Evaluating Knowledge Types in Design Science Research: An Integrated Framework

Jacky Akoka[1], Isabelle Comyn-Wattiau[2(✉)], Nicolas Prat[2], and Veda C. Storey[3]

[1] CEDRIC-CNAM, Paris, France
jacky.akoka@cnam.fr
[2] ESSEC Business School, Cergy-Pontoise, France
{wattiau,prat}@essec.edu
[3] Department of Computer Information Systems, Georgia State University, Atlanta, GA, USA
vstorey@gsu.edu

Abstract. As the design science research (DSR) paradigm has evolved, several frameworks and taxonomies have been proposed to aid researchers in understanding and applying DSR principles and practices to evaluate the knowledge produced. This paper attempts to integrate two such efforts, the evaluation methods taxonomy and the genres of inquiry framework, in an attempt to derive a more complete evaluation standard. The integration is based on various knowledge types, using mapping and merging techniques. Doing so results in three artifacts. The first is the integrated framework which refines the genres of inquiry framework with six recognized knowledge types (definitional, descriptive, prescriptive, explanatory, predictive, and explanatory and predictive). Evaluation criteria for each knowledge type emerge from the integration. The second contribution is a guidance scheme that helps in determining relevant evaluation criteria, based upon the type of DSR contribution and the role of the evaluator (DSR researcher, author, reader, or reviewer). Finally, the approach to the integration, including the integration algorithm, is adaptable to other contexts where criteria to evaluate different knowledge types need to be integrated.

Keywords: Design science research · Nomothetic design · Nomothetic science · Evaluation criteria · Evaluation methods taxonomy · Genres of inquiry framework · Knowledge types · Mappings · Artifacts · Integrated evaluation framework

1   Introduction

Many efforts in design science research have focused on analyzing this research approach to propose ways to effectively carry out meaningful research and to guide research initiatives. An important aspect of design science research is the evaluation of the results (e.g., [1]), although there is no generally-accepted, systematic way to do so. Prat et al. [2], for example, develop a detailed Evaluation Methods Taxonomy for information systems artifacts, including a hierarchy of criteria, derived from analysis of existing studies on design science research. The Genres of Inquiry Framework [3] is

comprised of two dimensions: knowledge goal (design or science), and knowledge scope (idiographic or nomothetic), from which general evaluation criteria are derived (based upon general scientific research) for each of the resulting four quadrants. These two approaches share the common goal of providing specific criteria for evaluation in design science research. The Genres of Inquiry Framework has a broad scope, proposing general evaluation criteria for the four combinations of knowledge goal and knowledge scope. The Evaluation Methods Taxonomy has a more specific scope as it focuses on the evaluation of artifacts contributed by design science researchers. In other words, the focus of the Evaluation Methods Taxonomy is on nomothetic design. The criteria proposed in this taxonomy are at a finer level of granularity than those of the Genres of Inquiry Framework and are only one among the six dimensions of evaluation methods for information systems artifacts. Thus, the two approaches are complementary in the sense that both attempt to provide specific criteria for artifact or knowledge evaluation in design science research, but differ in terms of scope and level of detail.

Since there is no generally accepted way to carry out the evaluation process in design science research, integrating these two complementary approaches might lead to insights on how a more general approach might be developed. The objective of this research, then, is to explore and integrate these two approaches in an attempt to develop a more complete set of evaluation criteria from which a set of actionable evaluation guidelines in design science research could be proposed.

To conduct the research, the integration is carried out by first considering the set of six types of design science research knowledge proposed by Johannesson and Perjons [4], to which the criteria can be associated. Then, mapping and merging rules are developed and applied to integrate the Evaluation Methods Taxonomy with the Genres of Inquiry Framework. The results highlight missing criteria, which are needed to develop a more complete set of evaluation criteria.

The mapping and merging results in an integrated evaluation framework, where the criteria depend on the knowledge type. Some criteria are organized hierarchically and decomposed into sub-criteria. This paper focuses on the results from the integration of criteria for nomothetic knowledge. Criteria for evaluating idiographic knowledge could be integrated following the same principles. Application of the integrated evaluation framework is facilitated by a guiding scheme that helps to identify relevant evaluation criteria, based on the type of design science research contribution and the role of the evaluator. To illustrate, the guidance scheme is then applied to four types of DSR contributions from articles published in journals from the AIS Senior Scholars' basket (www.aisnet.org).

The primary contribution of this research is the integrated evaluation framework. In addition, the guidance scheme for the identification of relevant criteria is detailed as well as the integration approach, which includes an integration algorithm.

The next section reviews work in design science research evaluation, the Evaluation Methods Taxonomy, and the Genres of Inquiry Framework. The integration approach is described in the following section, followed by the results of carrying out the integration. Then, the insights gained are presented as a new guidance scheme that considers the knowledge type and actor's role. This is illustrated by four examples where the role is that of a researcher. The conclusion summarizes the research and suggests avenues for future work.

2   Related Research

We briefly review previous work on evaluation in DSR, focusing specifically on evaluation criteria. Then, the essence of the Evaluation Methods Taxonomy and the Genres of Inquiry Framework are presented to identify their complementarities, which provides the basis for performing the integration.

2.1   Evaluation in DSR

Evaluation is a critical part of design science research [2, 5]. According to Goes [6, pp. v–vi], a major concern of design science research is to "*create knowledge through meaningful solutions that survive rigorous validations through proof of concept, proof of use, and proof of value*." As a result, design science researchers need to understand and apply acceptable criteria to evaluate the outcomes of their research [3].

