.) and how it was cooked (the process leading to the developed artifact and choices made)—needs to be understood to make the recipe understandable and thereby usable. The views expressed by [20], [21] thereby—to some extent—oppose the work by [19] on theory in IS.

2.2 Critical Realism

Philosophical assumptions are the basics for scientific paradigms, which are important because they legitimize how knowledge may be created and what constitutes valid knowledge [22]. Critical realism as a philosophical assumption was introduced by [8] and can be seen as an alternative to the most prominent research paradigms in information systems, i.e. of positivism and interpretivism [23].

According to [24], three aspects can distinguish a research paradigm: ontology, epistemology and methodology. It can be argued that a fourth aspect, i.e. axiology, should also be considered in research paradigms that concern inquiry [25]. Axiology concerns the nature of value and the value question of what is intrinsically worthwhile [25]. Regarding these four aspects, ontology is the most important concept in critical realism and also the importance of separating ontology from epistemology. Critical realism ontology argues that a real world exists, apart from our human experience and knowledge, i.e. realism ontology. Opposed to positivism, critical realism, however, states that human understanding and knowledge of the real world are socially constructed and fallible, thereby adopting a more interpretive epistemology [26]. Critical realism also contains an emancipatory axiology often associated with a critical research paradigm, e.g. [27]. According to [23], critical realism “offers exciting prospects in shifting attention to the real problems that we face and their underlying causes, and away from a focus on data and methods of analysis” [23, p. 795]. In regards to methodology, critical realism favors using a variety of methods to identify underlying mechanisms of observed events [26].

The basic theoretical building blocks for critical realism are objects (entities) and the external and visible behavior of people, systems, and things as they occur, or as they have happened. Another key aspect in critical realism is retrodiction [28], and the question asked is, “What must be true in order to make this event possible?” A closely related concept is abduction, which advocates the observation of phenomena through theory to derive explanatory theory about the phenomena [26]. Critical realism is thereby “a powerful tool in understanding the interplay of structure and agency in design activity dependence, and in theorizing generative mechanisms, well suited to theorize aspirations to a more humane and equal society” [5, p. 606].

Another important part of critical realism is the notion of different domains in the world that is the real, the actual and the empirical [8], and the accompanying levels, i.e. mechanisms, event and experiences [29]. Mechanisms cause events in the actual domains, which in turn become experiences in the empirical domain [30]. Outcomes thereby arise from the multifaceted relationship between powers, structures, and predispositions that create generative mechanisms [31]. As humans we can experience outcomes in the empirical domain, whereas the event is actually happening in the real domain and is caused by mechanisms operating in the actual domain [26]. Since human experience is limited to outcomes

located in the empirical domain, they are always open to different interpretations. Thereby, observations are always fallible, the understanding of social situations remains incomplete, and there can be no definite criteria to judge the truth of explanations [26]. In order to find the best or most suitable approximation of truth, critical realism trusts the researcher to accumulate data that facilitate differentiation among alternative explanations, and the research community to debate them thoroughly [26]. Critical realism differs in this aspect from established paradigms, for example positivism, where domains and layers are flattened, implying that a single empirical experience gives direct access to outcomes, including those that certainly arise in the real and actual domains, and in the events and mechanisms layers [28].

CR thereby offers a perspective upon which to base reflective and critical analysis. The methodological support in CR is, however, considered limited and needs more investigation. Examples of research that suggests methodological guidelines for doing CR research in IS are [32] for data analysis and [33] for case study research. Also, [34] discusses implications of CR in mixed methods research. These authors do not, however, address CR in DS.

3. Considering Evaluation in DSR

Several authors have proposed methods for evaluating DSR artifacts (e.g. [35]; see also [36] for a comprehensive review of existing artifact evaluation methods). The scientific evaluation of DSR should include both the utility of the designed artifact, and the quality of the design science knowledge outcomes [35]. The evaluation of knowledge outcomes should, according to [35], provide evidence that the theory is correct, i.e. leads to development of artifacts that solve problems or make an improvement. In regards to evaluation of the artifact as an output, [9, p. 85] writes, “A design artifact is complete and effective when it satisfies the requirements and constraints of the problem it was meant to solve,” this is articulated as the utility of the artifact. Thereby the most common evaluation type is a means-end-oriented evaluation [37], mostly interested in how effectively the artifact helps achieve the given goals or ends [37]. This is also the type suggested by [1], [2] and also [11], [13], [17] who state that the outcome of the DSR evaluation is a measure of how well the artifact solves the defined problem. The means-end-oriented types of DSR also have an epistemological standpoint of positivism [3] in common, although both qualitative and quantitative research methodologies are allowed to evaluate the utility of the artifact.

