AIS SIGPRAG
Pre-ICIS Workshop 2018
"Practice-based Design
and Innovation of
Digital Artifacts"

A Pragmatic Approach for Identifying and Managing Design Science Research Goals and Evaluation Criteria

Alan R. Hevner[a], Nicolas Prat[b], Isabelle Comyn-Wattiau[b], and Jacky Akoka[c]

[a] Muma College of Business, University of South Florida, Tampa, FL, USA
ahevner@usf.edu

[b] ESSEC Business School, Cergy-Pontoise, France
{prat, wattiau}@essec.edu

[c] CEDRIC-CNAM & IMT-BS, Paris, France
jacky.akoka@cnam.fr

Abstract

The effectiveness of a Design Science Research (DSR) project is judged both by the fitness of the designed artifact as a solution in the application environment and by the level of new research contributions. An important and understudied challenge is how to translate DSR project research goals into discrete and measurable evaluation criteria for use in the DSR processes. This position paper proposes an inclusive approach for articulating DSR goals and then identifying project evaluation criteria for these goals. The goals are organized hierarchically as utilitarian goals, safety goals, interaction and communication goals, cognitive and aesthetic goals, innovation goals, and evolution goals. Goals in a DSR project are identified pragmatically by considering the components of the context coupled with the hierarchy of goals. Based on the identified goals, the associated evaluation criteria are determined and organized along the same hierarchy. These criteria measure the ability of the artifact to meet its goals in its context (immediate fitness). Moreover, our approach also supports the innovation and research contributions of the project. The apex of the goal hierarchy addresses the identification of criteria measuring the fitness for evolution of the designed artifact, to accommodate for changes in goals or context.

Keywords: Design Science Research, Goals, Fitness, Evolution, Innovation

1 Introduction

Design science research (DSR) projects seek to solve interesting problems via innovative artifacts that contribute new knowledge to the world. Among the key research challenges is the ability to understand and define the goals of the DSR project and how to determine *whether* and *to what extent* such goals are met. Multiple design cycles of *build* and *evaluate* activities strive to produce the design artifacts that provide both a satisfactory solution to the problem and the contribution to the research knowledge base (Simon 1996, Hevner et al. 2004). It is the rigorous evaluation of the DSR artifacts that determines if research goals are met.

The identification of DSR goals and their transformation into well-defined and measurable evaluation criteria remains an understudied topic. How do we evaluate the "goodness of fit" of a designed artifact as a solution in an application environment? How do we rank potential design candidates so as to select the best one for implementation as a solution? How do we capture and represent the new research contributions (e.g. design theories) of an innovative artifact in order to add them to the knowledge base? Such questions require the DSR project team to define the evaluation criteria for the project and design the rigorous methods for evaluating the design artifacts under these criteria. The rigor and credibility of a DSR project is determined by these evaluation decisions.

In the Information Systems field, several models of DSR evaluation methods have been proposed. Venable et al. (2016) present a framework of artifact evaluation comprising two dimensions – naturalistic vs. artificial and ex ante vs. ex post. Naturalistic evaluation is done in real world environments while artificial evaluation is performed in more controlled "laboratory" settings. Ex ante evaluation is formative, performed during the build and evaluate design cycles, while ex post is summative, performed after the artifact build is complete. The authors propose and plot evaluation strategies along the two-dimensional grid. Sonnenberg and vom Brocke (2012) propose a framework with four different stages of evaluation in a DSR project. Different evaluation methods are used in problem identification, solution design, artifact instantiation, and solution in use. Prat et al. (2015) perform an extensive literature review of DSR projects and produce a taxonomy of evaluation methods. A key insight from this study is the limited scope of evaluation criteria applied across the reported DSR projects. Most previous DSR studies have stated high-level evaluation criteria in such terms as artifact effectiveness, utility, validity, or other general concepts that are difficult to define and measure (Prat et al. 2015). We observe that none of the studies provide detailed guidance to determine the evaluation criteria for a given DSR project's goals.

Our objective in this paper is to present initial thinking toward the development of a pragmatic approach for determining the research goals of a DSR project and transforming those goals into a well-defined set of criteria for use in the formative and summative evaluations of the design artifacts. A DSR project aims at building an artifact that provides a solution to a problem while adding new knowledge to the world. A problem is characterized by goals identified in a context. Evaluation criteria not only assess to what extent the artifact meets the goals in the context (immediate fitness), but also to what extent it can accommodate changes in goals or context (fitness for evolution). Central to our approach is a hierarchy of DSR goals. Coupled with the context, this hierarchy guides the identification of DSR goals. From there, evaluation criteria are identified via use of the same goal hierarchy.

The rest of the paper is organized as follows. Section 2 presents the basic building blocks of our approach: problem context, goals, and evaluation criteria. We propose a six-level hierarchy of goals as a way to organize the new approach. Section 3 reviews past literature for examples of evaluation criteria that fit into the goal hierarchy. Section 4 presents a model of the DSR processes that incorporates our new pragmatic approach for applying the goal-driven evaluation criteria in DSR project design cycles. We complete the paper with our conclusions and future research directions in Section 5.