Traditionally, DSR is decomposed into two activities: build and evaluate [7, 8]. However, artifact building and evaluation are intertwined, with several micro-evaluations [9] carried out during design. Evaluation is central to the DSR process presented by Hevner [10]. This process comprises three inter-related cycles: relevance, design (build and evaluate), and rigor. The DSR methodology of Peffers et al. [11] comprises an activity dedicated to evaluation. Sonnenberg and vom Brocke [12] distinguish four different evaluation activities. Each of these four activities has specific goals, evaluation criteria, and evaluation methods. Some evaluations are carried out early in the DSR process. The benefit of these early evaluation activities should be outweighed by their cost [13]. Moreover, coherence should be ensured between the evaluation activities [14]. In addition to evaluations carried out as part of the DSR process, retrospective evaluation has been suggested to gain knowledge from both successful and unsuccessful DSR projects [15].

A Framework for Evaluation in Design Science (FEDS) [1] has been proposed that encompasses many aspects of evaluation in DSR. It comprises two dimensions: the purpose of the evaluation (formative or summative) and the paradigm of the evaluation (artificial or naturalistic). From this framework, evaluation strategies may be defined. An evaluation strategy is a planned trajectory along the two dimensions of the framework. There are also steps identified for the evaluation process, including the choice of the evaluation strategy. However, FEDS is a high-level framework. March and Smith [7] list some evaluation criteria for the four types of artifacts that they define: constructs, models, methods, and instantiations. This list does not necessarily aim at completeness, similarly for that proposed by Hevner et al. [5] and Sonnenberg and vom Brocke [12]. The criteria by Aier and Fischer [16] draw on those defined by Kuhn [17] for traditional science, but are specific to design theories. Thus, the literature lacks a complete list of evaluation criteria for DSR, covering the different types of knowledge and artifacts. Two systematic approaches to developing criteria for DSR have been proposed [2, 3] and are reviewed below. These approaches are complementary, in that one is based on the analysis of existing work on artifact development in design science research and the other started from the general literature on scientific research. Combining these approaches leads to the derivation of an integrated evaluation framework

that systematically specifies applicable evaluation criteria for the different types of DSR knowledge.

2.2   Evaluation Methods Taxonomy

The development of the Evaluation Methods Taxonomy [2] was motivated by the need to investigate the "what" and the "how" of information systems artifact evaluation. That is, what the artifacts are, and what the criteria of evaluation are, as well as the relationship between the "what" and the "how." By doing so, the research identifies important relationships between different dimensions of design science research artifact evaluation. The taxonomy of evaluation methods comprises six dimensions, including criterion and evaluation technique. The dimension "criterion" proposes a complete hierarchy of evaluation criteria for information systems artifacts. Based on a systems view of artifacts, the first level of the hierarchy is formed by the five aspects constituting the canonical form of systems: goal, environment, structure, activity, and evolution, each of which has criteria, sub-criteria, and sub-sub-criteria.

2.3   Genres of Inquiry Framework

The Genres of Inquiry Framework for design science research [3], recognizes four modes of reasoning that can exist in design science research as derived from analyzing the knowledge goals (design versus science) and knowledge scope (idiographic and nomothetic) in knowledge production. The result is the four genres of idiographic design (ID), idiographic science (IS), nomothetic design (ND), and nomothetic science (NS), each of which has its own evaluation criteria.

3   Integration of Evaluation Approaches

The Evaluation Methods Taxonomy is intended to express evaluation methods, including evaluation criteria, in a systematic way. The focus is on evaluating artifacts contributed by design science researchers, primarily constructs, models or methods. This corresponds to the quadrant of nomothetic design in the Genres of Inquiry Framework. The latter framework has coarse categories. Thus, the Genres of Inquiry Framework has a broader scope, in that it considers all types of DSR knowledge and proposes general criteria to evaluate the four types of knowledge (ND, NS, ID and IS). The Evaluation Methods Taxonomy has a more specific scope, presenting detailed criteria, and more generally, evaluation methods, for artifacts that are typically constructs, models, or methods (ND). Given these complementarities, integrating the Evaluation Methods Taxonomy with the Genres of Inquiry Framework should result in a more complete approach to evaluation in DSR, at a finer level of granularity. The process for integrating these two approaches involves mapping from one approach to the other. This process requires the following steps: mapping between genres of inquiry and knowledge types, artifact positioning, criteria mapping, and criteria integration.

These four steps are detailed and applied below, resulting in an integrated framework of evaluation criteria organized by knowledge type.

3.1   Step 1. Mapping Between Genres of Inquiry and Knowledge Types

The Evaluation Methods Taxonomy deals mainly with artifact evaluation. The Genres of Inquiry Framework focuses on knowledge evaluation and characterizes knowledge by its scope and goal. In order to integrate both evaluation mechanisms, we rely on a pivot concept based on the knowledge types proposed by Johannesson and Perjons [4]. The authors distinguish six knowledge types: definitional, descriptive, prescriptive, explanatory, predictive, and explanatory and predictive. These knowledge types are closely related to the types of theories defined by Gregor [18]. We should note the distinction between definitional and descriptive knowledge, refining Gregor's notion of "theories for analyzing." For example, the terminological box in an ontology is definitional knowledge, and the assertion box is descriptive knowledge.