The claim that most DSR research is positivistic is further supported by, for example [22], who see an implicit positivistic paradigm in both [38] and [1]. Consistent with a positivistic view, the preferred methods for evaluation in design science is, according to [1], quantitative methods such as computational and mathematical methods, although empirical methods are also allowed.

Evaluation in general has also had a strong emphasis on positivistic measures of success [39]. Current evaluation outcomes are thereby considered restricted and their value less, due to the narrow ontological and epistemological foundations and because they only assess outcomes of interventions in the experience layer and empirical domain [39], [40]. Positivistic evaluations can thereby only provide a measure for comparing achievement against pre-determined objectives (i.e. means-end) but fail to define the causal influences (mechanisms) that can generate the observed outcomes. Insights into plausible explanations of mechanisms and relationships behind an observed event occurring within an intervention such the implementation of a DSR artifact in a public organization remain implicit and unknown. Following the argumentation made by [39], [40] utility as a measure for success in DSR evaluation is considered limited. One example is that utility, as often used in DSR evaluation, only considers the capacity or usefulness [1] of the object and actively disregards the ability of the users and the organizational context as a “why” question to be solved by behavioral science [1]. Although the necessity of considering more aspects than the capacity of the artifact in isolation has been acknowledged [11], they remain uncommon in DSR.

Utility in relation to DSR artifacts also constantly changes. Ten years ago, it was important to keep communication interfaces compressed and small, and even XML was considered too chatty due to the limited bandwidth; today, we easily stream motion pictures. When an evaluation is carried out also impacts the utility, for example, ex-ante (measuring the potential of an artifact) or ex-post (measuring the actual benefits). More recently, the need to address the limitations in the concept of utility has also been acknowledged (e.g. [3], [35], [41]). For instance, [41] suggests a fitness-utility model for evaluation in DSR; to accommodate utility in the future, the guiding philosophical assumption is, however, still positivistic and does not address the issues of extracting knowledge from failures or include values that fundamentally differ from utility.

Basing the generation of knowledge only on success is, however, not the only way, and sometimes not even the most preferable one [42], [43]. In line

with this traditional, evaluation methods have also been criticized by [42] since they only consider intended or expected outcomes and fail to detect outcomes or consequences that were not included in the logical framework or research hypothesis. In [42] risks are also identified when ignoring negative consequences that can affect the weak and vulnerable groups. Research programs thereby risk failing on equity objectives, allowing richer and more powerful groups to enjoy a disproportionate share of program benefits, while resulting in serious negative consequences for some groups—usually the poorest and most vulnerable [42]. Current IS design science discourse for socio-technical design projects do not, however, encourage, or even guide, researchers to prevent their artifacts from triggering negative consequences for society [3].

Relating the argumentation made by [42], [43] to DSR would imply that the focus on defining a DSR problem that has a possible solution limits the scope of possible problems and also puts unsolvable problems out of scope for DSR. There is also a risk with the means-end-oriented evaluation that problems are simplified to make it possible to solve them, since the value of the research is defined as the solving of the specified problem. One example would be to limit the number of stakeholder perspectives since they often differ and thereby complicate the problem definition. This in turn could lead to excluding stakeholder groups with less influence, while letting strong stakeholder groups define the problem. There is also a risk of omitting the publication of design science failures or at least trying to hide such failures by only evaluating specific features that are expected to be successful. Support for the argumentation is found in [3], who state that current evaluation methods in DSR, including [1], [2], [16], [35], can lead to limited perspectives and misleading ex-ante and ex-post evaluations of potential and real impacts of designed artifacts. Furthermore, [13] argues that innovation potential can be lost when the focus is only on problem-solving of pre-specified problems in DSR. It seems obvious that the process of knowing through making can be hindered by the need to prove the utility of the result, i.e. the developed artifacts. What if an artifact cannot be proven to deliver enough utility? Should the research process, and the knowing that was produced during the making of the artifact also be judged as not good?

The goal of utility in DSR can also be related to how DSR research is evaluated by the research community, by using the notion of falsification as defined by Popper [44] as the demarcation line between science and pseudoscience. How do you judge whether an artifact provides better utility, or

functions better than any other artifact? According to [35] one possibility is to compare to other artifacts with a similar purpose. To do this in a real setting is, however, challenging. For instance, how should you define which artifacts to compare? How do you gain access to them? How do you eliminate variables such as marketing or country specific circumstances? Even if you manage to make a comparison, how do you define what and how to compare? It is also a risk that aspects for comparison are chosen based on the intention to prove the utility of the designed artifact.