2 Context, Goals, and Evaluation Criteria

DSR projects solve an important problem in a defined application domain. A detailed understanding and description of this problem are essential to provide a convincing relevance for the research project. Two key components comprise a project's problem space – the application context (including the environment) and the goals for solution acceptance. In the remainder of this section, we establish a foundation for our proposal by providing a brief review of these basic terms in a DSR project.

2.1 Context

The environmental context provides a rich understanding of the problem space. Typically, design artifacts are situated in an organizational context, and their immediate environment is composed of people, organizations and technology (Hevner et al. 2004; Sjöström 2010). Important pragmatic questions are: What is the application domain and what is the current state of design knowledge (e.g. existing artifacts and design theories in use) for this domain? Who are the key stakeholders in the problem space who will impact and be impacted by the design solution?

The context includes information on the time and location of the research project. Design solutions reflect the point in time when they are designed. Available technologies, scientific theory bases, government regulations, national and international laws, and societal mores change over time. Therefore, a clear fixing of the time during which the research is conducted is essential to support both immediate fitness and fitness for evolution. Contextual aspects of location include the relevant geographic particulars such as rural vs. urban environments or developed vs. developing countries. As an example, one formal approach for describing the problem context is the PESTLE tool for business analysis (Cadle et al. 2010) with contextual categories of Political, Economic, Socio-cultural, Technological, Legal, and Environmental.

2.2 Hierarchy of DSR Goals

In a DSR project, the value of a design artifact is defined by the goodness of its fit as a solution for the problem or opportunity presented. In addition, the design must be novel in a way that adds new knowledge to a growing domain knowledge base. The research challenge is how to express these high-level DSR goals in appropriate ways that can be understood, measured, and communicated to the various stakeholders. Unfortunately, there is little research into the characteristics that provide good estimates of design fitness and novelty. This naturally presents a substantial obstacle to any DSR project that attempts to predict or estimate the success of a particular artifact. The largely unexplored forces driving artifact design fitness and novelty represent a considerable opportunity for research into how best to identify DSR research goals and transform them into defined evaluation criteria.

To guide the identification of DSR goals, we propose to define major categories of goals and organize these categories into a hierarchy. Not only will this hierarchy facilitate the identification of goals, it will also provide a structure for classifying evaluation criteria. The idea here draws on the balanced scorecard (Kaplan and Norton 1996), which defines four perspectives for identifying business objectives. These perspectives form a hierarchy, so that meeting the objectives at one level contributes to meeting those at the upper level (the top level of the hierarchy being the financial perspective). The perspectives are also used to organize the indicators measuring to what extent the objectives are met.

To define our hierarchy of goals for socio-technical system solutions, we adapt Maslow's (1943) hierarchy of human needs. Although this hierarchy has sometimes been criticized as stereotypical, it continues to be used frequently in research. The ultimate purpose of artifacts is to serve human needs. However, we need to adapt Maslow's hierarchy to take into account the specificities of DSR artifacts with their socio-technical natures. Maslow's initial hierarchy of needs contains the following levels: physiological, safety, belonging and love, esteem, and self-actualization. These needs form a hierarchy in the sense that a human being needs to fulfill the needs of one level to focus motivation on the needs of an upper level. The original hierarchy was refined into the following hierarchy (Maslow 1998) via applications and analyses: physiological needs, safety needs, social needs, esteem needs, cognitive needs, aesthetic needs, self-actualization, and self-transcendence.

In the field of information systems, Urwiler and Frolick (2008) use Maslow's hierarchy as a metaphor to gauge the maturity level of Information Technology (IT) in organizations. They propose the hierarchy: infrastructure and connectivity needs, stability and security needs, integrated information needs, competitive differentiation, and paradigm shifting. Moving up this hierarchy means moving from commodity IT to innovative IT.

In marketing, Srinivasan et al. (2012) define the total product design concept (TPDC) as consisting of three elements, namely functionality, aesthetics, and meaning; each of which arises from more elemental product characteristics. Although the TPDC is not directly related to Maslow's pyramid, it provides insights for its adaptation. Functionality refers to the utilitarian dimension of a product. It arises from the product's features and related benefits. Aesthetics is the product's sensorial characteristics and meaning refers to the associations of the product in the minds of its customers.

Thus, based on Maslow's pyramid and the above-mentioned papers, we define the following six-level hierarchy (Figure 1) of DSR goals: (1) utilitarian goals, (2) safety goals, (3) interaction and communication goals, (4) cognitive and aesthetic goals, (5) innovation goals, and (6) evolution goals. We note that physical needs are generally not relevant to DSR artifacts. Instead, utilitarian goals form the base of the hierarchy, corresponding to functionality in the TPDC mentioned above. Placing utilitarian goals at the base of the hierarchy implies that they should be fulfilled in order for upper-level goals to be considered. Interaction and communication goals (similar to the integrated information needs above) correspond to Maslow's social needs. Cognitive and aesthetic goals are placed at the same level (we do not establish a hierarchy between the "beauty" of a demonstration in mathematics and the beauty of a piece of art).