We position the six knowledge types into the four quadrants of the Genres of Inquiry Framework. We justify this matching by the fact that both structures (quadrants and Johannesson and Perjons' [4] categories) represent different knowledge. The resulting knowledge types are shown in bold characters in Fig. 1 (the rest of the figure
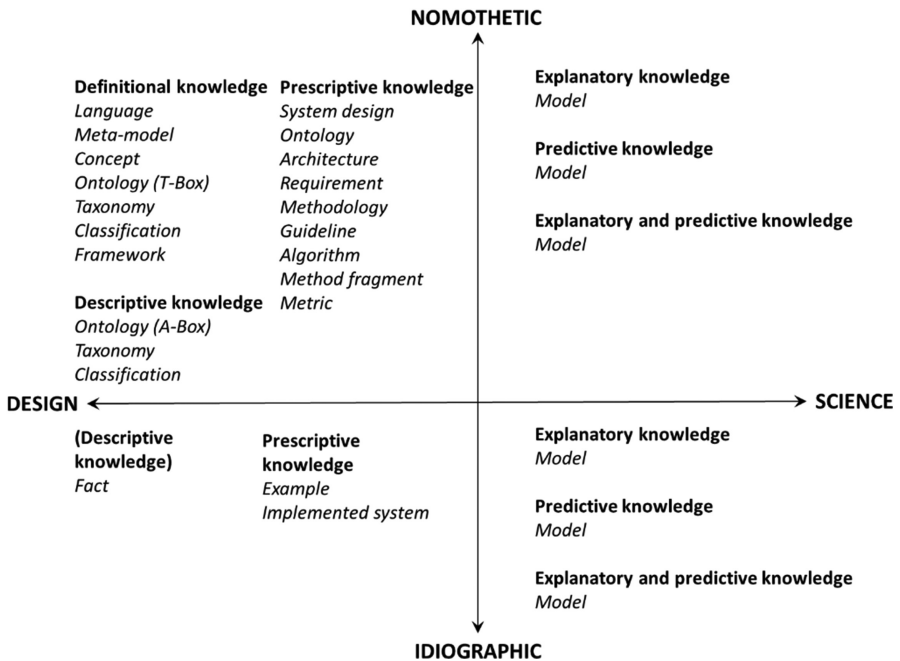

Fig. 1. Towards an integrated evaluation framework: mapping knowledge types and artifact types to genres

is explained in Sect. 3.2). Explanatory, predictive, and explanatory and predictive knowledge pertain to the "science" dimension. The other types of knowledge are related to design. Each knowledge type is defined at the nomothetic and idiographic level. There is no definitional knowledge at the idiographic level. Descriptive knowledge at the idiographic level may be considered outside the scope of research. Thus, this integration results in ten knowledge types.

3.2   Step 2. Artifact Positioning

From the results of step 1, we identify the types of artifacts associated with each knowledge type. Employing the typology of artifacts proposed by Sangupamba Mwilu et al. [19], we enrich it to include models generated by quantitative or qualitative research. We also extend the typology of artifacts by adding classification [20], as distinguished from the concept of taxonomy. The result of this step is presented in Fig. 1. The mapping and integration, described below, focuses on the integration of criteria for knowledge types in the nomothetic quadrants, because the latter deal with general theories or concepts that cover a set of classes.

3.3   Step 3. Criteria Mapping

First, for each knowledge type, we specify the applicable criteria from the Genres of Inquiry Framework and from the Evaluation Methods Taxonomy. To decide what criteria to apply from the Genres Framework, for each knowledge type, we consider the criteria from the relevant quadrant (as identified by Baskerville et al. [3]) and select the subset of those that are relevant depending on the knowledge type. To decide what criteria to apply from the Evaluation Methods Taxonomy (as identified by Prat et al. [2]), we consider those that are relevant for the artifact types associated to the knowledge type.

Second, we define the mappings between criteria. When the criteria are the same or similar (but represented differently), then they are mapped to each other. For example, generalizability in the Genres Framework is equivalent to adaptability in the Evaluation Methods Taxonomy. More generally, five types of mappings between concepts are adopted to perform the mappings: more abstract, less abstract, equivalent, compatible, and disjoint [21].

Mappings, in general, make explicit a relationship between elements, or set of elements, of different conceptualizations and/or instantiations [22]. Mappings have been used in the past to support integration. Choi et al. [23], for example, identify three broad categories of mappings to support ontology integration. Noy [24] proposes the use of ontologies and mapping to a common ontology to deal with issues of heterogeneity in structured data. A discussion of mapping-based merging, as required in this research, is found in [25]. The methods focus on the scalability issue when defining mappings requires a combinatory explosion of two-by-two concept comparison. In our

context, scalability is not an issue. Thus, we manually completed the mapping matrices (one per knowledge type). The mapping matrix for predictive knowledge in the nomothetic science quadrant is shown in Table 1. The applicable criteria from the Genres Framework are shown in the columns. Those from the Evaluation Methods Taxonomy appear in the rows.