It is thereby difficult to provide a convincing argument or “proof,” while at the same time the possibility of falsification [44] is hampered since it is equally difficult to prove that the artifact does not provide utility, using the same arguments. The opportunity to disregard the outcome of design science research due to the lack of utility is thereby always present because objectively proving the utility of an artifact without doubt is not possible. Instead of providing universal truth, [45] states that the goal of evaluation is to identify what works for whom in what circumstances. To guide this work, an evaluation framework based on CR is presented where context and mechanisms lead to outcomes.

This can be related to the food analogy where the meals and ingredients are part of what is behind (the mechanisms) of any recipe. Transferring the argumentation of [39], [40] and [45] to DSR, it can be argued that CR-guided DSR evaluation retrospectives provide the tool to generate understanding from DSR projects as advocated by [20], [21], and also addresses the need for meals and ingredients, rather than recipes as expressed by [21].

4. Retrospective

To illustrate the value of a DSR retrospective, a re-visit to two DSR projects in an e-government context is made. Due to space limitation, a complete CR retrodution where underlying mechanisms are identified and elaborated to explain the findings was not, however, possible; but still, the re-visitation is considered to provide a foundation and motivation to the argumentation made in the paper since it visualizes the need to extend what is regarded as valuable in DSR.

4.1. The Open Social E-service Project

The first case is a three-year e-government project funded by the Swedish Governmental Agency for Innovation Systems, named Vinnova. Other publications within the project include [46]–[51]. The

project, named “Open Social e-Services,” took place between 2009 and 2012 with an extension of 6 months in 2012. The project involved representatives from the public sector, i.e. from one municipality, and IT vendors, as well as academics from the disciplines of information systems, social work, and law. The purpose of the project was to develop open social e-services to improve citizen services and decrease authoritative barriers to assisted living services. The result was a number of new e-services, i.e. artifacts for assisted living, such as emergency help telephones, part-time successors and companions. Citizens apply for services online, and if they meet predefined eligibility criteria, they receive immediate approval and can start using the service. Integration towards existing administrative systems as well as service providers ensured a digital and automated application process.

Several types of evaluations were conducted during the project lifetime; examples are process verification tests, usability tests, and benefit analysis. The results of the evaluations show that most of the goals set for the project are being met [48]. Examples of measures that were evaluated and met are that the administrative work time can be reduced by 85% and the lead-time from application submission to decision received was reduced from seven working days to four minutes [49]. Business models for ensuring the afterlife of the artifacts were also developed.

4.2. Revisiting the Open Social E-service Project

A re-visitation of the project in 2014 showed that no other municipality had started to use the service, and the Swedish municipality where the project took place did not yet have well-functioning routines for handling the open social e-services; instead, a number of problems of getting the solution into use were still present to a large extent. At present (2016), the services have been taken down from the web page of the municipality and is thereby no longer in use. The main reason is said to be because no cases were submitted through the e-services. One explanation to why the service was not used is that the target group of elderly citizens did not have the ability to use the e-service since it required the user to, for instance, have an e-legitimation.

The result of the project has, however, been successfully used for educational purposes by several researchers participating in the project. Examples include; teaching business process modeling, e-government service design and examples of how IS artifacts can reduce time consuming tasks by automation. As a teaching resource the results from the project have thereby been highly valued.

Knowledge gained from the project include for instance the following: (i) the knowledge on how to legally automate positive decisions based on predefined criteria, (ii) an understanding of the time-consuming tasks by employees at the municipality, (iii) understanding that IS can remove repetitive and simple tasks and instead facilitate professionalism for case handlers within the social service department. (iv) The understanding that several of the developed services were similar and could be carried out in a similar fashion. Other benefits include negotiation power towards IT-vendors and improved competence of procurement of IT services. From the non-existent usage, the knowledge of considering all stakeholder groups was also gained as well as the importance of the users. Most of the envisioned benefits were not realized, although initial evaluations indicated that effectiveness and efficiency could be greatly improved by using the designed artifacts. The re-visitation does, however, also show that several other benefits (see examples above), not envisioned, have been attained. The values of these benefits are difficult to measure and were not part of the initial goals of the project, but they are still valuable. Regarding the project as a complete failure, therefore, provides an erroneous picture.