Innovation goals (contributing creative knowledge and artifacts) come next, corresponding to self-actualization in Maslow's hierarchy. Note that self-actualization is the highest level of Maslow's hierarchy of needs that is relevant in DSR (self-transcendence is not applicable). Placing innovation at this high level is also consistent with the hierarchy of Urwiler and Frolick (2008), which reflects increasingly innovative applications of IT in organizations. Finally, evolution (ability to manage change) goals constitute the upper level of the hierarchy. Although Maslow's original hierarchy of needs does not consider evolution, this concept appears in adaptations of the hierarchy. For example, Kiel (1999) proposes to represent the hierarchy of human needs as a triangle that is open at the top, reflecting the "boundlessness" of self-actualization: self-actualization is a continuously evolving process.

Figure 1. Hierarchy of DSR Goals

2.3 DSR Project Evaluation Criteria

Upon identifying the goals of a DSR project, the next step is to find evaluation criteria that are instantiated in the artifacts and subsequently measured in the evaluations. In the following section, we pay particular attention to the understudied goals of innovation and evolution (Gill and Hevner 2013). Consequently, one objective of this paper is to remediate this imbalance and investigate evaluation criteria for innovation and evolution more fully. To this aim, we start by a brief review of evaluation criteria for DSR artifacts, and then dig deeper into the past forms of evaluation for innovation and evolution from multiple fields. We end with an initial proposal for a partial listing of the six levels of evaluation criteria that can be used to measure goal achievement in a DSR project.

3 Evaluation Criteria for DSR Artifacts

The DSR literature contains a number of proposed approaches for the construction and evaluation of design artifacts and the reflection and guidance on resulting design theories (Winter 2008). Examples of evaluation criteria proposed in DSR projects are found in (March and Smith 1995; Hevner et al. 2004; Sonnenberg and vom Brocke 2012; Gregor and Hevner 2013; Baskerville et al. 2015). In particular, Prat et al. (2015) propose a taxonomy of evaluation methods for IS artifacts, with a special focus on evaluation criteria. This taxonomy is based on a literature review of design science papers, coupled with an in-depth content analysis of 121 DSR papers. The content analysis reveals a focus on a few assessed criteria being (in descending order) efficacy, usefulness, technical feasibility, accuracy, and performance. We find that the evaluation criteria addressed in these papers primarily address the lower levels of

DSR goals in our goal hierarchy, more specifically utilitarian goals with some discussion of safety and interaction/communication goals. In the next sections, we focus on the underexplored goal categories of innovation and evolution.

3.1 Innovation Evaluation Criteria

The innovation goals of a DSR project are especially challenging to identify and structure as evaluation criteria. Important guidance for understanding the DSR problem space and establishing innovation goals comes from Gregor and Hevner (2013; 2014) via the Knowledge Innovation Matrix (KIM). KIM (Figure 2) is structured on the most fundamental feature of innovation – new knowledge created and applied in some tangible form to achieve human goals. The tangible forms that innovations take include products, processes, and services. Taking knowledge and its application (to human needs) as the keys to understanding innovation means that many other labels and categorizations that only partly deal with the innovation space are encompassed: ideas, creativity, technological know-how, products, competencies, organizational learning, and exploration versus exploitation (March 1991; Tidd and Bessant 2013). That knowledge is the key feature in all of these terms is apparent on reflection. Companies do not value the innovative products they produce so much in themselves as the knowledge assets they embody. After all, they sell these products to consumers. What they value is the knowledge asset represented in the product – the intellectual property that may be worth protecting by patents, trade secrets, and/or copyrights – including the need it meets.

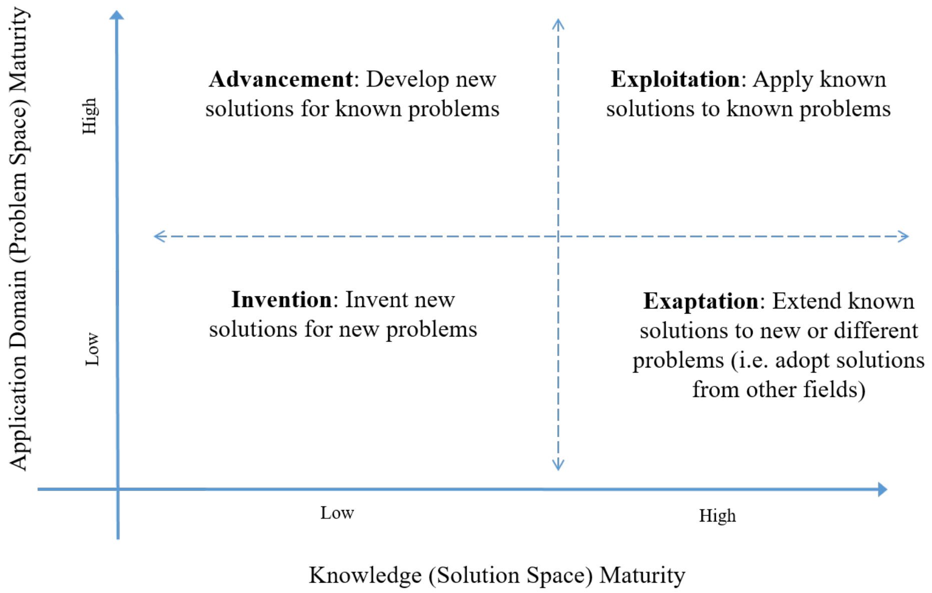

Figure 2. Knowledge Innovation Matrix (from Gregor and Hevner 2014)