Table 1. Mapping matrix for NS predictive knowledge

NS predictive	validity	utility	generality	completeness	simplicity	understandability	consistency	fidelity to modeled phenomena	accuracy	adaptability
applicability	⊥	⊥	*	⊥	⊥	⊥	⊥	⊥	⊥	≡
generalizability	⊥	⊥	*	⊥	⊥	⊥	⊥	⊥	⊥	≡
external validity	⊆	⊥	⊥	⊥	⊥	⊥	⊥	⊆	⊥	⊥
consistency	⊥	⊥	⊥	⊥	⊥	⊥	≡	⊥	⊥	⊥
reliability	⊥	⊥	⊥	⊥	⊥	⊥	⊆	⊥	⊥	⊥
internal validity	⊆	⊥	⊥	⊥	⊥	⊥	⊥	⊆	⊥	⊥
objectivity	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥

Legend:
⊇ more abstract than ≡ equivalent to ⊥ disjoint from
⊆ less abstract than * compatible with

3.4 Step 4. Criteria Integration

This step is dedicated to the integration of criteria based upon mapping matrices. Based upon these mappings, the criteria are integrated using the algorithm shown in Fig. 2. This algorithm enables the integration of the criteria for each mapping matrix in a systematic and replicable way, defined specifically for this research. The main principles underlying the algorithm are: (i) order the five mapping types as follows: equivalent, more abstract, less abstract, compatible, disjoint; (ii) examine, step by step, all the mapping types; (iii) merge the equivalent criteria; (iv) transform more abstract or less abstract mappings into generalization links between criteria; (v) create a common criterion for compatible criteria; and (vi) remove disjoint links and potential multiple inheritance cases. Thus, we obtain a unique hierarchy of criteria for each knowledge type.

```
  INTEGRATION ALGORITHM
  {This algorithm builds a list of criteria which may contain
ISA links}
  *** Step 0 – Initialization – Graph construction
  Each criterion is a node, each "equivalent" (resp. "less
abstract" or "compatible") link is an edge labeled with
"equivalent" (resp. "less abstract" or "compatible"). Only
"less abstract" labeled edges are directed. "More abstract"
links lead to inverse edges labeled "less abstract".

  *** Step 1 – Transformation of "equivalent" edges
  For each "equivalent" edge
  Do  Merge all the nodes belonging to the corresponding
equivalence class
      Call routine 1 for naming the resulting node
      Call routine 2 for transferring edges from initial
nodes to the resulting merged node
      Delete the "equivalent" edge
  End For {notice that, at the end of this step, there is no
remaining "equivalent" edges}

  *** Step 2 – Transformation of "less abstract" edges
  For each "less abstract" edge
  Do  Create an ISA link between corresponding nodes
      If   it generates a multiple inheritance
      Then  call routine 3 for deleting the multiple inher-
itance
      End If
      Delete the "less abstract" edge
  End For {notice that, at the end of this step, there is
neither remaining "equivalent" nor "less abstract" edges}

  *** Step 3 – Transformation of "compatible" edges
  For each "compatible" edge
If    the two corresponding nodes and possibly others also
      linked to them by "compatible" edges have no common ge-
      neric node
  Then      Create a generic node
      Call routine 1 for naming this new node
      If   it generates a multiple inheritance
      Then  call routine 3 for deleting the multiple inher-
itance
      End If
  End If
  Delete the "compatible" edge
  End For

  *** Step 4 – end of process
  The remaining criteria are added to the resulting graph.
  The expert can improve this result, especially by examining
in detail the hierarchies notably those where the specific
criteria do not completely cover the generic criterion.
```

Fig. 2. Integration algorithm

```
  Routine 1 {naming a node resulting from the mapping pro-
cess}
  If  the initial nodes share the same name
      Then  the new node will also share the same name
      Else  concatenate the names of initial nodes
  End If     {An expert may also suggest a better naming}
  End routine 1

  Routine 2 {in case of node merging, the routine transfers
the edges connected to the initial nodes}
  For  each edge connecting the merged node to another node N
  Do  Transfer this edge such that it now links the new node
resulting from the merging to node N
  End For
  If   at the end of the step, two edges share the same label
between two nodes
      Then  merge these edges
  End If
If     at the end of the step, two edges exist between two
      nodes with one label "less abstract" and the other la-
      bel "compatible"
      Then  merge these edges and assign the label "less ab-
stract"
  End If
  If  at the end of the step, two inverse edges with label
"less abstract" link two same nodes
      Then  merge the two nodes
      Call routine 1 for naming the new node
  End If
  End routine 2

  Routine 3 {deletion of multiple inheritance}
  If  there exists a multiple inheritance (i.e. two ISA links
start from the same node)
      Then  the expert must check the definitions of crite-
            ria associated to the target nodes in order to
            decide whether they have to merge
      If    he/she decides to merge them
            Then  merge the two nodes
                  Call routine 1 for naming the new node
            Else  the expert must check the definitions of
                  criteria associated to the source node and
                  to all cousin nodes in the ISA hierarchy
                  in order to decide which link is more rel-
                  evant
                  Delete the less relevant ISA link
      End If
  End If
  End routine 3
```

Fig. 2. (continued)

To further enrich the hierarchies of criteria, we used two other sets of criteria, complementing those of the Evaluation Methods Taxonomy and the Genres of Inquiry Framework: the criteria proposed by Weber [26] for evaluating theories in information systems, and those proposed by Kuhn for general scientific research [17]. Thus, for each knowledge type, we applied steps 3 and 4 in three successive iterations:

- *Iteration 1*: map and integrate the applicable criteria from the Genres Framework and the Evaluation Methods Taxonomy.
- *Iteration 2*: map and integrate the criteria resulting from iteration 1 and the applicable criteria from Weber [26].
- *Iteration 3*: map and integrate the criteria resulting from iteration 2 and the applicable criteria from Kuhn [17].