The re-visitation also puts focus on the problem space and problem definition in conjunction with evaluation. If we regard the problem space as the problem of improving the present case handling routine, and the artifact only as an isolated IT artifact, then we can evaluate the capacity of the IT artifact to solve the problem without considering the context and excluding the stakeholders and we arrive at a high utility. If we instead consider the problem as that of improving the case handling process from start to end and the artifact as an IS artifact, we also have to evaluate all aspects including context and stakeholders. The result of such an evaluation instead yields almost zero utility.

4.3. The Munizapp Project

The second case is a research project set up in 2010 and also funded by Vinnova. The goal of the project was to develop an m-government solution that enables citizen sourcing, and thereby facilitates collaboration between local governments and citizens. Other publications within the project include [52]–[56]. Two software development companies developed the artifact; one built a front-end app, and the other developed an e-platform and also integration towards case handling systems in municipalities. Eight municipalities in Sweden of different size and demography represented the user

side of the municipal systems in the project. Reference groups representing citizens as users were also part of the project. Within the project, a business model for the continuation of the solution into a commercial phase was developed. The project ended in June 2013, and the solution is now commercialized through one of the SME partners that participated in the project.

The solution was evaluated through experimental and testing evaluations. Controlled experiments, functional tests, and simulations were performed through iterative prototype testing by the project participants. Improvements were identified during the evaluations and fixed by the developing companies. In the evaluation of the final prototype version, thirty-five citizens tested the solution and then responded to a survey. Citizens demonstrated a positive attitude toward the solution. A number of citizens stated that they found the app intuitive and easy to use. Citizens also gave suggestions for improvements to the app. A theoretical evaluation of the solution [52] was performed using a citizen-sourcing framework. The findings from the theoretical evaluation of the solution show that the solution includes functionalities supporting open government and citizen sourcing [52].

4.4. Revisiting the Munizapp Project

A re-visitation in 2015 made through interviews with IT vendors, municipalities, and a web page survey showed that the app is offered by nine Swedish municipalities out of the total of 290 municipalities in Sweden. Five of these municipalities are using a case handling system with integration towards the Munizapp artifact. These five have received several hundreds of reports made by their citizens. None of these five municipalities, however, participated in the design of the solution. The owner of Munizapp considers the adoption rate very low, but admits that no marketing efforts have been made to sell the solution to Swedish municipalities, except by the only IT-vendor selling a case handling system integrated with the solution. The Munizapp app was found to be the best with regard to ease of use as well as functionality. It was also the only app providing two-way communication between municipalities and citizens.

The re-visitation of the project shows that artifacts can be considered as partly fulfilling the initial goals of the project. The artifact has been shown to enable citizen sourcing and collaboration, which is taking place but to a very limited extent. Several knowledge claims can, however, be made from the project. Examples include the need to

consider IT-vendors as an important stakeholder in citizen-sourcing, the difficulties in balancing government and citizens needs, and also the impossibility of designing for citizens as a homogenous group. Knowledge was also gained in risks associated with the solution, such as the exclusion of groups who do not own smartphones, privacy risks and information quality risks.

4.5 Summary

The re-visitation to the two e-government projects provides a scattered picture of the utility of the artifacts designed. The initial evaluations during testing and design show great potential for improvements both in regards to effectiveness and efficiency and also in regards to political goals of openness. Ex-post evaluations provide a different picture; most of the benefits envisioned at the beginning have not been realized at all by the developed artifacts, especially with regard to effectiveness and efficiency. It could, therefore, be argued that the artifacts have not addressed the identified problems.

With regard to design science evaluations, this picture would suggest that the utility of the artifacts could not be proved, at least not continuously. It could, therefore, be argued that the artifacts do not yield utility, which in turn could be regarded as a failed design science effort [1]. According to CR, non-occurrences of expected events also should, however, be of interest and can be learned from [26]. As shown above, values other than effectiveness and efficiency, or even other than problem solving, have been realized. It can also be argued that problem solving, or rather problem-understanding, has been improved by the design efforts—in other words, knowing through making has taken place. It can also be seen that change has happened, although not the changed envisioned.

5. Discussion

According to [39] and [40], CR should complement existing evaluation approaches in evaluations that are “multi-faceted, mixed mode empirical methodology embracing pluralistic perspectives and multiple stakeholders” [40, p. 3]. This can be related to the socio-technical view of DSR, which according to [15], differs from traditional DSR. The reasons are that effects can be more difficult to isolate and evaluate, and knowledge is more deeply embedded in the context [15]. This

also makes developing and evaluating the design knowledge a complex task [15].