A DSR project will position its innovation goals in one of the three research quadrants of invention, advancement, and exaptation. The selected quadrant provides guidance on the types of research contributions to be made and how these contributions will be communicated and evaluated. A recent editorial presents a detailed discussion of DSR contributions on finding the right balance between the design artifact and the design theories (Baskerville et al. 2018). The key observation is that DSR contributions form a continuum on at least two dimensions: from very novel artifacts to rigorous theory development and from early visions of technology impact to studies of technology impact on users, organizations and society. Thus, there can be multiple types of published contributions depending on the novelty of the artifact and the phase of the research project. Evaluation criteria for innovation will be the evidence of new artifacts, new theories, and the impacts of these novel contributions. We find several examples of proposed evaluation criteria for DSR innovation in Baskerville et al. (2015) with innovativeness and inventiveness and Gill and Hevner (2013) with novelty.

3.2 Evolution Evaluation Criteria

As a central premise of our research direction we contend that over time the evolutionary goals of design artifacts become far more interesting than the use fitness of a particular artifact at a static point in time. The validity of this premise is likely to depend on the problem space in which it is situated. For very static environments, for example, a particular artifact may exist for a very long time with high capability. In this situation, the immediate fitness of the artifact is a matter of considerable interest. In a highly dynamic environment, however, the artifact's potential to evolve needs to be given much greater weight. The benefits of encouraging diversity and openness of designs as insurance against major changes in the problem (goals or context) would be of paramount importance. Our belief is that such dynamism describes most environments facing designers of socio-technical systems today and forces such as globalization, energy conservation, social media, and advances in telecommunications will likely serve to increase such environmental turbulence.

Closely aligned to our approach, Gill and Hevner (2013) explore a design fitness and utility function of DSR artifacts, drawing on the fields of biology and economics. Fitness is not directly observable and is assessed with a utility function, based on a set of attributes of design artifacts, termed *design traits* (in essence, evaluation criteria). The authors identify and illustrate eight key design traits: usefulness, decomposability, malleability, openness (e.g., open source), embeddedness in a design system (e.g., the design ecosystem of Apple), novelty, interestingness, and elegance.

In the field of software engineering, Lehman (1996) identifies eight laws of software evolution: (I) continuing change, (II) increasing complexity, (III) self-regulation, (IV) conservation of organizational stability (invariant work rate), (V) conservation of familiarity of all involved with the software's goals, (VI) continuing growth, (VII) declining quality, and (VIII) feedback system. These eight laws are grounded empirically, i.e. resulting from observation. We find limited follow up in the software engineering literature to these early empirical results.

In the systems engineering field, Sauser et al. (2008; 2010) contend that there is a molecular level below General Systems Theory that is not fully discovered or defined. They call it *systomics*, the study of a system's DNA. Fundamental system characteristics in systomics are autonomy, belonging, connectivity, diversity, and emergence. In marketing, Srinivasan et al. (2012) propose the concept of "total product design" which contains elements of continuous evolution of the product.

In IS, Agarwal and Tiwana (2015) explore evolvable systems, understood as software-based information systems. They define the evolvability of a system as its capacity to efficiently serve new purposes and emerging possibilities. They argue that systems evolvability is an important and underexplored topic since most IS research generally focuses on the first use of a software system and the phases that precede it, not on evolution that takes place afterwards. They propose some proxies to assess systems evolvability: plasticity, boundary morphing (temporal change in the proportion of functionalities implemented within the system vs. those accessible via APIs), mortality, market durability, mutation, pleiotropy (production of more than one effect, e.g. genes having multiple phenotypic expressions), and interoperability.

3.3 Mapping of Evaluation Criteria to DSR Project Goals

In this position paper, our objective is not to produce an exhaustive list of all possible evaluation criteria for a DSR project to consider. In the Appendix, we present a proposed list of criteria that have been identified in the above literature review as a starting point for transforming DSR goals into specific evaluation criteria. For each evaluation criterion, Table A shows the definition of the criterion with the source or sources from which the definition was taken or adapted. In this section, we bridge the gap between the hierarchy of goals and evaluation criteria identified above, by using the hierarchy of goals to organize the sample of evaluation criteria, as illustrated in Figure 3. Beyond the criteria listed in Table A, we also use other criteria (security, privacy, research contribution, etc.) to illustrate the hierarchy.