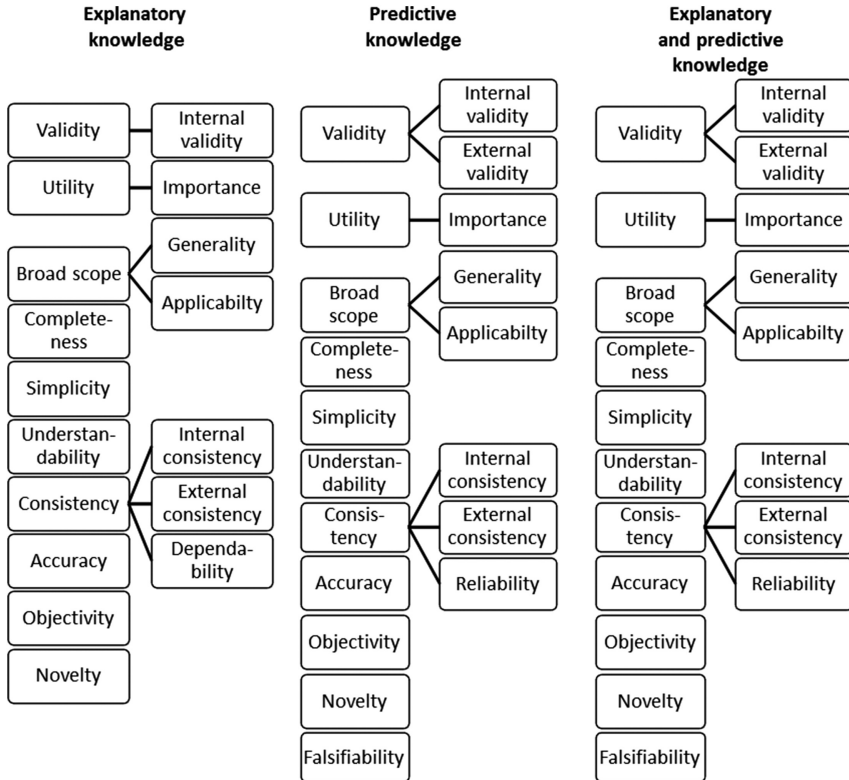The six resulting hierarchies for the nomothetic quadrants are depicted in Figs. 3 and 4.

Fig. 3. Criteria for knowledge types in the nomothetic science quadrant

Definitional knowledge

- Goal attainment
  - Efficacy
  - Effectivenes
- Validity
- Utility — Importance
- Broad scope
  - Generality
  - Adaptability
- Usefulness
- Ease of use
- Completeness
- Simplicity
- Understandability
- Consistency
  - Internal consistency
  - External consistency
- Modifiability
- Novelty
  - Innovativeness
  - Originality

Descriptive knowledge

- Goal attainment
  - Efficacy
  - Effectiveness
- Validity
- Utility — Importance
- Generality
- Usefulness
- Ease of use
- Completeness
- Simplicity
- Understandability
- Consistency
  - Internal consistency
  - External consistency
- Accuracy
- Novelty — Originality

Prescriptive knowledge

- Goal attainment
  - Efficacy
  - Effectiveness
- Validity
- Utility — Importance
- Feasability
  - Technical feasability
  - Operational feasabilty
  - Economic feasability
- Broad scope
  - Generality
  - Applicability
- Usefulness
- Ease of use
- Alignement with business
- Fit into technical IS architecture
- Completeness
- Simplicity
- Understandability
- Consistency
  - Internal consistency
  - External consistency
- Correspondence with another model
- Accuracy
- Performance
- Robustness
- Scalability
- Modifiability
- Learning capability
- Novelty
  - Inventiveness
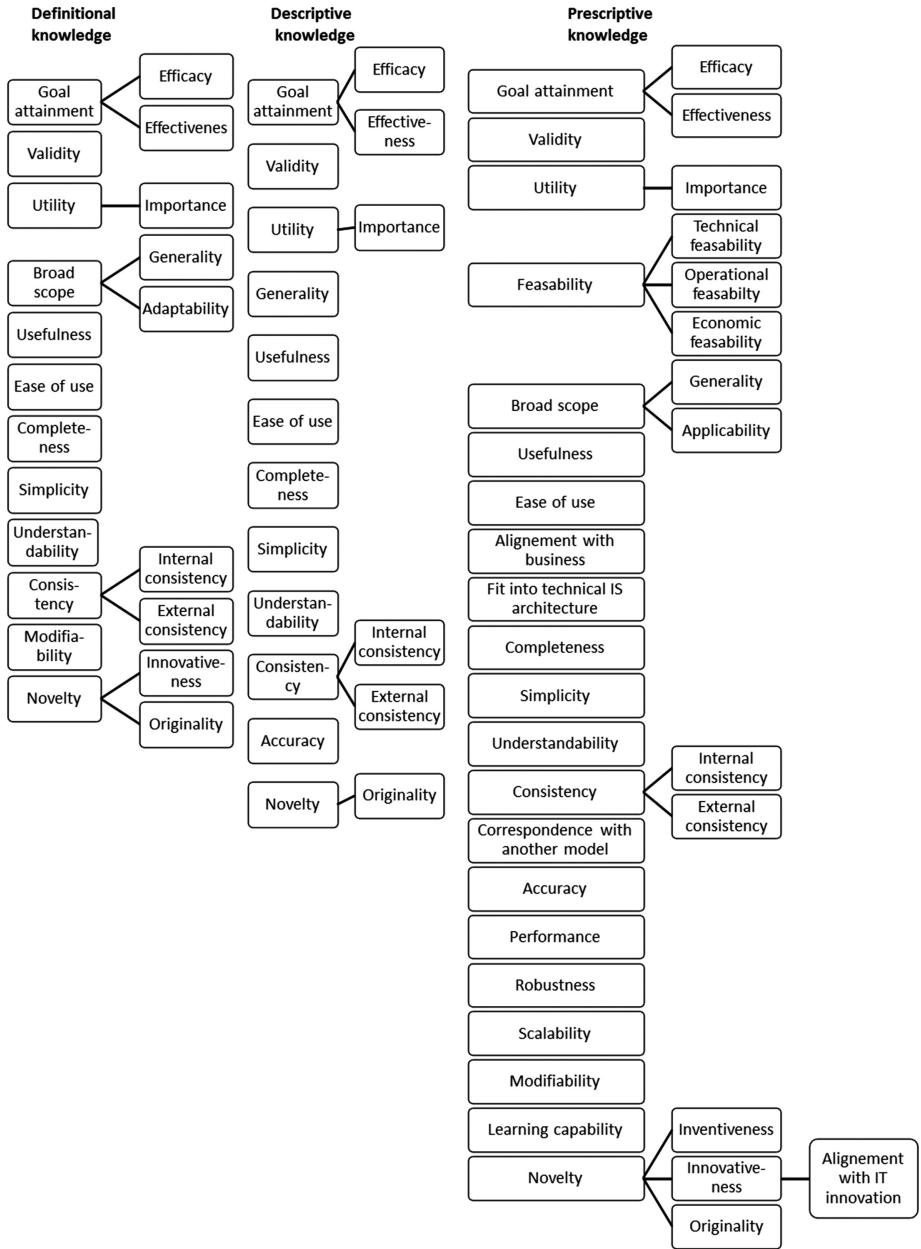 - Innovativeness — Alignement with IT innovation
  - Originality