Compared to previous research in DSR where the focus remains on problem-solving (although to a varied extent), this research argues for a wider focus in DSR that also includes other values such as problem-understanding and ethics. I also argue, in line with [57], that prescriptions are not the only—and perhaps not a sufficient theoretical outcome of socio-technical design science research. Furthermore, it can be argued that the problem can never be known beforehand, but is instead iteratively understood as the development of an artifact unfolds, similar to the reasoning by [11], [12]. By designing the problem beforehand and then performing evaluations that match the initial problem understanding, knowledge is reduced to recipes rather than meals and ingredients. To avoid focusing only on problem-solving [13] suggests an agile design science methodology but does not include retrospectives that can be considered a very important part of any agile methodology to continuously, and after each project, get feedback, and learn from the experiences of different participants.

To address the above-mentioned weaknesses in DSR while still acknowledging the value of current DSR methods is to perform CR-guided retrospective evaluation of both artifacts and DSR projects. This research thereby extends and complements the agile DSR method suggested by [13]. The CR-guided retrospective is especially considered beneficial for socio-technical DSR endeavors due to the arguments put forth by, e.g. [15]. The CR-guided retrospectives can be carried out in parallel and/or after the project is complete in order to both expand knowledge creation opportunities and to identify mistakes made to improve the research process—i.e. to do better next time. Transferring the argumentation of [39], [40] and [45] to DSR, it can be argued that CR-guided DSR evaluation retrospectives provide the tool to generate understanding from DSR projects as advocated by [20], [21], and also addresses the need for meals and ingredients rather than recipes as expressed by [21]. Furthermore, the use of CR theory other than prescriptive theory can be generated, which would make DSR more applicable to management research as stated by [20].

CR regards non-occurrences as interesting as expected outcomes of events [26]. Outcomes of DSR projects that turn out to be unusable or that are used in a fashion that differs completely from what was intended at the start, are thereby considered as important for knowledge generation as successful outcomes. Other researchers in DSR (e.g. [11], [12]) have also brought up the need to investigate

unintentional effects of DSR artifacts, but the wording implies that they are less important than the main (intended) effects. Following CR, this article argues that the understanding of the problem is as important as solving it in regards to learning and knowledge production. Therefore, knowledge can be produced regardless of the utility of the artifact.

The idea of adopting a CR philosophy when doing design science research is not completely new. In regards to the ideas put forward by [4]–[6]—that CR and DSR have similarities and can benefit each other—I agree. This paper does, however, differ in scope and area since [4]–[6] focus on DSR and CR on a philosophical level and do not mention DSR methods. Furthermore, the authors [4]–[6] do not consider evaluation in DSR, and the focus is mainly on organizational identity and management, although the critical and emancipatory element in CR is pointed out. In [15] as well as in [4]–[6] it is noted that CR contains critical and emancipatory elements and that they are important, but the focus remains solidly on the management and profit (utility) perspective, while the critical issues are left for future exploration and development. This paper argues that a retrospective guided by CR provides the opportunity to focus a DSR project on goals separate from utility. Examples are; goals discussing ethical issues, facilitating participation, transparency and collaboration (i.e. open government issues) or understand and explain organizational phenomena. Thereby the benefits of the critical and emancipatory element of CR for DSR are highlighted.

The theory-generating DSR method suggested by [12] is close to introducing a retrospective in the form of an extended theory generation phase. The problem of mixing utility with knowledge generation, as pointed out in this research, is not as prominent in the theory-driven type of DSR, for instance [11], [12]. An explicit recognition that knowledge production might be possible, although the designed ensemble artifact did not address the class of problem identified is, however, still missing in [11]. The evaluation part of [12] is also still means-end-oriented. It is clear that the main objective of evaluation remains, “proving or disproving the design theory and/or the utility of the DSR artifacts” [35, p. 81]. Thereby, the CR-guided retrospective evaluation could potentially be applied to complement the action design method proposed by [11] and the theory-generating DSR method suggested by [12]. The knowledge creation process could thereby be enhanced to include an explanation of underlying mechanisms and the inclusion of, for instance, ethical perspectives. It would also allow failures in the artifact outcomes without diminishing the knowledge outcomes.

6. Conclusions

This article re-evaluates the outcomes of two different design science research projects and argues that there are several weaknesses in the present evaluation methods of design science research, especially regarding the outcome of design science research as socio-technical artifacts. As a solution, critical realism is proposed as a philosophical ground to improve design science evaluation in the form of a stand-alone retrospective.