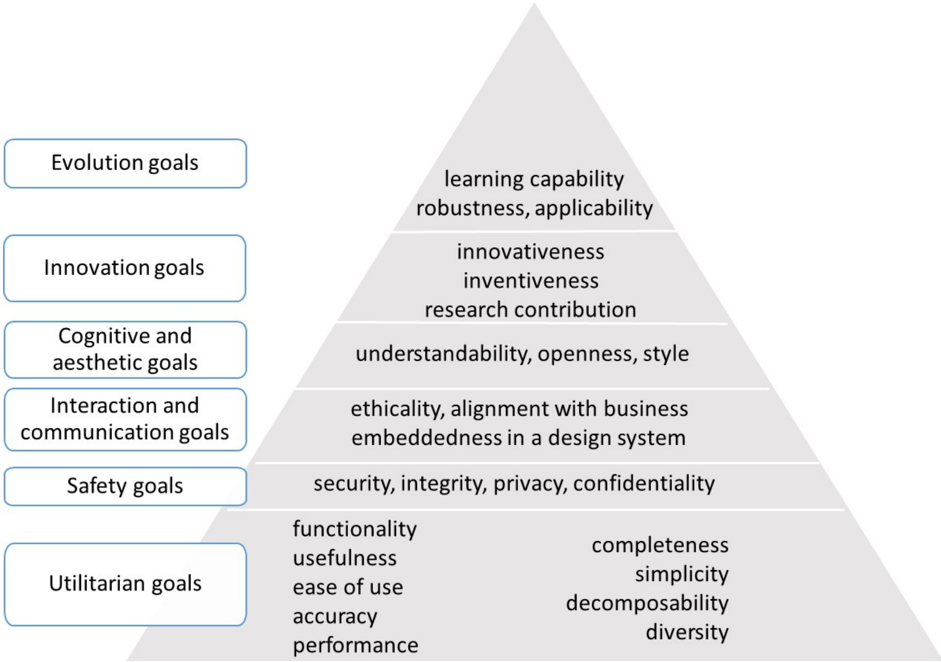

Figure 3. Hierarchy of DSR Goals and Evaluation Criteria

As mentioned above and as appears in Figure 3, many of the evaluation criteria appear at the bottom level of the hierarchy (utilitarian criteria). These criteria are inherently associated with the artifact: they qualify its behavior, e.g. functionality, and/or its structure, e.g. decomposability. As we move upwards in the hierarchy, subjectivity plays a greater role, the criteria being more related to the perception of individuals (e.g. style). At the upper level, innovation criteria measure the new knowledge added to the knowledge bases of the application domain and the degree to which the artifact can be generalized to new situations. At the top level, evolution criteria measure the extent to which the artifact is capable of accommodating changes in its context (e.g., malleability) and/or its goals (e.g. learning capability).

4 DSR Processes for Identifying and Managing Goals and Evaluation Criteria

The iterative DSR project process is composed of three key stages: 1) defining the problem, 2) building a solution, and 3) assessing DSR goals attainment. Figure 4 expands the framework via its interactions in a DSR project with process and information flows. Evaluation criteria link the three steps by providing a means to assess the fitness between problem components (context and goals) and designed components (structures and behaviors) of the solution (the DSR artifact). This framework provides an organization of the DSR project not previously seen in the literature and sheds new light on the richness of artifact construction and evaluation. In this section, we discuss the processes and information found in each of the three dimensions of the framework. Our focus is on the roles of DSR goals and corresponding evaluation criteria throughout the framework.

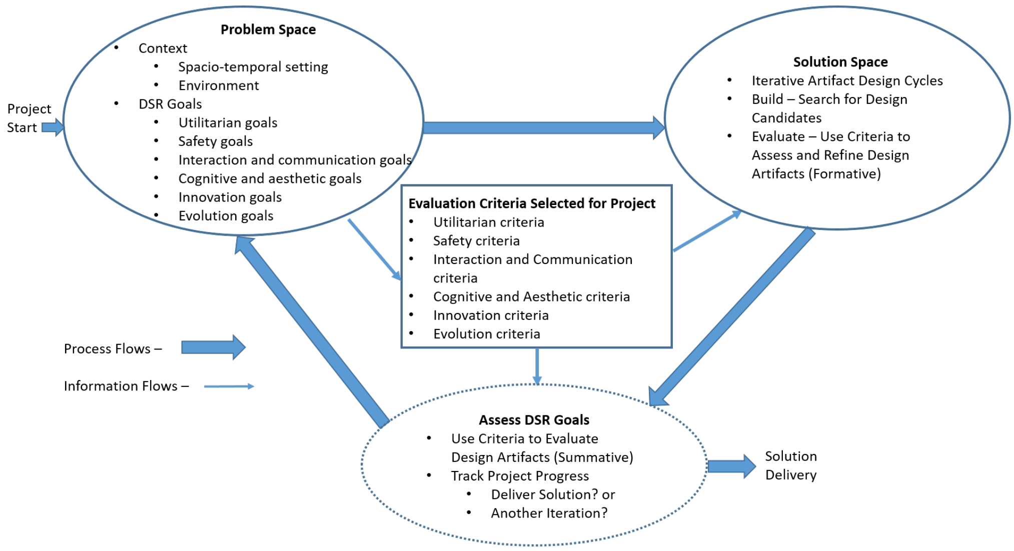

Figure 4. DSR Process Model with Goals and Evaluation Criteria

4.1 DSR Problem Space

DSR projects solve important problems in a defined application domain. A detailed understanding and description of the problem are essential to provide a convincing relevance of the research project. There are two key components that describe a project's problem – the environment (i.e. application context) and the DSR goals for solution acceptance.