Fig. 4. Criteria for knowledge types in the nomothetic design quadrant

4   Application of the Integrated Evaluation Framework

We have defined integrated evaluation criteria for nomothetic knowledge in DSR. From the results of integrating the two approaches, we derive a criteria selection scheme, the purpose of which is to provide guidance for selecting applicable evaluation criteria for DSR projects. The overall goal of the guidance scheme is to facilitate, enrich, and transform the evaluation process. The guidance scheme and its application to four types of DSR contribution are each illustrated using a published DSR paper.

4.1   Guidance Scheme

Applying the integrated framework requires characterizing the type of DSR contribution which means placing the research in a quadrant of the framework (at least at a given moment). Since a quadrant may contain several knowledge types, placement within a quadrant leads to the selection of a specific knowledge type from among those in that quadrant.

The framework may be used in different ways, depending upon the roles of the actors in DSR. We consider the following four roles (stakeholders):

- Researcher: framework helps define a research path for evaluation.
- Author: framework provides a set of criteria to apply in the evaluation.
- Reader: framework facilitates understanding of the research activities.
- Reviewer: framework suggests appropriate criteria for evaluation.

An author is a subtype of researcher and a reviewer is a subtype of reader.

DSR contributions can now be defined as chronological ordered sets of knowledge types where the evaluation should focus on certain knowledge types in the ordered set.

Summarizing, we provide a guidance scheme for aiding in the definition of an evaluation. It consists of: (i) defining the knowledge path as illustrated below; (ii) choosing the focus of the evaluation; (iii) deriving the sets of criteria that may be evaluated; and (iv) building the corresponding evaluation methods. The detailed description of the last step is beyond the scope of this paper.

In the next section, we derive four types of DSR contributions, illustrating each with a published DSR paper. We take the point of view of the DSR researcher. For the fourth type of DSR contribution, we also illustrate examples of criteria from the point of view of the reviewer. The proof of concept consists of comparing the criteria evaluated in the paper with the criteria suggested by our framework. Recall that the suggested criteria are deduced directly from the knowledge type.

4.2   Application to Four Types of DSR Contributions

Type 1: ND definitional → *ND prescriptive* → ID prescriptive, where the focus of evaluation is on ND prescriptive.

This path occurs when the researcher: (i) proposes an artifact among (language, concept, meta-model, ontology, taxonomy, framework, classification), belonging to

definitional knowledge, (ii) proposes an artifact among (methodology, algorithm, method fragment, guideline, etc.) based on this artifact, generating prescriptive knowledge, (iii) applies the artifact, generating idiographic prescriptive knowledge. Moreover, this type is related to the case where the focus of evaluation is on the second artifact (ND prescriptive).

Arnott [27] perfectly exemplifies this type. He proposes a taxonomy of cognitive biases (ND definitional) with an evolutionary DSS development methodology that uses cognitive bias theory as a focusing construct (ND prescriptive). The methodology is applied to a strategic DSS project (ID prescriptive). The focus of the contribution and evaluation is the DSS development methodology that uses the taxonomy of cognitive biases. Arnott [27] evaluated the effectiveness (degree to which the artifact achieves its goal in a real situation) and the operational feasibility (integration of the artifact in the daily practice of users) of the methodology. He could also have evaluated the simplicity and the understandability of the methodology by conducting another case study in which the analyst would have been different.

Type 2: *ND definitional* → ND prescriptive → ID prescriptive, where the focus of evaluation is on ND definitional.

Adomavicius et al. [28] develop the REQUEST query language and associated RA algebra (ND definitional), a mapping algorithm from REQUEST to RA (ND prescriptive), with an application to examples (ID prescriptive). The focus of the research is on the evaluation of the query language and associated algebra. The researchers evaluate the efficacy of REQUEST by applying it to example queries. They also evaluate the expressive power (completeness) of REQUEST and RA. They could also have applied other criteria relevant for ND definitional knowledge, as suggested in Fig. 4. For example, they could also have evaluated the simplicity of REQUEST and RA, or the understandability of REQUEST by conducting a laboratory experiment. Note that in this type of DSR contribution, the ordered set of knowledge types is the same as that for Type 1. However, for Type 2, the evaluation focuses on ND definitional, as opposed to ND prescriptive for Type 1.