The main contribution of this work is to complement existing DSR methods with a philosophically grounded retrospective that can be used to improve both the knowledge creation and the research process. The choice of CR as guiding philosophy for the retrospective is considered beneficial since it enhances the reach of DSR by providing a value-aware evaluation that is not limited to means-end-oriented research, but focuses on the knowing through making aspect of DSR. Thereby, CR guided DSR retrospectives recognize organizations’ obligations towards local and global societies, people outside the organizations (customers, citizens, etc.) as co-creators of value, and the duality of information systems to both co-create and co-destruct value simultaneously [58].

A possible weakness of critical realism is that as argued by [8], CR only operates by retrodution, which implies that events must have happened in order to be understood and explained. Thereby retrospective evaluations could require longitudinal studies that might be difficult to carry out. I would, however, argue that this should not be regarded as a weakness, but instead as support for an iterative and agile design science methodology. Therefore, short-term projects can also deliver designs that can be evaluated in a real setting and then improved in steps in conjunction with a better understanding of the problem space. Retrospectives can then be performed after each iterative cycle and not only after the completion of a step-wise DSR process. Thus, the suggested approach is more appropriate to DSR methods such as [11]–[13] compared to [1], [2].

7. Future Research

Future research should develop the ideas suggested in this paper in order to provide more detailed guidelines for how to conduct CR-guided retrospectives in DS. One possibility in this work would be to depart from the work by [33] on case study research guided by CR, and develop the

guidelines provided there to fit retrospective evaluation in DSR. Also, a deeper CR-guided retrospective on the cases presented here should be of interest in order to increase the understanding of how to do DSR in e-government as well as to develop design theories for e-government.

The use of agile methodologies in conjunction with critical realism also provides a prosperous way forward for creating improved design science guidelines—something that could also be undertaken in future research. Lastly, it would be interesting to further investigate inherent contradictions in IS artifacts in conjunction with value creation and collaboration, and adding critical realism as a perspective on existing research in this area.