DSR goals address the meaning and requirements for how well a design solution solves the problem in context. Once the goals for the problem are understood, we identify the evaluation criteria that are relevant and capture the essential requirements for a design solution. The explicit steps for identifying a good set of evaluation criteria are left as future research directions. As discussed in the previous section, any practical solution must include evaluation criteria that focus broadly on all six-levels of the goal hierarchy. Thus, evaluation criteria for satisfactory solutions will include a rich mix of measures from the categories of utility, safety, interaction/communication, cognitive/aesthetic, innovation and evolution. We note that a particular DSR project may not include criteria from all six levels but it would be important to consider all the levels for a full understanding of the problem. The description of these evaluation criteria provides a rigorous set of acceptance criteria for the evaluation of potential design solutions and establishes guidance for the design of both formative and summative evaluation methods (Venable et al. 2016).

4.2 DSR Solution

A DSR project creates a solution via build and evaluate sub-processes (Hevner et al. 2004). The project produces design artifacts in the form of constructs, models, methods, and instantiations. Artifacts must be complete representations possessing a structure (i.e. form) and a set of behaviors (i.e. functions). Information on context, goals, and evaluation criteria provides essential guidance to both the build and evaluate sub-processes. During construction of solution artifacts, search processes identify the best design candidates. Information on environmental context and evaluation criteria is used to guide a goal-driven search to maximize value while being constrained by the availability and feasibility of resources, including project time deadlines. Evaluation criteria particularly come into play in the design and execution of formative evaluations of the artifacts. Iterative build-evaluate cycles will refine the structure and activities of the design artifacts to best meet the goals.

4.3 Assessing DSR Goals

The third step develops the crucial bridge of summative evaluation between the problem and the solution. We look for compelling evidence from these evaluations of the innovative design artifacts to provide confirmation of the success or failure of the DSR project to satisfactorily solve the problem. We note that not all DSR projects have the opportunity to test the new design artifacts in realistic environments. In those cases, opportunities for summative evaluations in artificial environments should be considered (e.g. simulation, predictive analytic modeling) (Prat et al. 2015).

4.4 DSR Project Flows and Evaluation Criteria

Figure 4 illustrates the flows of process and information in a DSR project. The DSR process flow moves in an iterative cycle from Problem Space to Solution Space to Project Assessment. As the project progresses through multiple iterations of this cycle, a repository of project design knowledge accumulates rapidly. Emerging understanding of the problem produces a detailed description of the environment and a growing set of evaluation criteria. The Solution Space inputs this information and performs iterative artifact design cycles of build and evaluation sub-processes. In turn, the Solution Space produces design knowledge in the form of artifacts with well-defined structures and activities. Then the assessment step evaluates the artifact for goodness of fit to the problem environment and goals. A decision is made on whether the current artifact fit is satisfactory for solution delivery or if additional iterations of the overall DSR project cycle are required to grow needed design knowledge to better solve the problem.

4.5 Illustrative Example

To illustrate the proposed DSR process, consider a research context where a thesis supervisor assigns a literature review task to her PhD student. The student must identify all the machine learning papers regarding consumer behaviours in a specific application context. The supervisor suggests building a taxonomy of the different solutions. The goals hierarchy facilitates the supervisor's validation effort. Building the taxonomy forces the PhD student to study deeply the set of papers. This is clearly related to the utilitarian goal. Besides, the taxonomy is delivered with dimensions and characteristics which improve the understandability of the field (cognitive goal) and represent a research contribution (innovation goal). The PhD supervisor must verify that the taxonomy is robust and may survive during the thesis period (evolution goal). This illustrative example is a first step towards a more comprehensive validation, to be developed in our further research.

5 Conclusion and Further Research

The effectiveness of a Design Science Research (DSR) project is judged both by the fitness of the designed artifact as a solution in the application environment and by the level of new research contributions made to the appropriate knowledge bases. An important and under-studied challenge is how to translate DSR project research goals into discrete and measurable evaluation criteria for use in the DSR project processes. We propose an initial structured approach for articulating DSR goals and then identifying project evaluation criteria for these goals. The DSR goals are organized hierarchically: utilitarian goals, safety goals, interaction and communication goals, cognitive and aesthetic goals, innovation goals, and evolution goals. Goals in a DSR project are identified pragmatically, i.e. by considering the components of the context, coupled with the hierarchy of goals. Based on the identified goals, the associated evaluation criteria are determined and organized along the hierarchy of goals. These criteria measure fitness, i.e. the ability of the artifact to meet its goal. This goal-driven approach provides pragmatic insights for managing a DSR project to a successful conclusion with the achievement of measurable goals.