Type 3: *ND prescriptive* → ID prescriptive → *ND definitional*, where the focus of evaluation is on ND prescriptive and ND definitional.

Nickerson et al. [29] present a methodology for taxonomy development (ND prescriptive) and application of the methodology (ID prescriptive). The application of the methodology results in a taxonomy of mobile applications (ND definitional). The authors evaluate the usefulness of the taxonomy using a laboratory experiment. They could also have applied other criteria, such as completeness or modifiability, as well as the other ND definitional criteria. By building this taxonomy, they validated operational feasibility of the methodology (evaluation of the ND prescriptive knowledge). The authors listed a set of desirable properties for such methodologies that they evaluated using an informed argument. These properties correspond to performance, simplicity, and utility in our ND prescriptive quadrant. Note that the authors mention two properties very specific to taxonomy building, namely the possibility of taking into consideration alternative approaches and the reduction of arbitrariness.

Type 4: *NS explanatory and predictive → ND prescriptive → NS predictive →* **IS prescriptive, where the focus of evaluation is on NS explanatory and predictive and ND prescriptive.**

Arazy et al. [30] propose elements for a design theory of social recommender systems, based on the components of information system design theories [31]. The researchers introduce the concept of "applied behavioral theory," making the link between Walls et al.'s [31] kernel theories and meta requirements. In their case, the applied behavioral theory is a theory that explains and predicts willingness to accept advice. The applied behavioral theory (NS explanatory and predictive) leads to the meta requirements and meta design of social recommender systems (ND prescriptive), followed by testable product hypotheses (NS predictive), and then, a system implementation (IS prescriptive). The applied behavioral theory is carefully tested for reliability and validity (e.g., discriminant validity). It is also judged too complex (simplicity) for practical use, and is simplified from a PLS model (explanatory and predictive) into a regression model (essentially predictive). Our criteria for NS explanatory and predictive knowledge suggest other possible criteria for evaluating the applied behavioral theory, e.g. importance (this criterion may be evaluated ex-post, e.g. based on use of the applied behavioral theory in other DSR projects). With respect to the evaluation of the meta design, the authors focus on technical feasibility (through the implementation of a system) and accuracy. Our criteria for ND prescriptive knowledge suggest other possible criteria, e.g., utility or innovativeness.

Other paths, highlighting DSR contributions, can be defined. However, whatever the path, the integrated evaluation framework can potentially help in the evaluation process for all stakeholders. The framework offers the possibility of providing the researcher with a more complete set of relevant criteria, including three parameters: the scope of knowledge (nomothetic or idiographic), the goal of knowledge (science or design), and the type of knowledge (definitional, prescriptive, etc.). For each triple of parameter values, a hierarchy of criteria is derived.

5   Conclusion and Future Research

Despite the fact that design science research evaluation of artifacts has been addressed from different aspects (e.g., tactical and operational [2] versus strategic [1]), there is still a need for a comprehensive approach to artifact and knowledge evaluation. This research has attempted to integrate two complementary approaches, the Evaluation Methods Taxonomy and the Genres of Inquiry Framework, to derive a more complete set of integrated evaluation criteria. The result is an integrated evaluation framework. This framework:

1. refines the Genre of Inquiry Framework with the six types of design science research knowledge proposed by Johannesson and Perjons [4], as illustrated in Fig. 1; and
2. proposes a hierarchy of criteria for each knowledge type in the nomothetic quadrants, by mapping and integrating the criteria from the Genres of Inquiry Framework and the Evaluation Methods Taxonomy (Figs. 3 and 4).

To guide the choice of applicable evaluation criteria in the integrated evaluation framework, a guidance scheme is proposed. This scheme considers: (1) the type of DSR contribution, defined as a chronological ordered set of knowledge types where the evaluation should focus on certain knowledge types in an ordered set; and (2) the role of the actor (researcher, author, reader, or reviewer). To evaluate the efficacy and utility of the integrated evaluation framework and associated guidance scheme, they are applied to four studies (DSR papers published in the AIS basket of journals).

The benefit of the approach can be realized by a researcher, author, reader or reviewer. If they can identify where a DSR project or paper fits in terms of knowledge types, then they can use the identified criteria, depending on their role. One limitation of the approach is that different research paradigms (e.g., positivist versus interpretivist) have different views on criteria for evaluating knowledge, and on whether it is possible to objectively evaluate scientific knowledge. Consequently, the epistemological challenges of combining different criteria in a single evaluation framework deserve further consideration. Moreover, even if we contend that DSR should benefit from comprehensive hierarchies of evaluation criteria, some criteria may not be defined a priori and are specific to particular research endeavors.

Future research can proceed in several directions. The guidance scheme can be extended by identifying other types of DSR contributions, in addition to the four types illustrated. Our approach needs to be evaluated more extensively and expanded to deal with the idiographic quadrants [3]. Although the guidance scheme supports the identification of relevant evaluation criteria, it does not suggest when these criteria should be evaluated (e.g., formative versus summative evaluation). To assist in the definition of an overall evaluation agenda, we may combine our approach with evaluation strategies [1]. Another possibility is to adapt and extend the evaluation methods identified by Prat et al. [2].