8. References

- [1] A. R. Hevner, S. T. March, J. Park, and S. Ram, "Design Science in Information Systems Research," *MIS Q.*, vol. 28, no. 1, pp. 75–105, 2004.
- [2] K. Peffers, T. Tuunanen, M. Rothenberger, and S. Chatterjee, "A Design Science Research Methodology for Information Systems Research," *J. Manag. Inf. Syst.*, vol. 24, no. 3, pp. 45–77, 2007.
- [3] A. Drechsler, "A Postmodern Perspective on Socio-technical Design Science Research in Information Systems," in *New Horizons in Design Science: Broadening the Research Agenda.*, 2015, vol. 9073, pp. 152–167.
- [4] G. P. Hodgkinson and K. Starkey, "Not Simply Returning to the Same Answer Over and Over Again: Reframing Relevance," *Br. J. Manag.*, vol. 22, no. 3, pp. 355–369, 2011.
- [5] G. P. Hodgkinson and K. Starkey, "Extending the Foundations and Reach of Design Science: Further Reflections on the Role of Critical Realism," *Br. J. Manag.*, vol. 23, no. 4, pp. 605–610, 2012.
- [6] G. P. Hodgkinson, "Organizational Identity and Organizational Identification: A Critical Realist Design Science Perspective," *Gr. Organ. Manag.*, vol. 38, no. 1, pp. 145–157, 2013.
- [7] H. Simon, *The Sciences of the Artificial*. Cambridge, Massachusetts: MIT Press., 1996.
- [8] R. Bashkar, *A Realist Theory of Science*. Hemel Hempstead, 1978.
- [9] B. Stahl, "Responsible Research and Innovation in Information Systems," *Eur. J. Inf. Syst.*, vol. 21, no. 3, pp. 207–211, 2012.
- [10] G. Goldkuhl, "E-government Design Research: Towards the Policy-ingrained IT Artifact," *Gov. Inf. Q.*, vol. In press, 2016.
- [11] M. K. Sein, O. Henfridsson, S. Purao, M. Rossi, and L. Rikard, "Action Design Research," *MIS Q.*, vol. 30, no. 3, pp. 611–642, 2006.
- [12] R. Beck, S. Weber, and R. W. Gregory, "Theory-generating Design Science Research," *Inf. Syst. Front.*, vol. 15, no. 4, pp. 637–651, 2013.
- [13] K. A. Conboy, R. B. Gleasure, and E. A. Cullina, "Agile Design Science Research," in *New Horizons in Design Science: Broadening the Research Agenda*, 2015, vol. 9073, pp. 168–180.
- [14] S. Purao, "Truth or dare: The Ontology Question in Design Science Research," *J. Database Manag.*, vol. 24, no. 3, pp. 51–66, 2013.
- [15] S. Carlsson, S. Henningsson, S. Hrastinski, and C. Keller, "Socio-technical IS Design Science Research: Developing Design Theory for IS Integration Management," *Inf. Syst. E-bus. Manag.*, vol. 9, no. 1, pp. 109–131, Oct. 2011.
- [16] D. Jones and S. Gregor, "The Anatomy of a Design Theory," *J. Assoc. Inf. Syst.*, vol. 8, no. 5, pp. 312–335, 2007.
- [17] B. Kuechler and V. Vaishnavi, "On Theory Development in Design Science Research: Anatomy of a Research Project," *Eur. J. Inf. Syst.*, vol. 17, no. 5, pp. 489–504, 2008.
- [18] S. Gregor and A. R. Hevner, "Positioning and Presenting Design Science Research for Maximum Impact," *MIS Q.*, vol. 37, no. 2, pp. 337–355, 2013.
- [19] S. Gregor, "The Nature of Theory in Information Systems," *MIS Q.*, vol. 30, no. 3, pp. 611–642, 2006.
- [20] K. Pandza and R. Thorpe, "Management as Design, but what Kind of Design? An Appraisal of the Design Science Analogy for Management," *Br. J. Manag.*, vol. 21, no. 1, pp. 171–186, 2010.
- [21] A. Woolrych, K. Hornbæk, E. Frøkjær, and G. Cockton, "Ingredients and Meals Rather Than Recipes: A Proposal for Research That Does Not Treat Usability Evaluation Methods as Indivisible Wholes," *Int. J. Human-Comput. Interact.*, vol. 21, no. 10, pp. 940–970, 2011.
- [22] B. Niehaves and B. Stahl, "Criticality, Epistemology and Behaviour vs. Design-Information Systems Research Across Different Sets of Paradigms," in *European Conference on Information Systems*, 2006.
- [23] J. Mingers, A. Mutch, and L. Willcocks, "Critical realism in information systems research," *MIS Q.*, vol. 37, no. 3, pp. 795–802, 2013.
- [24] E. Guba and Y. Lincoln, "Competing Paradigms in Qualitative Research," in *Handbook on qualitative research*, N. Denzin and Y. Lincoln, Eds. Thousand Oaks, CA: Sage, 1994, pp. 105–118.
- [25] J. Heron and P. Reason, "A Participatory Inquiry Paradigm," *Qual. Inq.*, vol. 3, no. 3, pp. 274–294, 1997.
- [26] G. Easton, "Critical Realism in Case Study Research," *Ind. Mark. Manag.*, vol. 39, no. 1, pp. 118–128, 2010.