Future work will refine the hierarchy of DSR goals and associated criteria. First, to refine the hierarchy of goals, we will complement the top-down approach used in this paper (adaptation of a generic hierarchy of goals, namely Maslow's hierarchy) with a bottom-up approach (i.e., an approach starting from goals in actual DSR studies). Second, we will expand our literature survey and investigation into the top two levels of the goal hierarchy – innovation goals and evolution goals. As these are the least understood areas, we hope to propose new approaches for linking innovation and evolution goals to specific evaluation criteria. Recent thinking on the building of design theories for complex systems development will provide insights to our research directions (e.g. Hanseth and Lyytinen 2016; Demetis and Lee 2016). Third, we will provide researchers with more intuitive and detailed guidelines facilitating the identification and management of goals and evaluation criteria. Finally, we will explore different usages of the framework, including how a change in the solution impacts its fitness to the problem. This corresponds to tackling the variability of the artifacts. Although this variability concept has been originally developed in software engineering (software product lines), there is research value for its application to DSR project fitness evaluation.

References

Agarwal, R., and Tiwana, A. 2015. "Editorial – Evolvable systems: Through the Looking Glass of IS," *Information Systems Research* (26:3), pp. 473-479.

Baskerville, R., Kaul, M., and Storey, V. 2015. "Genres of Inquiry in Design-Science Research: Justification and Evaluation of Knowledge Production," *MIS Quarterly* (39:3), pp. 541-564.

Baskerville, R., Baiyere, A., Gregor, S., Hevner, A., and Rossi, M. 2018. "Design Science Research Contributions: Finding a Balance between Artifact and Theory," *Journal of the AIS* 19(5), 308-327.

Bondi, A. 2000. "Characteristics of Scalability and their Impact on Performance," in Proceedings of the 2nd International Workshop on Software and Performance, Ottawa, Canada, ACM, pp. 195-203.

Buchanan, R. 1992. "Wicked Problems in Design Thinking," *Design Issues* (8:2), pp. 5-21.

Cadle, J., Paul, D. and Turner, P. 2010. *Business Analysis Techniques – 72 Essential Tools for Success*. British Computer Society, Swindon, UK.

Checkland, P. and Scholes, J. 1990. *Soft Systems Methodology in Action*. Wiley, Chichester, UK.

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

Demetis, D. and Lee, A. 2016. "Crafting Theory to Satisfy the Requirements of System Science," *Information and Organization* 26, pp. 116-126.

Gill, T. and Hevner, A. 2013. "A Fitness-Utility Model for Design Science Research," *ACM Transactions on Management Information Systems* (4:2), pp. 1-24.

Glinz, M. 2007. "On Non-Functional Requirements," in Proceedings of the 15th IEEE International Requirements Engineering Conference, A. Sutcliffe and P. Jalote (eds.), New Delhi, India, pp. 21-26.

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

Gregor, S. and Hevner, A. 2014. "The Knowledge Innovation Matrix (KIM): A Clarifying Lens for Innovation," *Informing Science: The International Journal of an Emerging Transdiscipline*, 17, 217-239.

Hanseth, O. and Lyytinen, K. 2016. "Design Theory for Dynamic Complexity in Information Infrastructures: The Case of Building Internet," In: Willcocks L., Sauer C., Lacity M. (eds) *Enacting Research Methods in Information Systems*. Palgrave Macmillan.

Henderson, J. and Venkatraman, N. 1993. "Strategic Alignment: Leveraging Information Technology for Transforming Organizations," *IBM Systems Journal* (32:1), pp. 472-484.

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

ISO/IEC/IEEE: Systems and Software Engineering – Vocabulary. ISO/IEC/IEEE 24765:2010(E).

Kaplan, R. and Norton, D. 1996. *The Balanced Scorecard: Translating Strategy into Action*, Harvard Business School Press, Boston.

Kiel, J. 1999. "Reshaping Maslow's Hierarchy of Needs to Reflect Today's Educational and Managerial Philosophies", *Journal of Instructional Psychology* (26:3), p. 167.

Lehman, M. 1996. "Laws of software evolution revisited," in Proceedings of the 5th European Workshop on Software Process Technology, C. Montangero (ed.), Springer Verlag, pp. 108-124.

March, J. 1991. "Exploration and Exploitation in Organizational Learning," *Organization Science*, 2(1), 71-87.

March, S. and Smith, G. 1995. "Design and Natural Science Research on Information Technology," *Decision Support Systems* (15:4), pp. 251-266.

Martin, R. 2009. *The Opposable Mind: Winning Through Integrative Thinking*. Harvard Business School Press, Boston.

Maslow, A. 1943. "A Theory of Human Motivation," *Psychological Review* (50:4), pp. 370-396.

Maslow, A. 1998. *Toward a Psychology of Being*, 3rd edition, Wiley and Sons, New-York.

Prat, N., Comyn-Wattiau, I. and Akoka, J. 2015. "A Taxonomy of Evaluation Methods for Information Systems Artifacts," *Journal of Management Information Systems* (32:3), pp. 229-267.

Sauser, B., Blair, C. and Boardman, J. 2008. "Systomics: The Ultimacy of Design," in Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, Singapore, pp. 24-29.