References

1. Venable, J., Pries-Heje, J., Baskerville, R.: FEDS: a framework for evaluation in design science research. Eur. J. Inf. Syst. **25**, 77–89 (2016)
2. Prat, N., Comyn-Wattiau, I., Akoka, J.: A taxonomy of evaluation methods for information systems artifacts. J. Manag. Inf. Syst. **32**, 229–267 (2015)
3. Baskerville, R.L., Kaul, M., Storey, V.C.: Genres of inquiry in design-science research: justification and evaluation of knowledge production. MIS Q. **39**, 541–564 (2015)
4. Johannesson, P., Perjons, E.: Knowledge types and forms. In: Johannesson, P., Perjons, E. (eds.) An Introduction to Design Science, pp. 21–38. Springer, Cham (2014)
5. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design science in information systems research. MIS Q. **28**, 75–105 (2004)
6. Goes, P.B.: Design science research in top information systems journals. MIS Q. **38**, iii–viii (2014)

7. March, S.T., Smith, G.F.: Design and natural science research on information technology. Decis. Support Syst. **15**, 251–266 (1995)
8. Winter, R.: Design science research in Europe. Eur. J. Inf. Syst. **17**, 470–475 (2008)
9. Vaishnavi, V., Kuechler, B.: Design science research in information systems, pp. 1–45 (2004). http://desrist.org/design-research-in-information-systems
10. Hevner, A.R.: A three cycle view of design science research. Scand. J. Inf. Syst. **19**, 87–92 (2007)
11. Peffers, K., Tuunanen, T., Rothenberger, M.A., Chatterjee, S.: A design science research methodology for information systems research. J. Manag. Inf. Syst. **24**, 45–77 (2007)
12. Sonnenberg, C., vom Brocke, J.: Evaluations in the science of the artificial – reconsidering the build-evaluate pattern in design science research. In: Peffers, K., Rothenberger, M., Kuechler, B. (eds.) DESRIST 2012. LNCS, vol. 7286, pp. 381–397. Springer, Heidelberg (2012). doi:10.1007/978-3-642-29863-9_28
13. Braunnagel, D., Leist, S.: Applying evaluations while building the artifact - experiences from the development of process model complexity metrics. In: 49th Hawaii International Conference on System Sciences (HICSS), pp. 4424–4433. IEEE (2016)
14. Abraham, R., Aier, S., Winter, R.: Fail early, fail often: towards coherent feedback loops in design science research evaluation. In: 2014 International Conference on Information Systems (ICIS), pp. 1–12. Association for Information Systems (2014)
15. Uppström, E.: Re-visiting IS design science artifacts: making a case for critical realism. In: 50th Hawaii International Conference on System Sciences (HICSS), pp. 4675–4684. IEEE (2017)
16. Aier, S., Fischer, C.: Criteria of progress for information systems design theories. Inf. Syst. e-Bus. Manag. **9**, 133–172 (2011)
17. Kuhn, T.S.: The Essential Tension: Selected Studies in Scientific Tradition and Change. University of Chicago Press, Chicago (1977)
18. Gregor, S.: The nature of theory in information systems. MIS Q. **30**, 611–642 (2006)
19. Sangupamba Mwilu, O., Comyn-Wattiau, I., Prat, N.: Design science research contribution to business intelligence in the cloud—a systematic literature review. Future Gener. Comput. Syst. **63**, 108–122 (2016)
20. Parsons, J., Wand, Y.: Emancipating instances from the tyranny of classes in information modeling. ACM Trans. Database Syst. **25**, 228–268 (2000)
21. Bouquet, P., Giunchiglia, F., van Harmelen, F., Serafini, L., Stuckenschmidt, H.: C-OWL: Contextualizing Ontologies. In: Fensel, D., Sycara, K., Mylopoulos, J. (eds.) ISWC 2003. LNCS, vol. 2870, pp. 164–179. Springer, Heidelberg (2003). doi:10.1007/978-3-540-39718-2_11
22. Pérez, A.G., Gargantilla, J.Á.R.: Semantic mappings: out of ontology world limits. In: International Conference on Complex, Intelligent and Software Intensive Systems (CISIS), pp. 907–912. IEEE (2008)
23. Choi, N., Song, I.-Y., Han, H.: A survey on ontology mapping. SIGMOD Rec. **35**, 34–41 (2006)
24. Noy, N.F.: Semantic integration: a survey of ontology-based approaches. SIGMOD Rec. **33**, 65–70 (2004)
25. Pottinger, R.: Mapping-based merging of schemas. In: Bellahsene, Z., Bonifati, A., Rahm, E. (eds.) Schema Matching and Mapping, pp. 223–249. Springer, Berlin (2011)
26. Weber, R.: Theory building in the information systems discipline: some critical reflections. In: Gregor, S.D., Hart, D.N. (eds.) Information Systems Foundations: Theory Building in Information Systems, pp. 1–20. Australian National University, Canberra (2012)
27. Arnott, D.: Cognitive biases and decision support systems development: a design science approach. Inf. Syst. J. **16**, 55–78 (2006)

28. Adomavicius, G., Tuzhilin, A., Zheng, R.: REQUEST: a query language for customizing recommendations. Inf. Syst. Res. **22**, 99–117 (2011)
29. Nickerson, R.C., Varshney, U., Muntermann, J.: A method for taxonomy development and its application in information systems. Eur. J. Inf. Syst. **22**, 336–359 (2013)
30. Arazy, O., Kumar, N., Shapira, B.: A theory-driven design framework for social recommender systems. J. Assoc. Inf. Syst. **11**, 455–490 (2010)
31. Walls, J.G., Widmeyer, G.R., El Sawy, O.A.: Building an information system design theory for vigilant EIS. Inf. Syst. Res. **3**, 36–59 (1992)