- [27] B. Stahl, "The Ethical Nature of Critical Research in Information Systems," *Inf. Syst. J.*, vol. 18, no. 2, pp. 137–163, 2008.
- [28] A. Sayer, *Method in Social Science*, 2nd ed. London: Routledge, 1992.
- [29] R. Groff, *Critical Realism, Post-positivism and the Possibility of Knowledge*. Abingdon UK: Routledge, 2004.
- [30] B. Danermark, M. Ekström, L. Jakobsen, and J. Karlsson, *Explaining Society: Critical Realism in the Social Sciences*. Abingdon UK: Routledge, 2002.
- [31] S. Fleetwood, "Powers and Tendencies Revisited," *J. Crit. Realis.*, vol. 10, no. 1, pp. 80–99, 2011.
- [32] B. Bygstad and B. E. Munkvold, "in Search of Mechanisms. Conducting a Critical Realist Data Analysis," in *International Conference on Information Systems*, 2011, pp. 1–15.
- [33] D. Wynn Jr. and C. K. C. Williams, "Principles for Conducting Critical Realist Case Study Research in Information Systems," *MIS Q.*, vol. 36, no. 3, pp. 787–810, 2012.
- [34] M. Zachariadis, S. Scott, and M. Barrett, "Methodological Implications of Critical Realism for Mixed-Methods Research," *MIS Q.*, vol. 37, no. 3, pp. 855–879, 2013.
- [35] J. R. Venable, J. Pries-Heje, and R. Baskerville, "FEDS: a Framework for Evaluation in Design Science Research," *Eur. J. Inf. Syst.*, vol. 25, no. 1, pp. 77–89, 2014.
- [36] K. Peffers, M. Rothenberger, T. Tuunanen, and R. Vaezi, "Design Science Research Evaluation," in *Design Science Research in Information Systems: Advances in Theory and Practice*, 2012.
- [37] J. Iivari, "A Paradigmatic Analysis of Information Systems as a Design Science," *Scandinavian J. Inf. Syst.*, vol. 19, no. 2, p. 5, 2007.
- [38] S. March and G. Smith, "Design and natural science research on information technology," *Decis. Support Syst.*, vol. 15, no. 4, pp. 251–266, 1995.
- [39] P. Jennings, "Hidden Treasures: The Contribution of Critical Realism to Evaluation," in *ESRC Research Capacity Bu*, 2013.
- [40] P. L. Jennings, "Critical realism: an alternative perspective on evaluation methodology," *University of Warwick*, 2015.
- [41] T. G. Gill and A. R. Hevner, "A Fitness-Utility Model for Design Science Research," *ACM Trans. Manag. Inf. Syst.*, vol. 4, no. 2, p. Article 5, 2013.
- [42] M. Bamberger, M. Tarsilla, and S. Hesse-Biber, "Why so many 'rigorous' evaluations fail to identify unintended consequences of development programs: How mixed methods can contribute," *Eval. Program Plann.*, vol. 55, pp. 155–162, 2016.
- [43] E. Bamberger, M., Vaessen, J., & Raimondo, *Dealing with Complexity in Development Evaluation: A Practical Guide*. Thousand Oaks, CA: Sage.: Sage, 2016.
- [44] M. Curd, J. Cover, and C. Pincock, *Philosophy of science: The central Issues*. W. W. Norton & Company, Inc, 1998.
- [45] R. Pawson and N. Tilley, *Realistic Evaluation*. London UK: Sage Publications Ltd., 1997.
- [46] E. Uppström, "The Promise of Public Value Co-Creation in Open Government: Designing IT Services for Open Government and Public Value Co-Creation," *Stockholm University*, 2014.
- [47] G. Juell-Skielse and P. Wohed, "Design of an Open Social e-Service for Assisted Living.," in *Electronic Government*, 2010, pp. 289–300.
- [48] P. W. T. Juell-Skielse, "Business Process Management for Open E-Services in Local Government Experience Report," in *Enterprise, Business-Process and Information*, vol. 81, 2011, pp. 1–15.
- [49] J. Helfrich and A. Helfrich, "Öppna Sociala eTjänster ett Vinnova Projekt 2009-2011 Slutrapport," 2011.
- [50] C. M. Lönn, E. Uppström, P. Wohed, and G. Juell-Skielse, "Configurable Process Models for the Swedish Public Sector," in *Advanced Information Systems Engineering: Proceedings of the 24th International Conference (CAISE)*, 2012, vol. 7328 LNCS, pp. 190–205.
- [51] G. Juell-Skielse, T. Mattsson, V. Persson, and E. Uppström, "Goal Model for Open Social e-Services," in *eChallenges e-2011*, 2011.
- [52] E. Uppström and C. M. Lönn, "The Promise of a Crowd," in *Proceedings of the 19th Americas Conference on Information Systems*, 2013.
- [53] C. M. Lönn and E. Uppström, "Government 2.0 Challenges in Swedish Public Sector.," in *eChallenges e-2013*, 2013.
- [54] C. M. Lönn and E. Uppström, "Process Management Challenges in Swedish Public Sector: A Bottom Up Initiative.," in *Electronic Government*, 2013, pp. 212–223.
- [55] E. Uppström and C. M. Lönn, "Core Aspects for Value Co-Creation in Public Sector," in *Proceedings of the 21th Americas Conference on Information Systems*, 2015.
- [56] C. M. Lönn, E. Uppström, and A. Nilsson, "Designing an M-Government Solution: Enabling Collaboration through Citizen Sourcing," in *European Conference on Information Systems*, 2016.
- [57] J. R. Venable, "The Role of Theory and Theorising in Design Science Research," in *Proceedings of DESRIST*, 2006, pp. 24–35.
- [58] T. Vartiainen and T. Tuunanen, "Value Co-Creation and Co-Destruction in an IS Artifact: Contradictions of Geocaching," in *Proceedings of the 49th Annual Hawaii International Conference on System Sciences*, 2016, pp. 1266–1275.