Sauser, B., Boardman, J. and Verma, D. 2010. "Systomics: Toward a Biology of System of Systems," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* (40:4), pp. 803-814.

Simon, H. 1996. *The Sciences of the Artificial*, 3rd edition, The MIT Press, Cambridge, MA.

Sjöström, J. 2010. Designing Information Systems – a Pragmatic Account. PhD thesis, Uppsala University.

Sonnenberg, C. and vom Brocke, J. 2012. "Evaluations in the Science of the Artificial – Reconsidering the Build-Evaluate Pattern in Design Science Research," in Proceedings of the 7th International Conference on Design Science Research in Information Systems and Technology, K. Peffers, M. Rothenberger, T. Tuunanen, and R. Vaezi (eds.), Las Vegas, NV, Springer Verlag, pp. 381-397.

Srinivasan, R., Lilien, G., Rangaswamy, A., Pingitore, G. and Seldin, D. 2012. "The Total Product Design Concept and an Application to the Auto Market," *Journal of Product Innovation Management* (29:S1), pp. 3-20.

Tidd, J. and Bessant, J. 2013. *Managing Innovation Integrating Technological, Market and Organizational Change* (5th edn.). Chichester UK: John Wiley & Sons.

Urwiler, R. and Frolick, M. 2008. "The IT Value Hierarchy: Using Maslow's Hierarchy of Needs as a Metaphor for Gauging the Maturity Level of Information Technology Use within Competitive Organizations," *Information Systems Management* (25:1), pp.83-88.

Venable, J., Pries-Heje, J. and Baskerville, R. 2016. "FEDS: a Framework for Evaluation in Design Science Research," *European Journal of Information Systems* (25:1), pp. 77-89.

Williams, K., Chatterjee, S. and Rossi, M. 2008. Design of Emerging Digital Services: a Taxonomy. *European Journal of Information Systems* (17:5), pp. 505-517.

Winter, R. 2008. Design Science Research in Europe. *European Journal of Information Systems* (17:5), pp. 470-475.

Appendix

Table A. List of Evaluation Criteria (Excerpt)

Source	Design trait	Definition
Prat et al. (2015)	Accuracy	The degree of agreement between outputs of the artifact and the expected outputs (ISO/IEC/IEEE)
	Adaptability	The ease with which the artifact can work in environments other than those for which it was specifically designed, or change according to evolutions in environment (ISO/IEC/IEEE)
	Alignment with business	The congruence of the artifact with the organization and its strategy (Henderson and Venkatraman 1993)
	Completeness	The degree to which the artifact contains all necessary elements and relationships between elements
	Consistency	The degree of uniformity, standardization, and freedom from contradiction among the elements of the artifact (ISO/IEC/IEEE)
	Ease of use	The degree to which the use of the artifact is effort free (Davis 1989)
	Efficacy	The degree to which the artifact achieves its goal considered narrowly, without addressing situational concerns (Checkland and Scholes 1990; Venable et al. 2016)
	Ethicality	The degree to which the artifact complies with ethics
	Functionality	The capability of the artifact to provide functions which meet stated and implied needs when the artifact is used under specified conditions (ISO/IEC/IEEE)
	Learning capability	The ability of the artifact to learn from experience
	Modifiability	The ease with which the artifact can be changed without introducing defects (ISO/IEC/IEEE)
	Performance	The degree to which the artifact accomplishes its function with given constraints of resources (ISO/IEC/IEEE; Glinz 2007)
	Robustness	The ability of the artifact to handle invalid inputs or stressful environmental conditions (ISO/IEC/IEEE)
	Scalability	The ability of the artifact to handle growing amounts of work in a graceful manner, or to be readily enlarged (Williams et al. 2008; Bondi 2000)
	Simplicity	The degree to which the artifact contains the minimal number of elements and relationships between elements (ISO/IEC/IEEE)
	Style	The elegance with which the artifact has been built (March and Smith 1995; Hevner et al. 2004)
	Understandability	The degree to which the artifact can be comprehended, both at a global level and at the detailed level of the elements and relationships inside the artifact (ISO/IEC/IEEE)
	Usefulness	The degree to which the artifact positively impacts the

		task performance of individuals (Davis 1989)
Baskerville et al. (2015)	Applicability	The ease with which the artifact can achieve other goals than the one for which it was specifically designed, or change according to evolutions in goal
	Innovativeness	Requires "inventive leaps of generative reasoning" which facilitates trial and error, crucial to creative resolution" (Martin 2009)
	Inventiveness	"The inventiveness of the designer lies in a natural or cultivated and artful ability to return to those placements and apply them to a new situation, discovering aspects of the situation that affect the final design" (Buchanan 1992)
Gill and Hevner (2013)	Decomposability	The degree to which the artifact can be disassembled into nearly independent sub-artifacts
	Embeddedness in a design system	The degree to which the artifact is the product of a sustainable design system environment
	Openness	The degree to which the artifact lends itself to inspection
Sauser et al. (2010)	Diversity	The degree to which the artifact is composed of differing elements